Journal of Environmental Management 364 (2024) 121264

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

A GA-stacking ensemble approach for forecasting energy consumption in a smart household: A comparative study of ensemble methods

Mahziyar Dostmohammadi [a], Mona Zamani Pedram [a,**], Siamak Hoseinzadeh [b,*], Davide Astiaso Garcia [b]

[a] Energy System Division, Department of Mechanical Engineering, K.N, Toosi University, Tehran P.O.Box 19395-1999, Iran
[b] Department of Planning, Design, Technology of Architecture, Sapienza University of Rome, Via Flaminia 72, 00196, Rome, Italy

A R T I C L E I N F O

Keywords:
Building energy performance
Energy efficiency
Forecasting energy
Eco-friendly
Consumption pattern clustering
Machine learning
Ensemble technique

A B S T R A C T

The considerable amount of energy utilized by buildings has led to various environmental challenges that adversely impact human existence. Predicting buildings' energy usage is commonly acknowledged as encouraging energy efficiency and enabling well-informed decision-making, ultimately leading to decreased energy consumption. Implementing eco-friendly architectural designs is paramount in mitigating energy consumption, particularly in recently constructed structures. This study utilizes clustering analysis on the original dataset to capture complex consumption patterns over various periods. The analysis yields two distinct subsets that represent low and high consumption patterns and an additional subset that exclusively encompasses weekends, attributed to the specific behavior of occupants. Ensemble models have become increasingly popular due to advancements in machine learning techniques. This research utilizes three discrete algorithms, namely Artificial Neural Network (ANN), K-nearest neighbors (KNN), and Decision Trees (DT). In addition, the application employs three more machine learning algorithms bagging and boosting: Random Forest (RF), Extreme Gradient Boosting (XGB), and Gradient Boosting Trees (GBT). To augment the accuracy of predictions, a stacking ensemble methodology is employed, wherein the forecasts generated by many algorithms are combined. Given the obtained outcomes, a thorough examination is undertaken, encompassing the techniques of stacking, bagging, and boosting, to conduct a comprehensive comparative study. It is pertinent to highlight that the stacking technique consistently exhibits superior performance relative to alternative ensemble methodologies across a spectrum of heterogeneous datasets. Furthermore, using a genetic algorithm enables the optimization of the combination of base learners, resulting in a notable enhancement in prediction accuracy. After implementing this optimization technique, GA-Stacking demonstrated remarkable performance in Mean Absolute Percentage Error (MAPE) scores. The improvement observed was substantial, surpassing 90 percent for all datasets. In addition, in subset-1, subset-2, and subset-3, the achieved R^2 scores were 0.983, 0.985, and 0.999, respectively. This represents a substantial advancement in forecasting the energy consumption of residential buildings. Such progress underscores the potential advantages of integrating this framework into the practices of building designers, thereby fostering informed decision-making, design management, and optimization prior to construction.

1. Introduction

The residential sector occupies a substantial portion of global energy consumption, accounting for 40% of the total energy utilized (Kang et al., 2022). As the depletion of fossil fuel reserves escalates and concerns regarding environmental degradation amplify (Hamidinasab et al., 2023), it becomes imperative to employ accurate energy forecasting methods and practical management approaches to reduce consumption, improve overall efficiency, and promote the adoption of renewable energy resources (Ciupageanu et al., 2020). Accurate energy consumption forecasting is paramount, as it enables thorough evaluation of diverse building design alternatives, the optimal configuration of energy systems, and successful integration of renewable energies into supply and demand management frameworks (Zhang et al., 2020). Furthermore, the emergence of the COVID-19 pandemic has highlighted the

* Corresponding author.
** Corresponding author.
E-mail addresses: m.zpedram@kntu.ac.ir (M.Z. Pedram), siamak.hosseinzadeh@uniroma1.it, hoseinzadeh.siamak@gmail.com (S. Hoseinzadeh).

https://doi.org/10.1016/j.jenvman.2024.121264
Received 9 December 2023; Received in revised form 21 May 2024; Accepted 26 May 2024

Nomenclature

Artificial Neural Network	ANN
Balance Point Temperature	BPT
Coefficient of Determination	R^2
Decision Trees	DT
Deep Forest	DF
Deep Neural Networks	DNN
Deep Recurrent Neural Networks	DRNN
Ensemble Learning	EL
Extreme Gradient Boosting	XGB
Gaussian Process Regression	GPR
Genetic Algorithm	GA
Genetic Programming	GP
Gradient Boosting Machines	GBMs
Gradient Boosting Trees	GBT
Heating, Ventilation, and Air Conditioning	HVAC
K-nearest neighbors	KNN
Least-Squared Boosted	LSB
Linear Regression	LR
Machine learning	ML
Mean Absolute Percentage Error	MAPE
Mean Squared Error	MSE
Multi-linear Regression	MLR
Multiple Regression	MR
Particle Swarm Optimization	PSO
Random Forest	RF
Recursive Feature Elimination	RFE
Seasonal-Trend decomposition using LOESS	STL
Solar Radiation	SR
Support Vector Regression	SVR
Wind Speed	WS

importance of forecasting household energy consumption due to the significant increase in residential energy usage resulting from the widespread implementation of remote work practices (Zapata-Webborn et al., 2023). Accurate prediction of future outcomes is paramount in enhancing energy efficiency and maximizing architectural design and construction performance. Utilizing this technology empowers individuals in positions of authority to make well-informed decisions, thereby improving the overall performance of the building. Predictive models used to evaluate energy demands, optimize resource allocation, and pinpoint opportunities for improvement (Kumar et al., 2022). In addition, they facilitate anticipating periods of maximum demand and evaluating the feasibility and effectiveness of energy-conserving technologies. The accuracy of these prognostications is crucial for enhancing financial planning and assessing the effectiveness of energy conservation initiatives. Detailed forecasts ultimately aid the construction industry in environmentally friendly and cost-efficient buildings (Li et al., 2021b). In the determination of energy consumption in buildings, three main approaches employed: physical models (white box models), data-driven models (black box models), and gray box models that combine elements of both (Chen et al., 2022). White box models are highly effective during the design phase as they necessitate reliable data about a structure's physical attributes and environmental features. They offer high precision but may need more computationally efficient and require expert knowledge (Bourdeau et al., 2019). On the other hand, gray box models provide higher accuracy by using simplified physical models, but their performance limited due to the quality of available data (Li et al., 2021a). In contrast, black box models, solely reliant on mathematical models and measurements, offer simplicity, rapid development, and the ability to handle complex variable interactions. They particularly favored for forecasting building energy usage as they can identify patterns and correlations through training on extensive datasets (Bourdeau et al., 2019). The illustrated benefits and drawbacks of these approaches are outlined in Fig. 1

Data-driven methods utilize machine learning (ML) algorithms such as Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), Extreme Gradient Boosting (XGB), and Gradient Boosting Machines (GBMs) to extract valuable insights from historical energy data. These algorithms enable researchers to identify patterns and characteristics within the data, leading to significant advancements in energy consumption forecasting. Ensemble techniques have gained substantial traction in energy consumption prediction owing to their ability to improve reliability and alleviate the constraints of individual models. Ensemble methods, such as bagging, boosting, and stacking, harness the strengths of multiple models to improve overall performance. Bagging techniques create an ensemble by training independent models on different subsets of the data and aggregating their predictions. This approach reduces overfitting and improves generalization. Boosting techniques, such as GBMs, iteratively train weak models by emphasizing previously mispredicted data points, effectively focusing on challenging samples and enhancing accuracy (Ganaie et al., 2022). Stacking methods entail the aggregation of diverse base models through a meta-model capable of weighting and integrating individual model predictions, thereby yielding a more refined ensemble (Yao et al., 2022). Incorporating ensemble techniques in data-driven forecasting facilitates the proficient handling of limitations inherent in fundamental models, thereby leading to enhanced accuracy in energy consumption prognostications.

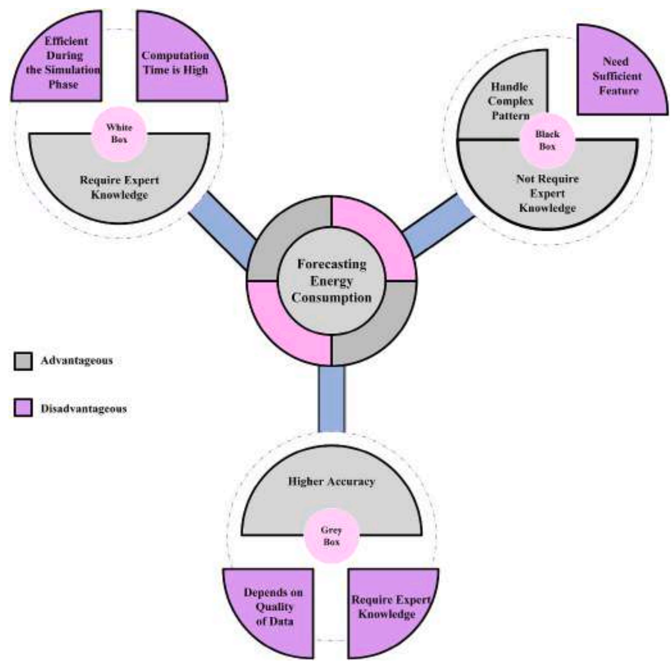

Fig. 1. Exploring forecasting approaches for energy consumption: Unveiling the advantages and disadvantages.

2. Literature review

The present study aims to achieve two key objectives. The primary objective involves employing clustering analysis to identify distinct consumption patterns within the initial dataset, offering a novel approach to pattern recognition that is independent of pre-existing knowledge. Concurrently, the research employs the Seasonal-Trend decomposition using LOESS (STL) method to extract three distinct temporal aspects or time series features. Incorporating temporal features into the examination of sequential datasets has resulted in significant advantages, resulting in a notable improvement in the overall efficacy of the prediction model. Furthermore, the research introduces a stacking method that combines three individual algorithms: DT, KNN, and ANN. The primary aim of this approach is to enhance the precision of energy consumption predictions. It is essential to emphasize that the research utilizes three additional ensemble methods, specifically bagging (RF) and boosting (XGB & GBT), as foundational models to further enhance forecast accuracy.

2.1. Decision tree (DT)

DT are a widely used as a fundamental ML method to regression tasks. Although initially linked to the prediction of continuous values, their versatility and performance in regression make them adaptable instruments in data analysis. DT is a hierarchical framework that effectively divides the feature space into distinct segments, progressing from the root to the leaf nodes. The partitioning process in regression tasks is guided by attribute values, tailored specifically for this purpose. The primary aim of this approach is to predict and simulate continuous outcomes, rather than categorizing individual samples. Regression DT retain their well-known benefits, including the capacity to analyze the model and make efficient predictions. DT are an essential asset in regression within ML due to their ability to produce straightforward division rules through recursive tree construction. This characteristic makes DT self-explanatory and enhances their usefulness. (Olu-Ajayi et al., 2022) explored various ML methods, including DT and other relevant techniques, to predict energy consumption patterns in residential buildings. (Zhou et al., 2021) utilized DT as a predictive model for short-term building occupancy, resulting in a notable enhancement in feature selection. The Deep Forest (DF) framework incorporated DT in conjunction with ACF, PACF, and MOTP, demonstrating superior performance compared to conventional models. This highlights the crucial importance of DT in improving the accuracy of predictions to optimize building occupancy. The study conducted by (Mienye et al., 2019) focused on applying DT algorithms, specifically ID3 and C4.5, in tackling prediction difficulties in various industries. The critical emphasis lies in evaluating the predictive capabilities of different decision tree algorithms. Additionally, it evaluates endeavors aimed at improving their efficacy. The primary objective of this investigation is to offer academics valuable insights regarding the effectiveness of DTs and their potential impact on predictive modeling in many fields.

2.2. K-nearestneighors (KNN)

The KNN algorithm is a flexible ML technique commonly employed for regression. The method utilizes proximity-based inference by selecting the K nearest data points in the feature space, typically quantified using Euclidean distance. KNN algorithm demonstrates exceptional performance in forecasting continuous outcomes. It is highly regarded for its transparency and interpretability within ML, specifically in the context of regression analysis. (Jiao et al., 2023) mainly focused on utilizing standard algorithms like KNN to enhance the accuracy of energy consumption predictions for Household energy management and demand-response systems. (Valgaev et al., 2017) considers the KNN' approach for short-term load forecasting in building uses. It automatically adjusts its parameters and accurately predicts future load based on past measurements from the building. The primary aim of (Yang et al., 2023) was to improve the accuracy of energy consumption forecasts, focusing specifically on utilizing the KNN model. The text introduces the notion of the Balance Point Temperature (BPT) label. It emphasizes the noteworthy performance of KNN, particularly in scenarios when there is a limited amount of data available. The statement as mentioned above highlights the strong efficacy of the KNN model in enhancing energy management systems in an academic setting. (Hong et al., 2022) utilized the KNN algorithm to forecast hourly energy consumption in a heterogeneous building community. The findings reveal that the predictions are accurate throughout the summer and fall seasons, while exhibiting a modest tendency to overestimate during the spring and winter periods.

2.3. Artificial Neural Network (ANN)

As evidenced by (Nabavi-Pelesaraei et al., 2023), ANNs have emerged as powerful tools in various prediction tasks across different fields, including energy consumption and environmental impact assessment. (Al-Mufti et al., 2023) have successfully utilized ANNs to forecast building energy consumption, as demonstrated in a study where a digital twin incorporating smart sensors at the University of Sharjah accurately predicted energy consumption 15 min ahead, yielding results comparable to experimental data. (Nainwal and Sharma, 2022) demonstrated that ANN methods outperform traditional multi-linear regression (MLR) techniques in accurately predicting monthly energy consumption, as evidenced by comparative studies conducted by various authors. (Ahmad et al., 2017) conducted a comparative analysis to assess the efficacy of two distinct algorithms, namely the feed-forward back-propagation ANN and the RF algorithm, in predicting the hourly HVAC energy consumption of a hotel situated in Madrid, Spain. (Amber et al., 2018) evaluated and conducted the efficacy of five intelligent system methods for forecasting the electricity usage of an administrative office in London, United Kingdom. The study assessed various methodologies, namely ANNs, Multiple Regression (MR), Genetic Programming (GP), and Deep Neural Networks (DNN). The researchers constructed the prediction models using a dataset spanning four years, which consisted of five features: Solar Radiation (SR), temperature, Wind Speed (WS), humidity, and weekday index. The findings revealed that the ANN outperformed the other techniques, demonstrating superior predictive capabilities for electricity consumption forecasting.

2.4. Ensemble models

Ensemble methods are commonly used in ML as a strategic approach that combines multiple predictive models to enhance prediction accuracy. The core methodology entails integrating the results of the different models to attain enhanced forecast precision. Ensemble approaches demonstrate efficacy in addressing intricate data interactions and possess the capacity to enhance model performance across a wide range of applications substantially.

2.4.1. Bootstrap Aggregating

In the domain of ensemble methods, the primary strategy employed is Bagging. The Bagging framework includes various algorithms, with RF being particularly notable for its utilization of multiple DT. Furthermore, (Wang et al., 2018) applied the RF algorithm for accurate hourly building energy forecasting. A comparative analysis of RF models with varying parameter configurations against alternative algorithms has been conducted in this study. The results demonstrated the robustness of RF models about the number of features. (Hadri et al., 2019) utilized an IOT and Big Data platform to collect real-time electricity and load consumption data. The focus was on developing predictive models using the RF algorithm and evaluating their accuracy in load forecasting. (P. et al., 2023) used RF to forecast upcoming electric power consumption and renewable energy generation, and their effectiveness in accurately predicting these factors has been demonstrated.

2.4.2. Boosting

Subsequently, the objective of boosting approaches in ML is to iteratively enhance the accuracy of a model by giving higher weights to examples that are mispredicted in each iteration, ultimately leading to the development of a final model that is both robust and precise. GBMs are a prominent illustration of boosting algorithms, and (Morteza et al., 2023) assessed the accuracy of energy usage forecasting. The Deep Recurrent Neural Networks (DRNN) model was selected as the final model for comparison with the GB technique. The study aimed to assess and contrast the efficacy of the two models in forecasting energy consumption patterns (Touzani et al., 2018).Utilized the powerful GBM algorithm to enhance energy consumption forecasts in a sample of 410 commercial buildings. The research conducted by the authors demonstrated superior performance compared to conventional approaches, highlighting the significant contribution of GBM in improving the precision of energy forecasting.

2.4.3. Stacking

Stacking is an ensemble technique employed in ML wherein the outputs of many models are combined. This approach frequently results in improved predictive accuracy, as it involves training a higher-level model to make final predictions by utilizing the outputs generated by a diverse set of base models. (Cao et al., 2023) combined eleven ML algorithms, including KNN, XGB, DT, RF, etc., for a stacked ensemble model and Applied Particle Swarm Optimization (PSO) to improve the performance of the model base on finding the optimal combination of based models and a meta-model within the Stacking framework. (Pachauri and Ahn, 2023) Introduced a novel ensemble predictive model, denoted as WGPRLSB, designed to provide precise estimations of energy consumption in both Heating and Cooling Load scenarios. The proposed combination of Least-Squared Boosted (LSB) and Gaussian Process Regression (GPR) through a weighted linear aggregation model. Marine predator optimization determines the design parameters' optimal values. Additionally, several traditional predictive models, such as support vector regression (SVR), Linear Regression (LR), DT, and generalized additive model, are developed for comparative analysis. (Qavidelfardi et al., 2022) considered meteorological conditions of Iran to predict residential electricity consumption. To achieve this, an Ensemble Learning (EL) framework is proposed. The framework incorporates fifty parameters: environmental, context and building, occupant, time-related, and additional inputs. This study aims to develop a practical approach that combines the strengths of multiple models within the EL framework t to forecast residential energy usage accurately (Lu et al., 2023) and several ML techniques, such as SVR, KNN, and ensemble inference models. These models were built to forecast cooling and heating load. The result indicates that the ensemble approach, which set SVR and KNN as a base model and set SVR as meta learner, demonstrated the highest accuracy for prediction purposes.

An extensive examination of the literature reveals a wide array of algorithms suitable for the specific task at hand in this research. Each algorithm presents its own set of advantages and disadvantages, contingent upon the characteristics inherent in various datasets. Consequently, careful consideration must be given when selecting among these algorithms to suit the unique requirements of different datasets.

However, the central focus of this study revolves around the implementation of stacking, an effective ensemble learning technique. Stacking allows for the integration of diverse algorithms into a unified framework, leveraging the strengths of each while mitigating their individual weaknesses. Notably, stacking eliminates the necessity for prior knowledge regarding which algorithm is best suited to a given dataset. Instead, through an automated process, stacking dynamically identifies the most appropriate algorithms from a predefined pool for each dataset encountered. By embracing stacking as the cornerstone of our methodology, this study endeavors to harness the collective capabilities of diverse algorithms without the constraints imposed by algorithmic selection biases, this approach not only enhances the resilience and adaptability of prediction.

This research paper presents a novel conceptual framework for managing energy in residential settings, specifically aiming to improve the precision of energy consumption predictions. The material in question differentiates itself from preexisting materials through a variety of notable means, placing particular emphasis on the subsequent advantages.

1. The dataset is partitioned into two subsets, one designated for weekdays and the other for weekends. The k-means technique is subsequently employed to cluster the patterns within the subset of weekdays, identifying two unique demand trends that correspond to periods of high and low demand.
2. The study utilized a novel hybrid feature selection, integrating Recursive Feature Elimination (RFE) with RF. The present methodology employed a systematic strategy to identify the most significant attributes through iterative evaluation of feature importance.
3. The primary objective of this work is to perform a comprehensive comparison analysis to assess the efficacy of different ensemble methodologies. This study aims to determine the optimal ensemble technique for the topic matter through a thorough investigation.
4. The present study introduces a Genetic Algorithm (GA) as a means of optimization to construct a hybrid model that enhances the performance of the ensemble model. This enhancement is achieved by optimizing the combination of base estimators.

3. Methodology

3.1. Research outline

The present investigation employs a dataset obtained from automated sensors installed in one house that captured data at 10-min intervals (Candanedo et al., 2017). The dataset consists of meteorological characteristics and historical energy usage records, as shown in Table 1. Initially, after the data cleansing process, the need for scaling arises due to the non-uniform distribution of the dataset. Subsequently, the K-mean clustering algorithm was utilized to cluster and extract essential demand

Table 1
Description of target and features in the dataset.

Features	Description	Units	Min	Max
Data Entries	Total Number of Data Points	19,735		
TEC	Total Energy Consumption	Wh	10	1100
T1	The temperature in the kitchen area	c	16.78	26.26
R1	Humidity in the kitchen area	%	27.02	63.35
T2	The temperature in the living room area	c	16.10	29.85
R2	Humidity in the living room area	%	20.46	56.02
T3	The temperature in the laundry room area	c	17.19	29.23
R3	Humidity in the laundry room area	%	28.76	50.16
T4	The temperature in the office room	c	15.10	26.19
R4	Humidity in office room	%	27.66	51.09
T5	Temperature in bathroom	c	15.33	25.795
R5	Humidity in bathroom	%	27.66	96.32
T6	The temperature outside the building (north side)	c	−6.06	28.28
R6	Humidity outside the building (north side)	%	1.00	99
T7	The temperature in the ironing room	c	15.39	26
R7	Humidity in the ironing room	%	23.19	51.39
T8	The temperature in teenager room 2	c	16.30	27.23
R8	Humidity in teenager room 2	%	29.60	58.78
T9	The temperature in the parents' room	c	14.89	24.5
R9	Humidity in parents' room	%	29.16	53.32
T_out	Temperature outside	c	−5.00	26.1
R_out	Humidity outside	%	24.00	100
P	Pressure	mm Hg	729.29	772.29
Ws	Windspeed	m/s	0.00	14
Vb	Visibility	Km	1.00	66
T_dp	Dewpoint	c	−6.59	15.5

time frames from the original dataset to predict diverse demand periods. Afterward, due to the features and complexity of the problem, this investigation employs a feature selection technique in the continuous domain to eliminate non-contributing features from the models' performance, thus simplifying the problem. The succeeding procedure entails the utilization of feature extraction methodologies, notably, the STL, as well as the extraction of lag features from the dataset. Various strategies are utilized to identify and analyze unique patterns within the dataset, improving the precision of predictions. Later on, the pre-processed data will be partitioned into three distinct subsets, specifically the training, testing, and validation sets. Two of the initials will be employed to evaluate the effectiveness of individual algorithms and ensemble models. The third data set will be utilized for hyperparameter optimization to determine meters for each model. The objective of the optimization procedure is to augment the exactness, reliability, and resilience of the models. The achievement of optimal combinations of base estimators in stacked models will be facilitated through a GA. The proposed methodology prioritizes the utilization of base algorithms that exhibit exceptional efficacy on the given dataset during the training of the meta-learner. As a result, the precision of the meta-learner is expected to be enhanced. The framework employed in this study will be illustrated in Fig. 2.

3.2. Data description

The present investigation employed a dataset about the energy consumption of a passive house in Stambruges, Belgium, roughly 24 km from the City of Mons (Candanedo et al., 2017). The residential structure's finalization occurred in December 2015, and the energy consumption monitoring was executed through M-BUS energy meters, which collected data at 10-min intervals. Measuring individual electrical loads involved various components, such as the heat recovery ventilation unit and the energy usage of appliances, lighting, and electric baseboard heaters. The dataset covers a temporal duration of 137 days, equivalent to 4.5 months. The primary emphasis of this analysis pertains to the recorded data of total energy consumption (Wh) at 10-min

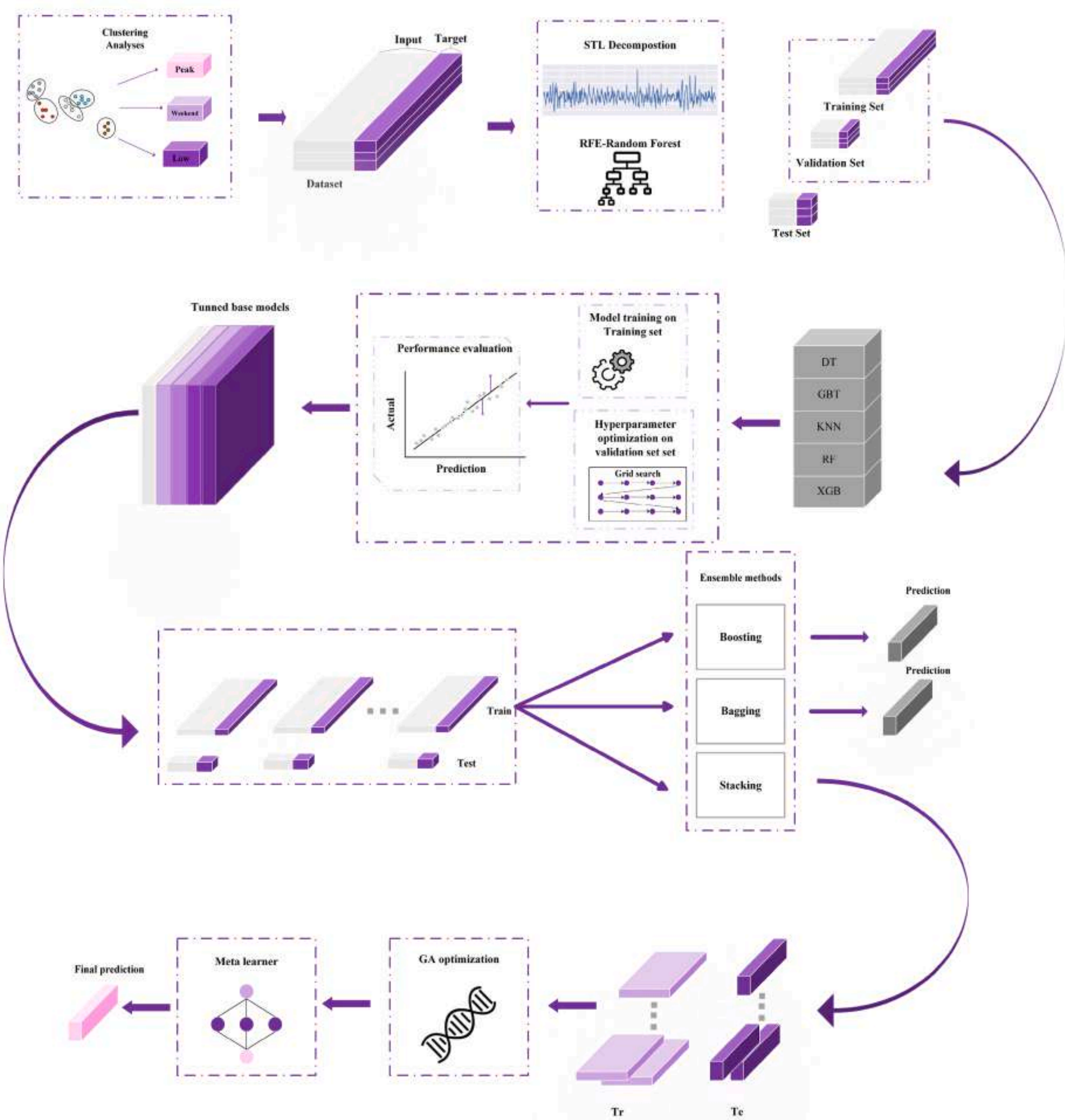

Fig. 2. Proposed framework for ML process design.

intervals. The 10-min reporting interval was chosen to ensure efficient capture of any rapid fluctuations in the target variable (energy consumption).

3.3. Data preprocessing

To improve the accuracy and reliability of the models utilized in this study, preprocessing techniques were applied to the dataset. Scaling techniques were also employed to maintain consistency among the input features (Nabavi-Pelesaraei et al., 2021). Notably, the scaling process was not applied to the target variable, the predicted variable in ML. Applying a scaler to the target variable would result in a modification of its intrinsic statistical significance. After implementing various scaling techniques, such as min-max and standard scaler (Ahsan et al., 2021), a comprehensive analysis was conducted to determine the optimal approach for the given dataset. It was concluded that standard scaling produced the most favorable outcomes.

Subsequently, box plots were used to detect any possible outliers in the input and output features (Krzywinski and Altman, 2014). Despite the potential benefits of excluding or managing outliers in the target variable to improve model accuracy, a decision has been made to maintain their presence in the dataset due to a convincing explanation. As mentioned earlier, the outliers represent the highest degree of energy utilization, a pivotal factor that influences the overall energy expenditure of the residents. Including these outliers in the dataset promises to enable the development of a model that exhibits improved reliability and accuracy and effectively captures the household's consumption behaviors. The elevated values of the target variable during peak demand periods have been recognized as a notable factor. To rectify this concern, a clustering algorithm will be employed to extract a new dataset, facilitating the modeling of distinct consumption behaviors individually.

3.4. Data splitting

Data splitting is a crucial aspect of ML as it plays a pivotal role in assessing the generalization capabilities of a model (Nguyen et al., 2021). The dataset is partitioned into distinct subsets, namely the largest subset known as the training set, which is utilized for pattern learning, and another subset used for model validation and hyperparameter tuning. Utilizing a distinct testing set allows for the evaluation of the model's capacity to accurately forecast novel data in an autonomous manner, separate from the training and validation processes. In this study, 20% of the data is allocated for hyperparameter tuning, while the remaining portion is separated into a 70% training set and a 30% test set. The selection of data splitting ratios and methodologies is of utmost importance, considering the dataset's specific properties and the modeling process's aims. In contrast to non-timeseries data, timeseries data should preserve its inherent sequence when splitting (Macas et al., 2016). This feature guarantees that the model can capture and effectively utilize temporal dependencies, trends, and distinctive patterns. This methodology follows the data's fundamental structure, hence improving the effectiveness and resilience of ML models when applied to timeseries datasets (Jin et al., 2021).

3.5. Clustering analyzes

Using the k-means clustering method to divide datasets that exhibit non-normally distributed target variables shown in Fig. 3 is an essential analytical technique in a wide range of scientific and business fields (Nepal and Sahashi, 2019). In situations where the distribution of the target variable deviates from a typical Gaussian distribution, the conventional practice of categorizing data into 'high' and 'low' groups may not be suitable (Metzler, 2020).

K-means clustering is an unsupervised learning technique that is driven by data. It is widely recognized as a helpful tool for identifying intrinsic groupings within a dataset without relying on strong assumptions about the distribution of the underlying data. The K-means algorithm seeks to optimize the within-cluster sum of squares by repeatedly assigning data points to clusters indicated by their centroids (Cui et al., 2023). This process effectively groups observations that are comparable to each other. The selection of the ideal number of clusters, denoted as K, equal to two in this work, is a crucial stage in the analysis.

As depicted in Fig. 3, the initial dataset has been divided into two distinct subsets, namely peak and low demand. Additionally, a third subset representing weekends was extracted from the original dataset, representing occupant behavior and its impact on the target variable. Henceforth, the three models mentioned above shall be referred to as

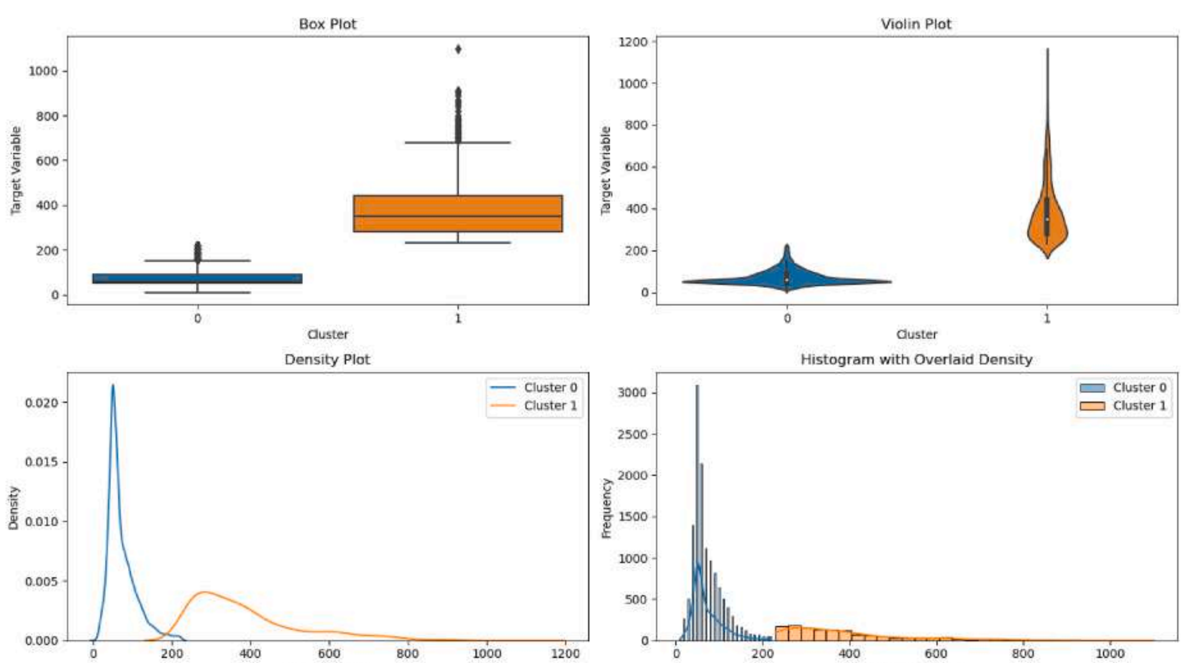

Fig. 3. Clustering analysis results for peak and low energy consumption patterns in buildings.

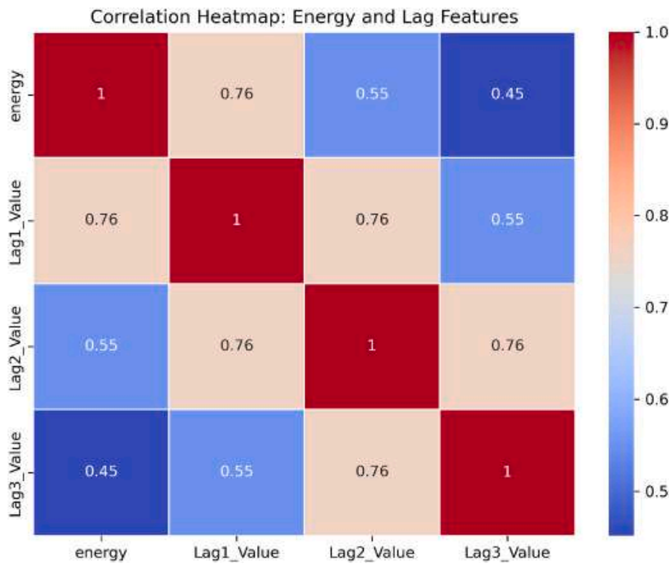

Fig. 4. A comparative analysis of the correlation between target variable and lag features.

subset-1, denoting the high demand set; subset-2, representing the low demand data set; and subset-3, signifying the weekend data set.

3.6. Feature engineering

3.6.1. Leveraging lag features

Incorporating lag characteristics, which capture past 'energy' values, enhances our analysis by introducing a temporal aspect. These characteristics enable the examination of temporal patterns, seasonal variations, and autocorrelations in the data. As shown in Fig. 4, the correlation between the lag feature and the target variable underscores these relationships. Within the domain of time series analysis, the temporal context holds significant importance in ensuring precise forecasting and uncovering latent patterns in energy usage that are peculiar to our research.

3.6.2. STL decomposition

Extracting significant characteristics from data is crucial in selecting important features to develop reliable models to predict the target variable. Time series features are noteworthy features that possess considerable value in the present context, wherein a robust association exists between temporal features and target variable patterns (Zhang et al., 2023). This is due to their ability to capture patterns that are temporally linked. Incorporating temporal features into prognostic models facilitates the identification of fluctuations in target variables over different time frames, such as hourly, daily, or monthly intervals (Sun et al., 2020). Within the domain of time series analysis, the Seasonal-Trend decomposition utilizing the LOESS (STL) method is widely utilized as a fundamental approach for unraveling complex patterns within temporal data (Phu, 2021). STL is a method that separates the data into three distinct components, which are demonstrated in Fig. 6. These components include the Seasonal Component, which represents recurrent patterns within the data; the Trend Component, which exposes long-term underlying trends; and the Residual Component, which captures any unexplained noise or abnormalities in the data. Valuable qualities are derived from these components (de Rautlin de la Roy et al., 2023). Seasonal characteristics comprise both the amplitude and phase of seasonal patterns, providing valuable insights into the intensity and timing of these patterns. Trend aspects encompass data about to the inclination and consistency of the fundamental trend. Residual features offer an evaluation of the fluctuation of data and the presence of measurement inaccuracies. This study involves the analysis of seasonal patterns, detecting anomalies, and identifying periodicity within different time intervals (1 h, 5 h, and 4 h) in three different models (subset-1, subset-2, and subset-3).

Additionally, the trend component, as represented by STL compositions, is used to showcase the outcomes of this decomposition, as depicted in Fig. 5. The STL decomposition technique and its associated feature extraction methods are significant in understanding temporal data structures. They enable various tasks such as forecasting, anomaly detection, and trend analysis in various academic and practical fields.

3.6.3. Hybrid feature selection: RFE and Random Forest

The process of feature selection using the RFE-RF hybrid method involves a carefully planned sequence that combines the systematic approach of RFE with the advanced modeling capabilities of RF (Jeon and Oh, 2020). The process of RFE is initiated by systematically removing attributes of lesser significance using an iterative pruning mechanism. This mechanism is based on the importance scores of these attributes, which are typically determined from the performance of a selected ML model (Yin et al., 2023). Concurrently, the RFE technique produces a prioritized list of features, offering a valuable understanding of their importance. After reducing the dimensionality of the dataset, RFE proceeds to transfer the remaining feature subset to RF, an ensemble of DTs renowned for its capability to capture complex data relationships (Niquini et al., 2023). The feature ranking process in RF involves refining criteria such as Gini impurity or information gain. This refinement takes into account the relevance of individual features as well as their complex relationships. Utilizing a collaborative technique guarantees that the ultimate feature subset consists of qualities that exhibit both individual importance and relevance within intricate data contexts.

The versatility and adaptability of RFE-RF enable its application to a wide range of ML tasks, hence serving as a helpful tool for improving the performance and interpretability of models in numerous domains. In brief, the RFE-RF hybrid approach facilitates a systematic and cooperative procedure for selecting features, leading to a compact yet effective subset of features that enhances the performance of ML tasks (Olu-Ajayi et al., 2023).

The selected feature in Fig. 7 indicates that many relative humidity features have been identified through the feature selection process. One possible explanation for this phenomenon could be that the significance of humidity features rests in their influence on Heating, Ventilation, and Air Conditioning (HVAC) systems. Elevated humidity levels necessitate increased exertion from HVAC systems, resulting in heightened energy usage. The correlation mentioned above highlights the significance of humidity in predicting energy use. Furthermore, it is worth noting that the seasonal humidity patterns frequently coincide with regional climate fluctuations. In numerous geographical areas, the level of humidity experiences an increase throughout the summer season. Energy usage exhibits seasonal patterns, reaching its highest levels during extreme weather. The synchronization observed in seasonal trends establishes a robust correlation between humidity levels and energy use.

3.7. Model selection

Within the field of ML, a range of approaches are utilized to develop predictive models. Ensemble models have gained significant popularity as a commonly utilized technique that has the potential to improve accuracy and robustness in predicting tasks. Ensemble approaches accomplish this objective by amalgamating the predictions generated by many base models, capitalizing on the heterogeneity of these models to collectively produce forecasts that are more resilient and precise.

3.7.1. Stacking

The core focus of our methodology is on the concept of stacking, which is employed to enhance the accuracy of target variable projections (Wang et al., 2023). This is accomplished by amalgamating

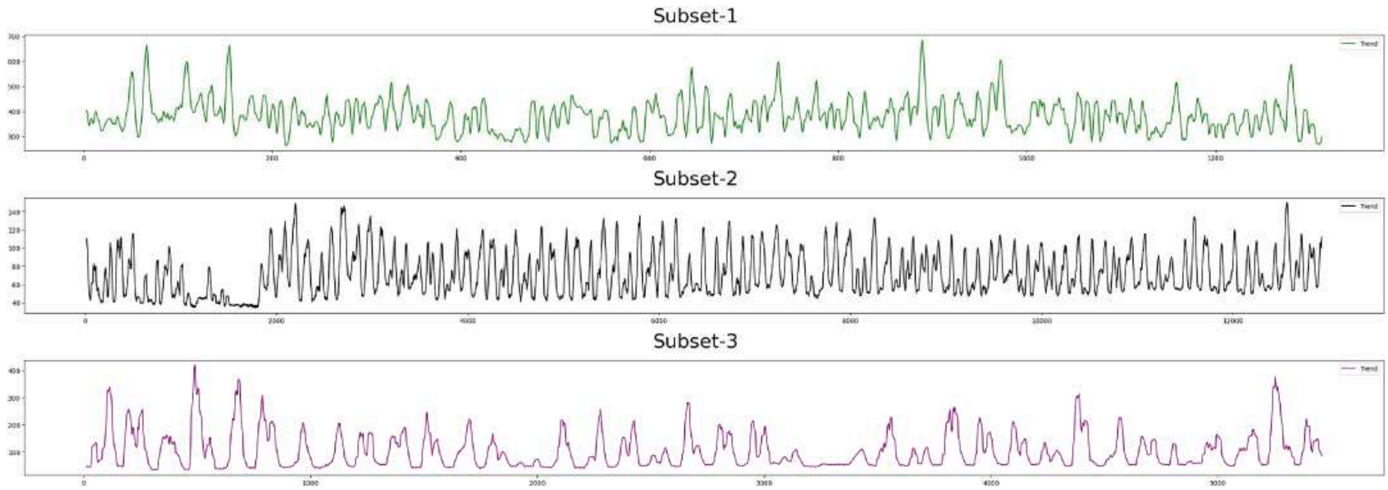

Fig. 5. Time series decomposition (STL): Trend composition.

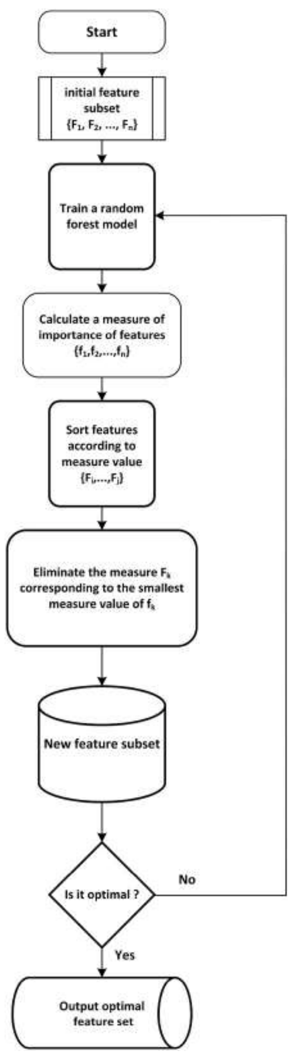

Fig. 6. The Sequential Process of hybrid-feature selection method.

results from a heterogeneous ensemble of foundational models, XGB, GBT, RF, DT, KNN, and ANN. Every individual base model utilizes unique algorithms and methodologies that have been trained independently using historical data on energy consumption. The stacking process comprises of two primary steps, as depicted in Fig. 8. Initially, the base models are trained autonomously using a comprehensive dataset, producing individual estimates for energy use. Following this, a meta-learner, namely a default ANN in our scenario, integrates these forecasts, hence enhancing the accuracy of the energy consumption forecast. Considerable focus is placed on the optimization of the stacking ensemble through the careful adjustment of parameters and the selection of relevant features (Huang et al., 2023).

The stacking technique presents numerous benefits, such as its ability to effectively capture intricate data relationships, improve the accuracy of predictions, and enable a comprehensive evaluation of model performance. Nevertheless, this approach is not devoid of constraints. One issue that arise in this context is the heightened level of complexity, which need careful consideration when selecting appropriate models (Pachauri and Ahn, 2023). To overcome the issue mentioned above, we utilize a GA optimization technique to determine the optimal combination of base models that can yield improved prediction accuracy.

3.7.2. Bootstrap aggregation

Bagging, often referred to as Bootstrap Aggregating, is a widely recognized ensemble approach renowned for improving models' resilience and raising their predictive accuracy. This is accomplished by mitigating variance through the process of resampling. The RF algorithm is at the forefront of bagging algorithms. The bagging technique is founded of bootstrapping, wherein several subsets, referred to as bags, are generated from the initial training data, as illustrated in Fig. 9. Every subset is created by randomly selecting data points with replacements to replicate the original distribution (Bühlmann and Yu, 2002). These bootstrapped subsets are then used to train a base model, often of the same method type. The result is a collection of models that have been trained individually (Mosavi et al., 2021).

Bagging is an ensemble technique that is widely recognized for its notable advantages. These advantages include the introduction of diversity through the creation of bootstrapped subsets of the training data. Each base model can capture distinct patterns and relationships (Jiang et al., 2011). Additionally, bagging reduces variance by combining predictions from multiple models, making a more resilient model less prone to overfitting (Wu et al., 2018). The RF algorithm demonstrates an exemplary illustration of the bagging technique, which enhances this notion by employing DT, commonly denoted as "tree learners," as fundamental models. The procedure entails the utilization of bootstrapping to create multiple subsets, constructing a decision tree for each subset, combining predictions through majority voting for classification tasks or averaging for regression tasks to derive the ultimate prediction,

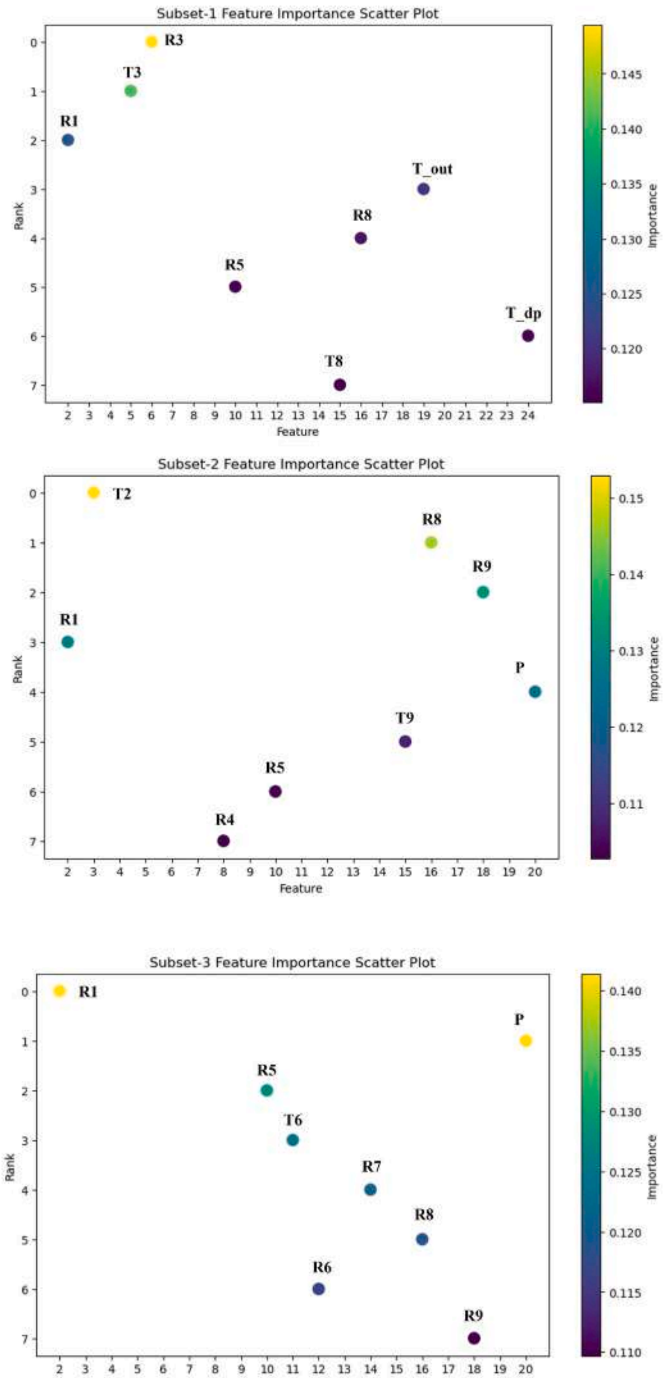

Fig. 7. Feature selection using RFE-RF: Importance ranking of selected features.

and incorporating feature randomization to augment diversity and alleviate overfitting (Feng et al., 2023). The RF algorithm provides several advantages, such as creating robust models, mitigating the risk of overfitting, and exhibiting versatility when applied to different types of data (Liu et al., 2021). Consequently, it is well-suited for performing classification and regression tasks in a wide range of domains.

3.7.3. Boosting

Boosting is a significant technique in the field of EL that plays a crucial role in improving model performance by employing a sequential training process (Mohammed and Kora, 2023). In contrast to parallel methodologies such as Bagging, the Boosting technique follows a sequential process in which a sequence of base models is constructed which is demonstrated in Fig. 11. (Wen et al., 2023). Each subsequent model in the sequence is designed to specifically target and rectify the faults made by its preceding models. The process of iterative refining leads to the creation of an ensemble that demonstrates exceptional proficiency in handling difficult circumstances and enhancing forecast accuracy (Nie et al., 2021).

The concept of boosting, as illustrated in Fig. 10, is based on attributing higher significance to instances that have been mispredicted, enabling succeeding models to prioritize these difficult cases. The mathematical representation of this concept is as follows: The initial training of base models, often consisting of shallow DT, aims to reduce regression errors on the dataset (Singh et al., 2023). Mispredicted instances are assigned higher weights, which increases their influence in following iterations. Successive base models are taught in a sequential manner, wherein the emphasis is placed on the mispredicted instances, and the errors made by previous models inform the learning process. The final ensemble prediction is generated by aggregating predictions from all base models, commonly by weighted majority voting or other suitable methodologies (Touzani et al., 2018). XGBoost, a notable example of the Boosting methodology, has garnered recognition due to its remarkable capability to improve the performance of models. XGB demonstrates exceptional performance in gradient-based optimization (Alshboul et al., 2022). It employs many strategies to effectively minimize loss functions and repeatedly enhance the model's accuracy.

3.8. Evaluation metric

This study employs various standard metrics, namely Mean Absolute Percentage Error (MAPE), mean square error (MSE), and coefficient of determination (R^2), to assess the precision and efficacy of the proposed models comprehensively (Figueiredo Filho et al., 2011; Chai and Draxler, 2014; Kim and Kim, 2016). The utilization of these metrics has been widely implemented in the assessment of predictive models. The error evaluation indices have been widely utilized to estimate forecasting models. The following three metrics are defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (A_i - P_i)^2}{\sum_{i=1}^{N} (A_i - \overline{A})^2} \tag{1}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (P_i - A_i)^2 \tag{2}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{A_i - P_i}{A_i} \right| \times 100\% \tag{3}$$

Evaluation metrics assess the adequacy of a regression model's fit by contrasting the predicted values (P) with the dependent variable's observed values (A). The parameter A denotes the actual values that are being predicted, whereas P denotes the values approximated by the regression model. The formula considers the dataset's complete set of observations (N). Furthermore, the mathematical average ($\overline{A}$) is computed by adding all the observed values and dividing the sum by the total number of observations.

3.9. Hyperparameter optimization

The process of optimizing hyperparameters is a crucial phase in the advancement of ML models. Hyperparameters refer to the parameters of a model that are not amenable to direct learning from the data and require specification before the training phase. Hyperparameters are exemplified by various factors such as the learning rate, regularization strength, and the number of estimators (Amir-Ahmadi et al., 2020). Determining the optimal values of hyperparameters depends on the

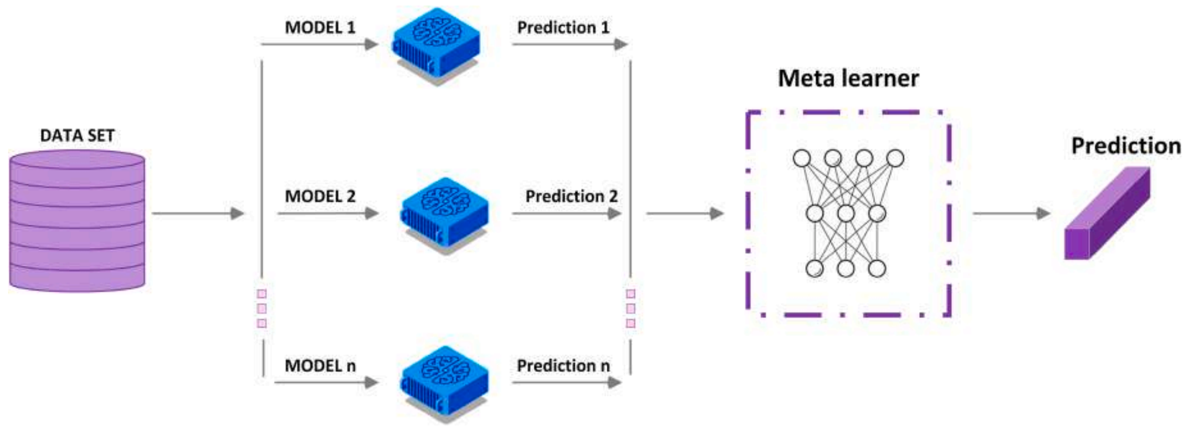

Fig. 8. Conceptual framework and structural overview of stacking models.

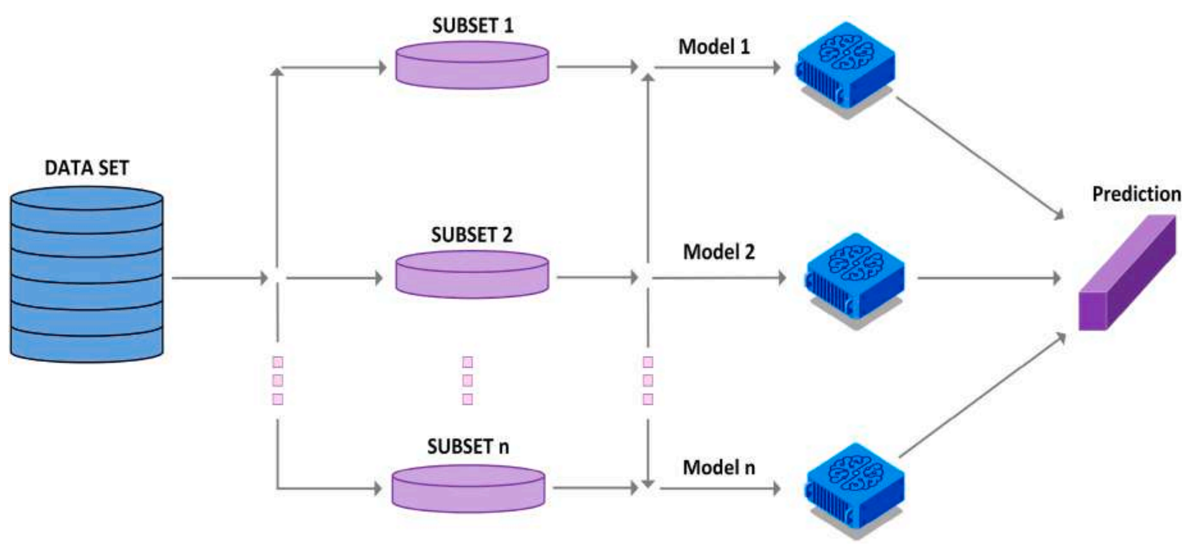

Fig. 9. Conceptual framework and structural overview of bagging models.

dataset and problem at hand, and discovering these parameters can have an enormous effect on the model's performance. The grid search technique is a widely used approach for optimizing hyperparameters, which involves specifying a grid of hyperparameter values to explore (Belete and Huchaiah, 2022). The model undergoes training and evaluation on the validation set for every hyperparameter combination within the grid. The evaluation metric determines the optimal set of hyperparameters, such as mean square error (Belete and Huchaiah, 2022). The present study employed grid search as a means of optimizing hyperparameters and the results of this is demonstrated in Table 2.

The hyperparameters of the models were explored through a predefined grid of values derived from the visualization of the effect of each hyperparameter on the metric evaluation. This approach was employed to determine the optimal range of each parameter for each algorithm. The model's performance was assessed; the optimal hyperparameters were chosen based on the mean squared error (MSE) criterion (Morteza et al., 2023). Grid search is a straightforward yet effective method for optimizing hyperparameters, which may significantly improve a model's performance. Nevertheless, the computational cost can be significant, particularly in cases where the hyperparameter space is extensive. Hence, it is crucial to meticulously choose the hyperparameters to explore and restrict the grid's magnitude to prevent overfitting and minimize computational expenses. Although the careful adjustment of base model hyperparameters continues to be a crucial task in stacking models, the comprehensive optimization of meta-learner hyperparameters may not always result in proportional enhancements (Palaniswamy and Venkatesan, 2021). After ensuring sufficient fine-tuning of underlying models, adopting default or simplified configurations for meta-learners represents a logical strategy. The careful arrangement of hyperparameter tuning within a nuanced perspective facilitates a balanced allocation of resources and optimization of predictive skills in ensemble modeling.

3.10. Genetic algorithm optimization

In the pursuit of enhancing the prediction accuracy of target variables within the residential construction sector, this study utilized GAs as an advanced optimization method to refine the combination of base learners (Shahhosseini et al., 2022). Building upon this premise, recent literature has highlighted the robustness of GAs in optimization. For example, a study on improving efficiency in Saffron farms showcased the efficacy of GAs (Saeidi et al., 2022), reaffirming their utility and adaptability across different domains and emphasizing their significance in advancing predictive modeling within residential construction. GAs, which draw inspiration from principles of evolution, provide a formal framework for this objective which is illustrated in Fig. 11. The utilization of binary strings to represent the combinations of basic learners enabled the investigation of a wide range of potential solutions. The fitness function utilized in this investigation was characterized by a unique emphasis, adopting the MSE as the key indicator for

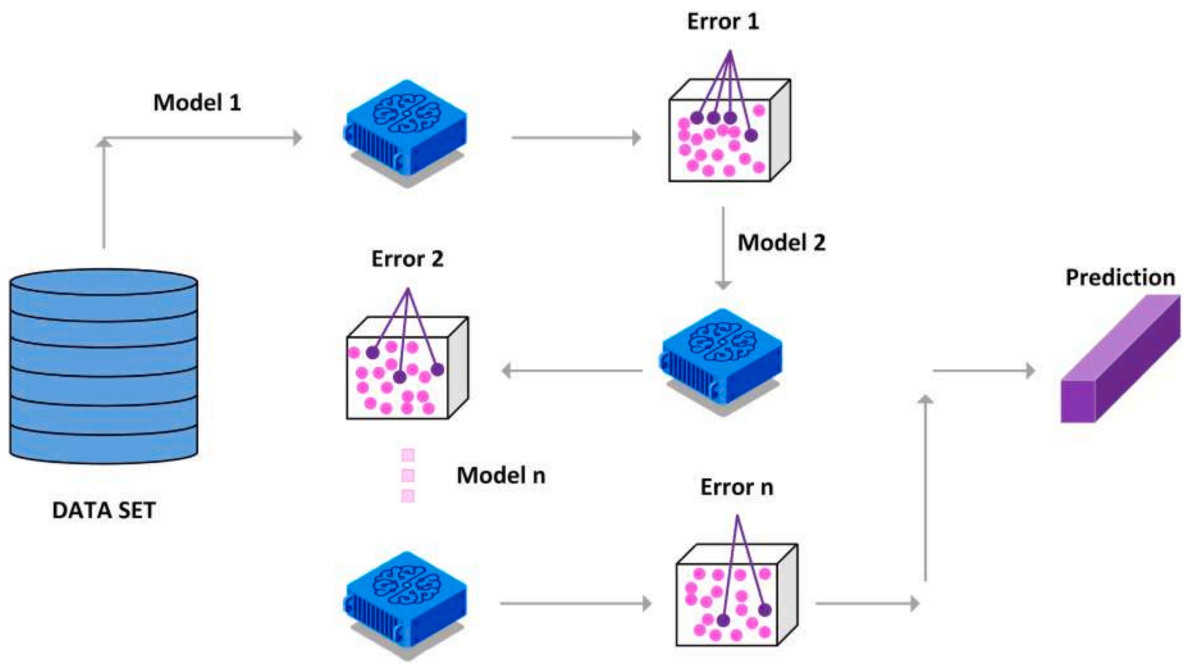

Fig. 10. Conceptual framework and structural overview of boosting models.

quantitatively evaluating the performance of each combination (Feng et al., 2021).

The methodology employed in this study involved generating an initial population including many potential solutions, each representing a unique combination of base learners. The population underwent a process of selection, crossover, and mutation, resulting in the evolution of subsequent generations. This evolutionary process favored combinations that exhibited enhanced prediction accuracy, as measured by the MSE. The convergence process denoted the culmination of the GA, unveiling the ensemble of base learners that encapsulated the most efficient amalgamation of several algorithms (Qu et al., 2021). Furthermore, a comprehensive overview of the essential parameters employed in this GA optimization procedure is presented, as outlined in Table 3.

This study presents a novel application of GAs that offers a promising approach for improving target variable predictions in residential buildings. By effectively integrating the collective intelligence of multiple ML algorithms, the predictive accuracy is enhanced, thereby making a valuable contribution to the field of energy-efficient residential building management.

4. Result and discussion

The results indicate that the model's performance depends on various factors. One of the most important factors is the dataset's characteristics, which directly influence our decision to model the dataset. The characteristics of the dataset have a significant impact on the performance of the algorithms. For instance, variations in the dataset can lead to different accuracy scores and affect our interpretation of the results (D'souza et al., 2020). Therefore, thoroughly understanding and analyzing the dataset is crucial for making informed decisions when modeling the data. The efficacy of single, bagging, and boosting algorithms varies across datasets due to their inherent limitations. Individually, each model's explanations for its results will be discussed. Stacking models consistently produce superior results despite the variable performance of single, bagging, and boosting algorithms across various datasets. Stacking models have combined six distinct algorithms as the base model, and ANN is set as the meta-learner for the models' final prediction. Then, ensemble models attain greater precision, robustness, and generalization across all models by stacking multiple models. Consequently, the efficacy of ensemble models will be discussed in a separate section.

4.1. *Prediction performance of single algorithms*

The KNN algorithm, known for its simplicity and versatility, exhibits different performance trends among the models examined. The results shown in Table 4 indicate that the KNN algorithm has a certain level of susceptibility to noisy data (Uddin et al., 2022). This observation is particularly clear in subset-1, which focuses on predicting peak energy demand. In the given setting, the KNN algorithm exhibits a somewhat elevated level of MSE and MAPE, suggesting its susceptibility to the impact of noisy or irrelevant information. This observation highlights a fundamental constraint in its relevance to situations characterized by variations in data.

Nevertheless, the study demonstrates a divergent pattern in subset-2, specifically about the forecasting of reduced energy consumption. Within the confines of this field, the implementation of the KNN algorithm effectively minimizes the occurrence of prediction inaccuracies, leading to a notable decrease in both MSE and MAPE. This implies that the KNN algorithm performs well in situations with prominent localized patterns in the data, but struggles to capture broader worldwide trends (Shapi et al., 2021). This is evident from the comparatively high MSE and MAPE observed in subset-3, which represents weekend energy usage. The constraint mentioned above is further intensified when dealing with datasets with many dimensions, as the computational complexity of the KNN algorithm becomes more prominent.

In sharp contrast, DT presents an alternate framework for predicting energy usage. The observations emphasize the consistent performance features displayed by the studied models. An important advantage of DT is their intrinsic capacity to choose features automatically, which is crucial for identifying relevant predictors and improving the accuracy of target variable estimates (Biresselioglu and Demir, 2022). Subset-2, which places significant emphasis on predicting situations with low consumption, provides as a strong example of this capability, as it exhibits minimal prediction errors and a correspondingly low MAPE. Moreover, DT demonstrate a remarkable ability to capture non-linear correlations in the data, making them a suitable option for situations

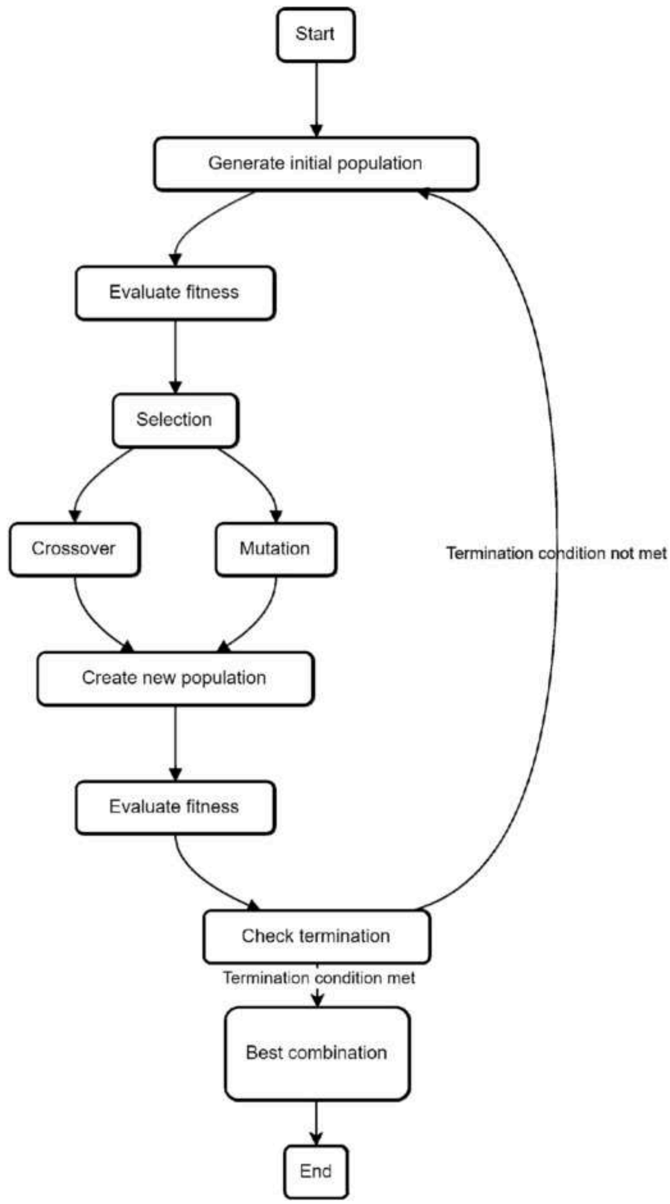

Fig. 11. Schematic representation of the GA optimization procedure.

when decision boundaries go beyond linearity (Czajkowski et al., 2023). The concept of interpretability emerges as a crucial factor, offering clarity on the factors that influence predictions. Additionally, DT include an inherent capability to mitigate overfitting, a prevalent issue in modeling initiatives, by skillfully employing pruning strategies and depth limitations (Mehdizadeh Khorrami et al., 2023). This fosters a nuanced equilibrium between predictive accuracy and the model's generalizability. The investigation concludes with a strong endorsement of ANNs as an excellent choice of algorithm, supported by its consistently outstanding performance across all models. The inherent complexity of ANN models enables them to effectively capture subtle correlations present in the data, resulting in consistently low MSE values (Truong et al., 2021). Subset-2 underscores the importance of accurate forecasts of reduced energy use and highlights the exceptional performance of the algorithm. The outstanding efficiency of this system is based on the algorithm's inherent ability to learn features automatically, thus reducing the requirement for costly manual feature engineering (Lu et al., 2022). Furthermore, the adaptive learning mechanism intrinsic to ANNs, assisted through the backpropagation process, allows for a continual adjustment of weights throughout training. This results in a

Table 2
Results of hyperparameter optimization: Optimal parameter Configurations for different algorithms.

Hyperparameters			
ANN	Subset-1	Subset-2	Subset-3
hidden_layer_sizes	(90, 39)	(94, 42)	(97, 40)
max_iter	954	1000	850
batch_size	17	20	14
n_iter_no_change	10	15	16
XGB			
n_estimators	450	500	485
max_depth	12	14	17
learning_rate	0.4	0.46	0.48
RF			
n_estimators	190	200	210
max_depth	50	55	48
KNN			
n_neighbors	3	5	7
p	1	2	2
GBT			
n_estimators	280	300	293
learning_rate	0.13	0.1	0.08
max_depth	10	15	11
DT			
max_depth	51	55	57

Table 3
GA parameters for base learner combination optimization.

Hyperparameter names	Value
Population size	6
Number of generations	10
Mutation Probability	0.2
Crossover Probability	0.2

gradual improvement in the accuracy of predictions (Runge and Zmeureanu, 2019).

4.2. Prediction performance of ensemble models

4.2.1. Bagging

The data pertaining to peak demand often exhibits substantial volatility and unpredictability, typified by abrupt surges and swift fluctuations, as depicted in Fig. 12. The presence of these inherent problems renders it a formidable testing ground for any forecasting model. In the present scenario, the RF algorithm exhibits outstanding performance, as evidenced in Table 5, by efficiently harnessing the collective potential of several DT (Wang et al., 2018). The system skillfully manages the intricacies associated with peak demand, successfully capturing the non-linear patterns and deep linkages inherent in the data. The marginally elevated MSE seen in this model does not indicate subpar performance. Rather, it underscores the algorithm's capacity to excel notwithstanding fluctuations in the data, leading to resilient predictions.

When focusing on predicting scenarios of low consumption, we are confronted with distinct aspects of data characteristics. The distinguishing features of low demand data include its inherent stability and predictability, exhibiting little fluctuations. In this context, the RF has significant success. The algorithm's ensemble nature, characterized by the combination of several DT, facilitates the creation of a very precise model. The extraordinary success of RF can be attributed to its versatility and its capacity to effectively handle the nuanced fluctuations present in low-demand data (Ahmad et al., 2017).

The energy usage trends observed during weekends exhibit a distinctive combination of data properties. The consumption of energy over weekends is subject to a variety of factors, encompassing individual behaviors, weather conditions, and special activities. The data

Table 4
Performance of evaluation metrics of single algorithms.

Algorithms	Subset-1			Subset-2			Subset-3		
	MSE (Wh)	R^2	MAPE %	MSE (Wh)	R^2	MAPE %	MSE (Wh)	R^2	MAPE %
KNN	2867.47	0.835	11.48	104.34	0.895	10.70	1044.67	0.904	18.99
DT	749.24	0.957	4.54	26.08	0.974	1.40	272.23	0.975	5.75
ANN	328.67	0.981	1.88	15.15	0.985	1.26	3.84	0.996	1.40

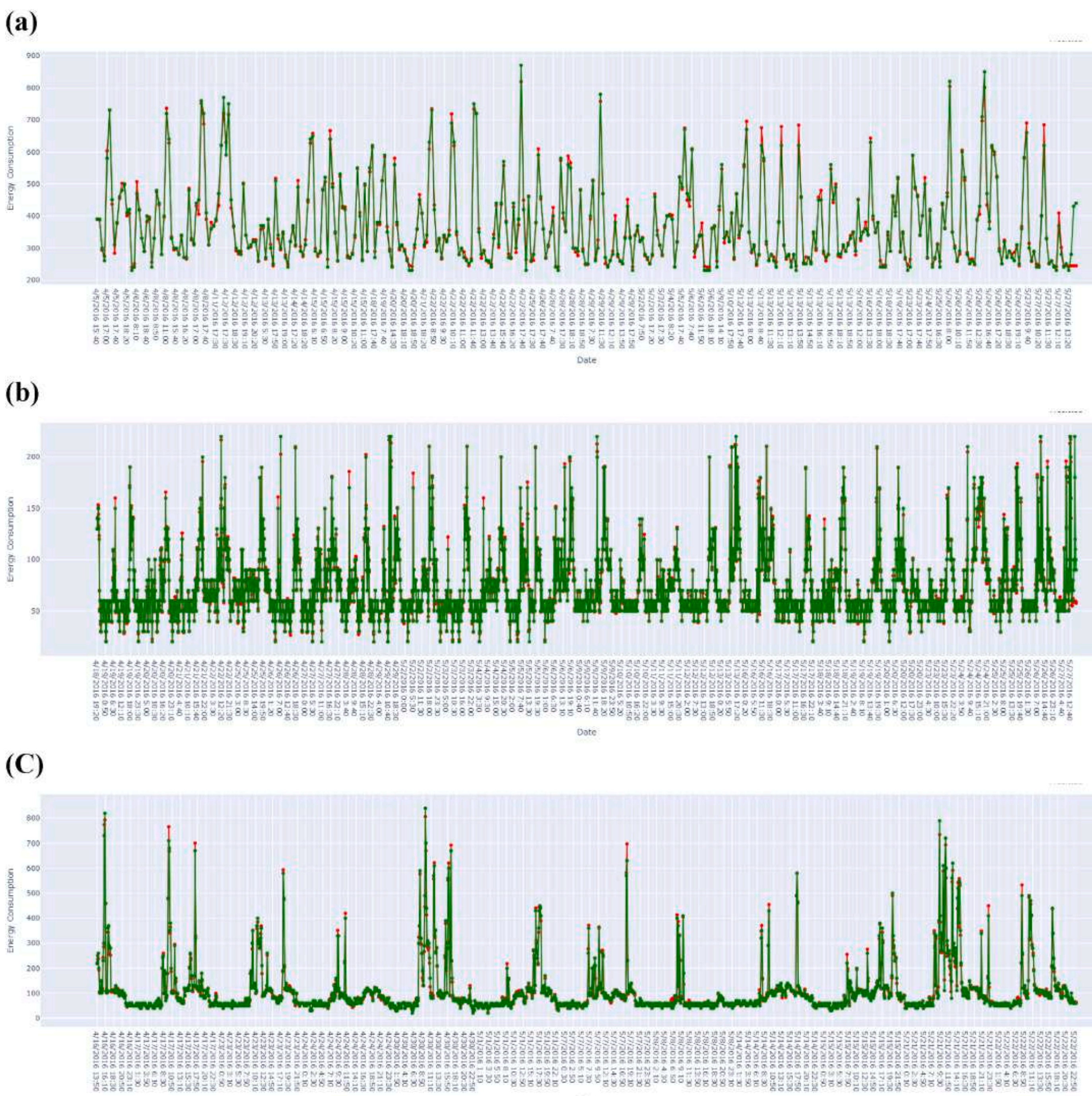

Fig. 12. Comparison of Actual and Predicted target variable in Buildings over Time Intervals: Assessing the Performance of Forecasting bagging (RF) Models. Predicted vs. actual energy consumption (a): Subset-1, (b): Subset-2, and (c): Subset-3.

Table 5
Performance of evaluation metrics of Bagging models.

Algorithms	Subset-1			Subset-2			Subset-3		
	MSE (Wh)	R^2	MAPE %	MSE (Wh)	R^2	MAPE %	MSE (Wh)	R^2	MAPE %
Bagging-RF	382.86	0.978	2.80	17.38	0.983	1.27	72.68	0.993	4.54

environment is characterized by intrinsic variability and frequent fluctuations. Despite the intricate nature of these difficulties, the RF algorithm consistently demonstrates a noteworthy degree of performance. The strong nature of its ensemble characteristics, along with its resilience to noise and dimensionality, is particularly evident in this scenario (Smarra et al., 2018). By combining numerous DT, the program effectively navigates the complex landscape of energy use patterns on weekends. The system demonstrates adaptability to the diverse and ever-changing elements that impact of target variable over the course of weekends.

4.2.2. Boosting

When prioritizing the prediction of peak energy usage, both XGB and GBT exhibit outstanding results. The remarkable accomplishment, as evidenced in Table 6, can be attributed to their proficiency in handling intricate and unpredictable data. The algorithms showcased in this study exhibit remarkable efficacy in capturing intricate, non-linear associations inherent in the dataset. Consequently, they are adept at properly managing abrupt fluctuations and volatility frequently encountered in scenarios characterized by peak demand, as illustrated in Fig. 13 (Bassi et al., 2021). Furthermore, the adaptive learning techniques utilized by these models are continuously improved to adapt to the ever-changing patterns of high demand (Chammas et al., 2019). The capacity to adapt is crucial in ensuring accurate predictions during periods characterized by heightened energy usage.

In the context of forecasting low energy demand scenarios in subset-2, both XGB and GBT algorithms consistently demonstrate remarkable performance. The efficacy of these entities in such situations can be ascribed to their adeptness in managing stability and the significance of their features. The aforementioned algorithms have the ability to sustain a substantial level of forecast precision, even when confronted with data that exhibits consistent and minimal energy use patterns (Huang et al., 2022). Further, these systems possess the capability to autonomously detect and assign priority to the most pertinent attributes, a characteristic that holds significant value in situations where minor fluctuations in data are pivotal for precise prognostications (Abediniangerabi et al., 2022).

In subset-3, the examination is directed towards target variable patterns during weekends. XGBoost and GBT algorithms exhibit robustness in the face of noise and fluctuations present in the dataset. The great performance of the subject under consideration can be attributed to their capacity to effectively adjust to the various elements that impact target variable on weekends, including individual behaviors and special events.

4.2.3. Stacking

The stacking approach effectively leverages the distinctive strengths of each component while simultaneously addressing their weaknesses, therefore offering a viable solution to the intricate challenge at hand (Sun et al., 2021). In the domain of peak demand forecasting, the utilization of Stacking demonstrates noteworthy efficacy. The process of stacking capitalizes on the combined intellectual capabilities of the six fundamental learners. The resulting predictive model has a notably low MSE, a high R-squared (R^2), and a remarkably low MAPE percentage. The statement above emphasizes the notable precision and adaptability of Stacking in handling the intricate and uncertain relationships found within the volatile data of maximum demand, as depicted in Fig. 14 (Cao et al., 2023). In settings characterized by low demand, Stacking continues to demonstrate its effectiveness. As mentioned above, the system consistently delivers accurate predictions, rendering it a compelling option for estimating household energy consumption. The stability handling skills of the model, derived from its different base learners, enable accurate predictions of energy use, even in settings characterized by consistent and low levels of target variable (Divina et al., 2018). Stacking effectively addresses the obstacles presented by target variable during weekends by exhibiting adeptness in handling noise and capturing intricate relationships among multiple factors that influence weekend energy usage. The characteristics demonstrate its reliability as a viable choice. Table 7 indicates its continuously low MSE values and a high R^2 when predicting weekend energy use (see Fig. 15).

The advantages of stacking become evident when it is compared to other methodologies, such as Single Algorithms (e.g., ANN, DT, and KNN), Bagging (RF), and Boosting (XGB&GBT). Stacking is a notable technique that combines the qualities of these six base learners to form a comprehensive and versatile strategy that effectively addresses diverse data features and obstacles. The consistent and accurate supply of extremely precise predictions, as demonstrated by low MAPE values across many situations, is of great value in the context of precise energy consumption forecasting. The adaptability of stacking is shown in its capacity to effectively manage a wide range of data properties, including the volatility associated with peak demand, the stability observed during periods of low demand, and the noise inherent in weekend consumption patterns. The model consistently has high R^2 values, indicating a strong fit and ability to explain the data variance effectively.

Additionally, it consistently achieves low MSE values, suggesting its proficiency in minimizing prediction mistakes. In summary, the use of Stacking exemplifies a resilient and adaptable methodology for predicting residential building energy consumption. This strategy is deemed highly suitable for this crucial undertaking due to its ability to amalgamate the advantageous attributes of six fundamental learners.

4.3. Performance of improved stacking models

To improve the precision of target variable prediction, the GA-Stacking method carefully chooses the most suitable combination of base models. Table 8 presents the compositions obtained for each scenario, namely subset-1 (Peak Demand), subset-2 (Low Demand), and subset-3 (Weekend Consumption).

Table 6
Performance of evaluation metrics of Boosting models.

Algorithms	Subset-1			Subset-2			Subset-3		
	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)
Boosting-XGB	542.64	0.969	3.68	19.86	0.980	2.12	123.58	0.989	5.39
Boosting-GBT	718.83	0.959	4.39	24.26	0.976	1.47	241.88	0.978	5.45

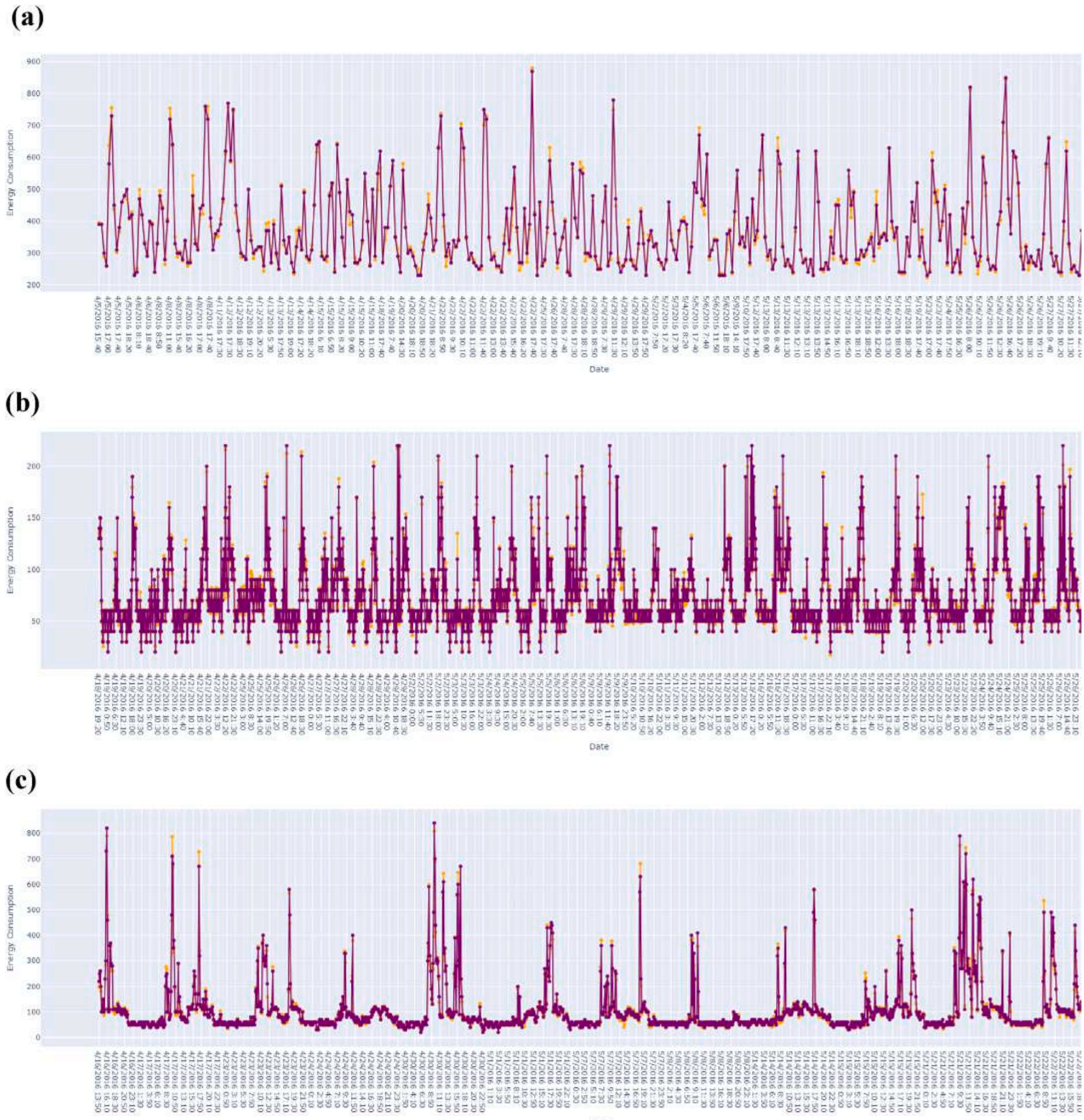

Fig. 13. Comparison of Actual and Predicted target variable in Buildings over Time Intervals: Assessing the Performance of Forecasting boosting (XGB) Models. Predicted vs. actual energy consumption (a): Subset-1, (b): Subset-2, and (c): Subset-3.

The process of choosing base models for different scenarios of projecting household energy use consistently exhibits a noticeable pattern in the field of GAs. The repetitive occurrence of this pattern underscores the adaptability of the GA in customizing combinations of base models to align with the unique attributes of each given scenario (Zheng et al., 2023). Adaptability is of utmost importance in the progression of Hybrid-Stacked models, as they provide enhanced accuracy in predictive capabilities. Furthermore, as seen from Table 9 and Fig. 15, using GAs for optimization led to a reduction in the quantity of base learners while maintaining a high level of accuracy. Additionally, there was a notable improvement in the MAPE metric.

Acknowledging that the traditional performance measurements, R^2 values, and MSE may exhibit minimal variations is important. Nevertheless, the significant alteration noticed in the MAPE data garners attention. The numbers mentioned above, which are utilized as a metric to gauge the precision of predictions, experience a significant decrease, ultimately reaching levels that are close to inconsequential. The substantial decrease indicates the models' impressive ability to generate target variable projections with high accuracy.

In conclusion, the combination of Stacking and GA optimization

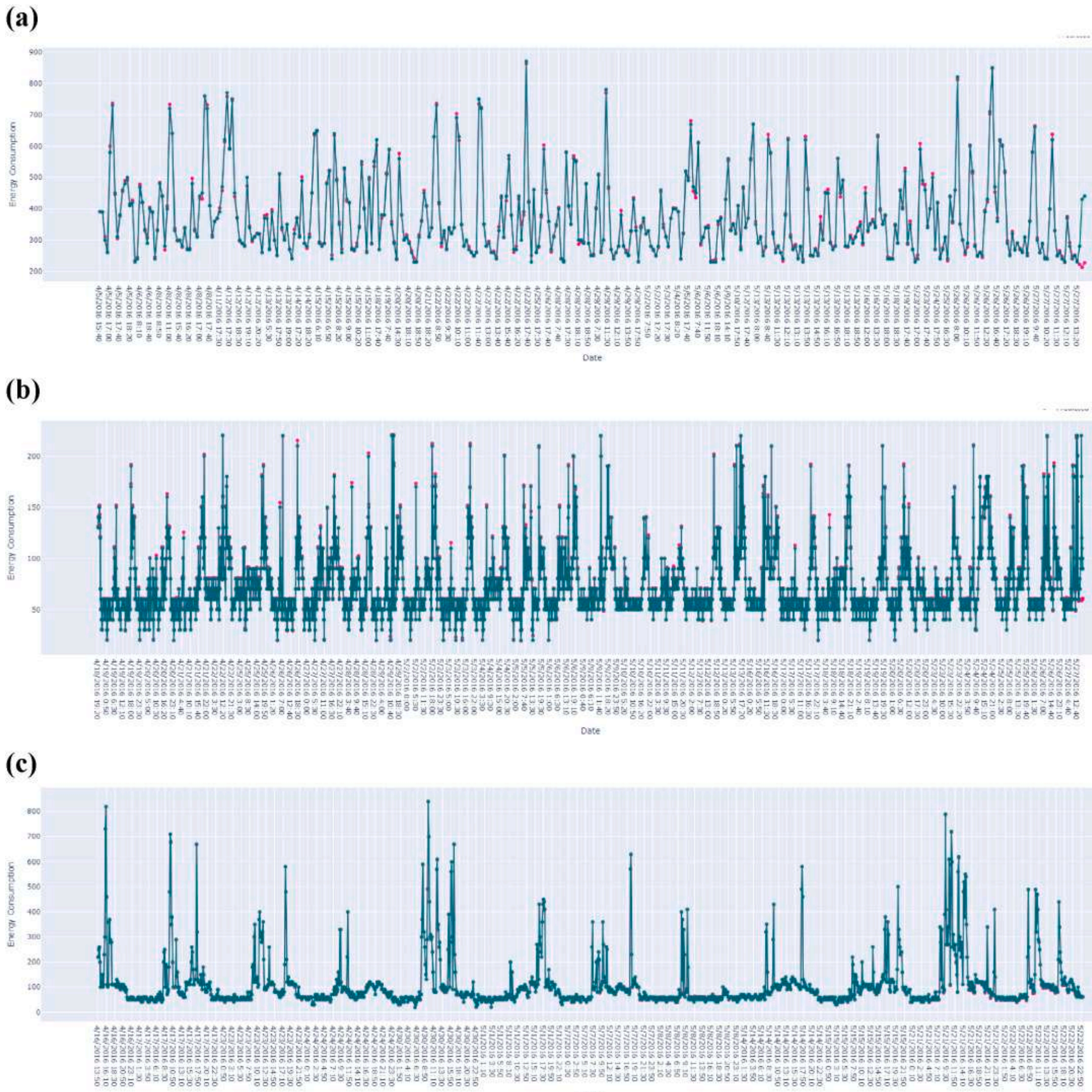

Fig. 14. Comparison of Actual and Predicted Target Variable in Buildings over Time Intervals: Assessing the Performance of Forecasting Stacking Models. Predicted vs. actual energy consumption (a): Subset-1, (b): Subset-2, and (c): Subset-3.

results in the development of Hybrid-Stacked models, significantly improved the accuracy of target variable forecasts. The significant decrease in MAPE demonstrates the accuracy and excellent performance of the models in providing precise estimates for household energy consumption. This establishes them as the ideal alternative for these crucial forecasting duties.

After conducting a thorough analysis of various algorithms and methods, it is evident that each approach presents a unique set of advantages and limitations. Table 10 offers a concise comparison to assist in selecting the most suitable model based on specific needs and constraints.

5. Conclusion

This paper addresses the significant challenge of accurately forecasting building energy consumption, which has garnered substantial attention in the scientific community in recent years. The results reported in this study have important significance for managers and policymakers in the global management and economics sectors. Through the utilization of sophisticated machine learning methods and predictive modeling, managers of organizations can get essential knowledge regarding energy consumption trends within buildings. This knowledge empowers them to make proactive decisions aimed at improving energy efficiency and optimizing resource allocation. This not only decreases

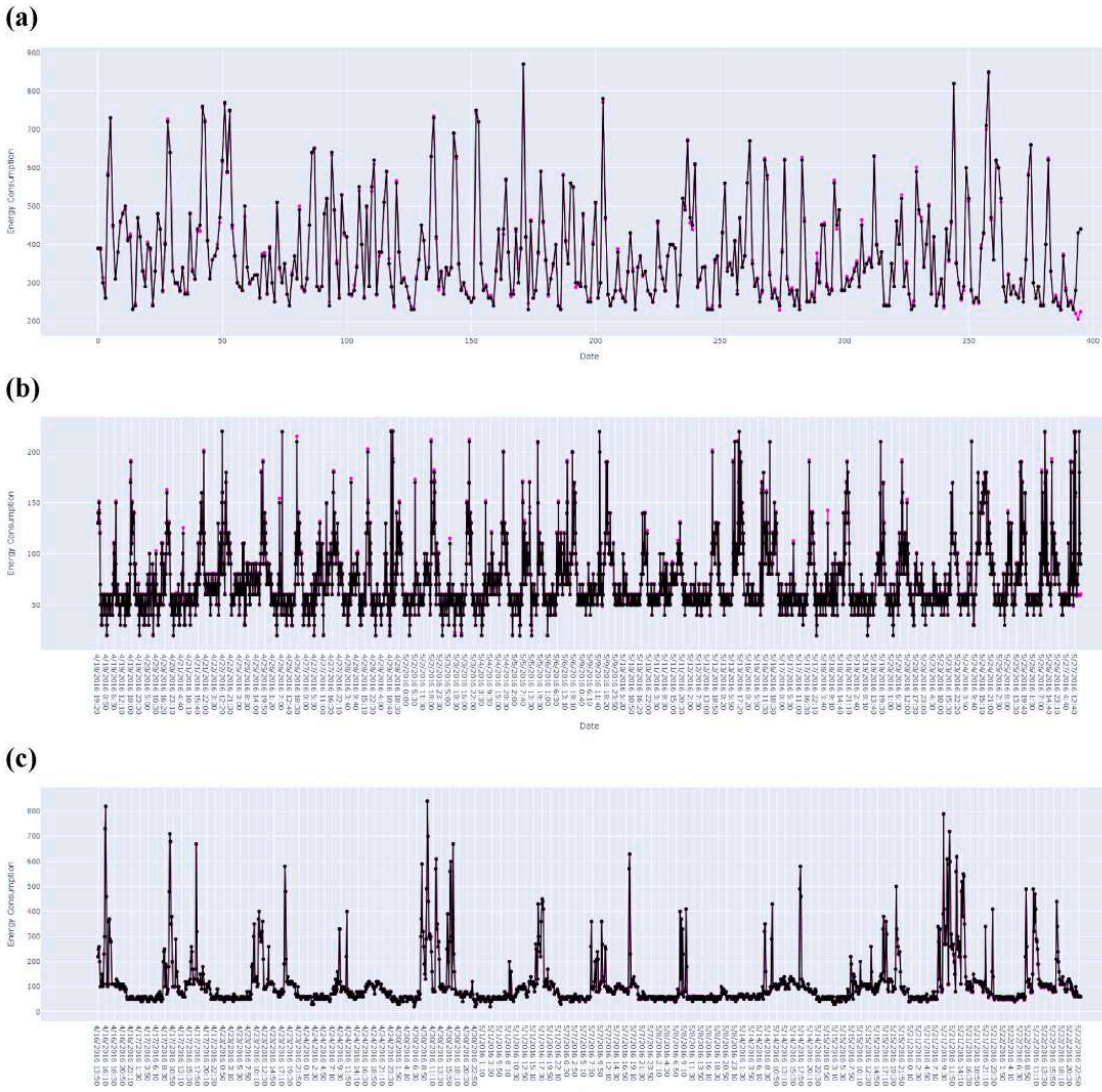

Fig. 15. Comparison of Actual and Predicted Target Variable in Buildings over Time Intervals: Assessing the Performance of Forecasting GA-stacked Models. Predicted vs. actual energy consumption (a): Subset-1, (b): Subset-2, and (c): Subset-3.

Table 7
Performance evaluation metrics of stacking models.

Algorithms	Subset-1			Subset-2			Subset-3		
	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)
Stacking	297.78	0.983	1.84	15.15	0.985	0.26	3.19	0.999	1.37

operational expenses but also corresponds with sustainability objectives, which have grown progressively crucial in the contemporary global economy. Policymakers can employ these prediction models to guide the development of regulations and incentives that encourage the adoption of environmentally friendly architectural designs and energy-efficient practices in different industries. It provides a framework that can be generally used to make educated decisions in building management and design. By incorporating these approaches into their operations, managers and policymakers of enterprises can promote sustainable development, reduce environmental impact, and optimize economic

Table 8
Optimal Base Model Combinations for Enhanced Target Variable Forecasting (GA outputs).

	Subset-1	Subset-2	Subset-3
Base-learners combination	GBT, XGB, ANN	RF, GBT, XGB, ANN	GBT, ANN, DT

outcomes worldwide. In this study the dataset is divided into weekday and weekend subsets to tackle this issue. Then, K-means cluster algorithms are employed on the weekday subset to identify patterns of low and peak energy consumption during a day on a 10-min basis, providing valuable insights into energy consumption within the study. ML methods, such as single algorithms and ensemble models, are utilized to achieve the best prediction performance. The single ML algorithms used in this study include KNN, DT, and ANN.

Furthermore, ensemble models such as bagging (RF), boosting (XGB&GBT), and stacking are employed to leverage the collective power of multiple models. Hyperparameter tuning is performed to optimize the performance of each model. Additionally, GAs is applied to the stacked model to improve the model's performance and determine the optimal combination of base estimators. Comparing the various algorithms indicates that the GA-stacked ensemble model outperforms other methods in all three models, with MAPE test scores of 0.016, 0.006, and 0.012% for the subset-1, subset-2, and subset-3, respectively. This finding suggests that the combination of GAs and the stacking approach leads to superior predictive capabilities in our study and can be an appropriate Framework to achieve the best result in the forecasting energy consumption field in buildings.

6. Future work

Forecasting energy consumption is crucial, so future work can be expanded to develop a comprehensive framework and establish proper guidelines for collecting efficient features. These features should be capable of aiding decision-makers and engineers in utilizing them effectively in forecasting energy consumption for different types of buildings, including commercial and residential structures, under diverse meteorological conditions. Another avenue worth exploring in future research is identifying and recommending efficient solutions for reducing and decentralizing energy consumption.

Although our study shows encouraging outcomes in forecasting building energy usage using ensemble models, it is important to acknowledge numerous limitations. The reliability and generalizability of our findings may be compromised by data restrictions, such as the extent and inclusiveness of the dataset, as well as concerns like missing or inaccurate data. Moreover, the inherent trade-offs between the accuracy and interpretability of ensemble approaches give rise to worries about the practical implementation of our findings. Ensemble models, although more accurate, tend to be more intricate and less explainable than simpler models such as linear regression. Furthermore, the computing resources needed for training and deploying ensemble models, particularly with extensive datasets.

CRediT authorship contribution statement

Mahziyar Dostmohammadi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mona Zamani Pedram:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Siamak Hoseinzadeh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project

Table 9
Performance evaluation metrics of hybrid-stacked models.

Algorithms	Subset-1			Subset-2			Subset-3		
	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)	MSE (Wh)	R^2	MAPE (%)
Hybrid-Stacked	292.18	0.983	0.016	14.86	0.985	0.006	2.75	0.999	0.012

Table 10
Comparative Analysis of Classical Algorithms and Ensemble Methods: Evaluating the Advantages and Disadvantages of Enhanced Predictive Modeling in Target variable Forecasting.

Algorithms	Advantageous	Disadvantages
ANN	1. Captures subtle correlations in data 2. Learns features automatically 3. Adaptive learning through backpropagation 4. Continuous improvement in prediction accuracy	1. Complex and may require significant computational resources 2. May need large amounts of data for training
KNN	1. Simplicity, versatility, and adaptability 2. Effective in capturing localized patterns 3. Can automatically choose relevant features	1. Susceptibility to noisy or irrelevant data 2. Performance varies based on data variations 3. Ineffective at capturing broad worldwide trends
DT	1. Intrinsic feature selection 2. Mitigates overfitting with pruning strategies 3. Balances predictive accuracy and generalizability	1. Limited to representing linear decision boundaries 2. May not perform well in complex situations
Bagging (RF)	1.Excellent performance in volatile, unpredictable data scenarios 2. Efficiently harnesses the potential of multiple decision trees 3. Captures non-linear patterns and deep linkages in data 4. Resilient to fluctuations and volatility in the data	1. Resilient to fluctuations and volatility in the data
Boosting (XGB&GBT)	1. Proficiency in handling intricate and unpredictable data 2. Adeptness in managing stability and relevant features 3. Adept at managing abrupt fluctuations and volatility 4. Adaptive learning techniques for ever-changing patterns	1. Computational Intensity 2. Difficulty in Parallelization
Stacking	1. Leverages strengths of individual components 2. Addresses weaknesses of the base learners 3. Precision and adaptability in handling intricate and uncertain data relationships 4. Handles noise and captures intricate relationships	1. Computational Intensity 2. Risk of Model Selection Bias

administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Davide Astiaso Garcia:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Abediniangerabi, B., Makhmalbaf, A., Shahandashti, M., 2022. Estimating energy savings of ultra-high-performance fibre-reinforced concrete facade panels at the early design stage of buildings using gradient boosting machines. Adv. Build. Energy Res. 16, 542–567. https://doi.org/10.1080/17512549.2021.2011410.

Ahmad, M.W., Mourshed, M., Rezgui, Y., 2017. Trees vs Neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy Build. 147, 77–89. https://doi.org/10.1016/j.enbuild.2017.04.038.

Ahsan, M., Mahmud, M., Saha, P., et al., 2021. Effect of data scaling methods on machine learning algorithms and model performance. Technologies 9, 52. https://doi.org/10.3390/technologies9030052.

Al-Mufti, O.A., Al-Isawi, O.A., Amirah, L.H., Ghenai, C., 2023. Digital twinning and ANN-based forecasting model for building energy consumption. In: 2023 Advances in Science and Engineering Technology International Conferences (ASET). IEEE, pp. 1–8.

Alshboul, O., Shehadeh, A., Almasabha, G., Almuflih, A.S., 2022. Extreme gradient boosting-based machine learning approach for green building cost prediction. Sustainability 14, 6651. https://doi.org/10.3390/su14116651.

Amber, K.P., Ahmad, R., Aslam, M.W., et al., 2018. Intelligent techniques for forecasting electricity consumption of buildings. Energy 157, 886–893. https://doi.org/10.1016/j.energy.2018.05.155.

Amir-Ahmadi, P., Matthes, C., Wang, M.-C., 2020. Choosing prior hyperparameters: with applications to time-varying parameter models. J. Bus. Econ. Stat. 38, 124–136. https://doi.org/10.1080/07350015.2018.1459302.

B, P., V, T., C, K., 2023. Machine learning based demand response scheme for IoT enabled PV integrated smart building. Sustain. Cities Soc. 89, 104260 https://doi.org/10.1016/j.scs.2022.104260.

Bassi, A., Shenoy, A., Sharma, A., et al., 2021. Building energy consumption forecasting: a comparison of gradient boosting models. In: The 12th International Conference on Advances in Information Technology. ACM, New York, NY, USA, pp. 1–9.

Belete, D.M., Huchaiah, M.D., 2022. Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. Int. J. Comput. Appl. 44, 875–886. https://doi.org/10.1080/1206212X.2021.1974663.

Biresselioglu, M.E., Demir, M.H., 2022. Constructing a decision tree for energy policy domain based on real-life data. Energies 15, 2420. https://doi.org/10.3390/en15072420.

Bourdeau, M., qiang Zhai, X., Nefzaoui, E., et al., 2019. Modeling and forecasting building energy consumption: a review of data-driven techniques. Sustain. Cities Soc. 48, 101533 https://doi.org/10.1016/j.scs.2019.101533.

Bühlmann, P., Yu, B., 2002. Analyzing bagging. Ann. Stat. 30 https://doi.org/10.1214/aos/1031689014.

Candanedo, L.M., Feldheim, V., Deramaix, D., 2017. Data driven prediction models of energy use of appliances in a low-energy house. Energy Build. 140, 81–97. https://doi.org/10.1016/j.enbuild.2017.01.083.

Cao, Y., Liu, G., Sun, J., et al., 2023. PSO-Stacking improved ensemble model for campus building energy consumption forecasting based on priority feature selection. J. Build. Eng. 72, 106589 https://doi.org/10.1016/j.jobe.2023.106589.

Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geosci. Model Dev. (GMD) 7, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014.

Chammas, M., Makhoul, A., Demerjian, J., 2019. An efficient data model for energy prediction using wireless sensors. Comput. Electr. Eng. 76, 249–257. https://doi.org/10.1016/j.compeleceng.2019.04.002.

Chen, Y., Guo, M., Chen, Z., et al., 2022. Physical energy and data-driven models in building energy prediction: a review. Energy Rep. 8, 2656–2671. https://doi.org/10.1016/j.egyr.2022.01.162.

Ciupageanu, D.-A., Barelli, L., Lazaroiu, G., 2020. Real-time stochastic power management strategies in hybrid renewable energy systems: a review of key applications and perspectives. Elec. Power Syst. Res. 187, 106497 https://doi.org/10.1016/j.epsr.2020.106497.

Cui, Y., Zhu, Z., Zhao, X., Li, Z., 2023. Energy schedule setting based on clustering algorithm and pattern recognition for non-residential buildings electricity energy consumption. Sustainability 15, 8750. https://doi.org/10.3390/su15118750.

Czajkowski, M., Jurczuk, K., Kretowski, M., 2023. Steering the interpretability of decision trees using lasso regression - an evolutionary perspective. Inf. Sci. 638, 118944 https://doi.org/10.1016/j.ins.2023.118944.

de Rautlin de la Roy, E., Recht, T., Zemmari, A., et al., 2023. Deep learning models for building window-openings detection in heating season. Build. Environ. 231, 110019 https://doi.org/10.1016/j.buildenv.2023.110019.

Divina, F., Gilson, A., Goméz-Vela, F., et al., 2018. Stacking ensemble learning for short-term electricity consumption forecasting. Energies 11, 949. https://doi.org/10.3390/en11040949.

D'souza, R.N., Huang, P.-Y., Yeh, F.-C., 2020. Structural analysis and optimization of convolutional neural networks with a small sample size. Sci. Rep. 10, 834. https://doi.org/10.1038/s41598-020-57866-2.

Feng, X., Zhao, J., Kita, E., 2021. Genetic algorithm-based optimization of deep neural network ensemble. Rev Socionetwork Strateg 15, 27–47. https://doi.org/10.1007/s12626-021-00074-9.

Feng, X., Bin, Zainudin E., Wong, H.W., Tseng, K.J., 2023. A hybrid ensemble learning approach for indoor thermal comfort predictions utilizing the ASHRAE RP-884 database. Energy Build. 290, 113083 https://doi.org/10.1016/j.enbuild.2023.113083.

Figueiredo Filho, D.B., Silva Júnior, J.A., Rocha, E.C., 2011. What is R2 all about? Leviathan 60. https://doi.org/10.11606/issn.2237-4485.lev.2011.132282.

Ganaie, M.A., Hu, M., Malik, A.K., et al., 2022. Ensemble deep learning: a review. Eng. Appl. Artif. Intell. 115, 105151 https://doi.org/10.1016/j.engappai.2022.105151.

Hadri, S., Naitmalek, Y., Najib, M., et al., 2019. A comparative study of predictive approaches for load forecasting in smart buildings. Procedia Comput. Sci. 160, 173–180. https://doi.org/10.1016/j.procs.2019.09.458.

Hamidinasab, B., Javadikia, H., Hosseini-Fashami, F., et al., 2023. Illuminating sustainability: a comprehensive review of the environmental life cycle and exergetic impacts of solar systems on the agri-food sector. Sol. Energy 262, 111830. https://doi.org/10.1016/j.solener.2023.111830.

Hong, G., Choi, G.-S., Eum, J.-Y., et al., 2022. The hourly energy consumption prediction by KNN for buildings in community buildings. Buildings 12, 1636. https://doi.org/10.3390/buildings12101636.

Huang, J., Algahtani, M., Kaewunruen, S., 2022. Energy forecasting in a public building: a benchmarking analysis on long short-term memory (lstm), support vector regression (SVR), and extreme gradient boosting (XGBoost) networks. Appl. Sci. 12, 9788. https://doi.org/10.3390/app12199788.

Huang, H., Zhu, Q., Zhu, X., Zhang, J., 2023. An adaptive, data-driven stacking ensemble learning framework for the short-term forecasting of renewable energy generation. Energies 16, 1963. https://doi.org/10.3390/en16041963.

Jeon, H., Oh, S., 2020. Hybrid-recursive feature elimination for efficient feature selection. Appl. Sci. 10, 3211. https://doi.org/10.3390/app10093211.

Jiang, T., Li, J., Zheng, Y., Sun, C., 2011. Improved bagging algorithm for pattern recognition in UHF signals of partial discharges. Energies 4, 1087–1101. https://doi.org/10.3390/en4071087.

Jiao, Y., Tan, Z., Zhang, D., Zheng, Q.P., 2023. Short-term building energy consumption prediction strategy based on modal decomposition and reconstruction algorithm. Energy Build. 290, 113074 https://doi.org/10.1016/j.enbuild.2023.113074.

Jin, Y., Yan, D., Kang, X., et al., 2021. Forecasting building occupancy: a temporal-sequential analysis and machine learning integrated approach. Energy Build. 252, 111362 https://doi.org/10.1016/j.enbuild.2021.111362.

Kang, Y., Ma, N., Bunster, V., et al., 2022. Optimizing the Passive House Planning Package simulation tool: a bottom-up dynamic approach to reduce building performance gap. Energy Build. 276, 112512 https://doi.org/10.1016/j.enbuild.2022.112512.

Kim, S., Kim, H., 2016. A new metric of absolute percentage error for intermittent demand forecasts. Int. J. Forecast. 32, 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003.

Krzywinski, M., Altman, N., 2014. Visualizing samples with box plots. Nat. Methods 11, 119–120. https://doi.org/10.1038/nmeth.2813.

Kumar, Y., Kaul, S., Hu, Y.-C., 2022. Machine learning for energy-resource allocation, workflow scheduling and live migration in cloud computing: state-of-the-art survey. Sustain Comput Informatics Syst 36, 100780. https://doi.org/10.1016/j.suscom.2022.100780.

Li, Y., O'Neill, Z., Zhang, L., et al., 2021a. Grey-box modeling and application for building energy simulations - a critical review. Renew. Sustain. Energy Rev. 146, 111174 https://doi.org/10.1016/j.rser.2021.111174.

Li, Y., Rong, Y., Ahmad, U.M., et al., 2021b. A comprehensive review on green buildings research: bibliometric analysis during 1998–2018. Environ. Sci. Pollut. Res. 28, 46196–46214. https://doi.org/10.1007/s11356-021-12739-7.

Liu, Y., Chen, H., Zhang, L., Feng, Z., 2021. Enhancing building energy efficiency using a random forest model: a hybrid prediction approach. Energy Rep. 7, 5003–5012. https://doi.org/10.1016/j.egyr.2021.07.135.

Lu, C., Li, S., Lu, Z., 2022. Building energy prediction using artificial neural networks: a literature survey. Energy Build. 262, 111718 https://doi.org/10.1016/j.enbuild.2021.111718.

Lu, C., Jing, Y., Lin, X., Li, K., 2023. Predicting building energy consumption using stacking-based ensemble model. Am. J. Energy Eng. https://doi.org/10.11648/j.ajee.20231102.14.

Macas, M., Moretti, F., Fonti, A., et al., 2016. The role of data sample size and dimensionality in neural network based forecasting of building heating related variables. Energy Build. 111, 299–310. https://doi.org/10.1016/j.enbuild.2015.11.056.

Mehdizadeh Khorrami, B., Soleimani, A., Pinnarelli, A., et al., 2023. Forecasting heating and cooling loads in residential buildings using machine learning: a comparative study of techniques and influential indicators. Asian J Civ Eng. https://doi.org/10.1007/s42107-023-00834-8.

Metzler, R., 2020. Superstatistics and non-Gaussian diffusion. Eur. Phys. J. Spec. Top. 229, 711–728. https://doi.org/10.1140/epjst/e2020-900210-x.

Mienye, I.D., Sun, Y., Wang, Z., 2019. Prediction performance of improved decision tree-based algorithms: a review. Procedia Manuf. 35, 698–703. https://doi.org/10.1016/j.promfg.2019.06.011.

Mohammed, A., Kora, R., 2023. A Comprehensive Review on Ensemble Deep Learning: Opportunities and Challenges, vol. 35. J King Saud Univ - Comput Inf Sci, pp. 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014.

Morteza, A., Yahyaeian, A.A., Mirzaeibonehkhater, M., et al., 2023. Deep learning hyperparameter optimization: application to electricity and heat demand prediction for buildings. Energy Build. 289, 113036 https://doi.org/10.1016/j.enbuild.2023.113036.

Mosavi, A., Sajedi Hosseini, F., Choubin, B., et al., 2021. Ensemble boosting and bagging based machine learning models for groundwater potential prediction. Water Resour. Manag. 35, 23–37. https://doi.org/10.1007/s11269-020-02704-3.

Nabavi-Pelesaraei, A., Saber, Z., Mostashari-Rad, F., et al., 2021. Coupled life cycle assessment and data envelopment analysis to optimize energy consumption and mitigate environmental impacts in agricultural production. In: Methods in Sustainability Science. Elsevier, pp. 227–264.

Nabavi-Pelesaraei, A., Ghasemi-Mobtaker, H., Salehi, M., et al., 2023. Machine learning models of exergoenvironmental damages and emissions social cost for mushroom production. Agronomy 13, 737. https://doi.org/10.3390/agronomy13030737.

Nainwal, R., Sharma, A., 2022. Comparison of multi linear regression and artificial neural network to predict the energy consumption of residential buildings. IOP Conf. Ser. Earth Environ. Sci. 1084, 012005 https://doi.org/10.1088/1755-1315/1084/1/012005.

Nepal, Yamaha, Sahashi, Yokoe, 2019. Analysis of building electricity use pattern using K-means clustering algorithm by determination of better initial centroids and number of clusters. Energies 12, 2451. https://doi.org/10.3390/en12122451.

Nguyen, Q.H., Ly, H.-B., Ho, L.S., et al., 2021. Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Math. Probl Eng. 2021, 1–15. https://doi.org/10.1155/2021/4832864.

Nie, P., Roccotelli, M., Fanti, M.P., et al., 2021. Prediction of home energy consumption based on gradient boosting regression tree. Energy Rep. 7, 1246–1255. https://doi.org/10.1016/j.egyr.2021.02.006.

Niquini, F.G.F., Branches, A.M.B., Costa, J.F.C.L., et al., 2023. Recursive feature elimination and neural networks applied to the forecast of mass and metallurgical recoveries in A Brazilian phosphate mine. Minerals 13, 748. https://doi.org/10.3390/min13060748.

Olu-Ajayi, R., Alaka, H., Sulaimon, I., et al., 2022. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. J. Build. Eng. 45, 103406 https://doi.org/10.1016/j.jobe.2021.103406.

Olu-Ajayi, R., Alaka, H., Sulaimon, I., et al., 2023. Building energy performance prediction: a reliability analysis and evaluation of feature selection methods. Expert Syst. Appl. 225, 120109 https://doi.org/10.1016/j.eswa.2023.120109.

Pachauri, N., Ahn, C.W., 2023. Weighted aggregated ensemble model for energy demand management of buildings. Energy 263, 125853. https://doi.org/10.1016/j.energy.2022.125853.

Palaniswamy, S.K., Venkatesan, R., 2021. Retracted article: hyperparameters tuning of ensemble model for software effort estimation. J. Ambient Intell. Hum. Comput. 12, 6579–6589. https://doi.org/10.1007/s12652-020-02277-4.

Phu, L.V., 2021. Nonlinear temperature response of electricity loads and implications for power development policies in Vietnam. Energy Build. 251, 111339 https://doi.org/10.1016/j.enbuild.2021.111339.

Qavidelfardi, Z., Tahsildoost, M., Zomorodian, Z.S., 2022. Using an ensemble learning framework to predict residential energy consumption in the hot and humid climate of Iran. Energy Rep. 8, 12327–12347. https://doi.org/10.1016/j.egyr.2022.09.066.

Qu, Z., Liu, H., Wang, Z., et al., 2021. A combined genetic optimization with AdaBoost ensemble model for anomaly detection in buildings electricity consumption. Energy Build. 248, 111193 https://doi.org/10.1016/j.enbuild.2021.111193.

Runge, J., Zmeureanu, R., 2019. Forecasting energy use in buildings using artificial neural networks: a review. Energies 12, 3254. https://doi.org/10.3390/en12173254.

Saeidi, E., Dehkordi, A.L., Nabavi-Pelesaraei, A., 2022. Potential for optimization of energy consumption and costs in saffron production in central Iran through data envelopment analysis and <scp>multi-objective</scp> genetic algorithm. Environ. Prog. Sustain. Energy 41. https://doi.org/10.1002/ep.13857.

Shahhosseini, M., Hu, G., Pham, H., 2022. Optimizing ensemble weights and hyperparameters of machine learning models for regression problems. Mach Learn with Appl 7, 100251. https://doi.org/10.1016/j.mlwa.2022.100251.

Shapi, M.K.M., Ramli, N.A., Awalin, L.J., 2021. Energy consumption prediction by using machine learning for smart building: case study in Malaysia. Dev Built Environ 5, 100037. https://doi.org/10.1016/j.dibe.2020.100037.

Singh, N.K., Fukushima, T., Nagahara, M., 2023. Gradient boosting approach to predict energy-saving awareness of households in kitakyushu. Energies 16, 5998. https://doi.org/10.3390/en16165998.

Smarra, F., Jain, A., de Rubeis, T., et al., 2018. Data-driven model predictive control using random forests for building energy optimization and climate control. Appl. Energy 226, 1252–1272. https://doi.org/10.1016/j.apenergy.2018.02.126.

Sun, Y., Haghighat, F., Fung, B.C.M., 2020. A review of the-state-of-the-art in data-driven approaches for building energy prediction. Energy Build. 221, 110022 https://doi.org/10.1016/j.enbuild.2020.110022.

Sun, J., Liu, G., Sun, B., Xiao, G., 2021. Light-stacking strengthened fusion based building energy consumption prediction framework via variable weight feature selection. Appl. Energy 303, 117694. https://doi.org/10.1016/j.apenergy.2021.117694.

Touzani, S., Granderson, J., Fernandes, S., 2018. Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy Build. 158, 1533–1543. https://doi.org/10.1016/j.enbuild.2017.11.039.

Truong, N.-S., Ngo, N.-T., Pham, A.-D., 2021. Forecasting time-series energy data in buildings using an additive artificial intelligence model for improving energy efficiency. Comput. Intell. Neurosci. 2021, 1–12. https://doi.org/10.1155/2021/6028573.

Uddin, S., Haque, I., Lu, H., et al., 2022. Comparative performance analysis of K - nearest neighbour (KNN) algorithm and its different variants for disease prediction. Sci. Rep. 1–11. https://doi.org/10.1038/s41598-022-10358-x.

Valgaev, O., Kupzog, F., Schmeck, H., 2017. Building power demand forecasting using K-nearest neighbours model – practical application in Smart City Demo Aspern project. CIRED - Open Access Proc J 2017, 1601–1604. https://doi.org/10.1049/oap-cired.2017.0419.

Wang, Z., Wang, Y., Zeng, R., et al., 2018. Random Forest based hourly building energy prediction. Energy Build. 171, 11–25. https://doi.org/10.1016/j.enbuild.2018.04.008.

Wang, Z., Liang, Z., Zeng, R., et al., 2023. Identifying the optimal heterogeneous ensemble learning model for building energy prediction using the exhaustive search method. Energy Build. 281, 112763 https://doi.org/10.1016/j.enbuild.2022.112763.

Wen, H.-T., Wu, H.-Y., Liao, K.-C., Chen, W.-C., 2023. JT9D engine thrust estimation and model sensitivity analysis using gradient boosting regression method. Aerospace 10, 639. https://doi.org/10.3390/aerospace10070639.

Wu, Z., Li, N., Peng, J., et al., 2018. Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings. Energy Build. 173, 117–127. https://doi.org/10.1016/j.enbuild.2018.05.031.

Yang, H., Ran, M., Feng, H., 2023. Improved data-driven building daily energy consumption prediction models based on balance point temperature. Buildings 13, 1423. https://doi.org/10.3390/buildings13061423.

Yao, J., Zhang, X., Luo, W., et al., 2022. Applications of Stacking/Blending ensemble learning approaches for evaluating flash flood susceptibility. Int. J. Appl. Earth Obs. Geoinf. 112, 102932 https://doi.org/10.1016/j.jag.2022.102932.

Yin, Y., Jang-Jaccard, J., Xu, W., et al., 2023. IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. J Big Data 10, 15. https://doi.org/10.1186/s40537-023-00694-8.

Zapata-Webborn, E., McKenna, E., Pullinger, M., et al., 2023. The impact of COVID-19 on household energy consumption in England and Wales from April 2020 to March 2022. Energy Build. 297, 113428 https://doi.org/10.1016/j.enbuild.2023.113428.

Zhang, G., Tian, C., Li, C., et al., 2020. Accurate forecasting of building energy consumption via a novel ensembled deep learning method considering the cyclic feature. Energy 201, 117531. https://doi.org/10.1016/j.energy.2020.117531.

Zhang, Q., Kong, J., Jiang, M., Liu, T., 2023. Building energy consumption prediction based on temporal-aware attention and energy consumption states. J Electr Eng Technol 18, 61–75. https://doi.org/10.1007/s42835-022-01159-3.

Zheng, J., Wang, M., Yao, T., et al., 2023. Dynamic mechanical strength prediction of BFRC based on stacking ensemble learning and genetic algorithm optimization. Buildings 13, 1155. https://doi.org/10.3390/buildings13051155.

Zhou, Y., Liu, Y., Wang, D., Liu, X., 2021. Comparison of machine-learning models for predicting short-term building heating load using operational parameters. Energy Build. 253, 111505 https://doi.org/10.1016/j.enbuild.2021.111505.