TELECOMMUNICATIONS AND RADIO ENGINEERING

UDC 629.3.025.2 (045)
DOI:10.18372/1990-5548.71.16825

N. V. Shapoval

TSK FUZZY NEURAL NETWORK FOR COVID-19 CLASSIFICATION

National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mail: shovgun@gmail.com, ORCID 0000-0002-8509-6886

Abstract**—*It is considered t the Takagi-Sugeno-Kang fuzzy neural network and its modern variations. The use of regularization, random exclusion of rules from the rule base allows solving the problem of excessive similarity of rules in the rule base. The use of batch normalization to increase the generalizing properties of the network allows to increase the accuracy of the model, while maintaining the possibility of interpreting the results, which is characteristic of fuzzy neural networks. It is proposed to use an ensemble of fuzzy neural networks to increase the generalizing capabilities of the network. Studies of the Takagi-Sugeno-Kang fuzzy neural network for the task of diagnosing the coronavirus disease show that the proposed model works well and allows to improve the result.*

Index Terms—Takagi–Sugeno–Kang fuzzy system; fuzzy neural networks ensemble; batch normalization.

I. INTRODUCTION

The Takagi–Sugeno–Kang (TSK) fuzzy system is widely known and is used for classification and regression problems in many areas – creditworthiness assessment, macroeconomic forecasting, stock market forecasting, assessment of the technical condition of construction, and others [1], [2], [3]. Since the TSK fuzzy system can be represented as a neural network with five layers, it is also known as the TSK fuzzy neural network. TSK fuzzy neural network parameters are adjusted using evolutionary algorithms, or gradient descent, or hybrid algorithms. Evolutionary algorithms require the support of a large population, can converge to different solutions and, accordingly, require high computational costs. Gradient descent requires the computation of the gradient over the entire data set to iteratively update the parameters. This can be quite slow when we have a large number of parameters and a large amount of data. Traditional approaches to solving the problem of too many parameters include reducing the number of features and adjusting the number of rules. Reducing the number of features is possible using traditional data analysis methods [4]. The number of rules can be adjusted: 1) by gradually increasing the rule base; 2) by gradually removing rules from the rule base by evaluating the quality of the fuzzy network; 3) by focusing on the firing level of the rule when adding a rule to the base [5]. Modern researchers, inspired by the successes in the field of deep neural networks, apply deep learning approaches in fuzzy neural networks as well. Batch normalization (BN), regularization, layer normalization (LN), random exclusion of a rule from the rule base during network training – all these techniques for training deep neural networks can be applied to the TSK fuzzy neural network. An important advantage of fuzzy neural networks is the possibility of using a trained rule base to interpret the result, which is especially important in a field such as medicine.

II. PROBLEM STATEMENT

A. Traditional TSK fuzzy neural network

Let the training dataset be $D = \{x_n, y_n\}_{n=1}^{N}$, where $x_n = \left[x_{n,1},\dots,y_{n,D}\right]^T \in R^{D_{x1}}$ is a D-dimensional feature vector and $y_n \in \{1,2,\dots,C\}$ is a class label for a problem with C classes or $y_n \in R$ for regression problems. Let our TSK fuzzy system have R rules. Each rule can be represented as:

$$\text{Rule}_r : IFx_1 \text{ is } A_{r,1} \text{ and ... and } IFx_D \text{ is } A_{r,D}$$

$$\text{THEN } y_r(x) = b_{r,0} + \sum_{d=1}^{D} b_{r,d} x_{d'}$$

where $A_{r,d}$ is a linguistic variable specifying the d-feature of an r-rule. $b_{r,d}$ is a linear parameter of the network for the d-feature of the r-rule. A Gaussian membership function is most often used to specify a linguistic variable, which has the form:

$$\mu_{A_{r,d}}\left(x_d\right) = \exp\left(-\frac{\left(x_d - m_{r,d}\right)}{2\sigma_{r,d}^2}\right),$$

where $m_{r,d}$ and $\sigma_{r,d}$ are the center and standard deviation of the Gaussian membership function of the linguistic variable $A_{r,d}$. The final output of the network will be

$$y(x) = \frac{\sum_{r=1}^{R} f_r(x) y_r(x)}{\sum_{r=1}^{R} f_i(x)},$$

where

$$f_r(x) = \prod_{d=1}^{D} \mu_{A_{r,d}}(x_d) = \exp\left(-\sum_{d=1}^{D} \frac{\left(x_d - m_{r,d}\right)^2}{2\sigma_{r,d}^2}\right)$$

the firing level of the rule r. The structure of such a fuzzy neural network is shown in Fig. 1.

The network is trained by a gradient descent algorithm. For classification problems, the cross-entropy loss function is taken. To improve the classification accuracy, it is necessary to overcome the traditional problems of gradient descent.

B. Application of deep neural network approaches to TSK fuzzy network

Overfitting is a well-known problem in machine learning. In such case, a model performs well on training data and performs badly on new data. This problem is solved by regularization – the loss function is adjusted for greater generalization of the model. In fuzzy neural networks, there is a problem when one rule is most often activated and it makes the biggest contribution to the response, as a result we will have a rule base with a large number of identical rules.

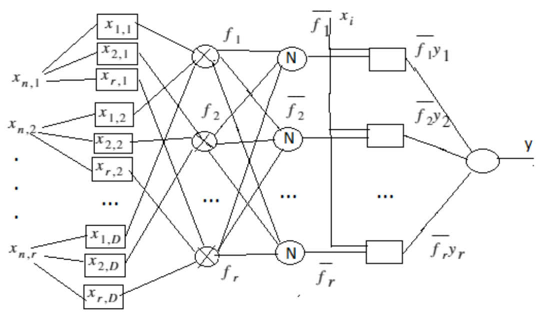

Fig. 1. The structure of the TSK fuzzy neural network

It is possible to check the similarity of the rules in the process of training the network, but this is an additional computational cost. In order for the rules in the rule base to have approximately the same firing level of the rule and make the same contribution to the response, the authors [6] suggested using regularization. To force all rules to have the same firing level of the rules, a uniform regularization (UR) is added to the loss function in gradient descent:

$$L_{UR} = \sum_{r=1}^{R} f_i(x) \left(\frac{1}{N} \sum_{n=1}^{N} \left(\overline{f}_r(x_n) - \tau \right)^2 \right),$$

where N is the number of training samples; $\tau \in [0,1]$ is the expected level of firing level of each rule, for C classes it is set to $1/C$.

Another approach to the regularization of deep neural networks is dropout. It is a regularization technique for reducing over fitting neural networks by randomly turning off neurons during training. In this way, several networks of different architectures are trained in parallel. This approach is also applicable to fuzzy neural networks [7]. It is called Drop Rule and involves randomly dropping a rule from the rule base, which can improve the performance of the TSK model.

Batch normalization (BN) is designed to speed up and stabilize the learning process of a deep neural network by normalizing the outputs of neurons in hidden layers. This technique can also be applied to TSK fuzzy neural networks [6]. At the training stage, the firing levels of the rules are first calculated, and then the BN is used to normalize the inputs, according to their mean and standard deviation in the current mini-batch. The normalized inputs are then used to compute the rule consequents. After training in the BN layer, we have the mean value $m = (m_1,...,m_D)^T$ and standard deviation $\sigma = (\sigma_1,...,\sigma_D)^T$, γ and β is parameters that are determined during training, e is set to 1e-8 to avoid division by zero. Then the output of the r-rule will be:

$$y_r\left(BN\left(x_n\right)\right) = b_{r,0} + \gamma \sum_{d=1}^{D} b_{r,d} \frac{x_{n,d} - m_d}{\sqrt{\sigma_d^2 + e}} + \beta D.$$

The structure of such a TSK fuzzy neural network is shown in Fig. 2.

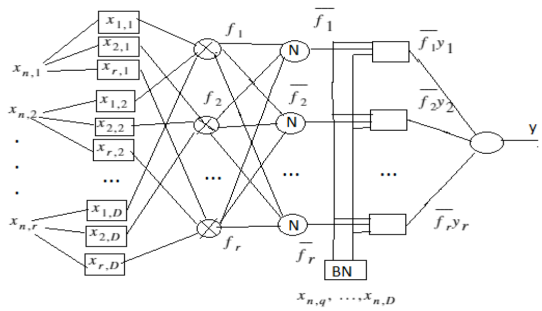

Fig. 2. The structure of the TSK fuzzy neural network with a BN layer

The use of the above approaches allows to increase the performance of fuzzy neural networks while maintaining the possibility of interpreting the result.

III. Problem Solution

A. *Ensembles of neural networks*

The statistical nature of learning neural networks leads to the fact that they have low bias and high dispersion. Each time during training, the network learns a slightly different input-to-output mapping function. Ensembles are a powerful approach in machine learning, when combining several different models can improve the accuracy of the overall result. This approach is used to build models for solving problems in many areas [8], [9]. Combining several neural networks into an ensemble allows to get a model with less variance. Combining multiple networks with different training parameters, different layers, etc. that learn a more heterogeneous set of mapping features reduces the variance and reduces the error of the generalized model.

B. *Experimental results*

Different configurations of TSK fuzzy neural networks were trained on 500 epochs to sample data on symptoms of the coronavirus disease [10]. The assessment of the accuracy of networks during training is shown in Figs 3 and 4. Comparison of the precision, recall and f1-score metrics for different configurations of the TSK fuzzy neural network are shown in Table I.

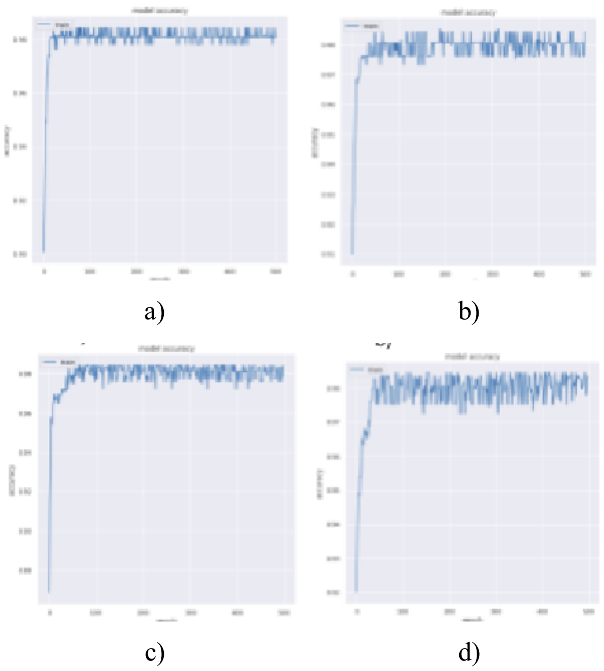

Fig. 3. Accuracy score. (a) TSK+BN; (b) TSK+UR+DropRule; (c) TSK+BN+UR+DropRule; (d) TSK+BN+DropRule

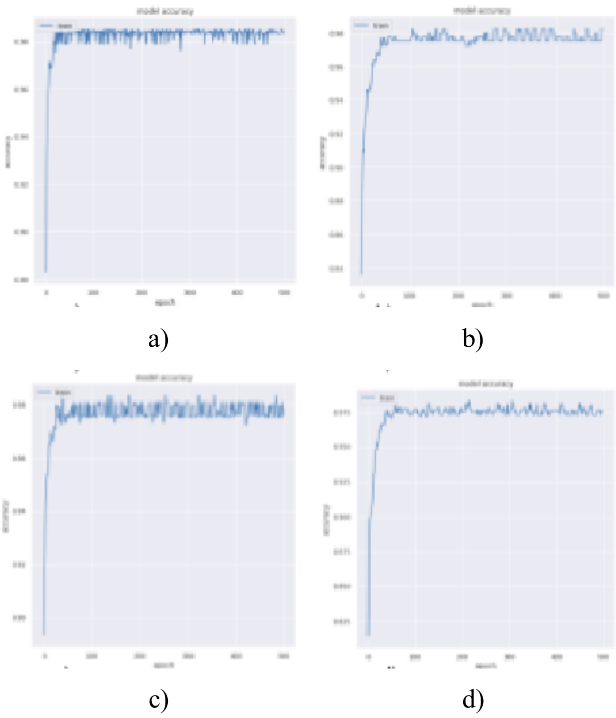

Fig. 4. Accuracy score. (a) TSK+BN+UR; (b) TSK+DropRule; (c) TSK+BN; (d) TSK

Table I. Comparison of the Results of Different Versions of the TSK Fuzzy Neural Network

	Precision	**Recall**	f_1
TSK	0.9774	0.9919	0.9846
TSK+DropRule	0.9885	0.9896	0.9891
TSK+BN	0.9817	0.9885	0.9851
TSK+UR	0.9896	0.9862	0.9879
TSK+BN+UR	0.993	0.9862	0.9896
TSK+BN+DropRule	0.993	0.9862	0.9896
TSK+UR+DropRule	0.9953	0.985	0.9902
TSK+UR+DropRule +BN	0.9953	0.985	0.9902

The best performance was achieved by the fuzzy TSK network with regularization. It can be explained by the presence of a diverse rule base. The best 5 TSK fuzzy neural network configurations were selected for the neural network ensemble. To increase the accuracy, not only the outputs of the basic models, but also the input data were taken as input data for the Result Aggregation block. The structure of the ensemble is shown in Fig. 5.

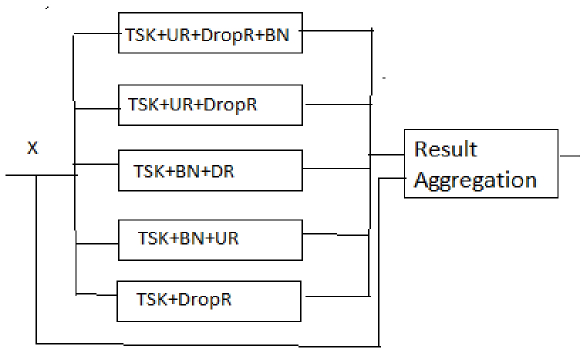

Fig. 5. The structure of the ensemble of fuzzy neural networks

As a result of using the ensemble, we will get an estimate of classification accuracy of 98.56%.

IV. Conclusion

It is proposed to use an ensemble of fuzzy neural networks using modern approaches in the field of deep neural networks. BN and uniform regularization have shown their effectiveness. Ensemble was applied to a medical dataset of symptoms of the coronavirus disease. The proposed ensemble increased the accuracy of problem solving.

References

[1] I. O. Kravets, and A. Yu. Timoshyn, "Exploring fuzzy neural networks for forecasting rapidly changing non-stationary time series," *Scientific works* [of the Petro Mohyla Black Sea State University]. Ser.: Computer technologies 191, vol. 179, 2012, pp. 72–79. [in Ukrainian]

[2] Natalia Shovgun, "Fuzzy Neural Networks for Evaluating the Creditworthiness of the Borrowers," *Information Theories & Applications*, vol. 21, pp. 54–59, 2014.

[3] O. Komandyrov, P. Kulikov, V. Ploskyi, and B. Yeremenko, "Application of the Takaga–Sugeno–Kanga Artificial Neuro-fussy Network to Assessment of the Technical Condition of Building Objects," *Management of the development of complex systems*, vol. 42, pp. 107–112, 2020. [in Ukrainian]

[4] M. Nilashi, O. bin Ibrahim, N. Ithnin, and N. H. Sarmin, "A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS," *Electronic Commerce Research and Applications*, vol. 14(6), pp. 542–562, 2015.

[5] N. V. Shovgun, "Increasing the effectiveness of fuzzy neural networks in the task of assessing the borrower's creditworthiness," *Bulletin of NTUU "KPI". Informatics, management and computer technology*, vol. 58, pp. 89–94, 2013. [in Ukrainian]

[6] Y. Cui, D. Wu, & J. Huang, "Optimize TSK fuzzy systems for classification problems: Minibatch gradient descent with uniform regularization and batch normalization," *IEEE Transactions on Fuzzy Systems*, vol. 28(12), 2020, pp. 3065–3075.

[7] D. Wu, Y. Yuan, J. Huang, & Y. Tan, "Optimize TSK fuzzy systems for regression problems: Minibatch gradient descent with regularization, DropRule, and AdaBound (MBGD-RDA)," *IEEE Transactions on Fuzzy Systems*, vol. 28(5), 2019, pp. 1003–1015.

[8] V. Sineglazov, & J. Smirnova, "Intelligence Diagnostics of Heart Disease Based on Neural Networks Ensemble Use," *Electronics and Control Systems*, vol. 3(69), 2021, pp. 9–15.

[9] S. Zhang, M. Liu, & J. Yan, "The diversified ensemble neural network," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 16001–16011.

[10] https://www.kaggle.com/code/midouazerty/symptoms-covid-19-using-7-machine-learning-98/data (accessed 15.07.22)

Received December 16, 2021

Shapoval Nataliia. ORCID 0000-0002-8509-6886. Candidate of Science (Engineering).
National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Politechnic Institute, Kyiv, Ukraine, (2010).
Research interests: computer vision, fuzzy neural network, deep neural network.
Publications: 5
E-mail: shovgun@gmail.com

Н. В. Шаповал. Використання нечіткої нейронної мережі ТСК для діагностування хворих на COVID-19

В статті розглядається нечітка нейронна мережа ТСК та сучасні її варіації. Використання регуляризації, випадкового виключення правил з бази правил дозволяє вирішити проблему завеликої схожості правил в базі правил. Використання пакетної нормалізації для збільшення узагальнюючих властивостей мережі дозволяє підвищити точність моделі, зберігаючи при цьому можливість інтерпретації результатів, яка властива нечітким нейронним мережам. Запропоновано використання ансамблю нечітких нейронних мереж для підвищення узагальнюючих можливостей мережі. Дослідження з великими наборами даних як для задачі діагностики хвороби серця так і для задачі діагностики коронавірусної хвороби показують, що запропонована модель добре працює і дозволяє покращити результат.

Ключові слова: нечітка система ТСК; ансамбль нечітких нейронних мереж; пакетна нормалізація.

Шаповал Наталія Віталіївна. ORCID 0000-0002-8509-6886. Кандидат технічних наук.
Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (2010).
Напрям наукової діяльності: комп'ютерний зір, нечіткі нейронні мережі, глибокі нейронні мережі.
Кількість публікацій: 5.
E-mail: shovgun@gmail.com

Н. В. Шаповал. Использование нечеткой нейронной сети ТСК для диагностики больных COVID-19

В статье рассматривается нечеткая нейронная сеть ТСК и ее современные вариации. Использование регуляризации, случайного исключения правил из базы правил позволяет решить проблему большого сходства правил в базе правил. Использование пакетной нормализации для увеличения обобщающих свойств сети позволяет повысить точность модели, сохраняя при этом возможность интерпретации результатов, свойственных нечетким нейронным сетям. Предложено использование ансамбля нечетких нейронных сетей для повышения обобщающих возможностей сети. Исследования с большими наборами данных как для задачи диагностики болезни сердца, так и для задачи диагностики коронавирусной болезни показывают, что предложенная модель хорошо работает и позволяет улучшить результат.
Ключевые слова: нечеткая система ТСК; ансамбль нечетких нейронных сетей; пакетная нормализация.

Шаповал Наталия Витальевна. ORCID 0000-0002-8509-6886. Кандидат технических наук.
Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина
Образование: Киевский политехнический институт, Киев, Украина, (2010).
Направление научной деятельности: компьютерное зрение, нечеткие нейронные сети, глубокие нейронные сети
Количество публикаций: 5.
E-mail: shovgun@gmail.com