Type-2 Fuzzy Classifier with Smooth Type-Reduction

Katarzyna Nieszporek[1[0000-0002-9424-2646]], Giorgio De Magistris[2[0000-0002-3076-4509]], Christian Napoli[2[0000-0002-3336-5853]], and Janusz T. Starczewski[1[0000-0003-4694-7868]]

[1] Department of Intelligent Computer Systems, Częstochowa University of Technology, Częstochowa, Poland
katarzyna.nieszporek@pcz.pl
[2] Department of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25 Roma 00185, Italy

Abstract. The defuzzification of a type-2 fuzzy set is a two-stage process consisting of firstly type-reduction, and a secondly defuzzification of the resultant type-1 set. All accurate type reduction methods used to build fuzzy classifiers are based on the recursive Karnik-Mendel algorithm, which is troublesome to obtain a feedforward type-2 fuzzy network structure. Moreover, the KM algorithm and its modifications complicate the learning process due to the non-differentiability of the maximum and minimum functions. Therefore, this paper proposes to use the smooth maximum function to develop a new structure of the fuzzy type-2 classifier.

Keywords: smooth type reduction, interval type-2 fuzzy logic systems

1 Introduction

In recent years, fuzzy logic methodology has shown to be very effective in solving complex nonlinear systems containing uncertainties that are otherwise challenging. However, it is also noted that fuzzy rules working in an uncertain or non-stationary environment require a higher order of fuzziness. This is due to the fact that type-1 fuzzy sets, whose membership grades are real numbers, could have limitations in minimizing the effect of uncertainty, whereas the membership grades of a type-2 fuzzy logic system are themselves fuzzy logic systems in [0, 1]. Describing a type-2 fuzzy set by a rectangular membership function sufficiently describes the uncertainty in modeling of most processes. However, for the output, there will need a type reduction to convert the output of the fuzzy inference engine into a type 1 fuzzy sets before defuzzification can be performed to obtain a crisp output. The center-of-sets iterative Karnik-Mendel (KM) approach to type reduction is of great interest. Over the years, modifications have been made to the basic KM algorithm, including Wu and Tan [17] who presented their concept using a genetic algorithm. In this paper, the smooth method is used to design efficient type reduction algorithms.

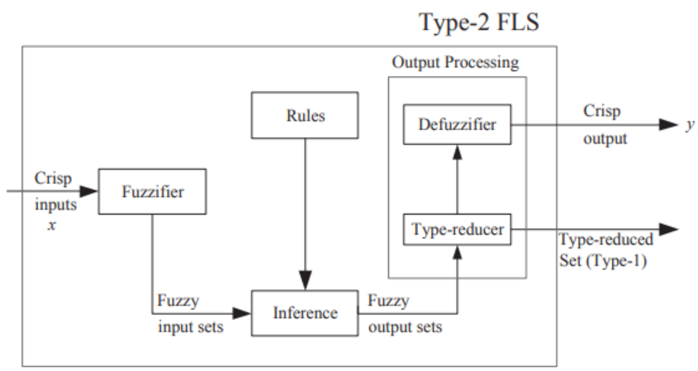

Fig. 1. Type-2 FLS (from Mendel [8])

2 An Overview

The Type-1 Fuzzy Set Let X be a universe of discourse. A fuzzy set A on X is characterised by a membership function $\mu_A : X \to [0,1]$ and can be represented as follows:

$$A = \{(x, \mu_A(x)); \mu_A(x) \in [0,1] \forall x \in X\} \tag{1}$$

The Type-2 Fuzzy Set Let $\tilde{P}(U)$, where $U = [0,1]$, be set of fuzzy sets in U. A type-2 fuzzy set $\tilde{A}$ in X is a fuzzy set whose membership grades are themselves fuzzy [23].

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)); \mu_{\tilde{A}}(x) \in \tilde{P}(U) \forall x \in X\} \tag{2}$$

where $\mu_{\tilde{A}(x)}$ is a fuzzy set in U for all x, i.e. $\mu_{\tilde{A}(x)} : X \to \tilde{P}(U)$.
It implies that $\forall x \in X\ \exists J_x \subseteq U$ such that $\mu_{\tilde{A}}(x) : J_x \to U$ [4].

$$\mu_A(x)) = \{(u, \mu_{\tilde{A}}(x)(u)) | \mu_{\tilde{A}}(x)(u) \in U \forall u \in J_x \subseteq U\} \tag{3}$$

where X is called the primary domain, J_x the primary membership of x, U is known as the secondary domain and $\mu_{\tilde{A}}(x)$ is the secondary membership of x.

In this paper, an interval singleton type-2 fuzzy logic system type is used. This means that the fuzzifier converts the fuzzy logic system input signals into fuzzy singletons and then the inference engine adjusts the fuzzy singletons with the fuzzy rules in the rule base.

Considering a type-2 fuzzy system with K rules will be used with the following scheme [7]:

$$\widetilde{R}^k\colon \text{IF } \widetilde{A}' \text{ is } \widetilde{A}_k \text{ THEN } \widetilde{B}' \text{ is } \widetilde{B}_k. \tag{4}$$

where $\widetilde{A}'$, $\widetilde{A}_k$, $\widetilde{B}'$ and $\widetilde{B}_k$ are type-2 fuzzy sets. In the interval case, they are subintervals of $[0,1]$ expressed by of upper and lower bounds, e.g. $\widetilde{A}_k = \left[\underline{\mu}_{A_k}(x), \overline{\mu}_{A_k}(x)\right] \subseteq [0,1]$ for each $x \in X$.

The output needs a type reduction to convert into a type 1 fuzzy sets before defuzzification can be performed to obtain a crisp output. This is the main structural difference between type-1 and type-2 logic fuzzy sets. One of the most common type reduction methods is the centroid type-reducer. The centroid of a type-1 fuzzy set when the domain X is discretised into k points is:

$$C_A = \frac{\sum_{i=1}^{k} x_i \mu_A(x_i)}{\sum_{i=1}^{k} \mu_A(x_i)} \tag{5}$$

Referring to the literature [6, 23] the centroid of a type-2 fuzzy set $\tilde{A}$ with domain X discretised into k points $x_1, ...x_k$ with $x_1 < ... < x_k$ as

$$C_{\tilde{A}} = \int_{u_1 \in J_{x_1}} ... \int_{u_k \in J_{x_k}} \left[\mu_{\tilde{A}}(x_1)(u_1) \cdot ... \cdot \mu_{\tilde{A}}(x_k)(u_k)\right] / \frac{\sum_{i=1}^{k} x_i u_i}{\sum_{i=1}^{k} u_i} \tag{6}$$

In case $\tilde{A}$ is interval type-2 logic fuzzy set, then the centroid is the crisp set:

$$C_{\tilde{A}} = \int_{u_1 \in J_{x_1}} ... \int_{u_k \in J_{x_k}} / \frac{\sum_{i=1}^{k} x_i u_i}{\sum_{i=1}^{k} u_i} \tag{7}$$

It has been shown that this iterative procedure can converge in at most K iterations [8]. Once y_l and y_r are available, they can be used to compute the approximate output. Since the reduced type set is an interval fuzzy set of type 1, the fuzzy output value is [17]:

$$y(x) = \frac{y_{max} + y_{min}}{2} \tag{8}$$

The KM type reduction in its simplest form can be summarized as follows in algorithm 1.1.

In Nowicki's work [10] on defuzzification for binary class membership of objects, it can be seen that the result does not require any ordering of $y_{j,k}$ as is done in the KM method.

According to a theorem stated in the literature [15] with a proof, it turns out that for given rough approximations, $\underline{\mu}_{j,k}$ and $\overline{\mu}_{j,k}$ of the binary set $y_{j,k} = 0, 1$ representing by a single rule class membership, where k is the index for rules $k = 1, ..., K$ and j is the index for classes $j = 1, ...J$, the lower and upper

1. Let the consequent values be aranged in the ascending order $y_1 < y_2 < \ldots < y_K$
2. calculate type-1 system output y_0 as an average of y_k weighted by mean membership grades, i.e., $\left(\underline{\mu}_k + \overline{\mu}_k\right)/2$,
3. set the initial values $y_{\min} = y_{\max} = y_0$,
4. for each $k = 1, 2, \ldots, K$, if $y_k > y_{\max}$, then $\overrightarrow{\mu}_k = \overline{\mu}_k$, otherwise $\overrightarrow{\mu}_k = \underline{\mu}_k$,
5. find the closest $y_{\text{next}} = \min\limits_{k=1,\ldots,K} y_k : y_k > y_{\max}$,
6. calculate $y_{\max}$ as an average of y_k weighted by new grades $\overrightarrow{\mu}_k$,
7. if $y_{\max} \le y_{\text{next}}$, continue, else go to step 4,
8. for each $k = 1, 2, \ldots, K$, if $y_k < y_{\min}$, then $\underleftarrow{\mu}_k = \overline{\mu}_k$, otherwise $\underleftarrow{\mu}_k = \underline{\mu}_k$,
9. find the closest $y_{\text{next}} = \max\limits_{k=1,\ldots,K} y_k : y_k < y_{\min}$,
10. calculate $y_{\min}$ as an average of y_k weighted by new grades $\underleftarrow{\mu}_k$,
11. if $y_{\min} \ge y_{\text{next}}$, finish, else go to step 8.

Algorithm 1.1: The KM type reduction

approximations of the object's class membership C_j are given by

$$y_{\min}(j) = \frac{\sum_{k=1}^{K} \underline{\mu}_{j,k} y_{j,k}}{\sum_{k=1}^{K} \underline{\mu}_{j,k} y_{j,k} + \sum_{k=1}^{K} \overline{\mu}_{j,k} \neg y_{j,k}}\ , \tag{9}$$

$$y_{\max}(j) = \frac{\sum_{k=1}^{K} \overline{\mu}_{j,k} y_{j,k}}{\sum_{k=1}^{K} \underline{\mu}_{j,k} \neg y_{j,k} + \sum_{k=1}^{K} \overline{\mu}_{j,k} y_{j,k}}\ . \tag{10}$$

3 Smooth Type Reduction

To characterize the smooth type reduction, assume that the described system has an output value V_p, where $p = 1 \ldots P$. Then its smooth maximum of $\upsilon_1, \ldots, \upsilon_p$ would be a differentiable approximation of maximum of a function with continuous derivatives. In addition, the universal smooth maximum/minimum function is defined as

$$y_\alpha(\upsilon_1, \ldots, \upsilon_P) = \frac{\sum_{p=1}^{P} \upsilon_p e^{\alpha \upsilon_p}}{\sum_{p=1}^{P} e^{\alpha \upsilon_p}} \tag{11}$$

which y_α has the following properties:

1. $y_\alpha \to max$ as $\alpha \to \infty$,
2. $y_\alpha \to min$ as $\alpha \to -\infty$,
3. $y_0 = \frac{\sum_{p=1}^{P} \upsilon_p}{P}$

Notably, this means that the values of $y_{-\infty}$ and y_∞ are the endpoints of the reduced set, respectively $y_{\min}$ and $y_{\max}$. In the search for end points, e.g., y_{max}, only those tuples should be considered that have lower memberships for

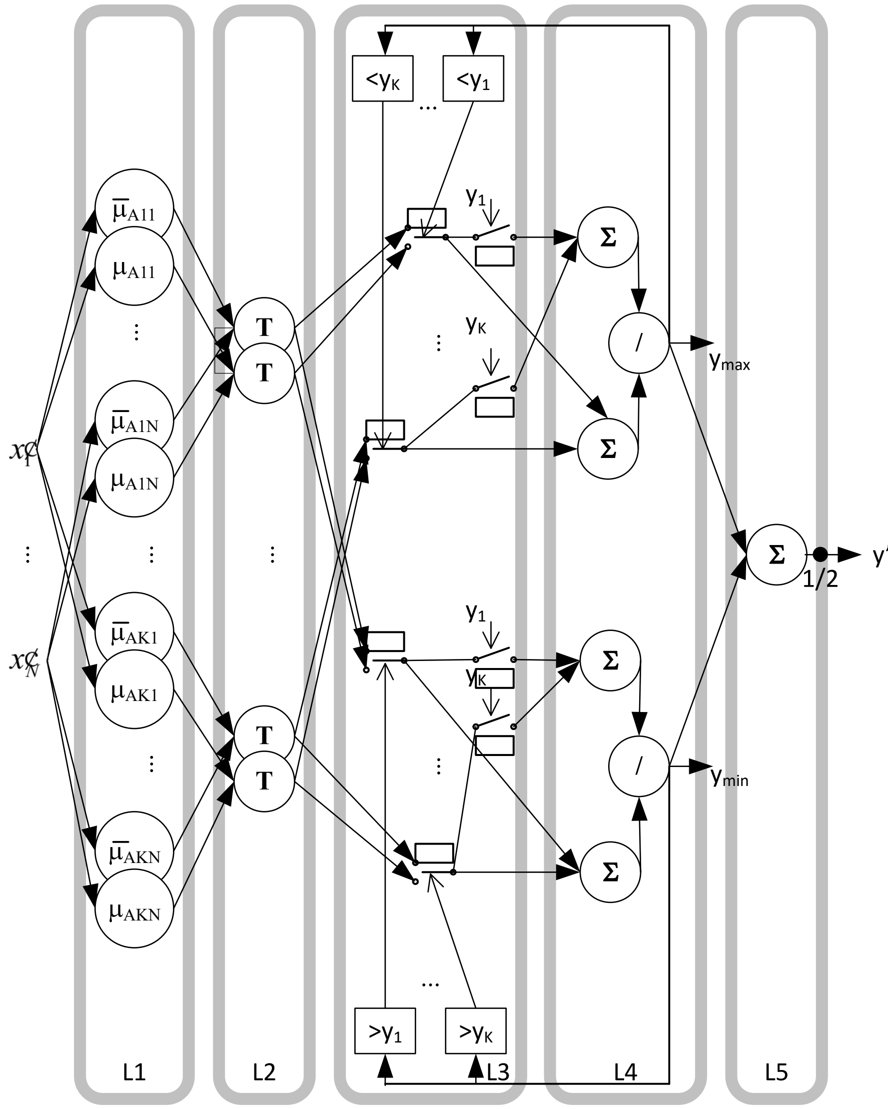

Fig. 2. Adaptive interval type-2 fuzzy logic system using KM type-reduction

consequents being no larger than y_0, which is the output of the type 0 fuzzy system in an interval fuzzy system of type 2. For values of v_k arranged in ascending order, we run the algorithm. Perform a right-shift operation to compute the output values v_p that maximize the result. An example shift is demonstrated in the table 1 and the proposed algorithm using the smooth extremum function is presented in algorithm 1.2.

r\k	1	2	R	R_{+1}	K_{-1}	K
R	0	0	1	1	1	1
R_{+1}	0	0	0	1	1	1
K_{-1}	0	0	0	0	1	1
K	0	0	0	0	0	1

Table 1. Right-shifted mask to calculate y_{max}

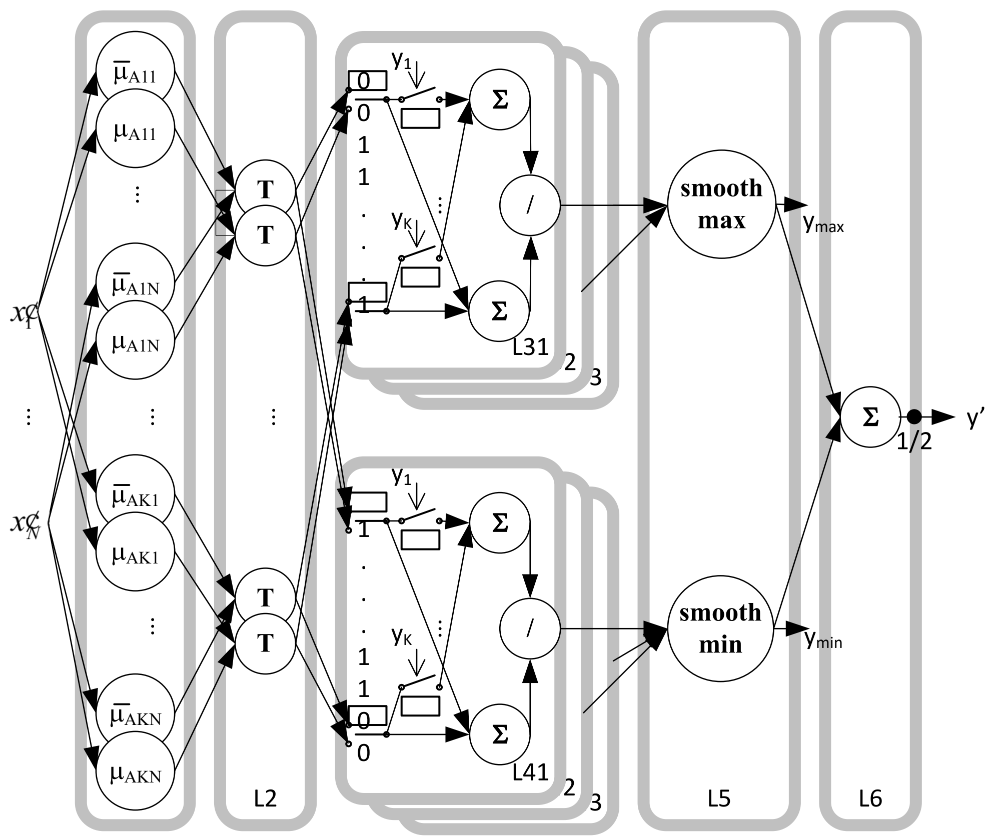

Fig. 3. Adaptive interval type-2 fuzzy logic system using smooth type-reduction

Another approach to smooth maximum is to use LogSumExp, which is as follows: $LSE(\upsilon_1, \ldots, \upsilon_P) = \frac{1}{\alpha} \log \sum_P \exp(\alpha \upsilon_p)$, which can be normalized for all non-negative V_P, yielding a function with domain $[0, \infty)^n$ and range $[0, \infty)$: $g(\upsilon_1, \ldots, \upsilon_P) = \log(\sum_P \exp(\upsilon_p) - (P-1))$. There is also another approach that uses the p-norm, $\|(\upsilon_1, \ldots, \upsilon_R)\|_p = (\sum_r |\upsilon_r|^p)^{\frac{1}{p}}$. The LogSumExp approach as well as the p-Norm approach generate similar results.

Let the consequent values be aranged in the ascending order $y_1 < y_2 < \ldots < y_K$ and the values in vector forms, i.e.,

$$\mathbf{y} = [y_1, \ldots, y_K]$$
$$\overline{\mu} = [\overline{\mu}_1, \ldots, \overline{\mu}_K]$$
$$\underline{\mu} = \left[\underline{\mu}_1, \ldots, \underline{\mu}_K\right]$$

To compute the right and the left endpoints of the type-reduced set, perform the following steps:

1. calculate type-1 system output y_0 as an average of elements of $\mathbf{y}$weighted by mean membership grades, i.e., $\left(\underline{\mu} + \overline{\mu}\right)/2$,
2. find index R of the closest $y_R = \min\limits_{k=1,\ldots,K} y_k : y_k > y_0$,
3. for $r = R, \ldots, K-1$:
   (a) set a mask $M_r = \underset{1\ \ldots}{0\ldots0}\underset{R\ \ldots\ K}{1\ldots1}$,
   (b) apply the mask to upper and lower memberships $\overrightarrow{\mu} = (1 - M_r) \odot \underline{\mu} + M_r \odot \overline{\mu}$ (where $\odot$ is the Hadamard product),
   (c) calculate $y_{\max,r}$ as an average of elements $\mathbf{y}$ weighted by $\overrightarrow{\mu}$,
4. return $y_{\max}$ as an aggregation of all $y_{\max,r}$ with the use of smooth maximum, $r = R, \ldots, K-1$,
5. find index L of the closest $y_L = \min\limits_{k=1,\ldots,K} y_k : y_k < y_0$,
6. for $l = 2, \ldots, L$:
   (a) set a mask $M_l = \underset{1\ \ldots\ L}{1\ldots1}\underset{\ldots\ K}{0\ldots0}$,
   (b) apply the mask to upper and lower memberships $\overleftarrow{\mu} = (1 - M_l) \odot \underline{\mu} + M_l \odot \overline{\mu}$,
   (c) calculate $y_{\min,l}$ as an average of elements $\mathbf{y}$ weighted by $\overleftarrow{\mu}$,
7. return $y_{\max}$ as an aggregation of all $y_{\max,r}$ with the use of smooth maximum, $r = R, \ldots, K-1$.

Algorithm 1.2: Smooth type reduction

4 Experimental Results

The source Wisconsin Breast Cancer data are reports of clinical cases [Mangasarian and Wolberg 1990] [18]. The original data set contained 699 cases divided into two categories: benign breast cancer (65.5% of instances) and malignant cancer (34.5%). Each case was described by nine attributes: clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, and mitosis, note that 16 individuals are missing the attribute.

The specificity of interval-valued fuzzy logic systems allows us for an analysis on a lower level of classification if only we make use of the interval outputs of the system: ymin and ymax. Using this information, instead of strict classification, we get three groups of objects classified with the following labels:

- certain classification if $y_{min} > 0.5$,

– uncertain classification if $y_{max} \geq 0.5 \geq y_{min}$,
– certain rejection if $y_{max} < 0.5$.

As a result, we get three rate groups: classified, misclassified, and unclassified ("NoClass.") when classification cannot be performed certainly. This can help in practical classification systems such as the medical diagnosis when uncertain classification cases can be again directed to a thorough examination. The classification results in the imputation of input values by means of rough-fuzzy sets are presented in Table 2

original data	Singleton Class./Misclass.	Interrval KM-T2FLC Class./NoClass./Misclass.	Interval T2FLC based Smooth Type-Reduction Class./NoClass./Misclass.
	0.988/0.012	0.975/0.10/0.015	**0.986/0.011/0.003**
σ_1			
1.0	0.978/0.022	0.964/0.019/0.017	**0.974/0.020/0.006**
5.0	0.931/0.069	0.673/0.315/0.012	**0.675/0.317/0.008**
σ_2			
1.0	0.977/0.023	0.963/0.016/0.020	**0.973/0.018/0.009**
5.0	0.960/0.040	0.647/0.336/0.017	**0.660/0.333/0.007**
σ_3			
1.0	0.970/0.030	0.912/0.074/0.014	**0.933/0.062/0.005**
5.0	0.911/0.089	0.589/0.406/0.005	**0.694/0.302/0.004**
σ_4			
1.0	0.977/0.023	0.975/0.009/0.016	**0.985/0.009/0.006**
5.0	0.967/0.033	0.838/0.152/0.010	**0.844/0.150/0.006**
σ_5			
1.0	0.977/0.023	0.970/0.011/0.019	**0.981/0.010/0.009**
5.0	0.962/0.038	0.795/0.195/0.010	**0.799/0.191/0.010**
σ_6			
1.0	0.978/0.022	0.948/0.034/0.019	**0.961/0.032/0.008**
5.0	0.938/0.062	0.824/0.166/0.010	**0.825/0.167/0.008**
σ_7			
1.0	0.980/0.020	0.961/0.024/0.015	**0.970/0.025/0.005**
5.0	0.965/0.035	0.634/0.360/0.006	**0.642/0.352/0.006**
σ_8			
1.0	0.978/0.022	0.968/0.012/0.020	**0.978/0.011/0.009**
5.0	0.970/0.030	0.854/0.137/0.009	**0.853/0.139/0.008**
σ_9			
1.0	0.980/0.020	0.969/0.014/0.018	**0.978/0.015/0.008**
5.0	0.944/0.056	0.717/0.272/0.011	**0.722/0.271/0.007**
σ_{all}			
1.0	0.973/0.027	0.518/0.480/0.002	**0.538/0.462/0.000**
5.0	0.749/0.251	0.001/0.999/0.000	**0.021/0.979/0.000**

Table 2. Wisconsin Breast Cancer classification with optional uniform noise applied to single input $X_1, \ldots, X_9$ as well as to all inputs X_{all}

5 Conclusion

In this paper, a smooth type-reduction method that is competitive with the KM type-reduction system is presented. It shows good results as it achieves low training error values. It is worth noting that both type-2 fuzzy systems significantly exceed the learning ability of the type-1 fuzzy system. The proposed system is worth considering for solving problems with increased model uncertainty or when there is uncertain input data.

The initial learning of type 2 systems treated as type 1 fuzzy systems, followed by the application of generating type-2 fuzzy rules methods for uncertain data using the fuzzy-rough approximation [13, 16] or possibilistic fuzzification [15], shows that fuzzy systems are important in the process of extracting explanatory fuzzy rules.

References

1. Bilski, J., Kowalczyk, B., Marchlewska, A., Zurada, J.: Local Levenberg-Marquardt algorithm for learning feedforward neural networks. Journal of Artificial Intelligence and Soft Computing Research **10**(4) (2020) 299–316
2. Bilski, J., Rutkowski, L.and Smoląg, J., Tao, D.: A novel method for speed training acceleration of recurrent neural networks. Information Sciences **553** (10 2020) 266–279
3. Chen, Y., Wang, D.: Study on centroid type-reduction of general type-2 fuzzy logic systems with weighted enhanced karnik–mendel algorithms. Soft Computing **22**(4) (2018) 1361–1380
4. Greenfield, S., Chiclana, F.: Type-reduction of the discretised interval type-2 fuzzy set: approaching the continuous case through progressively finer discretisation. Journal of Artificial Intelligence and Soft Computing Research **1**(3) (2011) 183–193
5. Greenfield, S., Chiclana, F., Coupland, S., John, R.: The collapsing method of defuzzification for discretised interval type-2 fuzzy sets. Information Sciences **179** (06 2009) 2055–2069
6. Karnik, N., Mendel, J.: Centroid of a type-2 fuzzy set. Inf. Sci. **132** (02 2001) 195–220
7. Liang, Q., Mendel, J.: Interval type-2 fuzzy logic systems: Theory and design. Fuzzy Systems, IEEE Transactions on **8** (11 2000) 535–550
8. Mendel, J.: Uncertain rule-based fuzzy logic systems. Introduction and New Directions (2001)
9. Mendel, J., John, R.: Type-2 fuzzy sets made simple. IEEE Transactions on Fuzzy Systems **10**(2) (2002) 117–127
10. Nowicki, R.: On combining neuro-fuzzy architectures with the rough set theory to solve classification problems with incomplete data. Knowledge and Data Engineering, IEEE Transactions on **20** (10 2008) 1239–1253
11. Nowicki, R., Seliga, R., Żelasko, D., Hayashi, Y.: Performance analysis of rough set–based hybrid classification systems in the case of missing values. Journal of Artificial Intelligence and Soft Computing Research **11**(4) (2021) 307–318
12. Nowicki, R., Starczewski, J.: On non-singleton fuzzification with DCOG defuzzification. In Rutkowski, L., et al., eds.: Artificial Intelligence and Soft Computing

— ICAISC 2010, Part I. Volume 6113 of Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin Heidelberg (2010) 168–174
13. Nowicki, R., Starczewski, J.: A new method for classification of imprecise data using fuzzy rough fuzzification. Inf. Sci. **414** (2017) 33–52
14. Runkler, T., Coupland, S., John, R.: Properties of interval type–2 defuzzification operators. (07 2015)
15. Starczewski, J.: Advanced Concepts in Fuzzy Logic and Systems with Membership Uncertainty. Volume 284 of Studies in Fuzziness and Soft Computing. Springer (2013)
16. Starczewski, J., Nowicki, R., Nieszporek, K.: Fuzzy-rough fuzzification in general FL classifiers. In Guervós, J.J.M., Garibaldi, J., Linares-Barranco, A., Madani, K., Warwick, K., eds.: Proceedings of the 11th International Joint Conference on Computational Intelligence, IJCCI 2019, Vienna, Austria, September 17-19, 2019, ScitePress (2019) 335–342
17. Tan, W.W., Wu, D. In: Design of Type-Reduction Strategies for Type-2 Fuzzy Logic Systems using Genetic Algorithms. Springer Berlin Heidelberg, Berlin, Heidelberg (2007) 169–187
18. Wolberg, W.H.: Breast Cancer Wisconsin. University of Wisconsin Hospitals Madison, Wisconsin, USA. (1990)
19. Wu, D., Mendel, J.: Recommendations on designing practical interval type-2 fuzzy systems. Engineering Applications of Artificial Intelligence **85** (2019) 182–193
20. Wu, L., Qian, F., Wang, L., Ma, X.: An improved type-reduction algorithm for general type-2 fuzzy sets. Information Sciences **593** (2022) 99–120
21. Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning. Information Sciences **8** (1975) 199–249
22. Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning-II. Information Sciences **8** (1975) 301–357
23. Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning-III. Information Sciences **9**(1) (1975) 43–80