HIGH-PERFORMANCE ADAPTIVE NEURO-FUZZY CLASSIFIER WITH A PARAMETRIC TUNING

Sergey Gorbachev [a], *Vladimir* Syryamkin

National Research Tomsk State University, 634050, Tomsk, Russia

Abstract. The article is devoted to research and development of adaptive algorithms for neuro-fuzzy inference when solving multicriteria problems connected with analysis of expert (foresight) data to identify technological breakthroughs and strategic perspectives of scientific, technological and innovative development. The article describes the optimized structural-functional scheme of the high-performance adaptive neuro-fuzzy classifier with a logical output, which has such specific features as a block of decision tree-based fuzzy rules and a hybrid algorithm for neural network adaptation of parameters based on the error back-propagation to the root of the decision tree.

1 Introduction

Foresight research is becoming a widespread tool for solving multicriteria problems of expert data analysis that allow revealing technological breakthroughs and strategic perspectives of scientific, technological and innovative development [1]. Foresight as a system of methods is being constantly developed and improved, and for the last twenty years, considerable experience of their practical application has been accumulated [2]. The analysis of publications shows that a combination of qualitative and quantitative methods is effective [3]. Moreover, constant complication of a system of methods requires development of intellectual approaches that ensure scientifically based transparent conclusions.

Thus, when evaluating the effectiveness of technologies, traditional statistical methods have their drawbacks, especially when two or more classes are connected by a chain of internally connected sampling objects, have a non-spherical form, are linearly inseparable sets, or when the density or volume of classes is different [4–6]. When classes intersect, contact or chain, the classification model is blurred, so the choice of mathematical methods for solving the problem of automatic classification is preceded by fuzzy procedures [7]. The theory of fuzzy sets operates with such concepts as a weighted membership of the object under investigation to the set and it offers a rather flexible apparatus for a formal description of similar situations. Therefore, the goal set transfers from question "Does the object belong to the class?" to identifying the degree of membership of an object to a

[a] Corresponding author: spp03@sibmail.com

particular class. Summarizing, we can say that a fuzzy approach to the solution of classification problems in certain cases allows us to separate clusters of complex shape, thereby opening up new possibilities for interpreting the results of classification.

Comparing the reliability of the results of classification obtained by different authors allows us to single out the most effective classifiers, which include neural networks [4, 8] and methods of cluster analysis [9, 10]. The best results were obtained using a three-layer perceptron and a probabilistic neural network, as well as Kohonen self-organizing maps. At the same time, in the studies known to us, classification of the effectiveness of scientific and technological solutions and technologies has not been considered, and a neural-fuzzy network has not been used to classify scientific projects.

The purpose of this study is to develop high-performance adaptive algorithms for neural-fuzzy inference when solving multicriteria problems of expert analysis (foresight), with high classification accuracy and logical inference based on a compact set of the classification rules.

2 A method for constructing an adaptive neural-fuzzy classifier

The construction of classical adaptive neuro fuzzy inference systems (ANFIS) is associated with formalization of input parameters and target indicators of the effectiveness of scientific (innovative) projects and technologies in the form of a vector of interval values (fuzzy interval); the hit in each interval of this vector is characterized by some degree of uncertainty [11]. This procedure is called “fuzzification”. Based on primary data, experience, and intuition experts usually do not have difficulties in making quantitative evaluation of the boundaries (intervals) of possible (permissible) values of parameters and the range of their most possible (preferred) values. For example, experts can characterize the primary data using triangular fuzzy numbers with the following membership function $\mu A\ (x)$ of the fuzzy set A:

$$\mu_A(x)=\begin{cases}0,\ if\ x\le a_{\min}\\ \dfrac{x-a_{\min}}{c-a_{\min}},if\ a_{\min}<x<c\\ \dfrac{a_{\max}-x}{a_{\max}-c},if\ c\le x<a_{\max}\\ 0,\ if\ x\ge a_{\max}\end{cases} \qquad (1)$$

This expression simulates the following expert position: “parameter A is approximately equal to c and uniquely enters the segment $[a_{min},\ a_{max}]$”.

In practice, an analytical representation of the membership function is often used [7], for example, the Gaussian membership function has the form:

$$\mu_A(x)=e^{-\frac{(x-c)^2}{2\sigma^2}}\ , \qquad (2)$$

where c is a center and σ is a standard deviation (width).

To form the base of fuzzy rules, the authors proposes to use the fuzzy decision tree [12, 13], which can be considered as a hierarchical network with direct signal propagation. Therefore, we propose to use a hybrid algorithm for back propagation of the error to the root node of the tree in order to adapt its parameters (centers and width of membership functions and parameters of rules consequents on tree leaves).

3 Results and Discussion

Let us consider a set of data *D* describing multidimensional feature objects (innovative technologies), where in each line there are *m* numerical values for the attributes $x_1,...,x_n$ and one class C_l among the set $C = \{C1, C2\}$.

Let x_k j be the value of the *k*-th input variable of the *j*-th observation;

C_j is a value of output variable of j-th observation;

N is a number of observations;

n is a number of input variables.

3.1 Initialization of fuzzy model (stage *1*)

Based on data fuzzification, we set the initial parameters of the term membership functions and construct a fuzzy solution tree whose leaves have output parameters r_{i1} and r_{i2} for each *i*-th branch. The structure of such a tree sets the basis for fuzzy rules such as a singleton with a complexly constructed consequent (conclusion) of the following form:

$$Rule\ i\text{: } IF\ x_1 = A_{1i}\ and\ x_2 = A_{2i}\ and\ ...\ and\ x_n = A_{ni}\ THEN\ y_1 = r_{i1}\ and\ y_2 = r_{i2}, \quad (3)$$

where A_{ki} is the linguistic term, by which the *k*-th input variable is estimated.

Figure 1 shows the structural-functional diagram of the constructed adaptive neural-fuzzy classifier with a logical conclusion that implements the mapping $F:\Re^5 \to \Re^2$. The module solves the problem of classifying multidimensional feature objects based on the construction of a fuzzy decision tree (stage *1*) and fine-tuning the parameters of membership functions and rules consequents with hybrid learning algorithms (stage *2*). It consists of *7* layers with the following basic functions of the nodes.

Layer 1 (input).

As inputs, indicators are considered, according to which experts evaluated innovative technologies, for example, the degree of technology influence on the solution of global problems that arose in the framework of the most acute global crises (ecological, food, geopolitical, and energy).

Nodes in the *1*-st layer accurately convey the input values of the attributes in the *2*nd layer.

Layer 2 (fuzzification).

Conversion of numerical values of attributes into linguistic terms in order to reduce information and present it in an understandable form for an expert. Each *k*-th input variable is estimated by the linguistic term A_{ki} .

This layer simulates the above-described membership functions.

Layer 3 (aggregation) determines the degree of truth conditions for each of the fuzzy rules, it is a non-adaptive layer of *AND*-neurons that simulate the logical connection *AND* by the multiplication of the membership functions:

$$\omega_{ij} = \prod_k \mu_{ki}(x_{kj}) = \prod_k \exp\left(-\frac{(x_{kj} - c_{ki})^2}{2\sigma_{ki}^2}\right),\ i=1,...,m, \quad (4)$$

where *i* is a rule number corresponding to the branch of the decision tree, in this case *i=1, ..., 10,*

m is a number of fuzzy rules.

The multiplication is computed over all the variables x_k on the *i*-th branch. ω_{ij} can be interpreted as a weight, or a degree of truth, of the premises of the *i*-th rule (the branch of the tree) in the *j*-th example.

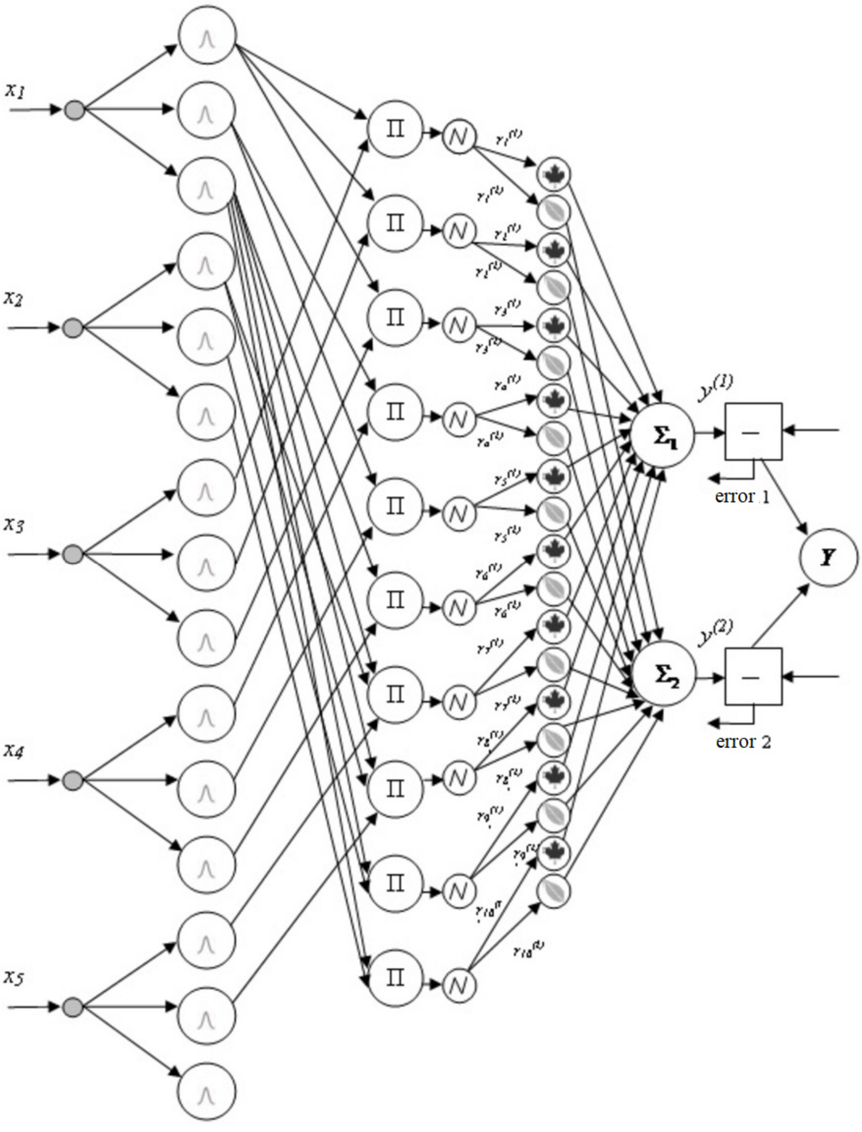

Fig. 1. Structural-functional scheme of the adaptive neural-fuzzy classifier based on a fuzzy decision tree.

Layer 4. Normalization of the degree of rules execution.

The nodes of this layer are nonadaptive; they calculate the relative degree (weight) of the fuzzy rule execution among all the rules in the *j*-th example using the formula:

$$\overline{\omega_{ij}} = \frac{\omega_{ij}}{\sum_{k=1}^{m} \omega_{kj}} \tag{5}$$

Layer 5 (activation). On the leaves of the tree, it determines the degree of truth of the consequents (conclusions) of the rules for each class using weight parameters r_i^1 and r_i^2 which are initialized, for example, by the method of finding the nearest neighbor.

Note that the procedure for initializing the values of the membership parameters (centers and widths), rules antecedents and rules consequents is preliminary, and they are subject to fine-tuning at the next stage.

For each class, the adaptive nodes of layer 5 calculate the degree of truth of the rule conclusions when passing the *j*-th signal:

$$y_{ij}^{1} = \overline{\omega_{ij}} r_{i}^{1},\; y_{ij}^{2} = \overline{\omega_{ij}} r_{i}^{2},\; i=1,\dots,m. \tag{6}$$

Layer 6 (accumulation) combines all the degrees of truth of the conclusions of each rule to obtain the membership function of the output variable. The nonadaptive nodes of this layer for each class form the output value of the variables y_j^1 and y_j^2, representing the degree of truth of assigning the *j*-th signal to the class as a sum of the degrees of truth of all rules:

$$y_j^1 = \sum_{i=1}^{m} y_{ij}^1,\; y_j^2 = \sum_{i=1}^{m} y_{ij}^2. \tag{7}$$

Layer 7 (defuzzification). After fine-tuning the parameters (stage 2), this layer calculates the total output Y_j and the forecast class l_j of the *j*-th signal by the method of the right modal value (based on the maximum degree of truth):

network output $Y_j = \max(y_j^1, y_j^2)$,

forecast class $l_j = \arg\max_{l=1,2} \{y_j^l\}$.

The accuracy of fuzzy inference is achieved by the joint work of rules pertaining to one class.

3.2 Adaptive parametric tuning of the neuro-fuzzy classifier (stage 2)

Let us consider the constructed network of fuzzy inference (Figure 1). It has 5 input variables, each of which is fuzzy in 3 linguistic terms. The rule base is formed by the weighted branches of the decision tree that connects layers 2 and 3. It is obvious that exactly one rule corresponds to each node of the layer 3. Thus, the rule base includes 10 rules.

We describe a gradient algorithm for learning the parameters of the centers and the widths of the Gaussian membership functions and the rule consequents. The error of the fuzzy decision tree function is defined as a differentiable function in terms of the root-mean-square deviation of E [14]:

$$E = \frac{1}{2N} \sum_{j=1}^{N} (Y_j - C_j)^2 \tag{8}$$

where Y_j is the output value obtained as a result of fuzzy inference,
C_j is the output value from the training sample,
N is a number of observations.

A necessary condition for minimizing the error is that its derivatives are zero in respect to Gaussian parameters (center c_{ki}, width σ_{ki} of the Gaussian function of membership of the *k*-th input attribute on the *i*-th branch) and to the parameters of the rule consequents. As a result, we obtain the following general rule for updating parameters by the back propagation method:

$$r(t+1) = r(t) - \eta \frac{\partial E}{\partial r}. \tag{9}$$

By the rule for differentiating a composite function, we write the partial derivative:

$$\frac{\partial E}{\partial r}=\frac{\partial E}{\partial Y}\frac{\partial Y}{\partial \mu}\frac{\partial \mu}{\partial r}=\sum_{j=1}^{N}(Y_j-C_j)\frac{\partial Y}{\partial \mu}\frac{\partial \mu}{\partial r}\cdot \tag{10}$$

We get the following rules for updating the parameters:

$$r_i^1(t+1)=r_i^1(t)+\frac{\eta}{N}\sum_{j=1}^{N}(Y_j-C_j)(\prod_k \mu_{ki}(x_{kj}))^2\cdot \tag{11}$$

The expression $\prod_k \mu_{ki}(x_{kj})$ is the degree of truth ω_{ij} of the *i*-th rule premises (branch of the tree) in the *j*-th example, which is computed by the layer 3. Finally, we have:

$$r_i^1(t+1)=r_i^1(t)+\frac{\eta}{N}\sum_{j=1}^{N}(Y_j-C_j)\omega_{ij}^2 \tag{12}$$

The condition for termination is reaching the counter of iterations of a given value or getting the error value less than the set value.

4 Conclusion

1. Instead of a classical *4*-layer network, we can use a neural-fuzzy system in order to solve the problems of classification of complex objects in a multidimensional weakly formalizable space of features using expert data. The structure of this neural-fuzzy system automatically determines the base of fuzzy rules, reverse cycle-iteration of neural network by the back propagation method, and may be applied for setting parameters (center and width of the membership function and weight parameters rules consequents on tree leaves).

2. The synthesized knowledge base can be interpreted as a partition of the space of influencing factors into regions with blurred boundaries, within which the response function assumes an indistinct value. The number of such fuzzy regions is equal to the number of rules.

3. The proposed neural network method of adapting the parameters of the decision tree (as a fuzzy hierarchical system) based on the back propagation of the error from the tree leaves to the root node, improves the accuracy of the fuzzy decision tree classification, without changing the tree structure while preserving the interpretation.

4. The method proposed increases the speed of operation of the classical algorithm because of the optimized network scheme and the reduced number of calculations. In comparison with another known method of synthesis of neural-fuzzy networks using the training sample by introducing rules into its structure, our method is advantageous because it does not require a parametrically complex construction and cumbersome model with a large number of neurons.

5. The described structural-functional scheme of an adaptive neuro-fuzzy classifier with a logical conclusion, built based on a combination of hierarchical methods (decision trees) with fuzzy logic and neural networks, has a significant cognitive potential for modeling sensations, perceptions, pattern recognition, learning, and memorizing patterns with the purpose of identifying knowledge from the data. It allows the complex use of strong qualities of fuzzy systems (interpretability of accumulated knowledge) and neural networks (the ability to learn from data).

6. From the point of view of expert methods, a new intellectual logically transparent foresight tool for foresight research has been obtained.

7. Such hybrid neural-fuzzy networks have great potential for practical use in intellectual systems for analyzing, managing and supporting multi-criteria decision-making process in the case of processing multidimensional weakly formalized expert data

describing advanced technologies, scientific and technical (innovative) objects, and territories.

Acknowledgements

The article is prepared within the framework of Competitiveness Enhancement Programme of National Research Tomsk State University and grant of the Russian Foundation for Basic Research No 16-29-12858.

The authors are grateful to Tatiana B. Rumyantseva from the National Research Tomsk State University for her assistance in preparing the paper.

References

[1] A. Sokolov, Fores. **1**, 8 (2007)
[2] M. Keenan, I. Miles *The Handbook of Technology Foresight* (Edward Elgar, 2008)
[3] N.V. Morozova, Bul. of the Chuv. Univ. **3**, 178 (2014)
[4] S.V. Gorbachev, V.I. Syryamkin, *Neuro-fuzzy methods in intelligent systems for processing and analyzing multidimensional information* (Tomsk State University, 2014)
[5] D.V. Shashev, S.V. Shidlovskiy, *J. Phy: Conf. Ser.* **881** (2017)
[6] D.V. Shashev, *MATEC Web Conf.* **79** (2016)
[7] D.A. Vyatchekin, *Fuzzy methods of automatic classification* (UE «Technoprint», 2004)
[8] V.I. Smirnov, *Forecasting and classification of economic systems under conditions of uncertainty by methods of artificial neural networks* (Orenburg State University, 2004)
[9] D.S. Khaydukov, Phil. of mathem: act. probl. **2**, 4 (2009)
[10] V.V. Boyarshinova, Eltyshev D.K. *Automated control systems and information technologies* (Publishing House of the Central Scientific and Technical University, 2015)
[11] D. Rutkovskaya, M. Pilinsky, L. Rutkovsky, *Neural networks, genetic algorithms and fuzzy systems* (Moscow, 2006)
[12] C.Z. Janikow, *Fuzzy Processing in Decision Trees* (University of Missouri, 1993)
[13] S.V. Gorbachev, *Designing a Hybrid Neural-Fuzzy System Based on Soft Computing* (Tomsk State University, 2014)
[14] V.V. Kruglov, M.I. Long, R.Y. Golunov, *Fuzzy logic and artificial neural networks* (FIZMATLIT, Moscow, 2001)