ISSN 0146-4116, Automatic Control and Computer Sciences, 2017, Vol. 51, No. 6, pp. 391–398. © Allerton Press, Inc., 2017.

Diagnostic Neuro-Fuzzy System and Its Learning in Medical Data Mining Tasks in Conditions of Uncertainty about Numbers of Attributes and Diagnoses[1]

Iryna Pliss and Iryna Perova*

Kharkiv National University of Radio Electronics, Kharkiv, 61166 Ukraine

**e-mail: rikywenok@gmail.com*

Received November 14, 2016

Abstract—Architecture and learning method for evolving diagnostic neuro-fuzzy-system for Medical Data Mining tasks in situation of uncertainty about quantities of attributes and diagnoses are proposed. Diagnostic neuro-fuzzy-system was approbated on data set, which present erosive ulcerous disease of the gastrointestinal tract and shown high quality of classification in condition of different quantity of input and output data.

Keywords: Medical Data Mining, classification, fuzzification, neuro-fuzzy system, learning algorithm, membership function

DOI: 10.3103/S0146411617060062

1. INTRODUCTION

For various Data Mining tasks, connected with diagnostics, classification, clusterization, pattern recognition etc. nowadays methods of Computational Intelligence, firstly Soft Computing and Machine Learning [1–8] are widely used.

Ones of the most effective now are neuro-fuzzy systems due to their learning abilities, including self-learning, universal approximative capacities, linguistic interpretability and "transparency" of results. ANFIS and TSK-systems of different orders, like approximators and extrapolators, and NEFCLASS [9] with its different modifications, oriented for classification (pattern recognition) tasks solving have the widest spread.

However, there exists a wide class of tasks where these systems are not sufficient. Primarily, there are the tasks where training set is short, data sets are fed to processing sequentially, in the form of data stream [10] and learning of system has to realize in parallel with analysis of input information.

This situation often appears in Medical Data Mining tasks [11, 12] and is complicated by the fact that data set under processing is nonstationary and dimensionality of input features space can be comparable with size of training data set. When it comes to diagnosis task, firstly, data set can have very small size in situation of rare diagnosis, and secondly, quantity of possible diagnosis (especially in situation of screening programs) can change during analysis. Naturally, that traditional diagnostic neuro-fuzzy system like NEFCLASS cannot overcome these problems.

2. FAST DIAGNOSTIC NEURO-FUZZY SYSTEM

Let's consider architecture of diagnostic neuro-fuzzy system (DNFS) that consists of six sequentially connected layers (Fig. 1) [13]. Here $(n \times 1)$ input vector of signals-attributes $x(k) = (x_1(k), x_2(k), \ldots x_n(k))^T \in R^n$, where $k = 1, 2, \ldots$ is current time, is fed to input layer of system. First hidden layer of system contains nh membership functions $\mu_{li}(x_i(k))$, $i = 1, 2, \ldots, n$; $l = 1, 2, \ldots, h$ and provides fuzzification of input feature space.

Because of this in system scatter partitioning of features space is realized as a membership functions standard bell shape constructions with unlimited supports are used. Most often, they are traditional Gaussians or more exotic functions, for example, derivatives of tangent hyperbolic function.

[1] The article is published in the original.

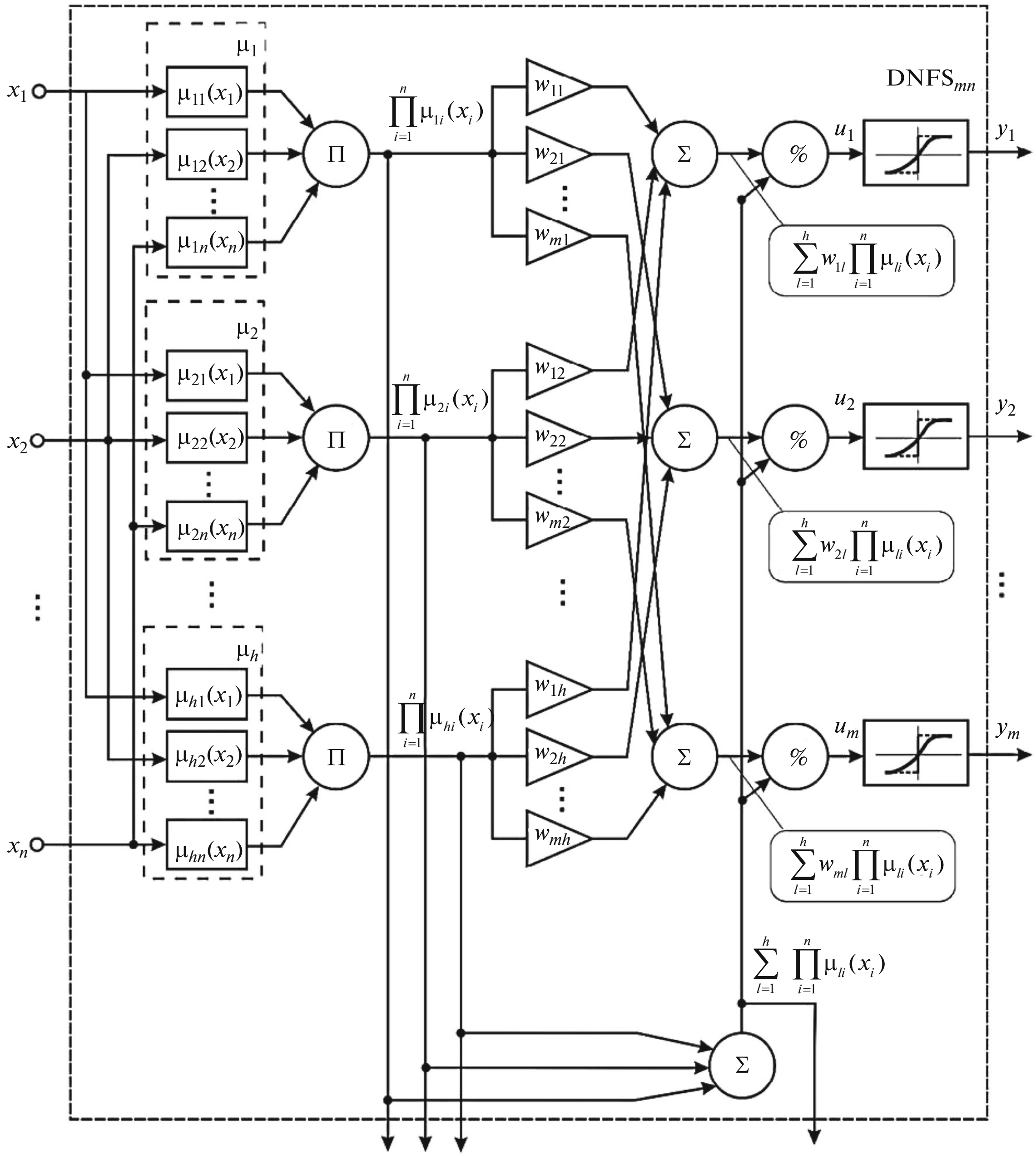

Fig. 1. Diagnostic neuro-fuzzy system DNFS$_{nm}$ with n inputs and m outputs.

Second hidden layer realizes aggregation of membership levels that were calculated in first layer and consists of h usual multipliers. Third hidden layer is a layer of synaptic weights w_{jl} ($j = 1,2,\ldots,m$ – number of possible diagnosis taken on the basis of empirical consideration) which have to be adjusted in training process. It is the most "responsible" layer of DNFS because effectiveness of whole system depends of precision and speed of training.

Common quantity of synaptic weights equals to mh. Fourth hidden layer is formed by $m+1$ adders, which realize elementary operations. In fifth hidden layer that is formed by m division units defuzzification of "gravity center" type is realized. At last output, (sixth) layer contains m nonlinear activation functions. In diagnostics task simple signum function is often used, which takes value $+1$ in case of true diagnosis and -1 – otherwise. That's why output signal of DNFS $y_j(k)$ can take only two values ± 1.

When feature vector $x(k) \in R^n$ is fed to the input of system,

in output of first hidden layer – *hn* values of $\mu_{li}(x_i(k))$ appear;

in output of second hidden layer – *h* signals $\prod_{i=1}^{n} \mu_{li}(x_i(k))$,

in output of third hidden layer – *mh* values $w_{jl}\prod_{i=1}^{n} \mu_{li}(x_i(k))$,

output of fourth layer – *m* + 1 signals: $\sum_{l=1}^{h} w_{jl}\prod_{i=1}^{n} \mu_{li}(x_i(k))$ and $\sum_{l=1}^{h}\prod_{i=1}^{n} \mu_{li}(x_i(k))$,

fifth layer –

$$u_j(k) = \frac{\sum_{l=1}^{h} w_{jl}\prod_{i=1}^{n} \mu_{li}(x_i(k))}{\sum_{l=1}^{h}\prod_{i=1}^{n} \mu_{li}(x_i(k))} = \sum_{l=1}^{h} w_{jl}\frac{\prod_{i=1}^{n} \mu_{li}(x_i(k))}{\sum_{l=1}^{h}\prod_{i=1}^{n} \mu_{li}(x_i(k))} = \sum_{l=1}^{h} w_{jl}\phi_l(x(k)) = w_j^T\phi(x(k))$$

and sixth – *m* diagnostics signals $y_j(k) = \text{sgn } u_j(k)$.

So, system under consideration is a modification of neuro-fuzzy system of Wang–Mendel [14] and intended for solving of diagnostic-classification tasks.

For training of this system in [13] traditional criterion of pattern recognition neural network was used [15]:

$$E_j(k) = e_j(k)u_j(k) = \left(d_j(k) - \text{sgn} w_j^T\phi(x(k))\, w_j^T\phi(x(k))\right) = d_j(k)u_j(k) - |u_j(k)| \quad (1)$$

and algorithm for synaptic weights matrix tuning:

$$W(k+1) = W(k) + \frac{(d(k) - \text{sgn } W(k)\phi(x(k)))}{\eta + \|\phi(x(k))\|^2}\phi^T(x(k)), \quad (2)$$

where $W(k) = \begin{pmatrix} w_1^T(k) \\ w_2^T(k) \\ \cdot \\ \cdot \\ w_m^T(k) \end{pmatrix}$ – $(m \times h)$-synaptic weights matrix;

$\text{sgn } u(k) = (\text{sgn } u_1(k), \text{sgn } u_2(k), \ldots, \text{sgn } u_n(k))^T$;

$d(k) = (d_1(k), d_2(k), \ldots, d_m(k))^T$ – reference signals vector, taking only two values ± 1;

$\eta \geq 0$ – momentum term.

Easy to see, that when $\eta = 0$ algorithm (2) can be rewritten in simple form

$$W(k+1) = W(k) + (d(k) - \text{sgn } W(k)\phi(x(k)))\phi^+(x(k)), \quad (3)$$

where $(\cdot)^+$ – symbol of pseudoinversion.

Elementary analysis of (1)–(3) shows, that training error $e_j(k)$ can take only three values: –1, 0, +1, that's mean the training process has oscillatory "bang-bang" nature. It can lead to its delaying and in situation when training is realized in tandem with processing in online mode these oscillations may never stop.

To exclude these oscillations we can introduce in sixth layer (instead of signum functions) activation function of hyperbolic tangent type that are often used in neural networks:

$$y_j(k) = \tanh \gamma u_j(k) = \frac{1 - e^{-2\gamma u_j(k)}}{1 + e^{-2\gamma u_j(k)}},$$

where gain parameter γ increasing leads to approaching of function $\tanh \gamma u_j$ to $\text{sgn}\, u_j$ without derivative discontinuity.

Using standard quadratic criterion of training

$$E_j(k) = \frac{1}{2}e_j^2(k) = \frac{1}{2}\left(d_j(k) - \tanh\gamma w_j^T\phi(x(k))\right)^2 = \frac{1}{2}\left(d_j(k) - \tanh\gamma u_j(k)\right)^2$$

we can write δ-rule of Rosenblatt's perceptron training

$$w_j(k+1) = w_j(k) + \eta(k)e_j(k)\gamma\left(1 - y_j^2(k)\right)\phi(x(k)) = w_j(k) + \eta(k)\delta_j(k)\phi(x(k)), \quad (4)$$

where $\eta(k) > 0$ – learning rate parameter, $\delta_j(k)$ – δ-error of training for jth output at kth time iteration.

Using ideas of quasi-Newtonian optimization [16] we can introduce optimized variation of (4) [17]:

$$w_j(k+1) = w_j(k) + \frac{\delta_j(k)\varphi^T(x(k))}{\eta + \|\varphi(x(k))\|^2}$$

or in matrix form:

$$W(k+1) = W(k) + \frac{\delta(k)\phi^T(x(k))}{\eta + \|\phi(x(k))\|^2}.$$

In situation when $\eta = 0$ we can write

$$W(k+1) = W(k) + \delta(k)\phi^+(x(k)), \quad (5)$$

where $\delta(k) = (\delta_1(k), \delta_2(k), \ldots, \delta_m(k))^T$;

$$\delta_j(k) = e_j(k)\gamma\left(1 - y_j^2(k)\right) = \left(d_j(k) - \tanh\gamma u_j(k)\right)\gamma\left(1 - \left(\tanh\gamma u_j(k)\right)^2\right).$$

By selecting of tuning gain parameter γ we can obtain necessary character of learning process convergence. Also, conspicuously, training algorithm (5) is very simple in numeric implementation.

3. EVOLVING DIAGNOSTIC NEURO-FUZZY SYSTEM

Diagnostic system under consideration is designed for using in conditions, when quantity of diagnostic features n and diagnosis quantity m is fixed, that is natural for neural networks and neuro-fuzzy systems, whose architecture is set a priori during synthesis.

In real medical tasks during training new diagnosis can appear, that was not previously involved. To enlarge quantity of possible diagnosis we can use ideas of evolving systems of computational intelligence [18, 19], that can tune their parameters and architecture. Architecture of evolving system $\text{DNFS}_{n,\, m+1}$ with n inputs and $m + 1$ outputs is shown in Fig. 2.

It is based on DNFS_{nm} system, shown on Fig. 1, where neuro-fuzzy-element NFE, that contains h synaptic weights $w_{m+1,l}$, one adder (summation block), one divider and activation function $\tanh\gamma w_{m+1}^T\phi(x(k))$ was introduced.

Rearranging training algorithm (5) for DNFS_{nm} in the form

$$W^m(k+1) = W^m(k) + \delta^m(k)\phi^+(x(k)),$$

we can introduce algorithm for $\text{DNFS}_{n,\, m+1}$:

$$W^{m+1}(k+1) = \begin{pmatrix} W^m(k+1) \\ ------ \\ w_{m+1}^T(k+1) \end{pmatrix} = \begin{pmatrix} W^m(k) \\ ---- \\ w_{m+1}^T(k) \end{pmatrix} + \begin{pmatrix} \delta^m(k) \\ ---- \\ \delta_{m+1}(k) \end{pmatrix}\phi^+(x(k)). \quad (6)$$

It's easy to see, that including of new NFE blocks in extended diagnostic system does not change original DNFS_{nm} training.

In real situations dimension of input vector can grow together with quantity of diagnoses. On Fig. 3 architecture of $\text{DNFS}_{n+1,\, m}$ based on DNFS_{mn} (Fig. 1) is presented. This system has one more input signal $x_{n+1}(k)$, that's mean that input vector became the form $x(k) = (x_1(k), \ldots, x_n(k), x_{n+1}(k))^T \in R^{n+1}$. Easy to see that system evolution is realized in first hidden layer, where membership functions

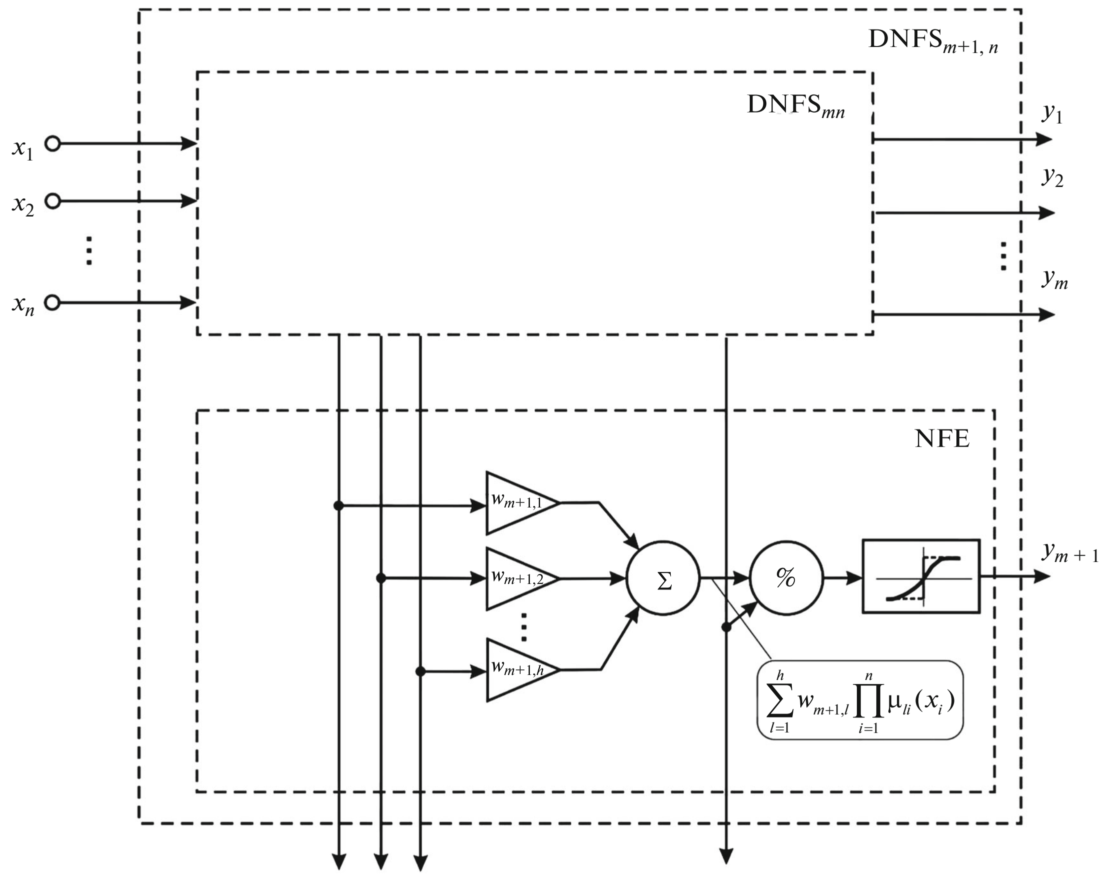

Fig. 2. Evolving diagnostic neuro-fuzzy system with n inputs and $m + 1$ outputs ($\mathrm{DNFS}_{n,\,m+1}$).

$\mu_{1,n+1}(x_{n+1}), \mu_{2,n+1}(x_{n+1}), \ldots, \mu_{h,n+1}(x_{n+1})$ were introduced. All other layers keep constant (unchangeable) and, it's very important, quality of tuned synaptic weights and its learning algorithms does not change.

By analyzing the architectures shown on Figs. 2 and 3 it's easy to conclude, that sequential diagnostic process must be started using few diagnoses and attributes, that can be increase during training process (evolution process). So system under consideration can grow its architecture until attainment of required quality of obtained results.

4. EXPERIMENT

Nowadays erosive ulcerous disease of the gastrointestinal tract is an actual problem of medicine because of increasing number of its pathology at young peoples in whole world. This disease has high and stable level of total lethality. Actuality of this pathology was confirmed by any research those results was discussed at 31-th Congress of Societies of critical medicine (San-Diego 2002), Congress of Anesthesiologists-Resuscitators of Southern Federal District of Russia (2005) [21]. But all this discussion does not solve the problem of syndrome of stomach acute lesion development mechanisms and schemes of medical drugs therapy.

Syndrome of stomach acute lesion has any special capacities, like:

—organs of gastrointestinal tract often were not influenced directly by affecting factor;

—development of this syndrome take a lot of time;

—any of patients has clinical manifestation of stress [22].

So all these patients have different quantity of input parameters. Generally, input information composes of 180 patients witch characterized from 50 to 54 symptoms. These symptoms consist of clinical, mechanical and biochemical blood and blood plazma parameters, parameters of blood circulation, blood pressure, parameters of drug therapy [23, 24]. These patients devised by 3 groups correspond to 3 type's

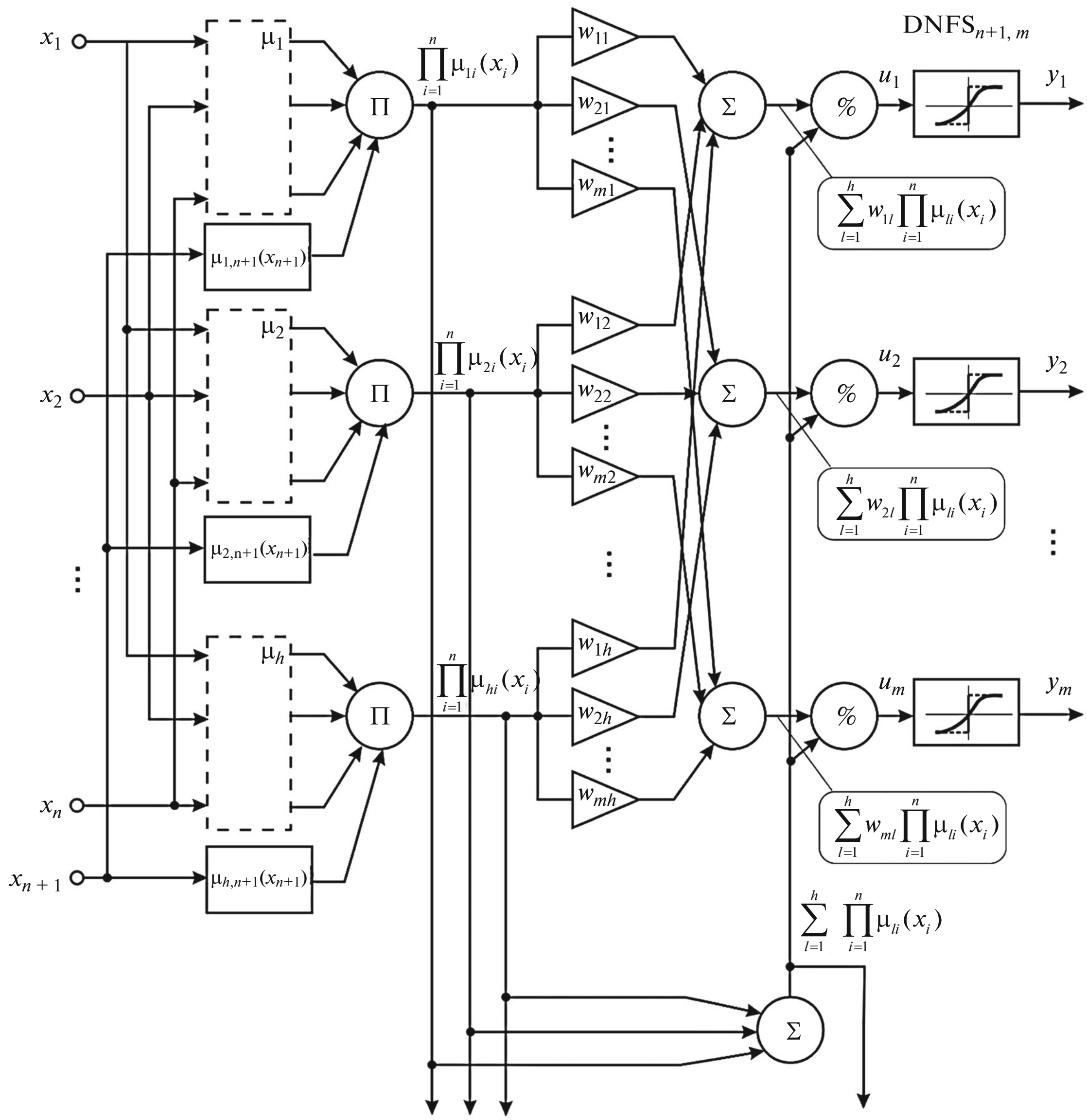

Fig. 3. Diagnostic neuro-fuzzy system DNFS$_{n+1,m}$ with $n + 1$ inputs and m outputs.

genesis of syndrome of stomach acute lesion: 1st group – acute gastrointestinal bleeding of I–II types ulcer genesis; 2nd group – acute pancreatitis nutritional origin; 3rd group – acute intestinal obstruction primary cancer genesis. Patients of 1st and 2nd group characterized by 50 symptoms, patients of 3rd group characterized by 54 symptoms. Each group contains patients with different type of drug therapy (4 types). So general quantity of output information is 12.

Input information can be presented using principal component analysis (Fig. 4) in the space of three first principal components, where we can see 3 types-clusters of syndrome of stomach acute lesion and each cluster contains 4 mini-clusters.

To process all these patients by system of computational intelligence we must delete some symptoms from patients of group 3 to make all input data set with same parameters. Then input data was divided to training data set (consist of 150 patients) and testing data set (consist of 30 patients). Comparative analysis of DNFS$_{mn}$, k-means clustering algorithm presents in Table 1.

Easy to see that DNFSmn realizes more qualitative classification on training and testing data sets.

But proposed clustering-classification was made in condition of incomplete input parameters. To process all parameters in whole system we need to use evolving system those can change quantity of inputs

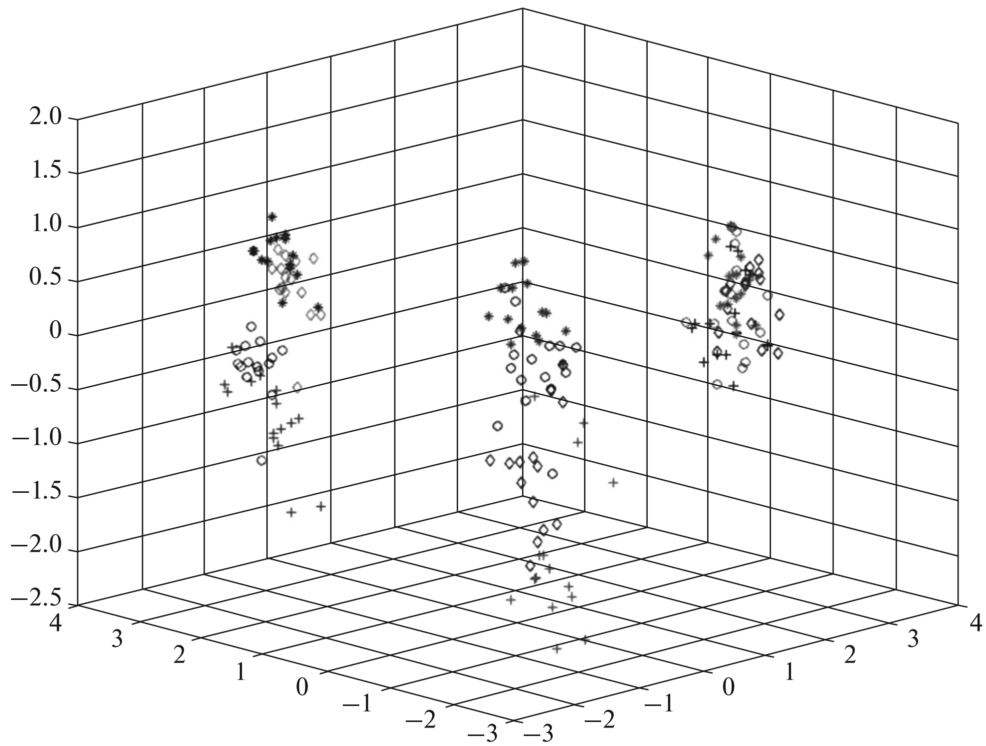

Fig. 4. Visualization of input information.

and outputs like $\mathrm{DNFS}_{m,\ n\ +\ 1}$. In Table 2 was presented results of $\mathrm{DNFS}_{m,\ n\ +\ 1}$ classification in training and testing data sets on 50 input parameters and training and testing data sets on 54 input parameters (evolving process).

Capacity of $\mathrm{DNFS}_{m\ +\ 1,\ n}$ was approbating on proposed data set. At first step, system was trained on 11 outputs (group 1, group 2 and group 3 without 12-th diagnosis – 90 patients) and checked its work on 20 patients. After that at system was introduced 12-th output using (6) and system was trained again with 12 outputs (50 patients) and checked its work on 20 patients. The analysis of classification results presented in Table 3.

Easy to see high quality of classification results on train and test sets. We can note that other known systems cannot make this classification without retraining its parameters on new quantity of outputs.

Table 1. The comparative analysis of clustering-classification results

Neural network	Error of clustering- classification, train set	Error of clustering- classification, test set
DNFS_{mn}	0.55%	10%
k-Means clustering algorithm	42.76%	47.22%

Table 2. The analysis of classification results

Neural network	Error of classification (50 inputs), train set	Error of classification (50 inputs), test set	Error of classification (54 inputs), train set	Error of classification (54 inputs), test set
$\mathrm{DNFS}_{m,n+1}$	1.11%	9.8%	2.66%	16.25%

Table 3. The analysis of classification results

Neural network	Error of classification (11 outputs), train set	Error of classification (11 outputs), test set	Error of classification (12 outputs), train set	Error of classification (12 outputs), test set
$\mathrm{DNFS}_{m+1,n}$	0%	2.24%	6.66%	15.84%

5. CONCLUSION

In this paper architecture and learning method for evolving diagnostic neuro-fuzzy-system are proposed. This system is designed for broad class of Data Stream Mining tasks, especially Medical Data Mining ones in online mode in situations of unknown quantity of possible diagnoses that can change during training-diagnostics process. Proposed system is simple in numeric realization and characterized by a high learning rate, that make possible to use it in conditions of small training sets and on big data sets, coming to processing in online mode. Diagnostic neuro-fuzzy system $\mathrm{DNFS}_{n+1, m+1}$ with $n + 1$ inputs and $m + 1$ outputs was approbate on real medical data set and shown high level of classification capacities.

REFERENCES

1. Jang, J.-S.R., in *Neuro-fuzzy and Soft Computing,* Jang, J.-S.R., Sun, C.-T., and Mizutani, E., Eds., Prentice Hall, Upper Saddle River, NJ, 1997.
2. Friedman, J., in *The Elements of Statistical Learning. Data Mining, Inference and Prediction,* Friedman, J., Hastie, T., and Tibshirani, R., Eds., Berlin: Springer, 2003.
3. Osowski, S., *Sieci neuronowe do przetwarzania informacji*, Warszawa: Oficina Wydawnicza Politechniki Warszawskiej, 2006.
4. Ruthowski, L., *Computational Intelligence. Methods and Techniques,* Berlin-Heidelberg, 2008.
5. Mumford, C.L. and Jain, L.C., *Computational intelligence. Collaboration, fusion and Emergence,* Berlin: Springer-Verlag, 2009.
6. Kruse, R., Borgelt, C., Klawonn, F., Moewes, C., Steinbrecher, M., and Held, P., *Computational Intelligence. Methodological Introduction,* Berlin: Springer-Verlag, 2013.
7. Kroll, A., *Computational Intelligence. Eine Einführung in Probleme, Methoden und technische Anwendungen,* Munchen: Oldenbourg Verlag, 2013.
8. Du, K.-L. and Swami, M.N.S., *Neural Networks and Statistical Learning,* London: Springer-Verlag, 2014.
9. Nauck, D. and Kruse, R., NEFCLASS—a neuro-fuzzy approach for the classification of data, *Proc. of the 1995 ACM Symposium on Applied Computing, "Applied Computing 1995,"* George, K.M., Ed., Nashville, 1995, 461–465.
10. Bifet, A., *Adaptive Stream Mining. Pattern Learning and Mining from Evolving Data Streams,* Amsterdam: IOS Press, 2010.
11. Rizzo, R., Computational intelligence methods for bioinformatics and biostatistics, *7th International Meeting, CIBIB 2010, Palermo, Italy, September 16–18, 2010,* Rizzo, R., Eds., Springer, 2011.
12. Kountchev, R., *Advances in Intelligent Analysis of Medical Data and Decision Support Systems* (Studies in Computational Intelligence), Kountchev, R., et al., Eds., Springer-Verlag, Berlin Heidelberg, 2013.
13. Bodyanskiy, Ye.V., Vynokurova, O., Mulesa, P.P., and Perova, I.G., Diagnostic neuro-fuzzy system and its adaptive learning algorithm in the medical and biological data mining tasks, *Syst. Technol.,* 2014, no. 2 (91).
14. Wang, L.X. and Mendel, J.M., Fuzzy basis functions, universal approximation, and orthogonal least squares learning, *IEEE Trans. Neural Network,* 1992, no. 3, pp. 807–814.
15. Shynk, J.J., Performance surfaces of a single-layer perceptron, *IEEE Trans. Neural Networks,* 1990, no. 1, pp. 268–274.
16. Shepherd, A.J., *Second-Order Methods for Neural Networks,* London: Springer-Verlag, 1997.
17. Otto, P., Bodyanskiy, Ye., and Kolodyazhniy, V., A new learning algorithm for forecasting neural-fuzzy network, *Integr. Comput.-Aided Eng.,* 2003, vol. 10, no. 4, pp. 399–409.
18. Kasabov, N., *Evolving Connectionist Systems: Methods and Applications in Bioinformatics, Brain Study and Intelligent Machines,* London: Springer-Verlag, 2003.
19. Lughofer, E., *Evolving Fuzzy Systems—Methodologies, Advanced Concept and Applications,* Berlin-Heidelberg: Springer-Verlag, 2011.
20. Perova, I. and Bodyanskiy, Ye., Fast evolving diagnostic neuro-fuzzy system and its learning in medical data mining tasks, *Syst. Technol.,* 2015, no. 6 (101), pp. 109–116.
21. Gelfand, B.R., *Sepsis in the beginning of XXI century*, in *Guidelines,* Moscow, 2004.
22. Rauws, E.A.J., Langenberg, W., Houthoff, H.J., Zanen, H.C., and Tytgat, G.N.J., *Campylobacter pyloridis*-associated chronic antral gastritis. A prospective study of its prevalence and the effects of antibacterial and antiulcer treatment, *Gastroenterol.,* 2005, vol. 94, pp. 33–40.
23. Bodyanskiy, Ye., Pavlov, O., and Vynokurova, O., in *Outliers Resistant Learning Algorithm for Radial-basis-fuzzy-wavelet-neural Network in Stomach Acute Injury Diagnosis Tasks,* Markov, K., Ivanova, K., and Mitov, I., Eds., Sofia: Institute of Information Theories and Application FOI ITHEA, 2008.
24. Boyko, V.V., Sushkov, S.V., Pavlov, O.O., and Vynokurova, O.A., *The Anesthesia and Intensive Therapy Strategy in the Acute Gastro Ileum Hemorrhage,* Kharkiv: Eksklusiv, 2007.