Zhengbing Hu[1], Yevgeniy V. Bodyanskiy[2], Oleksii K. Tyshchenko[3], Viktoriia O. Samitova[4]

[1]Ph.D., Associate Professor of School of Educational Information Technology, Central China Normal University, Wuhan, China
[2]Dr.Sc., Professor, Professor of AI department, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[3]Ph.D., Senior Researcher, Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[4]Post-Graduate student, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

FUZZY DATA CLUSTERING IN THE RANK SCALE BASED ON A DOUBLE NEO-FUZZY NEURON

Context. A task of data classification under conditions of clusters' overlapping is considered in this article. Besides that, it's assumed that information to be processed is given in the rank scale.

Objective. It's proposed to use a double neo-fuzzy neuron for classification which is a modification of a traditional neo-fuzzy neuron with specially designed asymmetrical membership functions and improved approximating properties.

Method. The double neo-fuzzy neuron (just like the traditional one) is designated for processing data given the scale of natural numbers. However, the situation may become complicated greatly if source data is not given in the numerical scale but in the ordinal one which is a quite common case for a wide variety of practical tasks.

Results. A gradient minimization procedure with a variable learning step parameter was used for learning the double neo-fuzzy neuron. The proposed approach to fuzzy classification for data given in the ordinal scale based on the double neo-fuzzy neuron which is learnt with the help of a high-speed algorithm possesses additional smoothing properties. The clustering accuracy for a training sample and the test one as well as the system's learning speed were measured during experiments. The proposed architecture of the double neo-fuzzy neuron is a sort of compromise between a traditional neo-fuzzy neuron and its extended modification. This architecture demonstrates good performance in those cases when the results' accuracy has more influence compared to the elapsed time used for data processing.

Conclusions. Experimental implementation (for both artificial and real-world data) proved efficiency of the proposed techniques. During the experiments, properties of the proposed system were studied which confirmed usability of the proposed system for a wide range of Data Mining tasks.

Keywords: neuro-fuzzy system, Computational Intelligence, membership function, fuzzy clustering, neo-fuzzy neuron, rank scale.

NOMENCLATURE

ANFIS is an adaptive neuro-fuzzy inference system;

$x(k)$ is a vector of input values;

$x_i(k)$ is the i-th component of $x(k)$;

$y(k)$ is an output signal of the double neo-fuzzy neuron;

μ_{ji} is a membership function of the i-th component in a vector of input values for the j-th centroid;

$w_{ji}(k)$ is a synaptic weight for the j-th membership function of the i-th component;

NS_i is the i-th neo-fuzzy synapse;

$e(k)$ is an error at a time step k;

E an objective function;

$d(k)$ is an external reference signal;

η is a learning rate parameter that usually defines a convergence speed of a learning process;

n is a number of neo-fuzzy synapses;

m_i is a number of membership functions;

$u(k)$ is an output signal of the first layer of the double neo-fuzzy neuron;

c_{ji} is a center of the j-th membership function of the i-th component;

$x_i^{r_i}$ is a linguistic variable;

r_i is a corresponding rank;

D a sample of learning signals.

INTRODUCTION

Hybrid systems of Computational Intelligence and especially adaptive neuro-fuzzy systems are currently popular for solving pattern recognition, classification etc. tasks under conditions of substantial uncertainty about a data distribution's nature and mutual classes' overlapping [1–4].

Main drawbacks of these systems are their bulkiness (for example, a five-layer ANFIS [2] and other similar systems) and a low convergence speed for corresponding learning algorithms that requires big volumes of training samples.

To overcome these shortcomings, an architecture of a neo-fuzzy neuron was introduced for neuro-fuzzy systems [5, 6] which is similar in its architecture to a traditional formal neuron with n inputs, although it contains nonlinear synapses NS_i, $i = 1, 2, ..., n$ (instead of traditional synaptic weights). These nonlinear synapses are formed by a set of triangular symmetrical membership functions μ_{ji} ($j = 1, 2, ..., m_i$) which are equally distributed in the interval $[0,1]$. Each membership functions is connected to its own adjustable weight w_{ji}.

An output of the neo-fuzzy neuron in response to a presented image $x(k) = (x_1(k), x_2(k), ..., x_n(k))^T$ (here $k = 1, 2, ..., N$ stands for either a number of an image $x(k)$ in a training set or a current discrete time item) can be written down like

$$y(k) = \sum_{i=1}^{n} f_i(x_i(k)) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \mu_{ji}(x_i(k)) w_{ji}(k),$$

where $w_{ji}(k)$ is a current meaning of an adjustable synaptic weight in a time moment k for the j-th membership function of the i-th component of an input signal.

The standard square error is commonly used as a learning criterion of the neo-fuzzy neuron

$$E(k) = \frac{1}{2}(d(k) - y(k))^2 = \frac{1}{2}e^2(k) =$$

$$= \frac{1}{2}\left(d(k) - \sum_{i=1}^{n}\sum_{j=1}^{m_i}\mu_{ji}(x_i(k))w_{ji}\right)^2.$$

Minimizing this function with the help of a gradient procedure leads to a learning algorithm [6]

$$w_{ji}(k+1) = w_{ji}(k) + \eta e(k)\mu_{ji}(x_i(k))$$

where $d(k)$ is an external reference signal, η is a learning rate parameter that is usually chosen empirically and defines a convergence speed of a learning process.

Centers of triangular membership functions μ_{ji} are located in such a way that the Ruspini (unity) partition is provided, i.e.

$$\sum_{j=1}^{m_i}\mu_{ji}(x_i(k)) = 1, \; i = 1, 2, ..., n \qquad (1)$$

that makes introduction of a hidden defuzzification layer unnecessary (one can usually find this layer in most of neuro-fuzzy systems).

An architecture of a double neo-fuzzy neuron was introduced in [7] to improve approximating properties of the neo-fuzzy neuron.

1 PROBLEM STATEMENT

Initial data for solving the fuzzy clustering task is a set (a sample) of images formed by N n − dimensional feature vectors $X = \{x(1), x(2), ..., x(k), ..., x(N)\}$ (here $x(k) = \left\{x_i^{r_i}(k)\right\}$, $i = 1, 2, ..., n$, $r_i = 1, 2, ..., m_i$ is a rank of a specific value of a linguistic variable in the i-th coordinate of the n − dimensional space for the k-th observation) and a sample of reference signals where $d(k) = d^{r_0}(k)$, $r_0 = 1, 2, ..., m_0$, is a rank of a reference signal's value in the sample D. After the double neo-fuzzy neuron has been learnt in data ranking, a partition of an initial data array X into m_0 overlapping classes with membership functions μ_{l0} of the k-th image to the l-th class should be provided.

2 REVIEW OF THE LITERATURE

A new relational fuzzy clustering method (FRFP) was proposed in [8]. This procedure is based on determining a fixed point of a function of the desired membership matrix. The produced membership matrices are in some sense less crisp than those produced by NERFCM and more representative of the proximity matrix that is used as input to the clustering process.

An algorithm for finding a fuzzy membership matrix in case of numerical and categorical features is described in [9]. A set of feature vectors with mixed features is mapped into a set of feature vectors with only real-valued components. The only condition for this situation is that a new set of vectors has the same proximity matrix as the original feature vectors. And then FCM is used for clustering of the new set of vectors.

An article [10] introduces another fuzzy relational clustering method for finding a fuzzy membership matrix. Objects' clustering may be performed by describing objects in terms of feature vectors as well as on the basis of relational data. And the relational data may be considered in terms of proximities between objects. Finally, proximities between membership vectors should be proportional to proximities between objects. Components' values of a membership vector corresponding to an object are membership degrees of an object in various clusters. In other words, a membership vector is just a sort of a feature vector.

A method where unknown numeric variables are assigned to ordinal values is described in [11]. Minimizing a clustering objective function means finding numeric values for these variables. The proposed method clustering utilizes the same objective function which is used by the FCM algorithm except the fact that both a membership function and an ordinal to real mapping are determined by the gradient descent method.

An introduced clustering method [12] is based on a modified FCM version. Input data features are considered as linguistic variables. A set of fuzzy numbers actually describes any feature. The modified method makes it possible to find an appropriate number of clusters. Besides that, it's claimed that this algorithm possesses some improved robust characteristics when compared to the traditional FCM.

A new algorithm for developing a mapping of ordinal values into numerical ones is presented in [13] for which a measure of dissimilarity exists. This method is a part of the well-known FCM procedure. The modified algorithm demonstrates better data partition into clusters as well as an ordinal-numerical mapping that reveals hidden structural knowledge of an ordinal feature.

A clustering algorithm is presented in [14] for data sets of mixed features (nominal, numerical and ordinal). The algorithm aims at reducing a negative effect from noise. An optimization function utilizes the likelihood for each individual feature as an optimization criterion for similarity or likeliness between patterns and clusters which is a quite opposite idea compared to FCM based on distances or the EM clustering algorithm. It's claimed that this method can quickly find fuzzy clusters with different distributions in each feature level.

An article [15] discusses a method for representation of ordinal values by fuzzy sets on a fixed interval of real values. This procedure contains centroids determined by a frequency distribution on ordinal values. Then triangular and trapezoidal fuzzy sets can be found that have these centroids by means of the gradient descent method. The obtained fuzzy sets do not share parameters which would be the case if an end point of one fuzzy set was a vertex of an adjacent fuzzy set.

That's why we would like to propose a new fuzzy clustering method based on the double neo-fuzzy neuron designated for clustering rank (ordinal) data. At the same time, this method possesses robust characteristics.

3 MATERIALS AND METHODS

An architecture of the double neo-fuzzy neuron is shown in Fig. 1.

It consists of two layers: the first layer is formed by n nonlinear synapses NS_i with m_i membership functions and synaptic weights for each nonlinear synapses; the output layer is formed by a nonlinear synapses NS_0 with m_0 membership functions μ_{l0} ($l = 1, 2, ..., m_0$) and synaptic weights w_{l0}.

If an image $x(k)$ is fed to the input of the double neo-fuzzy neuron, an output signal is

$$y(k) = f_0(u(k)) = f_0\left(\sum_{i=1}^{n} f_i(x_i(k))\right) = \sum_{l=1}^{m_0} \mu_{l0}(u(k)) w_{l0} =$$

$$= \sum_{l=1}^{m_0} \mu_{l0}\left(\sum_{i=1}^{n}\sum_{j=1}^{m_i} \mu_{ji}(x_i(k)) w_{ji}\right) w_{l0}.$$

It can be seen that a value of the output signal of the double neo-fuzzy neuron is defined by components' $x_i(k)$ values for an input image as well as by values of $\left(\sum_{i=1}^{n} m_i + m_0\right)$ membership functions and their corresponding synaptic weights.

Triangular membership functions that are equally distributed in the interval $[0,1]$ and meet the condition (1) can be written down like

$$\mu_{1i}(x_i) = \begin{cases} \dfrac{c_{2i} - x_i}{c_{2i}}, & x_i \in [0, c_{2i}], \\ 0, & x_i \notin [0, c_{2i}], \end{cases} \tag{2}$$

$$\mu_{ji}(x_i) = \begin{cases} \dfrac{x_i - c_{j-1,i}}{c_{ji} - c_{j-1,i}}, & x_i \in [c_{j-1,i}, c_{ji}], \\ \dfrac{c_{j+1,i} - x_i}{c_{j+1,i} - c_{ji}}, & x_i \in [c_{ji}, c_{j+1,i}], \\ j = 2, ..., m_{i-1}, \\ 0 \text{ otherwise}, \end{cases} \tag{3}$$
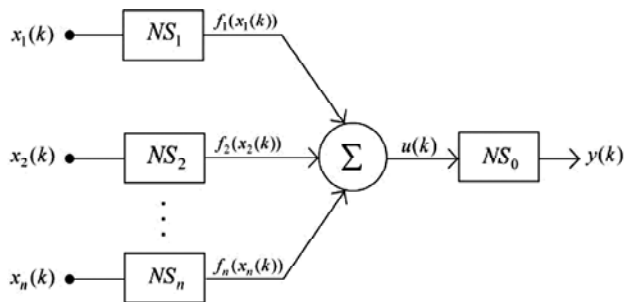
$$\mu_{m_i i}(x_i) = \begin{cases} \dfrac{x_i - c_{m_i-1,i}}{1 - c_{m_i-1,i}}, & x_i \in [c_{m_i-1,i}, 1], \\ 0, & x_i \notin [c_{m_i-1,i}, 1], \end{cases} \tag{4}$$

$$\mu_{10}(u) = \begin{cases} \dfrac{c_{20} - u}{c_{20}}, & u \in [0, c_{20}], \\ 0, & u \notin [0, c_{20}], \end{cases} \tag{5}$$

$$\mu_{l0}(u) = \begin{cases} \dfrac{u - c_{l-1,0}}{c_{l0} - c_{l-1,0}}, & u \in [c_{l-1,0}, c_{l0}], \\ \dfrac{c_{l+1,0} - u}{c_{l+1,0} - c_{l0}}, & u \in [c_{l0}, c_{l+1,0}], \\ l = 2, ..., m_{0-1}, \\ 0 \text{ otherwise}, \end{cases} \tag{6}$$

$$\mu_{m_0 0}(u) = \begin{cases} \dfrac{u - c_{m_0-1,0}}{1 - c_{m0-1,0}}, & u \in [c_{m_0-1,0}, 1], \\ 0, & u \notin [c_{m_0-1,0}, 1] \end{cases} \tag{7}$$

(here c_{ji} and c_{l0} are centers of corresponding membership functions). Wherein a distance between centers is constant for each nonlinear synapse and can be written like

$$c_{ji} - c_{j-1,i} = (m_i - 1)^{-1},$$

$$c_{l0} - c_{l-1,0} = (m_0 - 1)^{-1}.$$

Using the Ruspini (unity) partition leads to the fact that only two neighboring membership functions are activated at every learning step. Denoting these membership functions μ_{pi}, $\mu_{p+1,i}$ ($i = 0, 1, ..., n$), we can write down

$$f_i(x_i(k)) = \sum_{j=1}^{m_i} \mu_{ji}(x_i(k)) w_{ji}(k) = \mu_{pi}(x_i(k)) w_{pi}(k) +$$

$$+ \mu_{p+1,i}(x_i(k)) w_{p+1,i}(k) = \frac{c_{p+1,i} - x_i(k)}{c_{p+1,i} - c_{pi}} w_{pi}(k) +$$

$$+ \frac{x_i(k) - c_{p,i}}{c_{p+1,i} - c_{pi}} w_{p+1,i}(k) = a_i(k) x_i(k) + b_i(k),$$

where

$$a_i(k) = \frac{w_{p+1,i}(k) - w_{pi}(k)}{c_{p+1,i} - c_{pi}},$$

$$b_i(k) = \frac{c_{p+1,i} w_{pi}(k) - c_{pi} w_{p+1,i}(k)}{c_{p+1,i} - c_{pi}}$$

and

$$u(k) = \sum_{i=1}^{n} a_i(k) x_i(k) + b_i(k),$$

Figure 1 – An architecture of the double neo-fuzzy neuron

$$y(k) = \sum_{l=1}^{m_0} \mu_{l0}(u(k))w_{l0}(k) = \mu_{p0}(u(k))w_{po}(k) +$$

$$+ \mu_{p+1,0}(u(k))w_{p+1,0}(k) = \frac{c_{p+1,0} - u(k)}{c_{p+1,0} - c_{p0}} w_{p0}(k) +$$

$$+ \frac{u(k) - c_{p0}}{c_{p+1,0} - c_{p0}} w_{p+1,0}(k) = a_0(k)u(k) + b_0(k).$$

where

$$a_0(k) = \frac{w_{p+1,0}(k) - w_{p0}(k)}{c_{p+1,0} - c_{p0}} ,$$

$$b_0(k) = \frac{c_{p+1,0}w_{p0}(k) - c_{p0}w_{p+1,0}(k)}{c_{p+1,0} - c_{p0}} .$$

Thus, the double neo-fuzzy neuron provides piecewise-linear approximation of some nonlinear separating function in the form

$$y(k) = a_0(k)(\sum_{i=1}^{n} a_i(k)x_i(k) + b_i(k)) + b_0(k)$$

where parameters $a_i(k)$ and $b_i(k)$ to be tuned are defined by both values of corresponding membership functions and trained synaptic weights.

The double neo-fuzzy neuron (just like the traditional one) is designated for data processing given in the scale of natural numbers. A situation is getting substantially more complicated when initial data are given in the ordinal (rank) scale. This sort of cases can be often encountered in sociology, economics, medicine, education, etc. [16]. For a one-dimensional case this information is given in the form of an ordered sequence of linguistic variables

$$x_i^1, x_i^2, ..., x_i^{r_i}, ..., x_i^{m_i}, 1 \leq ... \leq r_i - 1 \leq r_i \leq r_{i+1} \leq ... \leq m_i \leq ... \leq N$$

where $x_i^{r_i}$ is actually a linguistic variable, r_i is a corresponding rank.

It was offered in [12, 15, 17] to perform fuzzification for initial data based on the occurrence frequency distribution analysis of specific linguistic variables for processing data given in the ordinal scale. It was also assumed that these distributions were subject to the Guassian law. There was an approach in [18] that was not associated with the normal distribution hypothesis which we will use in the future. Thus, initial data for solving the fuzzy clustering task is a set (a sample) of images formed by N n–dimensional feature vectors $X = \{x(1), x(2), ..., x(k), ..., x(N)\}$ (here $x(k) = \{x_i^{r_i}(k)\}$, $i = 1, 2, ..., n$, $r_i = 1, 2, ..., m_i$ is a rank of a specific value of a linguistic variable in the i-th coordinate of the n–dimensional space for the k-th observation) and a sample of reference signals $D = \{d(1), d(2), ..., d(k), ..., d(N)\}$ where $d(k) = d^{r_0}(k)$,

$r_0 = 1, 2, ..., m_0$ is a rank of a reference signal's value in the sample D. After the double neo-fuzzy neuron has been learnt in data ranking, a partition of an initial data array X into m_0 overlapping classes with membership functions μ_{l0} of the k-th image to the l-th class should be provided.

The fuzzification procedure for a sequence of rank linguistic variables should be considered by an example of a one-dimensional sample $x_i(1), x_i(2), ..., x_i(N)$ where every observation $x_i(k)$ may be assigned to one of the ranks r_i, $r_i = 1, 2, ..., m_i$.

Let's assume that a value $x_i(k)$ (that corresponds to the r_i-th rank) is met N_{r_i} times in the dataset. Then a relative occurrence frequency may be introduced for the r_i-th rank, $r_i = 1, 2, 3, ..., m_i$. In this case, this condition should be met

$$\sum_{r_i=1}^{m_i} f_{r_i} = 1.$$

Based on the obtained frequencies, asymmetrical unevenly located membership functions μ_{ji} and μ_{l0} are formed with centers which are calculated with the help of recurrent ratios

$$c_{1i} = 0.5f_{1i}; \ c_{ji} = c_{j-1,i} + 0.5(f_{r_i-1} + f_{r_i}) , \ j = 1, 2, ..., m_i.$$

These membership functions are computed with the help of expressions similar to the expressions (2)–(7). The only difference is that one should use

$$\mu_{1i}(x_i) = \begin{cases} 1, x_i \in [0, c_{1i}], \\ \dfrac{c_{2i} - x_i}{c_{2i} - c_{1i}}, x_i \in [c_{1i}, c_{2i}], \\ 0, x_i \notin [0, c_{2i}] \end{cases} \quad (8)$$

instead of the expression (2),

$$\mu_{m_i i}(x_i) = \begin{cases} \dfrac{x_i - c_{m_i-1,i}}{c_{m_i i} - c_{m_i-1,i}}, x_i \in [c_{m_i-1,i}, c_{m_i i}], \\ 1, x_i \in [c_{m_i i}, 1], \\ 0, x_i \notin [c_{m_i-1,i}, 1] \end{cases} \quad (9)$$

instead of the expression (4),

$$\mu_{10}(u) = \begin{cases} 1, u \in [0, c_{10}], \\ \dfrac{c_{20} - u}{c_{20}}, u \in [c_{10}, c_{20}], \\ 0, u \notin [0, c_{20}] \end{cases} \quad (10)$$

instead of the expression (5),

$$\mu_{m_0 0}(u) = \begin{cases} \dfrac{u - c_{m_0-1,0}}{c_{m_0 0} - c_{m_0-1,0}}, u \in [c_{m_0-1,0}, c_{m_0 0}], \\ 1, u \in [c_{m_0 0}, 1], \\ 0, u \notin [c_{m_0-1,0}, 1] \end{cases} \quad (11)$$

instead of the expression (7).

A way of setting membership functions (3), (6), (8), (9), (10), (11) also provides meeting the condition (1).

A gradient minimization procedure with a variable parameter of a search step $\eta_i(k)$ should be used for learning the double neo-fuzzy neuron. An algorithm for tuning an output synapse NS_0 can be written down in the form

$$\begin{cases} w_{l0}(k+1) = w_{l0}(k) + \eta_0(k)e(k)\mu_{l0}(u(k)), l = p, p+1, \\ w_{l0}(k+1) = w_{l0}(k), \forall l \neq p \neq p+1. \end{cases} \quad (13)$$

Thus, the weights are adjusted at every iteration step. These weights correspond to activated membership functions μ_{p0} and $\mu_{p+1,0}$. To increase the convergence speed and to introduce additional smoothing properties, it's expedient to use an algorithm in the form of [19]

$$\begin{cases} w_{l0}(k+1) = w_{l0}(k) + \eta_0^{-1}(k)e(k)\mu_{l0}(u(k)), l = p, p+1, \\ \eta_0(k+1) = \alpha\eta_0(k) + \mu_{p0}^2(u(k+1)) + \mu_{p+1,0}^2(u(k+1)), \\ w_{l0}(k+1) = w_{l0}(k), \forall l \neq p \neq p+1, \\ 0 \leq \alpha \leq 1 \end{cases} \quad (14)$$

which coincides (when $\alpha = 0$) with the one-step optimal algorithm by Kaczmarz-Widrow-Hoff [20, 21] and when $\alpha = 1$ it coincides with the stochastic approximation procedure by Goodwin-Ramage-Caines [22, 23].

To tune the synaptic weights of the first layer, a learning criterion can be written down like

$$E(k) = \frac{1}{2}(d(k) - f_0(u(k)))^2 = \frac{1}{2}(d(k) - f_0(\sum_{i=1}^{n}\sum_{j=1}^{m_i}\mu_{ji}(x(k))w_{ji})). \quad (15)$$

Introducing a derivative

$$\frac{\partial E(k)}{\partial w_{ji}} = -e(k)\frac{\partial f_0(u(k))}{\partial u(k)} \cdot \frac{\partial u(k)}{\partial w_{ji}} = -e(k)a_0(k)\frac{\partial u(k)}{\partial w_{ji}},$$

the gradient minimization procedure (15) can be written down in the form

$$\begin{cases} w_{ji}(k+1) = w_{ji}(k) + \eta_i(k)e(k)a_0(k)\mu_{ji}(x_i(k)), \\ j = p, p+1; i = 1, 2, ..., n, \\ w_{ji}(k+1) = w_{ji}(k), \forall j \neq p \neq p+1 \end{cases}$$

where $\eta_i(k)$ is a step parameter to be defined.

Introducing a designation

$$\mu_{ji0}(x_i(k)) = a_0(k)\mu_{ji}(x_i(k))$$

and using optimization techniques from [24, 25], we come to a simple and effective learning procedure for nonlinear synapses of the first layer

$$\begin{cases} w_{ji}(k+1) = w_{ji}(k) + \eta_i^{-1}(k)e(k)\mu_{ji0}(x_i(k)), \\ j = p, p+1; i = 1, 2, ..., n, \\ \eta_i(k+1) = \alpha\eta_i(k) + \mu_{pi0}^2(x_i(k+1)) + \mu_{p+1,i0}^2(x_i(k+1)), \\ w_{ji}(k+1) = w_{ji}(k), \forall j \neq p \neq p+1, \\ 0 \leq \alpha \leq 1 \end{cases}$$

which completely structurally coincides with the procedure (14).

4 EXPERIMENTS

We chose the sample Nursery from the UCI Repository [26] for the first experiment to demonstrate effectiveness of the proposed neuro-fuzzy system based on the double neo-fuzzy neuron and its learning method for our first experiment. The sample contained 12958 observations with 8 ordinal attributes for each one. The sample was divided into a training one and a test one with a ratio 70/30.

A set of experiments was performed for the second experiment for an adaptive fuzzy clustering method based on the double neo-fuzzy neuron. A synthetic sample was created for this purpose that consisted of 3 non-overlapping clusters, and 20% outliers were additionally added. The generated sample was ranked in heterogeneously broken ranges from 1 to 7 in order to obtain ordinal attributes. The sample is shown in Fig. 2. There's a separate enlarged sampling area on the right where outliers have been previously cut off.

5 RESULTS

We were comparing results obtained from architectures based on the traditional neo-fuzzy neuron, the double neo-fuzzy neuron and the extended neo-fuzzy neuron [27–32]. During the experiment, we were measuring the clustering accuracy for the training sample and the test one as well as the system's learning speed. The results are demonstrated in Table 1.

One can see clustering results of the second experiment with the help of the FCM algorithm (Fig. 3) that's known for its instability to outliers and partition with the help of the adaptive fuzzy clustering methods based on the double neo-fuzzy neuron. Outliers don't influence the result when we apply the adaptive procedure while traditional methods are hypersensitive to observations that lie far enough from all prototypes (centroids).

Table 1 – Clustering results for the dataset Nursery

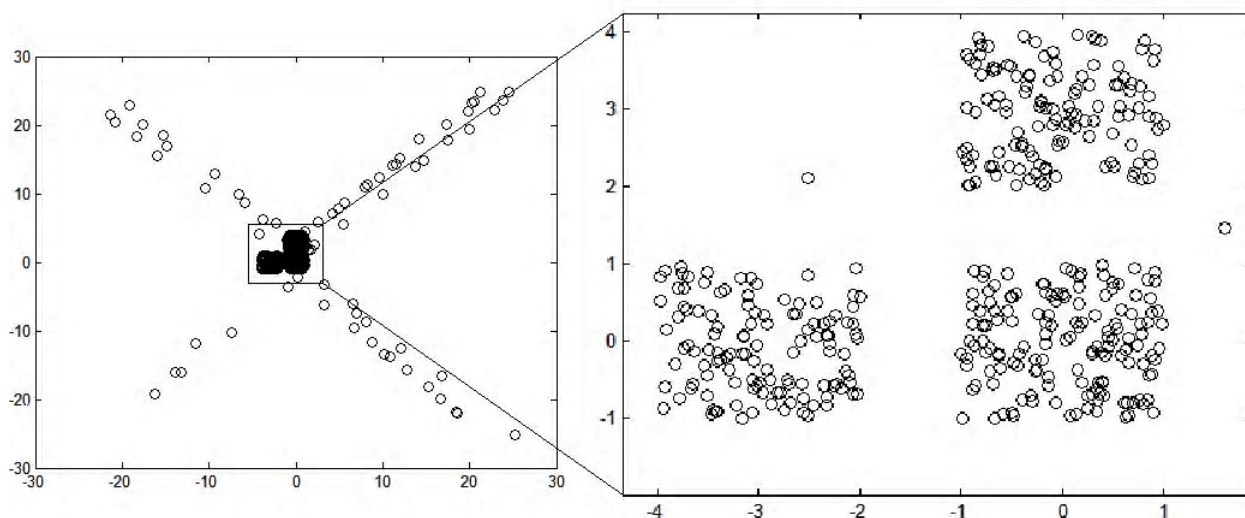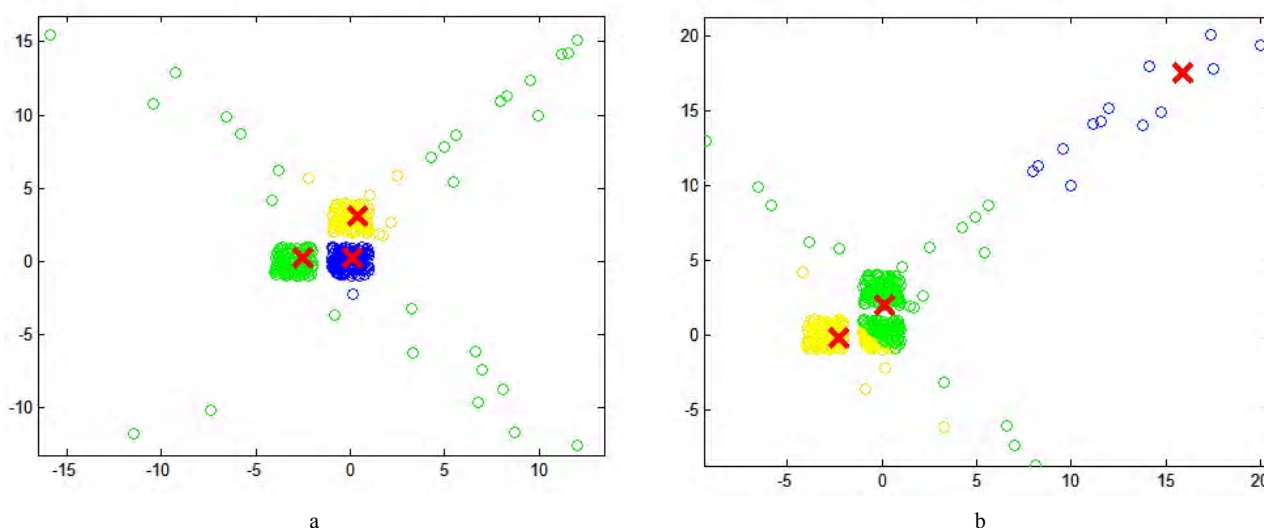| Neuro-fuzzy Systems	Learning time, sec	Training error	Test error
NFN	1,30	0,0167	0,0171
DNFN	1,45	0,0121	0,0127
ENFN	1,74	0,0076	0,084

Figure 2 – An initial dataset

a
b

Figure 3 – Comparing methods: a – with the help of FCM; b – a result of the adaptive fuzzy clustering based on the double neo-fuzzy neuron

6 DISCUSSION

We could come to a conclusion that if we need higher accuracy we should use either ENFN or DNFN (here everything depends on a level of accuracy you'd like to achieve). But if there's a system that's characterized by some speedy reasoning, probably the traditional NFN would be of a great benefit. Speaking of the demonstrated performance, DNFN was 12% slower than the conventional NFN and 20% faster than ENFN. Speaking of forecasting accuracy, DNFN was 35% better than NFN and about 50% worse compared to ENFN.

As it can be seen in Fig. 3, FCM and the proposed method demonstrated different partition. Although both algorithms divided all the data into 3 clusters but FCM assigned all outliers/tails into one cluster that was previously determined (Fig. 3a). At the same time, the proposed method carried out a more reasonable partition. A part of outliers/ tails was assigned to a new cluster. Frankly speaking, the adaptive fuzzy clustering method performed more like what a human expert would do. It demonstrated the fact that it was more robust compared to FCM.

Considering the results, we can assert that the proposed system provides high-quality clustering for ordinal data. The additional layer if the proposed neuro-fuzzy system makes it possible to increase the clustering accuracy, although it may take a little longer if we talk about the learning time.

CONCLUSIONS

It has been widely known that pattern vectors to be clustered may have attributes of various types including ordinal. Ordinal attributes with values such as «poor», «very poor», «good», and «very good» are neither entirely numerical nor entirely qualitative. That's why it leads to some troubles with clustering since it doesn't make any sense to take differences of values for these ordinal attributes as it is required for finding distances between pattern vectors.

An approach to fuzzy data clustering in the ordinal (rank) scale based on the double neo-fuzzy neuron is proposed. The double neo-fuzzy neuron is learnt by a fast learning algorithm that possesses additional smoothing properties.

The obtained results make it possible to process data (given in both the numerical and rank scales) effectively. Multiple simulations demonstrated that the algorithm was quite efficient.

REFERENCES

1. Bezdek J. C. Pattern Recognition with Fuzzy Objective Function Algorithms / J. C. Bezdek. – N.Y. : Plenum Press, 1981. – 272 p.
2. Jang J.-Sh. R. Neuro-Fuzzy and Soft Computing / J.-Sh. R. Jang, Ch.-T. Sun, E. Mizutani. – Upper Saddle River, NJ : Prentice Hall, 1997. – 614 p.
3. Hoeppner F. Fuzzy Cluster Analyse / F. Hoeppner, F. Klawonn, R. Kruse. – Braunschweig : Vieweg, 1999. – 280 p.
4. Nelles O. Nonlinear System Identification / O. Nelles. – Berlin : Springer, 2001. – 785 p.
5. A neo-fuzzy neuron and its applications to system identification and prediction of the system behavior / [T. Yamakawa, E. Uchino, T. Miki, H. Kusanagi] // IIZUKA – 92 : 2nd Int. Conf. on Fuzzy Logic and Neural Networks, Iizuka, Japan, 1992 : proceedings. – P. 477–483.
6. Miki T. Analog implementation of neo-fuzzy neuron and its on-boarding learning / T. Miki, T. Yamakawa // Ed. by N. S. Mastorakis «Computational Intelligence and Application». – Piraeus : WSES Press, 1999. – P. 144–149.
7. Evolving network based on double neo-fuzzy neurons / [Ye. Bodyanskiy, Ye. Gorshkov, V. Kolodyazhniy, P. Otto] // Computer Science Meets Automation : 52nd Int. Scientific Coll., TU Ilmenau (Thuer), 2007: proceedings. – P. 35–40.
8. Brouwer R. K. Fuzzy Relational Fixed Point Clustering / R. K. Brouwer // International Journal of Computational Intelligence Systems. – 2009. – Vol. 2, Iss.1. – P. 69–82.
9. Brouwer R. K. Clustering feature vectors with mixed numerical and categorical attributes / R. K. Brouwer // International Journal of Computational Intelligence Systems. – 2008. – Vol. 1, Iss. 4. – P. 285–298.
10. Brouwer R. K. Fuzzy relational clustering based on comparing two proximity matrices with utilization of particle swarm optimization / R. K. Brouwer, A. Groenwold // Soft Computing. – 2009. – Vol. 13, Iss. 6. – P. 577–589.
11. Brouwer R. K. Fuzzy clustering of feature vectors with some ordinal valued attributes using gradient descent for learning / R. K. Brouwer // International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. – 2008. – Vol. 16, No. 2. – P. 195–218.
12. Butkiewicz B.S. Robust fuzzy clustering with fuzzy data / B. S. Butkiewicz // Lecture Notes in Computer Science. – V. 3528. – Berlin-Heidelberg : Springer-Verlag, 2005. – P. 76–82.
13. Lee M. Fuzzy Clustering and Mapping of Ordinal Values to Numerical / M. Lee, R. K. Brouwer // IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007), 2007. – P. 538–543.
14. Lee M. Likelihood Based Fuzzy Clustering for Data Sets of Mixed Features / M. Lee, R. K. Brouwer // IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007), 2007. – P. 544–549.
15. Brouwer R. K. Fuzzy set covering of a set of ordinal attributes without parameter sharing / R. K. Brouwer // Fuzzy Sets and Systems. – 2006. – Vol. 157, № 13. – P. 1775–1786.
16. Лбов Г. С. Методы обработки разнотипных экспериментальных данных / Г. С. Лбов. – Новосибирск : Наука, 1985. – 160 с.
17. Brouwer R. K. A feedforward neural network for mapping vectors to fuzzy sets of vectors / R. K. Brouwer, W. Pedrycz // Processing ICANN/ICOMIP 2003: Int. Conf. on Artificial Neural Networks and Neural Information, Istanbul, Turkey, 2003. – P. 45–48.
18. Бодянский Е. В. Нечеткая кластеризация данных, заданных в порядковой шкале / Е. В. Бодянский, В. А. Опанасенко, А. Н. Слипченко // Системы обработки информации. – 2007. – Вып. 4(62). – С. 5–9.
19. Бодянский Е. В. Многошаговые оптимальные упредители многомерных нестационарных стохастических процессов / Е. В. Бодянский, И. П. Плисс, Т. В. Соловьева // Доклады АН УССР. – 1986. – Сер. А, №12. – С. 47–49.
20. Kaczmarz S. Approximate solution of systems of linear equations / S. Kaczmarz // Int. J. Control. – 1993. – № 53. – P. 1269–1271.
21. Widrow B. Adaptive switching circuits / B. Widrow, M. E. Hoff Jr. // 1960 IRE Western Electric Show and Connection Record. – 1960. – Part 4. – P. 96–104.
22. Goodwin G. C. Discrete time stochastic adaptive control / G. C. Goodwin, P. J. Ramadge, P. E. Caines // SIAM J. Control and Optimization. – 1981. – № 19. – P. 829–853.
23. Goodwin G. C. A globally convergent adaptive predictor / G. C. Goodwin, P. J. Ramadge, P. E. Caines // Automatica. – 1981. – № 17. – P. 135–140.
24. Бодянский Е. В. Адаптивные алгоритмы идентификации нелинейных объектов управления / Е. В. Бодянский // АСУ и приборы автоматики. – Харьков : Выща школа. – 1987. – Вып. 81. – С. 43–46.
25. Bodyanskiy Ye. An adaptive learning algorithm for neuro-fuzzy network / Ye. Bodyanskiy, V. Kolodyazhniy, A. Stephan // Ed. by B. Reusch «Computational Intelligence. Theory and Applications». – Berlin-Heidelberg, – New York : Springer, 2001. – P. 68–75.
26. UCI machine learning repository [Electronic resource]. – Access mode: http://archive.ics.uci.edu/ml/datasets/.
27. Bodyanskiy Ye. Adaptive learning of an evolving cascade neo-fuzzy system in data stream mining tasks / Ye. Bodyanskiy, O. Tyshchenko, D. Kopaliani // Evolving Systems. – 2016. – Vol. 7, No. 2. – P. 107–116.
28. Bodyanskiy Ye. An Extended Neo-Fuzzy Neuron and its Adaptive Learning Algorithm / Ye. Bodyanskiy, O. Tyshchenko, D. Kopaliani // I. J. Intelligent Systems and Applications (IJISA). – 2015. – Vol. 7, No. 2. – P. 21–26.
29. Bodyanskiy Ye. A hybrid cascade neural network with an optimized pool in each cascade / Ye. Bodyanskiy, O. Tyshchenko, D. Kopaliani // Soft Computing. A Fusion of Foundations, Methodologies and Applications (Soft Comput). – 2015. – Vol. 19, Iss. 12. – P. 3445–3454.
30. An Evolving Cascade System Based on a Set of Neo-Fuzzy Nodes / [Zhengbing Hu, Yevgeniy V. Bodyanskiy, Oleksii K. Tyshchenko, Olena O. Boiko] // International Journal of Intelligent Systems and Applications (IJISA). – 2016. –Vol. 8, No. 9. – P. 1–7.
31. Бодянський Є. В. Еволюційна каскадна система на основі нейро-фаззі вузлів / Є. В. Бодянський, О. К. Тищенко, О. О. Бойко // Радіоелектроніка, інформатика, управління. – 2016. – №2(37). – С. 40–45.
32. Bodyanskiy Ye. A Multidimensional Cascade Neuro-Fuzzy System with Neuron Pool Optimization in Each Cascade / Ye. Bodyanskiy, O. Tyshchenko, D. Kopaliani // Int. Journal of Information Technology and Computer Science (IJITCS). – 2014. – Vol. 6, No. 8. – P. 11–17.

Женьбінг Гу[1], Бодянський Є. В.[2], Тищенко О. К.[3], Самітова В. А.[4]

[1]Канд. техн. наук, доцент Школи освітніх інформаційних технологій, Central China Normal University, Вухань, Китай

[2]Д-р техн. наук, професор, професор кафедри штучного інтелекту Харківського національного університету радіоелектроніки, науковий керівник Проблемної НДЛ АСУ, Харків, Україна

[3]Канд. техн. наук, старший науковий співробітник Проблемної НДЛ АСУ, Харківський національний університет радіоелектроніки, Харків, Україна

[4]Аспірантка кафедри штучного інтелекту Харківського національного університету радіоелектроніки, Харків, Україна

НЕЧІТКА КЛАСТЕРИЗАЦІЯ ДАНИХ У РАНГОВІЙ ШКАЛІ НА ОСНОВІ ПОДВІЙНОГО НЕО-ФАЗЗІ НЕЙРОНУ

Актуальність. У статті розглянуто задачу кластерування даних за умов кластерів, що перетинаються. Передбачено, що оброблювану інформацію задано в ранговій (порядковій) шкалі.

Мета. Для кластерування запропоновано використати архітектуру спеціального вигляду – подвійний нео-фаззі нейрон, який є модифікацією класичного нео-фаззі нейрона зі спеціально сконструйованими несиметричними функціями належності та володіє поліпшеними апроксимуючими властивостями.

Метод. Подвійний нео-фаззі нейрон, як і звичайний нео-фаззі нейрон, призначений для обробки інформації, що її задано у шкалі натуральних чисел. Однак ситуація істотно ускладниться, якщо вихідні дані будуть задані не в числовий, а в порядкової шкалі, що досить часто зустрічається в різних практичних завданнях.

Результати. Для навчання подвійного нео-фаззі нейрона використовувалася градієнтна процедура мінімізації із змінним параметром кроку пошуку. Запропонований підхід до нечіткої класифікації даних у порядковій шкалі на основі подвійного нео-фаззі нейрона, що його навчають за допомогою швидкодіючого алгоритму, має додаткові згладжуючі властивості. Точність кластеризації для навчальної та тестової вибірок, а також швидкість навчання системи були виміряні в ході проведення експериментів. Запропонована архітектура подвійного нео-фаззі нейрона є своєрідним компромісом між традиційним та розширеним нео-фаззі нейронами. Дана архітектура гарно себе проявила у тих випадках, коли точність результатів має більше значення, ніж час, який витрачається на обробку даних.

Висновки. Експериментальне моделювання (на штучних і реальних наборах даних) довело ефективність запропонованого підходу. У ході експериментів було проведено дослідження властивостей запропонованої системи, що ще раз підтвердило доцільність використання даного підходу для вирішення широкого кола завдань інтелектуального аналізу даних.

Ключові слова: нейро-фаззі система, обчислювальний інтелект, функція належності, нечітке кластерування, нео-фаззі нейрон, рангова шкала.

Женбинг Ху[1], Бодянский Е. В.[2], Тищенко А. К.[3], Самитова В. А.[4]

[1]Канд. техн. наук, доцент Школы образовательных информационных технологий, Central China Normal University, Вухань, Китай

[2]Д-р техн. наук, профессор, профессор кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники, научный руководитель Проблемной НИЛ АСУ, Харьков, Украина

[3]Канд. техн. наук, старший научный сотрудник Проблемной НИЛ АСУ, Харьковский национальный университет радиоэлектроники, Харьков, Украина

[4]Аспирантка кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники, Харьков, Украина

НЕЧЕТКАЯ КЛАСТЕРИЗАЦИЯ ДАННЫХ В РАНГОВОЙ ШКАЛЕ НА ОСНОВЕ ДВОЙНОГО НЕО-ФАЗЗИ НЕЙРОНА

Актуальность. В статье рассмотрена задача кластеризации данных в условиях перекрывающихся кластеров. Предполагается, что обрабатываемая информация задана в ранговой (порядковой) шкале.

Цель. Для кластеризации предлагается использовать архитектуру специального вида – двойной нео-фаззи нейрон, который является модификацией классического нео-фаззи нейрона со специально сконструированными несимметричными функциями принадлежности и обладающий улучшенными аппроксимирующими свойствами.

Метод. Двойной нео-фаззи нейрон, как и обычный нео-фаззи нейрон, предназначен для обработки информации, заданной в шкале натуральных чисел. Однако ситуация существенным образом усложнится, если исходные данные будут заданы не в числовой, а в порядковой шкале, что довольно часто встречается в различных практических задачах.

Результаты. Для обучения двойного нео-фаззи нейрона использовалась градиентная процедура минимизации с переменным параметром шага поиска. Предложенный подход к нечеткой классификации данных в порядковой шкале на основе двойного нео-фаззи нейрона, обучаемого с помощью быстродействующего алгоритма, обладает дополнительными сглаживающими свойствами. Точность кластеризации для обучающей и тестовой выборок, а также скорость обучения системы были измерены в ходе проведения экспериментов. Предложенная архитектура двойного нео-фаззи нейрона является своеобразным компромиссом между традиционным и расширенным нео-фаззи нейронами. Данная архитектура хорошо себя показывает в тех случаях, когда точность результатов имеет большее значение, чем время, затрачиваемое на обработку данных.

Выводы. Экспериментальное моделирование (на искусственных и реальных наборах данных) доказало эффективность предлагаемого подхода. В ходе экспериментов было произведено исследование свойств предложенной системы, что еще раз подтвердило целесообразность использования предложенной системы для решения широкого круга задача интеллектуального анализа данных.

Ключевые слова: нейро-фаззи система, вычислительный интеллект, функция принадлежености, нечеткая кластеризация, нео-фаззи нейрон, рангова шкала.

REFERENCES

1. Bezdek J. C. Pattern Recognition with Fuzzy Objective Function Algorithms. N.Y., Plenum Press, 1981, 272 p.
2. Jang J.-Sh. R., Sun Ch.-T., Mizutani E. Neuro-Fuzzy and Soft Computing, Upper Saddle River, NJ, Prentice Hall, 1997, 614 p.
3. Hoeppner F., Klawonn F., Kruse R. Fuzzy Cluster Analyse. Braunschweig, Vieweg, 1999, 280 p.
4. Nelles O. Nonlinear System Identification. Berlin, Springer, 2001, 785 p.
5. Yamakawa T., Uchino E., Miki T., Kusanagi H. A neo-fuzzy neuron and its applications to system identification and prediction of the system behavior, *IIZUKA – 92: 2nd Int. Conf. on Fuzzy Logic and Neural Networks, Iizuka*. Japan, 1992, proceedings, pp. 477–483.
6. Miki T., Yamakawa T. Analog implementation of neo-fuzzy neuron and its on-boarding learning, *Ed. by N. S. Mastorakis «Computational Intelligence and Application»*. Piraeus, WSES Press, 1999, pp. 144–149.

7. Bodyanskiy Ye., Gorshkov Ye., Kolodyazhniy V., Otto P. Evolving network based on double neo-fuzzy neurons, *Computer Science Meets Automation: 52nd Int. Scientific Coll., TU Ilmenau (Thuer), 2007*, proceedings, pp. 35–40.

8. Brouwer R. K. Fuzzy Relational Fixed Point Clustering, *International Journal of Computational Intelligence Systems,* 2009, Vol. 2, Iss.1, pp. 69–82.

9. Brouwer R. K. Clustering feature vectors with mixed numerical and categorical attributes, *International Journal of Computational Intelligence Systems,* 2008, Vol. 1, Iss. 4, pp. 285–298.

10. Brouwer R. K., Groenwold A. Fuzzy relational clustering based on comparing two proximity matrices with utilization of particle swarm optimization, *Soft Computing,* 2009,Vol. 13, Iss. 6, pp. 577–589.

11. Brouwer R. K. Fuzzy clustering of feature vectors with some ordinal valued attributes using gradient descent for learning, International Journal of Uncertainty, *Fuzziness and Knowledge-Based Systems*, 2008, Vol. 16, No. 2, pp. 195–218.

12. Butkiewicz B. S. Robust fuzzy clustering with fuzzy data, *Lecture Notes in Computer Science*, V.3528. Berlin-Heidelberg, Springer-Verlag, 2005, pp. 76–82.

13. Lee M., Brouwer R. K. Fuzzy Clustering and Mapping of Ordinal Values to Numerical, *IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007),* 2007, pp. 538–543.

14. Lee M., Brouwer R. K. Likelihood Based Fuzzy Clustering for Data Sets of Mixed Features, *IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007),* 2007, pp. 544-549.

15. Brouwer R. K. Fuzzy set covering of a set of ordinal attributes without parameter, *Fuzzy Sets and Systems,* 2006, Vol. 157, No. 13, pp. 1775–1786.

16. Lbov G. S. Methods of processing different-type experimental data. Novosibirsk, Nauka, 1985, 160 p. (in Russian)

17. Brouwer R. K., Pedrycz W. A feedforward neural network for mapping vectors to fuzzy sets of vectors, *Processing ICANN/ICOMIP 2003: Int. Conf. on Artificial Neural Networks and Neural Information, Istanbul.* Turkey, 2003, pp. 45–48.

18. Bodyanskiy Ye. V., Opanasenko V. A., Slipchenko A. N. Fuzzy clustering for ordinal data, *Sistemy Obrobky Informacii,* 2007, Iss. 4(62), pp. 5–9. (in Russian)

19. Bodyanskiy Ye. V., Pliss I. P., Solovyova T. V. Multistep optimal predictors of multidimensional nonstationary stochastic processes, *Doklady AN USSR*, 1986, Iss. A, №12, pp. 47–49. (in Russian)

20. Kaczmarz S. Approximate solution of systems of linear equations, *Int. J. Control*, 1993, No. 53, pp. 1269–1271.

21. Widrow B., Hoff Jr. M. E. Adaptive switching circuits, *1960 IRE Western Electric Show and Connection Record*, 1960, Part 4, pp. 96–104.

22. Goodwin G. C., Ramadge P. J., Caines P. E. Discrete time stochastic adaptive control, *SIAM J. Control and Optimization*, 1981, No. 19, pp. 829–853.

23. Goodwin G. C., Ramadge P. J., Caines P. E. A globally convergent adaptive predictor, *Automatica*, 1981, No. 17, pp. 135–140.

24. Bodyanskiy Ye.V. Adaptive algorithms for identification of nonlinear control objects, *ASU and automation devices*. Kharkov, Vyshcha skola, 1987, Iss. 81, pp. 43–46. (in Russian)

25. Bodyanskiy Ye., Kolodyazhniy V., Stephan A. An adaptive learning algorithm for neuro-fuzzy network, Ed. by B. Reusch «Computational Intelligence. Theory and Applications». Berlin-Heidelberg, New York, Springer, 2001, pp. 68–75.

26. UCI machine learning repository [Electronic resource]. Access mode: http://archive.ics.uci.edu/ml/datasets/.

27. Bodyanskiy Ye., Tyshchenko O., Kopaliani D. Adaptive learning of an evolving cascade neo-fuzzy system in data stream mining tasks, *Evolving Systems*, 2016, Vol. 7, No. 2, pp. 107–116.

28. Bodyanskiy Ye., Tyshchenko O., Kopaliani D. An Extended Neo-Fuzzy Neuron and its Adaptive Learning Algorithm, *I. J. Intelligent Systems and Applications (IJISA)*, 2015, Vol. 7, No. 2, P. 21–26.

29. Bodyanskiy Ye., Tyshchenko O., Kopaliani D. A hybrid cascade neural network with an optimized pool in each cascade, *Soft Computing. A Fusion of Foundations, Methodologies and Applications (Soft Comput),* 2015, Vol. 19, Iss. 12, pp. 3445–3454.

30. Zhengbing Hu, Yevgeniy V. Bodyanskiy, Oleksii K. Tyshchenko, Olena O. Boiko An Evolving Cascade System Based on a Set of Neo-Fuzzy Nodes, *International Journal of Intelligent Systems and Applications (IJISA),* 2016, Vol. 8, No. 9, pp. 1–7.

31. Bodyanskiy Ye., Tyshchenko O., Boiko O. An evolving cascade system based on neuro-fuzzy nodes, *Radio Electronics, Computer Science, Control*, 2016, No. 2(37), pp. 40–45. (in Ukrainian)

32. Bodyanskiy Ye., Tyshchenko O., Kopaliani D. A Multidimensional Cascade Neuro-Fuzzy System with Neuron Pool Optimization in Each Cascade, *Int. Journal of Information Technology and Computer Science (IJITCS)*, 2014, Vol. 6, No. 8, pp. 11–17.