SAPIENZA
UNIVERSITÀ DI ROMA

Enhancing Decision-Making with Data-Driven Insights in Critical Situations: Impact and Implications of AI-Powered Predictive Solutions

Faculty of Economics
Department of Management

PhD in Management, Banking and Commodity Sciences – Cycle: XXXVII

PhD Candidate
Davide Liberato lo Conte

Supervisor
Prof. Giuseppe Sancetta

A. Y. 2024 - 2025

STVDIVM VRBIS

Abstract

Identifying early signals of crisis or insolvency is essential, for firms, to enable timely interventions, preserve financial health, and sustain competitive advantage. Traditional financial models, while foundational in this domain, often fall short in complex environments and there is a widening gap between the theoretical advancements in predictive modeling and their practical application in real-world decision-making, particularly for small and medium-sized enterprises, that may lack the resources to implement sophisticated risk management tools. Considering the Italian business landscape, this research examines the integration of advanced AI-driven techniques, showcasing their ability to handle vast and diverse data sources while improving predictive accuracy. By employing an inductive, multi-layered approach to model implementation and refinement, the study demonstrates how AI-based models can identify early indicators of distress up to five years before insolvency, thus providing a robust framework for preventive strategies. A key added value is the focus on enhancing explainability and understandability within AI models, delivering critical insights for both academia and managers; by improving transparency, these models not only strengthen predictive capacity but also enable stakeholders to better understand and interpret the drivers of financial and economic distress.

Furthermore, the study highlights the challenges of adopting new technologies in organizational contexts, addressing issues such as data ethics, managerial literacy, over-reliance risks and the alignment of AI-based decision tools with regulatory standards. The findings contribute both to academic discourse and practical applications by advocating for a cultural shift toward data-driven decision-making in critical phases of business life; by balancing the precision of Machine Learning with the interpretability required for managerial planning, it is underscored that, while AI just complements human expertise, it is, today, instrumental in surviving in complex economic landscapes with greater foresight and resilience.

Acknowledgments

I extend my profound gratitude to Prof. Giuseppe Sancetta, whose exceptional guidance, intellectual rigor, and steadfast support have been crucial to the completion of this dissertation. His mentorship has significantly shaped my academic and professional path, for which I am deeply appreciative.

My sincere thanks go to the Department of Management at Sapienza University of Rome for its essential support and commitment to academic excellence; I am also grateful to the Professors, Researchers, and Colleagues who contributed to an enriching and collaborative academic environment.

Additionally, I am grateful for the research opportunities provided by Queen Mary University of London and London School of Economics; I extend special thanks to Prof. Giuliano Maielli and Prof. Saqib Shamim for their invaluable insights and support, which greatly contributed to the quality and depth of my research.

I am further indebted to the colleagues from Dublin City University, whose expertise and insights have enriched this work.

To all who have supported this endeavor, my sincere appreciation.

STVDIVM VRBIS

Index

Personal Journey and Reflections on the Research Process 1

Dissertation Style 2

A Critical Partnership in Decision-Making: Prelude 3

Introduction and Context 4

Research Objectives Overview 7

Chapters Outline 8

Chapter 1

Integrating Historical Foundations with Future Directions in Corporate Crisis and Insolvency Prediction: a Beginning 9

Chapter Prelude 10

1. Background and Significance of the Study 11

1.1. Historical perspectives: a brief run-through 11

2. Relevance of an Ever-Evolving Predictive Landscape 16

2.1. Future-Forward Approach from a Managerial Perspective 18

2.1.1. New Managerial Literacy Approaches 19

2.1.2. Development of a Data-Centric Culture 21

2.2. Ethical and Social Consequences 22

2.2.1. Privacy Concerns and Data Security 23

2.2.2. Transparency and Accountability in AI Decision-Making 25

2.2.3. Socioeconomic Implications of AI-Driven Predictions 26

2.2.4. Transforming Financial Resilience and Socioeconomics Dynamics 27

2.3. Predictive Analytics in Corporate Strategy 28

2.3.1. Aligning Predictive Insights with Business Objectives 28

2.3.2. Sustainability and CSR Alignment 29

2.3.2.1. In crisis and insolvency situation 30

2.4. Special Considerations for SMEs: Strategic Integration and Scalability 30

3. Research Objectives 31

3.1. Main Purposes 32

3.1.1. Bridging Traditional and Modern Predictive Models 33

3.1.2. Enhancing Predictive Accuracy and Explainability through AI Integration 34

3.1.3. Strategy for Crisis Management and Sustainability 35

3.2. Other Objectives 36

3.2.1. Identifying Key Indicators and Variables 36

3.2.2. Addressing Ethical Considerations in a Multi-Algorithmic Scenario 37

3.2.3. Providing a Comprehensive Guide for Managers 38

3.2.4. Bridging the Gap between Academia and Real-World Applications 38

4. Research Questions: Some Thoughts 39

4.1. Research Questions: a List 40

4.2. Methodological Framework for Addressing Research Questions 40

Chapter 2

Understanding the Fundamentals: Key Concepts and Terminologies 42

Chapter Prelude 43

1. The Importance of Grasping the Basics 44

1.1. Core Concepts 44

1.2. The Dynamic Nature of Core Concepts 46

2. Defining Corporate Crisis 46

2.1. Firm-Level Focus 47

2.1.1. Drawing the Boundaries 47

2.1.2. Key Differences 48

2.2. Strands of Literature 49

2.3. Between Normalcy and Dysfunction 50

2.3.1. Crisis as a Breakdown of Coherence in the Corporate Context 51

2.3.2. Determining Factors: Analysis and Classification 54

2.3.3. Dysfunction and Corporate Crisis: Differences and Stages 57

2.3.4. Phases 58

2.3.5. Corporate Decline and its Quantitative Assessment 60

2.3.6. Evolution of Italian Scholarly Perspectives on Business Crisis & Insolvency 61
2.4. Insolvency & Financial Distress 62
2.4.1. Relationships and Differences: A Theoretical Evolution 63
2.4.2. The Concept of Equilibrium in Financial Distress and Stages of a Progressive Pathology 64
2.5. Bankruptcy 65
2.6. Key Concepts in Theory: Business Life Cycle Theory 66
2.7. The Reform of the Italian Insolvency System: A New Paradigm in Crisis Management 67
2.7.1. Cultural Shift and Path Forward 69
2.7.2. Between Gaps and Potentials 70
3. Seeing the Unseen: "*Predictive*" Key Concepts Explained 70
3.1. Some Considerations 71
4. Understanding Artificial Intelligence and Machine Learning in Predictive Studies: a Brief Guide 72
4.1. Machine Learning as Backbone of Predictive Analytics 73
4.2. Value Proposition in Corporate Crisis Management 74
5. Wrapping Up: Final Reflections and Forward Perspectives 74

Chapter 3

Predictive Analytics in Business Adversity: A Journey Through Theory and Methodology 75

Chapter Prelude 76

1. Turning Challenges into Opportunities: an Introduction 77
1.1. Insights for Strategic Realignment and Other Trends 77
1.2. Sustainability and Predictive Analytics for a Long-Term Realignment 78
1.2.1. An Holistic Ecosystem Approach 79
2. The Evolution of Corporate Crisis Prediction and Managerial Theories: From Traditional Financial Models to Modern Analytics 80
2.1. The Role of Management in Crisis Prediction 82
2.1.1. Early Views: the Mid-20th Century 82
2.1.2. The Impact of Financial Models on Management's Role in Crisis Prediction 82
2.1.3. Cross-Country Differences in Managerial Approaches 83

2.1.4. The Shift Toward Crisis Management as a "Managerial Responsibility" 84
2.2. State of the Art 84
2.2.1. The Italian Context 86
2.2.2. Conclusion: The Future of Insolvency Prediction in Italy 88
2.3. Rising Publication Trends on the Topic 89
3. Traditional Models, an Overview: the Altman Z-Score 94
3.1. Elements of Multivariate Discriminant Analysis 96
3.2. The Score 97
3.3. Some Considerations 99
3.4. Other Traditional Models 100
4. Innovative and "*Intelligent*" Predictive Analytics 102
4.1. An Emerging Paradigm: Intelligent Big Data Analytics (IBDA) 103
4.2. Business Intelligence and Managerial Literacy in Data-Driven Frameworks 105
4.3. IBDA in Insolvency and Crisis Prediction 106
4.4. Machine Learning Models 107
4.4.1. Shrinkage Methods: an Hint 109
4.4.2. Decision Trees 110
4.4.3. Random Forest 112
4.4.4. Logistic Regression 114
4.4.5. Boosting Methods 115
4.4.5.1. Gradient Boosting 116
4.4.5.2. XGBoost 117
4.4.5.3. Managerial Implications 119
5. Bridging AI Interpretability and Managerial Insights: Wrapping-Up and Final Considerations 120
5.1. Creating Trust and Accountability 121
5.2. SHAP in Literature 122

Chapter 4

***Data-Driven & AI-Driven Decision-Making:* an Empirical Investigation** 124

Chapter Prelude 125

1. Introduction 126

1.1. Preliminary Fundamental Considerations 127
1.2. Inductive Perspective 129
2. Research Design for Preliminary Analysis 130
2.1. Data Collection and Feature Selection 132
2.2. Data Preprocessing and Optimization 136
2.3. Descriptive Statistics and Correlation Analysis 137
3. Multi-Algorithmic Approach Implementation and Model Evaluation 141
3.1. Model Explanation and Results Interpretability 146
3.2. Conclusion and Remarkables 150
4. Research Design for Core Empirical Analysis: an Introduction 151
4.1. Setting the Stage 152
4.2. Data Collection and Feature Selection 153
4.3. Data Preprocessing and Optimization 158
4.3.1. A Critical Consideration 158
4.4. Descriptive Statistics and Correlation Analysis 160
5. Multi-Algorithmic Approach Implementation and Model Evaluation 163
5.1. Some Remarks 164
5.2. Starting Point and Models Deployment 166
5.2.1. Comparison and Key Insights 168
5.3. Confusion Matrix Results Evaluation 170
5.3.1. Confusion Matrix Results for SMEs 174
5.3.2. Confusion Matrix Dynamic Analysis 175
6. Model Explanation and Results Interpretability 177
7. Conclusion and Remarkables 181

Chapter 5

Conclusion: Key Findings and Implications 184

Chapter Prelude 185

1. Introduction 186
2. From Theoretical Framework to Empirical Evidence: The Research Trajectory 188
2.1. Structuring the Work: A Rigorous, Data-Driven Approach 189

2.2. A Balanced Method: From Passion to Analytical Rigor 190
3. Key Findings 190
3.1. Decision Support 193
3.2. AI-Driven Financial Management in Times of Crisis 194
4. Scientific and Practical Implications 195
5. Limitations and Future Directions 197
5.1. Data Limitation 197
5.2. Data Integrity and Value 198
5.3. Models Generalizability 198
5.4. Technological Constraints 198
5.5. Managerial Literacy and Competency 198
5.6. Ethical Concerns in the Age of AI and Effective Explainability of Models 199
5.7. Over-Reliance Risks 199
5.8. Cost and Resource Challenges 200
6. Future Research 200
7. Conclusion: AI and the Evolution of Management Practices 201

References 203

List of Tables

Table 1: Prediction Models Periods 14

Table 2: Determining Factors of Crises 56

Table 3: Composizione Negoziata della Crisi d'Impresa – Test Pratico 69

Table 4: Types of Documents 90

Table 5: Altman's & Ohlson's Models Comparison (Accuracy Performance) 101

Table 6: Core Analysis Selected Features 157

Table 7: Performance Assesment Using Classification Metrics of ML Models in Identifying in Crisis or Insolvent Italian Companies 167

Table 8: XGBoost confusion matrix (*Accuracy: 0.87, F1-Score: 0.88, AUC: 0.87*) 171

Table 9: SMEs XGBoost Confusion Matrix (*Accuracy: 0.8582996, F1-Score: 0.85, AUC: 0.86*) 175

List of Figures

Figure 1: Research Objectives 39

Figure 2: Key Differences - Firm-Level vs Macro-Level 49

Figure 3: Incoherences Types 53

Figure 4: Phases of Corporate Crisis 59

Figure 5: Business Life Cycle 66

Figure 6: A Comparative View of AI, ML, Deep Learning, and Generative AI 73

Figure 7: Annual Growth of Scientific Papers Regarding Insolvency and Crisis Prediction 90

Figure 8: Insolvency & Crisis Prediction Research Bibliometric Map (Used Keywords) 91

Figure 9: Network Visualization of Countries 93

Figure 10: Linear Discriminant Analysis 95

Figure 11: Altman Z-Score Areas 98

Figure 12: The "*Age of Trinity*" 104

Figure 13: Managerial Framework for IBDA 105

Figure 14: Random Forest Trees Illustration 114

Figure 15: Flow Diagram of Gradient Boosting ML Method 117

Figure 16: Differences Between Single Decision Tree and Ensemble 118

Figure 17: SHAP Values Explanation 120

Figure 18: Empirical Research's Fundamental Elements 128

Figure 19 & 20: Distribution Analysis 138

Figure 21 & 22: Relationship Between Debt to Equity Ratio and Debt to EBITDA Ratio Against the "*In-Crisis*" Status 139

Figure 23: Correlation Analysis 141

Figure 24: Random Forest Error Rate Graph 143

Figure 25: K-Means Clustering Results 147

Figure 26: Feature Importance in SHAP Values for Gradient Boosting Classifier 148

Figure 27: Violin Summary Plot 149

Figure 28: Distribution of Considered Companies by Region 161

Figure 29: Evaluation Metrics Synthesis 166

Figure 30: Violin Summary Plot 178

Figure 31, 32, 33: Distribution Plots for "*Revenues for Sales and Services Year -4"; "Number of Current Directors and Managers"; Total Production Costs Last Year"* 181

Glossary of Acronyms

AI – *Artificial Intelligence*

AIDA – *Analisi Informatizzata delle Aziende Italiane*

ATECO – *Attività Economiche*

AUC – *Area Under the Curve*

BA – *Business Analytics*

BI – *Business Intelligence*

CART - *Classification and Regression Trees*

CCII – *Codice della Crisi d'Impresa e dell'Insolvenza*

CDOs - *Collateralized debt obligations*

CEO - *Chief Executive Officer*

CNN - *Convolutional Neural Networks*

CSR – *Corporate Social responsibility*

CSRD - *Corporate Sustainability Reporting Directive*

DMI - *Digital Maturity Index*

EBITDA - *Earnings Before Interest, Taxes, Depreciation and Amortization*

ESG - *Environmental, Social, and Governance*

EU – *European Union*

GB – *Gradient Boosting*

GDPR – *General Data Protection Regulation*

GPT - *Generative Pre-trained Transformer*

GPU - *Graphics Processing Unit*

GPUI - *Global Political Uncertainty Index*

GRI - *Geopolitical Risk Index*

HR – *Human Resources*

IT – *Information Technology*

KDE - *Kernel Density Estimate*

KPIs - *Key Performance Indicators*

LDA - *Linear Discriminant Analysis*

LIME - *Local Interpretable Model-agnostic Explanations*

LR – *Logistic Regression*

LSE – *London School of Economics*

MBS - *Mortgage-Backed Securities*

MDA - *Multiple Discriminant Analysis*

ML – *Machine Learning*

OLS - *Ordinary Least Squares*

OOB - *Out-of-Bag*

PII – *Personally Identifiable Information*

QMUL – *Queen Mary University of London*

R&D – *Research & Development*

RF - *Random Forest*

RMBIPB - *Reference Model for Business Intelligence to Predict Bankruptcy*

RQ – *Research Question (s)*

RSS - *Residual Sum of Squares*

SDGs - *Sustainable Development Goals*

SGD - *Stochastic Gradient Descent*

SHAP - *SHapley Additive exPlanations*

SMEs – *Small and Medium-sized Enterprises*

SVM - *Support Vector Machines*

TMT – *Top Management Teams*

TRLs - *Technology Readiness Levels*

UN - *United Nations*

XAI – *Explainable Artificial Intelligence*

XGB - *Extreme Gradient Boosting*

STVDIVM VRBIS

"Forecasting is a notoriously underpaid profession, and extremely risky to boot, so I avoid it"

John Kenneth Galbraith

Personal Journey and Reflections on the Research Process

Embarking on this PhD journey has been much more than a purely academic endeavor—it has been a deeply personal and transformative experience. I found myself not only growing as a researcher but also evolving as a thinker, a strategist, and a professional. The challenges I encountered along the way—from grappling with intricate theoretical frameworks to interpreting vast datasets—served as invaluable lessons that have shaped my understanding of the world of management and business. Therefore, this dissertation is a reflection of not just the knowledge I've gained, but the persistence, curiosity, and creativity that fueled my research.

This journey has also been a powerful reminder of the importance of resilience, both in academic pursuits and in the broader context of business management; just as businesses must adapt and innovate in the face of uncertainty, so too must researchers push the boundaries of knowledge, embracing new ideas and methodologies while remaining grounded in rigorous analysis. The late nights spent refining models, the countless revisions of manuscript drafts, and the many discussions with mentors and peers were all part of a larger narrative—one of growth, discovery, and unwavering commitment to excellence. Moreover, this journey has been a testament to the power of collaboration: the insights and feedback from colleagues, the support from academic advisors, and the exchange of ideas at Conferences have all contributed to the richness of this research. These interactions have not only broadened my perspective but have also deepened my appreciation for the collective pursuit of knowledge that lies at the heart of academia.

As I reflect on this path, I am filled with a profound sense of gratitude for the opportunities and experiences that have come my way. This dissertation, while a culmination of years of hard work, is also a starting point—a foundation upon which I hope to build a career focused on contributing meaningful insights and driving innovation in my field.

Dissertation Style

In an era where businesses face unprecedented challenges, the ability to predict and manage financial distress has never been more crucial. This dissertation does more than merely explore this vital area; it redefines it, offering fresh insights and practical strategies that are both timely and transformative. From the outset, my research has been guided by a singular vision: to bridge the gap between theoretical frameworks and the practical realities faced by managers and decision-makers.

The stakes in today's business world are extraordinarily high, with crises often striking without warning and with devastating consequences. Therefore, this dissertation embarks on a comprehensive exploration of how predictive analytics, powered by the latest advancements in Artificial Intelligence (AI) and Machine Learning (ML), can be strategically integrated into corporate decision-making processes to mitigate risks and seize opportunities. Throughout this work, I have meticulously examined the historical evolution of crisis and insolvency prediction, drawing from foundational theories while also embracing cutting-edge technologies that offer new possibilities: the shift from traditional financial ratios to advanced predictive models reflects a broader trend in the field—one that emphasizes not just the ability to react to crises, but to anticipate and prevent them. This is where the true value of predictive analytics lies, and it is a central theme that runs through the entire dissertation. The work presented here is then not merely theoretical; it is deeply rooted in empirical research and practical applications.

Over the course of my candidature, I had the opportunity to publish papers in esteemed academic journals and presented my findings at national and international Conferences. These contributions are not just milestones in my academic journey; they are integral parts of this dissertation, reflecting the evolution of my ideas and the impact of my research on the broader academic community. Including parts and considerations from these published papers and Conference materials within this Thesis ensures that the insights presented are both validated and relevant to real-world challenges; moreover, the Thesis took shape through focused research periods at prestigious, internationally renowned institutions such as Queen Mary University of London (QMUL) and the London School of Economics (LSE), where immersive academic environments and collaborative efforts further enriched the depth and scope of my work. Each Chapter represents a critical juncture in my research journey, reflecting the continuous process of learning, adaptation, and innovation. The methodological approaches adopted have been carefully selected to align with the complex nature of the research questions, ensuring that the findings are robust and applicable to the dynamic business environment. The philosophical stance of this dissertation is then rooted in pragmatism: while it employs rigorous experimental methods to test hypotheses and validate models, it also embraces the flexibility necessary to explore diverse perspectives and interdisciplinary connections. This approach has enabled me to maintain a balance between methodological consistency and the adaptability required to address real-world problems.

In terms of style, this work is crafted to engage the reader through a narrative that is both informative and compelling. The use of both first- and third-person perspectives throughout the dissertation has been a deliberate choice, allowing for personal reflection on my research journey while maintaining the objectivity necessary for academic discourse: this blend of perspectives helps to convey the depth and breadth of the research, making it accessible to a wide audience while maintaining the rigorous standards of academic inquiry.

A Critical Partnership in Decision-Making: Prelude

As the path unfolded during the writing, it became increasingly apparent that the intersection between AI and human decision-making is not only a central theme in the academic world but also reflects the very challenges facing modern businesses. AI, much like the tools and frameworks I relied on throughout my research, provides immense value in processing data, offering predictive insights, and crafting strategies. However, just as my academic progress was shaped by feedback, critical reflection, and iterative improvement, AI can only serve as an enabler—it remains (today) the human decision-maker who must ultimately interpret, adapt, and apply those insights in practice. So, much like in the academic setting, where deep collaboration and intellectual engagement shaped my thinking, the role of human insight in crisis and insolvency prediction is indispensable. AI-driven models create powerful frameworks for anticipating financial distress, but these tools cannot function in isolation. They offer a map, but it is the manager who must navigate the terrain, balancing algorithmic recommendations with an understanding of context, ethics, and the broader business landscape. This mirrors my own research process, where each new dataset, method, or theoretical framework was refined not simply through technical application but through the human processes of analysis, critique, and revision.

The "*black box*" nature of many innovative models further underscores the need for transparency and human oversight; while AI excels at identifying patterns and forecasting outcomes, it is vital that managers maintain visibility into how decisions are made. Just as I had to dig deeper into the assumptions behind my theoretical models and research methods, corporate leaders must engage critically with AI outputs, ensuring that they understand not just the results but the reasoning behind them. This "*human in the loop*" approach is essential in both academic research and business management, as it determines accountability, creativity, and adaptability in decision-making. Moreover, this reflects a broader theme in both my academic journey and in AI: the need for courage in decision-making. In fact, whether it was refining my dissertation in the face of setbacks or revisiting research questions, my experience parallels the role of managers who must make bold decisions, often in the face of uncertainty. AI may provide predictive insights, but it is the human strategist who must take responsibility for actions, embracing both the potential for success and the risks of failure.

Introduction and Context

In today's fast-paced and unpredictable business environment, the ability to anticipate and manage crisis and/or insolvency is more crucial than ever for firms. Companies face numerous challenges, from economic downturns and market volatility, to technological disruptions and evolving regulatory landscapes (Ferri et al., 2022; Floridi et al., 2018). In this context, accurate predictions are not just a tool for survival; they are strategic assets that can safeguard stakeholders' interests, ensure operational continuity, and enhance long-term sustainability (lo Conte & Sancetta, 2024; Sun et al., 2014). However, understanding corporate crisis, bankruptcy, and insolvency is essential for addressing financial distress effectively. A corporate crisis is a significant disruption that threatens the viability of an organization, arising from operational failures, strategic errors, or external shocks such as economic downturns or regulatory changes (Beaver, 1966). Bankruptcy, on the other hand, is a legal state where a company cannot meet its debt obligations and seeks relief through court intervention, while insolvency is the financial condition preceding bankruptcy, characterized by liabilities exceeding assets or an inability to pay debts as they come due (Sciarelli, 1995; Altman, 1968; Beaver, 1966). Consequently, the methodologies for predicting corporate crisis and insolvency have evolved significantly over the years (Jiang et al., 2020; Sun et al., 2014): traditional approaches, such as univariate and multivariate analyses (Altman, 1968; Beaver, 1966), relied heavily on financial ratios and static historical data to predict future outcomes. Although they provided valuable insights, their predictive power was limited by their inability to account for the complex and dynamic nature of modern financial markets (Ferri et al., 2022; Sun et al., 2014; Altman, 1968). For example, early models, such as those developed by Beaver (1966) and Altman (1968), laid a quantitative foundation for predicting financial distress using statistical analysis of financial ratios. However, they had limitations in their scope and adaptability to dynamic market conditions (Altman, 1968; Beaver, 1966).

With advancements in technology, the field has shifted towards more sophisticated techniques, leveraging Machine Learning (ML) and Big Data analytics (Aliaj & Piersanti, 2020; Vinuesa et al., 2020; Barboza et al., 2017). These contemporary approaches can handle vast datasets, identify intricate patterns, and adapt to new information in real-time, offering a significant improvement over traditional models (Lo Conte & Sancetta, 2024; Zhang et al., 2021; Sun et al., 2014). For instance, while Altman's Z-score and similar models laid important groundwork, modern AI-driven models enhance these by incorporating a broader range of variables, including non-financial indicators like social media sentiment, macroeconomic factors, and geopolitical events. Furthermore, the integration of Artificial Intelligence (AI) and Machine Learning into insolvency prediction models holds immense promise: these technologies not only enhance the accuracy and precision of predictions by capturing complex interactions among variables and analyzing extensive datasets that traditional methods struggle to handle, but also offer real-time insights and adaptive learning capabilities (Zhang et al., 2021; Vinuesa et al., 2020). These capabilities, in fact, represent a transformative advancement in the realm of predictive analytics as they enable AI systems to continuously learn from new data, dynamically updating their predictive models to reflect changing conditions and emerging patterns (Goodfellow et al., 2014). This is particularly crucial in the context of insolvency and crisis prediction, where economic landscapes, market conditions, and organizational and financial factors can shift rapidly (Salchenberger et al., 1992; Altman, 1968; Beaver, 1966). By leveraging adaptive learning, AI systems can provide more timely and relevant predictions, allowing businesses to anticipate and respond to potential crises before they escalate: a continuous learning process ensures that predictive models remain accurate and effective over time, accommodating the evolving nature of financial environments and organizational behaviors (Jiang et al., 2020).

In the dynamic landscape of 2024, where the pursuit of precision is both within reach and extraordinarily challenging, the role of adaptive learning becomes then paramount. The complexity and volatility of modern financial environments demand a level of agility and responsiveness that traditional static models simply cannot provide. AI systems, through their mechanisms, offer a distinct advantage by continuously updating and recalibrating their algorithms based on real-time data (Zhang et al., 2021). This ensures that predictions are not only relevant to the current context but are also progressively more precise as the system learns from past inaccuracies and successes. The efficacy of these predictive models is further enhanced by their ability to handle vast and diverse datasets, and their adaptive nature supports proactive rather than reactive management strategies. Businesses, in fact, can use these insights to implement preemptive measures, thereby mitigating risks before they develop into full-blown crises (Bottiglia & Paltrinieri, 2010). This shift from a reactive to a proactive stance represents a transformative change in strategic management practices, fostering a culture of foresight and resilience (Floreani et al., 2013): businesses can rely on these systems for more informed decision-making, proactive risk management, and strategic planning (Nguyen et al., 2019). This signifies a move away from static, one-size-fits-all models to dynamic systems that evolve with the business environment. Finally, this adaptability is a critical component in managing uncertainty and fostering sustainable growth, making AI an indispensable tool in contemporary strategic management (Brynjolfsson & McAfee, 2017).

All of this allows for proactive risk management and early intervention strategies, significantly improving the ability to foresee and mitigate potential crises (Zhang et al., 2021; Vinuesa et al., 2020). From a managerial perspective, AI serves as an invaluable support to human judgment, empowering managers with advanced tools to make informed, data-driven decisions (Dasilas & Rigani, 2024; Aly et al., 2022). In the future, we can expect intelligent systems to provide even more sophisticated solutions for advanced and proactive management of crises and performance, ensuring sustainability and resilience in an increasingly complex and dynamic economic landscape. This synergy between AI and human expertise then promises a new era of strategic management, where technology amplifies human capabilities to drive innovation and secure business stability (Ribeiro et al., 2016). However, the deployment of AI and ML is not without challenges: high-quality data is essential, as the effectiveness of AI models heavily depends on the quality and completeness of the data they are trained on (Chen, 2011). This challenge is particularly acute for small and medium-sized enterprises (SMEs), which may lack the resources and expertise to collect and manage large datasets. Additionally, the interpretability of AI models presents another critical issue. Unlike traditional statistical models, AI algorithms, especially deep learning techniques, often operate as "*black boxes*", making it difficult for managers to understand how predictions are generated (lo Conte et al., 2024; Ribeiro et al., 2016).

The pervasive use of *"new intelligences"*, today, also raises significant ethical considerations (Zhang & Susheng, 2021). AI, in fact, should not merely function as an automatic instrument; it should support human decision-making by enabling more accurate and informed decisions for managers, policymakers, accountants, controllers, and academia. This collaborative approach ensures that it complements rather than replaces human expertise, fostering a symbiotic relationship where technology and human insight coalesce to drive superior outcomes. New technologies certainly allow for more agile and responsive management practices, can automate routine tasks such as data entry and reconciliation, reducing the risk of human error and freeing up professionals to focus on higher-value activities such as strategic financial planning and analysis, and, from a merely academic perspective, can advance the frontiers of knowledge in fields such as economics, finance, and business management. However, over-reliance on AI can lead to complacency, where decision-makers blindly trust AI outputs without applying their own judgment and critical thinking, which can be particularly dangerous in

crisis situations where subtle understanding and human intuition are essential. Moreover, the implementation and maintenance of AI systems can be costly and resource-intensive, particularly for SMEs, and can be exploited for malicious purposes, such as manipulating financial markets, spreading misinformation, or invading privacy.

Therefore, this study aims to foster a more precise and unique relationship between the human side of firms and new technologies, promoting ethical leadership and responsible AI deployment (Lo Conte & Sancetta, 2023), bearing in mind that ethical implications of AI in business extend to issues of fairness, accountability, and transparency. This last element, in particular, is essential to build trust among stakeholders (Conte et al., 2023): robust governance frameworks and regular audits can help mitigate risks such as moral hazard and opportunism, where managers might manipulate data or engage in risky behaviors to align with AI-generated predictions (Floridi et al., 2018). The integration of AI in corporate management also holds significant potential for promoting sustainability. AI can enhance business ethics and sustainability by optimizing resource use, reducing waste, and improving operational efficiency. These improvements align with both national and global sustainability goals, such as those outlined in the United Nations Sustainable Development Goals (*SDGs*) and the Corporate Sustainability Reporting Directive (*CSRD*)[1]. In the context of global value chains, particularly involving developing countries, AI can play a critical role in ensuring ethical practices: intelligent systems can monitor compliance with labor standards and environmental regulations, ensuring that suppliers adhere to ethical practices. This not only benefits workers and communities in these regions but also enhances the overall sustainability of business operations (Vinuesa et al., 2020).

Beyond predicting crises, enhancing managerial capabilities in crisis management is crucial. The goal, in fact, is not to hide problems but to detect early signals, reallocate resources efficiently, and mitigate adverse impacts on stakeholders (Nguyen & Han, 2019). By fostering a proactive approach, managers can preserve and even enhance corporate value during periods of dislocation. For instance, a firm anticipating a potential liquidity crunch can proactively secure additional financing or renegotiate terms with creditors, thereby avoiding insolvency (Mitroff, 2005). In response to these issues, this study underscores AI's potential to revolutionize the concept of going concern: a proactive approach to risk management ensures that businesses can maintain continuity and stability even in the face of unforeseen challenges (Davenport & Ronanki, 2018). For example, SMEs, characterized by their smaller scale of operations and limited resources, face unique challenges and opportunities in crisis and insolvency management[2]. Their size often allows for quicker decision-making and implementation of changes, which can be advantageous in times of crisis; however, they typically have less financial cushion to absorb shocks, making them more vulnerable to economic downturns and disruptions. Predictive models can provide SMEs with the foresight needed to plan strategically for the future and by identifying potential risks early, these firms can develop contingency plans and allocate resources more effectively to mitigate these risks.

The models proposed in this study offer a cost-effective solution by providing actionable insights that help firms prioritize their risk management efforts. In particular, AI-driven predictive models can help SMEs streamline their operations by identifying inefficiencies and suggesting

[1] The Corporate Sustainability Reporting Directive (CSRD) is a legislative framework introduced by the European Union under Directive (EU) 2022/2464. It aims to enhance and standardize corporate sustainability reporting across member states. The CSRD mandates comprehensive disclosure of non-financial and sustainability information by large companies and certain public-interest entities, focusing on environmental, social, and governance (ESG) criteria. The directive seeks to improve transparency, accountability, and comparability of sustainability data, supporting responsible business practices and contributing to the EU's broader sustainability and climate objectives.

[2] The survival of struggling firms can negatively impact the broader production system, affecting prices, competition, and resource allocation efficiency.

improvements, leading to significant cost savings and increased productivity. By monitoring key performance indicators (*KPIs*) such as throughput, cycle time, and utilization rates, these firms can continuously improve their operational processes and achieve higher efficiency levels (Davenport & Ronanki, 2018). This proactive approach can lead to long-term growth and competitive advantage. Given the relatively simple organizational structures of SMEs, AI can achieve high levels of pervasiveness within these firms, allowing for the efficient streamlining of costs, literacy of all business elements and stakeholders, and the creation of cost-effective platforms for monitoring performance[3]. Consequently, AI, when used alongside human expertise, can be a powerful catalyst for small firms (Zhang & Susheng, 2021). Furthermore, this synergy also allows businesses to maximize their limited resources, enhance decision-making processes, and drive innovation.

Research Objectives Overview

The present study aims to achieve several key objectives, which are intricately linked to the aforementioned issues. In particular, it is essentially divided into two "*macro-parts*." The first part provides a comprehensive excursus of the literature, examining both traditional and innovative & intelligent models for predicting corporate crises and insolvency. This section investigates the theoretical underpinnings, practical implications, and advancements in predictive analytics, highlighting the evolution from classical statistical methods to contemporary approaches. The second part is empirical, offering practical validation of the theories, considerations, and implications discussed earlier. Through rigorous analysis of real-world data, the study provides concrete evidence supporting the efficacy of advanced predictive models in improving crisis management and insolvency prediction, particularly for SMEs. Then, the empirical investigation not only tests the theoretical constructs but also demonstrates their applicability and effectiveness in actual business scenarios.

By integrating these two parts, the study aims to advance the literature on the subject, adopting a critical yet scientifically rigorous approach to both theoretical and practical advancements in the field. Furthermore, the research underscores the importance of ethical leadership and responsible AI deployment, fostering a symbiotic relationship between human decision-making and AI support. In doing so, it seeks to contribute significantly to academic discourse while providing actionable insights that enhance business resilience, sustainability, and ethical governance. However, this work also seeks to provide a comprehensive guide for managers to utilize predictive insights for strategic decision-making and addresses ethical considerations while enhancing the explainability of predictive models. Additionally, it aims to bridge the gap between academia and real-world applications: by leveraging AI-based algorithms and Machine Learning techniques such as Logistic Regression, XGBoost, and Gradient Boosting, in fact, this Thesis aims to provide a deep understanding of the advantages and limitations of these innovative approaches. Lastly, the interdisciplinary and holistic approach ensures that the study is both scientifically rigorous and critically reflective, bringing added value to the existing body of research.

[3] The deployment of AI-driven mechanisms within SMEs, characterized by their lean organizational structures, facilitates an elevated pervasiveness of automation and process optimization. Such integration enables not only the cost-effective minimization of operational redundancies through advanced machine learning algorithms but also enhances cross-functional literacy across all business dimensions, including supply chain logistics, financial forecasting, and customer relationship management (CRM) metrics. Furthermore, the establishment of robust, data-driven platforms for real-time performance monitoring leverages predictive analytics and edge computing, thereby enabling precise benchmarking and dynamic adaptability in response to fluctuating market variables, ultimately supporting a scalable architecture for long-term strategic growth.

Chapters Outline

The structure of this dissertation is meticulously crafted to provide an in-depth examination of the multifaceted aspects of corporate crisis and insolvency prediction. Each Chapter is thoughtfully organized to offer a coherent and comprehensive analysis, ensuring a thorough understanding of the topic. In addition to providing a comprehensive and rigorous analysis of the literature, along with novel and interesting empirical implications, this Thesis includes references and materials derived from papers and research products developed during the doctoral journey. These contributions have been published in Journals, Books, and presented at national and international Conferences. Consequently, the works have been already validated and shared with the scientific community. This Thesis then represents a critical, rigorous, and insightful synthesis of this path, offering numerous future directions on a deeply examined topic that is currently undergoing further scientific investigation and development. The Chapters are as follows:

- *Chapter 1:* This Chapter introduces the foundational and distinctive aspects of predictive analysis in corporate crisis and insolvency. It concludes with a detailed overview of the Thesis's objectives and Research Questions, setting the stage for the work ahead;

- *Chapter 2:* Serving as a cornerstone, this Chapter explores the specific terminology and key concepts that frame the entire Thesis, offering a clear understanding of the foundational ideas that are unpacked throughout the work;

- *Chapter 3:* It provides an in-depth examination of the technical and empirical aspects of insolvency prediction. It closely aligns with the relevant literature and theoretical frameworks, connecting rigorous analysis with established academic references;

- *Chapter 4:* Here, the practical and empirical outcomes of the research come to full fruition. This Chapter substantiates the previously discussed perspectives and offers tangible insights on AI's achievable impact in decision-making. Special focus is given to the explainability of results, an AI-driven approach, and tailored solutions;

- *Chapter 5:* Acting as a synthesis, this Chapter revisits the insights and results, with a focus on scientific and managerial implications. It highlights research limitations and proposes ways to bridge the gap between research and practical application in the field.

Chapter 1

Historical Foundations and Future Directions in Corporate Crisis and Insolvency Prediction: a Beginning

Chapter Prelude

Weaving the wisdom of the past with the innovation of the future, this Chapter embarks on a journey through the evolving landscape of corporate predictive analytics. By integrating historical foundations with cutting-edge methodologies, it set the stage for strategic insights that will guide businesses through critical situations. In the end, this Chapter will shine a spotlight on the study's goals and ambitions, charting a path for a research journey that seeks to deepen how we understand and predict corporate crises and insolvency. It's where the blueprint of this exploration takes shape, promising insights that could reshape the future.

1. Background and Significance of the Study

This introductory Chapter will establish the foundational framework essential for comprehending the broader scope and implications of this PhD Thesis. This section serves as a gateway into the critical exploration of corporate crisis and insolvency prediction, a field that has been integral to both academic research and practical management for decades. In fact, the ability to accurately forecast financial distress and insolvency is not only vital for maintaining the stability of individual businesses but also for ensuring the integrity of the global economic system. By tracing the historical evolution of predictive methodologies—from the reliance on traditional financial ratios to the integration of advanced techniques—this work aims to underscore the transformative impact of these technologies on strategic decision-making. Furthermore, this Chapter will explore the managerial implications of adopting predictive analytics, highlighting how these tools are reshaping the landscape of corporate strategy and operational management.

As AI-driven predictive analytics becomes increasingly embedded in business practices, the role of managerial literacy in these technologies becomes also paramount: understanding and effectively utilizing predictive insights are now critical competencies for managers, influencing everything from risk management to long-term strategic planning. The literature provides a wealth of insights into these developments, offering a rich backdrop against which this Thesis will critically analyze the current state and future directions of predictive analytics in corporate crisis management. In essence, this Chapter lays the groundwork for the entire study, situating the research within a broader managerial and academic context and outlining the pivotal role that predictive analytics and AI play in modern business strategy.

1.1. Historical perspectives: a brief run-through

The study of corporate crisis and insolvency prediction has long been a cornerstone of strategic analysis, playing a critical role in the stability and integrity of the global economic system[4]. The ability to predict, anticipate and manage insolvency with accuracy and timeliness is of paramount importance for a myriad of stakeholders, including managers, investors, regulators, and policymakers (Altman et al., 2017), and this PhD Thesis embarks on a comprehensive exploration of the historical evolution and future directions of prediction methodologies, emphasizing the transition from traditional financial ratios to advanced ML & AI-based techniques. Historically, the prediction of corporate insolvency has been grounded in financial ratio analysis, with seminal works by Altman (1968) and Beaver (1966) laying the foundation for subsequent models[5]. These traditional approaches relied heavily on static financial metrics such as liquidity ratios, profitability ratios, and leverage ratios to gauge a firm's

[4] Historically, corporate crisis was considered an exceptional and negative event, akin to a business pathology, prevalent up until the 1970s. Over recent decades, it has become apparent that crises are not isolated incidents but rather a consistent feature within modern industrial systems (Mitroff, 2005). Crises can manifest in various forms:

- *Widespread Crises:* Affect entire sectors or parts thereof due to operational environment changes and pre-existing imbalances;
- *Particular Crises:* Impact individual firms that were once efficient and well-managed but now face reduced profitability due to specific vulnerabilities in response to external environmental changes (Altman et al., 2017).

Corporate crises are now recognized as a structural and pervasive phenomenon, intimately linked to economic instability factors such as prolonged economic downturns, increased market competition, and the influence of emerging markets. Many companies adopt short-term financial policies to resolve liquidity issues, which, coupled with high short-term debt, can expedite crises when profitability declines or interest rates rise.

[5] It's important to note that these models were developed in a different economic context and their applicability can vary across different industries and time periods.

financial health (Ohlson, 1980). While they provided a useful starting point, these models were often limited by their inability to capture the dynamic and multifaceted nature of modern financial environments (Aydin et al., 2022)[6]. For instance, Altman's Z-Score model, which became a benchmark in the field, utilized a linear combination of five financial ratios to predict the probability of bankruptcy (Altman et al., 2017; Altman, 1968). Despite its widespread use and robustness in various contexts, the model and similar methodologies struggled to adapt to the increasingly complex and volatile financial markets of the late 20th and early 21st centuries (Bellovary et al., 2007; Ohlson, 1980). These limitations are rooted in their static nature and reliance on historical data that may not adequately reflect current or future market conditions (Aydin et al., 2022). Financial ratios derived from historical balance sheets can, in fact, fail to account for the rapid changes in market dynamics, regulatory environments, and technological advancements that characterize the business landscape (Wilson & Sharda, 1994). As such, these models often provide a retrospective rather than a predictive view, limiting their effectiveness in real-time crisis management and decision-making (lo Conte et al., 2024; Altman et al., 2017; Bellovary et al., 2007)[7].

Over time, the field has witnessed a gradual shift towards more sophisticated techniques, incorporating statistical methods and, more recently, intelligent algorithms (Elhoseny et al., 2022). Machine Learning, for example, with its capacity to process vast datasets and uncover intricate patterns, has significantly enhanced the accuracy and precision of predictions (lo Conte & Sancetta, 2024; Nguyen et al., 2019), and marked a significant turning point in the evolution of prediction methodologies and studies (Bellovary et al., 2007)[8]. These techniques, such as decision trees, random forests, support vector machines, and neural networks, for this reason, have been employed not only to enhance predictive accuracy, but also to improve model robustness, making them more suitable for dynamic and volatile financial environments (Barboza et al., 2017; Agarwal & Taffler, 2008)[9].

The estimation of insolvency and crisis probability then represents a fundamental pillar in the financial and economic domains, with far-reaching impacts across various actors and sectors. Predicting insolvencies is obviously an intriguing topic, yet the creation of highly accurate predictive models remains a challenge, as these models often specialize in specific sectors and tend to be influenced by the characteristics of the data sample on which they are trained (Zhang et al., 2021; Vinuesa et al., 2020). For financial institutions, credit risk assessment is crucial. Insolvency estimates are used to evaluate the likelihood that a debtor will default on their payment obligations. These evaluations directly influence

[6] The multifaceted nature of modern financial environments encompasses various elements across different contexts. For example, fluctuating global markets can drastically impact financial stability, as seen in the 2008 financial crisis and its aftermath (Reinhart & Rogoff, 2009). Technological innovations in fintech, such as blockchain and cryptocurrencies, introduce new dimensions of risk and opportunity that traditional models struggle to account for (Gomber et al., 2017). Regulatory changes, like the implementation of the Sarbanes-Oxley Act in the United States, have profoundly affected corporate governance and financial reporting practices (Coates, 2007). Additionally, geopolitical events such as Brexit have added layers of complexity to financial predictions, impacting trade, investment, and economic policy in the UK and EU (Dhingra et al., 2017).

[7] This backward-looking approach can be informative but falls short in predicting future crises, as it does not account for real-time changes and emerging risks (lo Conte et al., 2024). For example, by focusing on historical balance sheets, these models may miss early warning signs of financial distress that are only detectable through current market data or forward-looking indicators such as market sentiment or geopolitical shifts (Beaver et al., 2005). In contrast, predictive models aim to anticipate future events by continuously updating with new data, thereby enhancing the ability to manage crises proactively and make timely decisions (Aydin et al., 2022).

[8] Historically, ML began to gain traction in predictive analytics during the late 20th century, with its applications spanning various domains, including finance, healthcare, and marketing (Goodfellow et al., 2016). In the realm of financial distress and insolvency prediction, ML techniques started to be employed more extensively in the early 2000s, offering a robust alternative to traditional statistical models (Agarwal & Taffler, 2008; Barboza et al., 2017).

[9] Support Vector Machines (SVM) and Random Forests (RF) further advanced the field by providing high-dimensional data handling capabilities and improved predictive performance (Shin & Lee, 2002; Breiman, 2001). For instance, RF models, with their ensemble learning approach, combined multiple decision trees to reduce overfitting and enhance generalizability, making them particularly effective in crisis prediction scenarios (Barboza et al., 2017).

decisions on loan issuance, capital management, and investment strategies (Altman et al., 2017; Beaver, 1966). Similarly, companies and financial institutions managing diversified credit portfolios must understand the overall risk associated with these portfolios and take preventive measures to mitigate potential losses. Investors, both individual and institutional, also consider insolvency estimates carefully in their investment decision-making processes. These estimates influence the assessment of expected returns and the level of risk associated with financial instruments (Aly et al., 2022). Moreover, companies must make critical decisions regarding the extension of credit to their customers, with the probability of insolvency determining whether and how to extend credit, payment terms, and credit limits (Mitroff, 2005). Finally, the estimation of insolvency probability helps companies to detect early signals of a potential financial crisis, allowing them to proactively take corrective actions to avoid insolvency or, if necessary, to initiate restructuring processes (Jiang et al., 2020).

From a merely historical and academic perspective, the analysis of prediction models dates back to the 1930s and can be divided into four main periods[10], each characterized by a different methodological approach and increasing complexity in models:

- *First Period (1930s):* During this period, researchers focused on analyzing individual financial indicators of distressed firms, calculating their average values. These values were then compared with the values of the same indicators for each firm in the sample, highlighting trends that differentiated healthy companies from those in crisis. However, this method did not allow for the evaluation of the predictive capacity of individual indicators. In 1966, William Beaver introduced the binary classification test to assess the predictive effectiveness of indicators, paving the way for the simultaneous use of multiple indicators (Beaver, 1966);

- *Second Period (1968):* Initiated with the article by E. Altman, this period saw the introduction of the multivariate model of linear discriminant analysis (LDA)[11]. This approach spread and was applied to companies of various sizes and sectors. Significant contributions were made by Alberacci and Appetiti, who focused on predicting the insolvency of small and medium-sized Italian enterprises. In particular, Alberacci aimed to develop a model capable of predicting insolvency well in advance (Altman, 1968)[12];

- *Third Period (1980):* Characterized by the use of probabilistic models, such as the logistic function and the probit function, to measure the probability of business failure rather than classifying firms as either distressed or healthy. During this period, there was a

[10] 4 main plus recent developments.

[11] The multivariate model of linear discriminant analysis (LDA) is a statistical method used to predict group membership based on multiple predictor variables. It is particularly useful in contexts where the goal is to distinguish between two or more classes or groups, such as predicting financial distress or insolvency in companies (Altman, 1968). The key steps involved in LDA are:

1. Calculating the Means and Covariance Matrices: The process starts by calculating the means and covariance matrices of the predictor variables for each group.
2. Forming the Discriminant Function: LDA forms a discriminant function (or a set of functions, in the case of multiple groups) that is a linear combination of the predictor variables. This function maximizes the separation between the groups.
3. Classification: New observations are classified based on the value of the discriminant function. The classification is done by assigning the observation to the group with the closest mean on the discriminant function scale.

[12] Alberacci's methodologies have included the adaptation and refinement of traditional predictive models to better suit the financial realities of SMEs. For instance, while earlier models like Altman's Z-Score were highly effective for larger firms, they did not always capture the nuances of smaller enterprises. Alberacci addressed this gap by incorporating a broader range of financial and non-financial indicators into the predictive framework, enhancing the models' accuracy and reliability.

comparison between the predictive effectiveness of logistic analysis and discriminant analysis, but the results did not clearly indicate the superiority of one over the other (Zmijewski, 1984; Ohlson, 1980)[13];

- *From 1990 onwards:* There was an increase in the use of models based on neural networks, which demonstrated a high degree of accuracy in predictions[14]. These advanced models benefited from the ability to process larger datasets, significantly enhancing the predictive capabilities of insolvency models (Altman et al., 2017; Barboza et al., 2017);

- *Recent Developments* (*Intensive Use of AI and ML from 2020s*): The recent period has been marked by the extensive integration of Artificial Intelligence (AI) and Machine Learning (ML) in predictive studies regarding crisis and insolvency. These advanced methodologies allow for the continuous updating of models as new data becomes available, making them highly adaptable to changing market conditions and economic environments (lo Conte et al., 2024). Moreover, AI and ML can incorporate a wide range of different features (Zhang et al., 2021; Jiang et al., 2020).

Table 1 offers a brief and compehensive view of these periods.

Table 1: Prediction Models Periods

Period	Main Features
First Period (1930s)	Analyzed individual financial indicators; Introduced binary classification test by Beaver (1966).
Second Period (1968)	Introduction of multivariate model of linear discriminant analysis (LDA) by Altman (1968); Contributions by Alberacci and Appetiti in the Italian context.
Third Period (1980)	Use of probabilistic models (logistic function, probit function); Comparison between logistic analysis and discriminant analysis.
From 1990 onwards	Increase in use of neural networks; High accuracy in predictions; Ability to process large datasets and identify complex patterns.
Recent Developments (from 2020s)	Extensive integration of AI and ML; Continuous updating of models; Adaptability to changing market conditions; Incorporation of wide range of features.

Source: own elaboration

In conclusion, the evolution of corporate insolvency and crisis prediction models from early financial ratio analysis to sophisticated AI and ML techniques highlights significant academic and managerial implications. This progression, in fact, not only enhances our theoretical understanding of financial distress but also provides powerful tools for strategic decision-making in business environments marked by uncertainty and rapid change. From an academic perspective, the shift

[13] Ohlson (1980) developed a logistic regression model that included various financial ratios and company-specific variables to predict the likelihood of bankruptcy. Zmijewski (1984) extended this work by incorporating macroeconomic indicators, thus broadening the scope and applicability of insolvency prediction models. This period also saw the development of early warning systems, which aimed to provide timely alerts regarding potential financial distress, thus allowing for preemptive managerial actions.

[14] The advent of neural network models in the early 1990s introduced a transformative approach to crisis and insolvency prediction, offering a level of predictive accuracy and adaptability previously unattainable with traditional statistical methods. Neural networks, inspired by the human brain's structure and functioning, are capable of modeling complex, non-linear relationships within data, making them particularly well-suited for the multifaceted nature of financial distress.

towards incorporating AI and ML in predictive models has opened new avenues for research. These advanced methodologies enable researchers to explore complex, nonlinear relationships between variables that traditional models could not capture (Barboza et al., 2017; Agarwal & Taffler, 2008). This has significant implications for the development of new theories and models that better reflect the realities of modern financial environments. Moreover, the continuous learning capabilities of AI models allow for real-time updates and improvements, ensuring that predictions remain relevant as new data becomes available (Zhang et al., 2021; Jiang et al., 2020)[15].

For managers, the integration of innovative approaches into predictive models offers substantial strategic advantages. By leveraging these advanced tools, in fact, managers can gain deeper insights into the financial health of their organizations and the broader market environment. This enables them to anticipate potential crises and take proactive measures to mitigate risks before they escalate (Vinuesa et al., 2020). Furthermore, the application of AI and ML in insolvency prediction aligns with the broader trend towards data-driven decision-making in business (Narvekar & Guha, 2021)[16]. This helps to prevent the cascading effects of financial crises, which can have severe economic and social consequences (Altman et al., 2017; Dellepiane et al., 2015; Beaver, 1966)[17]. Moreover, by providing a more reliable assessment of credit risk, these models support more efficient capital allocation across the economy, fostering sustainable growth and development (Mitroff, 2005). In addition to these practical benefits, the integration of AI and ML in insolvency prediction raises important ethical considerations. As these technologies become more pervasive, ensuring their responsible and transparent use is crucial. This involves addressing issues related to data privacy, algorithmic bias, and the potential for misuse in financial markets (Floridi et al., 2018; Zhang & Susheng, 2021).

Looking ahead, the continued evolution of predictive models will likely involve further integration of AI and ML with other emerging technologies, such as blockchain and quantum computing[18]. For instance, a study by Li et al. (2023) discusses how blockchain can enhance the transparency and security of AI-driven predictive models by providing immutable and verifiable data records[19]. Additionally, Patel and Zhang (2022) explore the synergy between quantum computing and machine learning, suggesting that quantum algorithms could significantly boost the computational efficiency of predictive models, allowing for faster and more accurate predictions. These innovations have the potential to enhance the accuracy and efficiency of predictive models even further, providing new tools for managing financial distress in an increasingly complex and interconnected global economy (Elhoseny et al., 2022)[20]. For instance, blockchain technology could improve the transparency

[15] They can also handle missing or incomplete data effectively, enhancing the reliability of predictions in real-world business contexts.

[16] As organizations adopt AI and ML for crisis prediction, they are moving towards a proactive, data-driven approach that aims to save lives, safeguard assets, and create a more resilient future. The effectiveness of this shift hinges not just on the technology but also on the ethical and strategic frameworks guiding its use (Smith, 2020).

[17] For example, predictive models can help regulators identify and address fraudulent activities early, protecting consumers and maintaining trust in the financial system; furthermore, businesses can use AI to forecast disruptions in supply chains, enabling them to take preemptive measures and avoid cascading failures that could affect multiple industries (Dellepiane et al., 2015).

[18] Quantum computing can enhance insolvency prediction by rapidly analyzing complex financial datasets, such as transaction histories and market trends, identifying early warning signs like sudden liquidity shifts and unusual trading patterns that traditional models might overlook.

[19] Some examples include:

- *Supply Chain Management:* Blockchain can track each transaction and shipment, ensuring the data fed into AI models is accurate and tamper-proof, leading to better demand forecasting and inventory management;
- *Healthcare:* Patient data stored on a blockchain can be securely accessed and verified, allowing AI models to predict disease outbreaks or patient health trends with confidence in the data's integrity;
- *Financial Services:* Blockchain can ensure the accuracy of transaction records used by AI for fraud detection, reducing false positives and increasing the model's reliability.

[20] Elhoseny et al. (2022) explore the transformative effects of advanced technologies such as AI and quantum computing on the interconnected global economy. The study emphasizes that these technologies significantly improve the accuracy and efficiency

and traceability of financial transactions, reducing the risk of fraud and enhancing the reliability of data used in predictive models (Nguyen & Han, 2019)[21]. In summary, as the field continues to evolve, ongoing research and innovation will be essential to fully realize the potential of these advanced predictive tools, ensuring their responsible and effective use in the challenges of the modern managerial landscape.

2. Relevance of an Ever-Evolving Predictive Landscape

The significance of these studies lies mainly in their ability to provide foresight and strategic insights that are crucial for decision-making, risk management, and long-term planning, but also for an ethical and sustainable business management (Sun et al., 2014). One of the primary benefits of predictive studies is their role in risk management for bigger firms and also for SMEs. In an example, manufacturing companies use predictive maintenance models to forecast equipment failures (lo Conte & Sancetta, 2022). By monitoring machine data and identifying patterns that precede breakdowns, companies can schedule maintenance before a failure occurs, thereby reducing downtime and maintenance costs. This not only enhances operational efficiency but also extends the lifespan of equipment. In the realm of strategic planning, predictive studies provide businesses with a competitive edge (Porter, 1985)[22]. Retail companies, for example, use predictive analytics to forecast consumer demand. By analyzing historical sales data, market trends, and seasonal patterns, retailers can in fact optimize their inventory levels, ensuring they have the right products in stock at the right time. This reduces excess inventory costs and increases customer satisfaction by meeting demand promptly (Grant, 1991).

Predictive analytics also support sustainability efforts by helping companies forecast and manage their environmental impact (Kim & Upneja, 2014), and, also in industries where regulatory compliance is critical, predictive studies play a crucial role. For example, energy companies use predictive analytics to ensure they meet environmental regulations. By forecasting emissions and energy consumption patterns, these companies can implement measures to reduce their environmental impact and avoid regulatory fines. However, one of the most critical applications of predictive studies is in forecasting financial distress and crises (Addo et al., 2018; Begley et al., 1996). Financial distress in a company can lead to severe consequences such as bankruptcy, loss of jobs, and negative impacts on stakeholders and predictive models help in identifying the early signs of financial trouble, allowing businesses to take proactive measures to avoid or mitigate the impacts of such crises (Agostini & Nosella, 2017). By identifying trends such as declining revenue, increasing debt, or shrinking profit margins, these models can forecast potential financial distress, and this allows management to implement corrective actions, such as restructuring debt, reducing costs, or seeking additional capital, before the situation becomes critical (Sciarelli, 1995; Altman, 1968). For these reasons, the field of

of data analysis, enabling better prediction and management of financial risks. This integration helps in identifying early warning signs of financial instability, optimizing resource allocation, and enhancing overall economic resilience.

[21] One of the key features of blockchain technology is its ability to create a transparent and tamper-proof record of transactions. Each transaction is linked to the previous one, forming a chain of blocks, hence the name "blockchain." This structure makes it extremely difficult for any single entity to alter the transaction history without being detected by other participants in the network. As a result, blockchain can significantly reduce the risk of fraud and errors in financial transactions, which is crucial for maintaining the accuracy and reliability of predictive models (Yermack, 2017).

[22] A "*competitive edge*" is a term used in the business context to describe a factor or combination of factors that allow a company to produce goods or services better or more cheaply than its competitors. This advantage enables the company to generate more sales or superior margins compared to its competition. For Porter (1985) and Barney (1991), It can stem from various sources, such as technological innovation, superior product quality, exceptional customer service, cost efficiency, or brand reputation.

predictive analytics is continuously evolving, driven by advancements in technology, changes in market conditions, and the increasing availability of data. Several factors contribute to the dynamic nature of this field:

- *Technological Advancements*: Innovations in AI, ML, and big data analytics have significantly enhanced the capabilities of predictive models[23];
- *Data Availability:* The proliferation of digital technologies has led to an explosion of data generated by businesses, consumers, and markets. This data, when harnessed effectively, provides valuable insights that can improve the accuracy of predictive models;
- *Changing Market Dynamics:* Globalization, technological disruptions, and economic fluctuations constantly reshape the business environment. Predictive models must adapt to these changes to remain relevant and effective. For instance, the covid-19 pandemic introduced unprecedented economic challenges, prompting the development of new models to forecast its impact on businesses and economies[24];
- *Regulatory Changes:* Regulatory environments also evolve, impacting how businesses operate and manage risks and strategy. Predictive models must then incorporate regulatory changes to ensure compliance and optimize decision-making processes.

Along with these factors, we can define also several benefits for businesses involved:

- *Early Detection of Crises:* By forecasting potential financial distress, businesses can identify problems early and take corrective actions. For example, a retail company might use predictive analytics to detect a trend of declining sales and adjust its marketing strategies or inventory management accordingly;
- *Informed Decision-Making:* Predictive models provide valuable insights that inform strategic decisions;
- *Resource Optimization:* By anticipating market trends and consumer behavior, businesses can allocate resources more efficiently. For example, an airline might use predictive analytics to forecast travel demand and adjust its flight schedules and pricing strategies to maximize revenue;
- *Improved Financial Performance:* Ultimately, predictive studies help businesses improve their financial performance by enabling them to adapt to changing conditions, optimize operations,

[23] Artificial Intelligence (AI), a term coined by John McCarthy in 1956, refers to the capability of machines to perform tasks that typically require human intelligence, such as learning, reasoning, and problem-solving (McCarthy, 2007). This conceptual leap marked the birth of a field dedicated to creating intelligent systems capable of mimicking human cognitive functions.

[24] The integration of predictive analytics with ML and AI can provide several benefits during a health crisis such as covid-19:

- Early Detection and Monitoring: ML algorithms can identify early signs of outbreaks and monitor the spread of the virus in real-time, enabling swift public health interventions;
- Resource Allocation: Predictive models help determine the need for medical resources, ensuring that hospitals and health facilities are adequately equipped to handle surges in covid-19 cases;
- Policy Formulation: By analyzing various "what if" scenarios, predictive models assist policymakers in evaluating the potential impact of different public health strategies, such as lockdowns and social distancing measures;
- Supply Chain Management: AI-driven models ensure the efficient distribution of essential supplies, such as PPE and ventilators, by forecasting demand and managing supply chains effectively.

and make data-driven decisions. This proactive approach enhances profitability and long-term sustainability.

The ability to predict financial distress and crises is then particularly valuable, as it allows businesses to take proactive measures to avoid severe consequences. As technological advancements continue to enhance the capabilities of predictive models, their importance is likely to grow, further solidifying their role in strategic decision-making and economic stability. Furthermore, the relevance of a proactive and future-forward vision in crisis and insolvency prediction and management is underscored by the numerous examples in scientific literature and practice (Bellovary et al., 2007). Thanks to a proactive vision, businesses that adopt this approach are better equipped to withstand economic shocks and uncertainties and a proactive crisis management and strategic forecasting lead to more efficient operations and optimized resource allocation (Altman et al., 2017; Altman, 1968). Demonstrating a proactive and future-forward approach also enhances stakeholder confidence: investors, customers, and partners are more likely to trust and engage with businesses that are seen as forward-thinking and resilient[25]. Finally, on a macroeconomic level, widespread adoption of these practices contributes to overall economic stability. By preventing widespread insolvencies and financial crises, in fact, businesses play a crucial role in maintaining the health and resilience of the entire economy (Aly et al., 2022)[26].

2.1. Future-Forward Approach from a Managerial Perspective

Managers play a crucial role in implementing proactive crisis management strategies. However, they often require tailored approaches and continuous support to effectively utilize predictive analytics[27]. One-size-fits-all strategies are inadequate in today's diverse business landscape (lo Conte et al., 2022). Instead, customized solutions that consider the unique financial structures, market conditions, and operational dynamics of each organization are necessary (White & Green, 2021). Moreover, managerial literacy in advanced analytics and strategic foresight is paramount[28]. In fact, managers need to be adept at interpreting predictive data and integrating it into their decision-making processes. This necessitates comprehensive training programs and ongoing education to keep pace with technological advancements. As stated by Pinski et al. (2024), despite recognizing the potential of AI through strategic AI orientation, many firms still struggle to realize its full value, and the critical role of top management team (TMT) AI literacy and orientation in successful AI adoption cannot be overstated. In fact, AI-

[25] A future-forward vision integrates holistic analysis and continuous adaptation into financial distress management. This entails considering a wide array of variables, including macroeconomic indicators, industry-specific factors, and even social media sentiment, to develop a nuanced understanding of financial health and stability (Miller & Davis, 2019).

[26] The concepts of health and resilience are increasingly recognized as vital to the stability and sustainability of economies. Health, both at an individual and societal level, directly influences economic productivity, workforce stability, and long-term growth. A healthy population is more capable of contributing to economic activities, fostering innovation, and sustaining consumption, all of which are critical drivers of economic resilience. Resilience, on the other hand, refers to the capacity of an economy to absorb shocks, adapt to changes, and recover swiftly from crises. In the context of economic systems, resilience is built through diversification, robust public health systems, and policies that promote equitable access to resources and opportunities (Aly et al., 2022).

[27] To maximize the benefits of predictive analytics, SMEs frequently need customized strategies and ongoing guidance, as their limited resources and specialized needs demand a more focused and sustained implementation approach.

[28] AI literacy among top management is crucial for effectively implementing AI strategies within organizations. According to Pinski et al. (2024), AI literacy encompasses the ability to critically use, evaluate, and communicate AI technologies. This proficiency enables executives to identify AI value potentials and realize these values through effective AI implementation. The study highlights that firms with AI-literate top management teams (TMTs) are better positioned to leverage AI for strategic advantage.

literate executives not only contribute to strategic insights but also effectively manage AI-driven projects, facilitating smoother implementation processes and mitigating risks. Furthermore, AI orientation plays a significant role in AI implementation (Pinski et al., 2024). Effective communication of AI strategies, including through official documents, ensures that AI implementation is not just a goal but a means to facilitate overall organizational effectiveness. For example, the partial mediation of AI orientation between TMT AI literacy and HR-related AI implementation ability underscores that TMT AI literacy directly and indirectly influences AI implementation through strategic alignment[29].

That said, and considering the points raised in the previous paragraphs, it is clear that new perspectives must be adopted by businesses and managers operating in various capacities within firms today. Although enterprises still rely heavily on historical operations, addressing the challenges and obstacles of the modern era requires a proactive and forward-looking approach to fully understand the opportunities and threats surrounding their activities[30]. As managerial literature on the subject teaches us, contemporary managers need well-informed structures to support every decision. More importantly, they require skills, resources, and competencies tailored to the characteristics of their specific market and business environment. Indeed, a future-forward approach would be ineffective without adequate technical expertise and deep literacy, not just superficial understanding[31]. Scientific and academic methodologies, traditionally used to measure economic and financial performance and to evaluate firms going-concern, have been applied in managerial practice over the years. However, the reality of rapid technological advancements and exogenous upheavals has necessitated significant changes in this area. Often, there has been a temporal mismatch between the skills possessed by management and those required by the context. Today, however, AI and ML are promising and practically demonstrating significant advancements in this regard. Systems and tools based on these innovations enable the reliable, swift, and resource-effective measurement, management, analysis, visualization, and understanding of data and results. Furthermore, limitations due to inadequate managerial literacy and limited financial resources are being, and will increasingly be, overcome by the efficient implementation of AI-based technologies (Pinski et al., 2024)[32].

2.1.1. New Managerial Literacy Approaches

As said before, the advent of advanced technologies such as AI and ML has profoundly reshaped the landscape of financial distress prediction, necessitating a paradigm shift also in managerial literacy. As AI and ML become integral, today, to predictive analytics, managers must cultivate a comprehensive understanding of these technologies to effectively leverage their capabilities in strategic decision-making. Traditional management skills, which focused primarily on financial metrics and historical data analysis, now require augmentation with expertise in AI-driven predictive models. This evolution in managerial literacy is critical for navigating the complexities of modern environments and

[29] Research indicates that the impact of TMT AI literacy on AI implementation varies by firm type, with startups benefiting more due to their agility and resource configuration (Davenport & Bean, 2018). Furthermore, the human resource aspect of AI implementation is pivotal. TMTs with a strong grasp of AI can drive HR policies that support AI adoption, from recruitment to training, thus building a robust foundation for AI-driven innovation (Rana & Sharma, 2019).

[30] Despite the continued dependence on historical operations, businesses must adopt a proactive, forward-looking strategy to navigate the complexities of the modern era and fully comprehend the emerging opportunities and risks inherent in their activities (Altman et al., 2017; Davenport & Harris, 2017).

[31] Without this level of expertise, organizations risk misinterpreting data and making poorly informed decisions that could undermine their competitive advantage (Davenport & Harris, 2017; Cao et al., 2015).

[32] As AI becomes more accessible and integrated into business operations, its role in overcoming these traditional barriers is expected to grow, allowing firms to enhance their competitive positioning and operational efficiency (Nguyen et al., 2019; Brynjolfsson & McAfee, 2017).

ensuring organizational resilience, and encompasses several dimensions, including technical proficiency, strategic insight, and ethical considerations. Managers, for example, must be adept at interpreting the outputs of ML algorithms, understanding the underlying mechanisms of predictive models, and integrating these insights into broader strategic frameworks[33]. This involves not only recognizing patterns and trends in financial data but also anticipating potential risks and opportunities. As stated by Pinski et al. (2024), these new paradigms also require managers to develop a strategic foresight that extends beyond traditional economic-financial analysis. This foresight is characterized by the ability to foresee the implications of emerging technologies, regulatory changes, and market dynamics on the firm's financial health. Managers equipped with this foresight can proactively adjust their strategies to mitigate risks and capitalize on new opportunities. For instance, predictive analytics can identify early warning signs of financial distress, enabling managers to implement corrective measures before the situation escalates[34]. This proactive approach contrasts sharply with the reactive strategies of the past, underscoring the transformative potential of AI-driven management[35].

Moreover, the successful implementation of AI technologies in financial distress prediction hinges on continuous education and training. Managers must engage in ongoing learning to stay abreast of technological advancements and evolving best practices. This includes participating in specialized training programs, attending industry conferences, and collaborating with data scientists and AI experts[36]. In addition to this, the critical role of top management teams (TMTs) in driving AI adoption and implementation cannot be overstated. TMTs with high levels of AI literacy can effectively champion AI initiatives, communicate their strategic vision, and lead their organizations through the complexities of AI integration. This leadership is essential for overcoming resistance to change, securing buy-in from stakeholders, and ensuring the successful deployment of AI technologies. As AI technologies become more pervasive, issues related to data privacy, algorithmic bias, and transparency also gain prominence, but AI-literate managers are better equipped to address these challenges, ensuring that their organizations adhere to ethical standards and regulatory requirements. This ethical stewardship is vital for maintaining stakeholder trust and safeguarding the firm's reputation. In conclusion, the shift towards AI-driven predictive analytics in financial distress prediction necessitates a significant enhancement in managerial literacy. Managers must develop a deep understanding of new managerial technologies, cultivate strategic foresight, and engage in continuous learning to effectively leverage the

[33] For example, using tools like Microsoft Azure's ML platform can help managers create and interpret predictive models, which then inform strategic decisions, such as resource allocation or market entry strategies.

[34] Predictive analytics tools, such as SAP Predictive Analytics or IBM Watson, utilize historical and real-time data to detect patterns that indicate potential financial distress. These early warning systems allow managers to take preemptive actions, such as adjusting cash flows or renegotiating terms with creditors, thereby mitigating risks before they develop into critical issues (Jiang et al., 2020).

[35] Unlike traditional reactive strategies, which often involve responding to crises after they have fully materialized, AI-driven management leverages predictive models to foresee potential downturns. This shift towards a proactive management style represents a significant advancement, allowing firms to not only prevent crises but also to capitalize on emerging opportunities before competitors (Barboza et al., 2017).

[36] This ongoing education can take several forms:

- *Specialized Training Programs:* Managers can enroll in courses focused on AI, Machine Learning (ML), and data analytics. For example, programs like MIT's "Artificial Intelligence: Implications for Business Strategy" offer in-depth insights into how AI can be applied strategically within organizations (Agrawal, Gans, & Goldfarb, 2018);
- *Industry Conferences:* Attending conferences such as the AI Summit or Strata Data Conference allows managers to network with industry leaders, learn about cutting-edge technologies, and explore how other companies are integrating AI into their strategies (Brynjolfsson & McAfee, 2017);
- *Collaboration with Data Scientists:* By working closely with data scientists and AI experts, managers can gain a practical understanding of how predictive models function. This collaboration is crucial for integrating AI-driven insights into broader strategic frameworks, ensuring that the outputs of predictive analytics are effectively utilized in decision-making (Davenport & Harris, 2017).

capabilities of AI and ML. This evolution is essential for ensuring organizational resilience, optimizing decision-making, and maintaining a competitive edge.

2.1.2. Development of a Data-centric Culture

Expanding on the practical implications for executive roles and management approaches, it is essential for firms, today, to strategically prioritize the development of a data-centric culture (Chatterjee et al., 2021)[37]. A data-centric culture emphasizes the importance of data in decision-making processes, enabling firms to harness the full potential of innovative technologies. This shift is vital for ensuring that data-driven insights are integrated seamlessly into business strategies, thereby enhancing predictive analytics and other critical business functions[38]. To cultivate a data-centric culture, organizations must start by ensuring that all members possess adequate literacy. This comprehensive approach avoids the pitfalls of relegating knowledge to a single role. In the case of AI and new predictive technologies, shareholders and top executives should mandate AI literacy across all executive roles, tailoring this literacy to specific areas of responsibility. This requirement can be incorporated into hiring practices and role definitions, ensuring that every executive can contribute to and advocate for AI initiatives. Furthermore, the organizational resource configuration plays a significant role in the success of AI adoption[39]. For incumbent firms with less adaptable structures, TMTs must proactively address and eliminate obstacles that could impede AI implementation. This may involve developing strategies such as investing in data infrastructure, promoting data literacy at all organizational levels, and encouraging a mindset that values data-driven decision-making. Toutaoui et al. (2022), for example, highlight the importance of creating an environment where data is readily accessible, accurately interpreted, and effectively used to inform strategic decisions[40]. In addition to internal strategies, understanding firm type advantages can enhance the efficacy of AI adoption. Different firm types, such as startups and incumbents, possess unique strengths that can be leveraged through strategic partnerships and business networks[41]. For instance, startups often exhibit greater agility and innovation capabilities, while incumbents may have more extensive resources and established market presence. Executives can form alliances that combine these complementary strengths, creating synergistic opportunities for AI-driven growth and innovation (Steiber & Alänge, 2020).

Assessing competitors, suppliers, and acquisition targets through the lens of firm type advantages can also provide valuable insights. By understanding the AI adoption potential of these firms, executives can make informed judgments about their strategic fit and the benefits of potential partnerships. This knowledge can guide decisions on mergers and acquisitions, collaborative projects, and supply chain optimizations, ultimately enhancing the firm's competitive position. Furthermore, data-driven culture is considered to bring business-oriented cultural transformation to a firm. In fact, it

[37] For an in-depth understanding, see also Davenport & Harris (2017), which explores how organizations can systematically implement data-centric strategies to optimize decision-making processes. The text outlines specific methodologies for integrating data analytics into business operations and addresses the technical challenges that firms may encounter during this transformation.

[38] For further details on data-driven strategies and their implementation, see the following references: Chatterjee et al. (2024), Davenport & Harris (2017), Chen et al. (2012), Cao et al. (2015), and Gunasekaran et al. (2017).

[39] The configuration of an organization's resources, including technological infrastructure, human capital, and data management capabilities, is critical to the effective adoption of AI.

[40] Toutaoui et al. (2022) emphasize that for organizations to fully leverage the power of data, it is essential to establish a robust data infrastructure where data is not only easily accessible across departments but also processed and analyzed with precision.

[41] Leveraging the complementary capabilities of startups' agile innovation and incumbents' scalable resources through strategic alliances can optimize market positioning and accelerate technological adoption. For further insights, see the following works: Chatterjee et al. (2024) and Toutaoui et al. (2022).

is considered to provide substantial dividends to the firms' product and process innovations[42]. Recently, several firms have been using different advanced technology-embedded business analytics (BA) tools to improve their business performance[43]. Again, advancement of information and communication technology has helped firms to explore the option to use BA tools with AI[44]. This has brought radical change in the business-oriented cultural landscape of the firms to arrive at accurate decision-making to improve their innovation and performance. In conclusion, strategically prioritizing the development of a data-centric culture and leveraging firm type advantages are crucial steps for maximizing the benefits of AI and ML technologies, particularly in a predictive landscape. These strategies, in fact, not only improve decision-making but also position organizations to capitalize on emerging opportunities.

2.2. Ethical and Social Consequences

Nevertheless, ethical and social consequences and issues arising from these innovations must also be considered. In fact, as organizations increasingly integrate AI and ML into their financial distress management practices, it is imperative to address the potential negative impacts on various stakeholders and society at large. One of the primary ethical concerns with AI-driven predictive analytics is data privacy. The extensive use of personal and sensitive data in AI models necessitates robust measures to protect individuals' privacy. Ensuring that data is collected, stored, and utilized in a manner that respects user consent and confidentiality is then paramount. This requires strict adherence to data protection regulations, such as the General Data Protection Regulation (GDPR) in the European Union, which mandates transparency and accountability in data handling practices (Shin, 2021). Another critical ethical issue is algorithmic bias, where AI systems may inadvertently perpetuate or exacerbate existing biases present in the data they are trained on. This can lead to unfair outcomes, particularly in sensitive applications. To mitigate this risk, it is essential to implement rigorous auditing and validation processes to identify and rectify biases in AI algorithms. This includes diversifying training datasets and developing fairness-aware ML techniques that promote equitable outcomes (Binns, 2018). Furthermore, the social consequences of AI and predictive analytics extend beyond individual biases to broader societal impacts. The possible automation of decision-making processes, in fact, can lead to significant shifts in the labor market, potentially displacing workers and exacerbating economic inequality. As AI systems increasingly take on roles traditionally performed by humans, there is a pressing need to address the resulting employment challenges. This includes investing in education and retraining programs to equip the workforce with the skills needed to thrive in an AI-driven economy (Brynjolfsson & McAfee, 2014). Moreover, the deployment of AI in predictive analytics must consider the implications for social trust and public perception. Transparency in AI decision-making processes is crucial for maintaining trust among stakeholders and it involves clearly communicating

[42] This is particularly evident in studies that highlight the correlation between data-driven strategies and innovation outcomes, demonstrating how firms can leverage predictive insights to foster continuous improvement and maintain a competitive edge. See Chatterjee et al. (2024) for a detailed exploration of these dynamics.

[43] Examples of advanced technology-embedded business analytics (BA) tools that firms are increasingly utilizing include platforms like:

- *Tableau for data visualization;*
- *SAP BusinessObjects for comprehensive reporting and analytics;*
- *IBM Watson Analytics for AI-driven data exploration;*
- *Google Analytics 360 for in-depth web analytics.*

[44] Business Analytics (BA) refers to the systematic computational analysis of data and statistics within organizations, utilizing advanced algorithms and data processing techniques to support data-driven decision-making. BA tools, when integrated with AI, leverage machine learning and predictive modeling to enhance insights, optimize operations, and enable strategic foresight

how AI models work, the data they use, and the rationale behind their predictions. By determining an environment of openness and accountability, organizations can build public confidence in AI technologies and mitigate fears of opaque and untrustworthy systems (Floridi et al., 2018).

In addition to transparency, the ethical use of AI also requires a commitment to social responsibility. Organizations must ensure that their AI initiatives align with broader societal values and contribute positively to the community. By aligning AI deployment with social good, organizations can enhance their ethical standing and demonstrate their commitment to corporate social responsibility (CSR) (Vinuesa et al., 2020)[45]. Finally, the ethical and social consequences of AI-driven predictive analytics underscore the importance of interdisciplinary collaboration. Addressing these complex issues requires in fact input from diverse stakeholders, including legal experts and technologists. With a multidisciplinary dialogue, organizations can develop comprehensive strategies that balance technological innovation with ethical considerations and social impact (Mittelstadt et al., 2016).

2.2.1. Privacy Concerns and Data Security

Today's technological progression brings with it substantial concerns regarding privacy and data security. As organizations increasingly rely on vast amounts of data to train their predictive models, the need to address privacy issues and ensure robust data security measures becomes fundamental. Firstly, the data utilized in AI-driven predictive models often includes sensitive financial information about companies and individuals that can range from financial statements and credit histories to transactional records and market behaviors. The collection, storage, and processing of such data expose organizations to the risk of data breaches and unauthorized access. According to Kshetri (2022), the financial sector is particularly vulnerable to cyberattacks due to the high value of financial data and this vulnerability underscores the critical importance of implementing stringent data security protocols to protect against potential breaches. Moreover, privacy concerns are heightened by the inherent nature of AI and ML algorithms, which require extensive datasets to learn and make accurate predictions. These datasets often include personally identifiable information (PII), which, if mishandled, can lead to severe privacy violations. The General Data Protection Regulation (GDPR) and other regulatory frameworks emphasize the need for organizations to handle PII with the utmost care, ensuring compliance with legal standards and protecting individuals' privacy rights[46].

In particular, in the context of financial distress prediction, the use of AI can inadvertently lead to the profiling of companies and individuals based on their financial behaviors. This profiling can result in biased outcomes, where certain entities might be unfairly targeted or disadvantaged due to the patterns identified by the algorithms. As highlighted by Crawford and Calo (2016), the ethical implications of AI necessitate a careful examination of how data is used and the potential biases that may arise. To mitigate these risks, it is essential for organizations to implement transparency measures, ensuring that the data used in predictive models is representative and unbiased. Furthermore, the

[45] Embedding AI initiatives with a focus on societal values can be exemplified by organizations adopting AI systems that prioritize equitable outcomes and community benefits. For instance, by deploying AI to optimize resource allocation in public services, firms contribute to social equity and welfare. This approach strengthens Corporate Social Responsibility (CSR) as it demonstrates a commitment to ethical AI practices, addresses algorithmic fairness, and aligns business goals with broader societal interests, fostering public trust (Vinuesa et al., 2020).

[46] The General Data Protection Regulation (GDPR) mandates stringent protocols for handling Personally Identifiable Information (PII), requiring data controllers and processors to implement measures such as data minimization, pseudonymization, and encryption. Compliance involves maintaining detailed data processing records, ensuring lawful basis for data collection, and facilitating individuals' rights to access, rectify, and erase their data.

dynamic nature of AI and ML technologies means that models are continuously evolving based on new data inputs: this constant evolution necessitates regular updates to data security measures to address emerging threats. Following this, there is the need of a proactive approach to cybersecurity, incorporating advanced encryption techniques, secure data storage solutions, and regular security audits to safeguard their data assets. In this regard, Kshetri (2022) emphasizes that a robust cybersecurity strategy is not just a technical requirement but a critical component of organizational resilience in the digital age. Another critical aspect of privacy and data security in AI-driven predictive analytics is the principle of data minimization[47]. This principle advocates for the collection of only the necessary data required for a specific purpose, thereby reducing the risk of excessive data accumulation and potential misuse. In practice, this means that organizations should carefully evaluate the data they collect, ensuring that it is directly relevant to their predictive models and discarding any extraneous information. This approach not only enhances privacy but also aligns with regulatory requirements such as GDPR.

However, the implementation of data minimization can lead to inefficiencies and potential challenges. One significant issue is the possible loss of valuable information that could enhance the predictive power of the models. By discarding extraneous data, there is a risk of eliminating potentially useful variables that, while not immediately relevant, could provide critical insights when combined with other data points. This reductionist approach can hinder the model's ability to fully capture the complexity of financial distress scenarios, potentially leading to less accurate predictions. Additionally, the process of evaluating and filtering data to ensure compliance with privacy regulations can be resource-intensive. Organizations must invest in sophisticated data governance frameworks and employ skilled data scientists to continuously monitor and manage the data lifecycle. This necessity can strain organizational resources, particularly for smaller firms with limited budgets. Moreover, the dynamic nature of AI models requires constant updates and real-time data integration, which can complicate the adherence to data minimization principles. Another layer of complexity is introduced by the knowledge domain required to effectively manage AI and ML systems. The implementation of these technologies necessitates a deep understanding of both the technical aspects and the business context. Ensuring that all stakeholders, from data scientists to top management, are adequately trained and knowledgeable in AI applications is then crucial. However, achieving this level of literacy across the organization can be challenging. There are often gaps in understanding between technical experts who develop the models and business leaders who implement the insights derived from these models.

Furthermore, the interplay between data privacy and security and the effectiveness of predictive models highlights the importance of interdisciplinary collaboration. In addition to technical measures, a culture of data security within the organization is essential. This involves training employees on best practices for data handling, raising awareness about the importance of data privacy, and establishing clear protocols for responding to data breaches. As suggested by Westerman et al. (2019), a holistic approach to data security that includes both technological and human elements is crucial for effectively managing the risks associated with AI and predictive analytics.

[47] In AI-driven predictive analytics, the principle of data minimization plays a pivotal role; by limiting data collection to only what is strictly necessary, organizations can reduce vulnerability to data breaches, mitigate potential privacy violations, and align with regulatory mandates such as GDPR, fostering both operational security and ethical data stewardship.

2.2.2. Transparency and Accountability in AI Decision-Making

The rise of AI in predictive analytics also has undeniably transformed decision-making processes across various sectors, including finance, healthcare, and manufacturing. However, this transformation brings forth critical issues of transparency and accountability that need rigorous academic exploration and practical implementation. One of the fundamental concern is that AI systems often operate as "*black boxes*," where the decision-making process is not easily understood or interpretable by human users. This opacity can lead to distrust among stakeholders, including employees, customers, and regulatory bodies, potentially undermining the efficacy and acceptance of AI-driven decisions. Transparency in AI decision-making refers to the clarity and openness with which AI systems operate, including the rationale behind specific decisions and, according to recent studies, transparency is a key factor in fostering trust in AI systems. When stakeholders understand how and why AI systems arrive at particular decisions, they are more likely to trust and accept those decisions (Kshetri, 2022). This is particularly crucial in high-stakes environments like financial distress prediction, where the consequences of decisions can significantly impact organizational survival and stakeholder interests.

One approach to enhancing transparency is through Explainable AI (XAI), which aims to make the decision-making processes of AI systems more understandable to humans[48]. XAI techniques, in fact, provide insights into the internal workings of AI models, revealing how inputs are transformed into outputs. This transparency can help stakeholders to identify potential biases, understand the limitations of the models, and make more informed decisions based on AI predictions. Moreover, Explainable AI can facilitate compliance with regulatory requirements that mandate transparency in automated decision-making processes, such as the GDPR in Europe. Accountability in AI decision-making involves holding AI systems and their creators responsible for the outcomes of their decisions. This encompasses ensuring that AI systems are used ethically and that their decisions can be audited and traced back to specific algorithms and data inputs. Accountability mechanisms are essential for addressing potential misuse or errors in AI systems, which can lead to significant ethical and legal implications. For instance, if an AI system used in financial distress prediction makes an erroneous decision that results in unwarranted financial penalties for a company, there must be a clear framework for investigating and rectifying such errors.

The research by Qiu et al. (2022) underscores the importance of transparency and accountability in building trust in AI systems among employees. Their study highlights, in fact, that when AI systems are perceived as effective and their decision-making processes are transparent, employees are more likely to trust and feel comfortable using these systems. This trust is crucial for the successful implementation and adoption of AI technologies and, without it, even the most advanced AI systems may face resistance and skepticism, limiting their potential benefits. Furthermore, continuous monitoring and auditing of AI systems are necessary to maintain transparency and accountability. This involves regularly reviewing AI models to ensure they operate as intended and do not perpetuate biases or errors. It also includes providing channels for stakeholders to report concerns or issues related to AI decisions. This, in turn, can lead to more effective and widely accepted AI-driven decisions, ultimately contributing to the success and sustainability of firms.

[48] By decomposing complex algorithms, such as deep learning or ensemble models, XAI techniques (e.g., SHAP values, LIME) provide insights into feature importance and decision pathways, enhancing accountability and trust. XAI is critical in sectors with regulatory scrutiny, as it enables compliance with transparency requirements and supports ethical AI deployment by reducing '*black box*' effects.

2.2.3. Socioeconomic Implications of AI-Driven Predictions

As AI-driven predictions become more pervasive, they have also the potential to reshape economic landscapes, influence social dynamics, and alter firms behaviour. One of the most significant socioeconomic implications of AI-driven predictions is the potential for economic disruption. AI technologies, through their advanced data processing and predictive capabilities, can significantly enhance the accuracy of financial distress predictions: this, in turn, can lead to more timely and effective interventions by financial institutions, investors, and regulatory bodies (Manogaran et al., 2019). However, the same technologies can also exacerbate economic inequalities if their benefits are not equitably distributed. For instance, large corporations and financial institutions with substantial resources to invest in AI technologies may gain a competitive edge, leaving smaller firms and less technologically adept organizations at a disadvantage. This disparity could widen the gap between different economic actors, leading to increased market concentration and reduced competition (Schwab, 2017). Moreover, the predictive power of AI can influence investment decisions and capital allocation, potentially leading to more efficient markets. Investors, armed with accurate predictions of financial distress, can make better-informed decisions, thereby optimizing their portfolios and reducing risks. This efficiency, while beneficial for individual investors and the broader financial market, could also result in market volatility if predictive models are widely adopted and lead to homogeneous investment strategies. The herding behavior induced by AI-driven predictions could also amplify market fluctuations, posing systemic risks (Schwab, 2017). From a labor market perspective, the widespread adoption of AI-driven predictions could lead to significant shifts in employment patterns. The automation of predictive analytics and decision-making processes may reduce the demand for certain job roles, particularly those involving routine data analysis and financial assessment. Conversely, there will be a growing demand for roles that involve the development, implementation, and oversight of AI technologies. This shift necessitates a workforce equipped with advanced skills in AI and data science, highlighting the importance of continuous education and reskilling programs (Acemoglu & Restrepo, 2018).

However, the social implications of AI-driven predictions extend beyond the labor market to broader societal impacts. One concern is the potential for bias and discrimination embedded within AI models. If predictive models are trained on biased historical data, they may perpetuate existing inequalities, leading to discriminatory outcomes in areas such as credit scoring, hiring, and law enforcement (O'Neil, 2016): ensuring the fairness and ethical use of AI requires rigorous scrutiny of training data, transparency in AI decision-making processes, and the implementation of robust governance frameworks to mitigate biases (Chatterjee et al., 2021). On the other hand, AI-driven predictions also hold the promise of positive social impacts, particularly in enhancing public services and welfare programs. For example, predictive analytics can be used to identify individuals at risk of financial distress, enabling proactive interventions by social services and policymakers to provide support and resources. This can lead to improved social outcomes, such as reduced poverty rates and enhanced financial stability for vulnerable populations (Smith & Anderson, 2018). As stated before, because of these consequences, the deployment of AI-based technologies must be guided by principles of transparency, accountability, and inclusivity to ensure that their benefits are broadly shared and do not exacerbate social divides, and, in this situation, policymakers and regulators play a crucial role in establishing frameworks that promote ethical AI use and protect the rights of individuals affected by AI decisions (Floridi et al., 2018). In conclusion, the socioeconomic implications of AI-driven predictions are vast and complex, encompassing economic, social, and ethical dimensions; while these technologies offer significant potential for improving financial distress prediction and enhancing market efficiency,

they also pose challenges that must be carefully managed. However, addressing these challenges requires a collaborative effort from stakeholders across sectors, including businesses, governments, and civil society.

2.2.4. Transforming Financial Resilience and Socioeconomics Dynamics

Advanced predictive technologies not only transform the way organizations and firms anticipate and manage financial and economic instability, but also influence broader economic and financial structures. In fact, AI-driven predictions can significantly enhance the resilience of financial systems by providing early warnings of potential crises and this proactive approach enables firms to implement preemptive measures, thereby mitigating the adverse effects of financial distress on businesses and their stakeholders. For instance, timely predictions can help companies avoid bankruptcy by facilitating strategic adjustments such as restructuring debts, optimizing cash flow, and improving operational efficiencies. Consequently, these actions can preserve jobs, maintain supplier relationships, and sustain consumer confidence, which are critical for economic stability (Altman et al., 2017; Barboza et al., 2017). From a merely socioeconomic perspective, the ability to predict financial distress using AI and ML can also reduce the systemic risk within the financial sector supporting more informed decision-making by financial institutions, regulators, and policymakers. This can lead to a more stable financial environment, which is essential for economic growth and development (Floridi et al., 2018). Moreover, new predictive models can democratize access to sophisticated economic and financial analysis, enabling smaller businesses and startups to benefit from advanced predictive tools that were previously accessible only to larger corporations. This democratization can drive innovation and competition, leading to more dynamic and inclusive economic landscape[49]. Smaller firms, in particular, often constrained by limited resources, can leverage these technologies to identify potential financial risks and opportunities, thereby enhancing their competitiveness and sustainability (Oehmichen et al., 2023).

However, the socioeconomic implications of AI-driven financial predictions are not solely positive. Even in this case, there are significant ethical and social considerations that must be addressed to ensure equitable and just outcomes (Steiber & Alänge, 2020). In fact, the reliance on AI-driven predictions can lead to a reduction in human oversight and accountability in financial decision-making. This shift can create challenges in ensuring transparency and fairness, as the complex algorithms underlying AI models can be difficult to interpret and scrutinize. Another critical aspect of AI-driven predictions is their potential to improve disaster resilience. As noted by Sapountzaki et al. (2018), predictive models leveraging AI can anticipate the impacts of crises akin to natural disasters, thereby enhancing the preparedness and response strategies of both governments and private entities. This capability is crucial for mitigating the economic fallout from unforeseen financial shocks, ensuring that economies can recover more swiftly and sustainably. In addition, AI and ML have the potential to improve resource allocation, also during crises and bankruptcy. This aligns with the findings of Alfaro-Cortés et al. (2020), who highlight the role of advanced analytics in optimizing resource management in crisis situations.

[49] Through scalable cloud-based platforms and user-friendly interfaces, small businesses and startups can now leverage high-dimensional data processing, machine learning algorithms, and predictive analytics, traditionally reserved for large enterprises with extensive resources.

2.3. Predictive Analytics in Corporate Strategy

The integration of predictive analytics into corporate strategy represents a significant evolution in how businesses approach decision-making and overall strategic planning. As companies increasingly operate in complex and rapidly changing environments, the ability to forecast potential outcomes and trends becomes, as stated in previous paragraphs, not just a competitive advantage but a necessity for survival and growth. This section describes briefly the critical aspects of embedding predictive analytics into firms' strategy, exploring how aligning predictive insights with business objectives, enhancing stakeholder engagement, and learning from successful case studies can drive strategic success. Special attention is given to the unique challenges and opportunities faced by SMEs and the critical role of sustainability alignment in modern corporate strategy.

2.3.1. Aligning Predictive Insights with Business Objectives

Aligning predictive insights with business objectives, especially in the context of financial distress, crisis management, and insolvency prevention, goes beyond just technical precision; in fact, it involves embedding these insights into the very fabric of an organization's strategy and culture. This alignment is crucial to ensure that predictive analytics not only supports but drives the achievement of strategic goals during periods of financial uncertainty. As discussed in the literature, predictive models must be strategically relevant, translating critical business objectives—such as maintaining cash flow, managing debt levels, or avoiding insolvency—into specific, measurable key performance indicators (KPIs) that can be effectively monitored and forecasted (Davenport & Harris, 2017). To achieve this alignment, a data-centric culture within the organization is paramount. A data-centric culture, in this situation, prioritizes data as a fundamental asset in decision-making processes, ensuring that predictive analytics is not just a tool used by a few specialists but an integral part of the entire organizational workflow. This cultural shift requires, for sure, a concerted effort to enhance data literacy across all levels of management, enabling leaders to understand and leverage predictive insights effectively. Managerial literacy in data and analytics is then critical; without it, even the most sophisticated predictive models can fail to influence strategic decisions. Cao et al. (2015) emphasize that resistance to data-driven decision-making or a lack of understanding among top management can significantly hinder the effective use of predictive analytics, especially in crisis situations where rapid, informed decisions are crucial.

Moreover, a data-centric culture involves establishing continuous feedback loops where predictive outcomes are consistently evaluated against actual performance. This process allows for the recalibration of models and adjustments to strategies, ensuring that the organization remains agile and responsive to emerging risks. In this regard, Chen et al. (2012) highlight the importance of adaptive business processes that can swiftly integrate new insights, particularly during times of financial stress when the stakes are high, and the margin for error is low. Building managerial literacy in predictive analytics is then essential for overcoming the challenges associated with financial distress and insolvency situations. Leaders equipped with a deep understanding of data and predictive models are better positioned to understand the complexities of crisis management as they can anticipate potential pitfalls, make proactive adjustments, and guide their organizations through turbulent times with greater confidence. This literacy extends beyond basic familiarity with data as it involves a comprehensive understanding of how predictive models function, their limitations, and how to interpret their outputs in the context of the organization's strategic objectives. In the context of the

already mentioned data-centric culture, managerial literacy ensures that predictive insights are not only generated but are also effectively communicated and acted upon. This alignment of predictive analytics with business objectives determines a proactive approach to crisis management, enabling organizations to anticipate and mitigate risks before they escalate into full-blown crises. At last, as firms increasingly rely on data-driven strategies, the role of predictive analytics in shaping corporate strategy becomes even more critical.

2.3.2. Sustainability and CSR Alignment

In today's business environment, Corporate Social Responsibility (CSR) has become a crucial aspect of organizational strategy, reflecting a company's commitment to ethical practices and sustainable development. As stakeholders increasingly demand transparency and accountability, integrating predictive analytics into CSR initiatives offers a powerful tool for companies to align their operations with sustainability goals. This approach allows businesses to anticipate and mitigate environmental and social risks, making CSR not only a moral obligation but also a strategic advantage in maintaining competitiveness and building stakeholder trust. For example, by analyzing data related to resource consumption, waste production, and emissions, companies can forecast their carbon footprint and identify opportunities to mitigate negative environmental impacts (Gunasekaran et al., 2017). This forward-looking approach allows organizations to not only comply with regulatory requirements but also to position themselves as leaders in sustainability, which is increasingly valued by consumers, investors, and other stakeholders. Furthermore, by integrating predictive analytics into their sustainability strategies, companies can move beyond reactive compliance and adopt a proactive stance that anticipates and addresses potential challenges before they escalate into significant risks (Gunasekaran et al., 2017; Vermeulen & Witjes, 2016).

On the other side, in the realm of supply chain management, predictive analytics plays a pivotal role in ensuring that suppliers adhere to environmental and social standards. For instance, predictive models can be used to monitor supplier behavior and predict potential violations of sustainability standards, enabling companies to take corrective actions before these issues damage their reputation or disrupt their supply chains (Kiron & Unruh, 2018)[50]. This use of predictive analytics enhances transparency and accountability, both of which are critical in managing complex global supply chains. Moreover, predictive analytics facilitates the measurement and reporting of CSR initiatives, providing companies with the tools to quantify their efforts and communicate their impact effectively. This capability is particularly important in today's era. The alignment of predictive analytics with CSR also contributes to innovation within the organization. As companies use predictive tools to identify areas where they can improve their sustainability performance, they often uncover new opportunities for product and process innovation (Gunasekaran et al., 2017). For instance, predictive analytics can reveal inefficiencies in energy use or waste management that, once addressed, not only reduce environmental impact but also lower operational costs. This dual benefit of enhancing sustainability while driving efficiency is a powerful motivator for companies to invest in predictive analytics as part of their CSR strategy.

In particular, small and medium-sized enterprises face unique opportunities when it comes to integrating predictive analytics with CSR. Due to their size, they may find it more difficult to implement

[50] For example, firms like Unilever and Nestlé deploy predictive tools to assess suppliers' carbon emissions and labor practices, identifying risks of non-compliance early. Such proactive monitoring not only mitigates reputational damage but also minimizes supply chain disruptions, allowing for timely intervention and ensuring alignment with corporate sustainability commitments

sophisticated predictive models. However, the flexibility and agility that often characterize SMEs can be advantageous in quickly adopting and adapting to new technologies. Moreover, as larger companies increasingly demand that their suppliers adhere to stringent CSR standards, SMEs that can demonstrate their commitment to sustainability through predictive analytics will be better positioned to compete in the global market (Gunasekaran et al., 2017). However, the integration of predictive analytics into CSR also aligns with broader trends in corporate governance and sustainability reporting. As global initiatives such as the United Nations Sustainable Development Goals (SDGs) and the Corporate Sustainability Reporting Directive (CSRD) gain traction, companies are under increasing pressure to not only set ambitious sustainability targets but also to provide credible evidence of their progress (Vermeulen & Witjes, 2016): predictive analytics can support this by offering reliable, data-driven insights that help companies track their progress towards these goals, identify areas for improvement, and adjust their strategies accordingly.

2.3.2.1. In crisis and insolvency situation

Companies that proactively manage their CSR efforts through advanced analytics can better manage the financial and reputational risks associated with crises and potential insolvency. Predictive analytics, in fact, enables firms to assess and mitigate risks by providing insights into the long-term sustainability of their business practices, which is critical for avoiding situations that could lead to financial distress. Furthermore, integrating CSR with predictive analytics reinforces the company's ability to meet the expectations of increasingly conscientious consumers and investors who prioritize sustainability. This alignment not only safeguards the company's market position but also enhances its resilience in the face of economic uncertainties. The implications of this integration extend beyond immediate crisis management and future research and practice in this area will likely explore the development of more comprehensive models that incorporate a wider range of CSR metrics, including those related to social equity, community engagement, and ethical labor practices.

2.4. Special Considerations for SMEs: Strategic Integration and Scalability

Expanding on the strategic integration of predictive analytics for SMEs, it is crucial to recognize that while these enterprises often face constraints in resources, the judicious adoption of predictive analytics can significantly enhance their competitiveness and operational efficiency, primarly in crisis situations. The key lies in a phased and targeted approach that aligns with the specific needs and capabilities of SMEs themselves. This process starts with identifying high-impact areas where predictive analytics can offer the most value, ensuring that even limited investments yield substantial returns. Mainly, cash flow management, customer retention, and inventory optimization are the areas where predictive insights can directly influence financial stability and operational efficiency, which are vital for the survival and growth of smaller enterprises. For example, by leveraging predictive analytics for cash flow management, SMEs can forecast shortfalls and surpluses more accurately, allowing for better financial planning and more effective use of working capital. This proactive management can prevent liquidity crises, a common challenge for SMEs, particularly in volatile markets (Gartner, 2020). Customer retention is another domain where predictive analytics can have a transformative impact: by

analyzing customer behavior data, SMEs can identify patterns that indicate potential churn and intervene with targeted marketing strategies to retain valuable customers. This not only boosts revenue but also builds long-term customer loyalty, which is crucial for these entities competing against larger firms with more extensive resources.

Furthermore, inventory optimization, supported by predictive analytics, allows SMEs to maintain the right balance between supply and demand, reducing holding costs and minimizing stockouts. This is particularly important for SMEs that often operate with thinner margins and cannot afford the inefficiencies associated with poor inventory management. Predictive models, in this case, can forecast demand trends based on historical sales data, seasonal variations, and market conditions, enabling firms to align their inventory levels with expected demand more precisely (O'Leary, 2019). Moreover, the advent of cloud-based predictive analytics platforms has been a game-changer for SMEs. These solutions provide access to sophisticated analytics capabilities without the need for significant upfront capital investment in hardware and software: cloud-based platforms are scalable, allowing SMEs to start small and expand their analytics capabilities as their needs and resources grow. This scalability is critical, as it aligns with the incremental growth strategies that are often necessary for these realities. Furthermore, the pay-as-you-go pricing models typical of cloud services mean that SMEs can manage costs more effectively, paying only for the capacity and features they need at any given time (O'Leary, 2019). Cloud platforms also offer the advantage of continuous updates and improvements in analytics tools, ensuring that SMEs always have access to the latest technologies without the burden of managing and maintaining the infrastructure themselves. This allows SMEs to focus on leveraging the insights generated by predictive analytics rather than on the technical complexities of managing the analytics environment.

Finally, given the limitations in in-house expertise, SMEs can significantly benefit from strategic partnerships with technology providers, consulting firms, or academic institutions. These partnerships can in fact provide access to advanced predictive analytics tools and expertise that would otherwise be out of reach for smaller firms. For instance, collaborations with universities can offer SMEs access to cutting-edge research and talented graduates who can bring fresh perspectives and advanced analytical skills to the company. Similarly, partnering with technology vendors can help SMEs tailor predictive analytics solutions to their specific needs, ensuring that the tools they adopt are both effective and aligned with their strategic goals (Gartner, 2020). These partnerships can also play a critical role in building the managerial literacy needed to effectively implement and use predictive analytics. Training programs, workshops, and ongoing support can equip SME managers with the knowledge and skills necessary to interpret and act on predictive insights, thereby maximizing the value of their analytics investments.

3. Research Objectives

As already stated in previous paragraphs, this PhD Thesis seeks to delineate a comprehensive framework that integrates advanced predictive analytics with strategic management practices. The objectives of this study were then meticulously crafted to address the multifaceted challenges that Italian businesses face in predicting and managing these critical situations and can be considered not merely academic exercises but are grounded in the practical realities of today's business environment. They seek to bridge the gap between theory and practice, also providing innovative and *ready-to-use* insights that can be implemented by managers and studied, thoroughly, by academia.

In exploring these objectives, this work also raises critical Research Questions (RQ); these are designed to probe the accuracy of current predictive algorithms, the role of managerial literacy in adopting new technologies, the impact of AI in real-life decision making and the alignment of predictive analytics with broader corporate goals, keeping in mind that explainability and simplicity are the main features that a managerial model should have. Through these inquiries, the Thesis also endeavors to uncover the underlying mechanisms that drive successful crisis management and to propose innovative strategies that leverage the power of AI and ML for enhanced predictive accuracy and strategic foresight. The RQ, therefore, serve as the guiding framework for this investigation, directing the inquiry towards areas that are not only theoretically significant but also of paramount importance to practitioners. The ultimate aim is to develop a robust, data-driven and AI-based approach to crisis and insolvency prediction that is adaptable, scalable, and aligned with the ethical and strategic imperatives of modern businesses. In summary, this section of the Thesis outlines the core objectives and RQ that will shape the study, emphasizing the need for a forward-thinking approach that integrates cutting-edge technologies with a strategic, managerial, perspective.

3.1. Main Purposes

This research is structured to operate on multiple levels, each layer contributing to a comprehensive and deep understanding of corporate crisis and insolvency prediction. At its core, this research unfolds as a narrative that intertwines theoretical exploration, empirical investigation, and practical application. In fact, it begins by tracing the historical evolution of predictive methodologies, establishing a solid foundation upon which contemporary innovations in AI and ML are critically examined. This narrative arc then transitions into a detailed analysis of how these advanced technologies can be harnessed to address the limitations of traditional models, ultimately leading to the development of more effective and resilient predictive tools. This multi-layered approach reflects a deep commitment to advancing the field of predictive analytics, with each Chapter and section adding a new dimension to the overall understanding of how organizations can better anticipate and manage financial distress. So, the story of this Thesis is one of continuous inquiry, where each level of analysis supports and enriches the next, creating a cohesive and impactful scholarly work that speaks to multiple stakeholders.

The main objectives are designed to address the critical gaps in current research and practice and are not merely isolated goals but are interwoven to create a cohesive framework. Firstly, the work seeks to explore and bridge the divide between traditional predictive methodologies and the emerging potentials offered by AI and ML, focusing on enhancing the accuracy, applicability, and ethical considerations of these models. This involves a thorough evaluation of existing models, the identification of key indicators that can enhance predictive accuracy, and the integration of cutting-edge technologies into established frameworks. Additionally, the Thesis aims to translate these technical advancements into actionable strategies for Italian firms.

In the following subsections, each of the main, general, objectives will be analyzed and explored in depth: each objective will be dissected to reveal its underlying components, theoretical significance, and practical implications. Finally, this detailed analysis will be synthesized in articulated RQ.

3.1.1. Bridging Traditional and Modern Predictive Models

This objective seeks to develop a synergistic approach that marries the time-tested methodologies of traditional financial analysis with the innovative power of modern predictive technologies. The focus is on testing a more comprehensive predictive model that leverages the strengths of both approaches: this model aims to capture a wider array of financial indicators and incorporate dynamic, real-time data, thereby offering a more accurate forecast of corporate health. This purpose is not merely about juxtaposing two different approaches to predictive modeling; rather, it seeks to create a synergistic framework that leverages the best of both worlds. The relevance of this objective lies in its potential to transform the way financial distress, crisis, and insolvency are predicted, moving beyond the limitations of historical models while preserving their valuable insights.

Historically, financial ratio analysis has been a bedrock of predictive analytics in corporate finance. Pioneers like Altman (1968) and Beaver (1966) laid the groundwork for this, but the advent of ML and big data analytics has opened new avenues for enhancing the accuracy and robustness of predictive models (lo Conte & Sancetta, 2024; Nguyen et al., 2019). Considering this, the significance of this first main objective is multifaceted:

- First, it addresses the pressing need for more reliable predictive tools in the face of increasing volatility and uncertainty;

- Second, it highlights the importance of methodological innovation in maintaining the relevance of financial analysis in a rapidly evolving landscape;

- Finally, this objective serves as a foundational element of the entire research, as the insights gained from this analysis will inform and support the other objectives of the study, such as enhancing predictive accuracy through AI integration and better representing algorithmic results in order to be better understood by managers and practitioners.

Moreover, this objective is deeply intertwined with the overarching vision that guided this PhD Thesis. Each objective of this research has in fact been carefully designed to build upon the others, creating a cohesive and comprehensive strategy for advancing the field of predictive analytics. The bridging of traditional and modern predictive models is the first step in this process, laying the groundwork for a more integrated and effective empirical approach: it reflects a critical aspect of the PhD journey, which is to innovate while staying grounded in proven methodologies, ultimately contributing to both academic scholarship and practical management practices. In conclusion, the objective of bridging traditional and modern predictive models is not just a comparative exercise; it is a strategic endeavor to craft a more robust, adaptive, and accurate predictive framework. By doing so, this research aims to push the boundaries of what is possible, offering new tools and insights that are crucial today. This objective is, therefore, a fundamental building block of the entire work, setting the stage for the exploration of more advanced topics.

3.1.2. Enhancing Predictive Accuracy and Explainability through AI Integration

Imagine the power of blending time-tested financial wisdom with the cutting-edge capabilities of AI—a fusion that promises to revolutionize how we predict and manage corporate crises. This research embarks on a journey into this dynamic intersection, where innovative algorithms like XGBoost and Gradient Boosting, coupled with advanced representation methods, are set to redefine strategic decision-making and enhance the predictive accuracy of insolvency models. This second objective can be considered the cornerstone of the dissertation, embodying the confluence of technical rigor, empirical research, and strategic relevance. The core aim here is to explore the vast potential of AI and ML technologies to enhance the precision and reliability of predictive models, offering fundamental tools for managers to predict, inform, and manage critical situations. In fact, by leveraging AI's capabilities, new models can:

- Adapt to New Information, as they can continuously learn and evolve as new data becomes available, ensuring that predictions remain relevant and accurate;
- Process Large and Complex Datasets;
- Identify Non-Linear Relationships.

That said, the empirical component of this research will be particularly noteworthy. By utilizing extensive datasets encompassing a wide range of economic and financial Italian data, the study will test and validate AI-enhanced predictive models in real-world scenarios. These datasets will include diverse indicators to capture a holistic view of the factors influencing corporate stability[51]. The choosen models will be rigorously tested to ensure their reliability and applicability. Moreover, the integration of AI into these predictive frameworks will not be limited to enhancing accuracy alone; the dissertation will also explore how AI can improve the visualization, interpretation, and communication of predictive insights. By employing advanced AI-driven data visualization tools, the study will aim to make complex predictive models more accessible and actionable for non-technical stakeholders, particularly managers and executives. This aligns with the broader objective of bridging the gap between complex analytical tools and strategic decision-making.

The importance of this objective is intricately linked to all other objectives of this dissertation, serving as the technical and empirical foundation upon which subsequent goals are built. The insights gained will inform and refine the study's approach to strategic crisis management and prediction. In essence, the successful achievement of this objective will set the stage for the broader impact of the dissertation, enabling the development of considerations, insights and perspectives that are theoretically sound, but also practically invaluable. Key contributions of this objective are:

- *Advancing Academic Literature:* by exploring the integration of AI and ML with traditional financial analysis, this research will contribute to the evolving discourse on corporate crisis and insolvency prediction, offering new insights and methodologies;
- *Empowering Managers:* the practical tools and models developed and analyzed through this research will provide managers with enhanced capabilities to anticipate and mitigate financial distress;

[51] Eco-fin indicators and some qualitative ones.

- *Supporting Strategic Decision-Making*, giving important alerts regarding inefficiences and limitations of algorithms and human behaviors in critical situations.

However, this objective also lays the groundwork for future research and practical applications. In fact, the dynamic integration of AI in predictive models opens new avenues for continuous improvement, where models can be perpetually refined as more data and better algorithms become available. Furthermore, the implications of this research extend beyond crisis prediction, offering potential applications in other areas of risk management and strategic planning. In summary, the enhancement of predictive accuracy through AI integration is the linchpin of this dissertation as it encapsulates the essence of the research journey, combining theoretical exploration with empirical validation to produce a body of work that is both academically rigorous and practically impactful. Anyway, the additional value of this dissertation also lies in enhancing the explainability and understandability of AI and ML models, as well as the representation and application of their results—both within scientific research and practical settings: in academic research, this aspect addresses a common trend in recent literature that often employs AI in ways that lack transparency, sometimes offering sensational conclusions grounded in methodologically rigorous yet poorly explained statistical frameworks. This lack of clarity can make it challenging for managers, management scientists, or non-technical practitioners to understand and trust these AI-generated insights fully. In practice, enhancing explainability is essential to build trust in AI systems, reduce the stigma surrounding crisis prediction, and improve the preventive and proactive management of potential financial distress situations. To achieve this, the dissertation will explore specific AI-driven methods for interpreting and visualizing predictive models, aiming to make complex analyses more transparent and actionable. Techniques such as SHAP (SHapley Additive exPlanations) will be analyzed as potential tools to break down AI results into understandable, feature-level explanations that can be readily communicated to non-experts. Additionally, the study will investigate AI-driven data visualization approaches that highlight key factors and relationships in the models, making it easier for managers and decision-makers to grasp insights intuitively and act with greater confidence.

3.1.3. Strategy for Crisis Management and Sustainability

The objective of exploring *Strategic Implications for Crisis Management and Sustainability* serves as the critical bridge between academic inquiry and real-world application, anchoring theoretical insights in practical, actionable strategies. Derived from a foundation of rigorous empirical studies, this objective is a carefully crafted roadmap to understand the complex interplay between financial distress and long-term sustainability. By analyzing crisis management through a strategic lens, the research then aims to provide a set of optimized, relevant views that empower organizations to anticipate, manage, and mitigate crises effectively, while simultaneously aligning with broader sustainability goals.

This objective, often referred to as the "*strategic compass*" of the Thesis, synthesizes the understanding gained from advanced AI-driven predictive models with the practical need for resilience and adaptability in a volatile economic environment. The goal is to generate actionable insights that not only enhance organizational preparedness but also foster long-term value creation by embedding sustainability at the core of crisis management strategies. This integration of strategic foresight with sustainability is crucial in an era where the consequences of corporate missteps extend far beyond financial loss, impacting societal and environmental dimensions as well. As highlighted by Lozano (2020), in fact, companies that embed various kind of sustainability into their crisis management

frameworks not only enhance their resilience but also position themselves as leaders in an increasingly conscientious market.

To elevate the strategic relevance of this objective, the focus is squarely on harnessing the power of AI-driven tools and advanced predictive models to enhance decision-making in critical situations. In conclusion, through this objective, the research strives to transcend traditional crisis management approaches, offering a sophisticated blend of predictive analytics and strategic foresight that can be directly applied to real-world scenarios. In essence, on the other side, it represents the confluence of academia and practice, where theoretical models are tested, refined, and ultimately transformed into practical tools that managers can deploy to ensure both immediate stability and long-term sustainability.

3.2. Other Objectives

While the main objectives of this study lay the foundation for advancing the field of corporate crisis and insolvency prediction, it is the secondary objectives that provide the essential finishing touches—like the final brushstrokes on a masterpiece. These secondary, yet equally crucial, goals serve as the connective tissue that binds the primary objectives together, ensuring that the research is both comprehensive and applicable to real-world scenarios. These may be considered the unsung heroes of the research, quietly but significantly supporting the broader goals: they go into the finer details, offering insights that enhance our understanding of the main objectives and ensuring that no stone is left unturned. In synthetis, they, though secondary in name, play a primary role in shaping a well-rounded, insightful body of work that bridges theory and practice.

3.2.1. Identifying Key Indicators and Variables

This sub-objective, affectionately dubbed the "*Compass of Insight*," plays a pivotal role in the overarching narrative of this research. While the main objectives chart the course, this compass ensures that we are steering in the right direction, offering precise navigational points that guide us through the complex terrain of prediction. At its core, this objective is about pinpointing the crucial indicators and variables that enhance the accuracy and relevance of predictive models. It is the foundation that serves as a critical link between the theoretical insights derived from the literature and the empirical testing that brings those theories to life.

This process is deeply intertwined with the theoretical framework of the Thesis, drawing from the wealth of existing literature while pushing the boundaries by introducing and testing variables in an empirical setting. The importance of this sub-objective is underscored by its dual relevance:

- *Theoretical Relevance:* the literature reviewed offers a plethora of indicators and variables, but this objective seeks to distill these into a refined set of key indicators that are most impactful in predicting financial distress and insolvency;

- *Empirical Relevance:* by rigorously testing these indicators against real-world data, the study not only evaluates their performance but also uncovers new variables that could offer even greater predictive power. This empirical exploration is crucial for refining predictive models and ensuring they remain relevant in an evolving financial landscape.

In essence, the "*Compass of Insight*" is not just a sub-objective—it is the strategic linchpin that ties together the theoretical and empirical threads of this research, ensuring that the findings are both academically robust and practically valuable. However, not only will this dissertation analyze and discuss both financial and non-financial indicators, but it will also delve into specific results from multi-algorithmic analyses. Typically, certain outputs—often treated as errors—can serve as alert variables within the context of this study, offering a unique lens to deepen scientific inquiry and enhance our understanding of real-world phenomena: one such example is the number of false positives in confusion matrices generated by ML algorithms. Rather than dismissing these as purely erroneous classifications, the study will explore their potential to act as early warning signals, indicating areas that may require further examination.

3.2.2. Addressing Ethical Considerations in a Multi-Algorithmic Scenario

The sub-objective of addressing ethical considerations, enhancing explainability, and aligning with sustainability goals—dubbed the "*Simplicity, Integrity, and Sustainability Game*"—plays a pivotal role in balancing the complexity of diverse AI-driven models with the ethical responsibility and sustainable development imperatives that are essential in today's global landscape. The fundamental premise here is that the most effective predictive models are those that not only deliver accurate and actionable insights but also adhere to principles of transparency, ethical responsibility, and sustainability (Rudin, 2019; Doshi-Velez & Kim, 2017). So, the "*Simplicity, Integrity, and Sustainability Game*" extends beyond the confines of technical robustness to incorporate a broader vision of corporate responsibility. As AI and ML technologies become more deeply embedded in corporate strategies, the need for these tools to be ethically sound, transparent, and aligned with sustainability goals becomes increasingly critical, and this is not just an academic exercise but a strategic necessity for fostering long-term value creation, particularly in the face of mounting global pressures. This goal is then anchored on two intertwined targets:

- *Ethical Responsibility and Alignment:* the integration of AI in predictive analytics brings with it significant ethical considerations, including fairness, bias, and privacy concerns. However, in the context of sustainability, this sub-objective goes a step further by examining how AI models can create new and intelligent relationships between firms, stakeholders, legislators and practitioners;

- *Explainability and Transparency:* in the literature, there is a growing consensus on the importance of transparency and interpretability in AI models, particularly those used in high-stakes decision-making (Gilpin et al., 2018). This is especially relevant when considering the implications for sustainability, where the outcomes of predictive models can influence decisions with long-term environmental and social impacts. By enhancing the explainability of these models, this sub-objective aims to ensure that all stakeholders can understand and trust the insights provided, thereby facilitating more informed and responsible decision-making;

To summarize, this sub-goal is a strategic approach that underscores the importance of integrating ethical and sustainable considerations into the core of AI-driven predictive analytics.

Moreover, this objective is not an isolated concern but is intricately linked to the broader goals of the Thesis: whether the focus is on enhancing predictive accuracy through AI integration or bridging traditional and modern predictive models, the emphasis on ethical responsibility, transparency, and sustainability ensures that the outcomes are not only technically sound but also socially and environmentally responsible.

3.2.3. Providing a Comprehensive Guide for Managers

The goal here is to distill the advanced theoretical insights and empirical findings of this research into practical strategies that managers can implement to manage financial distress, crisis, and insolvency with confidence and precision. This sub-objective, aptly termed the "*Manager's Compass*," will help managers leverage the power of predictive analytics, AI, and ML to make informed decisions that not only mitigate risks but also uncover new opportunities for growth and sustainability. Key elements of the "*Manager's Compass*" include:

- *Decision-Making Tools*, as this work will translate complex predictive models into easy-to-understand tools that managers can use in their day-to-day decision-making processes. Whether it's identifying early warning signs of financial distress or forecasting the impact of strategic decisions, these tools will empower managers to act proactively rather than reactively;

- *Strategic Insights:* the guide will offer tailored strategies that align predictive insights with corporate objectives, ensuring that the integration of AI and ML technologies enhances, rather than complicates, the strategic planning process;

- *Adaptability for SMEs*: understanding the unique challenges faced by Small and Medium-sized Enterprises (SMEs), the "*Manager's Compass*" will include specific recommendations for resource-constrained organizations. This will help SMEs scale their use of predictive analytics in a manner that is both feasible and impactful (Gartner, 2020).

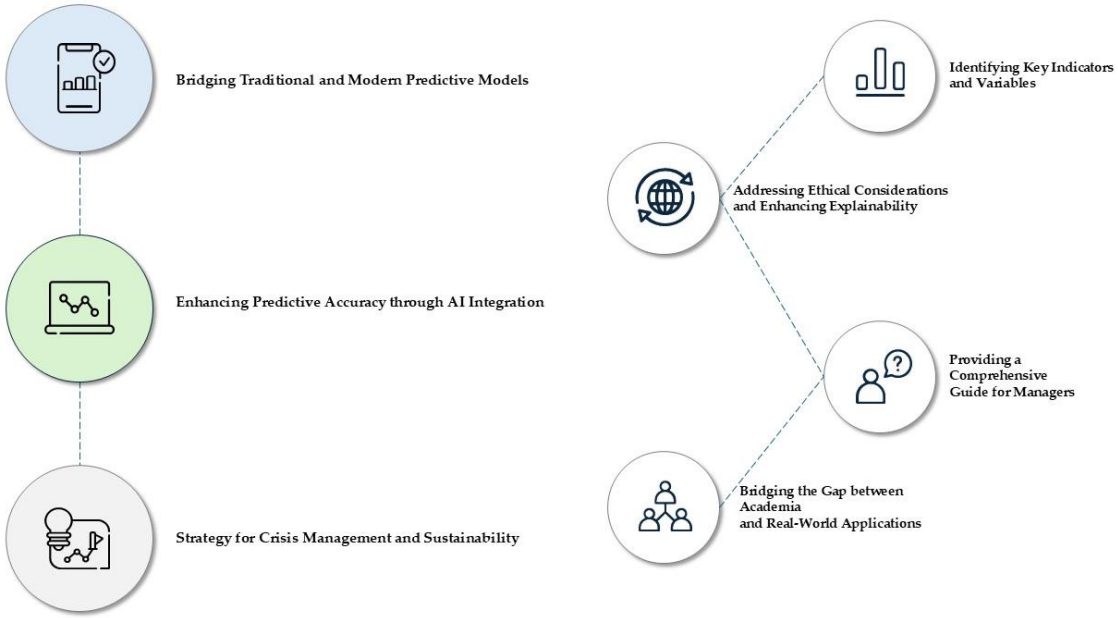
3.2.4. Bridging the Gap between Academia and Real-World Applications

This last objective focuses on fostering a meaningful connection between theoretical research and practical implementation, striving to make academic insights not just intellectually enriching but also practically applicable. The goal is to create a synergy where rigorous scholarly research informs and enhances real-world business practices, particularly in the challenging areas of crisis and insolvency prediction. Therefore, this study aims to validate its findings in a practical context, ensuring that they are not only academically sound but also actionable: this approach bridges the often-cited gap between academia and industry. All of this is integral to the overall purpose of the research, as it ensures that the theoretical advancements made through this study translate into tangible benefits for businesses.

Building on its dual academic and practical significance, this study will not shy away from conducting a critical and in-depth examination of the limitations and challenges inherent in AI models,

managerial thought, and economic-financial structures. Specifically, it aims to provide a thorough critical assessment of empirical complexities, linking alerts and results back to practical considerations and actions that a professional in the economic and managerial fields would undertake. This approach enhances the relevance of the study's conclusions, insights, and empirical findings by simplifying complex models, algorithms, and theoretical concepts into formats that are comprehensible and actionable for managers. Moreover, the research will integrate technical insights and practical validations, grounded in the realities of the professional world, to ensure that its contributions are not also aligned with Italian business needs. This lends a greater practical weight to the dissertation's findings, enabling managers to gain valuable, applicable knowledge supported by robust academic and empirical validation. Figure 1 offers a distinctive view of general research objectives.

Figure 1: Research Objectives



Source: own elaboration

4. Research Questions: Some Thoughts

Formulating Research Questions for a study of this magnitude is akin to distilling a vast ocean of knowledge into a few precise and targeted inquiries. The complexity and depth of the topics explored in this Thesis, from the integration of AI in predictive models to the ethical implications and strategic applications of these technologies, make it a formidable challenge to encapsulate the entire scope of the research in just a few questions. However, the following questions are not merely a simplification but are designed to serve as the guiding lights of this study, illuminating the path through the entire analysis. The following points are then crafted to reflect the multifaceted objectives of the study, drawing from the theoretical insights and practical considerations discussed in the preceding sections. They are designed to drive the inquiry forward, ensuring that the research remains focused on its core goals. Finally, these questions will serve as the foundation upon which the entire research is built, guiding the exploration and analysis.

4.1. Research Questions: a List

RQ1: *How do the implementation of AI-based techniques, improve the accuracy and the efficiency of insolvency and crisis prediction for Italian firms, and what refinements can be made to enhance their performance?*

- Sub-questions and topics:
    - *In what specific ways do AI-driven models outperform traditional financial ratio analysis, and how can these insights be leveraged for strategic advantage and crisis management?*
    - *How can advanced ML models, such as XGBoost and Random Forest, be optimized and taylored, mainly for Italian SMEs, to improve the precision and reliability of early warning signals for real-world applications?*
    - *What specific limitations do AI-driven predictive models face, and how can these challenges be addressed to ensure practical implementation?*

RQ2: *What are the potential real-life impacts of AI-driven insights on strategic decision-making, ensuring simplified applicability and enhanced explainability?*

- Sub-questions and topics:
    - *How can predictive insights from AI and ML be effectively communicated to managers and stakeholders, ensuring clarity and applicability in non-technical, real-world settings?*
    - *What limitations exist in managerial understanding and adoption of AI-driven insights, and what actions can bridge the gap between academic advancements and practical application in the corporate world?*
    - *How can organizations avoid the pitfalls of over-reliance on AI predictions, balancing AI insights with human judgment to support robust, ethically responsible decision-making in high-stakes scenarios?*

These RQ are grounded in an inductive approach, as suggested by Tharenou et al. (2018), aiming to build knowledge from observed data patterns rather than testing predefined hypotheses. This approach allows the study to better explore complex, real-world phenomena, The research will then employ a mixed-methods design, combining qualitative insights with quantitative analysis, which provides a comprehensive view of the effectiveness and practical applicability of AI-driven predictive models. Furthermore, a multi-algorithmic framework will be utilized, integrating various ML techniques to maximize predictive accuracy and robustness. This design allows for flexibility in identifying the most effective algorithms while ensuring a rich empirical foundation, supporting findings for academia and industry.

4.2. Methodological Framework for Addressing Research Questions

Today, the evolving landscape of business analytics necessitates a holistic, multi-methodological approach to fully capture the complexity of the challenges at hand. So, this Thesis embraces that complexity, drawing on a wide array of methodological techniques to build a robust and comprehensive understanding of the subject matter: the value of this approach lies in its ability to

synthesize a vast body of knowledge into a cohesive framework. The integration of AI and ML into predictive models for corporate crises is not a task that can be approached with a single lens; it demands a dynamic and versatile methodological strategy. Traditional financial ratio analysis, in fact, while foundational, must be complemented by cutting-edge technologies capable of adapting to the rapid changes in today's economic environment. Therefore, this work employs a range of methodologies, from basic quantitative data analysis to advanced ML algorithms, ensuring that each RQ is addressed with the appropriate tools. In synthesis, the methodological framework for this research primarily employs a quantitative approach but incorporates qualitative elements to delve deeply into specific aspects of predictive analysis in contexts of crises; the study's dual perspective is anchored in a solid and critical inductive approach, combining rigorous scientific insights with practical, professional observations.

This integration of theory and practice adds significant value to the entire dissertation. It, in particular, begins with a comprehensive and critical analysis of the literature, drawing from both traditional and contemporary sources through a multidisciplinary and dynamic approach. It then transitions to a detailed examination of the regulatory landscape surrounding crisis management, establishing a legislative context for the empirical findings. The empirical component relies on a multi-algorithmic approach connected to Intelligent Big Data Analytics, enabling robust predictive analysis grounded in real-world data. This methodological progression—spanning critical literature review, legislative examination, and empirical validation—ensures a holistic and impactful exploration of AI's potential.

Finally, the subsequent paragraphs will provide a comprehensive roadmap of the research design, offering insights into how these methods collectively contribute to achieving the study's objectives. However, before progressing to the detailed exploration of methodologies, it is essential to first examine the terminology and foundational concepts that are central to this study. A deep understanding of these elements is in fact critical to fully comprehend the objectives and implications of the research. Therefore, the next Chapter will be dedicated to a comprehensive examination of these fundamentals, laying a solid groundwork for the analyses that follow.

Chapter 2

Understanding the Fundamentals: Key Concepts and Terminologies

Chapter Prelude

Before diving into the advanced strategies and innovative techniques that will define this journey, it's crucial to get bearings with the key concepts and terminologies. Understanding these fundamentals is not just helpful—it's essential. This Chapter unpacks the core ideas that shape the study of corporate crisis and insolvency prediction, and can be thought of as "learning the language of the field", ensuring that as the work moves forward, every insight and conclusion is built on a solid understanding. This foundational knowledge will be a compass, guiding the reader through the complexities ahead.

1. The Importance of Grasping the Basics

In the dynamic and increasingly complex environment of predictive analysis, grasping the fundamental concepts and terminologies is paramount, particularly when engaging with intelligent and innovative tools. The methodologies and models that underpin these predictions, as explored in the works of Roubini (2010) and others, highlight the intricate balance between economic forces, and the structural vulnerabilities that can precipitate financial distress and the historical evolution of crises, detailed in extensive studies such as those by Gennaioli and Shleifer (2013), illustrates the critical need for a deep understanding of the mechanisms that drive these events. These scholars have in fact shown how financial turmoil often emerges from a confluence of factors, all of which must be carefully analyzed and understood to develop effective predictive models.

Today, the pervasive integration of AI and ML into financial analysis adds a new layer of complexity. These innovations promise unprecedented accuracy in forecasting but also require a solid foundation in both traditional financial principles and advanced algorithmic techniques. In fact, as recent research by Bacher et al. (2023) indicates, the success of AI-driven predictive models depends heavily on the quality of the data and the sophistication of the algorithms. However, without a thorough understanding of the underlying principles, these models risk delivering misleading results, which could lead to significant strategic missteps. In light of this, this Chapter is dedicated to exploring the key concepts and terminologies that form the backbone of predictive analytics and crisis-related environment. Drawing on the rich literature and case studies of past crises, research will trace the evolution of these concepts, examining how they have been shaped by historical events and how they continue to adapt to the demands of a rapidly changing global economy.

1.1. Core Concepts

In embarking on this exploration it is imperative to first lay a solid foundation by clearly defining and differentiating the key concepts that underpin this study. The field of corporate finance and management is rife with terminologies that, while seemingly synonymous, carry distinct implications depending on the context—be it legal, financial, or strategic. A deepen understanding of these terms is then essential not only for grasping the theoretical underpinnings of the entire research but also for applying these insights in practical scenarios. The first step in this journey involves a thorough examination of the concepts of corporate crisis, insolvency, and bankruptcy. In fact, although these terms are often used interchangeably in both scholarly literature and industry practice, they represent different stages and aspects of financial distress within an organization. Understanding these distinctions is then critical, as each term implies a different set of circumstances, legal frameworks, and strategic responses. In particular:

- *Corporate Crisis*[52]: at its core, a corporate crisis refers to a significant disruption in a company's operations or strategic direction, which may stem from internal operational failures or external shocks such as economic downturns or regulatory changes. This Chapter of the Thesis will dissect the multifaceted nature of corporate crises, distinguishing between those that arise from internal mismanagement—like strategic errors and operational inefficiencies—and those triggered by external forces. It is crucial to note that a corporate crisis does not necessarily lead

[52] In this Thesis, "*Corporate Crisis*" refers to firm-level crises, with a focus on enhancing managerial decision-making and strategy through AI integration, rather than addressing macroeconomic crises.

to insolvency or bankruptcy but often serves as a precursor, highlighting the need for early detection and management;

- *Insolvency:* moving beyond the general concept of crisis, insolvency is more specifically defined as the financial state where a company's liabilities exceed its assets, rendering it unable to meet its debt obligations. This section will then explore the financial indicators that precede bankruptcy, such as cash flow problems, excessive debt, and deteriorating profitability. It will also analyze the various legal interpretations of insolvency across different jurisdictions, emphasizing the importance of understanding these differences for effective crisis management and prediction;

- *Bankruptcy:* it represents a formal legal declaration of a company's inability to repay its debts, leading to either liquidation or reorganization under the supervision of a court. This section will differentiate between insolvency and bankruptcy, highlighting that while insolvency is a financial condition, bankruptcy is a legal process with specific procedures and outcomes. The discussion will extend to the varying legal frameworks and procedures that govern bankruptcy across different countries, particularly focusing on the distinctions between Italian and European bankruptcy laws.

Misunderstandings in these areas can lead to strategic missteps and inappropriate management responses. Furthermore, this section will also examine the terminologies related to prediction methodologies, particularly the differences between terms like "*predict*", "*forecast*", and *"anticipate"*. In predictive analytics, the distinction between these terms can significantly influence the interpretation of results and the strategic decisions that follow. So, this dissertation will clarify the usage of these terms within the context of the study, justifying the choices made and explaining their relevance to the methodologies employed. The discussion will also touch upon traditional and innovative models of crisis and insolvency prediction, setting the stage for a more detailed exploration in subsequent sections. Traditional models, such as those based on financial ratios, will be contrasted with newer, more sophisticated techniques that leverage ML and big data analytics.

In addition, a key aspect of this exploration will be the regulatory frameworks that shape the prediction and management of corporate crises and insolvency. Both Italian and broader European legal contexts will be scrutinized to provide a robust legal foundation for the predictive models discussed later in the work. This regulatory analysis is not merely supplementary but it is essential for contextualizing the practical implications of predictive analytics in different legal environments. In summary, this entire section on Core Concepts serves as a comprehensive primer, equipping readers with the necessary terminology and conceptual clarity to understand the evolving landscape of predictive analytics. However, previous ones were the principal terminologies essential for understanding the foundational concepts of this study; in the following sections, the work will outline additional key elements and terminologies that are equally crucial to the overall framework of this research.

1.2. The Dynamic Nature of Core Concepts

Today, the core concepts of crisis, insolvency, and bankruptcy are not static entities but are continuously shaped by emerging technologies, regulatory shifts, and the changing global economy. While the present study utilizes well-established definitions and frameworks that have been validated in both academic literature and real-world applications, acknowledging the dynamic nature of these concepts is necessary. The fluidity inherent in these terms, in fact, reflects the broader, rapidly changing environment in which they are applied, necessitating a forward-thinking approach that anticipates future developments. Therefore, one of the key strengths of this research lies in its ability to marry traditional wisdom with contemporary advancements. The integration of AI and ML into predictive models, for instance, has the potential to redefine what we understand by "*crisis*" or "*insolvency*" and what may constitute a crisis in one region or sector may be a routine fluctuation in another, necessitating a more flexible approach to defining and responding to such events. This research, therefore, not only adheres to tried-and-tested concepts but also critically examines them in the context of contemporary and emerging challenges.

As governments and international bodies adapt to the complexities of modern economies, the legal definitions and implications of insolvency, bankruptcy, and corporate crisis are also subject to change. In this regard, the dissertation aims to stay ahead of such developments, offering a forward-looking perspective that considers potential future shifts also in regulatory landscapes. This dynamic approach is further evidenced by the intention to contribute to the evolving discourse on these concepts.

2. Defining Corporate Crisis

Corporate crises, often characterized by the deterioration of a company's economic and financial performance, are critical events that can threaten the survival of an organization. The significance of understanding and defining these situation cannot be overstated, as they have been the subject of numerous definitions and interpretations, reflecting the complexity and multifaceted nature of these events. This dissertation considers a corporate crisis, from an Italian-context perspective, primarily as the significant deterioration of a company's economic and financial performance, which, if not addressed promptly, can escalate into insolvency or bankruptcy (CCII, 2022)[53]. This definition underscores the importance of early detection and strategic intervention, which are central to the predictive models and methodologies explored in later Chapters. However, understanding corporate crises requires more than a singular definition: a thorough examination from multiple perspectives, including both a business-centric and a managerial viewpoint is then necessary. The managerial perspective, in particular, holds significant importance within this work, as it provides insights into how decision-makers can effectively manage crises, ensuring the sustainability and resilience of their organizations.

[53] A notable example in Italy is Alitalia, which faced high debt levels, operational inefficiencies, and continuous cash flow erosion. Despite multiple restructuring efforts and bailout attempts, the lack of timely and effective corrective actions ultimately led the company into extraordinary administration, highlighting the critical importance of swift and structured turnaround strategies in severe crisis scenarios

2.1. Firm-Level Focus

In the often turbulent landscape of modern business, the term "*Corporate Crisis*" carries significant weight. In fact, it conjures images of unexpected challenges, financial instability, and the urgent need for decisive action. Yet, in the context of this dissertation, the concept of a corporate crisis is far more sharpened and specifically tailored to the dynamics of individual firms. So, the journey of understanding what constitutes a corporate crisis begins here, where we peel back the layers of complexity to reveal not just the challenges but also the strategic opportunities that arise when a firm faces the brink of instability.

2.1.1. Drawing the Boundaries

When speaking of a corporate crisis in this Thesis, it refers to crises that occur at the level of individual firms, distinguishing them from the broader macroeconomic crises that impact entire nations or global markets. This distinction is not just semantic but is critical to the focus and objectives of the research: while macroeconomic crises, such as the global financial crisis of 2008 or the economic downturns triggered by global pandemics, have far-reaching effects on multiple firms and industries, the crisis at the firm level involves unique challenges that require tailored managerial strategies[54]. A corporate crisis at the firm level often manifests as a significant disruption in the company's operations, financial health, or strategic direction; this disruption could be caused by internal factors, such as poor management decisions, operational inefficiencies, or financial mismanagement (Guatri, 1986). Alternatively, it might be triggered by external forces like regulatory changes, economic downturns, or competitive pressures. The key difference lies in the scope and scale of the impact—while macro-level crises affect the economy as a whole, firm-level crises are confined to the organizational boundaries of a single company but can be just as devastating in their effects (Guatri, 1986). To fully grasp the implications of a corporate crisis at the firm level, it's then essential to differentiate it from macro-level crises. The following examples and distinctions highlight the unique characteristics and challenges of each:

1. *Firm-Level Crises:*
    - *Operational Disruptions:* For instance, a company might face a crisis due to a critical failure in its supply chain, resulting in a halt in production and significant financial losses. This type of crisis requires immediate managerial intervention to restore operations and mitigate financial damage;
    - *Strategic Failures:* A firm might enter a crisis if its strategic decisions, such as an unsuccessful merger or market entry, fail to deliver expected results, leading to a loss of investor confidence and a sharp decline in stock prices;

[54] The 2008 global financial crisis, often regarded as the most severe economic downturn since the Great Depression, was triggered by the collapse of the U.S. housing market and the subsequent failure of large financial institutions, including Lehman Brothers. The crisis was fueled by complex financial instruments, such as mortgage-backed securities (MBS) and collateralized debt obligations (CDOs), which masked underlying risks and led to widespread market instability. As the contagion spread, liquidity dried up, and interbank lending froze, resulting in significant credit constraints across the global economy. Governments worldwide implemented unprecedented bailout packages and monetary easing measures to stabilize financial markets and prevent a systemic collapse, marking a pivotal moment in modern financial regulation, notably the introduction of the Dodd-Frank Act and the Basel III standards

- *Financial Distress:* A classic example of a firm-level crisis is when a company finds itself unable to meet its debt obligations, leading to insolvency or the threat of bankruptcy. It requires sophisticated financial restructuring and, sometimes, negotiation with creditors. This particular type of crisis is the central focus of this research, as it represents the most critical juncture for an organization—where the stakes are highest, and the need for effective intervention is most urgent.

2. *Macro-Level Crises:*

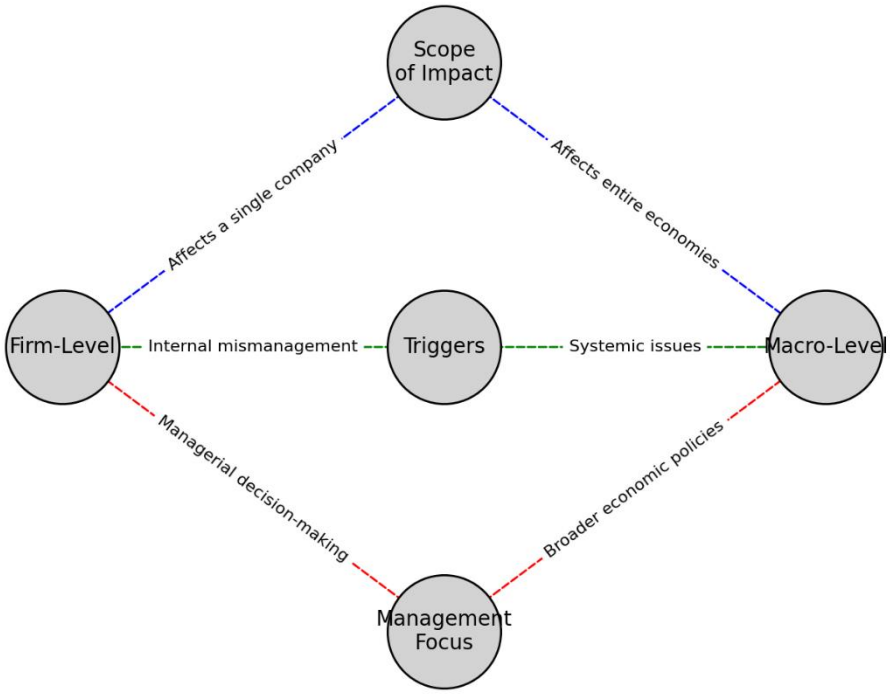
   - *Economic Recession:* A nationwide economic downturn that results in reduced consumer spending, widespread unemployment, and a contraction in economic activity. Such crises affect all firms within the economy, though their impacts may vary based on the industry and the firm's resilience;
   - *Global Financial Crisis:* The 2008 financial crisis is a prime example of a macro-level crisis that originated in the housing market and rapidly spread across global financial systems, leading to the collapse of major financial institutions and triggering recessions in multiple countries;
   - *Pandemics and Global Health Crises:* The covid-19 pandemic caused an unprecedented macroeconomic crisis, with widespread business closures, disruptions in global supply chains, and massive economic stimulus packages to prevent complete economic collapse.

2.1.2. Key Differences

Key differences, summarized in Figure 2, are the following (Guatri, 1986):

- *Scope of Impact:*
  - *Firm-Level:* Affects a single company, often requiring targeted, internal strategies for resolution;
  - *Macro-Level:* Affects entire economies, with impacts felt across multiple industries and regions.
- *Triggers:*
  - *Firm-Level:* Can be caused by internal mismanagement, operational failures, or competitive pressures;
  - *Macro-Level:* Often triggered by systemic issues, such as economic recessions, financial market collapses, or global crises like pandemics;
- *Management Focus:*
  - *Firm-Level:* Requires managerial decision-making and strategic planning to manage the crisis;
  - *Macro-Level:* Involves broader economic policies, government interventions, and sometimes global cooperation to mitigate the effects.

Figure 2: Key Differences - Firm-Level vs Macro-Level



Source: own elaboration

Furthermore, the managerial implications of a firm-level crisis are particularly significant, as they demand a proactive and strategic approach to decision-making. In this perspective, this work emphasizes the importance of understanding firm-level crises from a managerial perspective, where the application of AI and ML can play a pivotal role. Moreover, these technologies enable a more deepen understanding of the factors contributing to the crisis, whether they are internal inefficiencies or external threats, thereby facilitating a more targeted and effective response.

2.2. Strands of Literature

This research meticulously examines and analyzes the concept of corporate crisis, drawing on both practical experiences and an extensive review of the relevant literature from two crucial perspectives: the managerial and the business-oriented or corporate-oriented *("aziendalistico")* viewpoints. These dual perspectives are fundamental to the objectives of this study, offering a comprehensive and 360-degree understanding of the phenomenon and its far-reaching implications. By exploring the crisis through both lenses, the work also aims to provide a holistic insight into the dynamics of crisis management. Moreover, the research does not overlook the practical concerns of practitioners and policymakers, recognizing that a thorough and well-rounded analysis must address their needs as well. This approach guarantees that the study's conclusions and recommendations are both relevant and actionable.

The managerial literature, which often emphasizes strategic decision-making, leadership, and organizational behavior during crises, is juxtaposed with the other strand of literature, which focuses more on the operational, financial, and structural aspects of businesses in distress. This distinction is crucial as it highlights the different, yet complementary, ways in which corporate crises can be

understood, managed, and ultimately resolved. Finally, this alignment is essential for answering the RQ with depth and precision. The emphasis on crisis management from a managerial and *"corporate"* perspective also underscores the critical role of resilience and the integration of AI technologies; from a managerial standpoint, crises represent pivotal moments that test the leadership, decision-making processes, and adaptive capacity of an organization. It is during these times that the importance of a well-structured and agile management approach becomes evident, as leaders must make swift, informed decisions that can determine the survival and future success of the firm. AI-driven tools and predictive analytics are increasingly central to these processes, offering managers real-time insights and data-driven guidance that enhance their ability to anticipate challenges and respond effectively. The use of AI in crisis management, in fact, enables the identification of emerging risks, the modeling of potential scenarios, and the development of proactive strategies that can mitigate the impact of a crisis before it fully unfolds.

On the other hand, the corporate perspective focuses on the structural and operational dimensions of crises, analyzing how financial health, organizational processes, and resource allocation influence a company's ability to cope with and recover from adverse events. This viewpoint emphasizes the importance of maintaining a robust financial position, efficient operational systems, and strategic resource management to ensure that the organization can overcome crises without compromising its long-term viability. These perspectives highlight the need for a balanced approach that integrates both the strategic foresight and adaptability emphasized in managerial literature with the operational and financial rigor central to corporate analysis.

2.3. Between Normalcy and Dysfunction

Historically, corporate crises were viewed as rare and detrimental events, often perceived as symptoms of severe dysfunction within a business (lo Conte & Sancetta, 2024; Guatri, 1986). This perception prevailed until the 1970s when it became clear that crises were not mere anomalies but rather recurring features within modern industrial systems (Nguyen & Han, 2019). Today, the prevalence of corporate crises is undeniable as they are recognized as structural elements of the economic landscape, closely linked to the inherent volatility of modern economies. This volatility is driven by factors such as prolonged economic downturns, increased market competition, and external pressures from emerging economies. In many cases, companies respond to these challenges with short-term financial strategies aimed at resolving immediate liquidity problems, often neglecting long-term economic and financial stability. This reliance on short-term debt can exacerbate crises, particularly when combined with declining profitability and rising interest rates.

In the Italian context, the fragility of the banking system further complicates the corporate landscape. Italian banks, as the primary lenders to businesses, face significant challenges in assessing and managing corporate credit risks. The fragmentation of debt across multiple banks, many of which are small and tend to follow the lead of larger institutions, exacerbates the difficulty in preemptively identifying and managing potential crises. The lack of comprehensive oversight and the poor quality of financial information disclosed by companies further complicate the early detection and management of corporate crises. However, it is also crucial to recognize that these crises can be seen as a natural part of the economic cycle. In some cases, the failure of companies, especially those on the periphery of the market, may be a necessary cost for the rebalancing of industries or the broader economy. This process can ultimately lead to a more competitive and resilient market, where only the most efficient and innovative firms survive. Furthermore, the survival of a distressed company can have far-reaching

negative effects, disrupting the broader economic system by distorting prices, reducing competition, and misallocating resources.

In the context of this research, a corporate crisis is then defined as a critical disruption within a company that threatens its ability to maintain operational and financial stability[55]. This disruption may stem from internal management failures, strategic missteps, or external shocks. From a managerial perspective, a crisis represents not merely a financial or operational setback but a pivotal moment that demands decisive and strategic leadership to navigate the company through uncertainty and restore equilibrium. This definition could be considered tour cornerstone and guiding star, anchoring the analysis and driving the exploration of strategies that can enhance resilience and ensure sustainable success in the face of adversity. Furthermore, this managerial focus underscores the importance of understanding crises not just as challenges to be managed, but as opportunities to refine strategies, improve decision-making processes, and ultimately strengthen the organization's ability to thrive in a complex environment.

2.3.1. Crisis as a Breakdown of Coherence in the Corporate Context

Managerial literature also refers to the crisis concept as a "*breakdown of coherence*" (Passeri et al., 2011). For this strand of literature, in fact, in the pursuit of equilibrium within a company, it is essential to consider the coherence relationships that must exist both among the various internal components of the company and between the company and its external environment. This system of coherence is inherently dynamic and unstable, as the relationships among the elements within the company can change over time. Each moment presents a potentially different structural configuration, with coherence relationships in a constant state of flux (Passeri et al., 2011). This dynamism, already cited in previous paragraphs, can lead to the breakdown of existing coherence relationships, followed by their reconfiguration through adaptive processes. Therefore, it is normal for temporary dysfunctions to occur throughout the evolution of a company. The negativity of these dysfunctions should not be assessed in the short term but rather in terms of the persistence of incoherences over time, and a prolonged breakdown of coherence, one that cannot be reestablished, highlights the company's inability to adapt to new circumstances.

The presence of incoherences within the corporate system is a natural consequence of the dynamism inherent in business elements, and the resolution of these incoherences depends on the company's "*flexibility*"—its ability to adapt to new situations. It is important to note that the structure of the coherence system, which includes both internal and external coherence relationships, aids in identifying dysfunctions or incoherences. These incoherences can be categorized into two main types (Passeri et al., 2011) (Figure 3):

[55] Below are six types of crises:

- *Inefficiency crisis;*
- *Overcapacity/overproduction crisis;*
- *Product decay crisis;*
- *Crisis due to marketing deficiencies and errors;*
- *Crisis stemming from a lack of planning, strategy, and innovation;*
- *Financial/asset imbalance crisis.*

In this Thesis, each specific cause of the crisis will not be analyzed in detail. Instead, the focus will be on broader aspects of corporate crises, with reference made to the above categorization for those interested in exploring the underlying causes further see in text citations.

1. *Internal Incoherences:* These occur when coherence relationships break down within the company itself. Internal incoherences can be further subdivided into two primary categories:

    - *Incoherences Within Configurations:* These occur among the internal elements of a specific configuration. This type of incoherence is particularly relevant in the operational system of the company. For instance, problems of coordination or integration may arise between different functions within the operational system. A lack of alignment or synchronization within the company's internal processes can lead to inefficiencies and operational failures.

    - *Incoherences Between Configurations:* These manifest when there is an insufficient degree of integration between different configurations. These configurations may include ownership, top management, and the operational system. Inefficiency or a lack of coordination among these various decision-making levels can lead to both operational and administrative problems. For example, the ownership may not be adequately aligned with the strategy set by top management, or the operational system may not be effectively supported by ownership and top management. These incoherences between configurations can hinder the company's ability to adapt to changes and achieve its objectives[56].

2. *External Incoherences:* These occur when the dysfunction persists over time between the company and the external environment, highlighting the company's inability to adapt to market demands or changes in the surrounding environment. External incoherences can be described as "*rigidity*," indicating that the company is either incapable or unwilling to adapt quickly to external changes without incurring economic losses. The ability of a company to adapt to changes is contingent upon its flexibility. A highly flexible company has a greater capacity to respond to changes and, theoretically, has a lower likelihood of falling into an irreversible crisis.

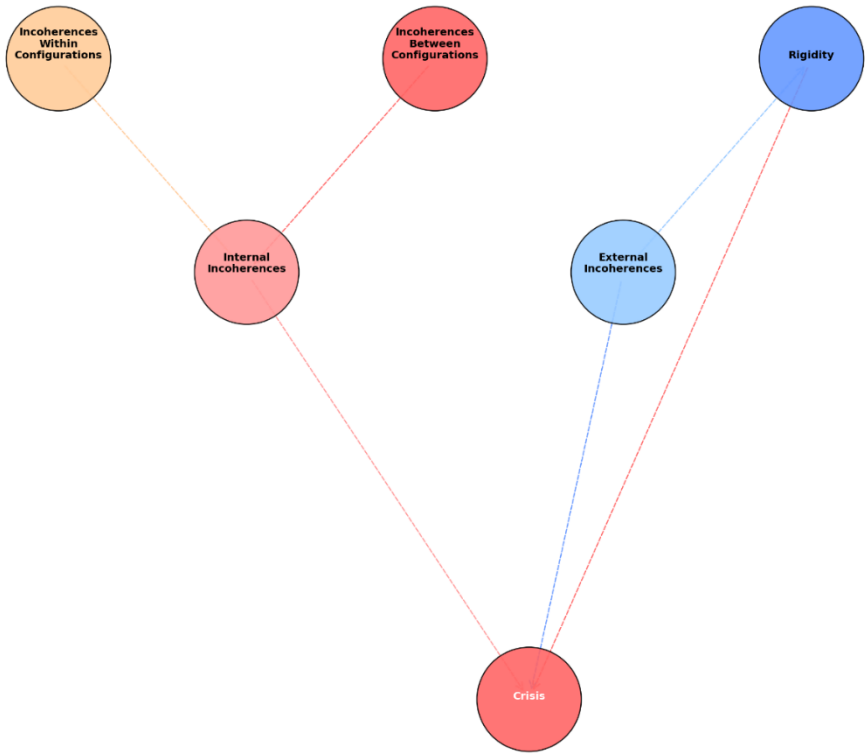
It is crucial to recognize that, in addition to direct incoherences, there are also indirect incoherences that arise from the breakdown of exclusively external relationships: these incoherences may not directly impact the company, but they can create conditions that, over time, influence the company's operational and strategic effectiveness[57]. The concept of coherence and the potential for its

[56] In Italy, the case of Telecom Italia illustrates these incoherences in a real-world setting. Over the years, Telecom Italia experienced shifts in ownership and conflicting strategic directives, resulting from frequent changes in majority shareholders, including significant influence from private equity and foreign investors. This variability at the ownership level has often been misaligned with the long-term strategic vision required by top management to foster sustainable growth. For instance, while certain shareholders advocated for rapid restructuring and short-term profitability, top management identified a need for substantial investment in digital infrastructure and innovation, requiring a long-term commitment. This disconnect was further exacerbated by the operational system, which faced challenges in executing strategy amid uncertain directives from both ownership and management. The resulting incoherences led to operational inefficiencies, high turnover among executive leadership, and delays in strategic initiatives, such as the deployment of next-generation broadband networks. The fragmented alignment among ownership, management, and operations restricted Telecom Italia's agility in adapting to competitive pressures and technological changes, ultimately impacting its market performance and ability to meet regulatory expectations.

[57] Indirect incoherences arise when external factors create latent misalignments that subtly impact an organization's strategic coherence and operational fluidity. While these do not immediately disrupt workflows, they introduce vulnerabilities that can accumulate over time, affecting key performance indicators (KPIs) such as Return on Assets (ROA), Operational Efficiency Ratios, and Strategic Alignment Indexes. These misalignments stem from complex interdependencies with external systems, and common indicators of such incoherences include:
Geopolitical Instability Indicators: Geopolitical Risk Index (GRI) and Global Political Uncertainty Index (GPUI) are critical in monitoring potential disruptions in politically volatile regions. Such instability can indirectly impact supply chains, forcing recalibration of sourcing strategies or inventory holding patterns;

breakdown form a central theme in the study of corporate crises. The persistence of incoherences, particularly when they go unaddressed, can lead to significant disruptions in a company's ability to function effectively. Understanding the nature and causes of these incoherences, both internal and external, is then crucial for developing strategies that enhance a company's resilience and ability to navigate periods of instability. This focus on the adaptability and flexibility of a company is also essential for preventing minor dysfunctions from escalating into full-blown crises. In this dissertation, this perspective on crisis as a breakdown of coherence underscores the importance of strategic management in maintaining alignment both within the company and with its external environment.

Figure 3: Incoherences Types



Source: own elaboration

Technological Obsolescence Metrics: Technology Readiness Levels (TRLs) and the Digital Maturity Index (DMI) measure the company's alignment with current technological standards. Indirect incoherences may emerge if external technology partners lag behind, leading to technical debt and compromised scalability;
Environmental and Climate Risk Assessment: Carbon Exposure Scores and Climate Risk Ratings assess vulnerability to climate-related externalities. Companies with suppliers in high-risk regions may face long-term incoherences in sustainability metrics, potentially raising Environmental, Social, and Governance (ESG) compliance costs and impacting CSR scores.

2.3.2. Determining Factors: Analysis and Classification

Usually, in the lifecycle of a company, periods of growth are often followed by phases of decline and crisis. These latter stages are typically preceded by warning signals that can be categorized into "*decline"* and "*imbalances.*" The signs of decline are frequently qualitative, while imbalances manifest as quantitative data, often measurable and indicative of underlying issues. The factors responsible for both decline and imbalances are considered as predisposing elements that can lead to corporate deterioration and crisis. This inclination toward decline can be traced back to several factors (Passeri et al., 2011):

- *Macroeconomic Factors:* These factors pertain to the broader economic, political, and social context in which firms operate. Macroeconomic phenomena that may contribute to corporate decline include:

    - *Lack of Creativity, Innovation, and Risk-Taking:* A deficiency in creative and innovative thinking, coupled with a reluctance to take risks, can impede a company's ability to adapt to changing market conditions;
    - *Inadequate Investment in Research and Development:* Insufficient investment in R&D can prevent a company from maintaining its competitive edge through innovation;
    - *Scarcity of Financial Resources:* Limited access to adequate funding can hinder growth and investment efforts;
    - *Weakness and Backwardness of the Banking System:* An underdeveloped or inefficient banking system can make it difficult for companies to obtain the necessary financing for their operations;
    - *Inadequate State Regulation:* The lack of effective state regulation can obstruct efficient corporate management and the creation of a stable business environment;
    - *Inefficiencies in Financial Markets:* Inefficiencies in financial markets can result in a lack of capital flow to businesses, making it challenging for them to secure funding through public offerings or other means;
    - *Political Instability and Excessive Bureaucracy:* Political instability, excessive and ineffective bureaucracy, and disproportionate tax burdens can create obstacles for businesses;
    - *Exchange Rate Fluctuations:* Uncertainty related to exchange rate fluctuations can negatively impact international trade and corporate profits;
    - *Excessive State Intervention in Industry:* Overreach by the state, especially in the manufacturing sector, can limit corporate efficiency and competitiveness;
    - *Labor Cost Disparities Between Countries:* Differences in labor costs between countries can influence corporate competitiveness and decisions regarding offshoring or maintaining local production;
    - *Fluctuations in Commodity Prices:* Major swings in commodity prices can significantly impact corporate costs;
    - *Competition from Developing Countries:* The entry of new developing countries into the market, often with lower labor costs and advanced technologies, can threaten the market shares of existing firms;
    - *Rapid Technological Evolution and Innovation in Production Processes:* Rapid technological advancements and innovations in production processes can render existing technologies and practices obsolete, pressuring companies to continuously adapt.

- *Sector-Specific Factors:* These are factors specific to the industry in which a company operates and can include:

    - *Different Stages of Sectoral Life Cycles:* Economic sectors often go through various stages of development, including emerging, expanding, and mature sectors. Decline tends to primarily affect mature sectors, though imbalances can impact all these groups. For instance, new sectors may initially experience low or negative profitability and significant variability in market shares. Imbalances may also arise during phases of rapid and sudden expansion;
    - *Sectoral Cyclicality*: Some sectors are subject to more or less intense cyclicality. Significant imbalances can occur during negative phases of the economic cycle, but these tend to be absorbed during positive phases. However, this cyclical condition should not necessarily lead to decline or crisis, as it is predictable. Therefore, companies can prepare during positive phases to face negative fluctuations, unless there are significant miscalculations or exceptional fluctuations in intensity and duration;
    - *Competitive Pressure:* In certain sectors, competitive conditions undergo periodic changes. These changes may include the entry of new competitors into the market, mergers and acquisitions that increase competitiveness through pricing strategies to gain market share, or the adoption of new technologies. Such changes can trigger sectoral imbalances and require an adequate response from companies to maintain their competitiveness.

- *Microeconomic Factors:* These are internal factors within the company that can trigger decline. It is important to note that this listing is purely indicative, as there is no universal list of such factors. However, the following elements are examined:

    - *Loss of Profitability and Reduced Investment:* A loss of profitability can lead to cuts in investments, particularly in intangible assets. This can alter the financial structure of the company and generate distrust towards the business;
    - *Negative Cash Flows:* Often, negative cash flows are correlated with losses incurred. However, in some situations, losses may not translate into negative cash flows. These cases are particularly dangerous because the management, lacking immediate financial pressures, might delay addressing the losses through questionable budget policies. Such behavior can further advance the company's decline, making any attempts at recovery too late;
    - *Reduction in Sales:* A decline in sales is a negative signal unless it is offset by an increase in prices and margins;
    - *Deterioration of Product Mix:* When there is a deterioration in the product mix, with a greater emphasis on low-margin products, this can erode overall profitability;
    - *Price-Based Sales Support:* If a company seeks to boost sales by reducing prices, this indicates a worsening of competitive conditions and may lead to margin reductions, risking negative margins;
    - *Worsening Relationships with Suppliers:* A decrease in supplier confidence in the company can lead to problems in sourcing and jeopardize the continuity of operations;

- *Loss of Key Managers and Skilled Personnel:* The loss of talent within the company, including managers and highly skilled personnel, accelerates the decline;
- *Loss of Market Share:* The loss of market share is a significant signal, especially when global demand is stable. In such times, it can lead to a reduction in sales and, therefore, revenue. During periods of strong demand growth, this signal may be less apparent as the overall sales increase may cover the loss of market share. However, if not addressed, the risk associated with a lower percentage growth than the global demand evolution may be underestimated, leading to increased competition and a collapse in sales when demand stabilizes;
- *Significant Increase in Debt and Reduction in Liquidity:* An increase in debt and a decrease in liquidity are clear signs of trouble. Once certain thresholds are surpassed, these factors become alarming, with the risk of reaching a point of no return. For example, a financial debt-to-revenue ratio exceeding one is a critical signal. Another indicator is the cover ratio, which measures the company's ability to cover financial charges with operating results. A higher value indicates greater financial security;
- *Deterioration of Financial Structure:* An inappropriate financial structure, such as financing long-term investments with short-term debt, can lead to financial imbalances;
- *Deterioration of Relations with the Financial Community: Financial* institutions and markets often anticipate corporate difficulties. This can lead to more restrictive credit conditions and make it difficult to access capital, jeopardizing the company's survival;
- *Increase in Bad Debts*: This is a common problem for companies trying to forcefully increase sales. It can result from choosing less reliable customers and negatively impact the company's financial situation.

These factors, synthetized in Table 2, collectively form the complex landscape of corporate decline, with each playing a significant role in driving a company towards financial distress and eventual crisis. Understanding these factors is crucial for developing effective strategies to prevent and mitigate the impact of such declines, thereby ensuring long-term corporate sustainability and resilience.

Table 2: Determining Factors of Crises

Macroeconomic Factors	Microeconomic Factors	Sector-Specific Factors
Lack of Creativity, Innovation, and Risk-Taking	Loss of Profitability and Reduced Investment	Different Stages of Sectoral Life Cycles
Inadequate Investment in Research and Development	Negative Cash Flows	Sectoral Cyclicality
Scarcity of Financial Resources	Reduction in Sales	Competitive Pressure
Weakness and Backwardness of the Banking System	Deterioration of Product Mix	
Inadequate State Regulation	Price-Based Sales Support	
Inefficiencies in Financial Markets	Worsening Relationships with Suppliers	
Political Instability and Excessive Bureaucracy	Loss of Key Managers and Skilled Personnel	
Exchange Rate Fluctuations	Loss of Market Share	
Excessive State Intervention in Industry	Significant Increase in Debt and Reduction in Liquidity	
Labor Cost Disparities Between Countries	Deterioration of Financial Structure	
Fluctuations in Commodity Prices	Lack of Strategic Capabilities	
Competition from Developing Countries	Tensions in Employee Relations	
Rapid Technological Evolution and Innovation in Production Processes	Deterioration of Relations with the Financial Community	
	Increase in Bad Debts	

Source: own elaboration based on Passeri et al. (2011)

2.3.3. Dysfunction and Corporate Crisis: Differences and Stages

Understanding the distinction between dysfunction and corporate crisis is crucial for a comprehensive analysis of a company's condition. Dysfunctions are interruptions or breakdowns in the relationships of coherence and coordination among the various components of the company, or between the company and its external environment. Given the ever-evolving dynamics of both the business environment and its subsystems, it is unrealistic to assume that a company can always maintain a state of perfect equilibrium. Consequently, the occurrence of dysfunctions is an inherent phenomenon in business management and can be considered part of the normal operational process. On the other hand, a corporate crisis is pathological in nature (Passeri et al., 2011; Altman, 1968). It arises from a development process of dysfunctions to which the company has failed to respond effectively and in a timely manner. The pathological aspect of corporate crises lies precisely in the company's inability to adequately react to the evolving process of deterioration in coherence relationships (Guatri, 1986). It follows that a single dysfunction, if identified promptly before it initiates a degenerative process, does not by itself threaten the overall equilibrium of the corporate system. Such dysfunctions can often be resolved through appropriate adjustments that restore the system to higher levels of equilibrium, preserving competitiveness, stakeholder satisfaction, and operational efficiency (Altman et al., 1994).

Conversely, a crisis, unlike dysfunctions, represents a genuine "*threat*" to the equilibrium of the entire corporate system (Aydin et al., 2022). It can trigger a downward spiral of negative synergies across three fundamental dimensions: competitiveness, stakeholder satisfaction, and operational efficiency (lo Conte & Sancetta, 2024). This regressive movement can even lead to the eventual closure of the business (Altman & Le Fleur, 1985). However, the difference between dysfunction and crisis, though significant, is not sufficient on its own to fully define the concept of a corporate crisis (Passeri et al., 2011). To qualify as a crisis, the company must find itself in a condition that is considered pathological, well beyond normal physiological equilibria (Kaczmarek et al., 2022). The analysis of the transition from a physiological equilibrium to a pathological one can be framed within the context of the already analyzed corporate coherence system, which can be categorized into four distinct states:

- *Physiological State:* The system of coherence is in balance. In this state, the company operates within its normal parameters, with all internal and external components functioning in harmony. Any minor disruptions are easily managed without significant impact on the overall system;
- *Tension State:* At this stage, some coherence relationships are disrupted, but the overall equilibrium is maintained through available compensatory mechanisms. These tensions signal potential vulnerabilities, yet the company is still able to operate effectively, utilizing its resources to counterbalance emerging issues;
- *Altered State:* The degenerative process begins here, compromising additional coherence relationships and producing noticeable effects. The company starts to experience tangible difficulties, such as declining profitability, reduced stakeholder confidence, and operational inefficiencies. At this point, the company's ability to self-correct diminishes, and the risks of entering a crisis increase;
- *Crisis State:* In this phase, the intensity and severity of the altered state, driven by the progression of the degenerative process, seriously threaten the company's objectives and purpose. The crisis state represents a critical juncture where the company faces existential risks, with its survival increasingly uncertain unless drastic and effective corrective measures are implemented.

The transition from dysfunction to crisis involves a gradual loss of control and the shift from a state of equilibrium to a pathological state, ultimately endangering the company's very survival. This progression underscores the importance of timely and effective intervention to prevent dysfunctions from escalating into crises. The delineation between these stages provides a framework for understanding how internal and external factors interact to influence a company's trajectory. Dysfunctions, when managed proactively, can be viewed as opportunities for improvement and strategic realignment. However, when neglected or inadequately addressed, they can evolve into crises with far-reaching consequences for the organization. Furthermore, this understanding of dysfunctions and crises is critical for both theoretical analysis and practical application in corporate management since it emphasizes the need for continuous monitoring, early detection, and strategic intervention to maintain corporate health and ensure long-term sustainability. The work of scholars such as Roubini (2010) and Gennaioli & Shleifer (2013) highlights, in fact, the complexities of managing these transitions, particularly in the context of an increasingly volatile global economic environment where AI and advanced predictive tools play a vital role in crisis management and decision-making.

2.3.4. Phases

A corporate crisis, as previously discussed, is a pathological condition that gradually develops through a process of organizational deterioration (Passeri et al., 2011). This progression is neither sudden nor arbitrary; rather, it follows a discernible pattern that can be broken down into distinct phases. Understanding these phases is critical for both diagnosing the onset of a crisis and implementing timely interventions to mitigate its impact. Below, the phases of a corporate crisis are outlined, providing a comprehensive framework that highlights the gradual decline and eventual crisis experienced by a company. Figure 4 visually represents this trajectory.

Decline:

- *Incubation Phase*: The incubation phase represents the initial stage of a corporate crisis, where early warning signs begin to surface. These signs are often subtle and manifest as inefficiencies and imbalances within the organization. During this phase, problems may include declining operational efficiency, increased internal conflicts, and a gradual erosion of competitive advantage. The inefficiencies may stem from misaligned strategies, poor management practices, or a failure to adapt to changes in the external environment. However, these issues are often overlooked or underestimated by management, who may regard them as temporary or insignificant;
- *Maturation Phase:* In the maturation phase, the issues identified during incubation intensify, leading to more pronounced symptoms of decline. The company begins to experience significant losses, both in terms of profitability and the overall value of its capital. The inefficiencies and imbalances that were once manageable now start to spiral out of control, causing a deterioration in financial performance. Operational inefficiencies become entrenched, market share erodes, and customer satisfaction declines. The company may also experience difficulties in maintaining relationships with key stakeholders, including suppliers, customers, and employees. At this stage, the company's financial statements typically reflect worsening profitability ratios, declining cash flow, and increasing debt levels, signaling the deepening of the crisis.

Crisis:

- *Phase of Severe Financial Repercussions and Deterioration in Corporate Credibility:* The crisis fully manifests in this phase, as financial difficulties become apparent and have a profound impact on the company's ability to manage cash flows. Liquidity shortages become acute, and the company struggles to meet its financial obligations. This phase is often characterized by an erosion of trust from creditors, investors, and other financial stakeholders. The company's creditworthiness deteriorates, leading to increased borrowing costs or an inability to secure additional financing. As the crisis deepens, the company may face significant losses in capital value, potentially resulting in a substantial or even total erosion of equity. The loss of financial stability can trigger a negative feedback loop, where declining investor confidence leads to further financial distress, exacerbating the crisis;

- *Phase of Stakeholder Impact:* The final and most severe phase of a corporate crisis is characterized by its widespread impact on stakeholders. At this point, the company may become insolvent, unable to honor its debt obligations, and potentially facing bankruptcy. The insolvency of the company has far-reaching consequences for a wide range of stakeholders, including employees, suppliers, customers, and investors. Employees may face job losses or reduced benefits, while suppliers may suffer from unpaid invoices. Customers may lose confidence in the company, leading to a decline in sales and market share. Investors, particularly shareholders, may experience significant financial losses as the value of their investments plummets. The company's inability to fulfill its obligations may lead to legal proceedings, such as bankruptcy filings or restructuring efforts, further complicating the resolution of the crisis. The collapse of trust in the company can have long-lasting effects, damaging its reputation and making recovery difficult, if not impossible.

Figure 4: Phases of Corporate Crisis

Source: Own elaboration of Passeri et al. (2011)

Figure 4 delineates the trajectory of a corporate decline and crisis, segmented into four critical phases: Incubation, Maturation, Severe Financial Repercussions, and Stakeholder Impact. The x-axis charts the temporal progression of the crisis, while the y-axis quantifies the intensifying severity of the corporate distress. The curve originates at the baseline, representing the early, often subtle, indicators of decline, and ascends sharply, reflecting the compounding and exacerbating effects of the crisis as it advances through each phase. All these descripted phases underscore the importance of early detection and proactive management in preventing a corporate crisis from reaching its most severe stages. By recognizing the incubation phase, management can take corrective actions before the situation escalates into a full-blown crisis. Furthermore, the maturation phase serves as a critical juncture where strategic interventions can either mitigate the decline or accelerate the company's descent into crisis. In the latter stages, particularly during the phases of severe financial repercussions and stakeholder impact, the focus shifts from prevention to crisis management and damage control. At this point, the use of advanced predictive analytics, including AI and ML models, becomes crucial. These tools, in fact, can provide real-time insights into financial performance, cash flow management, and stakeholder sentiment, enabling management to make informed decisions under pressure. However, the reliance on such tools must be balanced with sound managerial judgment and a deep understanding of the company's operational dynamics.

Moreover, the progression through these phases highlights the interconnectedness of different aspects of the company—financial health, operational efficiency, stakeholder relations, and strategic decision-making. A failure in one area can trigger a cascade of problems that ultimately lead to a crisis. This reinforces the need for an integrated approach to corporate management, where different functions and departments work cohesively to maintain organizational coherence and resilience. Finally, the understanding of these phases also has significant implications for corporate governance and regulatory oversight; boards of directors and regulatory bodies must be vigilant in monitoring the early signs of decline and ensuring that management takes appropriate actions. This may involve implementing stronger corporate governance frameworks, enhancing transparency in financial reporting, and promoting a culture of accountability and continuous improvement.

2.3.5. Corporate Decline and its Quantitative Assessment

The concept of corporate decline is frequently employed to describe the deteriorative process that can lead to a state of corporate crisis. In this context, decline is analyzed using a quantitative approach grounded in the theory of value creation (Nguyen et al., 2019). This theory posits that the fundamental objective of a company is the enhancement of its capital value—whether economic or market-based. This objective is rational, as it is intrinsically linked to medium- and long-term profitability, risk management, and is objectively measurable. The increase in the value of economic capital over a given period can be calculated using the pure income method, and, if this measure of corporate performance indicates insufficient, null, or negative growth, it signals that the firm is failing to achieve its objectives, potentially jeopardizing its medium- to long-term survival.

Decline, therefore, can be associated with negative performance in terms of changes in economic capital value over time. In other words, a company is in decline when it is consistently destroying value. This definition of decline is not solely dependent on expected income but also on the behavior of the capitalization rate, which is influenced by the level of risk the company has undertaken. In this regard, several key points merit consideration:

- Decline is not necessarily correlated with economic losses; it can also manifest when income streams continue to diminish, even if they remain positive;
- For a situation to be classified as decline, the loss of income streams must be systematic and irreversible, with no possibility for restorative or restructuring interventions. Losses due to extraordinary or cyclical factors do not fall under this definition;
- Decline can result not only from past losses but also from anticipated losses. The loss of the ability to generate future income above certain thresholds is also a contributing factor to decline.

In addition to income streams, corporate risks can lead to value destruction and thus contribute to decline. It is important to note that distinguishing between decline and corporate crisis is not always straightforward, particularly in the early stages. A crisis can be viewed as a further progression of decline, often characterized by an irreversible downward trajectory with no possibility for external intervention. At this point, the crisis becomes apparent even to external stakeholders. This understanding of corporate decline provides a critical foundation for the application of predictive analytics and AI in the context of crisis and insolvency forecasting. By quantifying and monitoring key indicators of decline, AI-driven models can identify early warning signs of financial distress, and, in this framework, the role of sustainability becomes paramount, as firms that can proactively address the factors contributing to decline are better positioned to maintain long-term viability and resilience.

Moreover, this approach aligns with the broader goals of corporate sustainability. In an environment where economic, social, and environmental factors are increasingly interlinked, the ability to anticipate and manage decline is essential for the long-term health of the organization. In conclusion, the quantitative assessment of corporate decline serves as a vital tool in the broader context of crisis and insolvency prediction.

2.3.6. Evolution of Italian Scholarly Perspectives on Business Crisis & Insolvency

In Italy, scholarly research on business crisis and insolvency began to take shape significantly during the economic boom that followed World War II (Passeri et al., 2011). This period marked a pivotal moment in the evolution of business management practices, particularly in response to the rapid industrialization and economic expansion that characterized the post-war era. The burgeoning complexity of corporate operations and the increasing interdependence of global markets necessitated a deeper understanding of the factors leading to business failure and the development of robust predictive models to anticipate such crises. From a managerial perspective, scholars like Argenti (1976),and Coda (1977, 1990) were instrumental in identifying the causes of business crises, focusing on the internal and external factors that could destabilize firms. Argenti's work, for example, introduced a framework for diagnosing corporate failure that emphasized managerial incompetence and organizational weaknesses as primary contributors to business crises. This approach underscored the importance of effective decision-making processes within the management structure, highlighting how managerial deficiencies could lead to operational inefficiencies, loss of competitive advantage, and ultimately, insolvency.

Coda (1990) expanded on this by exploring how strategic misalignment with market conditions could precipitate a crisis. His work emphasized the role of strategic planning and the need for continuous adaptation to the changing external environment. This managerial perspective on crisis emphasized that a firm’s resilience depended not only on its internal capabilities but also on its ability

to anticipate and respond to external shocks; the focus was on enhancing decision-making frameworks to incorporate a more dynamic and forward-looking approach, capable of mitigating potential risks before they escalated into full-blown crises.

On the other hand, the "*aziendalista*" (business-centric) perspective, deeply rooted in the Italian tradition of economic and business studies, contributed significantly to the discourse on crisis management. Scholars like Guatri (1986) and Sciarelli (1995) provided a systemic view of business crises, considering them as inherent to the business lifecycle. Their works argued that crises were not merely disruptions but rather potential turning points that could lead to either the rejuvenation or the decline of the business, depending on how they were managed. This perspective stressed the importance of viewing crises through a holistic lens, where the focus was on understanding the intricate interdependencies within the business structure and between the firm and its environment. Moreover, this approach also placed significant emphasis on the role of financial indicators in predicting crises. In this regard, Alberici's (1975) contribution to the development of predictive models for insolvency is particularly noteworthy. His model, which introduced different versions of a Score to be applied to specific years leading up to a crisis, provided a quantitative framework for assessing a company's financial health. This model laid the groundwork for subsequent studies that sought to refine predictive analytics in business management.

Recent advancements in this field have further integrated the "*aziendalista*" and managerial perspectives, particularly with the incorporation of AI and ML techniques into predictive models. Studies by Altman et al. (2013), and more recently by Brogi et al. (2022), have demonstrated the potential of AI-driven analytics in enhancing the accuracy and timeliness of crisis prediction. These technologies, in fact, enable managers to process vast amounts of data and make informed decisions more swiftly than traditional methods would allow, and this fusion of traditional management theories with cutting-edge technology represents a significant evolution in the field of crisis management, underscoring the need for a more integrated approach that combines strategic foresight with robust analytical tools.

2.4. Insolvency & Financial Distress

Insolvency and financial distress are critical stages in the financial deterioration of a company, often representing the tipping point between recovery and failure (Tron & Franceschi, 2021; Altman, 1968). These conditions are central to understanding corporate crises and are pivotal in the strategic decision-making processes that aim to mitigate their effects. Insolvency is defined as the state in which a debtor, typically a company, is unable to fulfill its financial obligations as they become due (CCII, 2022). It represents a critical juncture in the life of a company, where the entire organizational structure is deeply affected, making recovery exceedingly difficult with a low probability of success. So, addressing insolvency effectively requires radical interventions, particularly in the restructuring of the financial framework and the renewal of corporate management. Insolvency can be classified into two main types, temporary and permanent (Nigro & Vattermoli, 2023; Tron & Franceschi, 2021). The first type, is considered temporary if two crucial conditions are met:

- The company must still maintain a residual equity balance, meaning its net capital remains positive;
- There must be a reasonable expectation of improvement in the company's economic prospects, particularly following restructuring efforts.

The first condition suggests that the company can withstand the risks associated with recovery attempts without compromising the creditors' interests or the remaining capital. In such cases, creditors might be asked to extend the debt repayment terms without forgoing the amounts owed (Nigro & Vattermoli, 2023). The second condition has two potential outcomes: if economic equilibrium is not restored, it is unlikely that financial stability will be regained. However, a reasonable expectation of an improvement in economic balance is essential to argue that the company's net capital will be positive in the future. This condition implies that with proper management and strategic interventions, the company could potentially recover from its temporary financial distress, making this phase a critical period for decisive action. When insolvency is irreversible, it becomes definitive, or permanent. In this scenario, the company is in a state of financial distress so severe that the creditors are aware that their claims are likely already compromised unless preferential claims are applied. Any attempts at recovery or corporate restructuring are carried out at the expense and risk of the creditors. Definitive insolvency marks a terminal stage in the corporate crisis, where the prospects for recovery are minimal, and the focus often shifts to managing the dissolution or liquidation of the company's assets. Finally, insolvency, whether temporary or definitive, represents a critical phase in the lifecycle of corporate crises, as its nature is determined by various factors, including the residual equity balance and future economic prospects (Tron & Franceschi, 2023). Addressing insolvency requires drastic measures, which may involve partial debt forgiveness by creditors in an effort to salvage the company's survival (Nigro & Vattermoli, 2023).

2.4.1. Relationships and Differences: A Theoretical Evolution

Financial distress mainly refers to the broader condition of declining financial health that precedes insolvency. It encompasses a range of symptoms, such as diminishing profitability, shrinking cash flows, rising debt levels, and challenges in securing additional financing and is often an early warning sign of potential insolvency, signaling that the company is on a downward trajectory that could culminate in its inability to meet financial obligations. In financial terms, distress may not immediately lead to insolvency, but it sets the stage for the potential escalation of financial challenges. This stage is crucial for managerial intervention, where strategies to stabilize and recover the company must be implemented swiftly to prevent further decline. From a managerial perspective, understanding and identifying the signs of financial distress early is essential for mitigating risks and avoiding insolvency. Managers, in fact, should employ strategic tools, such as AI-driven predictive models, to monitor key financial indicators and anticipate potential issues before they become critical. Moreover, both financial distress and insolvency are not static conditions; they evolve over time, influenced by both internal management decisions and external economic factors. This dynamic nature requires a continuous reassessment of the company's financial health and the strategic agility to adapt to changing circumstances.

In literature, one of the earliest and most influential contributions came from Riparbelli (1950), who provided a foundational definition of corporate financial distress[58]. Riparbelli characterized financial distress as a pathological condition that emerges when unfavorable circumstances disrupt the normal functioning of a company. He also emphasized that financial distress occurs when a company

[58] Riparbelli (1950) also emphasized that the distinction between these two states is vital for understanding the severity of a company's financial distress. Temporary insolvency, if managed correctly, can be remedied through effective financial restructuring and management renewal. However, permanent insolvency represents a critical juncture where recovery is virtually impossible, and the company faces inevitable dissolution.

fails to maintain the minimum quantitative conditions necessary for equilibrium, thereby threatening its very existence. Riparbelli's work also introduced the idea that corporate financial distress is a dynamic process, not a static or sudden occurrence. He identified three primary types of disorders—economic, financial, and capital structure disorders—that serve as indicators or precursors to financial distress. Economic disorders arise when a company experiences a significant loss in profitability, while financial disorders are linked to a company's inability to meet its current liabilities. Capital structure disorders occur when there is an imbalance between equity and debt, leading to potential instability. These disorders, if left unmanaged, could collectively lead to a state of financial distress. However, he also noted that a short-term economic disorder, such as a temporary loss of profitability, should not automatically be equated with financial distress unless it persists over a long period, ultimately eroding the company's ability to cover its costs.

2.4.2. The Concept of Equilibrium in Financial Distress and Stages of a Progressive Pathology

Building on Riparbelli's insights, later scholars like Giannessi (1962, 1979) and De Sarno (1982) further refined the concept of equilibrium in the context of corporate financial health[59]. Giannessi, for instance, described equilibrium as a dynamic condition throughout a firm's life cycle, arguing that it is reflected in the firm's profitability and its ability to sustain economic vitality over time. This perspective aligns with the broader view that financial distress occurs when a company is unable to generate sufficient revenues to cover its costs, leading to a decline in profitability and, ultimately, its overall value. De Sarno (1982) added another layer to this understanding by distinguishing between endogenous and exogenous causes of financial distress. Endogenous distress originates from within the company, such as through poor management decisions or product obsolescence, while exogenous distress results from external factors like market shifts or unforeseen events. De Sarno's classification highlights the complexity of financial distress, noting that it can be triggered by multiple, interrelated factors.

Guatri (1986) expanded on these theories by framing financial distress as a multi-stage process[60]. He identified the initial stage as a period of disequilibrium, characterized by internal and external inefficiencies. If these inefficiencies persist, they lead to increased economic losses, marking the second stage. The third stage is insolvency, where the firm can no longer meet its obligations, and the final stage is one of permanent distress, where the company systematically destroys its economic capital. This progressive understanding of financial distress underscores its pathological nature, as described by

[59] According to these scholars, corporate financial distress is rarely the result of a single cause; instead, it often emerges from a confluence of several interrelated factors, both internal and external to the firm. External factors, which are largely beyond the control of a firm's management, play a significant role in influencing the life cycle of a company. As highlighted by De Sarno (1982) and further elaborated by Guatri (1986) and Vergara (1988), these external factors include the broader economic environment such as the growth rate of the economy, inflation, exchange rates, interest rates, and shifts in consumer behavior and market characteristics. These conditions directly impact a firm's profitability and market power, often creating a volatile backdrop against which businesses must operate. While external factors are significant, internal factors within the business are equally crucial in determining a firm's financial health. Riparbelli (1950) identified several internal factors that can destabilize a business. These include insufficient equity to finance growth, excessive reliance on leverage, an inability to meet customer expectations, and overinvestment in fixed assets.

[60] Guatri (1986) provided a detailed classification of financial distress, categorizing it into five distinct types, each with unique causes and implications.

Guatri as a company in financial distress is one that is not just experiencing temporary setbacks but is on a downward trajectory that, if unchecked, leads to the destruction of its value.

2.5. Bankruptcy

Bankruptcy, within the context of corporate financial distress, represents the terminal stage where a company is legally declared unable to meet its debt obligations. This condition, as outlined in the Italian *Codice della crisi d'impresa e dell'insolvenza*, is a formal legal process that signifies the end of the company's financial viability, either leading to liquidation or restructuring under judicial supervision (CCII, 2022). In academic literature, bankruptcy is often classified as either reversible or irreversible, depending on the potential for recovery and restructuring. The irreversible form of bankruptcy is often the focus of predictive models, where the objective is to identify the financial and operational signals that indicate a point of no return. Reversible bankruptcy, on the other hand, emphasizes the potential for turnaround[61], often with the support of advanced tools, which can assist in managing crises effectively before reaching the stage of irreversibility.

The historical evolution of bankruptcy prediction models, particularly in the United States, provides a foundational understanding of how financial distress is identified and managed. Early studies, such as those by Lough (1917) and Smith (1930), focused on static models that employed single financial ratios to predict distress. These studies laid the groundwork for understanding the financial indicators most closely associated with bankruptcy, such as working capital to total assets, net worth to debt, and current ratios. FitzPatrick (1932) and Merwin (1942) expanded on this by comparing distressed and healthy firms, identifying key financial ratios that were predictive of long-term distress. As the field evolved, scholars like Jackendoff (1962) and Horrigan (1965) emphasized the importance of liquidity ratios and the application of statistical models in predicting financial distress. Horrigan, in particular, highlighted the significance of parsimony in ratio selection due to high correlations among various financial indicators, a caution that remains relevant in modern AI-driven models.

In the context of this Thesis, bankruptcy is analyzed both as a condition to be predicted and as a process that can potentially be managed or even averted through strategic intervention. The research focuses on how AI and predictive analytics can be leveraged to provide early warnings of financial distress, enabling managers to make informed decisions that could prevent the progression from a state of financial difficulty to full-scale bankruptcy, and, because of this, AI's role in this context is twofold: firstly, it serves as a tool for predicting the onset of financial distress by analyzing vast datasets and identifying patterns that traditional models might miss; secondly, it acts as a decision support system in crisis management, offering recommendations for strategic actions that can mitigate the impact of distress or reverse the course of a failing business. This aligns with the goals of both European and Italian legislative frameworks.

In the methodological and empirical sections of this dissertation, specific legal situations and events will be considered, which, when categorized within the utilized databases, may be more or less directly associated with bankruptcy or, similarly, with specific crisis situations. Furthermore, it is important to note that bankruptcy is closely linked to the legal frameworks of individual countries (in Italy, for example, this term does not even exist in the new Legislation). Therefore, this Thesis will

[61] The term turnaround refers to a strategic process through which a struggling company seeks to reverse a crisis situation in order to restore profitability and operational stability. This process may involve various interventions, including financial restructuring, management overhaul, cost optimization, and the implementation of new market strategies.

primarily focus on the practical and technical characteristics of the phenomenon, taking into account the relevant normative dictates specific to each jurisdiction, in particular those of the Italian context.

2.6. Key Concepts in Theory: *Business Life Cycle Theory*

The financial health and sustainability of a company are inherently tied to the stages of its lifecycle, with crisis, default, and bankruptcy representing critical junctures that may arise as a firm progresses through its development (Wruck, 1990). For the Business Life Cycle Theory, the corporate lifecycle is typically segmented into four stages: birth, growth, maturity, and decline, each of which presents distinct challenges and opportunities for the organization (Anthony & Ramesh, 1992). Figure 5 offers a clear view of these.

Figure 5: Business Life Cycle

Source: Slater, 2023 based on Anthony & Ramesh, 1992

During the initial phase, often referred to as the birth stage, companies are generally small, informal, and relatively simple in structure. Decision-making is typically centralized, with power concentrated among a few individuals, often the founders. At this stage, there is a strong emphasis on innovation and product development as the company seeks to establish itself in the market (Adizes, 2004; Pashley and Philippatos, 1990). This phase is then characterized by high levels of uncertainty regarding future growth, leading to elevated firm-specific risks and a higher book-to-market ratio (Pastor and Veronesi, 2003).

As the company transitions into the growth phase, it begins to scale its operations, expand its market presence, and attract more significant resources. However, this phase can also bring about financial distress, often linked to liquidity constraints or concurrent cash flow challenges (Hasan et al., 2015). The rapid expansion during this period may necessitate increased capital, leading to a higher reliance on financial leverage, which can exacerbate financial instability if not managed properly (Agostini, 2018). The maturity phase is typically marked by a shift in focus from innovation and risky strategies to the stabilization and consolidation of the company's market position. During this stage, the firm's emphasis is on maintaining competitive advantage and ensuring steady, sustainable growth. However, the reduced focus on innovation can sometimes lead to complacency, making the company vulnerable to external market shifts and competitive pressures. Finally, the decline phase represents the period in which the company experiences a significant downturn in performance, often due to a

combination of internal inefficiencies and external market factors. If the underlying issues are not addressed, this decline can lead to a full-blown crisis, potentially culminating in definitive insolvency or bankruptcy. This stage, in particular, underscores the importance of timely and effective intervention to prevent irreversible damage to the company's financial health.

This Theory posits that the nature of the challenges faced by a company, as well as the strategies employed to address these challenges, vary significantly depending on the stage of the lifecycle (Koh et al., 2015). For instance, during the early stages, corrective actions might include cutting down on investments or reducing staff, while in the maturity stage, the focus might shift towards asset reorganization or financial restructuring. In this context, Sudarsanam & Lai (2001) categorize corporate restructuring into four main types: management, operational, asset, and financial restructuring. Managerial restructuring typically involves changes in the top management or CEO but does not directly involve cash flow changes. This type of restructuring is more likely to occur in later stages of the company's life cycle when the structure is more complex, and the separation between ownership and management is more pronounced. In contrast, during the early stages, such changes are less common due to the concentration of power among the founding members, who are often the primary decision-makers. Operational restructuring focuses on optimizing the company's operations by reducing costs and improving efficiency. This might involve selling off fixed assets, reducing expenses, or cutting back on capital expenditures and research and development (Lasfer & Remer, 2010). Asset restructuring, which may involve divestitures, spin-offs, or equity carve-outs, is typically seen as a more extreme measure. It is often undertaken when there is an urgent need to raise cash or when the company seeks to refocus on its core competencies by shedding non-core assets (Robbins and Pearce, 1993; Shleifer & Vishny, 1992).

Financial restructuring, on the other hand, involves changes to the company's capital structure or dividend policies (lo Conte et al., 2022). This may include issuing new securities, converting debt to equity, or altering dividend payments (Koh et al., 2015). Such strategies are crucial in maintaining the company's financial health, particularly during the decline phase, where the risk of insolvency is heightened. In summary, the Business Life Cycle Theory provides a framework for understanding the evolution of corporate strategies and decision-making processes as a company progresses through its lifecycle. Each stage presents unique challenges, and the appropriate response must be tailored to the specific circumstances of the company at that time.

2.7. The Reform of the Italian Insolvency System: A New Paradigm in Crisis Management

Recent reforms in the Italian insolvency system have introduced significant changes aimed at modernizing and enhancing the framework for managing corporate insolvency and financial distress. These reforms, encapsulated within the new *Codice della Crisi d'Impresa e dell'Insolvenza* (Crisis and Insolvency Code)[62], represent a comprehensive overhaul designed to align Italian practices with broader European directives, particularly the EU Insolvency Directive (Directive (EU) 2019/1023). This directive emphasizes early intervention and the preservation of viable businesses, which is a central theme in both European and Italian legislative efforts. One of the most transformative aspects of the reform was, certainly, the possible introduction of *strumenti di allerta precoce* (early warning tools), which were

[62] See Legislative Decree No. 14 of January 12, 2019.

intended to detect signs of financial distress at the earliest possible stage. However, these tools, crucial for enabling companies to take corrective measures before financial difficulties escalate into irreversible insolvency, are, today, eliminated from the reform[63]. The goal was to shift from a reactive approach, where intervention typically occurs only when a company is on the brink of bankruptcy, to a proactive stance that prioritizes the early identification and resolution of financial problems. However, under the new code, managers and directors are still legally obligated to monitor the financial health of their companies more rigorously and continuously as the legislation effectively mandates that corporate leaders adopt a more forward-looking approach to managing their companies, with a specific focus on sustainability and long-term viability.

Failure to heed these early warning signs can have severe legal consequences for company directors, who may be held personally liable for failing to prevent the company's descent into insolvency. This is a marked departure from previous norms, where the responsibility of directors was more narrowly defined and less proactive. The reform, therefore, reinforces the need for strategic decision-making and timely intervention, which are central themes in the broader discourse on corporate crisis management, and, obviously, in this dissertation. Furthermore, the new Code consolidates and streamlines the various insolvency procedures into a single, coherent legislative framework, replacing the previous, more fragmented system: by unifying the insolvency processes, the reform aims to make the system more accessible, reducing the complexity and inefficiency that have historically characterized Italian insolvency law. One of the most significant changes under is the replacement of the term "*bankruptcy*" (*fallimento*) with "*judicial liquidation*" (*liquidazione giudiziale*). This change is not merely semantic as it reflects a broader cultural shift away from the stigma traditionally associated with business failure. The term *liquidazione giudiziale* is, in fact, intended to convey a more neutral and procedural approach to insolvency, focusing on the orderly liquidation of assets rather than the punitive connotations historically associated with bankruptcy. This aligns with the reform's broader goal of encouraging a more pragmatic and less punitive approach to corporate distress, focusing on resolution and recovery rather than merely on winding up.

The Code also introduces the concept of *composizione negoziata della crisi d'impresa* (negotiated settlement of corporate crises), which serves as a mechanism for the private resolution of financial distress before formal insolvency procedures are initiated (See Table 3 for the tool scheme "*Test Pratico"*). This tool allows companies to negotiate directly with creditors and other stakeholders to reach an agreement that could avert formal insolvency. However, while it represents a significant step forward, it has been met with some skepticism within the literature. Critics argue that while the *composizione negoziata* is a valuable addition, it may require further refinement to be fully effective, particularly in the context of the Italian business environment, where there is still limited cultural and practical familiarity with preventive crisis management practices.

[63] Initially introduced as part of the "*Codice della crisi d'impresa e dell'insolvenza*" (Legislative Decree No. 14 of January 12, 2019), these tools were designed to detect early signs of financial distress, thereby allowing for timely interventions. However, they were removed due to concerns about their effectiveness and the potential for overburdening businesses with compliance obligations. The elimination was influenced by a recognition that the tools were not providing the expected benefits and were instead complicating the already challenging process of managing business crises. This decision reflects a broader reassessment of the balance between regulatory oversight and practical business needs.

Table 3: Composizione Negoziata della Crisi d'Impresa – Test Pratico

TEST PRATICO PER LA VERIFICA DELLA RAGIONEVOLE PERSEGUIBILITA' DEL RISANAMENTO			
L'entità del debito che deve essere ristrutturato			
debito scaduto		+	
(di cui relativo ad iscrizioni a ruolo)			Valore non utilizzato nel computo del TOTALE A
debito riscadenziato o oggetto di moratorie		+	
linee di credito bancarie utilizzate delle quali non ci si attende il rinnovo		+	
rate di mutui e finanziamenti in scadenza nei successivi 2 anni *(per le cooperative si tiene conto della probabile richiesta di romborso del prestito sociale secondo le evidenze storiche non precedenti a tre anni)*		+	
investimenti relativi alle iniziative industriali che si intendono adottare		+	
ammontare delle risorse ritraibili dalla dismissione di cespiti (immobili, partecipazioni, impianti e macchinario) o rami di azienda compatibili con il fabbisogno industriale		-	
nuovi conferimenti e finanziamenti, anche postergati, previsti		-	
stima dell'eventuale margine operativo netto negativo nel primo anno, comprensivo dei componenti non ricorrenti		-	
TOTALE A			
I flussi annui al servizio del debito			
stima del Margine Operativo Lordo prospettico normalizzato annuo, prima delle componenti non ricorrenti, a regime			
investimenti di mantenimento annui a regime		-	
imposte sul reddito annue che dovranno essere assolte		-	
TOTALE B			
Grado di difficoltà del risanamento		Grado di difficoltà non calcolabile	

Source: https://composizionenegoziata.camcom.it/

2.7.1. Cultural Shift and Path Forward

The Italian reform certainly reflects a significant cultural and procedural shift in how corporate crises are managed and, in some ways, predicted. By focusing on early detection, preventive measures, and the de-stigmatization of insolvency, the reform, in fact, aims to create an environment where companies are more likely to survive financial distress and where stakeholders are less likely to suffer catastrophic losses. This approach is consistent with the European Union's broader objectives of fostering business continuity and minimizing the negative social and economic impacts of corporate insolvency. However, its success depends not only on the legal and procedural changes it introduces but also on the willingness of Italian companies and their leaders to embrace these new tools and responsibilities. In fact, the Italian business culture has traditionally been less proactive in managing potential crises, often waiting until problems become severe before taking action. The new Code seeks to change this, by embedding a more preventive and forward-thinking approach into the legal and managerial frameworks of Italian companies.

So, to fully realize the benefits of the reform, it will be necessary to continue educating corporate leaders about the importance of early intervention and to encourage the adoption of advanced tools, such as AI-driven ones, which can provide the insights needed to dive better in different landscapes[64].

[64] AI can facilitate the process of negotiated solutions to crises, such as Italy's "*composizione negoziata della crisi d'impresa*" (negotiated resolution of business crises). Through it, companies can simulate different restructuring scenarios, assess their impacts on various stakeholders, and choose the most viable path forward. This technological support can improve the effectiveness of negotiations and increase the likelihood of reaching a consensual agreement that prevents liquidation.

Additionally, ongoing evaluation and refinement of the *composizione negoziata della crisi* and other elements of the reform will be crucial to ensure that these tools are effectively implemented and that they truly serve the needs of the Italian business community. However, as said before, the real impact of these changes will depend on how well they are understood, embraced, and implemented by the business community, making ongoing education and cultural change key components of the reform's success. This research on using AI and intelligent methods for predicting corporate insolvency and financial distress then aligns closely with both the existing literature and the evolving regulatory frameworks examined. In recent years, academic literature has increasingly recognized the potential of new models to enhance the accuracy and timeliness of financial distress predictions and, in terms of compliance, AI models can be designed to align with specific legal requirements, such as those related to the timeliness of interventions and the transparency of decision-making processes; for instance, AI algorithms can be calibrated to ensure that they provide clear, interpretable outputs that managers and regulators can use to justify their decisions.

2.7.2. Between Gaps and Potentials

Scientific literature supports the use of AI for enhancing the negotiation processes involved in crisis management as intelligent systems can simulate various restructuring scenarios, and assess their potential outcomes. However, despite these advancements, a significant gap remains in both European and Italian literature regarding the literacy of managers in AI technologies and the comparative effectiveness of different AI algorithms, particularly in the context of SMEs. While much research has focused on the technical capabilities of AI models, there is a notable deficiency in scientific studies that address the readiness of managers to adopt these technologies effectively. This gap is crucial because the impact of AI on crisis management, prediction, sustainability, and decision-making hinges not only on the technological prowess of AI systems but also on the ability of managers to understand and apply these insights strategically.

Indeed, AI has the potential to demystify and mitigate the stigma associated with insolvency or bankruptcy, aligning with the goals of legislators and certain academic perspectives that view financial distress not merely as a negative event but as a discontinuity that can lead to renewed business continuity. When managed correctly, these periods of difficulty can be leveraged to create both tangible and intangible value, enhancing the resilience and long-term sustainability of businesses. This approach reflects a shift in understanding corporate crises—not as terminal failures but as integral, manageable aspects of business life cycles that, through strategic intervention and advanced technologies like AI, can contribute to the overall health and innovation of the economic system.

3. Seeing the Unseen: "*Predictive*" Key Concepts Explained

Understanding and differentiating between the concepts of prediction, forecasting, and detection is fundamental. These concepts are pivotal for managing financial health and for anticipating crises, insolvency, and bankruptcy. While often used interchangeably, in fact, these terms reflect distinct methodologies and serve different purposes within the financial management framework. This Thesis, in particular, places particular emphasis on prediction, due to its comprehensive scope and strategic relevance, especially when integrated with advanced analytical tools.

- Prediction is defined as the process of estimating the likelihood of future events based on historical data, statistical models, and algorithms (Cox, 2006). In corporate finance, prediction models are used to assess the probability of a firm experiencing financial distress within a given timeframe (Altman, 1968). However, these models are not confined to historical financial ratios but also incorporate a broad range of variables, including market trends, macroeconomic indicators, and qualitative factors such as management quality and industry dynamics. The probabilistic outcomes generated are invaluable for stakeholders, particularly in enabling them to anticipate and preempt potential crises (Ohlson, 1980; Altman, 1968). Furthermore, by its nature, prediction in a strategic environment is proactive, as it equips managers with the foresight necessary to implement preventive measures well before financial difficulties escalate. This proactive stance is critical for effective crisis management and aligns closely with strategic decision-making processes, making prediction an essential tool for maintaining corporate stability over time;

- Forecasting, on the other hand, is more focused on projecting specific financial metrics into the future, such as revenues, expenses, and cash flows. Forecasting typically employs time-series analysis and trend extrapolation to generate these projections (Makridakis et al., 1998). Unlike prediction, forecasting generally assumes that past patterns will continue into the future, and as such, it is often used for routine financial planning rather than for anticipating disruptive events (Ohlson, 1980). Furthermore, while forecasting provides valuable insights into expected financial performance under normal conditions, it may not adequately capture the potential for abrupt changes or unforeseen crises, which are often the catalysts for financial distress. Therefore, while it is crucial for operational planning, it may not provide the early warnings needed for strategic interventions in the face of emerging risks;

- Detection involves identifying existing or emerging signs of financial trouble within an organization. Detection models, in fact, are designed to recognize anomalies and red flags in real-time or near-real-time, facilitating immediate responses to developing issues (Beaver, 1966). However, detection is fundamentally reactive, addressing problems as they arise rather than preemptively. Although detection is essential for prompt intervention, its reactive nature means it may not allow sufficient time for the strategic planning needed to prevent the escalation of financial distress into a full-blown crisis.

3.1. Some Considerations

The emphasis on prediction in this dissertation is then intentional and stems from several key considerations. First, prediction enables organizations to adopt a forward-looking stance, identifying potential financial distress well before it materializes. This proactive approach is essential for effective crisis management, as it allows ample time for developing and implementing strategic interventions that can prevent or mitigate adverse outcomes (Blazek, 2014). By predicting the probability of distress, firms can in fact allocate resources more efficiently or adjust their strategic plans to mitigate risks. This

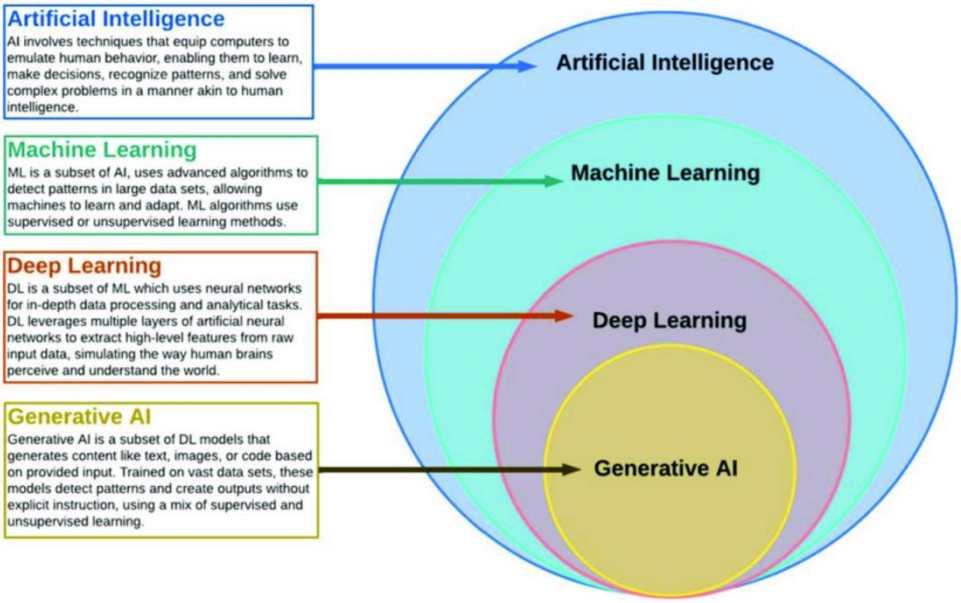
focus on early prediction is particularly relevant in today's business environment, where traditional reactive approaches are increasingly inadequate. Second, predictive models serve as critical tools in the decision-making process, providing managers with actionable insights grounded in data-driven analysis. The probabilistic nature of prediction accommodates uncertainty and supports informed judgments under varying scenarios (Sharda et al., 2020). This aligns closely with the strategic focus of this Thesis, which seeks to enhance managerial decision-making through the integration of new predictive analytics. Third, the effectiveness of predictive models is closely tied to managerial literacy in interpreting and utilizing these tools. Emphasizing prediction underscores the need for enhancing managers' competencies in data analysis, statistical reasoning, and technological proficiency (Guenster et al., 2011). This is particularly important in the context of AI-driven predictive models, which require a higher level of technical understanding to be used effectively.

Furthermore, the term *"prediction*" is extensively used and well-established in academic literature and industry practice when discussing financial distress and bankruptcy (Nguyen et al., 2019; Jones et al., 2017) as it encompasses a broad spectrum of methodologies and applications, often serving as an umbrella term that includes aspects of both forecasting and detection. This widespread usage enhances its understandability and acceptance among stakeholders, including policymakers, investors, and particularly managers, facilitating clearer communication and alignment across different domains. Prediction, as a term, is more understandable to non-specialist stakeholders, who are often responsible for implementing the insights generated by these models. This accessibility is crucial for ensuring that the outputs of predictive models are effectively integrated into organizational decision-making processes.

4. Understanding Artificial Intelligence and Machine Learning in Predictive Studies: a Brief Guide

Artificial Intelligence (AI) and Machine Learning (ML) have, today, become ubiquitous terms in both academic discourse and popular media. However, their overuse has led to a certain degree of misunderstanding, with "*AI"* often being employed as a buzzword to elevate the perceived value of research, products, or services. This paragraph seeks to clarify these concepts, particularly in relation to their application in predicting and managing corporate crises and insolvency, as well as enhancing decision-making processes. At its core, Artificial Intelligence refers to the capability of a machine to imitate intelligent human behavior (Breiman, 2001). This includes a broad spectrum of functionalities, from basic data processing to advanced problem-solving and decision-making (Nguyen et al., 2023). Within the scope of this Thesis, AI is not merely a theoretical construct but a practical tool that is deeply integrated into the strategic and operational aspects of business management. The focus, here, is primarily on the subfield of AI known as Machine Learning, which involves training algorithms to recognize patterns in data and make decisions based on those patterns (Figure 6). The AI explored in this study is not about creating autonomous systems that can think and act independently of human input. Rather, it is about leveraging sophisticated algorithms to support and enhance human decision-making, particularly in the context of anticipating, managing, and mitigating corporate distress. This is where ML plays a crucial role, and should not be taken for granted, cause the majority of new managerial literature on these topics often falls to explain this fundamental perspective.

Figure 6: A Comparative View of AI, ML, Deep Learning, and Generative AI



Source: Zhuhadar & Lytras, 2023

4.1. Machine Learning as Backbone of Predictive Analytics

ML is a subset of AI that has been around since the 1980s, though it has gained significant traction in recent years due to advancements in computing power and the availability of large datasets (Duan et al., 2019; Barboza et al., 2017; Tseng & Hu, 2010). It is essentially the process of training a machine to learn from data (Sigrist & Leuenberger, 2023). This involves feeding an algorithm with historical data so that it can identify patterns, learn from them, and apply this learning to new, unseen data. In the context of this research, ML algorithms are employed to predict corporate crises by analyzing historical financial and, in some cases, non-financial data. The strength of this usbset lies in its ability to process and analyze datasets far beyond human capabilities. However, ML is not without its limitations. The accuracy of its predictions, in fact, is heavily dependent on the quality and quantity of the data it is trained on, and there is always a risk of overfitting, where the model becomes too tailored to the training data and performs poorly on new data.

So, while there has been a surge of interest in Generative AI, especially with the advent of models like GPT-4 and other deep learning architectures, this dissertation deliberately does not focus on Generative landscapes[65]. Generative AI, in fact, involves creating new data instances based on existing data, such as generating text, images, or even synthetic financial data[66]. While this has exciting

[65] About this, see OpenAI (2023), GPT-4 Technical Report here: https://cdn.openai.com/papers/gpt-4.pdf.

[66] Common applications of Generative AI are, today:

- Image Generation: Generative AI can learn from large sets of images and generate new unique images based on trained data. This tool can generate images with creativity based on prompts like human intelligence;
- Video Synthesis: Generative models can create new content by learning from existing videos. This can include tasks such as video prediction, where the model generates future frames from a sequence of input frames. It can also perform video synthesis by creating entirely new videos. Video synthesis has entertainment, special effects, and video game development applications;

applications in fields like natural language processing and creative industries, it is less relevant to the objectives of predicting and managing corporate crises. Gen-AI also operates on principles of creativity and data generation, which, while valuable, do not directly align with the goals of this research. The focus here is on analyzing existing data to forecast potential crises, not generating new scenarios or datasets and the aim is to enhance decision-making by providing managers with insights based on real-world data, rather than speculative or synthetic scenarios. Therefore, while Generative AI represents an exciting frontier in AI research, it does not serve the predictive and managerial needs of this study.

4.2. Value Proposition in Corporate Crisis Management

In summary, this dissertation posits that AI and ML offer significant value in the realm of corporate crisis management and prediction. However, the successful implementation of AI and ML in this context requires a deep understanding of these technologies, both in terms of their capabilities and their limitations. So, the approach taken in this research is to apply ML as a tool for technical analysis, providing predictive insights based on historical data, while also considering the broader strategic and managerial implications of these insights. Furthermore, by focusing on AI in a supportive, rather than generative, role, the research aligns with the practical needs of businesses today, offering a balanced perspective that bridges the gap between technology and real business strategy.

5. Wrapping Up: Final Reflections and Forward Perspectives

This Chapter has laid a comprehensive foundation for understanding the key concepts central to this dissertation, particularly focusing on corporate crisis, insolvency, and bankruptcy within the context of predictive analytics. By differentiating between critical terms such as prediction, forecasting, and detection, the discussion has highlighted the strategic relevance of these concepts in managing critical situations. Furthermore, the emphasis on ML as a cornerstone of AI in this study is particularly significant, as it underscores better the technological evolution from traditional financial models to advanced AI-driven tools. Moreover, this Chapter has also explored the landscape of corporate crisis management, showing how AI and ML are not only tools for technical analysis but also integral to strategic decision-making processes. The discussion on the new Italian insolvency framework and its alignment with broader European regulations further contextualizes the practical application of these technologies, stressing the importance of early intervention of the management. These insights form the theoretical and practical backbone of the dissertation, setting the stage for the empirical analysis that follows.

- Social Media Content Generation: Generative AI can be leveraged to automate content generation for social media platforms, enabling the creation of engaging and personalized posts, captions, and visuals. By training generative models on vast amounts of social media data, such as images and text, they can generate relevant and creative content tailored to specific user preferences and trends.

Chapter 3

Predictive Analytics in Business Adversity: A Journey Through Theory and Methodology

Chapter Prelude

This Chapter embarks on a detailed exploration of predictive analytics within the context of business adversity, bridging foundational theories with contemporary methodologies. Examining the evolution of predictive models, from early financial ratios to the integration of AI, it also provides a rigorous framework essential for analyzing and forecasting critical corporate events. Then, the discussion not only solidifies the theoretical underpinnings of this study but also offers a full methodological clarity that will support the empirical analysis in subsequent parts of the dissertation. This structured approach positions the Chapter as a crucial bridge, seamlessly connecting the theoretical insights and foundational concepts with the practical, data-driven strategies that will be essential in the empirical core of this entire work. Furthermore, the insights provided here will be crucial for forming the basis of conclusions and specific considerations in the later stages of the research, ensuring a comprehensive and critical analysis.

1. Turning Challenges into Opportunities: an Introduction

In the lifecycle of any business or firm, periods of adversity and uncertainty are inevitable. Whether due to external shifts, competitive pressures, or internal mismanagement, companies often find themselves, as stated by Mbuyane et al. (2024), in "*turbulent waters*"[67]. While crises and insolvency are understandably sources of concern, they also present opportunities for growth, renewal, and strategic realignment—if managed with insight and foresight. So, the central thesis of this Chapter is that understanding these challenges deeply and in a timely manner can transform potential pitfalls into stepping stones for long-term sustainability. In fact, the ability to predict challenges before they fully manifest is a powerful tool that can make the difference between a temporary setback and a catastrophic failure (Duan et al., 2019). Predictive analytics, especially when "*enhanced*"[68], provides the framework for this proactive approach and a forward-thinking approach not only protects the firm from immediate threats, but also enhances its capacity to adapt to changing conditions, ultimately contributing to its long-term viability.

Moreover, the process of managing a crisis can lead to significant improvements in a company's operations and strategy: when faced with a crisis, businesses are often forced to re-evaluate their priorities, and focus on their core competencies (Gartner, 2020; Sapountzaki & Wanczura, 2018). This period of introspection can reveal inefficiencies that, when understood and managed effectively, can serve as catalysts for positive change. The main role of predictive analytics in this context is to provide the necessary insights that enable businesses to act with confidence and precision and the value of intelligent, AI-based tools, lies in their ability to complement, rather than replace, human decision-making. In the broader context of business management, the insights gained from these analytics can also ensure stronger relationships with stakeholders as investors, customers, and employees all have a vested interest in the stability and success of a company. In conclusion, while crises and insolvency are often viewed through a lens of negativity, they also offer unique opportunities for businesses to reassess and realign their strategies. The key, in this case, lies in understanding these challenges deeply and responding to them in a timely manner.

1.1. Insights for Strategic Realignment and Other Trends

Strategic realignment in the face of business adversity is a complex process that requires timely and accurate insights to be managed (Anderson & Zhang, 2021). The current scientific literature reveals a strong consensus on the value of innovative tools in providing actionable insights that enable firms to anticipate shifts, mitigate risks, and capitalize on emerging opportunities (Floridi et al., 2018). Recent studies emphasize the role of predictive analytics in facilitating proactive intervention: Nagarajan et al. (2024) argue that predictive models play a critical role in modern strategic decision-making processes, offering data-driven insights that enable firms to realign their strategies in response to external or internal forces. The researchers highlight how these models can transform raw data into strategic assets, allowing firms to forecast changes in consumer behavior, competitive pressures, and economic shifts, which are essential for maintaining competitive advantage over time. Idowu et al. (2023) build upon this foundation by examining the specific impact of AI-enhanced predictive tools on financial distress detection. Their research demonstrates that these tools can effectively identify early signs of financial

[67] About this, see also Allen et al. (2011).
[68] This refers to the use of AI and any "*intelligent*" method.

instability, mainly in volatile markets, where the ability to act quickly can mean the difference between recovery and failure.

Further supporting the integration of predictive analytics into strategic realignment, Yeh & Liao (2012) explore the technical evolution of AI models and their application in business strategy. They argue that advancements in ML have significantly enhanced the accuracy and reliability of predictive models using financial data. The authors also note that the strategic use of ML not only improves decision-making but also enhance agility. The benefits of predictive analytics are further illustrated in a study by Oluwaseun et al. (2023), who found that firms using AI-driven insights for strategic realignment were more likely to outperform their competitors in terms of market share and profitability. This study underscores the importance of incorporating AI into the strategic management process in industries characterized by rapid technological advancements and intense competition. Moreover, predictive analytics has been shown to play a crucial role also in enhancing stakeholder relations: according to Johnson & Phillips (2023), firms that effectively utilize AI in their strategic planning are better positioned to manage stakeholder expectations and build trust. The authors argue that predictive models provide transparency and foresight, which are essential for maintaining strong relationships with investors and employees. This alignment between predictive insights and stakeholder management not only supports business continuity but also fosters a culture of accountability and proactive governance.

Another critical aspect of strategic realignment is the ability to manage organizational change. Regarding this, Harris et al. (2022) discuss how predictive analytics can support change management by providing insights into resistance points within the organization. Their research in fact highlights the importance of data-driven decision-making in guiding the strategic realignment process, particularly during periods of significant transformation. By anticipating and addressing potential obstacles, firms can ensure a smoother transition and maintain operational stability. In addition to improving internal processes, predictive analytics also enhances external market analysis. A study by Chang & Lee (2022) demonstrates how AI tools can be used to monitor and analyze competitor activities, enabling firms to anticipate competitive moves and adjust their strategies accordingly.

However, literature also suggests that predictive analytics can significantly contribute to innovation management. According to Smith et al. (2021), AI-enhanced firms are better able to identify emerging trends and technologies that could disrupt their industry. Furthermore, the role of predictive analytics in enhancing financial performance cannot be overstated. Anderson & Zhang (2021) found that firms utilizing AI-driven financial forecasting models were more likely to achieve higher levels of profitability and financial stability. Their research indicates that predictive analytics enables firms to optimize their financial strategies, manage risks more effectively, and achieve better financial outcomes. In conclusion, the literature overwhelmingly supports the integration of predictive analytics and AI tools into the strategic realignment process.

1.2. Sustainability and Predictive Analytics for a Long-Term Realignment

As Idowu et al. (2023) emphasize, the integration of sustainability metrics into predictive models allows businesses to forecast not just financial performance but also the potential long-term

impacts of their operations on environmental and social outcomes[69]. This dual focus on profitability and sustainability is essential for companies aiming to thrive in an increasingly regulated and environmentally conscious market. One of the significant advancements in this area is the use of big data analytics to monitor and predict ESG-related outcomes (Idowu et al., 2023; Harris & Clements, 2022). For example, predictive models are now routinely employed to analyze data on energy consumption, carbon emissions, and resource utilization, providing actionable insights that help companies reduce their environmental footprint (Nagarajan et al., 2024). This study also highlights that by understanding and acting on these insights, businesses can make strategic decisions that enhance both their sustainability credentials and operational efficiency.

Moreover, predictive analytics is proving indispensable in supply chain management, particularly in optimizing operations for sustainability. Oyelere et al. (2023) discuss how predictive models are being used to forecast supply chain disruptions due to environmental factors such as natural disasters or resource shortages. By anticipating these disruptions, in fact, firms can make informed adjustments to their supply chains, ensuring continuity and reducing waste, which is critical for maintaining sustainable operations. However, the role of predictive analytics extends to regulatory compliance as well (Idowu et al., 2023). With increasingly stringent sustainability regulations being implemented globally, businesses face significant risks if they fail to comply. Idowu et al. (2023) point out that predictive models can help companies forecast the impact of new regulations on their operations, allowing them to adapt proactively and avoid potential fines or reputational damage. This approach to compliance obviously enhances its reputation as a responsible corporate citizen. Further reinforcing the strategic value of predictive analytics is its application in sustainability reporting. Accurate and transparent reporting is in fact crucial for maintaining stakeholder trust, as said in previous paragraphs, and predictive models can enhance this process by providing reliable forecasts of future sustainability performance based on current practices (Harris et al., 2022). This predictive capability allows companies to set realistic sustainability targets and demonstrate their commitment to achieving these goals, thus solidifying their position in the social and economic environment (Nagarajan et al., 2024; Idowu et al., 2023). In fact, as Nagarajan et al. (2024) suggest, the future of corporate sustainability will increasingly depend also on the ability to harness predictive analytics to manage new and innovative ESG or CSR-related regulations[70].

1.2.1. An Holistic Ecosystem Approach

The increasing complexity of business ecosystems, driven by rapid advancements in AI and big data, necessitates a more comprehensive and holistic approach to system design and integration. As technological capabilities have evolved, considerations such as sustainability, security, and ethical implications have taken center stage in the discourse around AI adoption (Perko et al., 2017). Traditional approaches that focused solely on the technical aspects of IT deployment are no longer sufficient, as modern business environments demand a more nuanced understanding of how AI systems interact with both operational and societal factors (SC 42, 2020; Duan et al., 2019). Perko et al. (2017) discuss the importance of adopting a holistic view when developing AI systems, emphasizing that modern business ecosystems are shaped by an increasing diversity of stakeholders, including regulators, social scientists,

[69] These impacts can include, but are not limited to, the depletion of natural resources, contribution to climate change, pollution, effects on biodiversity, as well as social consequences such as community displacement, changes in local economies, and broader societal well-being. About this, see also Harris & Clements (2022).

[70] About this, see also Oyelere et al. (2023).

and technologists. This sentiment is echoed by Thiebes et al. (2020), who argue that the inclusion of ethical and societal considerations is crucial for ensuring the trustworthiness and transparency of AI systems. In fact, as AI technologies become more embedded in critical decision-making processes, the need for early engagement with diverse stakeholders is paramount to developing systems that are both functional and socially responsible (Wamba-Taguimdje et al., 2020). The literature also highlights the need for AI systems to be adaptable and scalable across different domains. According to SC 42 (2020), the "*data ecosystem*" has become as critical as hardware and software in shaping business operations, with predictive analytics playing a key role in anticipating market trends and enabling strategic realignment. In this regard, Perko et al. (2017) note that AI systems must be maintainable and portable across different use cases, which is essential for ensuring long-term sustainability and efficiency. Duan et al. (2019) further elaborate on the necessity for AI systems to be integrated into larger technology deployments, arguing that this integration enhances operational resilience and supports business agility.

Trustworthiness remains a central theme in discussions on AI adoption. Thiebes et al. (2020) underscore the importance of transparency in AI systems, particularly in industries where regulatory compliance is critical. Building on this, SC 42 (2020) introduces a framework for AI governance that incorporates ethical and societal considerations across the entire lifecycle of AI systems. This governance framework aims to ensure that AI systems are not only technically sound but also aligned with broader societal goals, including sustainability and social responsibility (Duan et al., 2019). From a strategic perspective, predictive analytics has emerged as a powerful tool for overcoming business complexity. Wamba-Taguimdje et al. (2020) argue that predictive models provide actionable insights that help businesses anticipate shifts in different landscapes. Furthermore, Perko et al. (2017) discuss how the adoption of AI and big data in business ecosystems is increasingly driven by the need for sustainability.

In conclusion, the literature consistently supports the view that a holistic ecosystem approach is essential for the successful integration of AI and predictive analytics into business operations. By addressing both technical and non-technical requirements, businesses can develop AI systems that are trustworthy, adaptable, and aligned with broader societal goals. The work of Duan et al. (2019), Thiebes et al. (2020), and Perko et al. (2017) demonstrates that this integrated approach is key to ensuring the long-term sustainability and resilience of AI systems in complex business ecosystems. As businesses continue to adopt AI, the need for comprehensive governance frameworks, stakeholder engagement, and sustainability considerations will only grow, shaping the future of AI in business.

2. The Evolution of Corporate Crisis Prediction and Managerial Theories: From Traditional Financial Models to Modern Analytics

To fully appreciate the advancements brought by AI and intelligent tools in predictive studies, it is essential to first explore and understand the foundational methods and traditional literature in corporate crisis and insolvency prediction. In fact, understanding the strengths and limitations of traditional models provides a necessary context for discussing how modern predictive analytics, empowered by AI, builds upon and enhances earlier frameworks. The history of corporate crisis and insolvency prediction, particularly in its early stages, is closely tied to the development of simple financial models that were limited by the technological capabilities of the time. In the absence of advanced computational tools, early researchers like FitzPatrick (1932) focused on straightforward

univariate analysis, which compared key financial ratios—such as liquidity, profitability, and solvency—between firms that had failed and those that remained solvent. This method marked one of the earliest systematic attempts to predict corporate distress by relying on readily available accounting information. It also allowed for an accessible yet foundational approach to understanding financial health.

However, the simplicity of these early models was both their strength and their limitation. While univariate analysis provided crucial insights, it was constrained by its focus on individual financial metrics, which made it difficult to capture the multifaceted nature of business failure: financial distress, especially in complex industries, is rarely caused by a single factor, and thus a more sophisticated analytical framework was needed. Additionally, these models were backward-looking, relying entirely on historical data and offering no forward-looking perspective or adaptability to changing market conditions, which would become a key feature of later models. As such, while FitzPatrick and his contemporaries laid an important foundation for financial ratio analysis, they also underscored the need for the evolution toward more comprehensive, multivariate models capable of better addressing the complexities of corporate insolvency (Ohlson, 1980; Altman, 1968).

The limitations of early approaches became more apparent as the business environment grew increasingly complex, and it was clear that a more nuanced understanding of financial distress was necessary. Subsequent models, such as Altman's Z-Score, which combined multiple financial ratios into a single predictive model, marked a shift from the univariate techniques of FitzPatrick to multivariate analysis, allowing for a more comprehensive assessment of financial health. These advancements reflected a growing recognition that financial distress is a multi-dimensional phenomenon, influenced by a range of internal and external factors. As financial modeling advanced, researchers also began to incorporate broader business considerations into their analyses. In fact, the rise of professional management in the mid-20th century and the increasing separation between ownership and control brought new dimensions to the understanding of corporate governance and crisis management. Chandler (1962) noted that as companies grew in scale and complexity, the role of managers became more pronounced, particularly in overseeing operations and ensuring financial stability. This was a significant evolution from earlier periods, when the concept of "*management*" was less formalized, and business decisions were often made by owners or family members, particularly in countries like Italy, where family businesses dominated well into the 20th century (Colli, 2003).

In this context, financial distress came to be seen not just as a failure of financial performance but also as a failure of management (Colli, 2003). The literature of the time, influenced by management theorists like Fayol (1916) and Taylor (1911), emphasized the role of planning, organization, and control in maintaining business health, although these early theories focused more on efficiency than on crisis management. It was only later, as companies faced more frequent and complex crises due to globalization and market volatility, that management theories began to emphasize the importance of crisis prediction and proactive strategic interventions (Mintzberg, 1973). In summary, the early development of corporate crisis prediction models was constrained by technological limitations and a narrow focus on financial ratios, but it laid the groundwork for future innovations. The transition from univariate to multivariate models, alongside the evolving role of management, reflected a growing understanding of the complexities of financial distress and the need for more comprehensive tools to predict and manage corporate crises.

2.1. The Role of Management in Crisis Prediction

The role of management in corporate crisis prediction has evolved significantly over the past century, paralleling the development of financial models and the professionalization of management itself. In the early stages of corporate distress prediction—during the eras of FitzPatrick (1932), Altman (1968), and Ohlson (1980) the concept of management as a distinct, professional function was still in its formative stages, particularly in regions like the United States and Western Europe. However, understanding the evolving role of managers during this specific period provides key insights into how management's responsibility for crisis prediction and prevention developed, as well as how it differed across industries and countries.

2.1.1. Early Views: the Mid-20th Century

In the early 20th century, the concept of a professional manager as we understand it today was not yet fully established. Companies were typically owned and managed by the same individuals or families, particularly in industries that were not highly capital-intensive. In countries like Italy, where family-owned businesses dominated well into the 20th century, the separation of ownership from management was uncommon (Colli, 2003). Thus, corporate crises were often seen as failures of the entire organization, with little distinction between managerial oversight and ownership decisions (lo Conte et al., 2023). In contrast, by the mid-20th century, large corporations in the United States and the United Kingdom began to develop more formalized managerial roles. The work of Alfred D. Chandler (1962), in Strategy and Structure, highlighted the importance of the growing distinction between corporate ownership and management: Chandler argued that as companies expanded and became more complex, professional managers were increasingly responsible for the operational stability of firms, including the management of financial health and crisis prevention. However, crisis management itself was not a focal point of managerial responsibility during this period. Instead, early management theories, such as those proposed by Fayol (1916) and Taylor (1911), concentrated on efficiency, organizational structure, and productivity: the idea that managers should play a proactive role in anticipating and mitigating financial crises would not become mainstream until later decades, as firms became more exposed to global market forces and financial volatility.

2.1.2. The Impact of Financial Models on Management's Role in Crisis Prediction

As financial ratio models like those developed by FitzPatrick (1932) and Altman (1968) became widely adopted, the role of management in crisis prediction started to gain prominence. These models provided managers with a systematic way to assess financial health, enabling them to better understand the warning signs of financial distress (Altman, 1968). The development of multivariate analysis, which incorporated multiple financial indicators into a single predictive model, also increased the relevance of these tools for managers. Altman's Z-Score (1968) was particularly influential because it allowed managers to assess the likelihood of corporate bankruptcy with greater precision than ever before. This model, which combined ratios related to liquidity, profitability, and leverage, gave managers a quantitative tool that could inform their decision-making processes. As a result, crisis prediction became increasingly recognized as a managerial responsibility, particularly in capital-intensive industries like

manufacturing and heavy industry, where financial health was closely tied to operational success. Managers in these industries, particularly in the U.S., began to rely more heavily on financial ratios and other quantitative models as part of their broader strategic management frameworks (Altman, 1968).

Ohlson's O-Score (1980), which introduced more sophisticated logistic regression techniques into bankruptcy prediction, further reinforced the idea that managers could and should use financial data to anticipate potential crises. However, the adoption of these models was not uniform across industries or geographies. For example, while American corporations, particularly in manufacturing and finance, quickly integrated these predictive models into their managerial practices, European companies were slower to adopt them, particularly in regions like Southern Europe, where family-owned businesses and less formalized management structures remained dominant (Colli, 2003).

2.1.3. Cross-Country Differences in Managerial Approaches

The role of managers in crisis prediction varied significantly between countries during this period. In the United States and the United Kingdom, where managerial capitalism had taken firm root, the responsibility for anticipating and managing financial distress was increasingly seen as a core component of the managerial role. This shift was supported by the development of a robust financial infrastructure, which included not only predictive models like the Z-Score and O-Score but also a more developed system of financial reporting and corporate governance (Chandler, 1962). In contrast, countries like Italy, Germany, and Japan had different managerial traditions. In Italy, family ownership structures were predominant well into the 20th century, and the formal separation between owners and professional managers was less common. This affected the way crises were predicted and managed, as decision-making remained concentrated within a small group of owners who often lacked the sophisticated financial tools that were becoming standard in the U.S. and U.K. (Colli, 2003). In Germany and Japan, the role of managers in crisis prediction was more closely tied to the broader socio-economic systems in which firms operated. In Germany, for instance, the stakeholder model of corporate governance, which prioritized long-term stability and consensus among stakeholders, meant that crisis prediction was more closely linked to maintaining industrial relations and social stability (Wever, 1995). Similarly, in Japan, the "*keiretsu system of interlinked corporate groups*" provided a buffer against financial distress, reducing the pressure on individual managers to predict and prevent crises (Gerlach, 1992)[71].

[71] The *keiretsu* system is a distinctive organizational structure found in Japan, characterized by networks of interlinked corporate groups that share ownership stakes in each other and are often centered around a major bank or trading company. This system evolved from Japan's *zaibatsu* model of conglomerates, which was dissolved after World War II, and has been instrumental in shaping the country's post-war industrial and financial development (Lincoln, 1998; Gerlach, 1992). There are two primary forms of *keiretsu*: *horizontal keiretsu* (also known as financial keiretsu) and *vertical keiretsu* (or production keiretsu). In *horizontal keiretsu*, companies from diverse industries form close relationships through cross-shareholding, interlocking directorates, and coordinated financing, often centered around a major bank. Examples of these include the Mitsubishi and Sumitomo groups (Gerlach, 1992). These groups provided member firms with financial stability, reducing the risk of bankruptcy and fostering long-term strategic alignment. In contrast, *vertical keiretsu* are centered around a major manufacturing company, such as Toyota, and include suppliers, manufacturers, and distributors working together to streamline production processes (Dyer, 1996).
This structure enabled Japanese companies to manage financial risk more effectively, as *keiretsu* networks often acted as a safeguard against external economic shocks by providing mutual support within the group. For instance, banks would provide credit to struggling member companies, and these companies would prioritize business within the group, creating a closed-loop system that bolstered financial stability (Lincoln & Gerlach, 2004).
While the *keiretsu* system was highly effective in promoting industrial growth in Japan, particularly during the post-war economic boom, it has faced criticism for fostering inefficiencies and insulating companies from competitive pressures. In recent decades, globalization and deregulation have led to the gradual dismantling or weakening of many *keiretsu* structures, as companies seek more flexible and competitive strategies (Lincoln, 1998).

2.1.4. The Shift Toward Crisis Management as a *"Managerial Responsibility"*

By the late 20th century, the role of management in corporate crisis prediction had become more formalized, particularly as globalization and financial deregulation exposed companies to new risks. Scholars like Mintzberg (1973) began to argue that crises, far from being seen as inevitable failures, could represent opportunities for strategic realignment and renewal. This perspective shifted the focus of managerial responsibility from simply reacting to crises to proactively predicting and managing them. Predictive financial models, including those developed by Altman and Ohlson, became central tools for managers seeking to avoid insolvency, particularly in industries characterized by high levels of capital investment and financial risk (Altman, 1968). In addition to financial models, the development of more sophisticated management theories, such as those emphasizing strategic management and organizational agility, contributed to the growing recognition of crisis management as a key managerial function. Managers were no longer seen as mere overseers of operations but as strategic leaders responsible for guiding firms through periods of financial instability. This shift was reflected in both academic literature and corporate governance practices, particularly in Anglo-American countries, where managerial accountability for financial performance became more pronounced (Mintzberg, 1973; Chandler, 1962).

In conclusion, it is possible to affirm that the historical evolution of predictive studies reveals a growing convergence between financial modeling and managerial responsibility. Early models, such as those introduced by FitzPatrick, Altman, and Ohlson, provided managers with the tools to quantitatively assess financial health, but it was the broader socio-economic context and evolving management theories that solidified the role of crisis prediction as a key managerial responsibility. As businesses continued to globalize and industries became more complex, the need for predictive financial models that could accommodate a wide range of variables—including non-financial factors like market trends and regulatory shifts—became apparent. So, in the following sections, the focus will shift to a deeper examination of these traditional models, their limitations, and how they have paved the way for the more dynamic, AI-driven predictive tools used in modern business ecosystems[72].

2.2. State of the Art

While AI and new technologies introduce sophisticated mechanisms for analyzing complex datasets, human judgment remains crucial in interpreting and guiding the insights derived from these models. So, this review not only focuses on traditional models for crisis prediction but also reflects on the dynamic progression of the field, incorporating modern AI techniques that aim to support managerial decisions in crisis and insolvency contexts. As already stated, the field of corporate financial distress prediction has experienced significant advancements over the decades, evolving from traditional financial analysis methods to incorporating advanced computational techniques; these developments have led to increasingly sophisticated models that aim to predict financial distress more accurately and in a timely manner. The late 20th and early 21st centuries were marked by rapid globalization, which significantly altered the business landscape. According to Ko and Lin (2006), this period was characterized by turbulent changes as economies became more interconnected, leading to

[72] The present examination regards just some traditional methods. These were chosen based on study's objectives and degree of importance.

increased competition, market volatility, and the need for more dynamic financial strategies. Firms operating in this global environment face heightened risks of financial distress due to unpredictable economic shifts. Michalkova et al. (2021) also argue that successful business operations in such a context require a deep understanding of the global environment and continuous monitoring of financial performance. Financial analysis thus became essential for assessing a company's financial health, as highlighted by Kaczmarek et al. (2022), who emphasized that the evaluation of key financial indicators helps businesses anticipate potential crises and adjust their strategies accordingly.

The purpose of financial analysis has always been to assess the financial stability of a firm based on its historical performance, usually through the examination of financial statements. Balina et al. (2021) reinforce this idea by stating that such analysis encapsulates a firm's operational and financial performance in a compressed format, which allows managers and stakeholders to make informed decisions about the firm's future. However, while traditional financial analysis plays a vital role, the dynamic and unpredictable nature of the globalized economy demands more sophisticated approaches to accurately forecast corporate failure and insolvency. So, as already said, the origins of bankruptcy prediction models trace back to the early 20th century, with studies by Fitzpatrick (1932), Smith and Winakor (1935), and Merwin (1942). In particular, these early researchers demonstrated that certain financial indicators, such as profitability and liquidity ratios, began to diverge in financially distressed companies well before they became insolvent. Furthermore, Ansoff's (1965) concept of "*weak signals*" in strategic planning introduced the idea that early signs of corporate distress are often difficult to identify but can be detected through careful analysis[73]. This concept influenced the development of early warning systems and laid the groundwork for more predictive models. Beaver (1966) made a significant contribution by using financial ratios to predict bankruptcy and showing that these indicators could, indeed, forecast financial difficulties. However, Beaver also noted that not all financial ratios had the same predictive power. The landmark study by Altman (1968) built on this foundation by developing a multivariate model—the Z-Score—which combined several financial ratios to assess a company's likelihood of bankruptcy. Altman's model became the cornerstone for many subsequent models and proved invaluable for predicting bankruptcy with greater accuracy. As a result, it has been widely used by banks, investors, and financial institutions for assessing credit risk and determining the likelihood of loan defaults.

As data availability increased and computational methods advanced, prediction models began incorporating more sophisticated techniques. During the late 1970s and early 1980s, logit and probit models were introduced as alternatives to linear discriminant models. Ohlson (1980) was the first to apply logistic regression to predict bankruptcy, while Zmijewski (1984) employed probit analysis. These models improved predictive accuracy by addressing some of the limitations of earlier models, such as

[73] Ansoff's "*weak signals*", a concept originally introduced in his strategic management framework, refer to early, subtle indicators of potential changes in a business environment that, if detected early, allow organizations to preempt major disruptions. These signals are particularly relevant to insolvency prediction, where early signs of financial distress can be hidden in ambiguous or seemingly insignificant data points. Recognizing these signals enables companies to adjust their strategies in time to avert crises. According to Holopainen & Toivonen (2012) in *Futures*, weak signals have gained renewed relevance in today's volatile business climate. They describe how weak signals serve as precursors to major trends or crises, often emerging well before their full impact is felt. This makes them valuable for firms seeking to manage insolvency risks proactively. For example, a slight decline in customer satisfaction, a shift in supplier behavior, or minor liquidity issues could serve as early warnings of deeper financial problems to come. These signals, although difficult to interpret, can give firms the time they need to respond strategically to avoid insolvency. In the same vein, Choo (2011) discusses how weak signals can be systematically integrated into strategic issue management systems, offering a structured approach to detect and respond to emerging risks. This early detection capability aligns with Ansoff's original emphasis on strategic flexibility and foresight, which are critical in navigating business challenges, including insolvency. The combination of these perspectives reinforces the importance of incorporating weak signal detection into management practices. By doing so, organizations can enhance their resilience and agility in responding to financial distress, ultimately improving their ability to avoid insolvency.

the assumption of linear relationships between variables. Simultaneously, researchers began exploring theoretical models that did not rely solely on empirical financial data. Merton's (1974) model, which introduced the concept of corporate bankruptcy prediction based on the theory of options pricing, demonstrated how complex mathematical frameworks could be applied to assess a firm's financial risk. Although these methods showed promise, their reliance on advanced mathematical techniques limited their accessibility and adoption by practitioners[74]. So, the development of AI and ML technologies in the latest 1980s/1990s further transformed the landscape of insolvency and crisis prediction. With the advent of neural networks and genetic algorithms, models could process vast amounts of data and detect different patterns that were difficult to identify using traditional statistical methods (Barboza et al., 2017). Neural networks, for instance, were used to predict bankruptcy and insolvency with high accuracy by learning from historical data and adjusting their parameters to improve prediction over time (Paliwal & Kumar, 2009). Genetic algorithms, introduced by Tseng and Hu (2010), optimized bankruptcy prediction by simulating the process of natural selection, selecting the most relevant variables to predict financial distress.

However, these AI-driven models presented new challenges, particularly around interpretability. The "*black box*" nature of neural networks and genetic algorithms made it difficult for managers and decision-makers to understand how predictions were made, which raised concerns about the transparency and reliability of these models (Kapounek et al., 2022). As Jones (2017) highlights, many statistical methods used in prediction models are constrained by assumptions of linearity and normality, which do not always align with the realities of complex business environments. Furthermore, according to Kou et al. (2021), the success of a model largely depends on the accuracy and reliability of the data it processes. However, Kubenka et al. (2021) caution that input data can often be distorted due to accounting practices, economic changes, or corporate governance issues, which can affect the model's predictive power. On the theme, Michalkova et al. (2022) argue that a firm's success or failure is influenced not only by internal financial factors but also by external economic conditions. Therefore, models that do not account for the broader economic context may produce inaccurate predictions. Kliestik et al. (2022) further emphasize that prediction models must be updated regularly to reflect changes in the global economic environment, as static models based on outdated data may lead to misleading forecasts.

In conclusion, the field of insolvency and crisis prediction has progressed from basic financial ratio analysis to incorporating advanced statistical techniques and AI-driven models. However, while AI offers unprecedented predictive capabilities, its effectiveness is contingent on the quality of the data it processes and the context in which it is applied. Moreover, the "*black box*" nature of many AI models remains a significant challenge, as managers and decision-makers require transparency to trust and act upon the predictions.

2.2.1. The Italian Context

The study of insolvency prediction in Italy has evolved significantly over the decades. This evolution mirrors broader global trends, as companies worldwide seek more sophisticated tools to predict and manage financial distress (lo Conte & Sancetta, 2024; Ferri et al., 2022). However, the Italian context presents unique challenges and opportunities, particularly in terms of governance structures, family-owned businesses, and the specific regulatory environment. Understanding the trajectory of

[74] In the context of this dissertation, this element in the state of art can be considered as the starting point of the implementation of algorithms in predictive analytics regarding corporate crises and insolvency.

these research efforts is then essential to appreciate the current state of the art and the potential future developments in insolvency prediction.

Scholars like Argenti (1976) and Guatri (1986) focused on the internal and external factors contributing to business failures, emphasizing managerial and operational missteps as key causes of corporate decline. Alberici (1975) was a pioneer in developing quantitative models for predicting insolvency, introducing a multi-year scoring system that assessed a company's financial health over the five years preceding a crisis. This marked an important step in the evolution of bankruptcy prediction models in Italy, laying a foundation for more advanced approaches that followed. Furthermore, the Altman Z-score, originally developed for U.S. firms, was adapted for Italian companies by Altman, Danovi, and Falini (2013). With the advent of AI and ML, the field of insolvency prediction has witnessed transformative advancements even in the Italian context. AI models, particularly Random Forests and Logit regression, have been applied to predict insolvency with far greater accuracy than traditional financial models. A key study by Dallocchio et al. (2020) highlights the use of AI to predict financial distress in non-listed Italian firms classified as "*Unlikely to Pay*" (UTP). These firms, though not yet insolvent, are at a high risk of bankruptcy and require strategic restructuring to avoid collapse. Dallocchio et al.'s research demonstrated that corporate governance variables—such as CEO turnover, board composition, and the stability of the executive team—were strong predictors of financial distress. Notably, the Random Forest model outperformed traditional Logit models, offering greater accuracy and deeper insights into the non-linear relationships between governance factors and financial health (Dallocchio et al., 2020). This unconventional approach marked a departure from the exclusive focus on financial ratios, bringing corporate governance into the spotlight as a critical element in crisis prediction. However, while AI offers significant advantages in predicting insolvency, it also raises important ethical concerns, particularly regarding transparency and accountability. One of the central challenges with AI models, such as Random Forests and deep learning algorithms, is their "*black box*" nature, which makes it difficult for managers and stakeholders to understand how predictions are generated (Zhang & Susheng, 2021; Floridi et al., 2018). This lack of interpretability can undermine trust in AI systems, especially in high-stakes situations like financial distress prediction, where decision-makers need to justify their actions to shareholders and regulators.

Moreover, the pervasive use of AI in corporate decision-making could lead to a form of "*social scoring*" within companies, where decisions about creditworthiness or risk are made algorithmically without sufficient human oversight[75]. The AI Act, under development by the European Union, seeks to address some of these concerns by introducing regulations that ensure transparency, accountability, and fairness in AI systems. This legislation is particularly relevant to financial models, where the consequences of incorrect predictions can have severe ramifications for both businesses and

[75] The AI Act, recently adopted by the European Parliament, represents a landmark in the regulation of artificial intelligence within the EU, setting out strict rules aimed at ensuring AI systems are used ethically, transparently, and safely. One of the key concerns addressed by the Act is the prevention of "*social scoring*" practices, which involve AI systems systematically ranking or evaluating individuals based on their social behaviors or personal characteristics. This type of scoring is seen as problematic due to the risks it poses to privacy, equality, and human dignity. Social scoring refers to the use of AI algorithms to monitor, evaluate, and classify individuals, often resulting in unfair or discriminatory outcomes. Such practices can lead to significant social and legal consequences, as individuals may be categorized and treated differently based on AI-driven assessments of their behavior. In countries like China, social scoring has raised global concerns, as it can determine access to services, jobs, or loans. The AI Act prohibits any such practices within the EU, emphasizing the need to safeguard individuals' rights. According to Recital 31 of the AI Act, systems designed to evaluate or score individuals over time based on their behavior are prohibited due to their potential to violate fundamental rights. The Act classifies these systems as high-risk AI and mandates that they undergo strict oversight, transparency measures, and human intervention to ensure fairness and accountability. The law is intended to prevent situations where AI could be used to control or manipulate human behavior by assigning people social rankings that can impact their opportunities and freedoms. For further details, the AI Act can be explored in-depth on the European Parliament's official news page and the AI Act Law Recital 31.

stakeholders (lo Conte & Sancetta, 2023). In this context, AI should be viewed as a tool to support human decision-making rather than as an autonomous decision-maker. It is crucial that AI complements human expertise by providing data-driven insights that enhance managerial judgment. This approach fosters a symbiotic relationship where technology and human insight work together to drive better outcomes (Zhang et al., 2021). For instance, Dallocchio et al. (2020) showed that combining AI with corporate governance analysis not only improved prediction accuracy but also offered actionable insights for boards and executives, helping them make more informed decisions about restructuring and turnaround strategies.

However, despite the advancements in AI research, there is a notable gap in Italian literature concerning large-scale empirical studies on insolvency prediction, especially those involving small and medium-sized enterprises. Most studies focus on either a limited number of firms or are heavily technical, emphasizing statistical and Machine Learning models without integrating broader managerial and corporate governance considerations, and, mainly, without an effort to enhance explainability of the models and bridging the gap between scientific research and practical and professional world (lo Conte & Sancetta, 2024). Furthermore, research on SMEs, which form the backbone of the Italian economy, remains underdeveloped. These companies often lack the resources to implement advanced AI-driven models, yet they are the most vulnerable to economic downturns and market disruptions (Aliaj & Piersanti, 2020). This lack of empirical research presents a significant challenge for the broader adoption of AI in insolvency prediction. SMEs, in particular, could benefit greatly from AI-driven models that provide early warnings of financial distress, allowing them to take proactive measures to avoid insolvency. However, the high cost of AI implementation and the lack of accessible, high-quality data make it difficult for these firms to leverage such technologies effectively.

2.2.2. Conclusion: The Future of Insolvency Prediction in Italy

The integration of AI and ML into insolvency prediction models marks a significant evolution in the field, offering more accurate and actionable insights than traditional financial models. However, challenges remain, particularly in terms of transparency, ethical deployment, and the need for more large-scale empirical studies involving SMEs. Italian research has made substantial progress in adapting global advancements to the local context, as evidenced by studies like Tron (2021), but much work remains to be done. The future of insolvency prediction in Italy will depend on balancing the power of AI with the need for ethical, transparent, and interpretable models that complement human expertise. As AI continues to evolve, it will be crucial to ensure that these technologies are used responsibly, determining innovation while safeguarding the interests of businesses, stakeholders, and society at large, and avoiding critical perspectives like those of AI washing and algorithmic overtrust (Sundstrom, 2024; Giraud et al., 2022)[76].

[76] AI washing occurs when commercial and institutional entities misuse the term "*AI*" to describe technologies that lack advanced ML or deep learning techniques. This practice aims to leverage media enthusiasm and political support for AI, leading to an overestimation of the system's capabilities and fostering a false perception of innovation. In particular, AI washing could manifest through practices such as:

- *Inaccurate Labeling:* Using the term "*artificial intelligence*" to describe simple rule-based systems or heuristic algorithms;
- *Ambiguous Marketing Campaigns*: Promotional material that emphasizes terms like "*algorithm*" or "AI" without technical details, implying cognitive-like capabilities;
- *Omission of Operational Complexity:* Presenting tools devoid of machine learning as though they possess autonomous learning abilities.

On the other side, Algorithmic overtrust describes a psychological and behavioral phenomenon wherein users and decision-makers place excessive trust in the recommendations provided by algorithms, overlooking potential errors, biases, or model

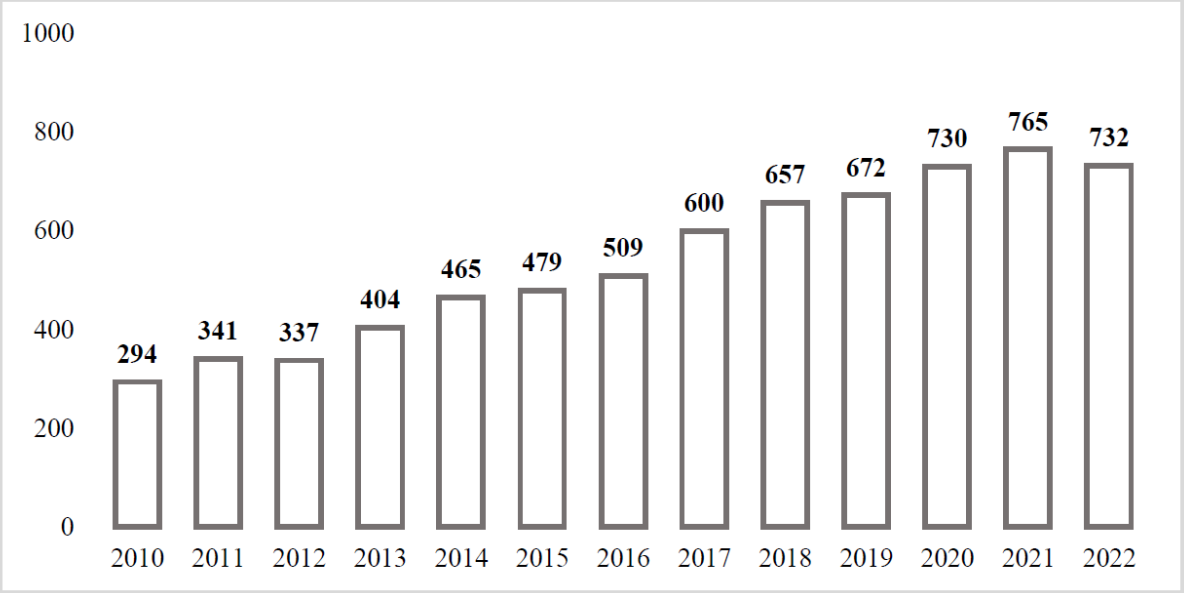
2.3. Rising Publication Trends on the Topic

The topic of insolvency and crisis prediction, which first emerged in academic literature in 1933, began to see significant scholarly attention worldwide only after 2005, when the number of published works surpassed 100 papers. This surge coincided with the evolving role of managerial oversight in businesses, as corporate structures became increasingly complex and intertwined with broader economic systems (Zhang et al., 2021). The evolution of Internet also enhanced the trend. By 2010, publications on this subject had tripled, signaling a clear pivot towards a deeper understanding of corporate failure. This year, in fact, serves as a key milestone for the current research paper, marking the beginning of a more focused examination of innovative prediction models. A closer look at the literature shows how the growing emphasis on managerial control and accountability has shifted the perception of bankruptcy from a purely financial concern to a more human and stakeholder-centered issue. As companies expanded in scope, their crises were no longer confined to balance sheets but became directly tied to the livelihoods of employees, suppliers, investors, and entire communities. So, the shift in focus has drawn increased attention to the social and human consequences of corporate collapse, which has, in turn, driven the growing academic interest in predicting financial distress.

Interestingly, despite this emphasis on human impacts, the latest research in bankruptcy prediction has paradoxically shifted towards AI-driven, data-centric models. While early studies largely relied on qualitative assessments and traditional financial ratios, recent advances have transformed the field. As highlighted in various studies, including those referenced before, AI tools are becoming indispensable for companies aiming to forecast financial health and make informed decisions long before the threat of insolvency becomes imminent. The evolution of insolvency prediction research reflects this duality—where the rise of managerial roles brought human and ethical dimensions to the forefront, modern AI-driven methods are pushing the field towards a more analytical, data-based approach. While the consequences of financial distress remain deeply personal for those affected, the strategies to predict and prevent such crises are becoming ever more technologically advanced. This contrast between the human stakes and the impersonal precision of AI underscores the complexity of contemporary corporate governance. As shown in Figure 7, Gajdosikova & Gabrikova (2023) demonstrates the steady increase in publications until 2016, followed by a surge in academic output, illustrates how the field has evolved in response to the changing needs of both businesses and society. In 2021, a record 765 papers were published on prediction, highlighting its growing importance, particularly as companies face increasingly unpredictable market conditions because of covid-19 pandemic. Though the number of publications dropped slightly to 732 in 2022, the continued focus on this area underscores its critical relevance in modern-day corporate strategy, where timely prediction can mitigate not just financial losses but also human and social repercussions.

limitations. This phenomenon is particularly concerning in high-stakes decision-making contexts, such as medical diagnosis and public resource allocation.

Figure 7: Annual Growth of Scientific Papers Regarding Insolvency and Crisis Prediction (2010-2022)[77]



Source: Gajdosikova & Gabrikova, 2023

Moreover, the study reveals a clear dominance of Articles, as depicted in Table 4. This is not surprising, as academic Articles often provide in-depth analysis, extensive empirical studies, and critical reviews of existing literature. On the other hand, conference proceedings represent 14.36% of all documents related to the prediction of corporate distress and insolvency. While smaller in proportion, these contributions are crucial for capturing the most up-to-date developments and cutting-edge research in rapidly evolving areas, such as AI-based predictive models in crisis management. Conference proceedings are often the first venue where new ideas and methodologies are tested and shared with the academic community. In fact, they provide a platform for researchers to present preliminary findings, foster academic discussions, and receive valuable feedback that can refine their work for future publication in full-length articles.

Table 4: Types of Documents

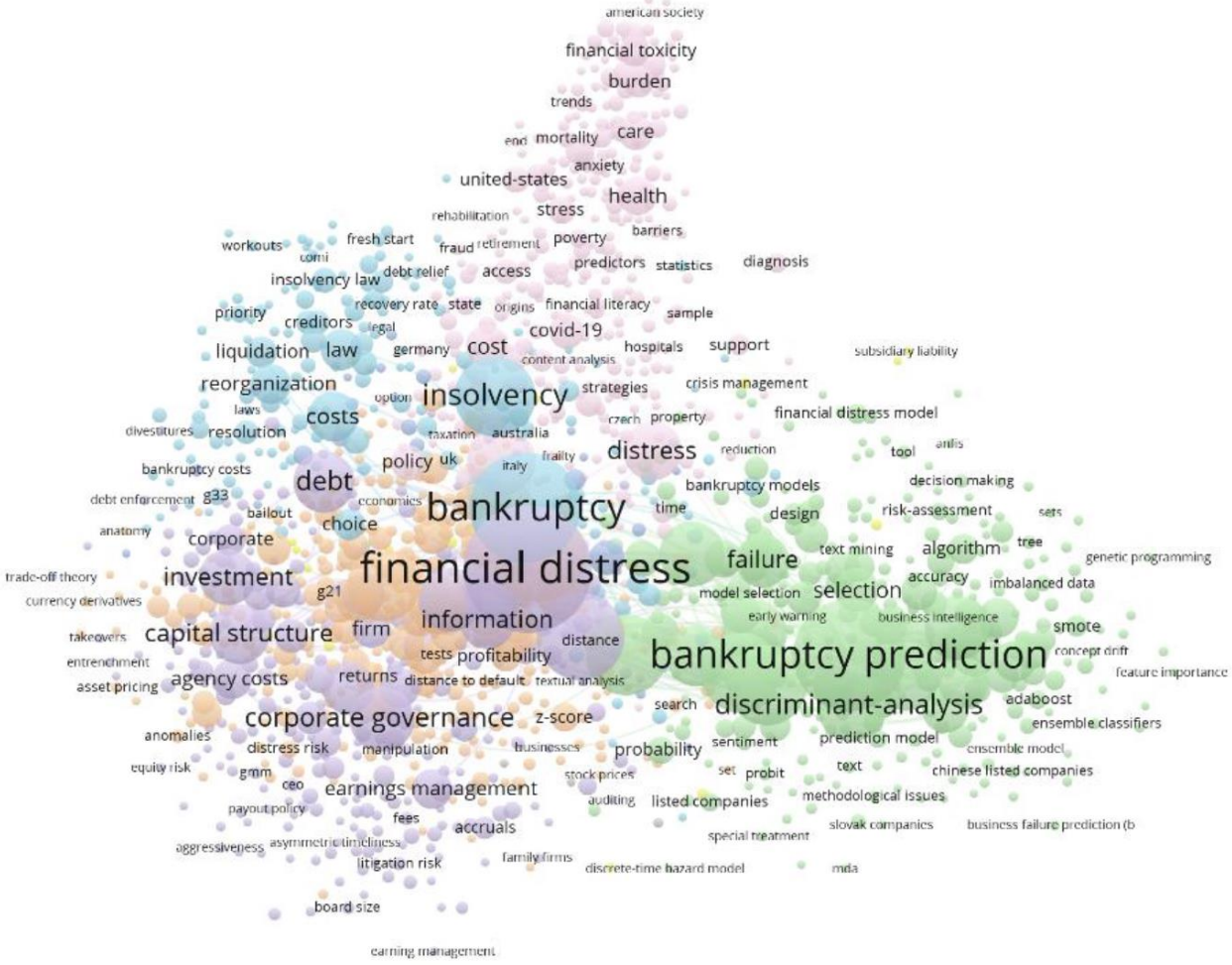
Document type	Frequency	Percent
Article	5,593	80.07
Proceedings Paper	1,003	14.36
Book Chapters	238	3.41
Review Article	84	1.20
Editorial Materials	64	0.92
Others	3	0.04

Source: Gajdosikova & Gabrikova, 2023

[77] Database used: Web of Science.

The bibliometric analysis conducted by Gajdosikova & Gabrikova (2023) also offers a detailed and comprehensive view of the global research network surrounding corporate insolvency prediction. This analysis not only sheds light on the interconnected nature of international research efforts but also highlights the most frequently used keywords and thematic clusters that dominate the academic discourse on the topic[78]. Through the mapping of 1,434 items divided into eight distinct clusters, their work underscores the complexity and breadth of research in this field, providing essential insights into the evolving body of knowledge[79]. In particular, the bibliometric map generated identifies the most frequently used keywords in prediction research (Figure 8). These keywords are organized into eight clusters, each representing a different thematic focus within the broader context of insolvency prediction.

Figure 8: Insolvency & Crisis Prediction Research Bibliometric Map (Used Keywords)



Source: Gajdosikova & Gabrikova, 2023

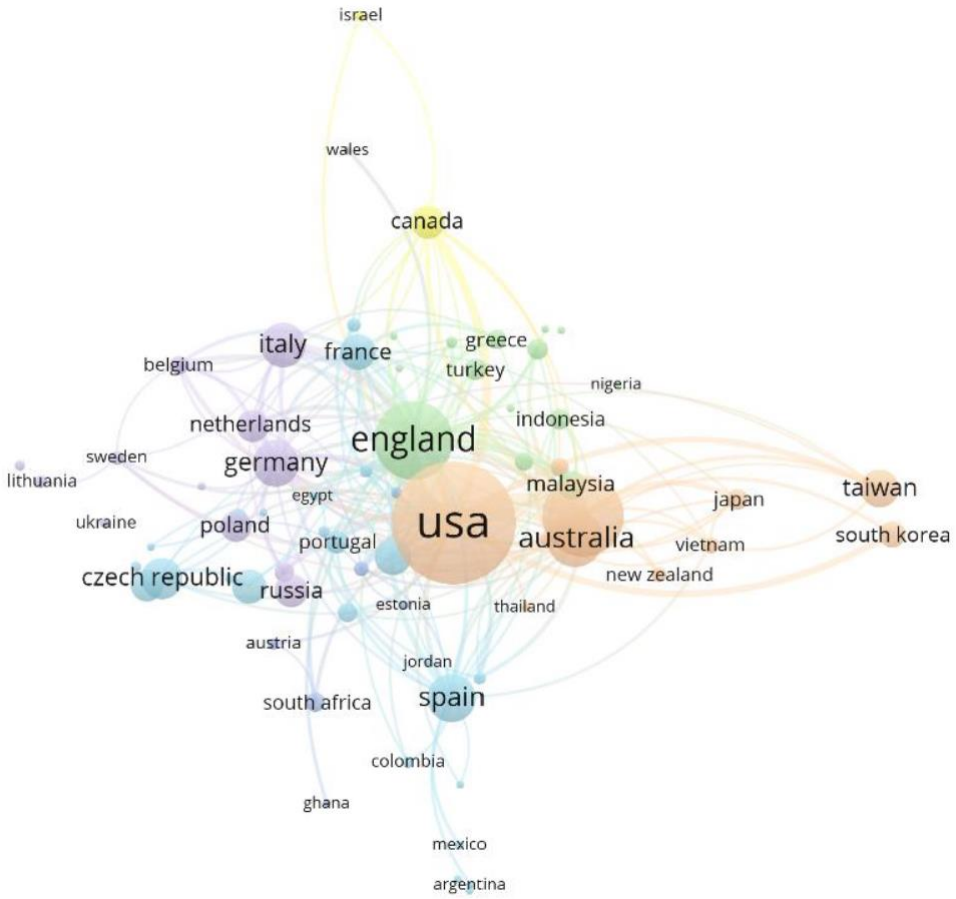
[78] The frequent use of specific keywords, combined with clustering, illustrates the concentration of scholarly attention in certain areas while also exposing the diverse approaches to addressing similar problems across different countries and disciplines. This kind of analysis, by bridging various thematic clusters, reinforces the importance of a collaborative and interdisciplinary approach to understanding complex global issues.

[79] Additionally, the division of research into these clusters highlights regional and thematic gaps, particularly in the application of modern techniques to smaller and medium-sized enterprises. While much of the research has historically concentrated on large corporations in developed economies, there remains a significant gap in the application of these predictive models in emerging markets and SMEs, which are often subject to different financial dynamics and risks (Aliaj & Piersanti, 2020).

Commenting the bibliometric network visualization, is it possible to affirm that the largest and most central cluster, highlighted in green, includes core topics such as *bankruptcy prediction*, *discriminant analysis*, and *algorithm*. These keywords point to the prominence of quantitative models and statistical techniques that are foundational to predicting corporate insolvency. Research in this domain has traditionally relied on financial ratios and discriminant analysis, as seen in seminal models like Altman's Z-score (1968) and Ohlson's (1980) logit model, which are both widely referenced for bankruptcy forecasting (Ohlson, 1980; Altman, 1968). In recent years, the application new techniques such as support vector machines, genetic programming, and neural networks has grown significantly, marking a shift towards more sophisticated, adaptive models capable of handling larger datasets and more complex interactions (Zhang et al., 2021; Barboza et al., 2017). The purple and blue clusters emphasizes the role of governance and restructuring in the context of financial distress, with keywords such as corporate governance, creditors, liquidation, and earnings management. Corporate governance, in particular, has become an essential variable in bankruptcy models, with studies highlighting the influence of board composition, CEO stability, and creditor relationships on firms' ability to avoid financial collapse (Tron et al., 2022). These findings are particularly relevant for family-owned businesses and SMEs, where governance structures are often less formalized and yet play a crucial role in crisis management and recovery (lo Conte & Sancetta, 2024).

The methodological evolution in bankruptcy prediction is well represented by the green cluster, which includes terms such as *data mining*, *probit, algorithm, decision-making, tree*, *tool*. This cluster reflects the adoption of ML and AI-based methods in recent literature, which have proven to significantly enhance prediction accuracy by accounting for non-linear patterns and interactions within large datasets (Barboza et al., 2017). Lastly, the study reveals that the United States leads the field in terms of the number of published studies on corporate bankruptcy prediction, as shown in Figure 9, followed by England. This dominance underscores the significant influence of Anglophone research in driving forward innovative methodologies in financial distress prediction. The leadership of the U.S. in this area is not surprising, as it aligns with the foundational contributions of American pioneers such as FitzPatrick (1932), Altman (1968), and Ohlson (1980), who have laid the groundwork for much of the predictive modeling still in use today. These early studies have established the United States as the bedrock of corporate insolvency research, setting a standard that other countries have since followed. England's strong position in this domain reflects the country's robust financial system and its long-standing role as a hub for both academic and practical financial research. The UK's academic institutions are well-known for their contributions to financial risk management and predictive modeling, often focusing on corporate governance and regulatory aspects in addition to traditional financial metrics. Germany and Italy rank after in published research, benefiting from a stable and sophisticated economic environment and a governance framework that prioritizes financial transparency and corporate stability. German scholars, in fact, have contributed significantly to the development of stakeholder-oriented approaches in crisis prediction, often incorporating long-term stability and industrial relations into their predictive models. In third position, we find Asian-based publications.

Figure 9: Network Visualization of Countries



Source: Gajdosikova & Gabrikova, 2023

This geographical distribution of research highlights how Anglophone countries remain at the forefront of bankruptcy prediction studies, particularly in frontier topics such as ML and AI-driven predictive models. The U.S. and UK, for instance, are spearheading advancements in the use of neural networks, random forests, and support vector machines, reflecting a broader trend toward integrating cutting-edge technologies into corporate finance (Barboza et al., 2017). On the other hand, countries like Italy, while catching up, are more focused on the legal and regulatory aspects of insolvency, as reflected in the changing legal frameworks such as the Codice della Crisi d'Impresa. These evolving laws are driving a new wave of research in Italy that seeks to incorporate both corporate governance and financial ratios into predictive models. Italian scholars, such as Guatri and Danovi, have contributed to this body of knowledge, yet the empirical studies often focus more on SMEs and family-owned businesses, leaving room for further exploration of large-scale predictive analytics using AI. In conclusion, the dominance of U.S. and UK research in predictive studies is a continuation of the historical trajectory set by early American pioneers. However, with the rapid advancements in AI, and the evolving regulatory landscapes in Europe, countries like Germany and Italy are contributing new perspectives that blend traditional financial analysis with modern governance frameworks, shaping a more global and interconnected research landscape.

3. Traditional Models, an Overview: the Altman Z-Score

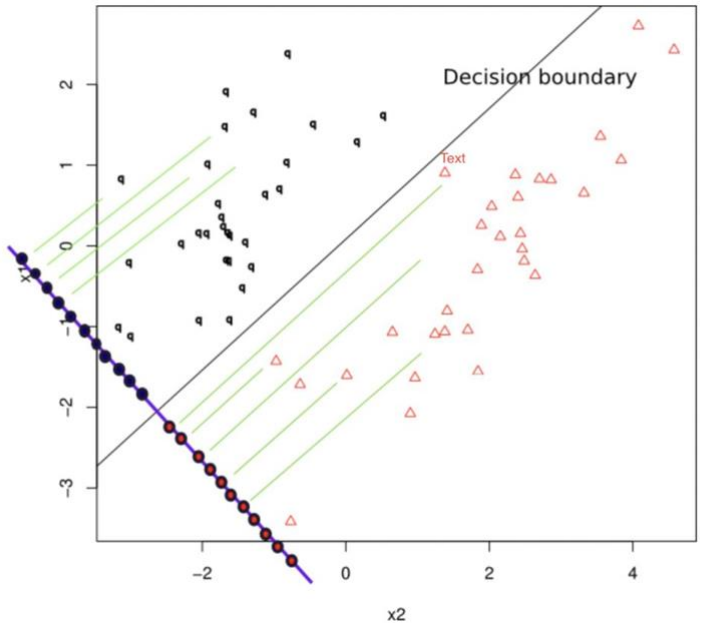
Credit scoring systems for assessing the likelihood of a firm's repayment capacity have evolved over centuries. Their origins can be traced back to early forms of risk assessment during the medieval period, when creditworthiness was based primarily on subjective judgments about borrowers. In the 19th century, particularly in the United States during the westward expansion, financial institutions began formalizing credit evaluation methods. These early systems were based on qualitative measures, such as ownership, management, and collateral, but lacked the data-driven sophistication of modern approaches. A key milestone in this evolution came with the introduction of Altman's Z-score in 1968, which transformed corporate bankruptcy prediction[80]. Altman's model was based on a dataset of 66 manufacturing firms—both bankrupt and solvent—and utilized five financial ratios to assess bankruptcy risk. The Z-score represented a shift toward multivariate analysis, where multiple financial indicators were combined to provide a more comprehensive and accurate measure of a company's financial distress.

In developing his predictive model for corporate bankruptcy, Altman utilized Linear Discriminant Analysis (LDA), a statistical technique initially introduced by Ronald Fisher in 1936 (Figure 10)[81]. Fisher's method, originally applied in biological and behavioral sciences, aimed to classify objects into predefined categories or populations based on specific characteristics (Fisher, 1959). Although not as prominent as linear regression in financial applications at the time, Altman recognized the versatility of LDA for classifying companies into distinct financial groups—those likely to remain solvent and those at risk of bankruptcy.

[80] In this dissertation, Altman's Z-score model is identified as a cornerstone in the field of corporate bankruptcy prediction. Understanding this seminal model is crucial for comprehending the evolution of newer, more intelligent predictive systems. Altman's contributions to the topic, particularly his application of Linear Discriminant Analysis (LDA), marked a significant leap forward in the literature and the practical application of financial distress prediction. His innovative integration of multiple financial ratios into a single predictive function set the groundwork for many subsequent models that followed.
Given its foundational status, this dissertation takes an in-depth look at Altman's model—examining its technical, managerial, and financial assumptions. By thoroughly unpacking the mechanics behind the Z-score, readers will be equipped to appreciate how this model not only revolutionized corporate risk analysis but also paved the way for modern AI-driven and machine learning models that have since built upon its principles. Understanding Altman's approach is then essential before delving into these newer methodologies, as it bridges the gap between traditional financial analysis and today's cutting-edge predictive analytics. Thus, the examination of the Altman model in this dissertation serves as a necessary prelude to the exploration of more advanced and adaptive models later in the work.

[81] What sets Altman's approach apart was not merely the application of LDA but the conceptual leap in using a multivariate framework to address corporate distress. Unlike earlier methods, which relied on single financial ratios, Altman's use of LDA synthesized multiple financial metrics into one discriminant function, allowing for a holistic view of a company's financial health.

Figure 10: Linear Discriminant Analysis



Source: Fisher, 1959

Altman's use of LDA in the financial sector represented a significant leap in the methodology for credit risk analysis. The core principle of LDA is to create a criterion that minimizes the misclassification of companies by using multiple financial ratios, rather than relying on a single indicator. In doing so, Altman's model assessed a company's likelihood of financial distress by combining a set of weighted accounting ratios into a discriminant function. This function then produced a score—often referred to as the Z-score—which determined whether a firm belonged to the "healthy" group or the "distressed" group. The discriminant function can be expressed mathematically as:

$$S_j = v_1 X_{1j} + v_2 X_{2j} + \ldots + v_n X_{nj}$$

Where:

- Sj is the score for the j-th company;
- vi represents the coefficient or weight for the variable Xi;
- $X\{ij\}$ is the value of the i-th characteristic for the j-th company.

By combining variables such as liquidity, profitability, and leverage, each weighted appropriately, the model generates a comprehensive score[82]. This score reflects a firm's overall financial health and its likelihood of bankruptcy. So, the power of LDA lies in its ability to process multiple variables simultaneously, providing a more nuanced and reliable classification than univariate methods. Altman emphasized that the variables considered in the LDA-based model must be

[82] These variables reflect critical aspects of a company's financial health:

- Liquidity measures a firm's ability to meet short-term obligations;
- Profitability reflects the efficiency with which a company generates profit from its operations;
- Leverage assesses the degree of a company's financial risk by examining its debt levels relative to equity.

contextualized over time, rather than viewed in isolation. Each financial ratio is subject to fluctuations, and its relevance may change depending on market conditions, industry trends, or company-specific factors. Thus, the model captures the dynamic nature of corporate financial health, ensuring that predictions are not based on static or one-time assessments. Given all these considerations, Altman's pioneering use of LDA in finance demonstrated its effectiveness as a robust tool for bankruptcy prediction, setting the stage for further advancements in financial modeling and risk management. Today, the use of LDA, alongside other ML techniques, continues to evolve, highlighting the model's enduring relevance in the modern financial landscape.

3.1. Elements of Multivariate Discriminant Analysis

Altman's exploration of corporate financial distress prediction methodologies led to the consideration of four key statistical tools:

- linear probability models;
- logit[83];
- probit;
- multiple discriminant analysis (MDA) (Altman & Saunders, 1997).

After assessing the problem's nature and the purpose of his research, Altman opted for MDA due to its unique ability to handle classification problems where the dependent variable is qualitative—such as distinguishing between "*bankrupt*" and "*non-bankrupt*" firms (Altman, 1968). Multiple Discriminant Analysis (MDA), although less common than regression analysis, has a long history of use across disciplines, initially in the biological and behavioral sciences. In financial studies, MDA became increasingly valued for its ability to classify observations into predefined groups based on a combination of quantitative variables. The key advantage of MDA is that it allows for the simultaneous consideration of multiple variables, providing a more holistic and accurate classification of firms' financial health than univariate methods, which examine one variable at a time.

In Altman's case, the goal was to derive a linear combination of financial ratios that would best differentiate between two groups: bankrupt and non-bankrupt firms. MDA provides a discriminant function that, when applied to a company's financial data, generates a score used for classification. This approach reduces the dimensionality of the dataset from multiple variables to a single dimension in cases of two groups, allowing for more streamlined analysis while still accounting for complex interactions among variables. From the perspective of predictive analytics, MDA's ability to synthesize a broad range of financial indicators makes it a powerful tool in bankruptcy prediction. It supports data-driven decision-making by offering clear, quantifiable insights into a company's financial status, which can be crucial for managers, investors, and creditors. Furthermore, its emphasis on data-centric culture aligns well with modern financial analysis approaches, which prioritize the use of extensive datasets and multivariate methods to improve the accuracy and reliability of predictions. So, the application of MDA by Altman is a pivotal moment in the development of predictive models for financial distress, bridging traditional statistical approaches with the evolving data-rich environments.

[83] Logit analysis began to appear in the academic and practical scenario in the late 1970's. Its main feature, at the time, was that it considered the probability that the firm go bankrupt.

3.2. The Score

Based on Fisher's technique, Altman created a model that combined several financial indicators into a single, unified score. The formula for this Z-score is represented as:

$$Z = a_1x_1 + a_2x_2 + \ \ldots\ldots\ + a_nx_n$$

where the variables *xi* represent different financial indicators (such as liquidity, profitability, and leverage), and *ai* are the weights assigned to these variables based on their significance in predicting corporate distress. The model was designed to distinguish between healthy and distressed companies by identifying which financial characteristics were most predictive of bankruptcy. Altman's challenge was not only technical—finding the right combination of variables—but also practical. He aimed to develop a tool that was both scientifically robust and easy for practitioners to apply in real-world settings (Altman, 1968). This was a critical consideration, as at the time, the role of management in using such quantitative tools for decision-making was still evolving (lo Conte et al., 2024). Although the United States was at the forefront of these developments, the model was initially met with hesitation in practical business environments, as the concept of structured management and quantitative financial tools was not yet fully embraced by the broader managerial community.

In the decades following its introduction, Altman continued to refine his model and collaborate with scholars internationally. In fact, the Z-score model has been successfully adapted for various economic contexts and industries, such as in emerging markets (Altman et al., 1995) and in the evaluation of corporate bonds. In many cases, the model has shown strong predictive power, affirming its versatility and relevance (lo Conte et al., 2023). Altman's collaborations with global researchers further demonstrated the model's ability to function effectively across different markets, highlighting its adaptability to diverse financial structures and, mainly, regulatory environments (Altman & La Fleur, 1985). For the first model, the final equation of the score, displayed as:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5$$

is a weighted linear combination of financial ratios used to predict the likelihood of a firm facing bankruptcy within two years. The coefficients are the weights assigned to five financial variables. The selected variables and their respective coefficients are the following:

- *X1: Working Capital / Total Assets* - This ratio measures the firm's ability to finance its operations through short-term assets. The coefficient of 1.2 signifies its moderate importance in assessing liquidity and financial health;

- *X2: Retained Earnings / Total Assets* - Reflecting a firm's profitability and reinvested earnings, this variable has a coefficient of 1.4, emphasizing its impact on long-term survival and growth;

- *X3: Earnings Before Interest and Taxes (EBIT) / Total Assets* - With the highest coefficient (3.3), this ratio highlights a firm's ability to generate profits from its assets, which is crucial in assessing ongoing financial performance;

- *X4: Market Value of Equity / Book Value of Total Liabilities* - This leverage ratio, with a coefficient of 0.6, provides insight into the firm's solvency by comparing the market valuation to its liabilities;

- *X5: Sales / Total Assets* - At 0.999, this variable measures asset efficiency, showing how well a firm uses its assets to generate sales[84].

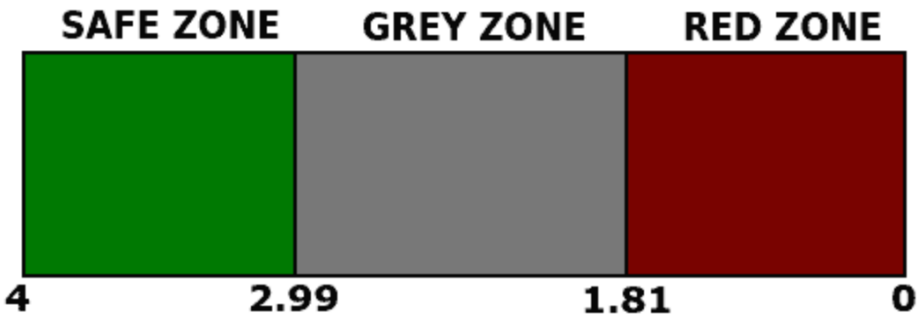
The Z-score calculation ultimately places a firm into one of three categories (Figure 11):

- *Z > 2.99:* Safe zone, where bankruptcy risk is considered low;

- *1.81 < Z < 2.99:* Gray zone, indicating a moderate risk of bankruptcy, suggesting the firm requires close monitoring;

- *Z < 1.81:* Distress zone, where the firm is at high risk of bankruptcy and likely to fail within the next two years.

Figure 11: Altman Z-Score Areas



Source: Altman, 1968 (design from ztable.net)

This model originally focused on public manufacturing firms but evolved over the years to accommodate different industries and company types, including private firms and non-manufacturing sectors, with adjustments made to the coefficients and variables to reflect the distinct financial characteristics of each sector (Altman et al., 1995; Altman, 1968). For example, Altman's adaptation for private companies features slight modifications in the weighting of variables like working capital, retained earnings, and sales, as these elements play different roles in privately held firms compared to public manufacturing companies. Similarly, the non-manufacturing and emerging markets models exclude the sales-to-assets ratio due to the variability in how assets are utilized across different industries (Altman et al., 1983). This adaptability and iterative refinement process illustrates Altman's awareness of sector-specific challenges and the evolving nature of financial distress modeling (Altman,

[84] In the majority of the studies, this value equals to 1 (Kuizinienė et al., 2022; Kim & Han, 2002).

2006; Altman, 1983). Furthermore, despite the simplicity of the three-zone categorization—safe, grey, and distress—the model remains relevant, though increasingly supplemented by newer techniques, which offer more granular insights and predictive power (Altman et al., 2017; Barboza et al., 2017). While Altman's ongoing refinements remain significant in financial and managerial literature, the primary interest here lies in how these foundational tools can be enhanced in today's data-driven era. Thus, the adaptations of the Z-Score, while relevant, take a secondary position to the exploration of modern, "*intelligent*" tools in the context of this dissertation.

3.3. Some Considerations

From a managerial perspective, the Z-score model was certainly transformative; as business practices evolved, particularly in the late 20th century, managers began to recognize the value of data-driven tools for decision-making in crisis situations. Predictive models like the Z-score offered a systematic way to assess a firm's financial health, providing insights that could inform restructuring efforts or pre-empt potential crises. Altman's work demonstrated how such models could serve as essential tools for proactive decision-making in the face of financial uncertainty. Furthermore, as the use of data and predictive analytics expanded, so too did the role of models like Altman's in supporting the development of a data-centric culture within businesses. By quantifying financial risk and enabling more informed decision-making, the Z-score laid the groundwork for modern predictive analytics tools that integrate advanced techniques. So, the evolution of Altman's model highlights the growing importance of blending technical sophistication with practical usability, ensuring that decision-makers have access to both the data and the tools necessary to navigate financial distress effectively.

Today, Altman's contributions remain foundational in the field of financial distress prediction. His collaborations with scholars and practitioners around the world have helped solidify the Z-score's relevance across various industries and markets. While newer models incorporating AI have emerged, Altman's work continues to inform the development of these tools, emphasizing the critical intersection of quantitative analysis, predictive accuracy, and managerial decision-making. In addition to his significant international collaborations, Altman has also established important connections in Italy, further demonstrating the global impact and adaptability of his bankruptcy prediction models. Over the years, Altman worked closely with Italian academics and professionals, sharing his expertise and contributing to the development of financial distress models tailored to the unique characteristics of the Italian market. Notably, Altman's participation in conferences and seminars at major Italian universities, as well as his collaborations with Italian scholars such as Danovi and Falini, have helped to contextualize his Z-score model within the framework of Italian corporate governance and financial systems.

In particular, Altman's work in Italy has contributed to a deeper understanding of how financial distress models can be applied in family-owned businesses, which are prevalent in the Italian corporate landscape. This collaboration has yielded valuable insights into how governance structures, like the board of directors and the renewal of top management, influence the financial health of firms. Such research has also provided practical guidance for Italian companies and financial institutions, allowing them to better anticipate insolvency and manage financial risks. Despite its importance, the Z-score model comes with limitations. Its primary weakness lies in its reliance on financial ratios from historical data, which might not always capture the complexity and dynamic nature of today's markets. The model assumes that financial distress can be predicted through static, linear relationships between variables, which can sometimes overlook nonlinear interactions or external shocks that affect firms'

financial health. Furthermore, the model's design focused predominantly on large corporations, which limits its applicability for SMEs, especially in Italy, where the majority of businesses are family-run and less structured than large corporations.

However, even today, Altman's Z-score remains a widely utilized tool in literature and practice, offering a quick and understandable metric for evaluating bankruptcy risk. Its simplicity allows managers, analysts, and investors to make informed decisions regarding a firm's financial trajectory. Yet, while the three distinct Z-score zones (distress, gray, and safe) provide general indicators of financial health, the static nature of the model highlights the need for complementary tools in today's volatile contexts. This is where AI and ML have stepped in, providing a more adaptable framework for decision-making, risk management, and crisis management. Unlike the Z-score, which offers a fixed score based on predefined variables, AI-driven models are able to continuously learn and adapt, incorporating real-time data and recognizing patterns beyond what traditional models can capture. The shift towards AI-enhanced predictive analytics then represents a fundamental shift in how financial distress is managed—focusing on proactivity, speed, and precision. In conclusion, this synergy between classic models and advanced technologies offers a future where decision-making is not just more informed, but more adaptive, ensuring that businesses can survive and thrive even in the most uncertain environments.

3.4. Other Traditional Models

Over the decades, several scholars have built upon Altman's framework, each introducing new nuances and variables to predict corporate distress across various industries and business environments. Among these is Springate's S-Score (1978), which closely follows Altman's methodology but adjusts the weights of financial ratios to suit the nuances of Canadian firms. Its distinction lies in the careful emphasis on profit before tax to current debt, offering an alternative lens to measure short-term financial solvency, especially for smaller, more volatile firms. Further, Zmijewski's X-Score (1984) took a different approach, using probit analysis instead of linear discriminant methods like Altman's. Zmijewski's focus on debt levels, with a simplified three-variable model, highlights the weight of leverage in predicting financial distress. Although Zmijewski's model remains cited in literature, its simplified form has been deemed less versatile, especially in more nuanced industries, leading to a decline in usage. Nevertheless, it remains a valuable tool in understanding leverage's critical role in corporate failure.

Similarly, Ohlson's O-Score (1980) applied logit regression to financial ratios, offering a statistical alternative that some argue is more suitable for categorical financial data, such as predicting whether a firm will fail or succeed[85]. Despite its mathematical robustness, Ohlson's model has seen limited adoption in practical settings, as its complexity and statistical orientation often overshadow the intuitive simplicity that Altman's model offers[86]. Moreover, Ohlson's model doesn't account for market sentiment—a critical aspect incorporated in Altman's approach via the market value of equity to total liabilities ratio[87]. Model's formula can be expressed as:

[85] While the Altman Z-score and Ohlson O-score share the common goal of predicting financial distress, they differ significantly in their methodologies and the variables they consider. Financial analysts often use both models in conjunction to gain a more comprehensive understanding of a company's solvency. By acknowledging the strengths and limitations of each model, analysts can make more informed decisions and provide better insights into the financial stability of the entities they evaluate.

[86] This model has contributed to shaping Italy's approach to financial risk, although its adoption remains limited in academic and managerial circles compared to more popular models like Altman's.

[87] If Altman (1968) and Beaver (1966) used data from Moody's Manual, then Ohlson (1980) got the data from the financial statement issued for taxes (10K-FinancialStatement).

$$T = -1.32 - 0.407\log(TA_t/GNP) + 6.03\frac{TL_t}{TA_t} - 1.43\frac{WC_t}{TA_t} + 0.0757\frac{CL_t}{CA_t}$$

$$- 1.72X - 2.37\frac{NI_t}{TA_t} - 1.83\frac{FFO_t}{TL_t} + 0.285Y - 0.521\frac{NI_t - NI_{t-1}}{|NI_t| + |NI_{t-1}|}$$

This specific equation incorporates multiple financial ratios and factors, including total assets (TA), gross national product (GNP), total liabilities (TL), working capital (WC), and several other variables like net income (NI) and cash flow (FFO). Each term in the formula carries a specific weight (or coefficient), and the overall purpose is to calculate a score (T; probability of default) that helps assess the likelihood of corporate financial distress or bankruptcy. Although Altman's Z-Score model remains widely used, Ohlson's logit regression model has demonstrated, in various contexts, unexpectedly better performance in predicting financial distress, including in manufacturing sectors. Numerous studies have shown that, even though the Ohlson model is applied less frequently in both Italy and internationally, its predictive accuracy can sometimes surpass that of the Z-Score, particularly when dealing with certain types of firms or datasets. For instance, a comparative study of Altman's Z-Score and Ohlson's logit model on Indonesian manufacturing firms revealed that Ohlson's model exhibited stronger performance in some financial environments (Imelda & Alodia, 2016) (Table 5). This suggests that Ohlson's model, which incorporates a more comprehensive range of financial indicators (such as firm size and leverage) and adjusts for non-linear relationships through logit regression, may offer advantages in specific scenarios where Altman's model, originally designed for manufacturing firms, might fall short.

Table 5: Altman's & Ohlson's Models Comparison (Accuracy Performance)

Year	Accuracy	
	Altman	Ohlson
1 Year Before	63%	65%
2 Years Before	60%	78%
3 Years Before	60%	75%
Average	61%	73%

Source: Imelda & Alodia, 2016

Another important contribution comes from Grover (2001), whose G-Score built on Altman's Z-Score but included 13 additional financial ratios, offering a broader, more detailed analysis of liquidity and profitability[88]. The Taffler model (1983), on the other hand, emphasizes pre-tax earnings in relation to current obligations, providing an impressive 96% accuracy in predicting bankruptcy in UK-listed companies. This precision, however, comes at the expense of a more general applicability, limiting its widespread use across other industries and regions.

[88] This study is an example of the intention, of all academic world, to add more and more financial and economic elements to predictive analytics and models in order to have better results. Since Altman's work, academia was interested in enhance relationship with industries in order to get particular hints about variables, elements and ratios to use in predictive models (lo Conte & Sancetta, 2024).

The continual refinement and development of these models have been largely iterative, responding to industry-specific needs and the increasing complexity of global markets. While these models have contributed to the academic literature on financial distress prediction, Altman's Z-Score remains the most extensively validated and widely used tool, especially in academia and practice. In fact, Altman's iterative revisions to the original model—incorporating non-manufacturing firms and emerging markets—reflect the evolving nature of financial modeling. However, in more recent years, the limitations of these models have become apparent, opening the door to new and innovative techniques. Given this shift, the next pages of this dissertation will focus into the methodological, technical, and managerial implications of new predictive models, emphasizing how traditional frameworks can be enhanced with AI tools to improve accuracy and responsiveness in predictive analytics.

4. Innovative and "*Intelligent*" Predictive Analytics

As we know, the rapid advancements in AI and ML have totally altered the way businesses approach financial distress prediction and crisis management[89]. However, one of the central findings in the literature is that while AI models outperform traditional methods, they also present a unique challenge: they often function as "*black boxes*," making it difficult for managers—especially those unfamiliar with AI processes—to understand how decisions are made (lo Conte et al., 2023; lo Conte & Sancetta, 2024; Duan et al., 2019). This is where Explainable AI (XAI), particularly tools like SHAP (SHapley Additive exPlanations), plays a crucial role[90]. By offering clear insights into how AI models arrive at their predictions, SHAP, in fact, allows managers to engage with AI outputs in a more informed way, reducing resistance to AI adoption and ensuring that AI serves as a support tool rather than a replacement for human judgment (Thiebes et al., 2020; Duan et al., 2019). Barboza et al. (2017) argue that ML models, particularly neural networks, can uncover better non-linear patterns in financial data; this ability to manage complex relationships among variables is echoed by Zhang et al. (2016), who highlight the performance of random forests in their ability to avoid overfitting and manage feature selection automatically. Both models, according to these authors, consistently outperform classic techniques like discriminant analysis and logistic regression when tested across firms of varying sizes.

In terms of financial sustainability, AI has significant potential to contribute to proactive risk management[91]. Anderson and Zhang (2021) discuss how the integration of ML models into insolvency prediction frameworks allows for early identification of distress, which in turn enables firms to take preventative actions; this approach certainly contrasts sharply with traditional models, which typically react to historical financial data rather than anticipate future crises (Brynjolfsson & McAfee, 2017). In large firms, where complex financial structures and vast datasets need to be analyzed in real-time, the speed and scalability of AI models are particularly beneficial (O'Leary, 2019). However, the literature also emphasizes the importance of retaining human oversight in AI-driven decision-making. While AI offers unprecedented computational power, it cannot replace the strategic judgment and ethical considerations that managers bring to crisis management. In this regard, Duan et al. (2019) highlight that, despite AI's predictive accuracy, human intuition is still needed to interpret the broader business

[89] Moreover, interpretability tools like SHAP or LIME (Local Interpretable Model-Agnostic Explanations) have become essential for data analysts. These methods help explain how and why an AI model made a particular prediction, ensuring that decision-makers trust the AI outputs.

[90] SHapley Additive exPlanations, more commonly known as SHAP, is used to explain the output of Machine Learning models. It is based on Shapley values, which use game theory to assign credit for a model's prediction to each feature or feature value.

[91] See lo Conte et al., 2023.

context, especially in crisis situations where data alone may not capture external market forces or cultural dynamics within the firm. This sentiment is echoed by Thiebes et al. (2020), who argue that explainable AI helps balance the strengths of Machine Learning with human accountability[92].

Moreover, as businesses increasingly operate in volatile markets, the real-time adaptability of AI models allows managers to continuously refine their strategies based on dynamic, up-to-date data. This capability is particularly beneficial in high-risk sectors or SMEs (Oluwaseun et al., 2023). In larger firms, AI's ability to process and analyze vast datasets provides enhanced granularity in assessing financial health, thereby supporting more sophisticated risk management strategies (O'Leary, 2019). So, the empirical portion of this research further explores and demonstrates how these predictive technologies can be tailored to meet the specific challenges faced by businesses of different sizes and industries. This includes not only assessing model performance across various firms but also investigating the specific managerial applications of these tools in crisis prediction. Through a deep dive into these frameworks, in fact, the following sections will address how AI technologies can be operationalized to meet real-world business needs, focusing on improving financial forecasting and enabling strategic interventions that prioritize long-term stability. Furthermore, this dissertation will introduce and explore technical, scientific, and literary concepts surrounding cutting-edge, data-driven methods for predictive analysis in scenarios of corporate distress. These approaches certainly represent a marked shift from traditional models, and several of them will be implemented in the empirical section of this PhD dissertation.

4.1. An Emerging Paradigm: Intelligent Big Data Analytics (IBDA)

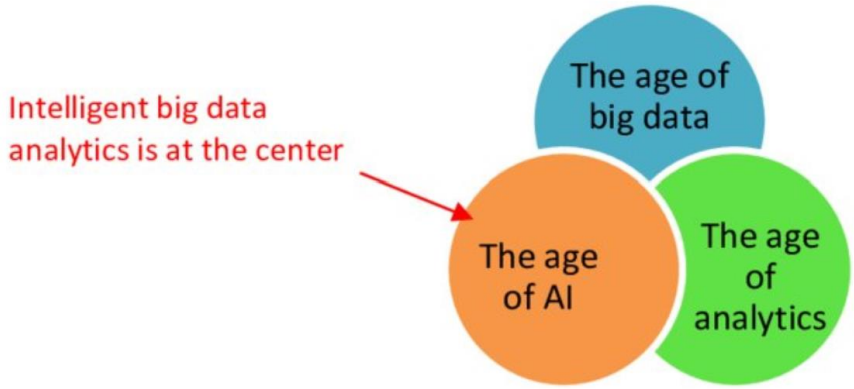
As already stated in this PhD Thesis, the field of predictive analytics, particularly concerning corporate crisis prediction, has witnessed significant evolution over the last decade (lo Conte & Sancetta, 2024). Researchers and practitioners alike are continuously seeking robust frameworks to anchor their methodologies, models, and implications in this rapidly changing environment, as predictive analytics frameworks have historically been fragmented (O'Leary, 2019). However, as the volume, variety, and velocity of data have exploded in recent years, a new paradigm has emerged: Intelligent Big Data Analytics (IBDA) (Sun, 2019). This framework, which integrates the fields of data science, ML, and business intelligence (BI), is rapidly gaining traction as the leading approach to addressing the complexities of predicting corporate crises and other critical business events (Sun, 2019). This, in particular, is distinct because it does not simply rely on large datasets; instead, it leverages advanced AI algorithms to analyze structured and unstructured data, providing real-time insights that enhance decision-making at both operational and strategic levels (managerial perspective, in fact, is always lightened in these studies) (Sun, 2019). As industries move towards data-centric decision-making, the adoption of IBDA tools is then becoming critical today. As stated by Sun, 2019, IBDA can be considered a paradigm created in an "*age of trinity"*, as it considers three different *"ages":*

- the age of Analytics;
- the age of Big Data;

[92] Human Accountability refers to the principle that human beings, rather than machines or algorithms, should ultimately be responsible for decisions and outcomes. Legal frameworks such as the EU's AI Act highlight the importance of human accountability by emphasizing that AI systems, especially those considered "*high-risk,"* should be designed to allow human intervention and oversight, ensuring that humans remain "*in the loop*."

- the age of AI (Figure 12)[93].

Figure 12: The "*Age of Trinity*"



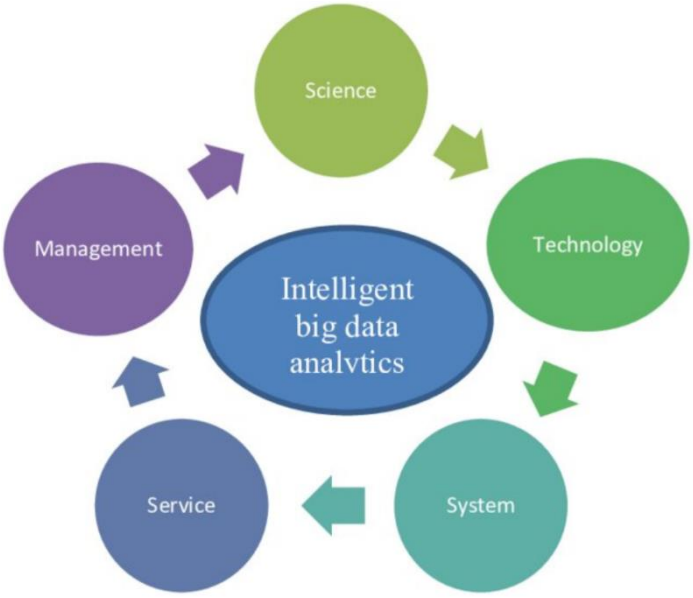
Source: Sun, 2019

Given this, both academia and industry practitioners are on a quest to define and refine analytical frameworks capable of integrating vast datasets and sophisticated analytics, still remaining simple, effective and explainable. But, what makes a good predictive framework today? First, it must support scalability—able to adapt from SMEs to large conglomerates (Sun, 2019). Second, it must integrate with various data sources, enabling businesses to take a holistic view of potential risks. Third, it must be adaptable, continually improving as more data becomes available, enhancing its predictive accuracy. Traditional models fail in several of these aspects, leading the academic and business community towards frameworks that leverage the already mentioned intelligent data analytics. In this context, IBDA represents the synthesis of technical data analytics and managerial decision-making. Unlike purely technical data frameworks, in fact, it includes managerial insights at its core, offering solutions that cater to the specific needs of business leaders. This dual focus is what distinguishes IBDA from other data-driven frameworks—it doesn't stop at delivering raw analytics but contextualizes it into actionable business intelligence.

Furthermore, as highlighted in recent literature, IBDA relies on key pillars such as data integration, advanced ML algorithms, and real-time analytics (Nguyen et al., 2023). The goal is to transform big data into business intelligence through systems that can support managers in making informed, data-driven decisions during times of crisis (Marjani et al., 2017). The framework also incorporates feedback mechanisms that allow businesses to continuously improve their strategies and operational plans, creating an ongoing cycle of learning and improvement (Sun, 2019). See Figure 13 for a contextual graph.

[93] The most recent one.

Figure 13: Managerial Framework for IBDA



Source: Sun, 2019

The managerial perspective embedded is then crucial because it recognizes that AI and data analytics are not standalone decision-makers. Instead, they serve as tools to augment human decision-making processes, offering managers insights that would otherwise be impossible to detect[94]. In this sense, IBDA becomes a decision-support tool, rather than an autonomous system, bridging the gap between data science and business operations, and, in the context of this Thesis, between tech-driven decisions and human ones. For instance, studies show that businesses utilizing IBDA are better equipped to manage financial crises, as predictive algorithms allow them to anticipate potential risks and develop strategies to mitigate those risks before they become unmanageable (Sivarajah et al., 2017). Given these considerations, it is possible to affirm that IBDA full encapsulates this dissertation aims.

4.2. Business Intelligence and Managerial Literacy in Data-Driven Frameworks

The role of business intelligence (BI) in IBDA cannot be understated. BI, as a subset of IBDA, refers to the tools and systems that play a pivotal role in the strategic planning process of corporations (Sun, 2019). These systems allow managers to gather, store, and analyze data to create actionable insights that improve decision-making. BI traditionally involved structured data sources, but with the rise of big data, BI tools have evolved to handle unstructured data, including text, images, and social media interactions. One of the main challenges facing managers today is their literacy in data analytics. The shift towards data-centric organizations requires managers to be well-versed in interpreting data outputs and understanding the predictive models that inform their decisions. Many traditional managers struggle with this transition, often relying on intuition rather than data-driven insights. IBDA offers, on the other side, a potential remedy for this by incorporating user-friendly interfaces and

[94] This reconnect with the aims and assumptions of this dissertation.

visualization tools that make data accessible, even for those without technical expertise[95]. However, there is still a gap in terms of training and knowledge dissemination, as many managers feel overwhelmed by the sheer volume of data and the complexity of predictive models (Chatterjee & Kar, 2024). As a result, managerial literacy in data science is becoming a key factor in the successful implementation of IBDA frameworks: the more managers understand how to use these tools, the more likely they are to trust and adopt them in their decision-making processes. In conclusion, AI and ML can be viewed as a support tool, enhancing decision making in critical cases, and not replacing humans, management and social feels.

4.3. IBDA in Insolvency and Crisis Prediction

The application of IBDA, and mainly its implications, in crisis prediction and distress management is one of the most promising areas of research today. With its ability to process vast datasets and generate insights in real-time, IBDA-based systems can forecast potential financial distress long before traditional methods, allowing businesses to take preemptive measures. For example, in the context of Italian firms, where financial distress is often tied to economic volatility and complex regulations, IBDA offers a unique advantage in identifying distress signals earlier and more accurately than traditional models. A key area where IBDA outperforms older models is its ability to integrate unstructured data such as news articles and macroeconomic indicators alongside traditional financial data. This holistic approach provides a fuller picture of a company's health and its external environment, leading to more accurate and timely predictions (Mariani & Jacob, 2019). Moreover, ML algorithms within the IBDA framework can self-adjust based on new data inputs, making the model more dynamic and flexible in comparison to static models like Altman's Z-score or Ohlson's logit. These advancements enable businesses to pivot strategies in real-time, offering a more proactive stance in crisis management. Machine Learning's ability to detect non-linear patterns within large datasets makes it particularly effective in complex, high-stakes environments.

In the literature, the Reference Model for Business Intelligence to Predict Bankruptcy (RMBIPB) by Aruldoss et al. (2015) provides an essential framework for understanding how BI can be used to predict corporate distress. This model incorporates a hierarchical structure that uses unit operations to select the most relevant techniques for bankruptcy prediction with minimal effort. By designing a flexible model, RMBIPB can adapt to various industries and economic environments, enabling it to serve as a comprehensive solution for corporate distress prediction. This model not only forecasts financial failure but also allows companies to implement a structured data-driven approach to decision-making, offering a way to align predictive analytics with real-world managerial needs. Aruldoss et al. (2011) further highlight the need for business intelligence systems that can adapt to these complexities, which is where models like the Hybrid Deep Learning Model proposed by Gaurav et al. (2024) come into play. The use of convolutional neural networks (CNN) and synthetic minority over-sampling techniques (SMOTE) in this model addresses issues related to high-dimensional data and imbalanced datasets, providing a predictive accuracy rate of over 81%—a notable improvement over traditional methods such as logistic regression and support vector machines (SVMs).

However, the literature also explores the need to go beyond purely technical models and address the managerial implications of these predictive tools. The Martin et al. (2012) model, for instance, stresses the importance of integrating qualitative factors—such as management decisions and

[95] One example is the already mentioned SHAP tool.

organizational behavior—into predictive frameworks, combining both quantitative financial data and qualitative non-financial data. The growing demand for real-time data integration and actionable insights makes business intelligence a critical tool in the hands of managers. These systems allow for more informed decision-making and proactive crisis management, helping firms navigate the complex terrain of financial distress and insolvency with more precision. On the other side, in Italy, where family-owned businesses and small to medium-sized enterprises dominate the economic landscape, the adoption of such intelligent systems remains a challenge. The study by Bragoli et al. (2019), which applied a full battery of bankruptcy forecasting models to Italian firms, demonstrates that ML models, particularly XGBoost, outperform traditional models in predicting firm distress. The inclusion of industrial and regional variables in the Italian context further reinforces the importance of adapting predictive models to specific economic environments. However, the study also reveals a significant gap in the application of business intelligence tools in smaller firms, particularly in terms of managerial literacy in data science and predictive analytics.

This strand of the research, however, is no free from challenges. While BI systems and AI-driven models provide more accurate predictions, one of the major challenges faced by traditional managers is, in fact, understanding and interpreting the results of these models. As also highlighted by Aruldoss et al. (2015), many predictive tools operate as "*black boxes*," offering results without clear explanations of how those results were derived (Chatterjee & Kar, 2024). This issue has led to some reluctance among traditional managers, especially in SMEs, to adopt these technologies. There is then a need for more transparent models that can explain their predictions in a way that is understandable to non-technical stakeholders. This is where models like SHAP come into play. In conclusion, the literature indicates that the field of business intelligence and predictive analytics is evolving towards more adaptive, data-driven frameworks that combine technical and managerial insights. Models such as RMBIPB and hybrid deep learning systems demonstrate the potential of integrating structured and unstructured data for more accurate predictions. However, there remains a gap in managerial understanding of these tools, particularly in the Italian context.

4.4. Machine Learning Models

As discussed in the previous Chapters, our focus on ML stems from its direct applicability to bankruptcy and crisis prediction, making it more aligned with our research goals compared to broader AI or generative AI applications[96]. Specifically, ML provides practical tools that, while rooted in mathematical rigor, have profound managerial implications, aiding decision-making processes in times of financial distress (Chatterjee & Kar, 2024). This part regards ML algorithms, some relevant & some not to the current analysis, exploring their underlying mechanics and their potential. We will then empirically evaluate these algorithms later in this dissertation, measuring their performance against a robust dataset to yield insights that may inform practical crisis management strategies. Unlike more general AI systems that might simulate human thought processes, ML models are often designed with specific problem-solving in mind—hence their utility in bankruptcy prediction, where they can sift through financial data to highlight key indicators of distress (Chatterjee & Kar, 2024). At the heart of ML is a process-oriented approach that involves multiple critical stages, which align with the steps often followed in traditional financial analysis, albeit with more sophisticated, data-driven methodologies:

[96] As we are talking about a PhD Thesis in management, information and data in this part regards, even technical, mostly managerial notions.

1. *Data Collection:* involves gathering relevant data reflective of the problem domain. For our purposes, this includes extensive financial metrics and other company-specific information;

2. *Data Preprocessing:* cleans and prepares the data for analysis, ensuring the removal of irrelevant or redundant information. Tasks here might include normalization and handling of missing & null values;

3. *Feature Extraction and Selection:* identifying the most relevant attributes from the dataset that will be used to train the model is fundamental. This ensures that the ML algorithm focuses on the most impactful variables;

4. *Model Training:* using the prepared data to train the model by minimizing the discrepancies between predicted and actual outcomes. In this case, training involves tuning the model to distinguish between financially healthy and distressed firms;

5. *Model Evaluation:* testing the trained model on new, unseen data (the test set) to assess its accuracy and generalizability. This stage is essential to ensure that the model can perform well in real-world scenarios;

6. *Model Deployment and Prediction:* finally, the trained model is deployed to make real-time predictions, providing decision-makers with actionable insights into potential insolvency risks.

Bankruptcy, crisis and insolvency prediction typically falls under supervised learning, a category of ML where models are trained on labeled datasets to predict specific outcomes[97]. More specifically, it is a binary classification task *(0=healthy;1=in crisis/insolvent)*, aiming to classify a company as either likely to go bankrupt or not. The algorithms used in this domain, known as classifiers, are mathematical functions trained to sort input data into pre-defined categories. Several algorithms have proven effective for these tasks, each offering unique strengths and trade-offs in terms of accuracy, interpretability, and computational complexity:

1. *Logistic Regression:* a foundational statistical method for binary classification problems, logistic regression applies a logistic function to model the probability of a company entering bankruptcy. While relatively simple, it offers transparency, making it a popular choice in managerial contexts where interpretability is crucial;

2. *Decision Trees:* these models segment data into branches based on specific features, making decisions that lead to classifications. Decision trees are widely appreciated for their simplicity

[97] In supervised learning, the primary objective is to create a model capable of accurately predicting an outcome based on input data. The model learns by identifying a function that maps inputs to known outputs in a way that minimizes error. For example, in image classification, the model would learn to associate pixel patterns with specific categories. Supervised learning frameworks focus on minimizing the discrepancy between actual and predicted outputs using a loss function, a measure of prediction error. This process allows the model to iteratively adjust its parameters until it reliably generalizes to new, unseen data. Training in supervised learning involves adjusting model parameters to minimize prediction error. A typical process includes initializing parameters, passing data through the model to generate predictions, calculating the error, and using backpropagation to update parameters. Optimization techniques like stochastic gradient descent (SGD) adjust parameters based on a random subset of data, accelerating convergence, while adaptive methods such as Adam dynamically modify the learning rate, which often improves training efficiency.

and visual representation, which aligns well with managerial decision-making, though they can sometimes lack precision in handling complex datasets;

3. *Random Forest:* building on decision trees, random forests aggregate multiple trees to improve prediction accuracy. By reducing the risk of overfitting, they are well-suited for bankruptcy prediction, offering both robustness and scalability;

4. *Support Vector Machines (SVM):* SVMs classify data by finding the optimal hyperplane that separates different classes. Their flexibility and high accuracy make them effective for complex datasets, though they require careful parameter tuning and are less interpretable than simpler models;

5. *Neural Networks:* these are advanced models inspired by the human brain, particularly effective in capturing complex relationships in large datasets. Neural networks are increasingly used in financial distress prediction, especially in conjunction with deep learning techniques. Their accuracy is unparalleled, but their "*black-box*" nature poses challenges for interpretability, often making them less suitable for decision-making in managerial contexts where transparency is required;

6. *XGBoost:* a popular algorithm in recent studies, XGBoost excels in structured datasets like financial records. It leverages gradient boosting to optimize model performance, making it particularly effective for tasks like bankruptcy prediction. The algorithm's flexibility and power have made it a frontrunner in many empirical applications.

4.4.1. Shrinkage Methods: an Hint

Predictive models are typically built on a range of distress predictors, predominantly derived from accounting variables and financial ratios. The selection of key variables is then essential in order to identify the most impactful predictors and enhance the accuracy of predictions. This process of variable selection is often paired with shrinkage techniques, which reduce the variance of the model by limiting the influence of less important variables (James et al., 2017) Shrinkage, or regularization, works by fitting a model that includes all the potential predictors and then adjusting the estimated coefficients to bring them closer to zero. This is done relative to the ordinary least squares (OLS) estimates, which are known to sometimes overfit models by placing too much emphasis on insignificant predictors. By shrinking some coefficients towards zero, regularization provides a form of implicit variable selection—essentially ignoring certain variables if their coefficients are sufficiently reduced. This helps streamline the model, making it more robust and generalizable. There are two primary shrinkage methods commonly used for this purpose: ridge regression and lasso (Least Absolute Shrinkage and Selection Operator). Ridge regression minimizes the residual sum of squares (RSS) in a similar fashion to OLS but adds a penalty proportional to the square of the coefficients. This penalization reduces the size of the coefficients without driving any of them entirely to zero. As a result, ridge regression retains all the predictors but shrinks their influence. In contrast, lasso regression not only penalizes the coefficients but can also shrink some of them to exactly zero, effectively removing less important predictors from the model. This characteristic makes lasso a more aggressive form of variable selection, which is particularly useful in identifying the most influential predictors in high-dimensional datasets. By

employing these regularization techniques, models become more robust and less prone to overfitting, ensuring that only the most relevant variables are retained for accurate bankruptcy prediction. Both ridge regression and lasso offer valuable tools in developing models that are data-driven, adaptable, and capable of improving prediction outcomes in financial distress scenarios.

However, these methods are still constrained by their reliance on a linear framework. In contrast, ML approaches such as random forests, gradient boosting machines, and neural networks allow for more flexible models that can adapt to non-linear data patterns without sacrificing accuracy. And they extend beyond classical linearity. In fact, ML algorithms such as decision trees and ensemble methods allow for intelligent analysis of complex financial data while maintaining interpretability, which is crucial for managerial decision-making (Brynjolfsson & McAfee, 2017). In this dissertation context, as already stated, the goal is to test these models' performance, identifying which is best suited for prediction in the Italian context. Then, there is also the aim to democratize ML tools for businesses of all sizes, from large corporations to SMEs, by emphasizing the need for managerial literacy in data science. This paradigm shift is not just about implementing new technology but about empowering managers to make informed, data-driven decisions that align with broader organizational goals, including sustainability and resilience (Lee et al., 2020). By combining the insights from business intelligence (Sharda et al., 2020) with ML models, managers develop more agile, intelligent strategies that can proactively address financial distress, ultimately defining a sustainable approach to crisis management (Vermeulen & Witjes, 2016).

4.4.2. Decision Trees

Beyond the typical linearity models, the tree concept is one the most prominent in literature on predictive analytics. In fact, decision trees are a pivotal tool in ML, frequently applied in both classification and regression problems. Their structure, which resembles a tree, is not only visually intuitive but also highly flexible, making them an effective method for solving a variety of prediction tasks. In a classification setting, decision trees help in assigning observations into predefined categories, whereas in regression tasks, they predict continuous values. This versatility is what makes decision trees applicable to diverse fields such as medical diagnostics, financial predictions, and bankruptcy analysis, among many others. The construction of a decision tree involves a recursive process known as recursive partitioning. At the start, the entire dataset is considered the root, and through successive splits, the data is partitioned into more homogeneous groups based on feature values. Each split is determined by evaluating which feature offers the most significant improvement in predicting the target variable—either a class label or a continuous value. These splits create branches, leading to terminal nodes or leaves, which represent the final decision points or predictions. This structure allows decision trees to develop clear "*if-then*" rules that guide decision-making, improving both interpretability and transparency. For instance, in a classification task, a path from the root to a leaf can be understood as a series of conditional statements that lead to the final classification of an instance.

There are two primary types of decision trees: regression trees and classification trees. A regression tree is used when the target variable is continuous, while a classification tree deals with discrete outcomes. The mathematical operation underlying both types is based on dividing the predictor space into a finite number of regions.

1. Regression trees:

- The predictor space, defined by variables X_1, X_2, ..., X_p, is partitioned into non-overlapping regions R_1, R_2, ..., Rj;
- Each region Rj corresponds to a predicted value, which is typically the mean of the response values in that region for the training data;
- The goal is to minimize the Residual Sum of Squares (RSS)[98]:

$$\sum_{j=1}^{J} \sum_{i \in R_j} (y_i - \hat{y}_{Rj})^2,$$

2. Classification Trees:

- These trees predict class labels by determining the most probable outcome for each region. The splits are typically chosen based on metrics like Gini impurity or entropy, which measure the purity of the splits;

- The aim is to maximize the homogeneity of the resulting classes within each partition, making the decision rules simple yet effective.

One of the primary strengths of decision trees lies in their interpretability. Unlike more complex models such as neural networks, decision trees allow stakeholders to trace the decision-making process. Each branch and node is essentially a decision rule that can be easily interpreted by managers, business analysts, or technical experts alike. This makes decision trees especially appealing in fields like finance, where understanding the logic behind predictions is critical for making informed business decisions.

Moreover, decision trees excel in handling both categorical and continuous data[99]. They can automatically determine which variables are the most important for prediction, enabling managers to focus on the key drivers of business performance. For instance, when applied to insolvency and crisis prediction, decision trees can help identify financial ratios or accounting variables that are most indicative of potential insolvency (lo Conte et al., 2024; Altman, 1968). Researchers have found that tree-based methods often outperform linear models when capturing non-linear relationships between financial variables and the probability of distress (Zhang et al., 2016). For example, in predictive models, decision trees are adept at hierarchically partitioning the data to focus on the most relevant variables, such as leverage ratios, profitability, and liquidity measures (Guenster et al., 2011). This adaptability makes them particularly suited for financial risk management, where early detection of distress is critical. Furthermore, along with preditive analytics regarding firms and businesses, in Italy researchers also adopted ML and classification trees to predict bankruptcy of local governments (Antulov-Fantulin et al., 2021). However, this strand of research i sas residual as the one regarding predictive analytics in SMEs. In this case, there are some studies (Tron et al., 2022; Ferri, 2022), but the most prominent ones are based on mathematical or data-mining related studies (DI Ciaccio & Cialone, 2019). Furthermore,

[98] This objective function ensures that the partitioning process reduces the prediction error, yielding more accurate results.

[99] A key challenge with decision trees is the risk of overfitting. As trees grow deeper, they tend to capture noise and idiosyncrasies in the training data, leading to poor generalization on unseen data. This issue can be mitigated using pruning techniques, where non-essential branches are trimmed to simplify the model. This not only reduces complexity but also helps improve the model's performance on new data. In addition to pruning, ensemble methods such as Random Forests and Gradient Boosting Machines (GBMs) offer powerful alternatives to single decision trees. These techniques build multiple trees and aggregate their predictions, effectively improving predictive accuracy and reducing variance (Barboza et al., 2017).

just a few studies focused on the explainability and enhanced interpretability of these methods, determining a real gap in Italian literature. However, focusing on this area is necessary, as it is an "*hot topic*" today, and in managerial applications, this combination of factors allow decision-makers to base their strategies on well-understood and actionable insights, leveraging decision trees as part of an overall business intelligence framework (Sharda et al., 2020).

4.4.3. Random Forest

Random Forest, is an ensemble learning method that is used for both classification and regression tasks. It constructs a multitude of decision trees during training and outputs the mode of the classes (classification) or the mean prediction (regression) of the individual trees. This technique mitigates the overfitting typically associated with single decision trees and improves predictive performance. A random forest operates by creating multiple subsets of the training data through bootstrap sampling. Each decision tree in the forest is trained on a different subset, and during the construction of each tree, a random subset of features is considered for splitting at each node. This process ensures that the trees are decorrelated, thereby enhancing the robustness and generalizability of the model. The final prediction of a random forest classifier is determined by a majority vote among the trees. For a classification problem, the formula can be expressed as:

$$\hat{y} = mode(\{h_t(X)\}_{t=1}^{T}),$$

where y is the predicted class, *ht(X)* is the prediction of the t-th decision tree, and T is the total number of trees (Breiman, 2001). Khan et al. (2021) explained well the functions of this algorithm, also focusing on managerial impacts and stakeholders impressions on it (Figure 14). Core mechanisms are:

- *Bootstrapping and Feature Randomization.* The first step in constructing a Random Forest is to create multiple bootstrapped samples from the original training data. Each sample is created by random sampling with replacement, ensuring that each tree in the forest trains on a unique subset of data. This resampling process, known as bootstrapping, typically uses around two-thirds of the original dataset for training, while the remaining one-third, referred to as the out-of-bag (OOB) data, can be used to assess each tree's predictive performance independently. After bootstrapping, Random Forest introduces another layer of randomness by selecting a random subset of features at each split within each tree. This feature randomization (or "*feature bagging*") prevents highly correlated trees, as each tree will not necessarily evaluate all features at every split. Instead, each node considers only a fraction of the total features, typically set to d square root for classification tasks and d/3 for regression, where d is the total number of features. This random subset approach creates uncorrelated trees, enhancing the forest's diversity and robustness while reducing the model's variance;

- *Tree Growth and Splitting Criteria*. Each decision tree in the Random Forest grows independently according to the CART (Classification and Regression Trees) methodology. For classification tasks, nodes are split by maximizing information gain or minimizing Gini impurity. In regression, the tree minimizes the mean squared error

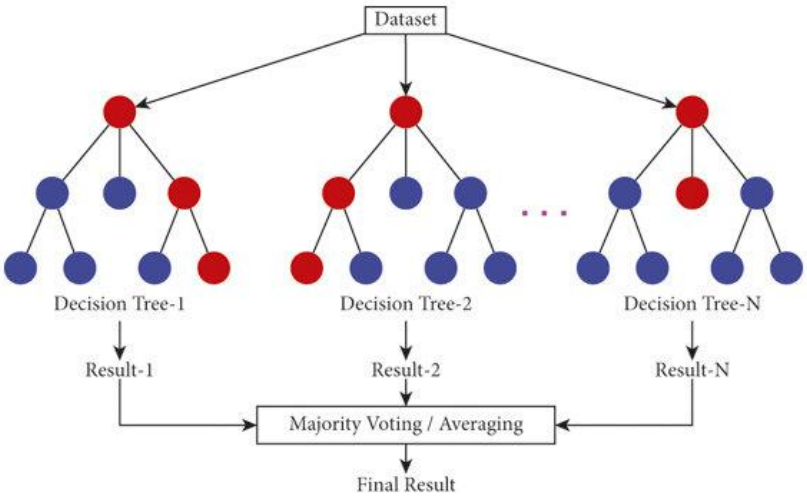
(MSE) at each split. Each node in a tree is split by evaluating the chosen metric over the subset of features until a stopping criterion is met, such as a maximum depth or a minimum number of samples per leaf. The unpruned trees in Random Forest are fully grown, meaning each tree explores its bootstrapped subset thoroughly. The depth and branching of these trees create unique decision boundaries, which, though individually overfit to each subset, collectively average out due to the ensemble, reducing overall model variance. Notably, this approach contrasts with traditional decision trees, where overfitting is a significant risk. Random Forest mitigates this risk by leveraging numerous overfit trees in combination, allowing individual biases to cancel each other out;

- *Aggregation and Voting Mechanism.* Once all trees are trained, Random Forest aggregates their outputs to arrive at a final prediction. For classification tasks, the forest uses majority voting across all trees in the ensemble: each tree "votes" for a particular class, and the most frequently chosen class becomes the final prediction. For regression tasks, predictions are aggregated by calculating the mean prediction of all trees, providing a stable and smooth output that averages out the noise from individual models. The aggregation process, whether through majority voting or averaging, significantly reduces both variance and bias, yielding a generalized model capable of strong performance on new data. This ensemble mechanism benefits from the diversity of predictions, where errors in individual trees are diluted through aggregation, and the law of large numbers stabilizes the output across varying subsets;

- *Out-of-Bag (OOB) Error and Performance Evaluation.* A distinctive feature of Random Forest is the use of out-of-bag error as an unbiased estimate of model accuracy. Since each tree is trained on a bootstrapped sample, the remaining third of the data not used in training can serve as a test set for that specific tree. The OOB error is calculated by averaging the prediction error for each sample over all trees where that sample was not included in training. This error estimate provides a nearly unbiased measure of the model's performance without the need for a separate validation set, allowing efficient assessment of model generalization during training. The OOB error offers insight into how well the model is likely to perform on unseen data and can be used to optimize parameters such as the number of trees, the maximum depth of each tree, and the number of features considered at each split[100].

[100] Random Forest requires the tuning of several hyperparameters to optimize performance:
• *Number of Trees:* Increasing the number of trees generally improves model performance, up to a point of diminishing returns, where the added computational cost does not significantly improve accuracy;
• *Max Depth:* Limits how deep each tree can grow. Controlling depth helps prevent overfitting, especially in high-dimensional data;
• *Minimum Samples per Split and Leaf:* Setting a minimum number of samples required to split nodes or form leaf nodes reduces overfitting by enforcing a constraint on tree complexity;
• *Number of Features per Split:* Determines how many features to consider at each split, balancing bias and variance. Lower values create more diversity among trees, while higher values lead to more accurate individual trees at the risk of higher variance.
These hyperparameters control the bias-variance trade-off, a fundamental aspect of Random Forest tuning. A larger number of trees and more features per split generally reduces bias, while shallower trees and fewer features reduce variance.

Figure 14: Random Forest Trees Illustration



Source: Khan et al., 2021

In managerial studies regarding insolvency and crisis prediction, as already stated, Italian academia focused on this method, but not too much was explained. On the other side, in the international context, there are some good examples of research in this way. For example, Makariou et al. (2021) focused on the us of Random Forest for predicting spreads in the primary catastrophe bond market; Gu (2021) uses the algorithm to capture early warnings of financial crises; Zhang (2024), on the other side, focuses more on the sustainability strand, demonstrating that the accuracy and recall rate of an improved algorithm based on Random Forest in identifying corporate financial distress are 98.03 % and 100 % respectively. Lastly, Hartini et al. (2021) estimate the probability of banking crises using this approach. Clearly, the use of Random Forest in the academic literature is broad, but the Italian context is still poor of contents.

4.4.4. Logistic Regression

Logistic Regression models the probability that a given input X belongs to a particular category (usually denoted as 1) using the logistic function. The model is defined as:

$$P(Y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k)}}$$

where *P(Y=1|X)* represent the probability of the dependant variable Y being 1, given the independent variables X1, X2, .., Xk. The parameters 0, 1, .., k are estimated from the data using maximum likelihood estimation. The logistic function ensures that the output is bounded between 0 and 1, making it suitable for probability prediction (Kleinbaum et al., 2012). One of the model's key advantages is its interpretability. Logistic regression, in fact, provides a clear, linear relationship between the independent variables and the log-odds of the event occurring. This interpretability is particularly beneficial in managerial contexts, where decision-makers need transparent models to

understand how specific variables (e.g., leverage or profitability) influence outcomes, such as the risk of insolvency. Additionally, the model's bounded output between 0 and 1 makes it ideal for predicting probabilities. This feature is advantageous for financial distress prediction, where the goal is to forecast binary outcomes (e.g., bankrupt vs. non-bankrupt).

Despite its advantages, logistic regression remains underutilized in Italian studies. One reason for this might be its perceived complexity, especially when compared to more straightforward methods like the Altman Z-Score, which remains more popular in practice. Studies have demonstrated the predictive accuracy of logistic regression, but it faces resistance due to the difficulty in interpreting log-odds and incorporating probabilistic predictions into traditional decision-making frameworks. Managers benefit from the ability to trace how individual predictors impact the likelihood of financial distress, offering a clear roadmap for interventions. However, despite these strengths, logistic regression has its limitations:

- *Linearity Assumption:* logistic regression assumes a linear relationship between the log-odds of the dependent variable and the independent variables. In real-world data, however, this assumption may not hold, especially in complex scenarios involving non-linear patterns (Floridi et al., 2018);

- *Lack of Flexibility:* in high-dimensional datasets or environments where complex interactions exist between variables, logistic regression may fail to capture these dynamics effectively. More advanced methods, such as Random Forests or Neural Networks, often outperform logistic regression in such cases (Robbins & Pearce, 1993);

- *Challenges in Italy:* despite its widespread use in academic research, logistic regression has not gained significant traction in the Italian corporate environment. Many businesses still prefer simpler, score-based models, which are easier to implement and explain (Oehmichen et al., 2023).

Despite the rise of these sophisticated techniques, this method remains a popular choice for predictive analytics in some financial settings due to its ease of use, interpretability, and well-established theoretical foundation (Kiron & Unruh, 2018). Moreover, it continues to be an integral part of predictive models, imainly, in risk management and national crisis prevention.

4.4.5. Boosting Methods

Boosting methods have emerged as one of the most powerful techniques in Machine Learning, particularly for enhancing model accuracy by transforming weak learners into strong learners (Perboli & Arabnezhad, 2021; Paliwal & Kumar, 2009). The core idea behind boosting is to combine the predictions of several weak models, which individually may perform slightly better than random guessing, into a robust and accurate predictive model. In technical terms, each model in the boosting sequence focuses on correcting the errors of its predecessor, thereby improving the overall prediction (Friedman, 2001). The concept of boosting was first introduced with AdaBoost (Adaptive Boosting) by Freund and Schapire (1997), which revolutionized the field of ensemble learning. AdaBoost works by assigning weights to misclassified instances and progressively improving the model by giving more emphasis to these errors. Over time, boosting methods have evolved, with notable variants such as

Gradient Boosting (Friedman, 2001), XGBoost (Chen & Guestrin, 2016), LightGBM (Ke et al., 2017), and CatBoost (Dorogush et al., 2018), each offering specific improvements in terms of computational efficiency, handling large datasets, and robustness to overfitting. Regarding the boosting method, it functions iteratively by fitting new models to the residuals of previous ones. The first step is then to fit a weak learner, often a decision tree, to the training data. Each subsequent model is then trained to correct the errors of the previous model by minimizing a loss function, typically by focusing on the most difficult instances. This process continues until no further significant improvements can be made, or the pre-defined number of iterations is reached. In mathematical terms, boosting can be understood as follows:

1. *Initial Model:* A weak learner, *f1(X)f_1(X)f1(X)*, is fitted to the dataset to predict y;

2. *Error Correction:* A second weak learner, *f2(X)f_2(X)f2(X),* is then trained to predict the residual errors (the difference between the actual and predicted y) from the first model;

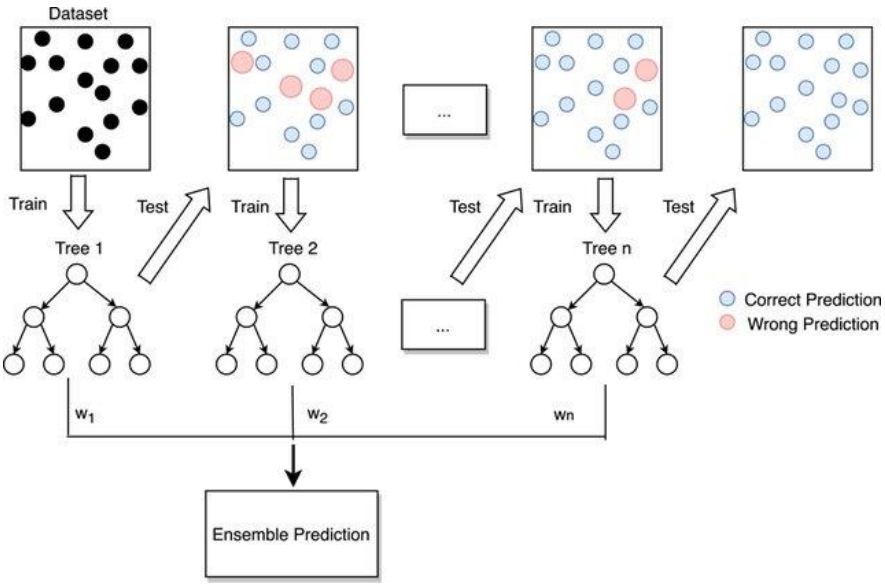
3. *Iterative Process:* At each iteration t, the new model *ft(X)f_t(X)ft(X)* is trained on the residuals from the previous combined model *Ft−1(X)F_{t-1}(X)Ft−1(X),* and the final model is the sum of all individual models.

Main models in predictive analytics are Gradient Boosting and XGBoost. A key difference between these lies in how they handle classification errors. While Gradient Boosting focuses on minimizing the overall error rate by constructing models to reduce errors in a gradual, stage-wise manner, XGBoost places more emphasis on penalizing complex models through its regularization mechanisms. This helps in preventing overfitting, which is a common issue in traditional boosting algorithms when working with noisy or imbalanced datasets (Chen & Guestrin, 2016). In terms of computational efficiency, XGBoost uses parallel processing and tree-based optimizations to handle vast datasets more effectively. It is designed to work well with both continuous and categorical variables, and its speed is significantly faster than traditional boosting methods. Moreover, gradient boosting methods are highly flexible and can work with different types of data, making them suitable for a wide variety of applications, including financial distress prediction, crisis management, and fraud detection (Gunasekaran et al., 2017; Friedman, 2001).

4.4.5.1. Gradient Boosting

Gradient Boosting, is an ensemble learning technique used for both classification and regression tasks. It builds models in a sequential manner, where each new model corrects the errors made by the previous ones. This method leverages decision trees as base learners and optimizes them by minimizing a specified loss function through gradient descent. In gradient boosting, the prediction model is composed of an ensemble of weak learners *ht(X),* typically shallow decision trees (Figure 15).

Figure 15: Flow Diagram of Gradient Boosting ML Method



Source: Zhang et al., 2021

The overall prediction y is a weighted sum of these weak learners:

$$\hat{y} = \sum_{t=1}^{T} \beta_t h_t(X),$$

where *Bt* are the weights determined during training, and T is the total number of trees. Each new tree is fitted to the residual errors of the combined ensemble of previous trees, effectively reducing the overall error in a stepwise manner (Natekin et al., 2013).

4.4.5.2. XGBoost

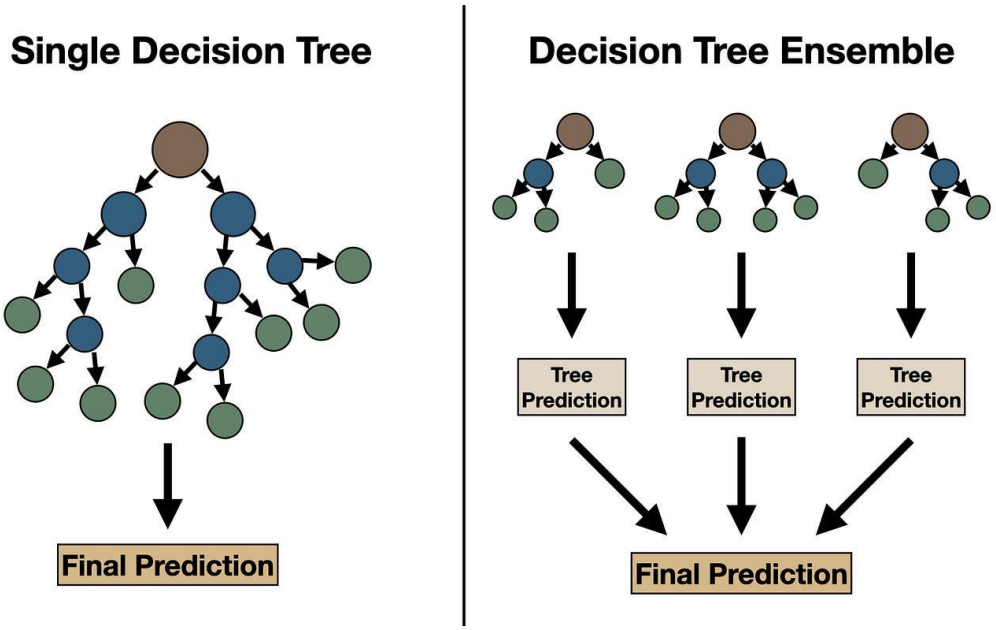
XGBoost, (Extreme Gradient Boosting) is an advanced implementation of gradient boosting designed for speed and performance. It builds an ensemble of decision trees sequentially, where each tree aims to correct the errors of the preceding ones by optimizing a specified loss function using gradient descent. XGBoost introduces several enhancements over traditional gradient boosting, such as regularization techniques to prevent overfitting, handling missing values, and parallel processing capabilities. The model prediction is formulated as:

$$\hat{y} = \sum_{t=1}^{T} \beta_t h_t(X),$$

where *Bt* are the weights assigned to each tree *ht(X)*, and T is the total number of trees (Chen et al., 2016). XGBoost works by creating an ensemble of decision trees sequentially, where each new tree tries

to minimize the errors (residuals) of the previous trees (Figure 16). The model then iteratively improves by optimizing a loss function—commonly a function like the mean squared error or log loss—using gradient descent.

Figure 16: Differences Between Single Decision Tree and Ensemble



Source: towardsdatascience.com

XGBoost refines this process with several optimizations:

- *Regularization:* One of the key enhancements in XGBoost is its built-in regularization techniques, which include L1 and L2 penalties (similar to Lasso and Ridge regression). These help to prevent overfitting, a common issue with complex models, especially when using a large number of features or decision trees. Regularization ensures that the model generalizes well on unseen data, which is particularly important for bankruptcy prediction, where avoiding overfitting is crucial due to the rare and unpredictable nature of financial distress events;

- *Handling Missing Data:* XGBoost can handle missing values during training, an essential feature in financial datasets where missing data is common. It can automatically learn which way to split based on missing data patterns, allowing it to impute missing values efficiently without requiring preprocessing;

- *Parallel Processing:* To further enhance performance, XGBoost leverages parallel processing. This allows the model to train more quickly on large datasets compared to traditional gradient boosting algorithms. Given that financial distress prediction often involves large datasets containing thousands or even millions of observations (e.g., balance sheets, cash flow statements), XGBoost's parallelism is a game-changer in improving the training time;

- *Gradient-Weighted Learning*[101]: In XGBoost, every tree is added in such a way that it "*corrects*" the mistakes of the previous one, but this correction is done by weighting each instance by the negative gradient of the loss function. This precise gradient-weighted approach ensures that the model focuses more on observations that are harder to predict, such as companies teetering on the edge of bankruptcy. This makes it particularly effective for financial predictions, where outlier firms often exhibit non-linear behavior in their financial metrics.

While other tree-based algorithms such as Random Forest or traditional Gradient Boosting also create ensembles of trees, XGBoost stands out because of its superior performance in terms of both speed and accuracy. Compared to Random Forest, which averages the predictions of multiple trees, XGBoost optimizes a loss function more aggressively by focusing on the hardest-to-predict cases. This enables XGBoost to outperform Random Forest in many prediction tasks, especially when precise and fine-tuned predictions are required, such as in financial forecasting. In the context of insolvency prediction, algorithms like logistic regression are often used due to their simplicity and interpretability. However, XGBoost has shown to be more accurate in predicting bankruptcy due to its ability to capture non-linear relationships and interactions between variables that simpler models might miss (Nguyen et al., 2023). Studies such as that by Veganzones & Severin (2023) have found XGBoost to handle imbalanced datasets—a common problem in bankruptcy prediction—better than SVMs, resulting in lower classification errors.

4.4.5.3. Managerial Implications

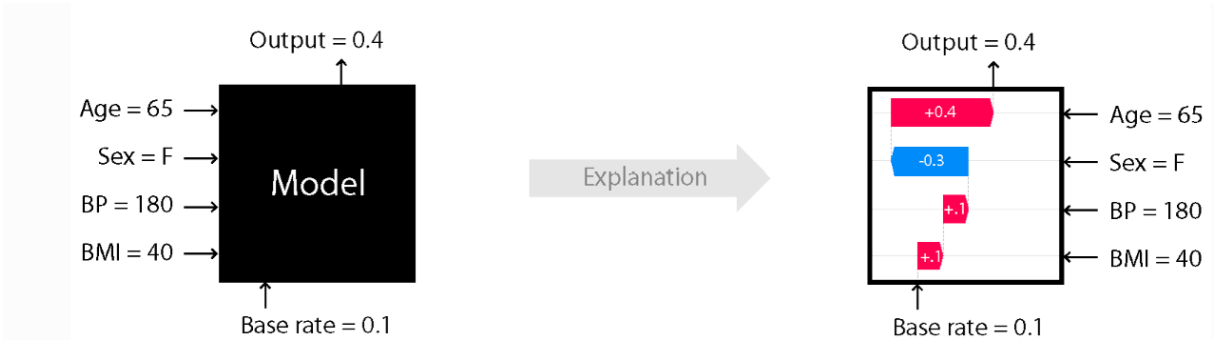
Boosting methods provide significant value to managers and decision-makers by offering data-driven insights that can inform strategic decisions. In the context of crisis management and financial distress prediction, boosting models help organizations anticipate risks and act preemptively to avoid crises. The interpretability of boosting models can be enhanced through techniques like SHAP (SHapley Additive exPlanations), which break down feature contributions and explain individual predictions, helping managers and stakeholders gain trust in the AI systems supporting their decisions (Chatterjee & Kar, 2024). However, while boosting methods are highly accurate and powerful, their complexity often presents a barrier to adoption, particularly in traditional management settings. Overcoming this barrier requires increased AI literacy among managers and a willingness to integrate advanced analytics into the decision-making process. By democratizing the use of ML tools like boosting, organizations can achieve agility, faster decision-making, and proactive risk management (Sudarsanam & Lai, 2001). In conclusion, boosting methods, particularly Gradient Boosting and XGBoost, represent some of the most advanced tools available for predictive analytics. These models, in fact, offer exceptional accuracy, scalability, and robustness, making them invaluable for business applications. Their ability to handle large, diverse datasets and extract meaningful insights makes them a powerful tool for organizations looking to enhance their predictive abilities.

[101] Gradient-Weighted Learning refers to an optimization technique commonly used in machine learning models, particularly in ensemble methods like gradient boosting. In this method, the algorithm builds a sequence of models, where each subsequent model focuses on correcting the errors made by its predecessor by applying weights to the instances based on the gradient of the loss function.

5. Bridging AI Interpretability and Managerial Insights: Wrapping-Up and Final Considerations

The integration of explainability and interpretability in all these models is a pivotal advancement today. These aspects are vital for ensuring that AI models can be trusted and adopted in real-world scenarios, especially in so technical fields, where understanding the rationale behind decisions is crucial. One of the most promising methods to enhance interpretability in ML models is the SHapley Additive exPlanations (SHAP) framework, which provides a clear, mathematically grounded approach to explain how individual features contribute to predictions (Lundberg & Lee, 2017). It is based on cooperative game theory, specifically Shapley values, which allocate credit for an outcome among multiple players. In the context of ML, SHAP assigns a value to each feature in a model, representing its contribution to the final prediction[102] (Figure 17).

Figure 17: SHAP Values Explanation



Source: https://shap.readthedocs.io/

This method then offers several advantages over traditional interpretability techniques, particularly in ensuring that complex models can be made transparent and understandable. SHAP, in fact, calculates the marginal contribution of each feature by considering the average contribution across all possible combinations of feature inputs. This offers a global and local interpretation, which helps managers understand how different inputs affect the model's predictions on both individual and aggregated levels.

Given this, from a managerial perspective and from the perspective used in this Thesis, SHAP represents a significant value proposition. By offering a breakdown of how individual features contribute to the prediction of outcomes like bankruptcy or financial distress, managers are better equipped to interpret results, understand risks, and make informed decisions. SHAP then demystifies the predictions by ML models, preventing them from being perceived as "*black boxes*," which often hinders trust and acceptance in traditional management environments (Chatterjee & Kar, 2024). As AI

[102] The fundamental concept behind Shapley value-based explanations in machine learning is to draw on the principles of fair allocation from cooperative game theory. This involves distributing the contribution of a model's output across its input features in a way that is considered equitable. To align game theory with machine learning, each input feature is treated as a "player" in the game, and the model's function is treated as the set of rules governing the game. This approach allows for an analysis that fairly attributes the impact of each feature on the model's prediction.

becomes more embedded in business operations, ensuring transparency is not just beneficial, but essential for widespread adoption and implementation (Floridi et al., 2018). Traditionally, the literature in this area relied heavily on statistical models that offered limited interpretability and had minimal practical implications. Standard models provide clear coefficients for their variables but often fail to capture the nuanced, non-linear relationships found in real-world financial data (Ohlson, 1980). With SHAP, the limitations of traditional models are mitigated. The method allows for the integration of a wider variety of variables and provides insights into feature importance, enabling managers to see which factors most significantly drive the context. This knowledge is invaluable in allowing firms to focus their efforts on the most critical variables and take proactive steps to avert crises. Furthermore, by providing a framework that is transparent, interpretable, and easy to use, SHAP allows decision-makers at all levels to leverage ML for financial risk management without needing deep technical expertise (Kiron & Unruh, 2018). This democratization is not just about increasing access but also about ensuring that AI-driven decisions are aligned with broader corporate strategies and stakeholder interests.

5.1. Creating Trust and Accountability

Another key benefit of SHAP and similar interpretability techniques is that they foster trust and accountability. As AI continues to make inroads into strategic decision-making, there is an increasing need for systems that offer not just predictions but explanations for those predictions. SHAP provides managers and stakeholders with granular insights into why a certain decision was made, allowing for a more informed and confident decision-making process. This is particularly important in crisis situations, where traditional management is often resistant to AI-driven suggestions due to a lack of understanding of the underlying mechanisms (Oehmichen et al., 2023). By making the decision process transparent, SHAP reduces the fear of the unknown that often accompanies AI, helping to integrate these tools more fully into the managerial toolkit.

In a broader sense, SHAP also encourages accountability. With clear explanations for each prediction, organizations can better understand the impact of their decisions and take responsibility for outcomes (Yeh & Liao, 2012). The use of SHAP and other interpretability frameworks also marks a significant shift in the relationship between academia and the business world. For years, the academic literature has focused on the development of statistical models with limited real-world applicability. However, as data-driven management practices become more widespread, there is a growing recognition that AI models must be both practically useful and understandable by the end-users (Idowu et al., 2023). SHAP provides a crucial link between the theoretical advancements and the practical needs of managers and stakeholders, helping to ensure that AI is not just a tool for data scientists but a strategic asset for organizations across industries. Then, one of the key contributions of this research is to demonstrate how SHAP can be integrated into a comprehensive AI-driven crisis management framework. By focusing on interpretability, the study aims to make ML accessible to managers, stakeholders, and policymakers alike, offering them a powerful tool for proactive crisis management. As AI literacy continues to improve, these tools will play an increasingly important role in shaping the future of management, allowing firms to anticipate risks, make data-driven decisions, and respond effectively to emerging threats (Nagarajan et al., 2024). In conclusion, explainability-enhancing models not only enhance the trustworthiness of AI models but also makes them usable and valuable for decision-makers, ensuring that AI is not seen as a rival but as a crucial ally for the human, and, mainly managerial, thinking.

5.2. SHAP in Literature

The SHAP methodology then marks a critical turning point in the relationship between AI and human decision-making[103]. In fact, it provides a level of transparency and explainability that was previously unattainable in traditional models. In today's business landscape, algorithms underpin virtually every aspect of decision-making, especially in large enterprises that have the resources to hire data scientists capable of leveraging these tools effectively. For smaller, less structured companies, achieving the same results requires simpler, cost-effective, and efficient methods—especially in terms of survival and investment, areas directly tied to anticipating and managing crises. SHAP, by enhancing explainability, addresses a key gap in literature and practice, particularly in prediction models where the lack of transparency in AI systems has long been a concern. This methodology is particularly promising because it democratizes the use of Machine Learning, allowing businesses without a dedicated AI team to understand and trust AI models. So, it breaks down complex black-box models, assigning a clear value to each input feature. This builds trust among managers and stakeholders, who are often skeptical of adopting AI-driven insights, particularly in times of crisis management.

Internationally, studies on SHAP as a tool for enhancing transparency and explainability in predictive models remain limited, and in Italy, the focus on this methodology is almost nonexistent. For instance, Veganzones & Seeverin (2018) explored, using SHAP, how the imbalanced distribution of datasets impacts bankruptcy predictions, noting that an imbalance where the minority class accounts for 20% severely impacts model performance. SHAP could play a critical role in other similar cases by providing insights into why certain features lead to misclassifications, thereby informing better decision-making processes. In another important study by Mahajan et al. (2023), SHAP was applied to analyze false positives in bankruptcy prediction, highlighting its role in enabling explainable AI to support decision-makers. This aligns with previous mentions in this dissertation about how explainability frameworks are now central to guiding managerial decisions. Similarly, Nguyen et al. (2023) evaluated the performance of ensemble machine learning models like XGBoost for French firms' bankruptcy prediction and applied SHAP to demonstrate how features influence the model's output. They concluded that SHAP made it possible to interpret black-box models, making them viable for credit risk management.

Enkhtuya & Kang (2023) further validate that SHAP fosters a transparent environment in financial distress prediction by offering model-agnostic explanations, crucial for stakeholders' trust[104]. Such research positions SHAP as a methodological guide for corporate crisis management, merging AI's technical sophistication with actionable managerial insights. Recent studies also defined SHAP as the only explanation method underpinned by solid theoretical foundations (Sigrist & Leuenberger, 2023; Perboli & Arabnezhad, 2021; Moen, 2020; Molnar, 2020), making it highly reliable for interpreting model outputs. However, in Italy there is a consistent gap in this area.

Beyond the realm of finance, SHAP has found widespread application across diverse fields, including drought prediction (Dikshit & Pradhan, 2021), power systems (Zhang et al., 2020), and air quality forecasting (Vega García & Aznarte, 2020), demonstrating its versatility in providing

[103] SHAP method is also used in model debugging, an essential process that involves pinpointing and rectifying issues that emerge during machine learning models' training and evaluation phases. See more on neptune.ai.

[104] The key idea behind model-agnostic explanation techniques is that they treat the model as a black box and work by analyzing the input-output relationships, generating explanations that highlight how each input feature influences the model's predictions. In the context of SHAP, this is achieved by calculating the contribution of each feature to the final prediction based on Shapley values, a concept borrowed from cooperative game theory.

transparency and interpretability in various predictive models[105]. So, through SHAP, AI-driven predictions move beyond raw statistical outputs, offering clear explanations that managers can leverage to not only identify potential financial distress but also take timely actions based on a deeper understanding of the variables at play. In summary, the transparency and accountability it offers to predictive analytics not only foster trust but also accelerate the adoption of intelligent decision-making tools.

[105] However, all these studies refer to more technical and data-related strand of research. This demonstrates the interdisciplinarity of this strand of research and its implications in different industrial, technical and managerial contexts. About this, see Molnar, 2020.

Chapter 4

Data-Driven & AI-Driven Decision-Making: an Empirical Investigation

Chapter Prelude

This Chapter introduces the empirical heart of this research, where the theoretical framework meets real-world data and application. Here, the Thesis investigate the dynamic relationship between AI-driven algorithms and human decision-making in the context of financial distress and crisis management. By utilizing extensive economic and financial datasets, alongside state-of-the-art methodologies, this part aims to uncover new perspectives on predictive analytics and strategic decision-making in the age of AI. The focus then shifts towards the practical integration of AI into managerial frameworks, showcasing how intelligent tools can enhance human judgment rather than replace it. This empirical investigation does not merely replicate international studies; instead, it seeks to advance the discourse on AI's role in the "hot topic" that is crisis and insolvency prediction today. Internationally, research is gaining momentum, but in Italy, these approaches are still in a nascent stage. This work then hopes to bridge this gap and open new avenues for both academic and managerial perspectives on AI-enabled financial forecasting. The insights derived from this analysis will offer a fresh lens on this evolving and, today, fundamental topics.

1. Introduction

In the upcoming empirical Chapter, the research transitions into its empirical core, focusing deeply into the dynamic and evolving relationship between AI-based technologies and human decision-making processes in the context of distress prediction and crisis management. First of all, AI (and ML), while often portrayed as a revolutionary force capable of transforming business operations, should not be viewed as an autonomous "*auto-pilot*" replacing human intervention[106]. Instead, the goal of this analysis is to emphasize AI's role as an "*intelligent support system*" that enhances the decision-making capabilities of managers, particularly in high-stakes situations where early warnings and proactive adjustments are critical to survival[107]. This empirical section then seeks also to address the growing need for managerial literacy in AI-driven tools, focusing on the principles of transparency and explainability, which are crucial to create trust in AI applications[108]. In fact, algorithms do not possess intuition or human judgment—they are just tools for augmenting managerial decision-making, not replacing it.

The ability of AI and ML to provide early, data-driven insights can significantly influence managerial strategies, offering crucial support in the early identification of financial instability and operational risks. Then, through the implementation of cutting-edge algorithms, this Chapter aims to showcase how AI can aid in improving performance by offering accurate predictive analytics that managers can use to adapt their strategies, redirect resources, and, if necessary, pivot toward more sustainable business practices. Yet, it is essential to recognize that AI lacks the capacity for ethical consideration or emotional intelligence—factors that are integral to human leadership. This is why AI should be seen as a partner in the decision-making process, providing actionable insights while leaving the ultimate strategic decisions in human hands. Furthermore, the empirical analysis explores some explainability frameworks, fundamental to mitigate concerns about usual AI's "*black-box*" nature. Explainable AI and algorithms, in fact, ensure that both managers and stakeholders can understand how predictions are generated, increasing transparency and trust in AI-driven outcomes: this transparency is particularly valuable in crisis management, where decisions must be swift, well-informed, and aligned with both short-term needs and long-term organizational goals.

The challenge, however, lies in democratizing these advanced tools across different business types and sizes[109]. Large corporations often have the luxury of dedicated data science teams capable of managing and interpreting these sophisticated systems[110]. SMEs, on the other hand, may lack these resources, making it essential to explore AI applications that are cost-effective, user-friendly, and adaptable to their specific needs[111]. This analysis, therefore, serves not only as a contribution to academic

---

[106] A significant limitation of viewing AI as an autopilot lies in its inability to replicate human intuition, emotion, or creativity. While AI excels in structured tasks, it relies on the data it's been trained on and struggles in novel situations that require out-of-the-box thinking.

[107] AI as an Intelligent Support System (ISS) offers a transformative role in enhancing human decision-making, particularly in complex business environments. Unlike traditional systems that follow static, rule-based algorithms, ISS uses machine learning (ML) and artificial intelligence (AI) to adaptively analyze large datasets and provide actionable insights. These systems are designed to complement human capabilities, providing support in decision-making processes (Shapley, 1953).

[108] This technical focus guarantees that models do not solely enhance accuracy but also usability and interpretability for a broad range of stakeholders.

[109] AI democratization, for example, involves making AI tools and platforms easier to use through low-code/no-code solutions, open-source libraries, and affordable cloud-based AI services. This not only enables companies without specialized data science teams to use AI but also brings AI into sectors where it can drive significant social and economic impact, such as healthcare, education, and public governance. The democratization of technology also brings about increased inclusivity and innovation.

[110] AI systems in big firms must comply with existing legal frameworks and industry regulations. This is especially true for industries such as finance, healthcare, and telecommunications, where regulatory requirements (e.g., GDPR, HIPAA) impose strict guidelines on data usage and privacy.

[111] For example, cloud-based AI platforms and user-friendly Machine Learning tools allow SMEs to access AI-driven insights without the need for in-house expertise. To understand better this topic, see lo Conte & Sancetta, 2024.

literature but as a practical framework for how AI and ML can be leveraged in diverse Italian business contexts to improve decision-making, performance, and, eventually, sustainability[112]. In conclusion, while disruptive AI-based technologies are reshaping the landscape of business analytics and strategic management, their true value lies in their ability to enhance—not replace—human decision-making, being support tools and elements useful to create an aligned, better business environment[113].

1.1. Preliminary Fundamental Considerations

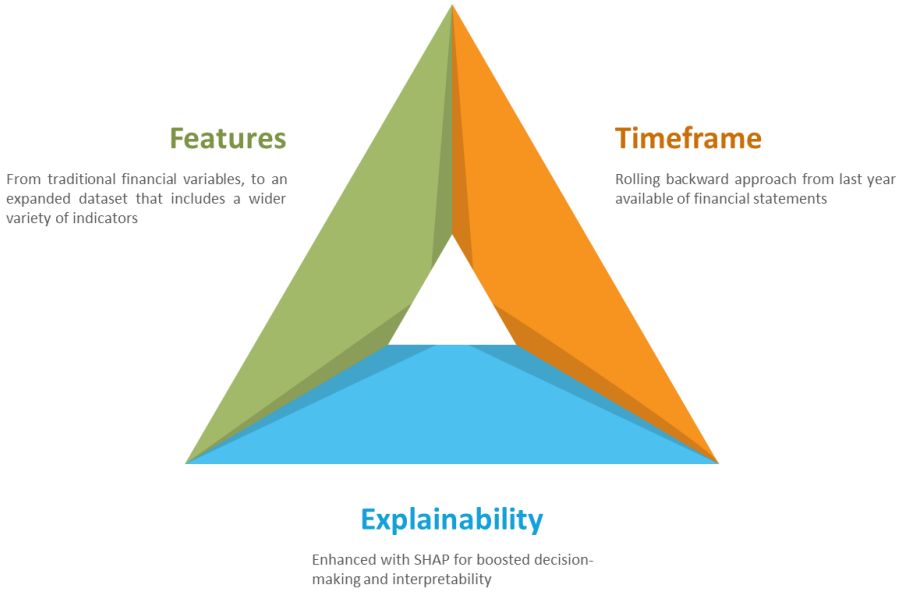
The empirical research is built around two distinct datasets, consisting of Italian firms, with differing sample sizes to evaluate the performance of different ML models: in fact, a key goal of the analysis is to assess how these predictive models perform in various contexts, using different types of data inputs. This three-step empirical approach is crucial in determining not only the overall accuracy of AI-driven predictions but also their adaptability to businesses of varying sizes and resource levels. In fact, it is important to remember that the algorithms are "*trained*" and their predictive power is highly dependent on the quality and volume of the data provided. In this study, several fundamental elements are highlighted: the timeframe considered, the number of indicators (or, hereinafter, "*features*", in technical terms), and the explainability of the results, especially through the use of the SHAP method[114] (Figure 18).

[112] The rapid growth in AI technologies has led to a proliferation of frameworks across industries and domains. Many organizations and developers have created specialized AI tools tailored to niche problems (e.g., healthcare, retail, finance). This can overwhelm decision-makers, who may struggle to identify the most suitable framework for their unique needs. The competition among frameworks, each with slightly different approaches or functionalities, can create fragmentation, reducing standardization and making adoption more complex. Conversely, there are still too few AI frameworks that address specific areas like crisis and insolvency prediction in a truly comprehensive way. AI frameworks in these areas often lack customization for different industries, company sizes, or decision-making processes.

[113] In the forthcoming analysis, AI's potential for enhancing business resilience will be juxtaposed with the limitations that arise when explainability is lacking.

[114] Existing studies using SHAP often fall short by merely identifying important features without offering further financial interpretation. About this, see Nguyen et al., 2023 and Sigrist & Leuenberger, 2023.

Figure 18: Empirical Research's Fundamental Elements



Source: own elaboration

Regarding the number of features, the literature has long debated the role of variables in bankruptcy and insolvency prediction, from FitzPatrick's (1932) foundational work and Altman's (1968) Z-Score model, which employed a limited set of financial ratios. As Bellovary et al. (2007) noted many years ago, prediction models have evolved significantly, with studies employing anywhere from 1 to 57 variables and more. Recent research has expanded the list of predictors to include more economic and financial variables, human resources, earnings management, and macroeconomic indicators, among others (Chen et al., 2023; Veganzones et al., 2023; Almaskati et al., 2021). This empirical investigation follows the same progression, with an initial focus on traditional financial variables, followed by an expanded dataset that includes a wider variety of indicators. This comparative analysis allows us to assess the benefits of using a more comprehensive set of features, especially given the availability, today, of large datasets and advancements in technological capabilities. The timeframe for this analysis is also critical. Unlike traditional statistical models, which often rely on fixed time periods, ML models allow for a more dynamic and iterative approach to understanding crises. The timeframe for this research is based on a rolling backward approach, where we look at the most recent available year of financial data and analyze the preceding years[115]. This method aligns with the existing literature, which suggests that predictive accuracy diminishes as the horizon lengthens (Nguyen et al., 2023; du Jardin & Séverin, 2011). While shorter horizons often yield obvious signs of impending bankruptcy, longer horizons introduce complexity, requiring more robust and sophisticated predictive techniques, with, sometimes, low-accuracy results (Altman et al., 2020). This study, therefore, emphasizes the importance of using, mainly, a -1 to -5-year horizon for predicting financial distress, a timeframe that is both practical and consistent with existing research. Based on these considerations, it is possible to affirm that this empirical analysis not only follows well-established methodologies in the literature but also introduces innovations under a managerial perspective, by applying AI-driven models that incorporate a broader range of variables, longer timeframes, and the use of SHAP for enhanced

[115] From financial statements.

explainability and boosted decision-making. In conclusion, the models used in this dissertation can be, in some cases, replicated easily also by market practitioners (Jones et al., 2017). However there is a trade-off between accuracy and interpretability called the "*black box problem*": in this regard, researchers, and in particular this specific research, try to overcome this problem by measuring the importance of features and the relationship between features and the predicted outcome.

1.2. Inductive Perspective

The empirical approach employed in this doctoral Thesis is rooted in an inductive methodology, which serves as a cornerstone of the research. This research, in fact, does not aim to test predefined hypotheses, a characteristic that aligns with Tharenou et al. (2018) and Azungah (2018), who highlight that in certain management studies, hypothesis testing is not always required. Given that the research area is relatively new, the focus is on inductive reasoning—starting from data to derive general conclusions and implications relevant to both academic scholarship and practical management applications (Woo et al., 2017). Furthermore, following the best practices recommended by Tharenou et al. (2018), this study collects and processes quantitative data, which is then subjected to statistical analysis to uncover underlying relationships, both theoretically and in practical contexts[116]. This approach contrasts with a deductive methodology, which typically begins with a well-defined theory and proceeds to test specific hypotheses (Ray, 1993). In this context, the inductive perspective allows for the development of theory directly from the empirical data, making it particularly suited to exploratory research in emerging fields. The research process adheres to key inductive research steps (Tharenou et al., 2018):

- Identifying the Research Questions to guide the investigation;
- Establishing the theoretical foundation and assessing the relevance of existing literature;
- Structuring a clear direction for the empirical analysis;
- Investigating the empirical results to explore relationships that are not only theoretical but also applicable in real-world scenarios.

This facilitates the generation of insights that are grounded in data, offering both theoretical and practical contributions to the field.

[116] Inductive reasoning plays a pivotal role in managerial studies, particularly when exploring relatively new or under-researched areas. As detailed by Tharenou et al. (2018) and other scholars, the inductive approach starts with the observation of data rather than preconceived hypotheses. This method allows researchers to develop general conclusions by examining patterns, trends, and relationships in the collected data. Unlike deductive reasoning, which begins with a theory and tests specific hypotheses, inductive reasoning is more exploratory and helps in theory-building rather than theory-testing (Tharenou et al., 2018). This approach is especially suitable for studies where theoretical foundations may not yet be fully established, and the goal is to derive meaningful insights directly from empirical evidence. The inductive method's strength lies in its flexibility, enabling researchers to adapt to findings as they emerge, refining conclusions as the data unfolds.

2. Research Design for Preliminary Analysis

The first analysis conducted in this study can be considered exploratory and preliminary in nature, determining a well suited basis for the next, solid, empirical part. In fact, it aims to identify early results on a restricted dataset of Italian firms[117], employing a limited, yet "*traditional*", set of features. This approach aligns with the iterative nature of the study, which starts with traditional assumptions from the existing literature—such as limited dataset size—and gradually expands the criteria to advance both scientific knowledge and managerial implications. The purpose of this phase is then to lay the groundwork for deeper empirical investigations by testing foundational methodologies. The research design for this initial analysis draws inspiration from previous studies on distress prediction, utilizing a basic structure to begin evaluating the effectiveness of ML models. The software used for statistical analyses is R, a powerful open-source programming language and environment widely used in data science, especially for statistical computing and graphical visualization. Furthermore, it offers various packages specifically designed for predictive modeling, making it a valuable tool for this type of empirical research.

The dataset referenced in these analyses comes from the AIDA database by Bureau van Dijk[118], one of the most commonly employed databases in the financial and economic literature. AIDA (also known as Diane in France, depending on the region) includes comprehensive balance sheets, financial data, and other relevant information for both listed and unlisted companies in Italy, spanning 10 years. The decision to use this database stems from several factors:

1. *Completeness:* AIDA provides access to all financial and economic records for Italian companies, covering a wide range of industries and firm sizes, including micro-enterprises. This is critical in the context of Italian corporate structures, where small and medium-sized enterprises (SMEs) dominate the business landscape;

2. *Literature Validation:* The most influential studies in the field of bankruptcy and insolvency prediction have relied on AIDA or similar databases. Its established use in the literature offers a degree of validation, providing consistency with other academic studies;

3. *Accessibility and Speed:* AIDA allows for efficient aggregation and retrieval of data directly from official sources, such as chambers of commerce, ensuring that the data is both up-to-date and comprehensive;

4. *Tailored Research Queries:* One of the strengths of the AIDA database lies in its flexibility. In fact, it offers a wide array of customizable search criteria, enabling researchers to fine-tune their datasets to specific needs. This is particularly useful for focusing on particular industries, financial ratios, or timeframes, which ensures a tailored approach to empirical analysis.

The research design for all three empirical studies follows a systematic, multi-phase, multi-algorithmic approach, reflecting best practices ML applications for predictive analytics, framed in a compehensive Intelligent Big Data Analytics approach. Furthermore, this empirical study adopts both a quantitative and qualitative perspective, especially in the phases of commenting on results and

[117] "*Restricted*" is referred to the entire study in dissertation. However, the number of firms considered, even in this first part, is scientifically significative.
[118] See https://login.bvdinfo.com/. AIDA stands for "*Analisi Informatizzata delle Aziende Italiane*".

implementing explainability frameworks. The quantitative aspect is reflected in the rigorous application of algorithms and statistical methods, while the qualitative aspect emerges during the interpretation and contextual analysis of the results, particularly when integrating the SHAP explainability model (Lundberg & Lee, 2017). Research design phases are, indicatively, the following[119]:

1. *Data Download and Collection:*
    - This involves the extraction of relevant data from the AIDA database using specific search criteria tailored to the research questions[120];

2. *Data Preprocessing:*
    - In this phase, raw data is cleaned and prepared for analysis. Tasks here include handling missing values, removing duplicates, normalizing financial ratios, and possibly transforming variables for consistency. This process may also include outlier detection and elimination, ensuring the quality of data is robust and reliable;

3. *Feature Selection and Optimization:*
    - Here, the most relevant variables (*"features")* for insolvency and financial distress prediction are identified;

4. *Dataset Preparation:*
    - The prepared dataset is split into training and test sets (20%; 80%), through cross-validation techniques, to ensure the models can generalize well on unseen data. This division is key in assessing the predictive power and reliability of ML models;

5. *Descriptive Statistical Analysis:*
    - Descriptive statistics such as mean, median, standard deviation, and correlation between features are calculated to better understand the data and provide initial insights. This step can reveal underlying trends and patterns that might inform the algorithm selection and modeling approach;

6. *Algorithm Implementation (Multi-algorithm Approach):*
    - In this phase, various ML algorithms are applied. Given the complexity of crisis and insolvency prediction, a multi-algorithmic approach is adopted, implementing models like Logistic Regression, Decision Trees, Random Forests, and XGBoost. These algorithms are selected based on their performance in prior studies and their ability to handle complex, high-dimensional datasets;

7. *Model Optimization:*
    - Techniques such as grid search or random search are applied to find the best combination of hyperparameters that maximize predictive accuracy;

[119] In the following sections these phases are explicitated, sometimes, in a mixed way, reflecting the strong relationships between each phase.
[120] *.csv* file.

8. *Model Evaluation and Validation:*
    - Each model's performance is evaluated using a combination of accuracy, precision, recall, F1-score, and AUC-ROC curve analysis. SHAP is used to provide interpretability, showing which features contribute the most to the model's predictions;

9. *Result Elaboration and Interpretation:*
    - The results of the predictive models are analyzed in detail, with a focus on managerial insights. The importance of each feature is evaluated, and the overall reliability of the predictions is assessed. SHAP is specifically employed to demystify the "*black box*" nature of ML models, providing transparent insights into why certain predictions were made;

10. *Assessment and Conclusion:*
    - Finally, the performance of the algorithms is compared across different datasets and scenarios. The results are assessed not only for their technical accuracy but also for their practical applicability to crisis management and crisis and insolvency prediction. The implications for managers and stakeholders are highlighted, especially in terms of early warning systems, decision-making, and crisis mitigation.

At last, in line with previous bankruptcy and insolvency prediction studies, the research opts to work with a subsample of non-in-crisis firms, a technique frequently employed by Altman (1968), Beaver (1966), du Jardin (2010), and more recently by Jabeur et al. (2021) and Smith & Alvarez (2022)[121]. This methodological choice allows for a more manageable dataset while ensuring the representativeness of the firms in the sample.

2.1. Data Collection and Feature Selection

For this first exploratory part, a total sample of 2.990 Italian firms was selected. So, the analysis focused on a limited sample of firms and a narrow timeframe. This targeted approach was chosen to ensure, and start, a more controlled exploration of the effectiveness of various algorithms, with particular attention given to the Random Forest algorithm. According to existing literature, in fact, Random Forest is highly suitable for studies involving smaller datasets due to its robustness and ability to handle feature selection and prevent overfitting in such contexts (Breiman, 2001). Furthermore, in this preliminary phase, the analysis did not include a dynamic and iterative examination over several years prior to the insolvency event. Instead, it concentrated on a single absolute year to derive initial results that align with the more "*traditional*" literature surrounding crisis and insolvency prediction (Altman, 1968; Beaver, 1966). This decision allowed for a more straightforward comparison with foundational studies, ensuring the research builds on established methodologies before advancing into more complex iterations.

This explained iterative nature of the research is a fundamental aspect, not just of this dissertation, but also of the broader scientific journey that has shaped the entire investigation. This iterative approach demonstrates the evolution of both scientific and practical considerations, showing how the research has been refined and expanded over time to meet evolving managerial needs and

[121] *"Healthy"* firms, active without events related to crisis or insolvency.

address emerging scientific complexities. This method also reflects a commitment to continuously improving the predictive capabilities of the models, marking a progression to more sophisticated, and explainable, data-driven and human-supporting techniques.

In this initial empirical step, the focus was specifically on identifying "*in crisis/insolvent*" firms, rather than other characteristics of the companies that could influence performance, such as revenue or profit margins. These additional features, while not the central focus of this phase, will become critical in subsequent analytical steps, where a broader array of factors will be incorporated to deepen the analysis. By beginning with a more targeted focus on insolvency, this stage is aligned with the primary objective of distress prediction, which forms the core of the research question. Later stages of the analysis will explore how different variables influence predictions and contribute to the development of a more robust predictive framework.

To create a balanced sample and ensure more accurate predictive results (Breiman, 2001), the number of "*healthy*" and active companies was matched to the number of distressed firms[122], allowing for a methodologically sound analysis[123]. To identify comparable *"healthy"* firms, an advanced R command to match the distressed firms with active ones based on their financial profiles was used[124]. This *1:1* matching process allowed to select companies that were similar in size, industry, and financial ratios but not in crisis or insolvent, ensuring comparability between the two groups. The following command was used:

```
matchit <- matchit(insolvency ~ financial_variables, data = dataset, method = "nearest")
```

This selection process ensured that the sample was heterogeneous, incorporating a wide range of firm types (e.g., S.R.L., S.R.L. a socio unico, S.P.A.) to represent the Italian corporate landscape fully, and equally[125]. The following specific criteria and parameters were employed in research:

- *Legal Form:* All available legal forms with financial statements were considered. These included various corporate structures such as *S.R.L* (Limited Liability Company), *S.R.L a socio unico* (Single-member Limited Liability Company), *S.R.L semplificata* (Simplified Limited Liability Company), *società consortile a responsabilità limitata* (Limited Liability Consortium Company), *S.P.A* (Joint Stock Company), *S.P.A a socio unico* (Single-member Joint Stock Company), *S.C.A.R.L.P.A* (Limited Partnership Company), *S.A.S* (Limited Partnership), and *S.C.A.R.L* (Cooperative Company with Limited Liability)[126];

- *Legal Events:* The dataset included companies that had undergone specific legal events indicative of crisis or insolvency in the last year of financial statement available. These events included: bankruptcy or liquidation due to bankruptcy, insolvency status, judicial liquidation,

[122] In this design, the companies classified as non-crisis firms, or "*healthy*" firms, are entities that are both active in the market and free of any historical crisis events or insolvency procedures. This classification is validated through objective sources, such as databases or records from Chambers of Commerce, which confirm the absence of any insolvency-related proceedings or similar indicators throughout the company's operational history. This criterion is crucial not only because it ensures a clear distinction between the two samples under consideration but also because it highlights that not all active companies in a dataset are necessarily "*healthy*." Active firms can also be undergoing crisis situations or be involved in judicial procedures for restructuring or liquidation.

[123] A balanced sample mitigates the risks of bias often associated with imbalanced datasets, ensuring that the model is more likely to generalize well across various scenarios.

[124] Propensity score matching *1:1.*

[125] By focusing on heterogeneous organizational structures, the analysis captures a broader spectrum of firm behavior, enhancing the external validity of the research.

[126] Listed and un-listed companies.

dissolution and liquidation, bankruptcy arrangement (concordato fallimentare), and other forms of liquidation[127]. These statuses reflect the actual insolvency or crisis-related condition of the firms, making them suitable for this predictive analysis[128];

- *Timeframe for Insolvency:* The time period for these events was from *January 1, 2016*, to *December 31, 2018*[129]. So, depending on the firms' year of last statement, the setting aims to predict the critical event 1, 2, or 3 years before. In the exploratory phase of this section, the years 2019 and beyond were excluded for practical reasons. Due to the relatively static nature of the exploratory analysis and the limited representativeness of the sample, including these years could have led to findings heavily influenced by the impact of the covid-19 pandemic, potentially distorting insights. However, this limitation was addressed in the second analytical step. In this predictive analysis phase, in fact, the inherently forward-looking nature of predictive modeling in 2024 necessitates moving beyond considerations overly bound to the effects of covid-19[130];

- *Test Balance Sheet Year:* The balance sheet year used for extracting financial indicators ("*features*") for predictive analysis was *2015*. This allowed for predicting insolvency based on financial information from the year preceding the insolvency events in the defined 3-year total timeframe;

- *Industry Sector:* All sectors were included in the analysis, classified according to the *ATECO 2007* industry code system, which ensures a comprehensive view across the business landscape[131];

- *Sample Size:* After applying the above filters, a total of 1.495 Italian companies were extracted from the database for analysis.

On the other side, the total sample of "*healthy*" companies was selected based on the following criteria:

- *Legal form:* All available legal forms with balance sheets were considered, including *S.R.L.*, *S.R.L. a socio unico*, *S.R.L. semplificata*, *società consortile a responsabilità limitata*, *S.P.A. a socio unico*, *S.P.A.*, *S.C.A.R.L.P.A.*, *S.A.S.*, and *S.C.A.R.L.* Using propensity score matching, the study matched the active companies 1:1 with distressed companies, arriving at a balanced and optimized sample

[127] The classification of events is based on practical concepts but is aligned with binding determinations imposed by the AIDA database, which is nonetheless validated by the existing literature.
[128] For the legal events in the research design, the focus is directly linked to the previously discussed concepts of corporate crisis and insolvency, which reflect the company's inability to meet its obligations. These events, viewed from the equity side of the Italian legal framework, align with the legal definition of insolvency under Italian law. Specifically, insolvency occurs when a firm is unable to honor its commitments, either by failing to repay debts or by ceasing operations altogether. From a legal perspective, the legislator considers insolvency as the condition where a company's liabilities exceed its assets, impeding the firm's capacity to fulfill its financial obligations. This legal viewpoint has been key in selecting specific bankruptcy-related events, such as liquidation, bankruptcy, judicial liquidation, and insolvency status, which offer concrete markers of financial distress.
[129] The years up to 2018 were selected for purely exploratory purposes. In the subsequent empirical sections of this dissertation, more current years will be considered. This approach allows for a comparative analysis across different time periods. By first exploring data from earlier years, particularly before the global disruptions caused by the pandemic, the research offers a foundation that can later be expanded with more recent data, ultimately strengthening the analysis by showing how crisis prediction models perform in different temporal contexts. This iterative process is essential for drawing robust, scientifically grounded conclusions.
[130] Also considering the limited availability of complete data just for the past ten years.
[131] For exploratory reasons.

of 1.495 active companies. This balanced dataset was essential to ensure that the ML algorithms produced the most accurate results;

- *Legal status:* Active, with no crisis-related and insolvency-related events;
- *Years with available financial statements:* 2015, 2016, 2017, and 2018;
- *Year from which the indices for the analysis were derived:* 2015;
- *Sector of operation:* All sectors were included, classified according to the ATECO 2007 codes.

This careful selection process ensured that the sample was balanced and ready for analysis, aligning with best practices in the literature on ML in crisis prediction (Nguyen et al., 2023; Breiman, 2001). A total sample of 2.990 Italian firms was extracted in an Excel format file. In this first exploratory step of the analysis, the selection of features was based on traditional financial analyses proposed by foundational literature (Ohlson, 1980; Altman, 1968; FitzPatrick, 1932) and, mainly, practical, real-world applications. The objective was to ensure that even this early stage of the research had both academic relevance and real-world applicability. As a result, the feature selection was not limited to theoretical paradigms but also included the actual criteria used by practitioners involved in monitoring and evaluating business performance (even in crisis situations). This approach aligns with best practices found in the literature and allows for an initial assessment before moving to the broader sample in the second empirical phase. The specific features selected for each company in this initial step, focused primarily on their economic and financial nature, were as follows:

- *Debt to equity ratio:* Measures a company's financial leverage, showing the proportion of equity and debt used to finance the company's assets. This ratio is essential for insolvency prediction as high leverage often signals distress;
- *Debt to EBITDA ratio:* Provides insight into a company's ability to pay off its incurred debt from its operational earnings, which is crucial for understanding its risk of default (Ohlson, 1980);
- *Current ratio:* This liquidity ratio indicates a company's ability to cover short-term obligations with its current assets, a direct signal of financial health;
- *Interest expense to revenue ratio:* This shows how much of the company's revenue is spent on paying interest, revealing how burdened the company is by debt, a critical factor in predicting financial instability;
- *Average receivable duration:* Assesses how long it takes the company to collect payments, which can be indicative of liquidity problems;
- *Average payable duration:* Indicates how long a company takes to pay its obligations. Prolonged payment periods could suggest cash flow difficulties;
- *Return on turnover:* Measures the efficiency with which a company uses its capital to generate sales, a low turnover might indicate inefficiency and impending trouble;

- *Legal form:* Although not economic or financial, the legal form of a company was included to assess whether different legal structures (e.g., S.R.L., S.P.A.) have an impact on financial outcomes and insolvency risk;

- *Industry sector (Ateco Code 2007):* Similar to legal form, the industry classification was added to evaluate whether certain sectors are more prone to economic and financial difficulties.

These features reflect both academic rigor and practical relevance, allowing for a primary test in this first step before transitioning to a more extensive sample. For legal form and industry sector, although not directly tied to financial outcomes, we included them to facilitate a descriptive analysis of the companies and assess how these variables might impact model performance in the later stages (Nguyen et al., 2023).

2.2. Data Preprocessing and Optimization

Data preprocessing was a critical step to ensure the accuracy and reliability of the analysis. Given the complexities of prediction, multiple preprocessing techniques were applied to cleanse and prepare the dataset that was. Below are the key operations performed:

- *Data Cleaning:* As part of the initial data cleaning process, Debt to EBITDA Ratio labeled as "*n.s.*" or "*N.A.*" were removed from the analysis[132]. This ensured that irrelevant or misleading data points did not skew the results. Additionally, non-numerical values in the some fields were converted to 0, as they did not hold meaningful numerical information for the analysis[133];

- *Handling Categorical Variables:* For categorical variables such as "*ATECO Code*" and "*Legal form*," Frequency Encoding was applied. This approach assigns a numerical label to each category based on its frequency in the dataset. For example, if 30 companies had the legal form "*S.p.A.*," each instance of "*S.p.A.*" was replaced by the value 30. This method was chosen, following Pargent et al. (2022), to maintain the ordinal significance of frequently occurring categories without introducing bias from arbitrary ordinal labels[134];

- *Feature Scaling and Transformation:* Min-Max Scaling was used to bring all features to the same scale, facilitating more efficient model convergence during training;

- *Data Balancing and Sampling.*

[132] This was done for every n.s or N.A. in the dataset, if without impact on results.
[133] Following conversion, statistical tests (such as variance checks) were performed to ensure that the inclusion of 0 in place of non-numeric entries did not skew the distribution of values or introduce bias into model training processes.
[134] In their 2022 paper titled "*Regularized Target Encoding Outperforms Traditional Methods in Supervised Machine Learning with High Cardinality Features*", Pargent et al. investigate techniques for encoding high cardinality categorical variables, which are often challenging for ML algorithms designed primarily for numerical inputs. The authors focus on methods that convert categorical variables into numeric representations, which can then be used in ML models. The study concludes that regularized target encoding outperforms traditional encoding techniques, particularly when dealing with high cardinality features.

After preprocessing, optimization techniques were applied to improve model performance:

- *Hyperparameter Tuning:* Was conducted using Grid Search to identify the best configuration for parameters like the number of trees, maximum depth, and learning rate[135];
- *Cross-Validation:* K-Fold Cross-Validation (K=5) was used to evaluate the model's performance, ensuring that it generalizes well to unseen data[136];
- *Regularization.*

At the end of the data preprocessing and optimization phase, a correctly balanced dataset was obtained. "*Healthy*" firms, hereinafter, are categorized as "*0*"; "*In-crisis*" one sas *"1"*[137].

2.3. Descriptive Statistics and Correlation Analysis

Descriptive statistics analysis is critical, as it serves as the foundational step for any empirical study, particularly in fields like crisis prediction and crisis management. In line with the best practices outlined by Tharenou et al. (2007), descriptive statistics are indispensable in management research as they provide a clear understanding of the data, offering both researchers and practitioners a context to interpret findings effectively. So, this study follows this way. Descriptive analytics for eco-fin features were the following (in R code):

```
Debt_Equity_Ratio = rnorm(2990, mean = 1.3, sd = 1.5),
Debt_EBITDA_Ratio = rnorm(2990, mean = 3.5, sd = 1.0),
Current_Ratio = rnorm(2990, mean = 1.6, sd = 0.6),
Interest_to_Revenue = rnorm(2990, mean = 2.2, sd = 1.5),
Avg_Receivable_Duration = rnorm(2990, mean = 120, sd = 30),
Avg_Payable_Duration = rnorm(2990, mean = 130, sd = 25),
Return_on_Turnover = rnorm(2990, mean = 1.3, sd = 0.5)
```

This provided insights into the central tendencies and dispersion of key financial metrics used in the prediction. For instance, the Debt to Equity Ratio reveals a moderately leveraged profile across firms (mean = 1.3), which, along with the Debt to EBITDA Ratio (mean = 3.5), indicates a mix of operational efficiency and financial risk. The Current Ratio (mean = 1.6) suggests most firms have sufficient liquidity to cover short-term obligations, while the Interest to Revenue Ratio (mean = 2.2) implies that interest expenses are a manageable part of revenue generation. The Average Receivable and Payable Durations reflect typical working capital cycles, providing a baseline for liquidity and cash flow assessments critical in crisis prediction. For Legal form count, results were the following:

---

[135] The Grid Search method employed cross-validation on each hyperparameter configuration to ensure the reliability of performance metrics and prevent overfitting to specific data subsets.

[136] Following steps were conducted:

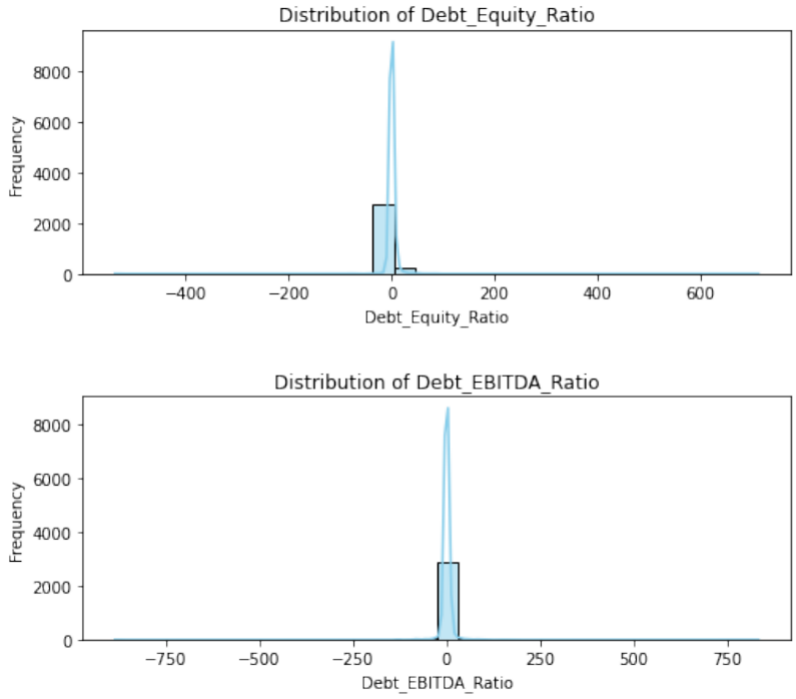
- *Dataset Partitioning;*
- *Iterative Training and Testing;*
- *Average Performance Check;*
- *Overfitting Check.*

[137] For simplicity, studies refer to firms in 1 as "*in-crisis*" entities, a term widely validated by the literature. This classification is applied even when firms under consideration have undergone formal judicial liquidation or insolvency-related procedures in the final year included in the analysis. These firms are considered in a critical situation, wherein the deterioration of their economic and financial performance has been captured by formalized judicial procedures. Given the varied levels of financial distress represented by these scenarios, this approach ensures that predictive analysis can effectively encompass both crisis and insolvency situations, allowing for a comprehensive assessment across different stages of financial deterioration. This terminology also enables consistent reference to firms experiencing significant distress, aligning with established definitions in academic research and enhancing model applicability.

```
legal_form_distribution <- data.frame( legal_form = c("S.R.L. a socio unico",
"S.C.A.R.L.P.A.", "S.R.L.", "S.P.A.", "S.R.L. semplificata"), count = c(1200, 1100, 550,
90, 50)
```

The majority of the companies in the dataset reflect a prevalent form of corporate structure with a restricted ownership and control. This confirms the well-established trend in Italy, where closely-held companies structures, dominate the entrepreneurial landscape. Such firms typically have more concentrated ownership, which can impact governance practices and influence their financial stability, especially in times of crisis. For some key financial features (Broby, 2022), a distribution analysis was conducted to understand the spread and identify potential outliers or anomalies. The majority of values cluster near zero, indicating that most firms maintain relatively balanced ratios, though a few extreme values suggest higher financial leverage in some companies, which could signal potential financial distress (Figure 19 & 20).

Figure 19 & 20: Distribution Analysis

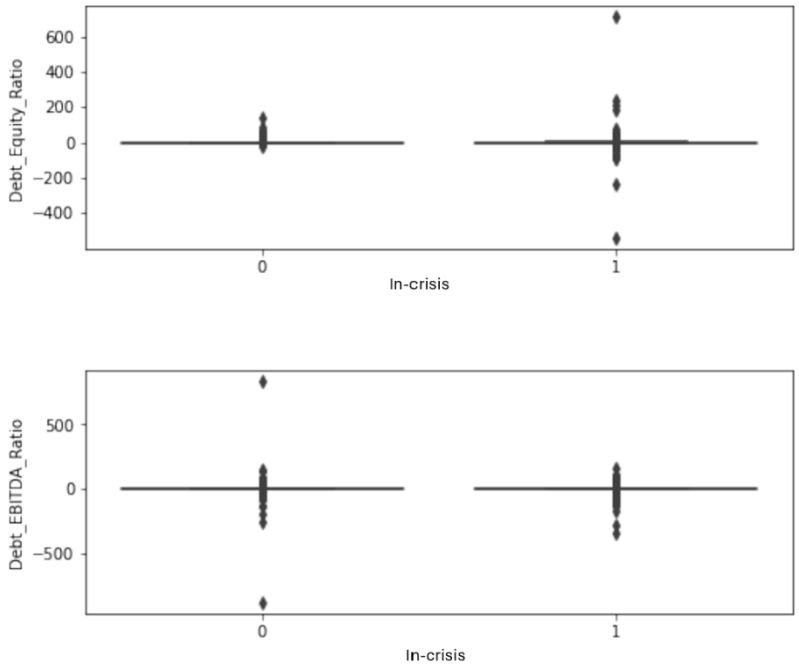


Source: own elaboration

Figure 21 & 22, instead, illustrate the relationship between Debt to Equity Ratio and Debt to EBITDA Ratio against the "*in-crisis*" status. First plot shows a greater spread of values, particularly for in-crisis or insolvent companies. Outliers can be observed both above and below the main distribution, indicating that distressed firms tend to have extreme leverage (either high debt or negative equity). Healthy firms generally cluster around lower, more stable values. Similarly, in-crisis firms display a wide distribution with several outliers for the other ratio, highlighting that companies with very high or negative Debt to EBITDA ratios are more likely to face financial distress. Healthy firms, on the other hand, maintain more moderate ratios, suggesting that they manage their debt more effectively in

relation to their earnings. In conclusioni, these boxplots suggest that higher values and extreme outliers in these key financial ratios are associated with distress situations, reinforcing their importance in insolvency prediction models.

Figure 21 & 22: Relationship Between Debt to Equity Ratio and Debt to EBITDA Ratio Against the "*In-Crisis*" Status



Source: own elaboration

As part of the empirical analysis, a *T-test* was conducted to investigate statistically significant differences between in-crisis and healthy companies for key variables. The T-test is a standard statistical tool used to determine whether there is a significant difference in the means of two groups. In this case, it was employed to assess differences in financial ratios and operational features, which are crucial for predicting financial distress:

- *Debt-to-Equity Ratio:* The results show that there is no significant difference in the Debt_Equity_Ratio between insolvent and non-insolvent companies, with a p-value of 0.54. This suggests that, for this dataset, the leverage ratio does not significantly distinguish between companies facing financial distress and those that are stable;

- *Average Receivable Duration*: A highly significant difference was observed between the two groups, with a p-value of 1.17e-12. This result indicates that the time it takes for companies to collect receivables is much longer for insolvent companies compared to those that are not insolvent. This is a strong signal that liquidity management could be a key factor in predicting insolvency;

- *Average Payable Duration:* Another highly significant difference was identified for this variable, with a p-value of 1.47e-38. Insolvent or in-crisis companies tend to take significantly longer to pay their debts, further highlighting liquidity problems as a major predictor of financial distress;

- *Interest Expense to Revenue Ratio*: Similarly, there was a significant difference between the two groups, with a p-value of 8.80e-7. Insolvent companies have a much higher proportion of their revenue allocated to paying interest, suggesting a higher debt burden that may contribute to their financial instability;

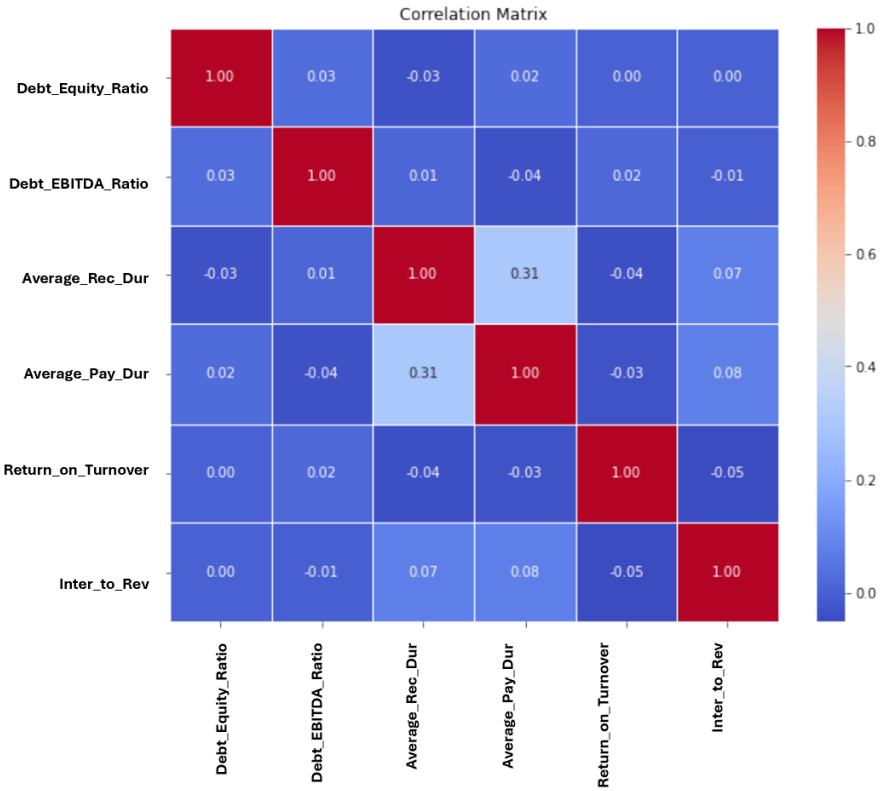
- *Current Ratio:* The test revealed a significant difference between insolvent and non-insolvent companies, with a p-value of 4.32e-9. Insolvent companies tend to have a lower Current Ratio, indicating a reduced ability to meet short-term obligations. This finding aligns with the notion that liquidity is often a precursor to financial distress, as companies facing insolvency are less capable of covering their liabilities with current assets;

- *Debt-to-EBITDA Ratio:* There was also a significant difference for this variable, with a p-value of 2.68e-5. Insolvent companies tend to have a much higher Debt-to-EBITDA Ratio, meaning that their operational earnings are insufficient to cover their debt load;

- *Return on Turnover*: The T-test results for Return on Turnover showed a significant difference, with a p-value of 3.54e-8. Insolvent companies displayed a lower Return on Turnover, suggesting that they are less efficient in utilizing their capital to generate revenue.

Following the descriptive analysis, a correlation analysis was conducted, crucial in understanding the relationships between the selected financial features and to identify any multicollinearity issues that may affect the performance of ML algorithms (Tharenou et al., 2007). Results can be seen in Figure 23.

Figure 23: Correlation Analysis



Source: own elaboration

The highest positive correlation is between Average Receivable Duration and Average Payable Duration (0.31), suggesting that firms that take longer to pay their creditors also tend to wait longer to receive payments from customers. This could reflect liquidity management strategies, especially in firms that balance incoming and outgoing cash flows. On the other hand, features like Debt to Equity Ratio and Return on Turnover show almost no correlation with other variables, suggesting that these metrics might capture independent aspects of financial health that don't necessarily align with the liquidity or operational metrics considered. This low correlation between most variables is beneficial for ML models, as it reduces multicollinearity and ensures that each variable provides unique information[138].

3. Multi-Algorithmic Approach Implementation and Model Evaluation

In this section, the description of the multi-algorithmic approach implemented in the first empirical analysis is explained. The entire strategy is designed to improve predictive accuracy and robustness by leveraging multiple ML algorithms, each with its own strengths and weaknesses, based on an inductive and data-driven approach. The rationale behind a multi-algorithmic approach lies in its

[138] Low correlation was also observed for Current ratio feature.

ability to address the limitations of individual models, combining their complementary advantages to create a more reliable and precise data-driven prediction of corporate distress. The Random Forest algorithm was selected as the primary model. According to the literature, Random Forest is particularly well-suited for datasets with the characteristics of our study, offering a balance between interpretability and predictive power (Breiman, 2001). It is also one of the more traditional ML algorithms used for classification problems such as insolvency prediction, and it has been widely adopted in the field due to its robustness and simplicity compared to more advanced models. In addition, Gradient Boosting and XGBoost were utilized, which represent more advanced and computationally demanding algorithms. Gradient Boosting is a sequential ensemble method where trees are built to correct the mistakes of their predecessors; XGBoost, on the other hand, is a highly efficient and scalable implementation of Gradient Boosting that leverages parallel computing and GPU capabilities, making it the most advanced algorithm in this study (Chen & Guestrin, 2016).

For Random Forest, a more in-depth analysis was performed, including attempts to improve prediction accuracy by tuning the hyperparameters. To begin, and in line with best practices (Kuhn & Johnson, 2013; Tibshirani et al., 2012), the dataset was split into two sets: a training set and a test set. This split is fundamental to ensuring that the model generalizes well to unseen data and provides reliable predictions. The training set was used to fit the model, while the test set was reserved for evaluating its performance. Importantly, the data was balanced, ensuring that the proportion of in-crisis (1) and healthy (0) firms was maintained in both the training and test sets[139]. So, the main objective of the algorithmic approach in this study is to evaluate the capacity and accuracy of these models in detecting early signals of financial distress and crisis well in advance. Specifically, the focus is on assessing the ability of these models to predict insolvency by identifying early warning signs in the financial data from 2015, preceding the actual crisis or insolvency events that occurred in 2016, 2017, or 2018. It is essential to reiterate here, as discussed in previous Chapters, that a crisis is a gradual and complex process that can unfold over several years, with various potential outcomes. A crisis may evolve into a declared and irreversible insolvency, or, in some cases, be reversible and lead to significant organizational restructuring. In this study, we focus on the period when the crisis becomes manifest through external communication, either via legal events or corporate disclosures, such as insolvency or judicial liquidation of company assets. For this reason, it is crucial to investigate the indicators that could provide early warning signs of distress. These signs, if identified and addressed in a timely manner, can enable a proactive and efficient management of the crisis, or even prevent its occurrence altogether. By understanding these early indicators and acting upon them, firms may avoid more severe outcomes,. This entire approach aligns with the goal of our empirical analysis, where the aim is to assess how innovative and intelligent models can support the timely detection of such warning signs, facilitating better decision-making for crisis management and prevention.

Starting with the multi-algorithmic approach, for the Random Forest model, all available variables were utilized except for the company name[140]. The number of variables chosen for each split was 2, instead of the default value of 3, which resulted in a lower error rate[141]:

---

[139] This balancing step is crucial for avoiding prediction bias, particularly in classification problems where one class (e.g., healhty firms) dominates the dataset.

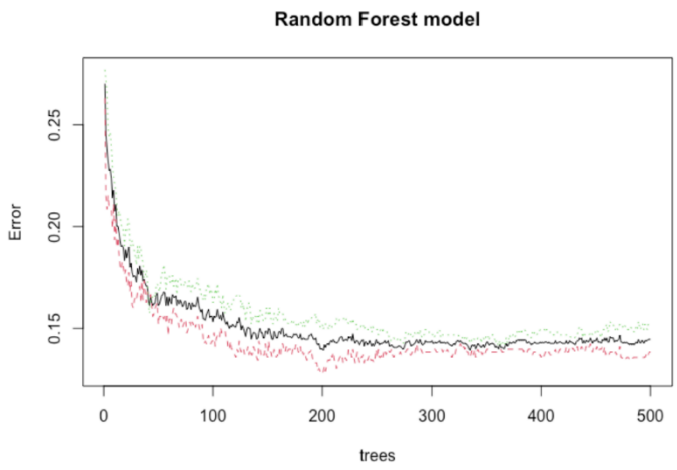
[140] Because it had no influence for the study's objective. However, it was present in the Excel file.

[141] The decision to select 2 variables per split, instead of the default 3, was driven by the objective of improving model performance. In Random Forest, selecting fewer variables at each split can help in reducing overfitting, especially in datasets where certain features may dominate splits, thus leading to a model that overemphasizes specific patterns. By lowering the number of variables per split, the model becomes more diverse, forcing it to consider a broader range of features across the trees. This approach often leads to a more balanced model and, as observed in this study, resulted in a lower OOB (Out-Of-Bag) error rate, improving predictive accuracy and reducing the model's tendency to overfit to the training data (Breiman, 2001).

```
random_model <- randomForest(formula = as.factor(in-crisis) ~ . - company name
data = data_training,
mtry = 2)

Type of random forest: classification
Number of trees: 500
Number of variables tried at each split: 2
OOB (Out-Of-Bag) estimate of error rate: 14.47%
```

The chosen number of trees in the model is 500. The final model, with 500 trees and 2 variables selected at each split, returns an error rate of 14.5%. This error rate represents the average prediction errors for each training instance, *xix_ixi*, calculated through the aggregation of predictions made by only the trees that did not contain *xix_ixi* in their training data (the Out-of-Bag error). The error rate is reasonably low, with the confusion matrix showing lower errors for predictions of the class "*0*" (*healthy companies*) and slightly higher errors for predictions of class "*1*" (*in-crisis companies*). This discrepancy in error rates for different classes is illustrated in Figure 24[142].

Figure 24: Random Forest Error Rate Graph[143]



Source: own elaboration

Technically, the Random Forest algorithm was then implemented to obtain the confusion matrix, which contains the following elements:

- *True Positives (TP):* Correctly predicted positive instances (in-crisis companies correctly classified as in-crisis);

- *True Negatives (TN):* Correctly predicted negative instances (healthy companies correctly classified as healthy);

[142] The prediction error decreases significantly up to 200 trees, after which it remains essentially constant. Overall, the estimation error is relatively low.
[143] Black line → estimated overall test error;
Green line → estimated test error for the positive class;
Red line → estimated test error for the negative class.

- *False Positives (FP):* Incorrectly predicted positive instances (healthy companies incorrectly classified as in-crisis)[144];

- *False Negatives (FN):* Incorrectly predicted negative instances (in-crisis companies incorrectly classified as healthy).

Obtained confusion matrix is the following:

```
[[228, 73]
[26, 271]
```

resulting in 228 True Negatives; 73 False Positives; 26 False Negatives; 271 True Positives. The model is then proficient at correctly identifying insolvent companies; tends to classify some healthy companies as insolvent, which could lead to unnecessary interventions but is often preferable to missing insolvency risks; a low number of false negatives is vital, as it means the model successfully captures most of the in-crisis companies and minimizes the risk of missing critical crisis signals; it also performs well in correctly identifying healthy companies. This result indicates that as early as 2015, there were already warning signs that the companies in question were not financially healthy, even though the crisis had not yet been externalized or formalized through specific insolvency procedures. The fact that these firms entered into insolvency or crisis procedures in the subsequent three years demonstrates the predictive power of the model. It effectively identifies early signs of financial distress, highlighting its ability to forecast a crisis one to three years before it fully manifests. The results were then evaluated using three specific and fundamental metrics (Hossin & Sulaiman, 2015):

- *AUC (Area Under the Curve)*[145] = 0.83: This metric assesses the performance of a classification model by evaluating the ability to distinguish between classes across different threshold levels. An AUC of 0.83 indicates strong model performance, meaning the model has a high probability of correctly classifying companies;

- *Accuracy = 0.84*: This is a strong result, suggesting that the model has a high overall predictive capability[146];

- *F1 Score = 0.84:* This metric represents the harmonic mean of precision and recall, providing a balance between false positives and false negatives. An F1 score of 0.84 suggests that the model maintains an excellent balance between identifying true positives and minimizing false positives and negatives.

Following the initial implementation of the Random Forest model, the multi-algorithmic analysis continued with the application of two additional ML algorithms: Gradient Boosting and XGBoost. Each of these algorithms was chosen for its specific characteristics and its potential

[144] As we will see in the next part of this Thesis, sometimes, False Positives are not, in reality, "*false*".
[145] "*Under the ROC curve*": the ROC curve illustrates how well a classifier distinguishes between classes by showing the trade-off between recall and the false positive rate.
[146] A high accuracy might still hide cases where the model fails to identify certain critical cases (such as false negatives), which are equally important for crisis prediction. Thus, while promising, this accuracy must be complemented by a deeper evaluation using other metrics to ensure robust model performance.

contribution to the predictive accuracy of the model. Also XGBoost model demonstrated strong performance in this analysis:

```
XGBoost Classifier:
AUC: 0.83305927491974
Accuracy: 0.8220
F1 Score: 0.8387096774193549
Confusion Matrix:
[[238 63]
[ 37 260]]
```

An AUC closer to 1.0 would represent perfect classification, so a score of 0.83 suggests the model is performing well but leaves some room for improvement in differentiating between the two classes. The F1 score of 0.84, suggests a good balance between these two metrics. It shows that the model is performing well in terms of both detecting the positive cases (true positives) and minimizing false positives. This high F1 score also suggests the model is well-calibrated for practical applications in financial distress prediction. Also Accuracy shows good results. Results seems similiar to Random Forest ones. On the other side, classification matrix demonstrates that the model has a relatively low number of false negatives, which is crucial, as failing to identify a firm in distress could lead to significant managerial oversights. All these results confirm the already spoken trends regarding the role of ML in predicting alerts. At last, Gradient Boosting was employed. Results are the following:

```
Gradient Boosting Classifier:
AUC: 0.8449332751658333
Accuracy: 0.85337
F1 Score: 0.8535433070866141
Confusion Matrix:
[[234 67]
 [ 26 271]]
```

Even if it is more simple than its upgraded version (XGBoost), Gradient Boosting performed better than the other two algorithms. The model shows strong predictive performance, with an AUC of 0.8449, indicating good discriminatory power between insolvent and non-insolvent companies. An accuracy of 85.34% confirms that the model correctly classified most instances. The F1 score of 0.8535 suggests a balanced trade-off between precision and recall, making the model well-suited for crisis prediction. The confusion matrix shows 234 true negatives, 271 true positives, and relatively few false positives (67) and false negatives (26). In these studies, false positives are crucial to carefully consider as they can lead to significant managerial and technical implications. This is because they may stem from algorithmic bias, but more importantly, they could represent errors due to the database or reflect real-world conditions. For instance, companies predicted as being in crisis may still be active and not yet initiated formal insolvency procedures. However, a deeper analysis of their financial statements and actual data could reveal underlying critical issues. This possibility highlights a key turning point in these analyses, making the confusion matrix an essential tool for understanding the accuracy and reliability of the model's predictions. It ensures that managers can identify potential risks early, even when formal signs of crisis have not yet emerged.

3.1. Model Explanation and Results Interpretability

As already stated, in this dissertation, the emphasis on model interpretability is twofold: first, to ensure that decision-makers and stakeholders can trust the predictions made by these models, and second, to provide actionable insights from the models' outcomes. This is where frameworks like SHAP play a crucial role. Furthermore, interpretability is critical for regulatory compliance and stakeholder trust. In situations of crisis management, where companies are making high-risk decisions, understanding the reasoning behind AI-driven predictions is vital. For these reasons, model explanation was implemented, and enhanced, using "*innovative*" methods, focusing more on SHAP, considered, in this context, simple and user-friendly. In the initial phase of our predictive analysis, *K-means clustering* was applied as an unsupervised ML technique (Abiodun et al., 2023). This method is essential for uncovering hidden patterns within the data. K-means clustering, in fact, allows to segment companies into distinct groups based on the similarity of their financial metrics, enabling to identify clusters of firms that exhibit similar financial behavior, such as those potentially on the brink of insolvency (Ikotun et al., 2023). From a predictive standpoint, clustering helps in identifying latent financial health categories, providing a foundational understanding of the dataset's structure. Specifically, it provides insight into how firms might naturally group themselves based on risk factors such as debt ratios, liquidity measures, and operational efficiency[147]. R Code used was basically this:

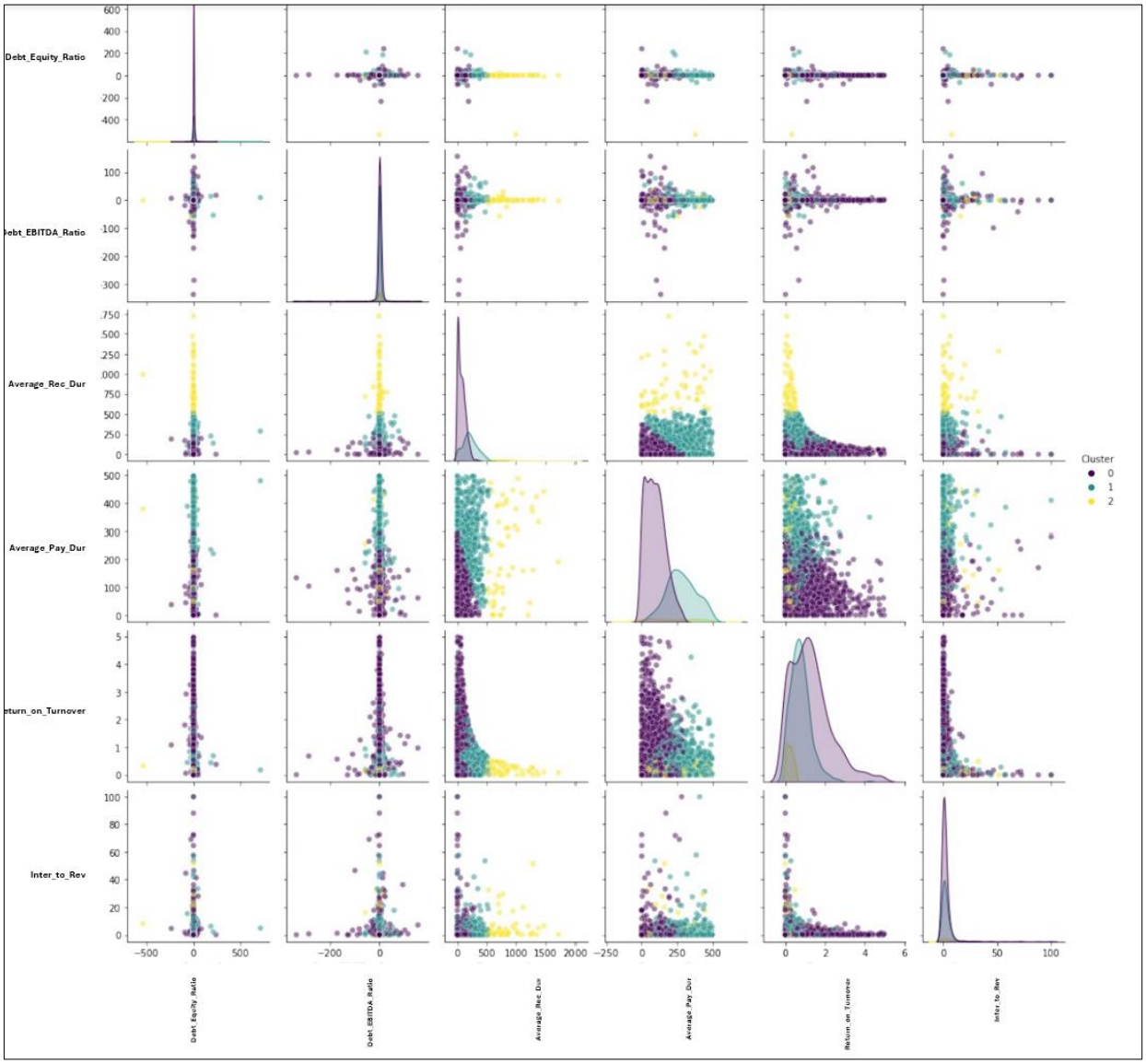
```
library(tidyverse)
library(cluster)

set.seed(123)
data <- select(dataframe, Debt_Equity_Ratio, Debt_EBITDA_Ratio, Current_Ratio,
Avg_Receivable_Duration, Avg_Payable_Duration, Return_on_Turnover, Interest_Exp_to_Revenue_Ratio)
scaled_data <- scale(data)
kmeans_result <- kmeans(scaled_data, centers = 3, nstart = 25)
dataframe$Cluster <- as.factor(kmeans_result$cluster)
fviz_cluster(kmeans_result, data = scaled_data, geom = "point", stand = FALSE, frame = TRUE)
print(kmeans_result$centers)[148]
```

Results are the following:

[147] The K-means algorithm minimizes the within-cluster variance by iteratively assigning data points to the nearest cluster centroids. Once clustering is complete, these groups can be analyzed to determine how closely they correlate with actual financial distress or insolvency, validating the approach's ability to segment companies according to risk.

[148] The number of clusters (k = 3) can be adjusted depending on the data structure and specific analysis goals.

Figure 25: K-Means Clustering Results



Source: own elaboration

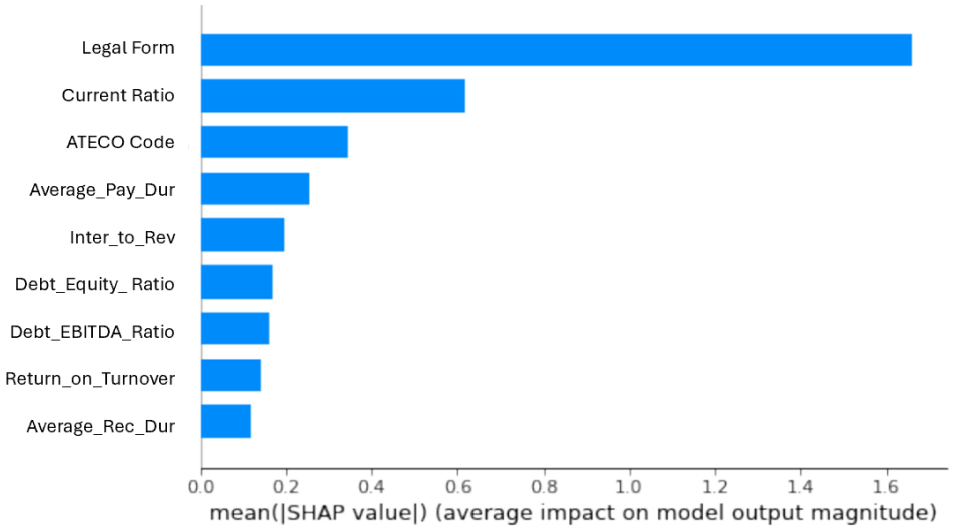
This pairplot provides a detailed view of the financial ratios used in the study, focusing on the clustering patterns derived from features such as Debt to Equity Ratio, Average Receivable Duration, and Interest to Revenue. The dataset is segmented into three clusters: Cluster 0 (purple), Cluster 1 (turquoise), and Cluster 2 (yellow). This clustering helps in uncovering key financial trends and insights related to the financial health of different companies. Each cluster's behavior shows noticeable divergence, particularly in areas like Debt to EBITDA Ratio and Return on Turnover, where outliers and trends are prominent. For example, companies with a higher Debt to Equity Ratio are spread more broadly in Cluster 2, which may indicate higher financial leverage and potential risk. Meanwhile, the clusters centered around Average Payable Duration show a concentration of firms with delayed payments, a known early indicator of financial distress. From a technical standpoint, the plot reveals underlying relationships and provides transparency into how the algorithm grouped companies, offering early detection of financial instability up to 1-3 years before the crisis event. This strengthens both decision-making and predictive accuracy, as managers and stakeholders can leverage these

explainable AI outputs to take timely, informed actions. The combination of visual clustering and explainable AI also ensures that the results are accessible and actionable. However, clustering results in a more technical difficulty for practitioners and management. SHAP, on the other side, is considered more and more user-friendly. Because of this, to enhance explainability and interpretability for practitioners and stakehodlers, the SHAP code was integrated into the best results, derived from the Gradient Boosting algorithm[149]. In particular, the SHAP values were calculated to provide a detailed explanation of feature importance and their individual contributions to the model's predictions. Consequently, a graphic visualization of these results was made, following SHAP usual graphic methods[150]:

```
library(shapviz)
library(xgboost)

shap_values <- shapviz(gb_model, X_test, model_type = "xgboost")
sv_importance(shap_values, kind = "bar")
sv_summary(shap_values, kind = "beeswarm")
```

Following Nguyen et al. (2023) and Mahajan & Shukla (2023) examples, the following "*Feature importance graph*" was prepared (Figure 26), including economic-financial features and non economic-financial features (legal form & ATECO code; since it is important to state if these variables have an impact on prediction and on crisis itself):

Figure 26: Feature Importance in SHAP Values for Gradient Boosting Classifier



Source: own elaboration

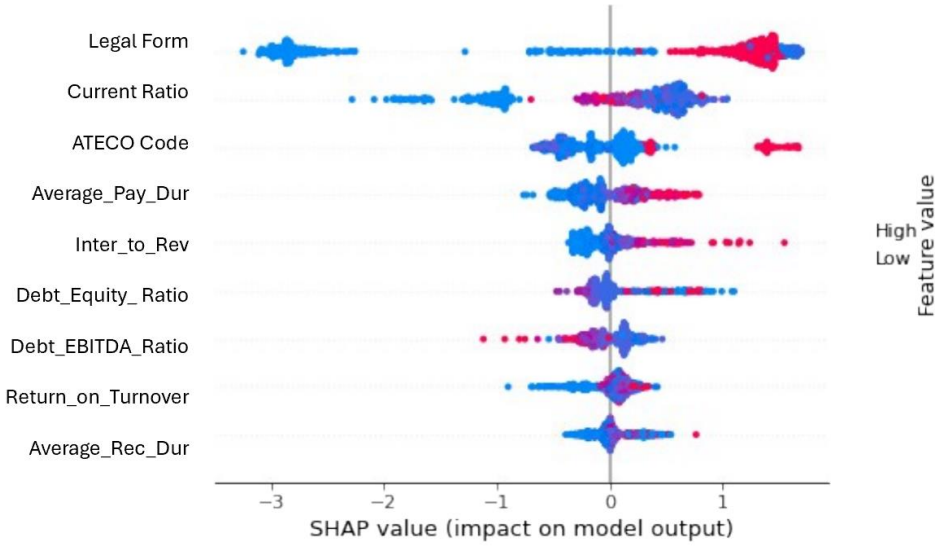
[149] By integrating SHAP with Gradient Boosting, we can break down the model's output into contributions from individual features, providing a transparent explanation for each prediction. Because of their nature, together, these tools create a powerful combination of predictive accuracy and explainability, making AI results more actionable for decision-makers.

[150] For detailed information regarding the package "*shapviz*" used, see Mayer & Stando's work at the following link: https://cran.r-project.org/web/packages/shapviz/index.html

Following this, the classical SHAP values graph (violin summary plot) was created in order to contextualize better the impact of each feature and to give important elements to understand better ML results and to enhance explainability out of a technical context (Figure 27)[151]:

```
sv <- shapviz(shap_values, X = X_test)
sv_summary(sv, "violin")
```

Figure 27: Violin Summary Plot[152]



Source: own elaboration

The SHAP plots provided here offer deep insights into the key financial indicators that affect the prediction in this exploratory analysis. From a technical perspective, these graphs allow for the visualization of the contribution of each feature to the model's prediction, while also demonstrating the consistency of SHAP in providing both local (individual) and global (overall) explanations for model outputs:

- The violin summary plot showcases how each feature impacts the model's output. Positive SHAP values indicate that the feature is pushing the prediction towards insolvency (or crisis), while negative values imply the opposite. In particular, variables such as Legal Form and Current Ratio have significant impacts. Regarding Legal Form, the use of frequency encoding in this study plays a crucial role. Since the descriptive analysis revealed that the majority of companies in the sample are small-sized firms with a narrow ownership base and limited

[151] A Violin Summary Plot is a data visualization technique that combines elements of a box plot and a kernel density plot, designed to summarize the distribution of a continuous variable across different groups or categories. This plot features a symmetrical, violin-like shape for each group, where the width of the shape at any given value represents the probability density of observations at that value, as estimated by a Kernel Density Estimate (KDE). KDE is a non-parametric method that smooths the observed data into a continuous curve, allowing the plot to reveal not only central tendencies but also the overall shape, spread, and multimodality of the distribution. Violin plots provide additional layers of information compared to standard box plots by displaying the density of values across the entire range of data, showing where values are concentrated or sparse. Typically, each *"violin"* is mirrored across a central axis to create a symmetrical appearance, with the most expansive parts indicating areas of higher data density.

[152] Hight Feature Value in Red; Low in Blue.

decision-makers (leading to high frequency in encoding), the elevated SHAP values for Legal Form indicate that firms with restricted ownership structures are more prone to crisis. This aligns with the prediction results, suggesting that for companies with fewer decision-makers and more concentrated control, there are earlier and clearer signals of distress. The SHAP values point out that these firms tend to display significant indicators of potential financial issues 1 to 3 years before the actual crisis becomes apparent. As for the Current Ratio, which measures liquidity, the SHAP values similarly highlight its strong influence in the model's predictions. A lower Current Ratio, often signaling potential liquidity problems, emerges as an early predictor of financial distress. The SHAP analysis reveals that companies struggling to cover their short-term liabilities with current assets exhibit a higher likelihood of entering a crisis phase within the subsequent 1 to 3 years. This aligns well with both financial theory and the empirical results of this study, emphasizing the importance of monitoring liquidity ratios to identify early warning signs and mitigate the risk of insolvency. Furthermore, the visual clarity provided by SHAP ensures that even for decision-makers without technical expertise, they can understand which factors are driving the model's predictions. The dispersion and concentration of certain features—like Debt to Equity Ratio or Average Payable Duration—further reveal clusters of firms that may exhibit early warning signs of financial distress;

- The bar chart ranks features based on their average contribution to the model's output. Legal Form emerges as the most critical factor, followed by Current Ratio and ATECO Code, indicating the relevance of both legal structure and liquidity. Interestingly, Return on Turnover and Average Receivable Duration, though typically strong financial metrics, show less importance in this specific context. The clear ranking helps managers prioritize which features to monitor closely when predicting financial instability.

3.2. Conclusion and Remarkables

This exploratory study set out to assess the predictive capabilities of various AI-based algorithms in identifying early warning signals of financial distress or insolvency in a limited sample of Italian firms. The research aimed to establish whether these algorithms—Random Forest, Gradient Boosting, and XGBoost—could effectively detect signs of crisis one to three years before any formal insolvency procedures or crisis events were declared. This focus on early detection aligns with the growing academic and practical understanding that corporate crises are rarely sudden, instead representing gradual processes that evolve over time. With this approach, the study sought to identify patterns and early signs that could signal impending trouble long before it becomes visible through traditional financial metrics or external events. Results showed high predictive accuracy across all models, particularly with Gradient Boosting, which outperformed the others. These findings underline the value of AI-driven predictive tools in crisis management, as they can detect early distress signals, such as liquidity shortages or increased debt levels, which often precede formal insolvency. Many of the firms flagged by the models as "*at risk*" showed early signs of distress in their financial data from 2015, yet did not formally enter insolvency procedures until the 2016-2018 period. This highlights the predictive power of the models and their ability to provide foresight, giving companies the ability to implement timely corrective actions.

However, a key takeaway from this study is that, despite the presence of these early warning signs, many firms fail to take proactive measures. The literature, as well as the practical reality, suggests

that while many companies exhibit early indicators of distress, they often lack the necessary skills, resources, or decision-making agility to respond effectively. As a result, they frequently enter formal insolvency procedures or face liquidation because they are unprepared to deal with the crisis as it escalates. The absence of forward-looking crisis management often leads to unnecessary business failures, and this study seeks to highlight the importance of proactive engagement with AI-driven tools that can help mitigate such risks. Furthermore, one of the most significant contributions of this study is the emphasis on explainability and transparency, particularly through the use of SHAP. Its values were integrated into the best-performing algorithm, allowing for an interpretation of the models that is both understandable and actionable for non-technical stakeholders. The SHAP analysis revealed key features that had the most significant impact on the model's predictions, such as Legal Form and Current Ratio:

- By applying frequency encoding, the model revealed that companies with a restricted ownership base were more prone to financial distress. These firms often lack the flexibility and resilience needed in periods of financial instability, and the elevated SHAP values for legal form highlight the importance of ownership structure as a key predictor of crisis. This is a critical insight for managers and policymakers alike, as it underscores the need for greater oversight and proactive risk management in closely held companies;

- The SHAP values also showed the Current Ratio to be a significant predictor. Firms with lower liquidity ratios—those struggling to meet short-term liabilities—are flagged as higher-risk entities. This aligns with both financial theory and the empirical findings of the study, emphasizing the importance of liquidity management in preventing insolvency.

From a scientific perspective, this empirical step contributes to the advancement of predictive modeling in the context of corporate insolvency: Random Forest was selected as a starting point due to its simplicity and wide adoption in literature, while Gradient Boosting and XGBoost added layers of complexity and precision to the analysis. The high AUC and F1 scores across all models reaffirm the reliability of the results, offering robust evidence that ML models can significantly improve the accuracy of distress predictions. Yet, beyond this technical achievement, the real value of this research lies in its practical implications for managers and practitioners. AI and ML offer more than just theoretical insights—they provide tangible, actionable tools that businesses can use to enhance their decision-making processes.

4. Research Design for Core Empirical Analysis: an Introduction

The preliminary analysis outlined in the previous sections was designed with two primary objectives: first, to establish a foundational base of considerations necessary for understanding the core empirical analysis that follows, and second, to offer key insights into the technical and theoretical approaches employed throughout the research. While it may include minor simplifications, its primary goal was to set the stage for the more comprehensive and advanced steps that will be presented next. Moreover, this initial phase provided a context grounded in both literature and technical elements, while also introducing the fundamental empirical phases that have been carried out: it has served to define the framework within which the following core analysis will be situated. Mainly, the purpose

was to generate a preliminary understanding of how ML algorithms function in crisis prediction, creating a controlled environment to test the early-stage predictive capabilities of these models.

As the flow now moves into the central part of the research, many elements introduced earlier will be considered as "*assumed knowledge*". This means that aspects such as data preprocessing, optimization techniques, and metrics will not be revisited in detail, as they were thoroughly addressed in the previous sections. Instead, the focus will shift toward the specific characteristics and innovations of this core study, such as the modifications made to the dataset and operational methodology, as well as the refined implementation of predictive systems and the interpretation of their results. In particular, the core analysis will place a stronger emphasis on overcoming the challenges posed by large datasets and the complexity of handling numerous variables from Italian firms. These challenges introduce new layers of difficulty that were not fully explored in the preliminary analysis: as a result, these paragraphs will highlight how these predictive systems and analytics can be adapted and optimized to handle the scale and intricacy of real-world data, determing, at the end, simple way to explain and understand results. So, at the end, analysis will focus into the interpretability and explainability, especially with the increasing dataset size. The iterative approach, which was initiated in the exploratory phase, will now evolve to tackle the broader dataset and variables, creating an important and fundamental parallel between scientific studies and real-world managerial practices, explicitly addressing critical issues that are often left underexplored in the existing literature. This parallel ensures that the research is not only theoretically sound but also highly relevant to practical applications.

4.1. Setting the Stage

As stated in the previous Chapters and paragraphs, today's literature about predictive analytics focus more on the quality and the number of data, given that it is possible to have high accuracy in predictions only having an accurate, well refined and well-adherent to reality dataset. So, for the core analysis of this dissertation, an extended dataset of Italian firms was used, sourced from the AIDA database, with an expanded timeframe: the dataset covers six total years of financial statements, starting from the most recent available financial year and going back five years. Furthermore, in line with recent literature that emphasizes the importance of a higher number of features to improve prediction accuracy, the analysis also employed a broader range of variables. The multi-algorithmic approach is here replicated, following an inductive and data-driven pathway, with the addition of the Logistic Regression algorithm: this brought the total to four AI-based algorithms used in the core analysis, allowing for a broader comparison of their performance and predictive accuracy.

Moreover, following the exploratory analysis discussed earlier, it became evident that the algorithms have the ability to conduct dynamic analyses over multiple years. These years are here treated as relative, rather than absolute: this means that the goal is to fully understand the type, characteristics, and timelines of crises and insolvencies without relying on fixed years (e.g., starting from 2018 and going backward). Considering absolute years would limit the sample size and the significance of the analysis, as specific years are often influenced by external macroeconomic dynamics. By treating the years as relative, the analysis captures a more comprehensive picture of crisis progression over time, enhancing the depth and generalizability of the findings[153]. Finally, the dataset

[153] In this study, given the heterogeneity of the companies involved, the extensive nature of the dataset, and the meticulous refinement processes undertaken during its preparation, it was decided not to account for the effects of the covid-19 pandemic. This decision is also tied to practical and technical reasons, primarily concerning the availability of data from the past ten years within the dataset. However, this approach aligns with methodologies found in the relevant literature, where similar predictive studies often exclude exceptional events that might skew the general trends being analyzed. In particular, by not explicitly

was meticulously composed by downloading financial data from companies of various sizes, all aligned with the size categories defined by the European Commission's Directive (EU) 2023/2775. This ensured that the it was heterogeneous yet, crucially, balanced between companies in crisis and healthy firms across each size class: as already stated before, maintaining this balance is fundamental for achieving high levels of accuracy, as well as scientific and practical relevance. The entire dataset was then analyzed using advanced and appropriate statistical techniques, ensuring robust and reliable findings. Additionally, a manual review was incorporated into the process, as certain biases—though replicable in statistical systems—can often be more effectively detected through human oversight: this combination of automated and manual approaches is applied to get a higher quality of data interpretation and minimizes the risk of overlooking critical patterns that may be obscured in purely automated processes.

The obtained results were subsequently analyzed through both qualitative and quantitative lenses, explicitly linking the findings to real-world decision-making processes in crisis or potential crisis situations. As outlined in the objectives of this dissertation, significant attention was devoted to explaining and representing these results to enhance explainability, interpretability, and practical application. Moreover, the empirical study emphasize the role of managerial literacy in AI and the importance of translating best practices into formalized business processes: these must not only automated through AI support but must also deeply connected to human intuition, judgment, and emotional intelligence. This integration of technology and human decision-making could represents a key pathway toward improving long-term business sustainability and resilience, enhancing Italian firms' skills to mantain going concern and to keep good relationships with stakeholders (based on trust, transparency and technicality).

4.2. Data Collection and Feature Selection

Data were meticulously collected from the AIDA database, following a rigorous process similar to the one used in the preliminary exploratory phase. The collection started by focusing on firms identified as being in serious crisis or insolvency, categorized across different size segments to ensure a comprehensive representation of the Italian corporate landscape: the entire dataset was stratified based on the four main categories of firms—micro, small, medium, and large enterprises—by adhering to the European Union's standard classification criteria, primarily based on two fundamental indicators: annual turnover and the number of employees, as specified by Directive (EU) 2023/2775. However, a critical observation emerged during the initial data search on AIDA: there was a significant imbalance in the number of available data points across different firm sizes, with micro and small enterprises being significantly more represented than medium and large enterprises: this phenomenon is not only reflective of the Italian industrial fabric but also highlights an important issue. In fact, micro and small enterprises dominate the Italian business landscape and, moreover, are more prone to experiencing financial distress or crises compared to their larger counterparts. Larger companies, on the other hand, tend to have more robust internal systems, better crisis management mechanisms, and are more resilient, which often allows them to overcome critical situations without entering formal insolvency

incorporating the pandemic's effects, the study aims to emphasize other factors related to decision-making in crisis situations that are independent of the pandemic, thereby focusing on more standard crisis indicators and patterns that persist beyond such an extraordinary event. Furthermore, considering that this research was conducted in 2024, the effects of the pandemic, as in other similar studies, are assumed to have largely dissipated, meaning that the long-term impacts have been resolved or normalized in most cases.

procedures. Consequently, the dataset revealed that the incidence of insolvency in micro and small enterprises was disproportionately high, further validating the need for precise prediction models tailored to this segment. Given this imbalance, it became imperative to ensure that the final dataset was balanced between firms in crisis and healthy firms to avoid skewed results and ensure that the models could provide accurate predictions. So, the dataset was structured to include a 1:1 ratio of firms in crisis and healthy firms within each size category (micro, small, medium, and large): to achieve this, propensity score matching code was employed. For in-crisis/insolvent firms, the following specific criteria and parameters were employed in research:

- *Legal Form:* All available legal forms with financial statements were considered. These included various corporate structures such as S.R.L (Limited Liability Company), S.R.L a socio unico (Single-member Limited Liability Company), S.R.L semplificata (Simplified Limited Liability Company), società consortile a responsabilità limitata (Limited Liability Consortium Company), S.P.A (Joint Stock Company), S.P.A a socio unico (Single-member Joint Stock Company), S.C.A.R.L.P.A (Limited Partnership Company), S.A.S (Limited Partnership), and S.C.A.R.L (Cooperative Company with Limited Liability)[154];

- *Number of emplyees & Revenue:* These elements change based on firms type (according to the European Union Directive (EU) 2023/2775, companies are classified into four categories: micro, small, medium, and large. Micro enterprises are defined as those with fewer than 10 employees and an annual revenue or total assets below €2 million. Small enterprises have fewer than 50 employees, with an annual revenue or total assets not exceeding €10 million. Medium-sized enterprises are characterized by fewer than 250 employees and an annual revenue below €50 million or total assets under €43 million. Large enterprises exceed these thresholds and are generally more resilient to economic shocks due to their scale and resources)[155];

- *Legal Events:* The dataset included companies that had undergone specific legal events indicative of crisis or insolvency in the last year available. These events included: bankruptcy or liquidation due to bankruptcy, insolvency status, judicial liquidation, dissolution and liquidation, bankruptcy arrangement (concordato fallimentare), and other forms of liquidation. These statuses reflect the actual insolvency or serious crisis-related condition of the firms, making them suitable for this predictive analysis[156];

- *Timeframe:* For this test, unlike the basic analysis conducted earlier, a more extended timeframe of six years was considered, meaning the last available balance sheet year and data going back five years (-5 from the current (and last) year). This approach aligns with current literature, which emphasizes the importance of including a comprehensive yet focused timeframe for crisis prediction. By limiting the analysis to five years before the latest available year, the study avoids the risk of introducing excessive noise into the results, which could distort the predictive accuracy of the models: in fact, years beyond this range are more likely to introduce outdated or irrelevant financial data, reducing the reliability of predictions. Furhermore, a key distinction of this study, compared to the previous analysis, is the use of relative years (-1; -2; - n …) rather

[154] Listed and unlisted companies.
[155] These indicators were considered "*locked*" for the entire timeframe studied, meaning that these elements remained consistent across all the years analyzed to ensure proper classification of firms within their respective size categories.
[156] As seen in the exploratory part.

than absolute years (2022; 2021; n …). This approach meets two critical needs: ensuring a greater availability of data and allowing for a dynamic, iterative analysis. In fact, by not confining the study to specific calendar years, the analysis avoids potential limitations posed by macroeconomic events that could distort the results. Instead, this methodology allows for a more comprehensive understanding of how Italian companies behave, generally, in periods of financial distress. The flexibility in using relative years also enables to assess patterns over a broader timeframe, without being constrained by particular events, and ensures a deeper exploration of the systemic behavior of firms in crisis; therefore, the six years considered in the study may be distributed in various ways across the 10-year period offered by the database, further enriching the analysis with diverse temporal insights[157];

- *Industry Sector:* All sectors were included in the analysis, classified according to the ATECO 2007 industry code system, which ensures a comprehensive view across the Italian business landscape[158];

- *Balance sheet availability:* For each of the six years considered. In fact, for the sake of data completeness and to avoid any missing information within the database that could potentially compromise the results, careful attention was paid to ensure that all available financial statements were collected for each of the years considered. This process of verifying the completeness of the dataset is critical, in order to maintain the integrity and robustness of the results, allowing for more reliable insights into the financial health and crisis dynamics of the firms examined.

At the conclusion of the implementation of this research strategy, the following numbers were obtained (given the limited number of large, medium-sized enterprises, and small, they were considered in their entirety. On the other hand, due to the high number of micro enterprises, a statistically significant sample was randomly selected. Care was taken to avoid any alphabetical order, to minimize the unlikely yet possible bias associated with the first letter of the company name. Similarly, the financial data and employee numbers were mixed without any specific order to ensure randomness and prevent any systematic biases): 59 large enterprises; 143 medium-sized enterprises; 798 small enterprises; 4.400 micro-enterprises. A total of 5.400 in-crisis firms was then downloaded[159]. For the selection of the sample of healthy and active companies, the same research criteria were applied across all size categories, with two key exceptions: legal event and legal status. In the first case, only companies that had not undergone any procedures related to crisis or insolvency, or experienced any related events, were selected. In the second case, the sample was further refined by excluding inactive companies, ensuring that only active enterprises were included in the analysis. Furthermore, to ensure balanced samples, the number of healthy companies in each size category was based on the number of companies in crisis for the same category. This means that the numbers are equal across all size

[157] Allowing the six years to be distributed variably across the 10-year database enriches the analysis with diverse temporal insights, and, this approach, is a fundamental feature of new intelligent prediction systems. It ensures that the predictive model is not rigid but is instead a dynamic tool capable of providing valuable, context-specific insights that enhance academic research and practical decision-making.

[158] In this context, conducting an analysis by industry sector would have been limiting, as it would not have achieved statistically significant numbers for certain sectors. Moreover, only a few industry sectors, such as agriculture, commerce, and construction, would have been more populated compared to others. This imbalance would have skewed the results and potentially reduced the robustness of the findings.

[159] Located on the entire Italian context.

categories. This balance results in a well-structured dataset, essential for achieving more accurate and reliable results with ML algorithms (50% active/healthy companies labeled as 0; 50% companies in crisis/insolvent labeled as 1). A key scientific observation from the AIDA database is that the number of active and healthy companies is significantly larger than the number of companies in crisis. This considerable discrepancy highlights the rationale behind using the number of firms in crisis as the proxy for balancing the dataset.

Regarding the feature selection for training the algorithm and proceeding with the predictive analysis, given the extensive availability of economic and financial data in the AIDA database, a broad set of variables was considered: these include variables that are well-established in the reference literature, metrics commonly used in practice to assess company performance, and a multitude of granular economic and financial data extracted from the financial statements of the firms under consideration (Nguyen et al., 2023). These statements were reported to the provincial chambers of commerce and follow strict legal and verifiability standards. Furthermore, in the downloaded Excel and CSV files, additional non-financial features were also included, such as the legal form, ATECO code, and province of registration: these were added primarily for statistical and informational purposes. Each economic and financial variable was also captured for each year in the time frame from the most recent available year to five years prior, ensuring a rich dataset with temporal depth. Here is the list of features extracted from the dataset (Table 6)[160]:

[160] For each feature, the data was reviewed and adjusted to the appropriate units of measurement (thousands or millions of Euros, percentages, or standard numbers).

Table 6: Core Analysis Selected Features

Selected Features	
Eco-fin	**Non Eco-fin**
Revenues from sales and services	Company name
Liquidity ratio	Province
Current ratio	Registered office address - Region
Current liabilities/Tot ass. %	Incorporation year
Long/med term liab/Tot ass. %	ATECO 2007 code
Tang. fixed ass./Share funds %	Number of current directors & managers
Depr./Tang. fixed assets %	Legal status
Leverage	Legal form
Coverage of fixed assets %	Number of advisors
Banks/Turnover (%)	Number of employees
Cost of debt (%)	
Interest/Operating profit %	
Interest/Turnover (%)	
Solvency ratio (%)	
Share funds/Liabilities %	
Net Financial Position	
Debt/Equity ratio %	
Debt/EBITDA ratio %	
Total assets turnover (times)	
Working cap. turnover (times)	
Operating working capital impact (%)	
Stocks/Turnover (days)	
Stocks/Cost goods sold (days)	
Average duration of receivables including VAT (days)	
Average duration of debts including VAT (days)	
Commercial cycle duration (days)	
EBITDA	
EBITDA/Sales (%)	
Return on asset (ROA) (%)	
Return on investment (ROI) (%)	
Return on sales (ROS) (%)	
Return on equity (ROE) (%)	
Net P&L / Operating P&L (%)	
Turnover per employee EUR	
Added value per employee	
Staff Costs per employee	
Turnover/Staff Costs	
Net working capital	
Gross profit	
Net short term assets	
Share funds - Fixed assets	
Cash Flow	
Profit (Loss)	
Total Production Costs	
Total depreciation, amortization and writedowns	
Operating Margin	
Total financial income and charges	
Total inventories	
Total receivables	
Total liquid funds	
Total shareholders' funds	
Capital stock	
Total provisions for risks and charges	
Total payables	
Total financial fixed assets	
Total current assets	
Total personnel costs	
Wages and salaries	
Current taxes	

Source: own elaboration

A total of 59 economic and financial features was selected; on the other side, 10 non eco-fin features were considered. The first group consisted of patrimonial (related to assets, liabilities, and shareholders' equity), financial (liquidity ratios, debt ratios, and interest coverage), margin-based (such as EBITDA and operating margin), and performance metrics, and the number of features, 59, aligns with the literature that suggests higher predictive accuracy in ML models when using a set of features close to 60 (Chatterjee & Kar, 2024). As said before, each eco-fin feature was taken for a 6 year timeframe, resulting in a total of 354 data points per company.

4.3. Data Preprocessing and Optimization

For this analysis, similar to the previous one, in-depth preprocessing and optimization of the database were carried out to eliminate noise, missing values, and potential outliers. Frequency encoding was once again applied for certain non-economic and financial variables to maintain consistency and reduce bias in categorical data representation. The data was then carefully processed using specific R packages, such as *dplyr* and *tidyverse*, to ensure optimal data cleaning and transformation[161]. Additionally, a manual review phase was incorporated into the process: this step was fundamental for several technical and scientific reasons. First, manual inspection enhances the reliability of the data by identifying anomalies that automated processes may overlook, especially in cases where unusual data patterns might still represent significant business insights; second, considering the large number of data points (354 features per company over multiple years), manual validation ensures that errors in the dataset, such as duplicated records or invalid entries, do not propagate and affect the final analysis.

4.3.1. A Critical Consideration

The manual control, combined with a numerical frequency check of the values present in the chosen dataset, led to a critical and practical consideration for the accuracy, explainability, and concreteness of the entire study. Specifically, it was observed that, for some companies in the sample of firms "*in-crisis*," certain key feature values, such as revenue per year and other financial indicators, were reported as zero (0) from the most recent year available up to several years before. This observation aligns with the core structure of our dataset, which is composed of "*healthy*" companies—those considered active and stable, as they are not involved in any insolvency procedures—and "*insolvent*" or "*in-crisis*" companies, that have either entered into formal insolvency proceedings, such as an arrangement with creditors or preventive agreements, or are still active but are financially unstable, potentially facing liquidation or asset disbursement. Crisis, by its nature, is a gradual process that deteriorates a company's performance and indicators over time. This degradation doesn't manifest abruptly, nor does it necessarily cause the company to cease operations at a precise moment, except for extreme events like the covid-19 pandemic. So, the crisis may last for many years, varying from firm to firm, and could lead to improvements in performance or insolvency and judicial liquidation (formerly bankruptcy under Italian law). In conclusion, the process could also have started earlier than the five-year timeframe we consider.

[161] *dplyr* is part of the tidyverse collection and offers a set of straightforward, efficient, and consistent functions to transform and summarize data; *tidyverse*, on the other hand, is an ecosystem of R packages designed for data science, which includes dplyr alongside other important packages like ggplot2 (for data visualization) and tidyr (for tidying data). Together, these packages provide a comprehensive toolkit for importing, cleaning, transforming, and visualizing data, making them especially useful for large datasets, ensuring a more structured and consistent approach to data management.

This study aims to determine whether predictive ML techniques can capture early warning signals or alerts years before the crisis or insolvency becomes evident, and these signals might indicate an impending crisis while the company is still active and operational (either unaware or not yet experiencing the full extent of the crisis). Thus, in our dataset—spanning the last six years of available financial data—there may be companies that started liquidation or ceased core business operations (e.g., sales, investments) several years ago but remain active to manage insolvency procedures, liquidate assets, or pay creditors. These firms might display zero values in key financial features (e.g., revenue, liquidity ratio, EBITDA, Return on Investment) from the most recent year to some earlier years within the five-year window, and this technical, yet practical, situation, could lead to less accurate or potentially skewed predictions due to the presence of companies that are technically still active (and considered "*healthy*" in the database) but in a clear state of crisis, where only activities related to liquidation or creditor payments are occurring. This finding is interesting because it reflects real-world business conditions, raising relevant considerations about specific features and the broader assumptions and implications of predictive studies on crisis and corporate insolvency. However, to address this challenge, one solution might be to filter the original sample of companies and reduce it to a new dataset that retains the same features but includes fewer companies. The eliminated firms would be those with a relevant number of zero values in certain key financial features, such as revenue, liquidity, and others, during years -1, -2, or -3, for example. This would remove companies that may have entered a crisis long before (e.g., eight years ago) and that now present severely deteriorated performance: these are firms that, while still active, are clearly in an advanced state of crisis. So, in the setting stage of the study, this reasoning supported the idea that a reduced dataset would likely improve prediction accuracy, aligning the results more closely with literature estimates. It also underscored the inevitable and necessary distinctions between crisis and active companies when explaining the study's qualitative aspects:in fact, extreme results could raise concerns about dataset errors or biases.

In conclusion, while a large dataset may initially appear to lend credibility and scientific relevance to a study, it almost certainly introduces more noise, as it is often less refined and carefully curated. Instead, a mixed approach that combines manual oversight with statistical analysis allows researchers to identify patterns in the data that could introduce noise and result in distorted, inaccurate findings. This is highly significant from a scientific perspective, as few studies in the predictive and managerial literature emphasize the dangers of over-reliance on large, pre-built datasets. Indeed, while such datasets may offer a wealth of structured data, they often require much deeper scrutiny beyond basic descriptive analyses. In the context of this dissertation, the critical approach merges scientific literature with real-world business insights, particularly the observation that companies can remain in latent crisis for many years before external conditions force them to disclose their struggles. This reality is compounded by the fact that managers—particularly of smaller firms—often lack the skills, competencies, or willingness to proactively address challenges. Instead, they may adopt a conservative or hidden approach to crisis management, which typically leads to negative outcomes for the company, society, and stakeholders: this is a common error in practice, and its consequences reverberate through the socio-economic fabric of entire regions or geographical areas.

Thus, as demonstrated in this study, real-world crisis management should prioritize a proactive and preventative approach, recognizing crises or critical situations as natural and almost inevitable aspects of a company's lifecycle. In fact, when addressed with timely interventions, these challenges can still ensure business continuity and, potentially, even create new opportunities. AI, along with ML-based preventive management systems, can play a crucial role in creating trust and transparency, addressing some of these challenges head-on. By providing data-driven insights and early warnings, these systems could offer managers the opportunity to intervene early, allowing for more effective

decision-making. This could significantly reduce the negative impacts of business distress, providing companies with the tools to overcome difficulties in a sustainable manner while improving trust across stakeholders. Ultimately, this approach—where crises are viewed as challenges to be managed rather than avoided—can transform business outcomes, ensuring both continuity and growth, while also reinforcing the value of AI-driven predictive systems as a means of supporting managerial decision-making in increasingly complex businesses.

Therefore, this reasoning was practically applied to the dataset, and the original sample of companies in crisis was resized to limit noise, data inconsistencies, and unreliable figures: only firms with complete data were considered, avoiding those that, while still active, showed clear signs of distress for 4 or 5 years (even if not formalized through specific events or procedures) that could have distorted the results and misled the algorithm. This approach was crucial in establishing not-so clear cut-off points or discriminating values to train the models. Thus, the final selection focused on companies that nonetheless had complete financial data, without a significant number of missing (*n.d. or N.A.)* or zero values in key financial features, such as revenue, liquidity ratios, and other indicators. To ensure a balanced dataset with an equal number of "0" (*healthy*) and "1" (*distressed*) labels, the number of active companies was adjusted accordingly, resulting in a more refined and reliable dataset for ML purposes. This guarantees that the algorithm's training process, at last, is based on accurate and meaningful data, leading to better prediction and analysis outcomes.

The final sample consists of 5.000 firms in crisis and 5.000 healthy firms (total 10.000)[162], with 59 economic-financial variables and 10 non-economic-financial variables per company. The total number of data points per company (in a 6 y. timeframe except for the non eco-fin features) is 364; the total feature number in the entire study is 3.640.000[163]. At the end of the entire data preprocessing and optimization phase, a correctly balanced dataset was obtained. "*Healthy*" firms are categorized as "0"; "*In-crisis*" ones as "1".

4.4. Descriptive Statistics and Correlation Analysis

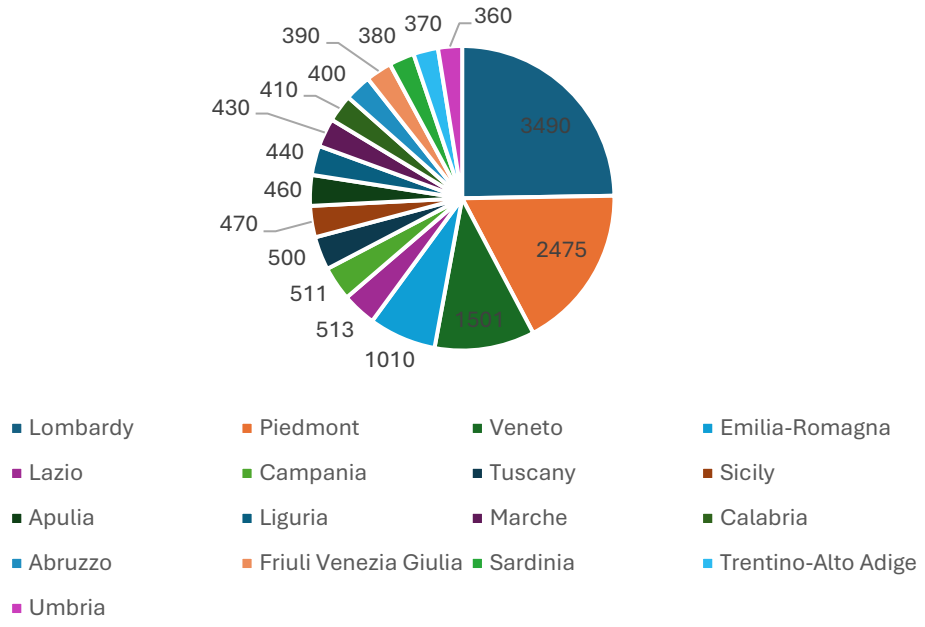
As in the previous analysis, a thorough descriptive statistical analysis was conducted to provide a detailed understanding of the dataset before applying any algorithms. Descriptive statistics were calculated for each of the 59 economic-financial variables across the 6-year timeframe for both sample companies. This analysis included measures such as mean, median, standard deviation, minimum and maximum values. Given the extremely high number of features in this phase of the study, the inclusion of exhaustive tables and detailed statistics has been omitted from this dissertation. However, the results showed that the values were appropriate for the study, with limited outliers and figures consistent with the respective units of measurement. The maximum, minimum, and average values aligned with those found in similar studies, indicating coherence and reliability within the dataset.

The dataset presents a rather heterogeneous distribution in terms of ATECO code and geographic origin. However, a majority of the companies, particularly micro-enterprises, are concentrated in northern Italian regions such as Lombardy (3.490) and Piedmont (2.475) (Figure 28).

[162] 59 big; 143 medium-sized; 698 small; 4.100 micro per group (0;1); -7,41% from the original, balanced, sample.

[163] The sheer volume of features allows the models to capture nuanced patterns and relationships between variables, leading to more accurate predictions and stronger generalizability. Moreover, the balanced nature of the data between healthy and insolvent firms mitigates the risk of bias and enhances the model's ability to make fair, representative decisions across different types of firms. Consequently, this large and well-structured dataset strengthens the study's predictive power, while also reinforcing the reliability of the results.

Figure 28: Distribution of Considered Companies by Region



Source: own elaboration

Lombardy accounts for the largest share, hosting 34.90% of the companies, followed by Piedmont with 24.75% and Veneto at 15.01%. Emilia-Romagna also holds a significant portion with 10.10% of the total. Moving southward, Lazio represents 5.13% of the companies, closely followed by Campania at 5.11%. Tuscany, with 5.00%, along with Sicily and Apulia, contributing 4.70% and 4.60% respectively, indicate moderate representation in central and southern Italy. Other regions like Liguria and Marche each account for around 4.4% and 4.3%. Calabria holds 4.10%, and Abruzzo follows closely with 4.00%. Friuli-Venezia Giulia contributes 3.90%, while Sardinia, Trentino-Alto Adige, and Umbria complete the distribution with 3.80%, 3.70%, and 3.60%, respectively.

The ATECO codes predominantly correspond to industrial and agricultural sectors, with fewer companies in the services sector. Additionally, larger and medium-sized enterprises are primarily located in the northern territories and the average age of active companies is approximately 12 years, consistent with estimates from the Chambers of Commerce, while the average age for companies in crisis is around 9 to 10 years, indicating a younger profile compared to healthy firms[164]. Notably, micro-enterprises in crisis tend to have significantly lower average ages. The dataset is then largely composed of small, non-listed companies, mirroring the typical structure of the Italian business landscape, where smaller enterprises prevail. S.R.L (Limited Liability Company) stands out as the dominant legal form, representing around 40% of the sample. This high prevalence underscores the popularity of S.R.L as a flexible and accessible structure, particularly appealing to small and medium-sized businesses that benefit from its limited liability protection without the complexities of public listing.

In addition to these basic statistics, correlation analysis was performed to explore the relationships between the selected features: this allows to identify multicollinearity among variables, which is crucial to avoid issues that could distort the predictions of used models. In particular, variables that are highly correlated (e.g., revenue and total assets) may introduce redundancy into the model,

[164] See *camcom.it* sites of each italian Region for more information.

leading to overfitting or bias. Therefore, identifying these correlations helped in deciding which variables to manage better for optimal model performance.

In this empirical context, among the features examined, no particularly high correlation values were observed between different variables, except for the expected strong correlations between the same feature across different years (e.g., "*Revenue from Sales and Services*" in the last year, and in the -1, -2, and -3 years, which exhibited correlations exceeding -0.75). A reasonable cutoff for considering high correlation would typically be around 0.7 or above, but in this case, the correlations between different features remained below this threshold. This outcome is consistent with the nature of the dataset and the variables: in fact, the features selected in this study encompass a wide variety of economic-financial indicators, each reflecting distinct aspects of a company's performance. Therefore, lower correlations between features are expected because they measure fundamentally different business aspects and this is scientifically significant as it suggests that the dataset offers a diverse and comprehensive representation of firm characteristics, reducing the risk of multicollinearity and ensuring that each feature contributes uniquely to the predictive power. Moreover, the strong correlations between the same feature across different years (e.g., revenue across multiple years) further validate the consistency of the dataset: these features tend to follow similar trends over time, which is natural in longitudinal financial data, where patterns of growth or decline are often persistent.

Another important step was that of evaluate the correlation between some key features and the "*in crisis/insolvent*" status of firms considered in the study. For this, the following R code was used:

```
correlations_with_insolvent <- sort(cor(correlation_matrix$Insolvent), decreasing = TRUE)
print(correlations_with_insolvent)
```

The importance of this step lies in understanding which feature have the strongest association with the likelihood of a firm being in a state of crisis: features such as "*Revenues from Sales and Services*," "*Operating Margin*," and "*Tang. fixed assets/Share funds*" are particularly critical in determining the probability of insolvency. For the micro-sized firms, that are the majority in the entire sample, results were the following:

```
correlations_with_insolvent <- c(
  "ATECO 2007 code" = 0.130,
  "Capital stock Year - 2" = -0.028,
  "Capital stock Last Year" = -0.032,
  "Operating Margin Year - 2" = -0.042,
  "Operating Margin Year - 3" = -0.108,
  "Tang. fixed ass./Share funds % Year - 2" = -0.102,
  "Operating Margin Year - 1" = -0.115,
  "Number of advisors" = -0.119,
  "Operating Margin Last Year" = -0.124,
  "Operating Margin Year - 4" = -0.145,
  "Tang. fixed ass./Share funds % Last Year" = -0.161,
  "Operating Margin Year - 5" = -0.170,
  "Number of current directors & managers" = -0.591,
  "Revenues from sales and services Year - 2" = -0.760,
  "Revenues from sales and services Year - 1" = -0.765,
  "Revenues from sales and services Last Year" = -0.784,
  "Revenues from sales and services Year - 5" = -0.611,
  "Revenues from sales and services Year - 3" = -0.720,
  "Revenues from sales and services Year - 4" = -0.710[165]
```

[165] These results show that, as the revenue from sales and services decreases, there is a higher likelihood that the company is in a state of crisis or insolvency. The strongest correlation is for Last Year, with a correlation of -0.784, indicating a highly significant inverse relationship between revenue decline and insolvency. This pattern suggests that revenues tend to deteriorate well before

In the correlation analysis, it is possible to notice that some financial metrics, particularly those related to revenues show a relatively strong negative correlation with insolvency. This means that as revenues decline (even after a preprocessing and optimization phase of the dataset), the likelihood of the firm being in crisis increases, which is consistent with the understanding that revenue reduction is a key indicator of financial distress. Correlation values such as -0.710 four years before crisis highlight a strong relationship between diminishing sales and the onset of insolvency. On the other hand, features like "*ATECO 2007 code*" show a relatively low correlation (0.130), suggesting that industry sector classification has some impact on the likelihood of insolvency but is not as strong a predictor as other financial variables like revenues or operating margins[166]. Variables such as "*Number of current directors & managers*" also show a moderately high negative correlation (-0.591), which may indicate that firms with fewer managerial resources tend to be more prone to financial distress[167]. This could imply that governance and leadership play a crucial role in maintaining a firm's financial health. Moreover, the fact that tangible fixed assets to share funds ratios and operating margins consistently show negative correlations, albeit lower than revenues, suggests that profitability and the effective use of fixed assets are also important factors in determining financial stability. The negative correlation of operating margins highlights that a decline in profitability, across multiple years, correlates with a higher likelihood of insolvency.

However, these correlations must be interpreted carefully. The strong correlations between similar metrics across different years (e.g., revenue in different years) are to be expected due to their temporal connection. Overall, this correlation analysis aligns with established financial theories, such as those that associate declining revenues and profitability with increasing insolvency risks and, moreover, it underscores the significance of a dynamic approach to insolvency prediction, where historical trends provide a clearer picture of a firm's trajectory into financial distress.

5. Multi-Algorithmic Approach Implementation and Model Evaluation

In this core section of the study, the multi-algorithmic approach is still implemented with the primary goal of refining and advancing the predictive capabilities of used models for crisis and insolvency prediction. This phase of the research builds on the methodological framework previously established, with a clear focus on dynamic, data-driven analysis. So, the objective is to push the boundaries of predictive accuracy while providing actionable insights that are directly applicable to both academic literature and real-world managerial practice. The ultimate aim is then twofold: to advance the state of the art in crisis prediction models and to equip decision-makers, particularly in small and medium-sized enterprises, with intelligent, data-driven tools to manage financial distress in

a company enters a formal insolvency or crisis situation. A steady decline in revenues over multiple years indicates financial instability and could serve as an early warning signal for management.

[166] This collides, in part, with previous results of SHAP in the exploratory analysis of this dissertation.

[167] From a scientific perspective, the number of directors and managers is often linked to the company's ability to handle complex decision-making, manage crises, and respond to external and internal shocks. Firms with a limited management structure may lack the diversity of expertise and the operational agility required to navigate challenges effectively. The negative correlation indicates that a lower number of managers and directors is associated with a greater risk of insolvency, highlighting the importance of managerial capacity in maintaining financial health. From a managerial perspective, this finding underscores the role of a strong and diverse leadership team in preventing financial distress. Companies with fewer leaders may suffer from a concentration of decision-making power or may not have sufficient bandwidth to manage various aspects of the business, such as financial oversight, strategic planning, and crisis management. This can lead to inefficiencies, delayed responses to warning signs, and ultimately, the inability to prevent or mitigate financial crises.

a timely and proactive manner. This approach well aligns with the main assumptions of this dissertation and with the broader goals set by both the Italian and European legislatures.

However, before starting with the technical procedures, it is possible to affirm that this analysis follows a similar inductive, data-driven methodology to that of the earlier exploratory study, but with a new layer of dynamism. In fact, the models, here, are more designed to analyze the influence, impact, and determinants of specific financial and economic features across a broader timeframe. So, unlike the initial study, this implementation incorporates a dynamic analysis of how certain variables evolve and interact over time, adding a temporal dimension that allows, well, for a deeper understanding of the triggers of financial distress. Moreover, as before, in addition to offering predictive insights, the algorithms utilized also well emphasize explainability and interpretability of results.

5.1. Some Remarks

The multi-algorithmic approach in this study, applied to an extensive dataset with a broader timeframe, is, obviously, scientifically expected to yield more accurate and refined results compared to the initial exploratory study, which was primarily focused on evaluating predictive margins for Italian companies. In contrast to the preliminary research, which provided an initial understanding of crisis prediction, this study leverages the larger dataset to generate fundamental insights, particularly through a comparative lens across companies of varying sizes, while focusing primarily on small and medium-sized enterprises: one key reason for focusing on SMEs is that ML algorithms require a minimum data volume, typically numbering in the thousands, to effectively train and produce reliable predictions (with an 80% training and 20% test sample split). By having such an extensive dataset, the ML models can be expected to train more thoroughly and make predictions theoretically accurate but also practically relevant for business continuity and crisis prevention. So, after evaluating the algorithmic performance, results, using both SHAP plots and confusion matrices to measure model accuracy and potential errors, will be analyzed. Specifically, confusion matrices provide a critical view of the model's performance, helping identify both false positives and false negatives, which are especially important from a managerial decision-making perspective. In our context, false positives—where the model incorrectly (?) predicts a company to be in crisis—are particularly noteworthy: although we are often inclined to trust AI-driven predictions, in fact, it is essential to maintain a scientific and critical approach, analyzing and reflecting on the implications and characteristics of each result carefully. Specifically, when dealing with false positives, predictive studies in general, and especially those focused on crisis and insolvency prediction, require rigorous contextualization and comprehension: crisis prediction studies rely heavily on data, and this necessitates precise, accurate, and comprehensive datasets, often involving large volumes of information.

However, it is now common practice in scientific research to depend on widely used and scientifically validated databases for the sake of optimizing both time and resources (Veganzones & Severin, 2018): a clear example is the database utilized in this dissertation—AIDA by Bureau van Dijk, which is both industry-standard and academically validated. So, in this case, we deal with two distinct datasets of companies: one consisting of firms in crisis and the other of active, healthy companies. Yet, despite companies being classified as "*healthy*" in the database, an analysis of their financial statements may reveal underlying situations of crisis or distress, as previously discussed in this dissertation. If this occurs with a significant number of firms in the sample, the model's results could be skewed, particularly the false positives: firms that are "*active*" yet experiencing financial difficulties could distort

the results, leading to what we might call "*false false positives*.[168]" These firms, though active, are indeed in crisis and should have already initiated recovery processes or sought to ensure business continuity through crisis management strategies. Furthermore, this is compounded by the reality that some business owners might manipulate financial data to conceal signs of distress, aiming to avoid legal repercussions or maintain a favorable perception among stakeholders: this deceptive practice could further skew the results, making it even more critical to interpret the outputs from AI-driven models with caution and a deep understanding of the practical business realities.

One critical aspect, frequently observed in financial analyses, is the practice of financial data manipulation among small businesses: this manipulation often stems from both technical and practical motivations, which make it an established, if unethical, practice with substantial implications for data accuracy and the interpretability of ML models. From a technical perspective, small business owners may adjust their financial statements to mask indicators of distress, strategically altering variables such as debt levels, liquidity ratios, and revenue figures. This often takes advantage of gaps in the auditing and monitoring processes, which tend to be less rigorous for non-listed, small firms compared to publicly traded companies: these adjustments may be subtle, targeting specific line items that influence perceptions of financial health without triggering immediate regulatory scrutiny. Scientific evidence supports the existence of such practices, especially in small firms with limited governance oversight. These firms may engage in "*earnings smoothing*" or selective disclosure to present a more stable financial profile, even when internal economic pressures are significant. Practical motivations further drive this behavior. For many small business owners, manipulating financial data serves as a protective mechanism to avoid triggering insolvency proceedings or credit restrictions that could result from openly displaying financial instability: by maintaining the appearance of stability, owners aim to retain the trust of creditors, investors, suppliers, and employees, delaying or preventing interventions that could impose significant personal and operational burdens. Furthermore, a favorable financial presentation helps secure better terms for trade credit and banking relationships, which are often lifelines for small companies. However, this deceptive practice introduces substantial risks to data-driven analysis: for AI-driven models relying on financial metrics, the presence of manipulated data can lead to skewed predictions, overestimating the financial health of certain companies or underestimating the likelihood of distress. As a result, AI models trained on data that contains misrepresented financial information may inadvertently propagate inaccuracies, making it crucial for analysts and practitioners to account for potential biases introduced by such manipulations. All these

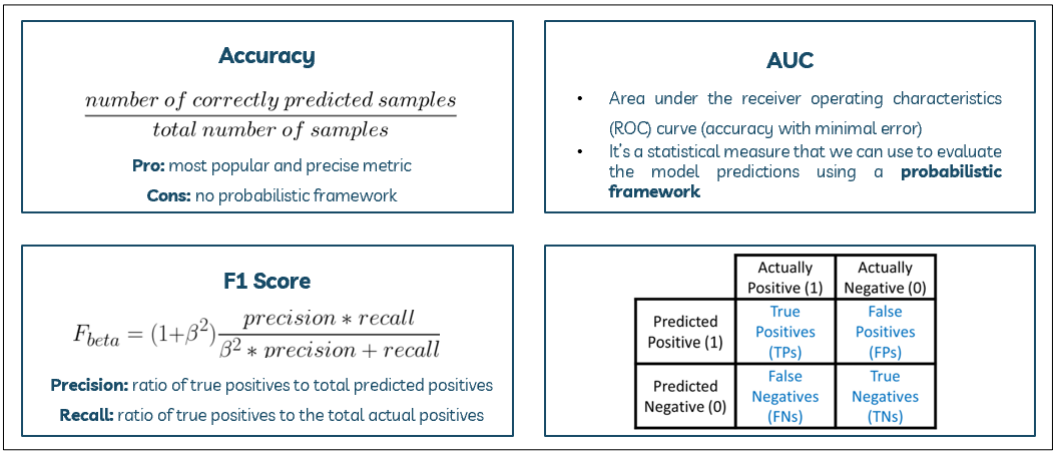
[168] This represents an important consideration in predictive analytics in general. In fact, as stated by Wedge et al. (2017) and the majority of literature about predictive studies, when dealing with a *"0;1"* results study, it is fundamental and even necessary to limit and reduce the number of false positive. This approach is based on some considerations that believe, sometimes wrongly, that results in confusion matrices need to be trust, and, just in some cases, refined. However, false positive, especially in predictive analytics about crises and insolvency, are, sometimes *"not false"* (Sandosham, 2024). In fact, in the context of crisis prediction, false positives are traditionally viewed as errors—instances where the model predicts a crisis that does not materialize. However, the nature of false positives in propensity models (like the one used in this study) offers a new perspective. Just because the state of *"crisis*" did not manifest as expected, it does not mean the prediction was entirely off: it may be that the company is showing early warning signs or subtle indicators of distress, but the crisis has not yet fully emerged, or the company managed to prevent the full manifestation of a crisis, perhaps through competitor intervention or managerial action. Thus, what the model flags as a false positive might actually reflect underlying issues that have yet to escalate to a visible crisis or may be hidden by managerial practices. In fact, as noted by Sandosham (2024), in some cases, false positives in predictive analytics are "*not false"*, because they still indicate an underlying propensity for crisis, even if the final outcome is not visible immediately. So, companies flagged as potentially at risk (if without managerial manipulation of data) may not be in visible distress but could benefit from proactive interventions. From an AI and data-driven standpoint, this highlights the need to interpret false positives not merely as errors but as valuable signals that the company might be approaching a tipping point. By intentionally generating more false positives, in fact, models can capture potential risks, offering opportunities for strategic interventions, and this approach is especially relevant in a competitive market, where understanding latent signals may provide a competitive edge.

considerations, underline the importance of critical analysis, even when using advanced predictive models, to avoid relying solely on seemingly objective data.

5.2. Starting Point and Models Deployment

Once the final dataset was refined and clearly defined, as discussed in previous sections, the next step involved deploying four AI-based algorithms using *R*: Logistic Regression, Random Forest, Gradient Boosting, and XGBoost. In addition to improving the predictive framework, this stage of the research introduces a more dynamic methodology that reflects real-world complexities: the capacity of these algorithms to adapt and refine their predictions based on iterative feedback and a large dataset provides invaluable insights into how financial crises unfold in real life, in different Italian contexts. The evaluation metrics employed in this core analysis are the same as those previously utilized and thoroughly explained in earlier sections, all of which have been extensively validated by existing literature. These include key metrics such as Accuracy, F1-Score, and AUC. For the sake of simplicity and clarity, and to maintain a smooth flow of the analysis, the following summary tables are provided for ease of understanding (Figure 29):

Figure 29: Evaluation Metrics Synthesis



Source: own elaboration

Following the deployment and evaluation of the Machine Learning models, SHAP was used as a tool for Explainable AI (XAI). So, the implementation process began with Logistic Regression, the simplest of the models, which is well-suited for binary classification. This was followed by Random Forest, which, though more complex, leverages ensemble learning by aggregating the outputs of multiple decision trees[169]. Once these foundational models were established, the analysis moved on to Gradient Boosting and XGBoost: this strategic order of implementation—progressing from simpler to more advanced models—ensures that the research provides a framework for comparative analysis, allowing for both transparent interpretation and the iterative maximization of predictive accuracy. In particular, for the multi-algorithmic approach, only the economic and financial data of the companies were used. Below, the results of the evaluation metrics and confusion matrices are presented, providing

[169] For this algorithm, the same settings as the previous exploratory analysis were used (n. of splits etc.).

insights into the accuracy, precision, recall, and overall performance of each model in the study (Table 7):

Table 7: Performance Assesment Using Classification Metrics of ML Models in Identifying in Crisis or Insolvent Italian Companies[170]

Model	Accuracy	F1 Score	AUC
Logistic Regression	0.64	0.68	0.63
Random Forest	0.81	0.83	0.80
Gradient Boosting	0.55	0.67	0.54
XGBoost	**0.87**	**0.88**	**0.87**

Source: own elaboration

Commenting in critical way these results, it is possible to affirm that:

- Logistic Regression, a classic linear model, delivers an accuracy of 0.64 and an AUC of 0.63, which is relatively low compared to other models. This is expected given the limitations of this model, which assumes linearity between independent variables and the log odds of the binary target variable (in-crisis vs. healthy). In datasets where the relationship between features is complex and non-linear, Logistic Regression struggles to capture these intricate patterns. Furthermore, the modest F1 score (0.68) shows that while it balances precision and recall reasonably well, its simplicity prevents it from excelling in more complex environments. So, from a technical perspective, Logistic Regression may underfit the data because it cannot capture non-linear interactions inherent in financial variables such as revenue, debts, and margins. On the other side, in a managerial context, the model's interpretability is a strength, as decision-makers can easily understand and explain the impact of each feature;

- Random Forest, on the other hand, significantly outperforms Logistic Regression with an accuracy of 0.81 and an AUC of 0.80. The key reason for this improvement could lie in Random Forest's ability to model non-linear interactions between features by combining the predictions of multiple decision trees, each trained on different subsets of the data: this allows the model to capture complex dependencies that linear models cannot handle. Moreover, Random Forest's method of aggregating predictions helps to mitigate overfitting, especially in high-dimensional datasets where there are numerous potential interactions among financial variables. The high F1 score of 0.83 then reflects the model's effectiveness at balancing precision and recall, making it a robust choice for detecting insolvency in firms. From a management perspective this algorithm not only provides strong predictive power but also offers insights into feature importance, helping decision-makers understand which financial metrics (e.g., long-term liabilities or current ratio) are most critical in predicting a company's likelihood of insolvency. This model, therefore, adds value both in terms of accuracy and explainability;

[170] Bold values indicate the model with the best overall performance.

- Gradient Boosting presents a surprisingly poor performance, with an accuracy of 0.55 and AUC of 0.54. Despite its theoretical advantages in building sequential models to correct previous errors, its performance in this context is hindered by several factors. In fact, it is known to be highly sensitive to hyperparameters, and suboptimal tuning can lead to underperformance, especially in datasets with a high degree of noise or non-linearity. Additionally, Gradient Boosting is prone to overfitting if regularization techniques are not applied effectively. This could explain why it performs worse than Random Forest despite its theoretical advantages in managing imbalanced data. From a managerial perspective, the poor performance of Gradient Boosting in this study suggests that managers need to exercise caution when using more complex models, especially if those models are not properly tuned. In real-world applications, improper configuration can lead to misleading results, emphasizing the need for rigorous validation before adopting any predictive tool for decision-making in critical situations;

- XGBoost exhibits the best performance across all metrics, with an accuracy of 0.87, F1 score of 0.88, and an AUC of 0.87. In fact, it benefits from advanced regularization techniques and parallel processing, making it faster and more effective at handling large datasets while reducing the risk of overfitting. So, XGBoost excels in this context due to its ability to efficiently manage missing data and handle imbalanced classes, which are common in insolvency prediction. Moreover, its regularization techniques help prevent the model from overfitting, especially when dealing with the wide range of economic-financial variables across multiple years in the dataset. For managers, and to enhance explainability and understandability of the models, XGBoost represents the best balance between predictive accuracy and computational efficiency. Based on these results, in conclusion, it is possible to affirm that its robustness and reliability make it ideal for real-world applications where predictive accuracy is critical, and make it ideal to be the base of the definition of an AI-based explanation tool for managers and stakeholders.

5.2.1. Comparison and Key Insights

The results of these models highlight several key takeaways:

- *Non-linearity and Feature Interactions*: Both Random Forest and XGBoost outperform simpler models like Logistic Regression due to their ability to handle non-linear interactions and complex feature dependencies. This demonstrates that, in financial datasets, where variables interact in non-linear ways, ensemble methods provide a clear advantage;

- *Practical Insights for Management*: Random Forest and XGBoost offer the best combination of predictive accuracy and insight into feature importance. These models, in fact, not only predict insolvency with high reliability but also allow managers to focus on critical financial metrics[171]. However, looking ahead, it is clear that the role of management will need to evolve. In fact, the

[171] The setting of these algorithms, particularly decision trees like, can be metaphorically compared to concrete internal processes that companies implement during times of crisis. For instance, decision trees simulate the decision-making pathways that a company's management might follow in evaluating various factors during a financial crisis: each split in the decision tree represents a critical decision point, much like how managers must assess key financial metrics, market conditions, or operational adjustments in real-time.

increasing reliance on AI and predictive models requires the emergence of algorithmic-oriented managers—leaders who are not only familiar with traditional financial metrics but also well-versed in algorithmic literacy[172]. These managers must understand how to interpret AI-generated insights and apply them in real-time to optimize strategies and mitigate risks. Developing this algorithmic mindset will be, then, essential, as AI becomes an integral tool in strategic planning;

- *Data Sufficiency and Model Requirements:* Logistic Regression's poor performance, as already stated, underscores the limitations of overly simplistic models when dealing with large, complex datasets. In practice and reality, this over-reliance on basic models can lead to significant inefficiencies, not just in scientific research but, more critically, in real-world businesses. In fact, when companies face financial distress, decision-makers often try to turn to predictive-like models to guide their strategy. However, if the underlying model is too simplistic, it may fail to identify early warning signs and, furthermore, the reliance on inaccurate or incomplete data, whether due to manipulations, poor data quality, or a lack of deeper financial understanding, further exacerbates this issue. In such cases, in fact, an erroneous interpretation of the data can lead to misinformed decisions that not only harm the immediate financial health of the business but also jeopardize long-term sustainability[173].

So, the analysis positions XGBoost as the most effective model for insolvency prediction, striking a balance between high predictive accuracy and computational efficiency. This robustness is achieved through advanced regularization techniques that mitigate overfitting. From both a scientific and practical standpoint, XGBoost holds significant importance. Scientifically, it offers a rigorous framework for insolvency prediction that advances methodological standards in financial analysis. Practically, it provides the predictive accuracy and interpretability crucial for real-world decision-making, helping to bridge the gap between academic research and actionable business insights. Furthermore, this model aligns particularly well with emerging regulatory standards, such as the AI Act of the European Union and the European Insolvency Directive, both of which underscore transparency, explainability, and accountability in AI applications impacting various kind of stability. Through its structured feature selection and regularization mechanisms, XGBoost can enhance explainability, which is critical for addressing ethical concerns in high-impact applications. These qualities make it an ideal candidate for building bespoke decision support systems that meet the regulatory expectations of modern AI ethics frameworks. Furthermore, it aligns with the IBDA Theory, which stresses the importance of data-driven insights that are both accurate and interpretable.

The key insights from these first empirical results not only raise important methodological concerns found in the literature but also offer crucial practical applications for real-world: a data-driven approach is undeniably essential in today's context; however, while academia and research institutions can develop highly sophisticated and refined models or frameworks, the managerial realm often lacks

[172] About this, see Jarrahi et al. (2023) and Nagtegaal (2020).

[173] For instance, if a company's financial reports do not fully reflect the actual state of the business—whether due to intentional misrepresentation, gaps in data, or incomplete knowledge—predictive models relying on such data could output misleading results. In critical situations, a poorly understood or misused dataset could lead management to overlook critical issues, implement ineffective strategies, or delay necessary interventions, ultimately compromising business continuity. The lesson for both researchers and managers is that simplistic models, while easy to implement, are often insufficient in these cases. More sophisticated models, such as Random Forest or XGBoost, are better equipped to handle higher-order interactions in large datasets. So, a solution to this issue could be to find a simplicistic way to represent, implement and evaluate technically difficult models.

either the intention or the skills necessary to effectively implement these advanced systems. This consideration highlights a critical gap that is frequently discussed but not always fully understood: the scientific community must strive not only to push the boundaries of knowledge and deliver advanced predictive models, but also to focus on developing pragmatic solutions. These solutions, in particular, should be grounded in reality and take into account the practical limitations and needs of businesses, especially in sectors where financial distress and crisis management are pressing concerns. This dissertation, in particular, as stated in each Chapter before, aims not just to advance the science of insolvency prediction, but also to bridge this critical gap. Consequently, the true value of predictive models and data-driven methodologies lies not just in their theoretical precision but in their ability to solve real-world challenges and make difficult situations more manageable for business leaders, managers, and stakeholders.

5.3. Confusion Matrix Results Evaluation

The confusion matrices represent a pivotal moment in this PhD research, as they lay the foundation upon which the predictive power of the models is truly tested and evaluated. These matrices, in fact, provide a granular view into how each algorithm performed in classifying companies into two categories. However, it's important to emphasize that the analysis is not just about "*how well*" these models perform; it's in fact fundamentally about "*how*" they arrive at their predictions. In fact, while confusion matrices are often viewed as unbiased and accurate representations of an algorithm's performance—particularly when models are well-refined and datasets are complete and well-structured—the reality is much more nuanced: in many cases, and perhaps in most cases, the results generated by confusion matrices need to be carefully understood, contextualized, and analyzed beyond their immediate statistical output, and this requires deeper technical insights and real-world knowledge, where specific details of the data, industry, or financial environment can reveal layers of complexity that are not immediately evident in the raw output. For example, practitioners, industry experts, or financial professionals might detect patterns or anomalies that go unnoticed in purely data-driven approaches: these professionals, armed with years of sector-specific experience, can interpret and contextualize what a model might flag as a "*false positive*" or "*false negative*," thereby refining how the results are understood and acted upon. This entire consideration, reinforces the idea that predictive analytics, especially when dealing with crises or insolvency, are not just about achieving high numerical accuracy but also about ensuring that the right insights are derived for meaningful, practical application. Repeating what was said before in this dissertation, the performance here is not just a "*numerical outcome*"; it's a reflection of the study's overall success in achieving its primary objective: offering a reliable, data-driven framework for predicting financial distress in real-world scenarios.

Despite these considerations, which may seem obvious to seasoned professionals but often perceived as a "*trade-off*" for theoretical perfection and harmony by academics, the multi-algorithmic implementation in this study has yielded fascinating results from the confusion matrices. In fact, the varied algorithmic performance, as also highlighted by the confusion matrices, further underscores that Italian businesses—with their inherent diversity in size, sectors, and financial management—require inner predictive tools capable of not just delivering accurate forecasts or predictions but also supporting timely and impactful managerial decisions. Matrix results for the best performing algorithm are the following[174]:

[174] Given the clear superiority of XGBoost in categorizing companies into "*in-crisis*" and "*healthy*" categories, as compared to the other algorithms used in this study, and considering the refined dataset with statistically significant sample size, the deeper

Table 8: XGBoost confusion matrix (*Accuracy: 0.87, F1-Score: 0.88, AUC: 0.87*)[175]

	Predicted Healthy (0)	**Predicted in-crisis (1)**
Actual Healthy (0)	795 (T.N.)	180 (F.P.)
Actusl in-crisis (1)	70 (F.N.)	955 (T.P.)

Source: own elaboration

The confusion matrix for XGBoost (representing the test on 20% of the entire sample; 2.000 firms), in fact, demonstrates its high accuracy in correctly predicting companies' health status, particularly its strength in identifying companies that are genuinely in crisis[176]. The model's ability to capture True Positives (955), businesses correctly identified as being in crisis, is especially notable. This indicates that this algorithm is highly proficient at detecting distress signals in firms with deteriorating financial performance, likely due to the vast, detailed dataset it was trained on. Similarly, the True Negatives (795) indicate that the model is equally effective in identifying healthy, and active, companies. This reaffirms the robustness of XGBoost in predicting firms that are on stable ground, where financial indicators suggest these businesses can continue operations without immediate risk of insolvency. The high accuracy of these results reflects the algorithm's ability to correctly distinguish between firms that are solvent and those that aren't, ensuring that management can trust the predictions for making informed decisions regarding business continuity. However, the False Positives (180) and False Negatives (80) require attention. Firstly, the false negatives, in this case, 80 companies, were predicted to be healthy but had actually encountered significant financial difficulties, having either entered into formal insolvency procedures or experienced institutionalized crisis events: this suggests that these companies, while facing financial difficulties, may not yet be in irreversible distress or insolvency; instead, they may still possess the potential for recovery, with the possibility of restructuring or other turnaround strategies. Such firms might have avoided terminal insolvency but are nonetheless on a precarious path that requires close monitoring and possible intervention and, this, also hints at the necessity for further refinement in identifying early-stage distress signals, even when a company appears operationally sound.

In addition, the outcome regarding False Positives in this study presents both a significant and intriguing result, being intuitive and counterintuitive at the same time: in predictive studies, especially those centered around crisis and insolvency forecasting, it is common practice to attempt to minimize the number of false positives by applying statistical techniques under the assumption that the results produced by AI-based algorithms must always be trusted and validated[177]. While this assumption may hold some truth, it is crucial to recognize the nuanced and often subtle nature of corporate crises and

analysis of feature importance and confusion matrices will focus on this particular algorithm. The limited performance of the other models, such as Logistic Regression and Gradient Boosting, makes further scientific investigations into those algorithms less fruitful in the context of this research.

[175] T.N. refers to True Negatives; F.P. refers to False Positives; F.N. refers to False Negatives; T.P. refers to True Positives. Accuracy was calculated as follows: Accuracy=TP+TN/TP+TN+FP+FN=950+790/180+70+950+790=0.875.

[176] This matrix refers to the test utilizing every data, per every year, in the dataset.

[177] In predictive analytics, especially in the context of crisis and insolvency forecasting, the principal objective is to accurately identify situations of distress or impending financial trouble. Therefore, while false negatives—where companies in crisis are misclassified as healthy—can often be rationalized and deemed physiological due to the nuanced and gradual nature of financial crises, false positives demand more scrutiny, as they are typically regarded as inefficiencies or errors in the predictive model, diverging from the primary goal of the analysis.

insolvency, as well as the manner in which these situations quietly infiltrate the fabric of business operations, sometimes without immediate external recognition. Therefore, it becomes essential to contextualize these results and reflect critically on their implications. In the previous confusion matrix, False Positives account for 180 instances, representing nearly 10% of total predictions. This figure is noteworthy and invites deeper scrutiny: while such a value could be considered normal and physiological in predictive models, it may, in fact, conceal underlying nuances of financial distress within companies whose deteriorating performance has not yet been formally acknowledged. These companies, in fact, might be facing internal crises, but for a variety of reasons, such as managerial reluctance, fear of reputational damage, or a desire to avoid triggering formal insolvency procedures, their crisis status remains unaddressed and unreported. One plausible interpretation of these 180 False Positives is that they could represent businesses with degraded performance indicators, where the financial health of the company is slowly eroding, but the crisis has not yet reached a level of severity that necessitates formal intervention or insolvency procedures. In such cases, the AI algorithm, equipped with the ability to detect subtle patterns in the data, may be identifying early warning signals—financial red flags that are indicative of potential crisis down the road.

However, these signals might not have yet materialized into formal insolvency events, leading to a classification of "*False Positive*" when, in reality, the company may already be on a precarious path toward financial instability. Consider SMEs that may appear "*healthy*" in public databases, such as those maintained by chambers of commerce, because there are no formal insolvency procedures initiated against them. In reality, however, these businesses may already be grappling with significant financial distress, but have chosen to address it through private resolutions or non-judicial pathways. For instance, under Italy's new Business Crisis and Insolvency Code (Codice della Crisi d'Impresa e dell'Insolvenza), introduced to modernize the country's insolvency regulations and offer more preventive measures, companies facing distress can access mechanisms such as the "*composizione negoziata*" (negotiated settlement): this process allows businesses to voluntarily seek the assistance of an independent expert who helps them resolve their financial difficulties without resorting to formal court proceedings. Because this is a private procedure, these companies may not appear in formal insolvency records, and thus they may still be classified as "*healthy*" in public databases, leading predictive models to flag them as False Positives when in fact, they are in latent crisis. In this regard, a practical exemple could be that of a family-owned manufacturing firm in northern Italy, which, like many SMEs in the region, has encountered severe liquidity issues due to a combination of rising raw material costs and delayed payments from key clients. The company has not yet initiated formal insolvency procedures but has quietly opted, last year, for negotiated settlement through the new Italian insolvency code. As part of this process, they have engaged an independent expert to help restructure their operations and renegotiate terms with creditors; however, this entire process remains largely hidden from public view—no formal filings with the court have been made, and stakeholders such as suppliers or even employees might not be fully aware of the severity of the firm's financial situation. In this case, for a predictive model, this firm would still appear as "*healthy*" due to its active status in public records, but an AI-based algorithm that takes into account deteriorating financial metrics like falling revenues, rising debt ratios, or worsening liquidity could rightly flag this company as a potential crisis case—thus leading to a False Positive.

In a more general way, it is plausible that the use of out-of-court solutions for crisis resolution, including negotiated settlements, could be also strategically exploited by management to delay the disclosure of financial distress. By opting not to externalize the crisis, managers might seek to maintain internal control over the situation, attempting to implement corrective measures with the hope—or even the informed expectation—of achieving a "*miraculous*" turnaround without external intervention. This

approach, however, carries significant risks. Firstly, continuously deferring the formal acknowledgment of distress may exacerbate the company's financial position, especially if internal recovery efforts ultimately prove ineffective; failing to communicate the company's difficulties transparently can also erode the trust of creditors, investors, and employees, thereby further compromising the company's reputation and stability; lastly, this behaviour may limit the effectiveness of potential interventions, particularly if the crisis escalates to a level that cannot be managed autonomously. In some instances, managers may adopt this strategy based on specific information or anticipated events—such as incoming financing, market shifts, or new contracts—that they believe might rescue the company. However, such expectations may not materialize, potentially worsening the firm's situation.

In such cases, the False Positive is not an error, but rather an insightful signal: in fact, it indicates that the company may be on the verge of formal insolvency if corrective actions are not successful, or if the financial strain becomes too great to handle privately. The company's decision to avoid formal procedures could stem, certainly, from concerns about damaging relationships with stakeholders, suppliers, or banks, or fear of harming its reputation: in Italy, where personal reputation and business relationships are deeply intertwined with economic stability, many entrepreneurs, particularly in smaller firms, prefer to "*conceal*" their crisis for as long as possible to avoid ripple effects that might lead to withdrawal of credit lines, worsening of supplier terms, or loss of key business partners[178]. From a managerial perspective, this outcome highlights a critical insight: False Positives could be interpreted as a signal that the company is experiencing latent crisis symptoms that have not yet been formalized or acted upon by the management. It may indicate that management has chosen not to externalize, knowingly or not, these warning signs, perhaps to avoid negative repercussions such as damage to relationships with stakeholders. Moreover, it is well-known that, especially in smaller firms, managers may lack the expertise, willingness, or proactive attitude to address early signs of distress, preferring instead to take a more conservative and hidden approach to the crisis. This, unfortunately, often leads to worsening conditions and an eventual collapse that could have been avoided with timely intervention. In this context, False Positives are crucial signals, as they suggest the presence of firms that should be closely monitored for early intervention, even if they have not yet crossed the threshold into formal insolvency.

In conclusion, this result is essential because it underlines the importance of contextualizing the outputs of predictive models and recognizing that AI algorithms may provide valuable early insights into the potential for future financial distress, even when the crisis is not immediately apparent. While the tendency in the field is to reduce false positives to improve the perceived accuracy of a model, in the specific case of insolvency prediction, False Positives can represent a key opportunity for preventive action as they offer a second chance for companies to address their financial vulnerabilities before the situation escalates to the point of no return[179]. Thus, the 180 False Positives identified in this study may not just represent algorithmic failures; rather, they could highlight latent risks that are essential for management to consider. To enhance comparison and oversight, future research could involve a comprehensive analysis of the financial statements of firms included in this sample of False Positives: by conducting detailed financial analyses, it would be possible to assess whether there are discernible

[178] In practical terms, imagine a supplier or financial institution reviewing these AI-generated predictions. A company flagged as a False Positive might warrant a conversation or closer examination of recent financial statements, perhaps prompting a meeting with management to assess the situation more deeply. If the signs of distress are caught early, this could lead to more favorable restructuring options, better terms with creditors, or a reworking of supply chain strategies before the company spirals into insolvency.

[179] This phenomenon is particularly relevant in Italy, where small and medium-sized enterprises form the backbone of the economy, and where informal, non-judicial crisis management strategies are increasingly being encouraged under the reformed insolvency law.

indicators of distress. Furthermore, methodologies such as Benford's Law could be applied to detect potential data manipulation. This statistical technique, in particular, could help uncover inconsistencies or irregularities that may indicate deliberate alterations by management to conceal the firm's distress. Such methods would allow for the identification of potentially unreliable data, offering a more robust basis for assessing the financial health of these firms and providing additional layers of scrutiny to future studies.

5.3.1. Confusion Matrix Results for SMEs

In order to deepen the analysis and evaluate potential differences, a second round of confusion matrix analysis was conducted, focusing exclusively on SMEs within the dataset, excluding the large enterprises. In fact, while large companies represent a small fraction of the sample, their inclusion might introduce noise into the results, potentially affecting the clarity of the findings. On the other side, by isolating SMEs, especially micro-enterprises—which dominate the Italian business landscape—it becomes possible to assess the true performance of the models in the context of smaller firms, whose crisis behaviors are distinctively different from those of larger corporations. From an intuitive perspective, SMEs, particularly micro-enterprises, are known to have unique managerial and financial structures, often influenced by family-run dynamics. Furthermore, these companies typically face resource constraints, both in terms of financial capital and human resources, which can significantly limit their ability to engage in preventive restructuring or turnaround efforts. Consequently, managers may lack the skills or expertise required for early crisis intervention, and, due to reputational concerns, they may delay formalizing or addressing their financial troubles, which can exacerbate the situation. All these things make SMEs particularly vulnerable, and, examining the model performance specifically for this subset of firms can offer valuable insights. Implementing the same R code for the XGBoost algorithm, and considering the following dataset: 143 medium-sized firmes; 698 small; 4.100 micro (4.941 total firms in the final dataset per sample (balanced); 9.882 firms in total; -1,18% with respect to the original dataset)[180], results are the following:

```
xgb_predictions <- predict(xgb_model, xgb_test_matrix)
xgb_pred_class <- ifelse(xgb_predictions > 0.5, 1, 0)

xgb_accuracy <- confusionMatrix(factor(xgb_pred_class),
factor(test_label))$overall["Accuracy"]
xgb_f1_score <- F1_Score(as.factor(test_label), as.factor(xgb_pred_class))
xgb_auc <- auc(test_label, xgb_predictions)

xgb_accuracy <- 0.86
xgb_f1_score <- 0.85
xgb_auc <- 0.86
```

Furthermore, on a test set of 1.976 firms (20% of the total number of observations), confusion matrix results are the following:

[180] The percentage difference is minimal, indicating that the noise within the data is also negligible. This proportion, however, broadly reflects the overall composition of businesses in Italy, where SMEs dominate the economic landscape.

Table 9: SMEs XGBoost Confusion Matrix (*Accuracy: 0.8582996, F1-Score: 0.85, AUC: 0.86*)

	Predicted Healthy (0)	**Predicted in-crisis (1)**
Actual Healthy (0)	790 (T.N.)	183 (F.P.)
Actusl in-crisis (1)	70 (F.N.)	933 (T.P.)

Source: own elaboration

In the revised analysis, the exclusion of large enterprises indeed reveal some differences in the results: the confusion matrix shows a slight increase in false positives (+3). This aligns, slightly, with the earlier hypothesis that SME managers, especially those in micro-enterprises, may often delay or avoid formalizing crises due to the personal and familial ties within the company, as well as fears of damaging their reputation with stakeholders. Additionally, the model's overall accuracy decreased slightly compared to the overall dataset (-0,01). So, the matrix reveals that, despite the reduced dataset, the false positives remain at approximately 10% of the total test set, which still makes them worthy of attention, especially since this type of analysis generally aims to minimize these occurrences. A further consideration is that the inclusion of larger enterprises has only introduced minimal noise, resulting in nearly identical performance metrics compared to earlier analyses. However, this consistency reaffirms the need to carefully interpret false positives within the context of predictive studies, suggesting they may not always be *"false"* in the traditional sense. In future research, these numbers could be further investigated by utilizing a more advanced AI-based or statistical system for detecting and categorizing companies within each area of the confusion matrix: this would allow for a deeper understanding of the classification outcomes.

5.3.2. Confusion Matrix Dynamic Analysis

Given the data-driven and inductive nature of the chosen empirical approach, as well as the potential of ML algorithms to conduct dynamic, time-sensitive analyses, it was deemed appropriate—consistent with the literature cited in Chapters 2 and 3—to evaluate the confusion matrix results and the accuracy of the best-performing algorithm (XGBoost) specifically at time intervals of -1, -3, and -5 years from the point of financial distress, which are considered significant by prior research[181]. This approach addresses the need to understand how predictive accuracy may vary over time and whether specific economic-financial patterns emerge that enhance the detection of financial distress.

[181] The choice to analyze years -1, -3, and -5 relative to the crisis event is grounded in both theoretical and practical considerations. The year -1 is expected to reveal a particularly critical situation, as firms are generally more visibly distressed just before insolvency. Indeed, only in rare cases do companies face sudden insolvency without prior signs of distress—this tends to happen either due to data manipulation by management or due to unforeseen exogenous shocks, such as a pandemic or other disruptive events. By contrast, year -3 serves as a midpoint, capturing financial signals from prior years that are neither too close to the crisis event nor too far removed. Lastly, year -5 represents the earliest temporal perspective in this analysis, offering insight into the longer-term evolution of financial distress. This divergence in time intervals provides valuable opportunities for comparative evaluations, as the heterogeneity in these timeframes allows us to assess how well the model performs across varying distances from the crisis. Naturally, in the -1 year analysis, we anticipate the highest accuracy in model predictions, as indicators of financial distress are likely more pronounced at this stage. Furthermore, these timeframes are not only scientifically relevant but also highly practical for management: strategic planning and decision-making typically follow similar temporal horizons, with firms often setting objectives and evaluating performance over one, three, and five-year spans. As such, the results of this analysis could prove instrumental for managers, offering guidance on when early warning signs may first emerge and how far in advance a proactive approach could be beneficial. This temporal sensitivity thus enhances the practical applicability of the entire study.

Furthermore, it seeks to determine if algorithms and new technologies are capable of identifying early warning signals of crisis well in advance of formal insolvency events.

This longitudinal analysis is essential for several reasons. In the initial phase of the study, a multi-algorithmic approach was applied to the entire dataset, utilizing all available data for each firm and each year. As a result, the model was "*omniscient*" in a statistical sense—it had access to every historical data point for each company. However, to gain a clearer understanding of the predictive power over specific periods leading up to a crisis, it is necessary to isolate and test individual years of data (-1, -3, and -5) and analyze the results independently at each time point. By restricting the analysis to these specific years, the work shifts from a holistic view to a temporal one, assessing the model's ability to forecast financial distress relative to time. This approach enables us to achieve more targeted results linked to specific years before the crisis, always moving from an absolute-year analysis to a relative, dynamic one based on the last available year in the dataset. In doing so, this part aims to answer several key questions: firstly, how does predictive accuracy evolve as the time interval before a crisis event increases? Specifically, does the model's predictive power diminish or improve as we move from -1 year (closer to the insolvency) to -5 years (further from the insolvency)? Implementing this temporal analysis involves running XGBoost on subsets of data that correspond exclusively to each specified year. For instance, in the -1 year model, only the financial features from the year immediately preceding the crisis are utilized; similarly, for the -3 and -5-year models, only data from three and five years prior to the crisis are included, respectively. This design allows for an isolated evaluation of the model's accuracy and robustness for each pre-crisis interval, free from potential contamination by data from other years. The implications of this analysis are also substantial: by focusing on distinct years before the crisis, this part provides a deeper insight into the temporal sensitivity of financial distress prediction models. Such results can inform practitioners and policymakers regarding the appropriate lead time for interventions, and they may indicate which financial indicators are most useful for forecasting at different stages before a crisis. Additionally, understanding the limitations and capabilities of ML models in this time-sensitive context contributes to the literature on predictive analytics, offering practical recommendations for the early detection of vulnerability in firms.

So, the implementation of this dynamic approach required some different steps:

- *Data preparation for different timeframes.* The goal, here, was to create subsets for year -1; -3; -5, based on the main dataset. The following R code was then used:

```
data_years <- lapply(years, function(y) {
df %>%
filter(year == (last_year + y)) %>%
select(-year)
```

- *XGBoost Model Setup for Dynamic Analysis.* In this step, the definition of a function that trains and evaluates the model for each dataset is needed. This function trains XGBoost on each specific year's data, calculates performance metrics, and stores the results. In the context of this empirical study, the following R libraries were used: *library(caret); library(MLmetrics).*

After the creation of different subsample, XGBoost was used to determine different outcomes for different years. So, each data subset was used to train an independent XGBoost model. The model parameters, chosen based on prior research in binary classification for financial prediction, were consistent across all three intervals to ensure comparability. For each subset, the dataset was split into training and testing sets, where the model was trained on 80% of the data and tested on the remaining

20%. Once trained, the model generated predictions, which were subsequently evaluated on accuracy. The analysis was implemented for the general sample (10.000 firms). Results were the following:

- At year -5, the model's Accuracy was recorded at 0.79, which is notably lower than the overall accuracy previously observed (0.87). This decrease in accuracy five years before insolvency indicates a predictably reduced performance of the algorithm at a greater temporal distance from the distress event. The model's ability to detect early signals is inherently challenged as the timeframe extends, which aligns with expectations, as early indicators are often subtler and harder to identify accurately. For a test sample size of 2,000 firms, True Positives (TP) are equal to 781; True Negatives equals to 788. False Positive are quite high (220);

- At year -3, the model's Accuracy was recorded at 0.845 for a sample of 2,000 firms. This accuracy, while improved compared to year -5, is still lower than the overall accuracy of 0.87, indicating a gradual increase in predictive power as the timeframe approaches closer to the distress event;

- At year -1, Accuracy equals to 0.86, in line with the general result. This confirm the robustness of the algorithmic prediction.

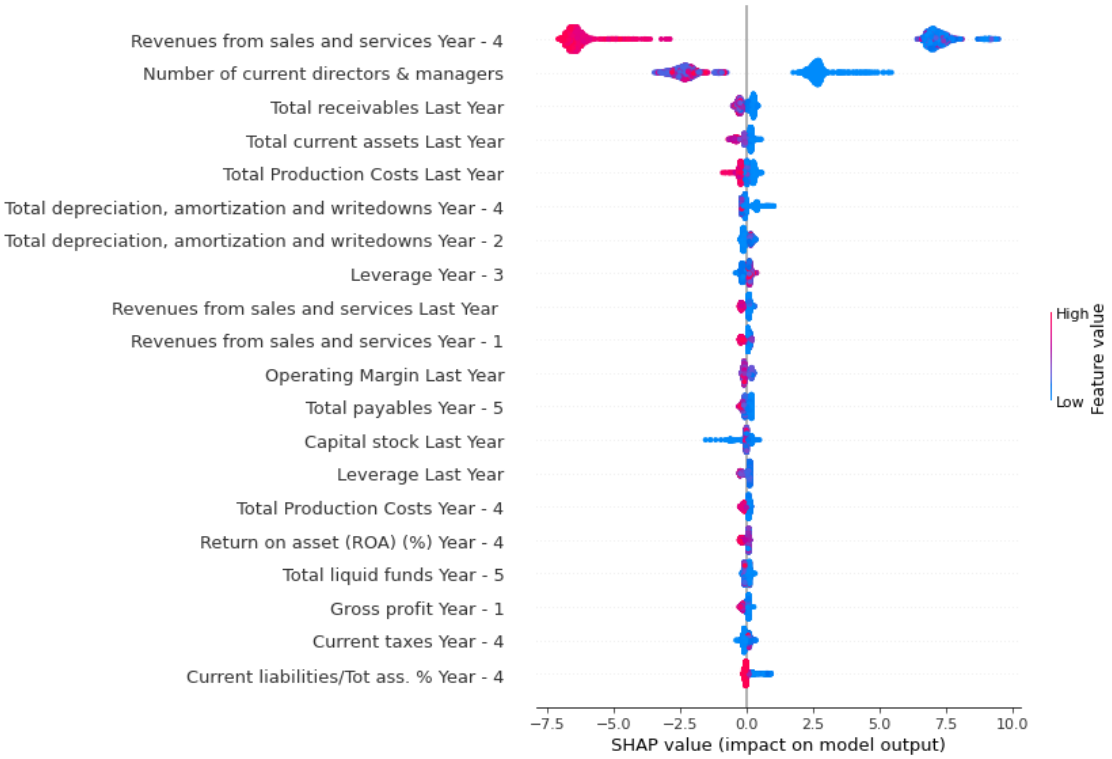
The dynamic approach with data segmented by individual timeframes not only confirms the robustness of the XGBoost algorithm but also provides unique insights into its predictive reliability over time, a method rarely explored in both international and Italian literature. Specifically, the False Positives at year -1 closely mirror the overall results, stabilizing around 9% of the test sample of 2.000 firms. This similarity indicates that, as expected, the model's accuracy is highest in the timeframe immediately preceding insolvency, where financial distress signals tend to be more visible and easier to detect. In technical terms, this study addresses a critical gap in predictive modeling, where models are often overfitted to static datasets, leading to high accuracy in controlled settings but limited generalizability in dynamic, real-world scenarios. By applying a dynamic analysis, it achieves a more temporally sensitive evaluation, thereby reducing overfitting and enhancing the model's practical applicability. This process reveals the algorithm's ability to distinguish genuine distress indicators across time—a valuable trait that bolsters its reliability in evolving financial contexts. However, such dynamic temporal analysis remains relatively uncharted territory in the literature, especially in Italian studies, where few analyses have employed this method to systematically evaluate predictive performance at multiple pre-insolvency intervals. In conclusion, the results provide actionable insights for targeted managerial interventions: notably, even at -5 years, the algorithm's accuracy aligns closely with classic findings, such as Altman's 1968 model, reaching a similar accuracy of 0.78. This consistency highlights the model's capacity to detect early signs even at a significant temporal distance from insolvency. At -3 years, the results are not only acceptable but remarkably robust, achieving accuracy levels that underscore the model's effectiveness in identifying at-risk firms well in advance of crisis onset.

6. Model Explanation and Results Interpretability

As in the exploratory empirical example, the SHAP tool was employed for this study, leveraging its data-driven and AI-powered capabilities to assess the model's decision-making process.

Specifically, SHAP was used to visualize the impact of various features classified by the XGBoost algorithm as the most significant in predicting insolvency or financial health and the graphical representation was conducted through violin diagram, which allow us to effectively measure the impact of each feature on the model's output, providing clear insights into the contribution of each variable. The results were calculated with careful consideration of the relatively small number of large enterprises in the dataset, as previous findings indicated that these did not introduce significant noise into the overall results; however, given their unique importance, the analysis also included qualitative variables. In particular, in this context, the feature "*number of managers and directors*," as classified by the dataset (AIDA), refers to the total number of executives, managers, and members of the board who are directly involved in overseeing business operations and strategic decisions. Results of the violin summary plot are reported in Figure 30.

Figure 30: Violin Summary Plot



Source: own elaboration

Each SHAP value represents how much a particular feature pushes the model's prediction towards insolvency or health: the higher the absolute value of a SHAP score, the more influence that feature has on the prediction. Moreover, features with positive values are associated with a higher likelihood of insolvency, while those with negative values are associated with a company being healthy. In the graph, the majority of SHAP values have a peculiarity: they tend to cluster towards the center, indicating that for most of the firms, these features generally have moderate or balanced effects on the predictions. Only in a few cases do we see extreme SHAP values, signaling that certain variables have significantly pushed the model's predictions in one direction. Commenting some results, it is possible to affirm that:

- Revenue-related variables, such as "*Revenues from sales and services*", stand out as some of the most impactful features in the model[182]. Unsurprisingly, declining revenues over multiple years signal financial distress and have a significant influence on insolvency and crisis predictions. Notably, the revenue from Year - 4 shows the highest impact compared to other years. This could suggest that, often, the signs of a company's crisis may begin to manifest long before they become apparent in formal proceedings, indicating a lag in the recognition of financial troubles[183];

- A particularly intuitive result is the observed impact of the "*Number of current directors & managers.*" Firms in crisis, particularly in the last available year (which "*current*" refers to in this context), may have already reduced the number of decision-makers, and this is often a consequence of restructuring, downsizing, or consolidating leadership to respond more efficiently to financial pressures. However, such reductions can lead to a narrower decision-making process, where fewer people are responsible for managing complex crisis scenarios, making it harder to execute broad-based, strategic interventions. This is especially relevant in family-owned businesses or SMEs, where leadership and decision-making are often concentrated within a small group or even a single family; in fact, in such cases, there may be emotional and personal factors influencing decisions, which can further complicate crisis management. On the left side of the SHAP plot, we see, instead, a mix of values, indicating that for some firms, both higher and lower numbers of decision-makers have a lesser impact on insolvency. This suggests that certain companies are able to maintain stability regardless of the number of directors or managers, likely because other factors (such as strong operational foundations or external financial support) play a more pivotal role in their ability to survive crises. Therefore, a smaller or larger leadership team alone doesn't guarantee stability but may interact with other internal and external dynamics;

- "*Leverage*", represented by the proportion of debt relative to total assets—shows a strong, but mixed, impact: higher leverage ratios make firms more vulnerable to financial instability, confirming that highly leveraged companies have limited flexibility in managing financial shocks, making them more prone to insolvency (Leverage Year -3); a reduced level of Leverage the year before crisis, instead, seems to have less impact on the crisis itself;

- One of the more surprising results involves the "*Total Production Costs*" (Last year; Year -4) and variables such as "*Total depreciation, amortization and writedowns*". These variables, despite their intuitive link to financial performance, do not show as strong a correlation with insolvency as revenue or leverage. This may be due to the fact that companies in distress often reduce production costs and slow down depreciation as part of survival strategies. As a result, their financial health deteriorates gradually, and these cost-cutting measures might obscure the real extent of their struggles;

[182] This confirms correlation results previously presented in this Chapter.
[183] This confirms considerations about crises evolution, already made in this Chapter.

- "*Gross profit"* (-1), which represents a company's earnings after deducting production costs, has a moderate influence in the model. A declining gross profit is typically a sign of inefficiency or shrinking market competitiveness. However, it is interesting to note that the model places greater emphasis on revenue and leverage. This suggests that in the Italian business context, gross profit margins alone may not be the best predictors of financial distress, and their importance varies depending on other financial factors;

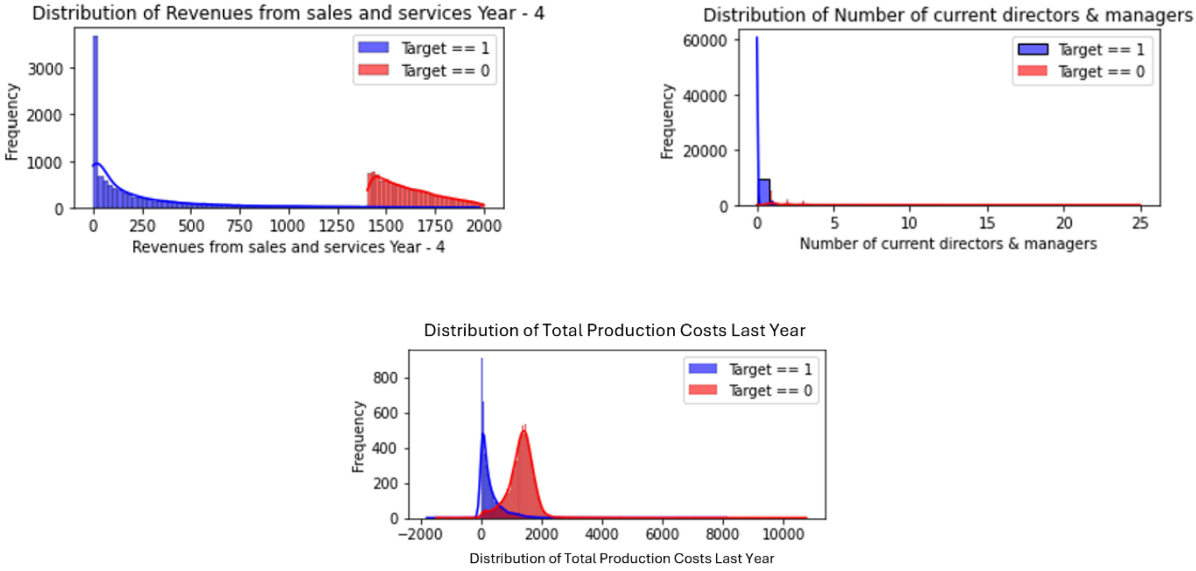
- For "*Total Receivables Last Y*ear," the values reveal a counterintuitive yet significant insight: firms with lower receivables tend to have a higher propensity for being classified as in crisis or insolvent (target = 1). This might seem unexpected, as we typically associate large outstanding receivables with liquidity issues. However, in this context, it suggests that firms with very low receivables may have already experienced a downturn in sales or a significant slowdown in operations, leading to fewer outstanding debts from customers, which is a red flag for declining business activity. In contrast, firms with higher receivables tend to be classified as healthy (target = 0) because they are still generating enough business to maintain a strong receivables portfolio, indicating ongoing commercial activity and a healthier financial state;

- For "*Total Current Assets Last Year*," the results are more straightforward and intuitive. The SHAP values indicate that firms with lower current assets in the year before crisis are more likely to be classified as in crisis or insolvent (target = 1). This makes sense since lower current assets suggest weaker liquidity, making it difficult for the firm to cover its short-term obligations, a typical precursor to financial distress. Conversely, firms with higher current assets tend to be classified as healthy, as they have sufficient resources on hand to meet immediate liabilities, ensuring better financial stability and solvency.

This SHAP analysis provides a clear, actionable roadmap mainly for micro and small enterprises, often lacking the resources for deep financial analysis or crisis management strategies. By utilizing AI-driven insights, these businesses can now identify critical early warning signs of financial distress—long before they reach a point of no return. Features like revenues, receivables, and current assets, which the SHAP analysis has flagged as pivotal, become the beacons for proactive management: for small businesses, where decisions are often made intuitively or informally, these insights provide an objective, data-backed guide to steer through turbulence. Furthermore, in order to deepen the peculiar relationships, the analysis also incorporates frequency distributions of key pivotal features, providing a deeper understanding of their trends and behaviors across different business states. Results for *"Revenues for Sales and Services Year -4"; "Number of Current Directors and Managers"; Total Production Costs Last Year"* are the following (Figure 31; 32; 33)[184].

[184] These features were specifically chosen because the results are clear-cut, with visible differences in SHAP values. Such sharp contrasts provide critical managerial insights and serve as effective facilitators for the algorithm, enabling it to categorize firms with greater precision.

Figure 31, 32, 33: Distribution Plots for "*Revenues for Sales and Services Year -4*"; "*Number of Current Directors and Managers*"; *Total Production Costs Last Year*"



Source: own elaboration

These distribution plots align with the insights observed in the earlier violin diagrams, confirming very clear-cut patterns:

- "*Revenues from Sales and Services Year -4":* The distribution shows that most firms in crisis (Target = 1) exhibit significantly lower revenues, clustering around zero. Meanwhile, healthy firms (Target == 0) are distributed across a wider range of revenues, with a clear separation between the two groups (confirming the SHAP interpretation);

- *"Number of Current Directors and Managers":* The majority of firms in crisis have a very low number of decision-makers, with the peak occurring at zero or close to zero. This suggests that the reduction of managers may occur in the later stages of a crisis, confirming the connection between governance structure and financial health;

- *"Total Production Costs Last Year":* The distribution for production costs is strikingly different between the two groups: firms in crisis tend to have negative or very low production costs, while healthy firms exhibit a more diverse and generally higher cost structure. This result suggests that companies in crisis drastically reduce their production activities, further reinforcing the algorithm's ability to distinguish between the two categories.

7. Conclusion and Remarkables

This empirical study offers a unique and profound contribution to both the academic and managerial realms by exploring the predictive power of AI-based models, particularly through the deployment of XGBoost in a dynamic way. The results presented not only underscore the scientific relevance of it but also highlight its practical importance for decision-makers. A key strength of this

study lies in its emphasis on manual checks and the thorough analysis of false positives: while they are often seen as inefficiencies in predictive models, this study advocates for a more deep interpretation. These cases, which at first glance appear to be algorithmic misclassifications, might in fact represent latent or unreported crises within companies. This perspective is especially important in light of the evolving legislative landscape, where both European and Italian regulators are increasingly favoring private, negotiated solutions to corporate distress in an effort to reduce the stigma associated with formal insolvency procedures. However, this can create "*grey areas*" in both scientific research and in real-world business operations as, on one hand, these frameworks allow managers or entrepreneurs the flexibility to resolve crises discreetly, potentially restoring financial health without attracting unwanted attention; on the other hand, this discretion may lead to inefficiencies, as unresolved or hidden crises can eventually culminate in severe, irreversible insolvency[185].

This delicate balance between discretion and transparency highlights a critical tension in crisis management: while concealing the early stages of distress may provide companies with a temporary shield to recover, it may also delay necessary interventions, prolonging the deterioration of performance and ultimately leading to unavoidable failure. In this regard, an example is that of the already mentioned Italian tool of negotiated settlement of corporate crises, that could be opportunistically utilized by certain entrepreneurs[186]. Another key insight from this study lies in the value of dynamic analysis, which bridges an important gap in the literature by providing objective, robust evidence of the algorithm's capability to identify financial imbalances as far as five years before an insolvency event, and with respectable accuracy. This finding is especially pertinent for Italian SMEs, where implementing tailored AI-driven systems for data analysis and corporate health monitoring could significantly benefit firms. Following appropriate training and awareness-raising programs for managers and employees, companies might explore the integration of these tools, leveraging advancements in predictive modeling along with the simplified practical test introduced by the Italian Business Crisis and Insolvency Code (CCII), the EU Insolvency Directive, and guidelines from the AI Act. However, it is essential to remain aware of risks such as AI washing (overstating AI's capabilities), algorithmic overtrust, and potential biases stemming from data manipulation. If implemented with these factors in mind, AI-driven predictive systems can lead to significant cost optimization and proactive, preventive strategic improvements and this approach not only enhances relationships with

[185] In the context of Italian businesses, especially those facing financial distress, the tendency to hide or downplay a company's crisis is often intertwined with more deliberate acts of financial statement manipulation. In fact, this is a well-documented practice where entrepreneurs, in an attempt to avert immediate repercussions or maintain a façade of stability, alter the presentation of key financial data. These manipulations go beyond mere behavioral tendencies, entering the realm of deliberate financial misrepresentation, where specific items in the balance sheet are adjusted to conceal the reality of the firm's declining health. A notable example of this can be found in the treatment of inventory: for many businesses, particularly those in manufacturing or retail, inventory represents a significant portion of their assets. So, in times of crisis, one common manipulative tactic is to inflate the value of inventory on the balance sheet. This can be achieved through various means—overestimating the value of unsold stock, failing to write down obsolete or unsellable goods, or delaying recognition of cost-of-goods-sold (COGS). By artificially inflating inventory, the business appears to have more tangible assets than it truly does, which can temporarily bolster financial ratios like the current ratio or the return on assets (ROA), presenting a healthier financial outlook to stakeholders, creditors, or potential investors. This kind of manipulation then allows the entrepreneur to defer dealing with the underlying issues and potentially avoids triggering crisis-related procedures, such as insolvency proceedings. However, it also creates a dangerous illusion of stability: the artificially propped-up numbers not only obscure the firm's true financial condition but can also lead to poor decision-making, worsening the company's actual situation in the long run.

[186] The Italian legislator expressly provides that, if the negotiations within the framework of the negotiated composition fail, the entrepreneur is granted the opportunity to access a simplified arrangement for the liquidation of assets. So, even if this simplified arrangement allows for the rapid sale of assets without requiring agreements with creditors, contrasting with the traditional liquidation procedures, it may also prolong the resolution process and worsen the company's financial condition. Thus, while the flexibility offered by the law can be beneficial, it may also create "*grey areas*" in real-world practices, as the absence of public crisis declarations can hinder accurate risk assessments and timely interventions. However, in this particular case of simplified arrangement, there are strict rules for managers and creditors in order to avoid opportunistic situations.

stakeholders but also strengthens competitive advantage and operational performance. Furthermore, it fosters trust and presents a clearer, more promising growth outlook for publicly listed firms, with positive ripple effects on the market and stock performance.

Furthermore, by visualizing the impact of various financial features on the prediction outcome, SHAP facilitates a deeper understanding of which elements are critical in distinguishing healthy firms from those at risk of insolvency: this underscores also the necessity of algorithmic literacy among managers (both the algorithms and the managers using them need to be finely tuned and adaptive to the specificities of the business context). A tailored input approach, in fact, would ensure that AI solutions are accurately aligned with the firm's operational realities, leading to more reliable outcomes. In conclusion, this research demonstrates the power of AI-driven tools like XGBoost and SHAP in delivering valuable insights that can be directly applied to improve managerial decision-making processes in critical situations. However, as these models evolve, it is critical that managers themselves are equipped with the knowledge to interpret these insights and implement tailored solutions for their specific business needs.

Chapter 5

Conclusion: Key Findings and Implications

Chapter Prelude

It's here that everything comes together, where all the theories, data, and analyses converge to offer something more than just conclusions. Key insights presented are not just findings in a study, but a moments of clarity in understanding the dynamics of modern businesses, especially in times of crisis. A look is given at what this research reveals about AI's potential to guide decision-making and how data-driven tools can become lifelines in uncertain times. But as with any meaningful journey, there are also imperfections, limitations that don't diminish the value of the work, but rather point the way forward, to the next steps in scientific exploration and refinement. This chapter, then, is not just about wrapping up, but it's about setting the stage for what's next. The work done here, though complete in its own right, aims to serve as a springboard for future questions, deeper investigations, and continued innovation in how we understand and handle the ever-changing managerial landscapes.

1. Introduction

Critical situations are an inevitable part of business cycles, a reality faced by companies of all sizes, across industries, in every Country. While the specific triggers of a financial or operational crisis may vary, what remains consistent is the disruptive impact on a company's stability and its long-term survival prospects (Cefis et al., 2020)[187]. For managers, in particular, dealing with a crisis is often about making decisions with incomplete or uncertain information, which is why the ability to predict potential distress before it manifests is becoming, today, increasingly vital. This is where predictive analytics, and more specifically AI-driven and Ml-based models, can offer transformative advantages. In fact, firstly, he adaptability of these systems allows them to continuously improve their predictions as they are exposed to new data, making them not only tools for forecasting but also for ongoing risk management: this dynamic learning process means that such models are not static but evolve alongside changing market conditions, internal company data, and even global economic shifts. So, in this context, AI and ML can serve as both predictive and prescriptive tools, helping businesses to design timely, data-driven interventions.

For businesses, particularly small and medium enterprises that often lack the resources to implement large-scale, costly risk management systems, these models can democratize access to sophisticated prediction tools. In fact, by integrating these technologies into their decision-making processes, managers can shift from a reactive stance—responding to crises after they unfold—to a proactive, preventative strategy where early signals are detected and addressed before they escalate into full-blown financial distress. This shift represents a profound evolution in crisis management, one where data becomes the foundation of strategic foresight and resilience. Furthermore, AI-driven models are highly customizable and can be tailored to the specific characteristics of an organization or industry, and this flexibility allows businesses to fine-tune predictions to their unique operational contexts, resulting in more accurate, reliable, and actionable insights. So, in an era where uncertainty and volatility are increasingly part of the business landscape, these models then offer a much-needed advantage by transforming vast amounts of data into meaningful, real-time decisions that support business continuity and sustainability. However, as shown in the SHAP analyses before, the ability to interpret model outputs and understand the driving forces behind predictions is crucial for actionable insights, as AI-based tools are not "*human*" and "*decision-makers*" but just tools that could allow real people to understand crisis alerts and to drill down into how some particular features impact their specific company. That said, it is glaring that an high level of explainability is essential in building trust in AI systems. In conclusion, the managerial, yet academic, relevance of the entire study, extends beyond the immediate benefits of the "*prediction*", as it also highlights the importance of continuous monitoring and the need for an adaptive, real-time taylored decision-making tools.

To futher illustrate the importance of this study, consider a scenario where an organization embarks on a comprehensive process of managerial and algorithmic literacy: this process is not merely about familiarizing managers with AI tools but fostering an environment where decision-makers fully understand also the potential and limitations of AI-driven systems. Through targeted investments in technology and a strong focus on building practical skills within the workforce, a business implements

[187] Crises, whether financial or operational, precipitate critical disruptions across organizational systems, destabilizing core operations and eroding both structural resilience and stakeholder. As identified by Cefis et al. (2020), such crises expose latent vulnerabilities in corporate governance and operational frameworks, impeding strategic adaptability and diminishing long-term viability. The destabilizing effects necessitate a refined, adaptive management response aimed at safeguarding operational continuity and reinforcing resilience mechanisms. This aligns with findings that underscore the importance of preemptive risk management and robust contingency planning to mitigate crises' pervasive impact on organizational survival.

an AI-based system for crisis prediction and management; this system is carefully integrated into the company's existing decision-making processes, enhancing the ability to detect early warning signals of financial distress. The key to success in this scenario is not just the implementation of the AI system but the creation of trust—both in the technology itself and in the organization's capacity to use it effectively. The management team, now equipped with algorithmic literacy, can interpret the AI's predictions, act on its insights, and make informed decisions in a timely manner: this newfound capability instills confidence not only within the organization but also among external stakeholders, shareholders, creditors, and regulators. The transparency and accountability that come with data-driven decision-making then foster a sense of trust in the company's long-term strategy and financial health and, from a broader perspective, this trust has a ripple effect. As stakeholders recognize the company's proactive approach to managing risks and ensuring business continuity, they become more willing to invest, extend credit, or enter long-term partnerships: this external validation can bolster the company's reputation, leading to an improved market position and greater stability. Moreover, the benefits extend beyond the individual organization: when multiple companies adopt AI-driven crisis prediction and management systems, an entire market ecosystem becomes more resilient. Businesses are able to avoid major financial collapses that can ripple through industries, thereby contributing to a more stable and sustainable economic environment. The ability to preemptively address crises then reduces the overall systemic risk, benefiting entire sectors.

Furthermore, the optimization of costs and resources through early intervention in difficulties can lead to more efficient business operations and the reduction in unnecessary expenditures associated with late-stage crisis management—such as legal fees, liquidation costs, or job losses—means that organizations can allocate resources more effectively. This, in turn, promotes not just financial sustainability but also social and environmental sustainability: as companies become more efficient, they are better positioned to invest in sustainable practices, reduce their environmental footprint, and contribute positively to their communities. In this way, the implementation of AI-based systems for crisis management and prediction, supported by a well-rounded managerial literacy in data science and algorithms, is not merely a technical advancement, as it represents a holistic transformation that touches on corporate governance, financial stability, market trust, and the broader social and environmental responsibilities of businesses.

In this context, it is also crucial to consider the evolving regulatory landscape, both in Italy and across Europe, where significant efforts have been made to streamline and simplify procedures related to business crises and insolvency: the legislative focus, exemplified by Italy's Codice della Crisi d'Impresa e dell'Insolvenza and the European Union's framework for crisis management, aims to reduce the stigma traditionally associated with financial distress and insolvency. These efforts are not merely bureaucratic; they are strategic interventions designed to enable businesses to recognize and address signs of financial discontinuity early, thus avoiding total collapse and ensuring continuity through proactive measures. So, the implementation of AI-driven and ML-based predictive systems becomes even more relevant as these systems, capable of detecting early warning signals, align with the legislative push towards crisis prevention rather than reaction. The legal simplification of insolvency procedures, particularly those allowing for more private and non-judicial resolutions like composizione negoziata della crisi d'impresa, can now work hand in hand with these advanced technological solutions, and the AI-based models help identify which companies may benefit from such procedures and provide data-driven justifications for initiating these strategic interventions.

Furthermore, the reduced stigma surrounding crisis management and insolvency, as promoted by both Italian and European legislators, is essential in fostering a healthier approach to business recovery. In fact, historically, the reluctance to acknowledge financial difficulties stems from fear of

reputational damage, particularly among stakeholders like investors, suppliers, and creditors. However, the new legislative frameworks encourage transparency and early resolution and, AI-driven predictive systems, when integrated within this legal context, minimize the risks of delays in addressing distress. For example, in the Italian context, when signals of disarray are detected, companies can firmly leverage legal tools such as the composizione negoziata before reaching a point of no return; by doing so, businesses have a structured opportunity to negotiate with creditors in a more favorable environment and, should these negotiations fail, the legal option of the concordato semplificato remains, allowing for the liquidation of assets without the burden of lengthy judicial proceedings. These legislative options, when supported by data-driven insights from AI, then enable companies to pursue solutions that are both timely and strategic, thus reducing the chance of an uncontrolled collapse. Ultimately, this synthesis of advanced AI technologies with the new legal frameworks designed to streamline crisis management can significantly enhance the effectiveness of both: the practical relevance of these tools goes high as they provide actionable insights that can help managers implement strategic interventions, optimize resources, and secure the entire business continuity. The cascading effect of such measures, as already stated, is clear.

Finally, the pages of this dissertation have demonstrated the power, the potential, and the implications of AI-driven predictive analytics in managing business critical situations. Through the deployment of sophisticated models and a rigorous methodological approach, it was shown that predictive models can provide both early warning signals and actionable insights, especially in the context of Italian SMEs. While there are still challenges and limitations to be addressed, particularly in terms of model refinement and data completeness, the findings provide a strong foundation for academic inquiry and practical application. In the next pages, each of these insights will be deepen.

2. From Theoretical Framework to Empirical Evidence: The Research Trajectory

What began as a set of research questions rooted in the shades of business crises, insolvency prediction, and the potential of AI-driven tools, has evolved, in these pages, into a robust framework that bridges theory with practical application. So, the trajectory of this research has not only tested established paradigms but has ventured into quite new territory, offering innovative insights that challenge conventional thinking: each step, from defining the core concepts to deploying sophisticated Machine Learning models, could represent a carefully constructed bridge between the usual academic rigor and real-world relevance. In following this path, in fact, the entire research has traversed complex landscapes—both theoretical and empirical—ultimately arriving at findings that hold significance also for practitioners. It is possible to affirm, then, that this PhD Thesis represents the culmination of a careful, methodical, and impassioned journey, one that began at the inception of the PhD program and unfolded through the subsequent years of rigorous research, reflection, and exploration. The core theme of predictive analytics, particularly within the context of corporate crisis and insolvency management, has been examined from multiple angles—methodologically, theoretically, and, mainly, empirically. Throughout this process, the entire 3-y work has been presented to the scientific community in the form of various academic outputs, including published papers, conference presentations, and workshops, each serving to disseminate key findings, share critical insights, and contribute to the broader discourse in the field.

So, the PhD research was a deeply personal exploration that sought to address critical gaps in the current body of knowledge: as part of this process, I critically examined existing methodologies,

identified state-of-the-art advancements, and pinpointed key turning points that could offer new contributions to the field. The research then evolved organically through a balance of passion for the topic and the structured guidance of my supervisor, whose input was instrumental in shaping the final outcomes of this work. In particular, rather than adopting an opportunistic approach aimed at securing superficial recognition, this Thesis sought to rigorously engage with the topic and to offer meaningful advancements to the academic community and real-world practitioners alike. Early on, the foundations were established through a structured plan of inquiry, and the entire trajectory was grounded in a scientific approach that balanced rigor with analytical creativity, aimed at producing results that were both reliable and forward-thinking.

2.1. Structuring the Work: A Rigorous, Data-Driven Approach

At the outset, the research was structured to follow a methodical, data-driven approach that would enable a deep dive into some critical aspects. This process required an open, analytical mindset, as well as the capacity to synthesize complex theoretical frameworks with empirical data. So, this dissertation was grounded in an inductive approach, one that prioritized not only a thorough analysis of existing literature but also a detailed examination of empirical data: the emphasis on data analysis was central to both the empirical component and the qualitative conclusions drawn from the study. The inductive methodology also allowed the research to challenge traditional frameworks that may no longer align with contemporary realities as, in contrast to the often-linear approaches of traditional models, this work embraced the complexity of modern business environments and the multifaceted nature of today's corporate criticalities. By leveraging a data-driven approach, the research then aimed to bridge the gap between theory and practice.

In Chapter 1, the foundation of the Thesis was laid out. This Chapter, in fact, critically introduced the core premises and objectives of the research. It underscored the necessity of predicting corporate crises and insolvency in a timely manner, especially in an era characterized by rapid economic changes and technological disruptions and also articulated the key research questions and objectives, focusing on how predictive analytics, supported by AI, could revolutionize traditional approaches to crisis management and could help managers to get alerts and keep going with their businesses.

On the other side, Chapter 2 followed by diving into the key elements necessary for a thorough understanding of predictive analysis, corporate crises, and insolvency. This section, in fact, offered a broad exploration of the theoretical underpinnings of the field while also considering practical, professional, and normative aspects. By doing so, it provided a comprehensive foundation for the rest of the Thesis, contextualizing the research within both academic literature and the real-world challenges faced by managers and businesses.

The third Chapter delved into the interdisciplinary literature that informs the current state of predictive analytics, and explored both the technical and managerial dimensions of the field, focusing on how advanced methodologies—particularly multi-algorithmic approaches and business intelligence theories—have reshaped the landscape of corporate crisis management and prediction. This Chapter, in particular, was essential in bridging the gap between theory and empirical research, showing how complex data models and innovative methods can be applied to real-world scenarios.

The fourth Chapter marked the empirical core of the work, representing a significant turning point in the research. This section presented the results of rigorous data analysis, using extensive datasets and cutting-edge methodologies: the empirical studies were designed to test the efficacy of various predictive models, and results were both enlightening and impactful, offering key insights into the strengths and limitations of different models.

Now, in Chapter 5, we arrive at the final stage, where the conclusions are drawn, and key insights are synthesized. This part discusses the limitations of the study, key insights, and suggests avenues for future research. The limitations identified in the predictive models, in particular, underscore the need for continued innovation and refinement in this area, while the key findings provide a valuable framework for businesses seeking to enhance their crisis management strategies.

2.2. A Balanced Method: From Passion to Analytical Rigor

In applying ML models like XGBoost and Random Forest, a conscious effort was made to refine each phase, from feature selection to model tuning, ensuring that the results were both reliable and robust. Equally important was the inductive reasoning that underpinned the study, allowing the data itself to reveal insights rather than being constrained by preconceived hypotheses. This data-driven approach brought with it the opportunity to discover patterns and relationships that traditional theories may overlook, and the empirical results provided deeper understanding of how various financial, operational, and managerial factors interplay to signal impending crises: this is where the passion for discovery intersected with the rigorous demands of scientific research—each insight represented a step forward in understanding complex business phenomena through an empirical lens. However, a key takeaway from this research was the need for a balanced approach to AI-driven tools. While advanced algorithms can process large datasets and detect subtle trends, the study highlighted the importance of managerial interpretation and human oversight. Predictive analytics, in fact, is not a standalone solution; it is a decision-support tool. In this context in particular, AI models must be complemented by human expertise, mainly in SMEs, where decisions are often made by a small number of individuals with deep institutional knowledge. Here, the ability of managers to interpret and act on predictive signals becomes just as crucial as the model's ability to produce those signals. Thus, the real power of AI lies in its capacity to augment human decision-making rather than replace it.

Furthermore, the research underscored the ethical dimensions of applying predictive models, particularly in ensuring transparency and fairness in their deployment. The "*black-box*" nature of certain algorithms, such as deep learning models, can raise concerns about explainability and accountability: therefore, integrating tools like SHAP was essential in demystifying model outputs, providing clear and interpretable insights that managers can trust. This focus on transparency, while maintaining predictive power, reflects, in particular, the dual commitment to ethical responsibility and scientific rigor throughout the research. Ultimately, this work demonstrates that a balance between passion and analytical rigor is not only possible but necessary in addressing complex, real challenges. The drive to push beyond established boundaries was always tempered by a disciplined focus on data accuracy, model reliability, and the practical implications of the findings. This balance, furthermore, led to a deeper understanding of the field and provided a meaningful scientific and practical contribution.

In conclusion, key findings of this research highlight the transformative potential of AI in crisis management, demonstrating how advanced predictive models can offer businesses a strategic advantage in distress situations. The next paragraphs will synthesize these findings, providing a roadmap for future research.

3. Key Findings

The research has been rooted in the idea that crises are not sudden, isolated events but processes that often begin to reveal themselves in subtle ways before they culminate in financial distress or insolvency. This understanding frames the importance of early detection, where predictive models have

the potential to identify at-risk companies well before they formally enter a state of crisis. However, the ability to predict crises hinges on the quality of data and the sophistication of the tools used to interpret that data: AI-based models like those deployed in this study—Logistic Regression, Random Forest, Gradient Boosting, and XGBoost—offer significant improvements over traditional methods, allowing for the capture of non-linear relationships and complex feature interactions that often characterize business environments. Moreover, these models go beyond simply flagging potential distress; they can offer insights into the specific variables driving a company's downturn, giving managers a roadmap for intervention.

The first crucial insight from this research is that, in today's landscape, where AI is becoming more advanced and integrated into business processes, it is imperative for companies of all sizes to equip themselves with the necessary competencies, skills, and tools. This preparation is essential for maintaining competitiveness in times of business-as-usual but becomes even more critical in situations where continuity is at risk due to financial distress or operational dislocation. Businesses that possess up-to-date technological tools and managerial literacy in AI are in a stronger position to anticipate challenges and ensure long-term sustainability. Furthermore, the analysis of the literature, the practical implications, and the evolving legal frameworks highlights an increasing focus on the concept of "*predictive capability". T*his term, which has gained substantial traction, reflects one of the core promises of modern AI: the capacity to foresee the future. Scholars and practitioners alike view AI's predictive capabilities as an extension of a timeless human desire—understanding and preparing for what lies ahead; in the corporate context, predictive analytics are then not only about anticipating supply chain needs or developing new products but also about safeguarding the company's ability to operate smoothly during periods of disruption.

The research has also revealed that much of the existing literature on predictive analytics and insolvency focuses on large corporations in non-Italian contexts, which presents a significant gap in the understanding of how small and medium-sized enterprises could benefit from these technologies. Many studies attempt to create theoretical frameworks that could be universally applied, often within the sphere of business intelligence or larger corporate strategy frameworks. However, these frameworks frequently overlook the unique challenges that SMEs face. Furthermore, in the Italian and broader European context, the legislative focus on easing the procedural burdens of business distress and removing the stigma surrounding insolvency has opened new avenues for incorporating AI-driven predictive models into legal and business practices. For example, the new insolvency frameworks introduced in both Italy and the European Union aim to simplify crisis management procedures and promote early intervention, thereby supporting the long-term viability of businesses; the concept of insolvency has shifted from being a signal of inevitable failure to being an opportunity for restructuring and renewal, and AI can play a pivotal role in facilitating this shift by providing fundamental early warning systems.

So, the practical and academic interest in AI's predictive capabilities is growing, especially as businesses and policymakers strive to understand the opportunities and limitations of AI in crisis prevention and management. Predictive analytics, once considered a luxury reserved for larger firms, has become a strategic necessity even for SMEs and, additionally, as national and European legislation continues to evolve, making it easier for businesses to access restructuring, the demand for advanced predictive models will likely increase. This shift towards more flexible and stigma-free insolvency frameworks makes predictive analytics even more essential as legislators have recognized that by allowing companies to enter into pre-insolvency frameworks without the burden of reputational damage, they can encourage earlier and more effective interventions. In conclusion, the growing emphasis on AI's predictive capabilities within the academic literature, legal frameworks, and practical

management points to an inevitable trend: businesses that invest in these technologies will be better equipped to navigate future uncertainties.

In the entire study, a critical managerial perspective was consistently applied, underpinned by an intelligent, data-driven, and above all, tailor-made approach; this, allowed to thoroughly address the specific needs, implications, and results derived from a meticulously refined and extensive dataset. The multi-algorithmic approach, combined with a critical, mixed-method analysis—both qualitative and quantitative—enabled the extraction of numerous key insights throughout the research process. One of the most critical takeaways has been the recognition of the central role that data plays in modern research and decision-making processes: today, in fact, data is often referred to as the "*new gold*" of our society, and it holds a profound significance not only in the scientific realm but also in the real world. However, data is often taken for granted, undervalued, and insufficiently scrutinized. The empirical analysis emphasized, instead, that while advanced AI-based and statistical techniques are invaluable for analyzing large datasets, these methods alone are insufficient. There is in fact the the necessity of adopting a manual approach as well—contextualizing the data to ensure that it reflects real-world conditions and complexities. A consequent key element revealed by the confusion matrices used in the study was the importance of understanding "*false positives*" and "*false negatives*." In particular, the results demonstrated how firms experiencing performance degradation due to various underlying causes might still be classified as "*active*" and "*healthy*," simply because they are still operating and have not yet fully revealed their distress. From a scientific and analytical standpoint, this scenario leads to the concept of "*false false positives*": these are instances where firms may be erroneously considered healthy when, in fact, they are on a clear downward trajectory—often because the crisis has not yet been fully externalized.

Unlike false negatives, which are expected and understandable given the nature of predictive insolvency studies, these "*false false positives*" can pose significant risks in a real-world context, as they could lead to dangerous delays in implementing critical interventions. Managers and business owners—particularly those lacking the necessary skills or unwilling to openly acknowledge distress—in fact, may defer essential corrective actions needed to ensure the continuity of the business. This procrastination, influenced by either ignorance or fear of the potential reputational damage that could result from disclosing financial distress, often leads to more severe consequences that could have been avoided with timely intervention. Another crucial issue that intersects with this analysis is the manipulation of financial data, which is especially relevant in cases where firms are under pressure. The multi-algorithmic, data-driven approach employed in this study places a strong emphasis on the integrity of the data; however, it is essential to recognize the reality that financial data may be intentionally manipulated by businesses, especially those facing distress, to mask their true financial state. This manipulation typically involves presenting an overly optimistic financial picture, concealing red flags that would otherwise signal a company's worsening condition. A common example in practice, particularly among SMEs, is the overstatement of inventory values in financial statements—one of the many manipulative techniques used to obscure underlying problems: by inflating inventory figures, companies create the illusion of solvency and health, when in reality they may be facing significant challenges. This manipulation can distort the analytical results of even the most sophisticated predictive models, as they rely on the accuracy and authenticity of the input data.

Furthermore, the study highlighted another dimension of potential distortion: the influence of legislative simplifications and out-of-court crisis resolution mechanisms, such as the negotiated settlement of business crises introduced by recent Italian and European legislative reforms. While these reforms are aimed at facilitating smoother crisis management processes and reducing the stigma associated with insolvency, they also carry risks if used opportunistically by business owners.

Specifically, tools such as the composizione negoziata allow firms to negotiate with creditors outside the court system, extending the time available to address their financial issues. However, when these tools are used in bad faith or as a stalling tactic, they may prolong a company's financial distress rather than resolving it, further exacerbating the crisis. In this context, and goven these considerations, the empirical analysis demonstrated how a multi-algorithmic and data-driven approach can help identify firms that are at risk of leveraging such mechanisms not to resolve their crisis, but rather to delay the inevitable.

In conclusion, the empirical part underscores the complexity, in real life, of predicting corporate insolvency and managing crises in a timely and effective manner. It also emphasizes the critical role that accurate, contextually analyzed data plays in this process and the study calls for a balanced approach, where advanced AI and ML techniques are combined with hands-on, context-specific insights to ensure that the predictive models developed are both accurate and actionable. This blend of technological innovation and managerial acumen can offer companies the foresight needed to make informed, timely decisions that safeguard their survival and their long-term growth and sustainability. The study's findings also highlight the importance of addressing the potential pitfalls of false positives and the manipulation of financial data, urging businesses to adopt more transparent, ethical practices in their reporting and crisis management strategies. Moreover, it underscores the importance of equipping managers with both the technological tools and the skills needed to interpret these predictive insights, ensuring that they can take decisive, informed actions.

3.1. Decision Support

This study adopted a dual approach: first, an overarching analysis utilizing the full spectrum of economic-financial data available (over 3 million data points in total), and second, a dynamic, year-by-year approach considering data points from specific years (-1, -3, and -5) relative to insolvency onset. Both approaches demonstrated robust predictive capacity, capturing even subtle signals of distress up to five years before a firm's definitive insolvency. The improvement in model accuracy—from 0.78 (Altman, 1968) to 0.87—represents a foundational achievement, advancing not only methodological rigor and predictive sophistication in scientific discourse but also exemplifying to professionals the concrete, data-driven capabilities of advanced analytics, which are too often confined to academic contexts. This shift in accuracy highlights the potential to bridge the gap between theory and application, proving to the professional and managerial communities the untapped potential of data-driven decision-making: while the development of tailored, high-performance decision support systems may carry prohibitive costs, especially for SMEs, it is essential to drive a mindset shift among corporate stakeholders. The imperative, in fact, is not merely about adopting AI but integrating it as a strategic tool that supplements, rather than supplants, human decision-making. AI systems, when effectively implemented, could serve opportunistic yet vital functions that ensure business continuity in increasingly challenging environments. Intelligent, fully digitalized systems can also help reallocate resources and personnel to areas where human insight is critical, thereby lowering costs associated with physical infrastructure and staff for maintenance and operations. By leveraging AI and ML for targeted operational efficiencies, companies can achieve a substantial reduction in overhead. Intelligent systems could in fact enable automated monitoring and predictive maintenance, reducing the costs of equipment upkeep, and replacing routine tasks in financial monitoring with algorithm-driven accuracy. Performance indicators like Return on Assets (ROA), Cost-to-Income Ratios (CIR), and Profit Margin

Improvement Rates could become benchmarks achievable through optimized, AI-powered operational models that integrate real-time insights and anomaly detection.

Another key contribution is that of of Explainable AI (XAI) in shaping the future of these intelligent systems. XAI, in fact, can provide tailored tools that align with the unique characteristics of individual businesses or even entire sectors, creating a customizable set of critical indicators that are specifically designed to highlight the early signs of distress. By integrating ML models with explainability tools like SHAP, managers can gain deeper insights into the variables that most significantly impact their firm's financial health: for example, the SHAP violin diagrams used in this study offered a visual representation of feature importance, giving a clear indication of which economic and financial variables played a decisive role in predicting crisis or insolvency. This form of representation serves as an advanced analytic tool and provides actionable insights—specific, tailored information that managers can use to guide decision-making during times of distress. Additionally, the use of XAI ensures that these advanced AI-driven models are not just black-box tools but are fully transparent and interpretable.

3.2. AI-Driven Financial Management in Times of Crisis

The importance of financial management during periods of crisis cannot be overstated, and AI-driven tools have become crucial also in this realm. This section highlights how AI, with its ability to tailor predictions to specific business needs, also supports strategic financial planning, fostering resilience for companies and instilling confidence among stakeholders, including shareholders in publicly listed firms. The implementation of AI-driven solutions, in fact, can profoundly transform the financial management of companies, especially when liquidity is critical: the ability to forecast financial distress, manage cash flows, and make informed decisions about resource allocation is crucial to maintaining business continuity and, in some cases, averting insolvency. In crisis scenarios, companies are vulnerable to disruptions in cash flow—whether due to delays in receivables, declining sales, or rising operational costs. So, a goo financial management is especially relevant for optimizing operational costs. AI models can identify spending patterns, detect inefficiencies, and recommend areas for cost reduction without jeopardizing the company's core operations, and, in the context of Italian SMEs, which often face high operational costs and slim margins, such insights can be crucial.

For publicly listed companies, the implications of AI-driven predictive models are even more significant: shareholders and the financial market are highly sensitive to financial disclosures and future projections and the ability to predict financial distress or liquidity issues in advance allows publicly traded companies to communicate more effectively with shareholders and the market. With AI-driven models, firms can provide more accurate and transparent financial forecasts, which enhances market confidence and reduces volatility. For example, during a financial downturn, an AI-driven predictive model could offer a clear projection of the company's liquidity needs and cost-saving measures, allowing management to present a well-considered recovery plan to the market; this transparency can stabilize stock prices and protect the company's market valuation, as investors and analysts are reassured by the firm's proactive and data-driven approach to crisis management. In this context, AI can play a vital role in shaping investor relations. Additionally, AI can improve access to financing: with the ability to offer detailed financial forecasts based on predictive analytics, companies can present a more compelling case to potential lenders or investors, whether through equity or debt markets. Transparent and accurate projections then build trust with financial institutions, potentially leading to better credit terms or higher confidence in secondary market trading, further strengthening the

company's financial position during times of distress. In conclusion, integrating AI into financial management processes could offer substantial benefits to companies, both private and publicly listed, by enhancing liquidity management and enabling smarter, more strategic decisions. Whether through improved cash flow forecasting, cost optimization, or providing shareholders with accurate and transparent guidance on financial health, AI-driven predictive solutions offer managers the tools needed to act swiftly.

4. Scientific and Practical Implications

Managers, today, are often faced with information overload, and this study's focus on feature selection helps streamline decision-making by emphasizing only the most critical metrics: this, prevents the distraction of excessive data and allows managers to take decisive actions based on clear insights. Furthermore, the empirical results illuminated several key areas where AI-driven predictive models can profoundly impact business practices. One of the central findings has been the boost provided by the synergy between AI tools and human decision-making:

- *Boost of AI and Human Decision Synergy;*
  AI should not be seen as a replacement for human judgment but rather as a tool that enhances the ability of managers to make informed decisions and it could serve as an early detection system, providing critical insights into areas that may otherwise be overlooked. However, these insights are only actionable when paired with managerial literacy—the ability of business leaders to interpret, contextualize, and apply these insights effectively. So, the human-AI collaboration is most effective when managers are equipped with the necessary skills to understand and leverage AI outputs and the need for managerial literacy in AI and data analytics is more pressing than ever. Training managers to be proficient in interpreting AI-driven insights is paramount to ensuring that AI tools are used to their full potential;

- *Operational Efficiency and Tailored Solutions;*
  Another practical implication of the study is the increased operational efficiency enabled by AI-driven predictive models. In fact, they are capable of identifying inefficiencies in a company's processes and suggesting targeted interventions. The ability to streamline operations is particularly beneficial for SMEs, which may lack the resources for comprehensive risk management strategies or in-depth operational analyses. Additionally, feature selection, as highlighted by the study, allows businesses to prioritize critical metrics and areas;

- Continuity Through Discontinuity;
  A particularly compelling insight from the study is also the concept of creating value through discontinuity: the ability of predictive models to foresee disruptions allows businesses to capitalize on these moments as opportunities for transformation and innovation. During periods of economic uncertainty, such as the covid-19 pandemic, businesses that could pivot quickly to new markets or adapt their operations to changing conditions were more likely to survive and thrive; as this, predictive analytics allows firms to identify emerging trends and criticalities and to adjust their strategies accordingly, ensuring that they remain competitive even during turbulent times. This proactive approach to crisis management enables firms to turn potential threats into opportunities for growth and strategic differentiation. For SMEs, in

particular, the agility inherent in their smaller size can be a significant advantage when combined with the foresight provided by AI-driven tools: by embracing discontinuity as part of the business lifecycle, they can develop resilience and flexibility, ensuring their long-term viability;

- *Managerial Literacy and Competency;*
  The study also sheds light on the managerial competency gap that exists within many SMEs, particularly when it comes to the implementation of AI-driven tools, emphasizing the need for education and training programs that equip managers with the skills needed to interpret AI outputs and integrate them into their strategic planning processes. However, also algorithms should be "*trained*" in order to create taylor-made tools for every situation. The study also advocates for a collaborative learning environment, where managers, scholars, data scientists, and AI systems work together to continuously improve both the models and real-world applications;

- *Ethical and Sustainable Management;*
  The findings highlight the importance of ethical AI deployment, particularly in ensuring transparency and accountability in decision-making processes. By optimizing resource use and improving operational efficiency, AI can in fact contribute to both financial sustainability and environmental responsibility. This alignment with broader global sustainability goals, such as those outlined by the UN Sustainable Development Goals, ensures that AI is used not just for profit but also for the long-term well-being of businesses and society;

- *Bridging the Gap Between Science and Practice;*
  This has emerged as a central theme of this Thesis, a call to action that addresses a persistent issue in academic research. The analysis of the existing literature has made it evident that scientific inquiry often emphasizes theoretical advancements, sophisticated models, and methodologies that, while intellectually stimulating, are frequently far removed from real-world applicability. These models, in fact, are useful for advancing knowledge within academic circles, but they tend to require significant investments in time, resources, and strategic alignment to be practically implemented in business environments. This disconnect between theory and practice means that many innovations remain confined to academic papers, without impacting real-world business operations in a meaningful way. This Thesis seeks to disrupt that paradigm: the aim here is not just to achieve theoretical breakthroughs—such as the AI model used in this research demonstrating an impressive 0.87 accuracy rate in predicting insolvency before its manifestation (in the main empirical test) —but to translate these scientific insights into actionable tools that can actually be used by businesses. The goal is then to offer more than just intellectual advancement: it's about providing SMEs, and businesses in general, with concrete tools and actionable insights that can improve decision-making and support long-term business health. Predictive models are only useful if they can be deployed in practice, offering value beyond academic journals and actually helping managers identify risks, manage crises, and maintain operational continuity. It's fundamental to turn scientific results into practical solutions and the findings of this study are designed with real-world application in mind: the models were not only constructed to push the boundaries of academic knowledge but to be useful in day-to-day business operations, allowing companies to identify early warning signs of distress and respond appropriately. It is then imperative that the academic world continues

to reflect on how its innovations can be made accessible to practitioners in a way that doesn't require years of integration or heavy financial outlays. In addition to closing the gap between theory and practice, this research also addresses the disconnection between academia and the professional world. Too often, the academic community assumes a detached position, using established theories to compensate for a lack of direct business experience or technical proficiency. This approach risks developing frameworks that do not fully consider the nuances of reality. In contrast, those who work within the business environment daily often bring practical knowledge that academics may overlook. This Thesis, therefore, calls for a deeper integration between academic insights and professional expertise, ensuring that research is grounded in practical realities and that those on the front lines of business crises can benefit from academic advancements.

In conclusion, the research represents both the culmination of a meticulous academic journey and a stepping stone for future work. While it serves as the conclusion of a formal research path, it simultaneously opens the door to new questions, new applications, and new partnerships between the scientific community and the business world. This journey from theory to practice, from academic rigor to real-world impact, is not just the end of a research project, but it is the beginning of an ongoing dialogue about how we can better align scientific advancements with practical business needs.

5. Limitations and Future Directions

It is essential to acknowledge the limitations that emerged during the research process as they can offer crucial insights for guiding future research efforts and enhancing the practical applicability of AI-driven predictive models.

5.1. Data Limitation

One of the core challenges identified is the reliance on historical financial data, which frequently lacks the real-time adaptability and contextual depth required. This issue becomes even more pronounced when considering the quality and timeliness of data submitted to official registries such as the Chamber of Commerce, where reporting delays, omissions, or inaccuracies can further distort the predictive accuracy of the models. In this dissertation's analysis, a clear discrepancy was often observed between the financial data available through traditional channels and the actual performance of firms. For example, the financial data used to train ML models often failed to reflect impending liquidity constraints or operational inefficiencies, as these were masked by delayed or manipulated reporting, especially in the case of firms experiencing distress. Furthermore, preprocessing of the dataset, which involved standardizing, cleaning, and removing outliers, revealed that data quality issues, such as incomplete or inconsistent entries, frequently skewed initial predictions. Without proper optimization during this preprocessing phase, these inconsistencies would have compromised the integrity of the models, leading to biased or inaccurate forecasts. Moreover, the matrices used, further highlighted limitations in relying solely on historical data. False positives and false negatives, in particular, illustrated how certain companies, flagged as healthy or in distress, often presented data inconsistencies that hindered clear interpretation. To enhance the robustness of future models, it is then essential to integrate more diverse and comprehensive data streams. Finally, optimization techniques, such as feature selection and dimensionality reduction, play a crucial role in refining these models: employing

methods like Recursive Feature Elimination or Principal Component Analysis can help identify the most significant predictors of financial distress, reducing the risk of overfitting while improving the model's generalizability across varying sectors and economic contexts. Additionally, advanced data preprocessing techniques, should be more considered to ensure the consistency and quality of the input data.

5.2. Data Integrity and Value

The reliability of the data used remains a crucial factor in determining their accuracy and effectiveness. Issues such as data manipulation, financial fraud, or moral hazard can in fact compromise the integrity of predictive models and opportunistic behaviors, such as manipulating financial reports, can distort the inputs and lead to inaccurate predictions. Future research should focus on developing methods to ensure data integrity, such as using anomaly detection techniques like Benford's Law, which can help identify irregularities in financial data.

5.3. Models Generalizability

The predictive models developed in this study were tailored to the unique context of Italian firms, a focus that, while valuable, presents a limitation in terms of generalizability. Different sectors and regions are characterized by distinct economic, regulatory, and cultural factors that influence corporate performance and risk differently. A model that performs well in the context of European manufacturing SMEs, for instance, might underperform when applied to tech startups in the United States or larger firms in Asia. Future research must address this issue by developing adaptable ML models capable of being better customized to fit the specific needs of different sectors, regions, and even individual firms. Doing so will ensure that these models have broader applicability and can be effectively deployed in a wider array of business environments.

5.4. Technological Constraints

A critical barrier to the implementation of advanced AI methodologies, particularly for SMEs, is the technological infrastructure required to support such systems. Many smaller firms lack the data management systems, IT infrastructure, and skilled personnel needed to effectively deploy and maintain AI-driven predictive models: this technological gap is particularly evident in businesses that may not have the financial resources to invest in cutting-edge technology. As a result, even the most accurate models can be underutilized due to a lack of implementation capacity. Future research should explore cost-effective, scalable AI solutions that are more accessible to smaller firms. This includes not only developing lightweight AI tools that require less computational power but also providing support structures such as training programs or partnerships with larger firms to build the necessary technical capabilities within SMEs.

5.5. Managerial Literacy and Competency

A key limitation is the varying levels of managerial literacy regarding AI, ML, and predictive analytics. Many managers may also harbor mistrust of AI systems, particularly if they feel these systems are "*black boxes*" that are difficult to understand or control. This gap in knowledge and trust can limit the successful deployment and use of AI-driven predictive models. Future research should then place

a greater emphasis on bridging this literacy gap by developing user-friendly AI tools that are accessible to managers without technical backgrounds. Additionally, it is crucial to offer practical training programs aimed at enhancing managerial capacity in utilizing AI for decision-making. The objective here is not to replace human judgment but to augment it, using AI as a supportive tool that enhances managerial insight and decision-making ability.

5.6. Ethical Concerns in the Age of AI and Effective Explainability of Models

As AI technologies continue to evolve and become more embedded in business processes, the ethical implications of their use become increasingly important. There is a growing concern that AI models trained on historical data may inadvertently perpetuate existing biases or introduce new ones: for instance, if a predictive model is trained on data that reflects societal inequalities, it may make decisions that unfairly disadvantage certain groups. Furthermore, there is a risk of moral hazard, where managers might manipulate data or engage in opportunistic behavior to exploit AI systems for personal or organizational gain. Ensuring the ethical deployment of AI in crisis prediction is then crucial, and future research should focus on developing frameworks that prioritize transparency, fairness, and accountability: bias detection and mitigation techniques, along with robust data privacy protocols, should be integral components of any AI-driven system used in corporate governance and crisis management. Furthermore, legislative world have a fundamental role in this context. Another limitation lies in the effective explainability of AI models. However, the growing field of explainable AI offers promising solutions to this problem, and future research should focus on developing techniques that make AI models more transparent and accessible.

5.7. Over-Reliance Risks

One important limitation in the practical implications of AI-driven predictive tools in management is the potential risk of over-reliance on these systems. Managers, who become adept at using innovative AI-based systems for predicting and managing crises, may develop, in fact, an increased level of trust in these technologies[188]. While AI can enhance decision-making, an over-reliance could lead to inefficiencies, particularly when managers assume that AI-generated alerts will always deliver fast and accurate results. This over-dependence is problematic because AI models, despite their sophistication, are not infallible; they may produce outputs based on incomplete, biased, or even manipulated data, which could lead to flawed predictions. Managers must then avoid the temptation to uncritically accept AI-generated insights without applying human oversight. AI-based systems, in fact, no matter how advanced, require results to be contextualized, verified, and tested in real-world scenarios. Moreover, over-reliance can diminish accountability: if managers place too much confidence in AI, they might neglect their responsibility to critically evaluate the data and assumptions underlying

[188] The adoption of advanced AI-based systems for crisis prediction and management can lead managers to develop a progressively high level of trust in these technologies. However, this trust, if not tempered by ongoing critical assessment, risks becoming the initial step toward AI washing. In this context, an uncritical reliance on AI's capabilities can drive behaviors that extend AI's role superficially, where the "*AI-driven*" label is applied strategically to attribute value to practices not genuinely supported by AI. This phenomenon manifests through the promotion of tools and initiatives presented as AI-driven but with only a superficial integration of AI, often aimed at reputational, economic, or performance-related benefits. Such misalignment between the actual technological capability and the externally communicated expectations can lead to a credibility gap, potentially undermining trust in the organization and raising ethical concerns about the authentic use of AI technology.

the predictions. Ethical concerns also arise when AI is used opportunistically, for instance, when models are manipulated to reflect overly positive forecasts based on distorted or inaccurate data.

5.8. Cost and Resource Challenges

Another significant limitation of AI-driven, data-centric research and its practical implementation in managerial contexts is the high cost and resource intensity associated with both data collection and multi-algorithmic processing[189]. Gathering diverse datasets—financial, operational, and external environmental data—can be a time-consuming and costly process, as it often requires access to proprietary databases, extensive data preprocessing, and the integration of heterogeneous data sources. Furthermore, managing a multi-algorithmic approach, which involves testing, validating, and optimizing various ML models, adds complexity, requiring specialized skills, computational resources, and longer timelines. These factors inevitably increase both the financial and time costs of conducting such research. From an inductive research perspective, the inherent inefficiencies in generating theories from data rather than testing pre-defined hypotheses also pose challenges: the exploratory nature of inductive reasoning necessitates continuous data refinement, model iteration, and statistical testing, which can lead to extended research cycles.

Furthermore, when translating this data-driven, AI-based approach to the managerial and corporate environment, additional costs arise, particularly in terms of managerial literacy and the implementation of technological infrastructure. For AI-based systems to be successfully integrated into organizational decision-making, companies must invest in training their managers to develop algorithmic literacy: this involves not only understanding how to use AI tools but also grasping the underlying data processes and model outputs—tasks that are not immediately intuitive and can require extensive training programs, even if there is the use of *"user-friendly"* tool as SHAP. Additionally, organizations need to establish the necessary IT infrastructure, including data management systems, cloud storage, and computational tools, to support the practical deployment of AI models. These investments in technology and human capital further increase the overall costs and complexity of implementation. Moreover, the time required to implement these AI-based systems and train personnel should not be underestimated. In conclusion, as data is increasingly seen as a critical asset—often referred to as the *"new gold"*—the costs associated with acquiring, storing, and processing large-scale datasets continue to rise: this gap between the academic potential of AI-driven approaches and their practical implementation highlights the need for cost-effective solutions, and bridging this gap requires both researchers and practitioners to develop more accessible frameworks that can be deployed without prohibitive costs, while maintaining scientific rigor and practical relevance.

6. Future Research

Future research in this domain must then build on the significant advancements made by this study, expanding both the academic and managerial implications of AI-driven predictive models. One of the most pressing areas for future exploration lies in diversifying the data sources used to train and refine these models. Historically, financial data has been the backbone of predictive analytics, but it

[189] Another key challenge regarding the multi-algorithmic processing involves the testing of multiple ML models (e.g., decision trees, neural networks, XGBoost) to optimize predictive accuracy (more than 4). This process demands substantial computational power, often requiring cloud computing or specialized hardware (GPUs, TPUs), as well as expertise in algorithm tuning, model validation, and ensemble methods. Moreover, algorithmic models need to be constantly refined and updated as new data becomes available, which adds to the ongoing maintenance costs.

often lacks the real-time fluidity and contextual depth required to capture the full spectrum of variables that influence a company's health. The complexities of modern business environments demand models that can incorporate not only traditional financial metrics but also dynamic data sources, and the integration of real-time data, processed through advanced algorithms, would allow businesses to respond more proactively, rather than reactively, to early signs of distress. Moreover, this approach could address one of the key limitations observed in the empirical part of this study: often, the financial data reported by companies, especially SMEs, is delayed or inaccurately presented due to factors like poor internal controls or intentional manipulation, creating a false picture of financial health. By improving data optimization and preprocessing, predictive models can become more robust and reliable.

In addition to broadening the data landscape, future research should focus on developing sector-specific models as the success of AI-driven models depends largely on how well they can be adapted to reflect the specific needs and challenges of different industries and geographic regions. This customization would not only improve the generalizability of the models but also make them more actionable, empowering managers to leverage precise, data-backed insights in real-time decision-making. Perhaps the most crucial and promising direction for future research, however, is the further exploration of the synergy between AI and human judgment: AI does not have a "*heart*"; it cannot fully comprehend the nuances of organizational culture, leadership dynamics, or the moral and ethical implications of business decisions. This is where the human element remains irreplaceable. The true value of AI, then, lies in its ability to be a complement of human decision-making—augmenting managerial capabilities with data-driven insights, but leaving the critical strategic decisions in the hands of human leaders who bring experience, intuition, and ethical judgment to the table. As we have seen in the empirical analysis, AI can detect insolvency risks with high accuracy years before they manifest in traditional financial indicators; however, AI's outputs are only as effective as the decision-making frameworks they inform. Explainable AI represents an exciting frontier in this respect.

Another area ripe for future exploration is the ethical deployment of AI in corporate crisis and insolvency prediction. This study has already identified potential ethical challenges, such as the risk of AI reinforcing biases or being manipulated to serve opportunistic behaviors; the development of robust ethical frameworks will be then essential, as they should ensure that AI is deployed in ways that promote fairness and accountability.

7. Conclusion: AI and the Evolution of Management Practices

This Thesis has demonstrated the promising potential of AI in both academic and managerial contexts, and the implementation of AI also holds far-reaching positive implications for trust, transparency, reliability, and market stability, as it improves stakeholder relationships and promotes greater transparency in public markets, especially for listed companies, where accurate predictions and early warnings can significantly enhance investor confidence and corporate governance (Moreno & Camacho-Minano, 2024). Furthermore, SMEs can leverage AI-driven systems to ensure business continuity and enhance profitability prospects without losing the distinct characteristics that define their operations: by offering tailored solutions, AI provides these businesses with a strategic edge that was previously accessible only to larger corporations with more extensive resources.

Nevertheless, this Thesis also highlights that these technological advancements come with significant costs, both in terms of financial investment and resource demands. The barriers to entry—including challenges related to data collection, algorithmic processing, and practical implementation—

are substantial and cannot be overlooked. It is within this context that the limitations of AI become clear, emphasizing the need for a collaborative, interdisciplinary approach that unites academia, the corporate world, legislators, and professionals to develop cost-effective, accessible, and streamlined solutions that position AI as a critical driver of corporate progress.

In conclusion, while new technologies offer immense potential for transforming business management and academic research, their success hinges on the development of a structured framework that addresses limitations and makes AI more widely accessible across different sectors and business scales. The capacity of new technologies to become a transformational force for organizations and industries alike depends, infact, on the active collaboration of multiple stakeholders, working together to create an ecosystem where data-driven and decision-making becomes a powerful tool for fostering growth: although, as Galbraith once noted, forecasting is underpaid and risky, it has become, in our era, the seed for business continuity and survival.

References

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488-1542.

Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks, 6*(2), 38.

Adelaja, O. A., & Alkattan, H. (2023). Operating artificial intelligence to assist physicians diagnose medical images: A narrative review. *Mesopotamian Journal of Artificial Intelligence in Healthcare*, Vol. 2023, 45–51.

Adizes, I. (2004). *Managing corporate lifecycles*. The Adizes Institute Publications.

Agarwal, V., & Tafler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance*, 32, 1541–1551.

Agostini, L., & Nosella, A. (2017). Enhancing radical innovation performance through intellectual capital components. *Journal of Intellectual Capital, 18*(4), 789-806.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). Artificial intelligence: Implications for business strategy. *MIT Sloan School of Management*.

Alberici, A. (1975). *Analisi dei bilanci e previsione delle insolvenze: affidabilità bancaria e informativa del mercato mobiliare*. Volume 4 di Studi di economia industriale e di mercato finanziario. ISEDI.

Alfaro-Cortés, E., Pérez-Fernández, R., & Alfaro-Navarro, J. L. (2020). Big Data and Analytics for Smart Cities. *Journal of Big Data*, 7(1), 1-19.

Aliaj, T., & Piersanti, S. (2020). *Firms default prediction with machine learning*. In Proceedings from the Conference on Machine Learning Applications.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589-609.

Altman, E. I. (1993). *Corporate financial distress and bankruptcy: A complete guide to predicting and avoiding distress and profiting from bankruptcy.* Wiley.

Altman, E. I., & Le Fleur, J. K. (1985). I modelli di previsione delle insolvenze: Le loro applicazioni alla gestione d'impresa. *Finanza, Marketing, Produzione, 4*, 77-93.

Altman, E. I., Danovi, A., & Falini, A. (2013). Z-score model's application to Italian companies subject to extraordinary administration. *Journal of Applied Finance, 23*(1), 1-10.

Altman, E. I., Hotchkiss, E., & Wang, W. (2017). *Corporate financial distress, restructuring, and bankruptcy: Analyze leveraged finance, distressed debt, and bankruptcy.* John Wiley & Sons.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting*, 28(2), 131-171.

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance, 18*(3), 505-529.

Altman, E. I., Narayanan, P., & Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance, 21*(11-12), 1721-1742.

Aly, S., Alfonse, M., & Salem, A. B. M. (2022). Bankruptcy prediction using artificial intelligence techniques: A survey. In D. A. Magdi, Y. K. Helmy, M. Mamdouh, & A. Joshi (Eds.), *Digital Transformation Technology* (Vol. 224, pp. 327-344). Springer.

Anandarajan, M., Lee, P., & Anandarajan, A. (2004). Bankruptcy prediction using neural networks. In *Business Intelligence Techniques*.

Ansoff, I. (1965). *Corporate strategy: An analytic approach to business policy for growth and expansion*. McGraw-Hill.

Anthony, J. H., & Ramesh, K. (1992). Association between accounting performance measures and stock prices: A test of the life cycle hypothesis. *Journal of Accounting and Economics, 15*(2-3), 203–227.

Antulov-Fantulin, N., Lagravinese, R., & Resce, G. (2021). Predicting bankruptcy of local government: A machine learning approach. J*ournal of Economic Behavior & Organization, 183*, 681-699.

Aoki, S., & Hosonuma, Y. (2004). Bankruptcy prediction using decision tree. In *The Application of Econophysics*. Springer.

Argenti, J. (1976). Corporate planning and corporate collapse. *Long Range Planning, 9*(6), 12–17.

Aruldoss, M., Manjula, M., & Venkatesan, V. (2011). *A business intelligence model to predict bankruptcy using financial domain ontology with association rule mining algorithm*. arXiv preprint.

Aruldoss, M., Travis, M. L., & Venkatesan, V. P. (2015). A reference model for business intelligence to predict bankruptcy. *Journal of Enterprise Information Management*.

Atiya, A. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks, 12*(4), 929-935.

Aydin, N., Sahin, N., Deveci, M., & Pamucar, D. (2022). Prediction of financial distress of companies with artificial neural networks and decision trees models. *Machine Learning with Applications, 10*.

Azungah, T. (2018). Qualitative research: Deductive and inductive approaches to data analysis. *Qualitative Research Journal*, *18*(4), 383–400.

Bacher, M., et al. (2023). Advanced algorithmic techniques for AI-driven predictive modeling. In *Proceedings of the International Conference on Machine Learning Applications*.

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 8*3, 405-417.

Barney, J. B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99-120.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71-111.

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies, 1*(4), 267-284.

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education, 33*, 1-42.

Ben Jabeur, S., & Serret, V. (2023). Bankruptcy prediction using fuzzy convolutional neural networks. *Research in International Business and Finance, 64*, 101-844.

Biau, G. (2012). Analysis of a random forests model. *The Journal of Machine Learning Research, 13*, 1063-1095.

Blazek, J. (2014). *Financial planning and analysis and performance management*. Wiley.

Bottiglia, R., & Paltrinieri, A. (2010). *Financial crisis and ownership of global banks.* In R. Bottiglia, E. Gualandri, & G. N. Mazzocco (Eds.), Consolidation in the European financial industry. Palgrave Macmillan, 91-112.

Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., & Mezzogori, D. (2019). Machine learning models for bankruptcy prediction in Italy: Do industrial variables count? *Dipartimento di Discipline Matematiche, Finanza Matematica ed Econometria Working Paper No. 19/3*. Università Cattolica del Sacro Cuore.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-32.

Broby, D. (2022). The use of predictive analytics in finance. *The Journal of Finance and Data Science, 8*, 145–161.

Brogi, M., Lagasio, V., & Porretta, P. (2022). Be good to be wise: Environmental, social, and governance awareness as a potential credit risk mitigation factor. *Journal of International Financial Management & Accounting, 33*(3), 522–547.

Brynjolfsson, E., & McAfee, A. (2017). The business of artificial intelligence: What it can—and cannot—do for your organization. *Harvard Business Review*.

Cefis, E., Bartoloni, E., & Bonati, M. (2020). Show me how to live: Firms' financial conditions and innovation during the crisis. *Structural Change and Economic Dynamics, 52*, 63–81.

Chandler, A. D. (1962). *Strategy and structure: Chapters in the history of the industrial enterprise*. MIT Press.

Chatterjee, S., Chaudhuri, R., & Vrontis, D. (2021). Does data-driven culture impact innovation and performance of a firm? An empirical examination. *Annals of Operations Research, 333*(2), 1-26.

Chatterjee, S., & Kar, A. K. (2024). Artificial Intelligence Decision-Making Transparency and Employees' Trust: The Parallel Multiple Mediating Effect of Effectiveness and Discomfort. *Sustainability*, 15(21).

Chen, M. Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers and Mathematics with Applications*, 62(10).

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Coda, V. (1977). Crisi d'impresa e comportamento direzionale. In AA.VV., *Crisi d'impresa e sistemi di direzione*. Etas, Milano.

Codice della crisi d'impresa. Testo del Decreto Legislativo 12 gennaio 2019, n. 14, aggiornato al D.Lgs. 13 settembre 2024, n. 136.

Colli, A. (2003). *The history of family business, 1850-2000*. Cambridge University Press.

Cox, D. R. (2006). *Principles of statistical inference*. Cambridge University Press.

Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. *Nature, 538*, 311–313.

Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. *Cambridge University Press.*

Dallocchio, M., Ferri, S., Tron, A., & Vizzaccaro, M. (2020). The use of Z-Score to predict UTP loans. *Corporate Ownership & Control,* 18(1), 163–178.

Dasilas, A., & Rigani, A. (2024). *Machine learning techniques in bankruptcy prediction: A systematic literature review*. *Expert Systems with Applications, 255*.

Davenport, T. H., & Bean, R. (2018). Big Data and AI Executive Survey 2018. MIT Sloan Management Review.

Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Review Press.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108–116.

Dellepiane, U., Di Marcantonio, M., Laghi, E., & Renzi, S. (2015). Bankruptcy prediction using support vector machines and feature selection during the recent financial crisis. *International Journal of Economics and Finance*, 7(2), 182–195.

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3–4), 197–387.

De Sarno, M. (1982). *Equilibrio e crisi delle imprese*. Volume 1 di Ist. ragioneria-Univ. Napoli; Volume 1 di Pubblicazioni dell'Istituto di ragioneria, Facoltà di economia e commercio, Università di Napoli. CEDAM.

Di Ciaccio, A., & Cialone, G. (2019). Insolvency prediction analysis of Italian small firms by deep learning. *International Journal of Data Mining & Knowledge Management Process, 9*(6), 1-12.

Dikshit, A., & Pradhan, B. (2021). Interpretable and explainable AI (XAI) model for spatial drought prediction. *Science of the Total Environment, 801*.

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: Gradient boosting with categorical features support*. arXiv preprint.

Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning*. arXiv.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.

du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing, 73*(10–12), 2047–2060.

Efron, B. (1975). The efficiency of logistic regression compared to normal discriminant analysis. *Journal of the American Statistical Association*, 70(352), 892–898.

Elhoseny, M., Metawa, N., Sztano, G., & El-hasnony, I. M. (2022). Deep learning-based model for financial distress prediction. *Annals of Operations Research*, 316(1), 11–23.

Enkhtuya, T., & Kang, D.-K. (2023). Bankruptcy prediction with explainable artificial intelligence for early-stage business models. *The International Journal of Internet, Broadcasting and Communication, 15*(3), 58–65.

Fayol, H. (1916). *Administration industrielle et générale*. Paris: Dunod et Pinat.

Ferri, G. (2022). Corporate crises and insolvency prediction: A critical review. *Journal of Business Research*, 140, 276–290.

Fisher, L. (1959). Determinants of risk premiums on corporate bonds. *Journal of Political Economy*.

FitzPatrick, P. J. (1932). A comparison of ratios of successful industrial enterprises with those of failed companies. *The Certified Public Accountant, 12*(10), 598–605.

Floreani, A. (2013). *Risk measures and capital requirements: A critique of the Solvency II approach*. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 38(2), 189–212.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., & Schafer, B. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines,* 28(4), 689–707.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.

Gajdosikova, D., & Gabrikova, B. (2023). Corporate bankruptcy prediction: A systematic literature review and comprehensive bibliometric analysis. *Economics & Law, 5*, 105–132.

Gartner Research. (2020). *A holistic ecosystem approach to AI*. International Electrotechnical Commission (IEC).

Gartner. (2020). *How SMEs can achieve a competitive advantage using predictive analytics*. Gartner Research.

Gaurav, A., Gupta, B. B., Bansal, S., & Psannis, K. E. (2024). Bankruptcy forecasting in enterprises and its security using hybrid deep learning models. *Cyber Security and Applications*.

Gelfand, S. B., Ravishankar, C. S., & Delp, E. J. (1991). An iterative growing and pruning algorithm for classification tree design. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(2), 163–174.

Gennaioli, N., Shleifer, A., & Vishny, R. W. (2013). A model of shadow banking. *Journal of Finance, 68*(4), 1331–1363.

Gerlach, M. (1992). *Alliance capitalism: The social organization of Japanese business*. University of California Press.

Giannessi, E. (1979). *Appunti di economia aziendale: con particolare riferimento alle aziende agricole*. Pacini.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). *Explaining explanations: An overview of interpretability of machine learning*. arXiv preprint.

Giraud, L., Zaher, A., Hernandez, S., & Akram, A. A. (2022). The impacts of artificial intelligence on managerial skills. *Journal of Decision Systems, 32*(3), 566–599.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. Advances in *Neural Information Processing Systems*, 2672–2680.

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review*, 33(3), 114-135.

Gu, H. (2022). Early warning algorithm of financial crisis based on random forest. In *The International Conference on Forthcoming Networks and Sustainability (FoNeS 2022)*, 1-6

Guatri, L. (1986). *Crisi e risanamento delle imprese* (Vol. 2). Giuffrè.

Guenster, N., Bauer, R., Derwall, J., & Koedijk, K. (2011). The economic value of corporate eco-efficiency. *European Financial Management, 17*(4), 679–704.

Gunasekaran, A., Subramanian, N., & Rahman, S. (2017). Supply chain resilience: Role of complex network design, disruption events, and strategic decision making. *International Journal of Production Research, 53*(22), 6822–6837.

Hartini, S., Rustam, Z., Saragih, G. S., & Segovia Vargas, M. J. (2021). Estimating probability of banking crises using random forest. *IAES International Journal of Artificial Intelligence,* 407-413.

Hasan, I., Kobeissi, N., & Wang, H. (2015). Corporate life cycle and financial distress. *Review of Quantitative Finance and Accounting, 44*(3), 463–488.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. IEEE Intelligent Systems, 13(4), 18–28.

Holopainen, M., & Toivonen, M. (2012). Weak signals: Ansoff today. *Futures, 44*(3), 198-205.

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process, 5*(2), 1–11.

Idowu, N., Maxwell, W. P., & Owoeye, O. (2023). Cognitive adaptability and sales growth of selected small and medium scale enterprises in Lagos State, Nigeria. *International Journal of Research and Innovation in Social Science*.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., & Abuhaija, B. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences, 622*, 178–210.

Imelda, E., & Alodia, I. (2017). Analysis of Altman model and Ohlson model in predicting financial distress of manufacturing companies in the Indonesia stock exchange. *Indian-Pacific Journal of Accounting and Finance, 1*, 51–63.

Jiang, Y., Shi, M., & Xu, L. (2020). Machine learning in corporate insolvency prediction: A review. *Journal of Risk and Financial Management*, 13(10), 249.

Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*(2), 193–228.

Jones, S., Johnstone, D., & Wilson, R. (2017). Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting, 44*(1-2), 3–34.

Joshi, S., Ramesh, R., & Tahsildar, S. (2018). A bankruptcy prediction model using random forest. In Proceedings of the IEEE Second International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE.

Kaczmarek, J., Kolegowicz, P., & Szymla, W. (2022). Financial analysis in modern enterprise management. *Journal of Business Research, 120*, 84–91.
Kamruzzaman, M. M., & Alruwaili, O. (2022). AI-based computer vision using deep learning in 6G wireless networks. *Computers and Electrical Engineering,* 102.

Kapounek, S., Hanousek, J., & Bily, O. (2022). Challenges of financial prediction models in the context of modern AI. *Finance Research Letters, 48*, 265–287.

Kim, M. J., & Han, I. (2003). The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. *Expert Systems with Applications*, 25(4), 637–646.

Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and ada-boosted decision tree models. *Economic Modelling*, 36, 354–362.

Kiron, D., & Unruh, G. (2018). The convergence of digitalization and sustainability. *MIT Sloan Management Review, 59*(4), 1–5.

Kleinbaum, D. G., Kupper, L., Nizam, A., & Rosenberg, E. (2013). *Applied regression analysis and other multi-variable methods*. Cengage.

Kliestik, T., Kovalova, E., & Lăzăroiu, G. (2022). *Cognitive decision-making algorithms in data-driven retail intelligence: Consumer sentiments, choices, and shopping behaviors*. *Journal of Self-Governance and Management Economics,* 10(1), 30–42.

Ko, W., & Lin, C. (2006). The impact of globalization on the business environment. *Global Journal of Business Research, 12*(3), 84–91.

Koh, S. C. L., Gunasekaran, A., & Saad, S. M. (2015). *Lean manufacturing: Perspectives, methods, and practices*. Springer International Publishing.

Kou, G., Xu, Y., Peng, Y., Shen, F., Chen, Y., & Chang, K. (2021). Bankruptcy prediction models: Data, model evolution, and impacts of COVID-19. *Journal of Business Research, 142*, 126–143.

Kshetri, N. (2022). *The global rise of online devices, cyber crime and cyber defense: Enhancing ethical actions, counter measures, cyber strategy, and approaches*. University of Missouri, St. Louis.

Kubenka, M., Capek, L., & Sejkora, J. (2021). Bankruptcy prediction using accounting data. *Business and Economic Horizons, 17*(2), 167–185.

Kuizinienė, D., Krilavičius, T., Damaševičius, R., & Maskeliūnas, R. (2022). Systematic review of financial distress identification using artificial intelligence methods. *Applied Artificial Intelligence*, 36(2).

Kumar, R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research*, 180(1), 1–28.

Lasfer, M., & Remer, R. (2010). Corporate life cycle, governance and firm value. *Journal of Corporate Finance, 16*(5), 621–633.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature,* 521(7553), 436–444.

Lee, S. M., Trimi, S., & Kim, C. (2020). Innovation and entrepreneurship in the digital age: Digital transformation strategies in business and society. *Business Horizons, 63*(6), 585–594.

Li, Y. C., & Wang, Y. F. (2018). Machine learning methods of bankruptcy prediction using accounting ratios. *Open Journal of Business and Management,* 6(1), 1–20.

Lin, B., Liu, C., Tan, K. J. K., & Zhou, Q. (2020). CEO turnover and bankrupt firms' emergence. *Journal of Business Finance & Accounting*, 47(9–10), 1238–1267.

lo Conte, D. L., & Sancetta, G. (2022). Predicting corporate bankruptcy using Altman Z'-Score and ratio analysis in the manufacturing sector: Evidence from the Italian case. In Proceedings of the 40th IBIMA Conference, 1565–1574.

lo Conte, D. L., & Sancetta, G. (2023). Ethical businesses in an ethical world: does the adoption of the Italian "legality rating" affect firms' insolvency risk?. 26th Excellence in Services International Conference (EISIC), 31st August and 1st September Paisley, Scotland, UK.

lo Conte, D. L., & Sancetta, G. (2024). Safeguarding Ethics, Sustaining Success: the Role of Italy's Legality Rating in Mitigating Firms' Insolvency Risk, in Ciasullo, M. V., Martin, J., Brunetti, F. *Embracing Sustainability Management Through Excellence in Services.* Springer. Cham, Switzerland. 174-193.

lo Conte, D. L., Antonini, V., & Sancetta, G. (2024). *Too Small to Fail? Leveraging AI for Early Insolvency Detection in Italian SMEs.* 27th Excellence in Services International Conference (EISIC), 29th and 30th August, Bergamo, Italy.

lo Conte, D. L., Ricotta, F., & Rinna, G. (2024). *From Data to Action: AI-Enhanced Prediction of Business Crises Integrating Financial and Social Media Insights.* Sinergie-SIMA Conference 2024 "Management of sustainability and well-being for individuals and society", June 13th and 14th, Parma, Italy

lo Conte, D. L., Sancetta, G., & D'Amore, R. (2023). Corporate social responsibility and financial performance: An empirical analysis of the Italian case. In Sinergie-SIMA Management Conference 2023—Rediscovering Local Roots and Interactions in Management, 531–546.

lo Conte, D. L., Sancetta, G., & D'Amore, R. (2024). Enhancing Insolvency Prediction Accuracy: a Random Forest-based Algorithmic Approach. In Sinergie-SIMA Conference 2024 "Management of sustainability and well-being for individuals and society", June 13th and 14th, Parma, Italy.

lo Conte, D. L., Shamim, S., & Sancetta, G. (2024). *Sustainability Alignment in Cross-Border Collaborations: Dynamics, Challenges and Limits.* Sinergie-SIMA Conference 2024 "Management of sustainability and well-being for individuals and society", June 13th and 14th, Parma, Italy.

Lough, W. H. (1917). *Business finance: A practical study of financial management in private business concerns*. Students' ed.

Lozano, R. (2019). Analysing the use of tools, initiatives, and approaches to promote sustainability in corporations. *Corporate Social Responsibility and Environmental Management*.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in *Neural Information Processing Systems*, 30, 4765–4774.

Mahajan, A., & Shukla, K. K. (2023). Analyzing false positives in bankruptcy prediction with explainable AI. In *Proceedings of the 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*.

Makariou, D., Barrieu, P., & Chen, Y. (2021). A random forest based approach for predicting spreads in the primary catastrophe bond market. *Insurance: Mathematics and Economics, 101*(Part B), 140-162.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications* (3rd ed.). Wiley.

Manogaran, G., Lopez, D., Thota, C., Sundarasekar, R., & Abbas, K. (2019). Big data, Hadoop and cloud computing in industrial applications. In Handbook of Research on the IoT, Cloud Computing, and Wireless Network Optimization (pp. 173-205). IGI Global.

Marjani, M., et al. (2017). Big IoT data analytics: Architecture, opportunities, and open research challenges. *IEEE Access, 5*, 5247–5261.

Martin, A., Lakshmi, T. M., & Venkatesan, V. P. (2012). An analysis on business intelligence models to improve business performance. In *IEEE-International Conference on Advances in Engineering, Science and Management.*

McCarthy, J. (2007). "What is Artificial Intelligence?". Stanford University. Retrieved from http://jmc.stanford.edu/articles/whatisai/whatisai.pdf.

Merwin, C. L. (1942). *Financing small corporations in five manufacturing industries, 1926-1936: A dissertation in economics*. National Bureau of Economic Research.

Michalkova, A., Stehel, V., Nica, E., & Durana, P. (2021). Financial analysis and its role in corporate decision-making. *Economics and Management, 9*(2), 276–295.

Mikalef, P., & Gupta, M. (2021). The influence of big data and analytics capabilities on supply chain sustainability. Transportation Research Part E: Logistics and Transportation Review, 145.

Mintzberg, H. (1973). *The nature of managerial work*. Harper & Row.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2).

MIT Sloan Management Review. (n.d.). Algorithmic management: The role of AI in managing workforces. *MIT Sloan Management Review*. Retrieved from https://sloanreview.mit.edu/article/algorithmic-management-the-role-of-ai-in-managing-workforces/

Mitroff, I. I. (2005). Why some companies emerge stronger and better from a crisis: 7 essential lessons for surviving disaster. AMACOM Div American Mgmt Assn.

Molnar, C., König, G., Herbinger, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., Casalicchio, G., Grosse-Wentrup, M., & Bischl, B. (2022). General pitfalls of model-agnostic interpretation methods for machine learning models. *International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers*, 39–68.

Moreno, M., & Camacho-Miñano, M. M. (2024). Impression management in bankruptcy: An analysis of the narratives in listed companies. *Revista de Contabilidad – Spanish Accounting Review*.

Nagtegaal, Rosanna. (2020). The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government Information Quarterly*. 38.

Narvekar, A., & Guha, D. (2021). Bankruptcy prediction using machine learning and an application to the case of the COVID-19 recession. *Data Science in Finance and Economics*, 1(2), 180–195.

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics, 7*,

Nguyen, T. T., Li, Z., & Han, J. (2019). A systematic review of big data-based business analytics in the context of disaster management. *Journal of Business Research*, 101, 627–635.

Nguyen, H. H., Viviani, J.-L., & Ben Jabeur, S. (2023). Bankruptcy prediction using machine learning and Shapley additive explanations. *Review of Quantitative Finance and Accounting*.

Nigro, A., & Vattermoli, D. (2023). *Diritto della crisi delle imprese. Le procedure concorsuali. Nuova ediz.* Il Mulino.

O'Leary, D. E. (2019). *Enterprise resource planning systems: Systems, life cycle, electronic commerce, and risk*. Cambridge University Press.

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.

Oehmichen, J., Puck, J., & Lindner, T. (2023). Collaboration challenges between firm types in AI adoption. *Journal of Business Ethics,* 176(3), 581-597.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131.

Paliwal, M., & Kumar, U. (2009). Neural networks and statistical techniques: A review of applications. *International Journal of Business Intelligence and Data Mining*, *4*(2), 2–17.

Pargent, F., Pfisterer, F., Thomas, J., et al. (2022). Regularized target encoding outperforms traditional methods in supervised machine learning with high cardinality features. *Computational Statistics, 37*, 2671–2692.

Pashley, M. M., & Philippatos, G. C. (1990). The asset-liability matching problem and corporate financial flexibility. *Journal of Banking & Finance, 14*(2-3), 307–321.

Passeri, R., Gentili, S., & Mazzi, R. (2011). *Finanza straordinaria per la crisi d'impresa: Il ruolo della gestione straordinaria negli stati patologici*. Franco Angeli.

Pastor, L., & Veronesi, P. (2003). Stock valuation and learning about profitability. *Journal of Finance, 58*(5), 1749–1789.

Patel, V., & Zhang, S. (2022). *Global convergence and stability of stochastic gradient descent*. arXiv preprint.

Pepe, M. S. (2000). Receiver operating characteristic methodology. *Journal of the American Statistical Association*, 95(449), 308–311.

Perboli, G., & Arabnezhad, E. (2021). A machine learning-based DSS for mid and long-term company crisis prediction. *Expert Systems with Applications, 174*.

Perko, I. (2017). Behaviour-based short-term invoice probability of default evaluation. *European Journal of Operational Research,* 257(3), 1045–1054.

Pinski, M., Hofmann, T., & Benlian, A. (2024). AI Literacy for the top management: An upper echelons perspective on corporate AI orientation and implementation ability. *Electronic Markets*, 34(1), 24.

Porter, M. E. (1985). Competitive Advantage: Creating and Sustaining Superior Performance. New York: Free Press.

Prahalad, C. K., & Hamel, G. (1990). The Core Competence of the Corporation. *Harvard Business Review*, 68(3), 79-91.

Pumplun, L., Tauchert, C., & Heidt, M. (2019). A new organizational chassis for artificial intelligence–exploring organizational readiness factors. International Conference on Wirtschaftsinformatik.

Qiu, M., Li, J., & Yin, G. (2022). Artificial Intelligence Decision-Making Transparency and Employees' Trust: The Parallel Multiple Mediating Effect of Effectiveness and Discomfort. S*ustainability*, 15(21).

Qu, Y., Quan, P., Lei, M., & Shi, Y. (2019). Review of bankruptcy prediction using machine learning and deep learning techniques. *Procedia Computer Science*, 162, 895–899.

Rana, N. P., & Sharma, S. K. (2019). Assessing the impact of AI-driven technologies on marketing capabilities. *Journal of Business Research*, 123, 237-246.

Ravazzin, C., & lo Conte, D. L. (2022). L'accesso alla composizione negoziata. Nomina e funzioni dell'esperto, in Sancetta, G., Baratta, I. A., Ravazzin, C. (2022). *La nuova composizione negoziata per la soluzione della crisi d'impresa.* Giuffrè, Milano. 63-104.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Robbins, D. K., & Pearce, J. A. (1993). Entrepreneurial retrenchment among small manufacturing firms. *Journal of Business Venturing, 8*(4), 301–318.

Roubini, N. (2010). *Crisis economics: A crash course in the future of finance*. Penguin Books Ltd.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215.

Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

Sabir, Z., Raja, M. A. Z., Alhazmi, S. E., Gupta, M., Arbi, A., & Baba, I. A. (2022). Applications of artificial neural network to solve the nonlinear COVID-19 mathematical model based on the dynamics of SIQ. *Journal of Taibah University for Science*, 16(6), 874–884.

Salchenberger, L., Cinar, E. M., & Lash, N. A. (1992). *Neural networks: A new tool for predicting thrift failures*. *Management Faculty Research and Publications*, Marquette University.

Sancetta, G., & Gasbarro, D. (2022). La crisi d'impresa e il piano finanziario. In G. Sancetta, A. Baratta, & C. Ravazzin (Eds.), *La composizione negoziata per la soluzione della crisi d'impresa*. Giuffrè.

Sapountzaki, K., & Wanczura, S. (2018). Disasters as a Trigger for Regime Change: The Impact of Disasters on Political Stability. *International Journal of Disaster Risk Reduction*, 31, 14-24.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.

Schwab, K. (2017). The fourth industrial revolution. *Currency.*

Sciarelli, S. (1995). Ristrutturazione e risanamento di un'impresa di distribuzione di prodotti e servizi finanziari: il caso Finnet. *La crisi d'impresa*, Cedam.

Shapley, L. S. (1953). A value for n-person games. In *Contributions to the Theory of Games* (Vol. 2, pp. 307–317). Princeton University Press.

Sharda, R., Delen, D., & Turban, E. (2020). *Business intelligence, analytics, and data science: A managerial perspective* (4th ed.). Pearson.

Shin, K. S., & Lee, T. S. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications,* 23(3), 321–328.

Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1), 127-135.

Shleifer, A., & Vishny, R. W. (1992). Liquidation values and debt capacity: A market equilibrium approach. *Journal of Finance, 47*(4), 1343–1366.

Sigrist, F., & Leuenberger, N. (2023). Machine learning for corporate default risk: Multi-period prediction, frailty correlation, loan portfolios, and tail probabilities. *European Journal of Operational Research, 305*(3), 1390–1406.

Slatercoach. (2020). *Guide to understanding and navigating the business life cycle*. Retrieved from https://slatercoach.medium.com/guide-to-understanding-and-navigating-the-business-life-cycle b86e8890b3db.

Smith, A., & Anderson, M. (2018). AI, Robotics, and the Future of Jobs. Pew Research Center.

Springate, G. L. V. (1978). Predicting the possibility of failure in a Canadian firm: A discriminant analysis. Theses (Faculty of Business Administration), Simon Fraser University.

Steiber, A., & Alänge, S. (2020). *The Silicon Valley model: Management for entrepreneurship*. Springer Nature.

Sudarsanam, S., & Lai, J. (2001). Corporate financial distress and turnaround strategies: An empirical analysis. *British Journal of Management, 12*(3), 183–199.

Sun, J., Li, H., Huang, Q. H., & He, K. Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems*, 57, 41-56.

Sun, Z. (2019). *Intelligent big data analytics: A managerial perspective*. IGI Global.

Sundström, A. (2024). AI in management control: Emergent forms, practices, and infrastructures. *Critical Perspectives on Accounting, 99*.

Tate, W. L., Ellram, L. M., & Kirchoff, J. F. (2019). Corporate social responsibility reports: A thematic analysis related to supply chain management. *Journal of Supply Chain Management, 55*(1), 3–28.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), 509-533.

Tharenou, P., Donohue, R., & Cooper, B. (2007). *Management research methods*. Cambridge University Press.

Thiebes, S., Lins, S., & Sunyaev, A. (2020). Trustworthy artificial intelligence. *Electronic Markets, 31*(2), 447–464.

Toutaoui, J., Benlian, A., & Hess, T. (2022). Managing paradoxes in bi-modal information technology functions: A multi-case study. *Information Systems Journal*.

Tron, A. (2021). *Corporate financial distress: Restructuring and turnaround*. Emerald Publishing Limited.

Tron, A., Dallocchio, M., Ferri, S., & Colantoni, F. (2023). Corporate governance and financial distress: Lessons learned from an unconventional approach. *Journal of Management & Governance,* 27(2), 425–456.

Tron, A., & Franceschi, L. F. (2023). *La finanza nella crisi reversibile di impresa. Gli strumenti negoziali nel Codice della crisi*. EGEA.

Tseng, F., & Hu, Y. (2010). Applying genetic algorithms in bankruptcy prediction. *Expert Systems with Applications, 37*(2), 1846–1853.

Vega García, M., & Aznarte, J. L. (2020). Shapley additive explanations for NO2 forecasting. *Ecological Informatics, 56*.

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems, 112*, 111-124.

Veganzones, D., Séverin, E., & Chlibi, S. (2023). Influence of earnings management on forecasting corporate failure. *International Journal of Forecasting, 39*(1), 123–143.

Vermeulen, W. J. V., & Witjes, S. (2016). On addressing the dual and embedded nature of business and sustainability. *Organization & Environment, 29*(1), 132–153.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Nerini, F. F. (2020). The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications,* 11(1), 1-10.

Walczak, S., & Cerpa, N. (2003). Artificial neural networks. In Encyclopedia of Physical Science and Technology (3rd ed.). *Cambridge: Academic Press*.

Wamba-Taguimdje, S.-L., Wamba, S. F., Kamdjoug, J. R. K., & Wanko, C. E. T. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal, 26*(7), 1893–1924.

Weber, E., Schmid, A., & Hein, A. (2022). Artificial Intelligence and the Changing Nature of Business Models. *Journal of Business Research*, 143, 161-174.

Wedge, R., Kanter, J., Moral, S., Iglesias Pérez, S., & Veeramachaneni, K. (2017). *Solving the "false positives" problem in fraud prediction*.

Wever, K. (1995). *Negotiating competitiveness: Employment relations and organizational innovation in Germany and the United States*. Harvard Business School Press.

White, R., & Green, K. (2021). Strategic Financial Management in Volatile Markets. *Business Strategy Review*, 29(2), 176-195.

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11(5), 545-557.

Woo, S. E., O'Boyle, E. H., & Spector, P. E. (2017). Best practices in developing, conducting, and evaluating inductive research. *Human Resource Management Review, 27*(2), 255-264.

Wruck, K. (1990). Financial distress, reorganization, and organizational efficiency. *Journal of Financial Economics*, *27*, 419–444.

Xie, Y., Le, L., Zhou, Y., & Raghavan, V. V. (2018). Chapter 10—Deep learning for natural language processing. In Handbook of Statistics (Vol. 38, pp. 317-328). Amsterdam: Elsevier.

Yang, Z. R., Platt, M. B., & Platt, H. D. (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research,* 44(2), 67-74.

Zhang, G., Hu, M., Patuwo, B. E., & Indro, D. C. (2016). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research, 252*(1), 280–293.

Zhang, J., Hu, W., & Susheng, W. (2021). The application of artificial intelligence technology in financial risk prevention and control. *Financial Innovation,* 7(1), 1-18.

Zhang, K., Xu, P., & Zhang, J. (2020). Explainable AI in deep reinforcement learning models: A shap method applied in power system emergency control. *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, 711–716.

Zhuhadar, L. P., & Lytras, M. D. (2023). The application of AutoML techniques in diabetes diagnosis: Current approaches, performance, and future directions. *Sustainability*, *15*(18).

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22.

Zou, J., Han, Y., & So, S. S. (2008). Overview of artificial neural networks. In *Artificial Neural Networks. Totowa: Humana Press*, Vol. 458, pp. 14-22.

STVDIVM VRBIS