DEVELOPING METHODS AND ALGORITHMS FOR A DECISION-MAKING INTELLECTUAL SUPPORT IN PERSONNEL MANAGEMENT SYSTEMS

Aleksandra Danilenko

Samara State Aerospace University

This paper proposes methods and algorithms for an intellectual support of decision-making management based on neuro-fuzzy networks. The algorithm for automatic generation of knowledge base is developed in accordance with the criteria of completeness and minimality. Reducing space is made on the basis of expert estimates.

1 INTRODUCTION

The management of modern enterprises demands more efficiency from their leaders. In the period of rapid changes in the market, a shorter rotation cycle of products and services, variability in consumer demand completeness and relevance of the knowledge base for strategic decision-making is quite important, as well as monitoring their implementation. In this regard, leaders have to make many decisions, based on the results of processing large amounts of data that must be carried out in a short time. At the same time, it is necessary to develop the systems of intellectual support of the managerial decision-making process based on modern information technologies.

The technologies of personnel decision-making are based on psycho-diagnostic data in a varying degree today: individual performance received by the diagnosis of individual employees or candidates for vacancies that allow determining the degree of the testee's compliance with the requirements of the profession or occupation. The possibilities of psycho-diagnostic analysis is expanding even more if a diagnostic system includes criteria oriented tests that detect the level of expertise and skills – in this way, diagnostic measurements test evaluate not only personal but also professional qualities of the testee.

The task of predicting the features of human behavior on the basis of psycho-diagnostics includes the following subtasks: diagnosis and interpretations of the data and predictions of certain events based on certain models. In this regard, there is a need to develop systems of decision-making intellectual support of managerial, which allow not only assigning data or conclusions of a certain probability, but also converting this probability as data processing and new revenues. None of the existing software on the market meets these requirements. The methods and the algorithms proposed in this paper to some extent overcome the disadvantages mentioned above.

2 METHODS AND ALGORITHMS

The task of making decisions about the advisability of hiring can be attributed generally to the problem of classification, namely the assignment of the applicant to one of the classes: the candidate is fully compliant with the profession, basically meet the requirements for profession, partially compliant, is not compliant. Because of the large number of parameters affecting each other, there is a multifactorial uncertainty and therefore it becomes impossible to determine the type of the objective function to determine the degree of influence of each factor on the classification result. The application of data factor analysis at the same time does not give the desired results, and does not reduce the dimension of the space of classification. Therefore there is a need for the method of expert estimates. This helps reduce feature space, but does not solve the problem of uncertainty. To solve this problem, in this paper we propose to use the device of neural networks.

Since the information according we make a decision about aptitude of a candidate is the result of various psychological tests, classified data are inaccurate or insufficiently specific. Mathematical tools for solving this type of problems are fuzzy logics and the theory of fuzzy sets. When solving the problem of classification fuzzy neural networks to conventional neural networks are preferred in the presence of the intersection of classes, as the behavior of such

networks may be described on the basis of fuzzy rules, and thus the notation is set in these rules. For today there are two main groups of methods of knowledge acquiring: direct (interviews, literature study and others) and indirect (analysis of a training set of examples, observation of an expert and others). Numerous studies have shown that decision-making under uncertainty, that is when dealing with semi structured, incomplete or unclear information, the methods of the second group are more preferable.

At present there are many different methods of solving the problem of classification, but initially the number of classes is set by the user, based on an approximate idea of the nature of a future decision. Virtually unknown methods, in which in addition to solving the problem of classification the assessment of the significance of signs was carried out. For problems with a small number of examples (less than 20) and 2-3 attributes according to the classification methods of calculation can be carried out manually. For the problems with large dimension, these methods are not suitable.

In this case, the number of examples and the number of features is large enough, that requires the development of methods and algorithms for automatic generation of a fuzzy rule base which is founded on quantitative assessments of classified attributes with the issuance of the result in the form of a set of production rules of the kind "IF –THEN".

2.1 KNOWLEDGE BASE

In this paper we initially assume that the fuzzy–"IF – THEN"– rules specified by the experts are already represented in the linguistic form or in the form of a clear clustering correct input-output data sets. This knowledge base is examined for completeness and the minimum number of rules in the paper. Based on these results the methods and the algorithms for the automatic generation of a knowledge base have been developed.

For using this knowledge base in this paper in the problem of classification we modify the method of Abe and Lan used in to solve the problem of approximating functions. In our case, each subject is characterized by one-dimensional output y and m-dimensional input vector x, where x-the test results of the test for a particular factor and y-class accessories test (the candidate is fully compliant with the profession, basically meet the requirements for profession, partially compliant, is not compliant).

Write down the problem in general terms. The definition of a domain y is divided into n intervals and is determined by the degree of meeting the candidate the requirements of profession.

$[y_0, y_1], (y_1, y_2], ... [y_{n-1}, y_n]$, where $y_0 = M_1$, $y_n = M_2$.

Let us call i interval $(y_{i-1}, y_i]$ an output interval i. Using the specified output data, for which the outputs are in the output interval i, it is needed to recursively define the scope of the input values corresponding to this input interval. Namely, first of all, we must find the areas of activation which determine the input area corresponding to the output interval i by calculating maximum and minimum values of the input data. This must be done for each output interval. If the activation area of the output interval i overlaps the activation area of the output interval j, then the overlap area is defined as prohibition. If the input data for the output intervals and/or located inside the prohibition area, then one or two additional activation areas are determined. In future, if such additional areas of activation also overlap, the prohibition area is also determined for them. This process is shown in Figure 1, it is repeated as long as the problem of the overlay areas is resolved.

Fuzzy rules are defined on the basis of the activation area or a reason the basis of activation and the corresponding prohibition areas (if generated). Consider the algorithm by which fuzzy input data membership degree to the appropriate output intervals are calculated, followed by reduction to the clear value. This approach allows creating the initial rule base and optimizing it in the course of working with data. In the automatically generated basis of production rules its completeness and consistency is controlled, which leads to improvement of mechanisms of fuzzy logical deduction. Structuring the knowledge base to enhance the quality

and effectiveness of decision-making and namely to increase the speed of training network and its accuracy.
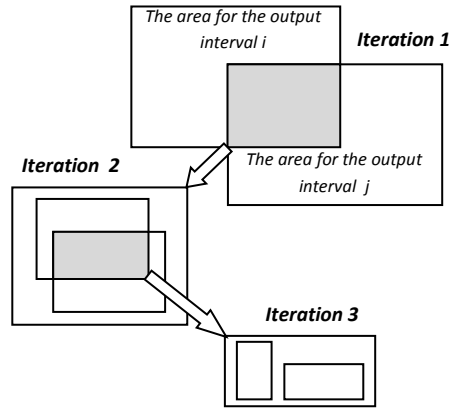

Figure 1: Recursive Identification of the Areas of Activation and Prohibition

The contents of the generated knowledge base is supplied to the neural network as a training sample. Network input vector consists of the test results of a particular testee, and the class of the candidate belonging to a particular profession is the output of the network.

For implementation of the intellectual support of the decision-making system the model of a hybrid neuro-fuzzy classifier and a fuzzy multilayer perceptron is developed.

2.2. HYBRID NEURO CLASSIFIER

The hybrid neuro classifier consists of four layers. Elements of the first layer implement the operation of fuzzification that is form of membership of the input variables to certain fuzzy sets A_{ij}:

$$\mu_{A_{ji}}(x'_j) = \exp\left[-\frac{1}{2}\left(\frac{x'_j - c_{ij}}{\sigma_{ij}}\right)^2\right]$$, where c_{ij}, σ_{ij}-the parameters of the membership function of bell type.

The initial values of these parameters are set so that the membership function satisfies the property of completeness, normality and convexity. The values c_{ij} should be uniformly distributed in the area of input vectors x. Values of these parameters are corrected during the training of the network based on the gradient method.

Each element of the second layer is neuron "I". They carry out the aggregation of the truth prerequisites degrees of each rule in the base in accordance with the operation of the T-norm by the formulas: $\alpha_1 = \min\{A_{11}(x_1), A_{12}(x_2),...,A_{1n}(x_n)\}$; ... $\alpha_n = \min\{A_{n1}(x_1), A_{n2}(x_2),...,A_{nn}(x_n)\}$

Elements of the third layer normalize and calculate the following values:

$$\beta_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2 + ... + \alpha_n}; \ ... \ \beta_n = \frac{\alpha_n}{\alpha_1 + \alpha_2 + ... + \alpha_n}$$

The elements of the fourth layer calculate the values of conclusions for each rule:

$$y'_1 = B_1^{-1}(\alpha_1) = a_1 + \frac{1}{b_1}\ln\frac{1-\alpha_1}{\alpha_1}; \ ... \ y'_n = B_n^{-1}(\alpha_n) = a_n + \frac{1}{b_n}\ln\frac{1-\alpha_n}{\alpha_n}$$, where a_i, b_i − the nonlinear parameters of membership functions $\mu_{B_i}(y)$ of the fuzzy sets of conclusions of the rules.

Fuzzy network outputs are calculated as follows: $y'_i = \beta_i B_i^{-1}(\alpha_i)$. These outputs are interpreted as membership degrees of the presented object to the corresponding class.

2.3. FUZZY MULTILAYER PERCEPTRON

The fuzzy multilayer perceptron consists of two parts: the "fuzzy layer" neurons and the actual multilayer perceptron. The activation function of the "fuzzy layer" neurons of such a network are radial basis functions (in the form of Gaussian function), modeling membership

functions. These neurons are used to determine the degree of membership of the input vector components (which may be fuzzy). In the neurons outputs of this layer the coefficients are formed in the required for the furture recognition form. The outputs of the "fuzzy layer" are used as the inputs to the traditional multilayer perceptron.

If the input of the network is supplied by $x = [x_1, x_2, ..., x_m]^T$, then the output "of fuzzy layer" is formed vector of degrees of x belonging to specific centers (radial basis functions):

$$\mu(x) = [\mu_1(x), \mu_2(x), ..., \mu_m(x)]^T$$

Specific components of the vector are calculated in such a way as to satisfy the condition of normalization $\sum_{i=1}^{L} \mu_1(x^{(k)}) = 1$, for each vector x^k, $k=1...K$, where K – the number of vectors in the training set.

The model of the fuzzy multi-layer perceptron contains the following operators:

1. the operator forming the membership degrees of the input variables to certain fuzzy sets:

$$\mu_1(x^{(k)}) = 0.5 + \frac{\exp(f(p_2 - p_1)/p_{med}) - \exp(-f)}{2(\exp(f) - \exp(-f))}, \quad \mu_2(x^{(k)}) = 1 - \mu_1(x^{(k)}),$$ where f – the constant selected experimentally (f>0); p_2 – the distance from the vector $x^{(k)}$ to the center of the class 2; p_1 – the distance from the vector $x^{(k)}$ to the center of the class 1 in accordance with a selected metric (e.g., Euclidean); p_{med} – the distance between the centers of the classes 1 и 2;

2. the operator of logical product conditions for aggregation of the fuzzy rules relative to the classified image.

The outputs of the fuzzy multilayer perceptron are interpreted as membership degrees of the presented object of the appropriate class.

2.4. RESEARCH OF CONSTRUCTING NETWORKS` QUALITY

In these neuro-fuzzy networks is used the algorithm of Tsukomoto, and as a training algorithm is realized Backpropagation. The result of research of constructing networks' quality as a universal classifier revealed the influence of the parameters of a hybrid neuro-fuzzy classifier (the volume of training sample, the step of network training, the type of the fuzzification function and the defuzzification method) and the fuzzy multilayer perceptron (volume of training sample, step training network, the type of the fuzzification function and the number of neurons in the hidden layer) on the rate of learning. These results can be summarized as follows:

1. Training sample high dimension does not lead to an improvement in quality of training of the neural network, but only causes its insensitivity to redundant data. For this task, the optimum is reached when the amount of training sample is 100 examples. However, if the number of training examples is limited, it is necessary to train the network to use the sample of 33 examples.

2. Quick network training increases the risk of deviations from the solutions followed by oscillations around it. Low values of the learning step eliminate this problem, leading to a slowdown in the process of network training. For this problem, in both networks 0.45 step is optimal.

3. Training of hybrid neuro-fuzzy classifier with Gaussian function is faster than modifying it with a triangular function. However, in the latter case, the network training algorithm is greatly simplified due to a simpler form of the fuzzification function, so that the system behaves more predictably. Training of the fuzzy multilayer perceptron with the hyperbolic tangent function is faster than it can be modified by the logistic fuzzification function. However, if the duration of the network training is not essential, it is advisable to use the fuzzy multilayer perceptron with the logistic fuzzification function because in this case the system behaves more predictably.

4. With the increase in the number of neurons in the hidden layer the fuzzy multilayer perceptron is trained quickly to the set acceptable error threshold. In this problem with the optimal number of neurons step should be chosen from 35 to 40.

5. The view of the defuzzification method does not affect the standard deviation and the training rate of the hybrid neuro- fuzzy classifier.

6. The hybrid neuro-fuzzy classifier training algorithm has a fast the network and good interpretability of the gained knowledge.

7. The fuzzy multilayer perceptron naturally behaves at all the stage of the network training.

Next, a comparative analysis of the constructed fuzzy-neural classifiers is presented (table 1). To solve the original problem of classification of candidates for a specific position on the basis of their professional competence it is advisable to use the hybrid neuro- the fuzzy classifier, because it has a fast network training algorithm and good interpretability of the acquired knowledge, which is useful for solving applied problems.

Table 1: Results of the Classifying Properties of Networks

Network	Fuzzification Function	Defuzzification Method	Classification Error, %
The Hybrid Neuro Classifier	Gaussian	minimal deviation	6
		SD	1
		maximum deviation	7
	Triangular	minimal deviation	12
		SD	3
		maximum deviation	7
The Fuzzy Multilayer Perceptron	Logistic	-	4
	Hyperbolic tangent	-	17

The upgraded structure of the hybrid neuro-fuzzy classifier and its training algorithm allow using the network to solve the problem of predicting, e.g., predicting the success of a professional. The ordinary approach in this case demands a plurality of measuring of individual parameters, the dynamics of which would allow building the trend, but in the actual practice of personnel management it is impossible. However, the neural network proposed above is able to maintain these forecasts, after training on a variety of examples.

3 CONCLUSIONS

Thus, in this paper we have proposed the method and the algorithm for automatic generation of knowledge in accordance with the requirements of the subject area, the neuro-fuzzy network model for solving the problem of the support of decision-making in personnel management systems based on psycho-diagnostics has been developed.

The practical value of the paper is the development of algorithms and automated software system of decision-making in personnel management systems, which allows solving the following problems:

1. support of decision-making about the compliance of the candidate with the requirements of the profession;

2. rotation of employees and the redistribution of functional responsibilities within the team of the enterprise;

3. multifactorial diagnosis of professionally important qualities of the person on the basis of an interactive list of requirements for the profession;

4. predicting professional success.

References

1. Novak, V., Perfilieva, I., Mochkorzh, I., 2006. Mathematical Principles of Fuzzy Logic. FIZMATLIT.

2. Borisov, V.V., Kruglov, V.V., Fedoulov, A.S., 2007. Fuzzy Models and Networks. Hotline-Telec