www.ijaceeonline.com ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 7, July 2017, pp. 1–10

Medical Online Neuro-Fuzzy Diagnostics System with Active Learning

Dr. Yevgeniy Bodyanskiy
Scientific Head of Control Systems Research Laboratory,
Professor of Artificial Intelligence Department, Prof., Dr. Sc. (Dr.-Ing. habil.),
Kharkiv National University of Radio Electronics, 14 Nauky av., Kharkiv, Ukraine 61166
Email: yevgeniy.bodyanskiy@nure.ua

Dr. Iryna Perova
Associate professor of Biomedical engineering department, PhD,
Kharkiv National University of Radio Electronics, 14 Nauky av., Kharkiv, Ukraine 61166
Email: rikywenok@gmail.com

Abstract: *Situations when in the medical data set some patients have known diagnoses and all other have unknown ones is spread wise problem of present-day medicine. Known systems of computational intelligence show mediocre level of diagnostics in these data sets. In this paper online neuro-fuzzy diagnostics system with active learning is proposed. This system allows to increase a quality of medical diagnostics under the condition of small number of known reference signals due to combination of special learning algorithms – active learning. The proposed online neuro-fuzzy system is based on popular neural networks as Self-Organizing Map (SOM) and Learning Vector Quantization network (LVQ). Active learning procedure permits to tune their synaptic weights using simple recurrent self-learning procedures (SOM) and controlled learning with teacher (LVQ). Neuro-fuzzy diagnostics system with active learningwas used for breast cancer in Wisconsin data set processing and showed higher level of classification-clusterization results comparatively with known systems.*

Keyword: *Medical Data Mining; active learning; neuro-fuzzy system; medical diagnostics; computational intelligence; learning vector quantization; self-learning procedure.*

1. INTRODUCTION

Nowadays, a computational intelligence systems [1-10] are widespread in medical diagnostics [11-14] due to its approximative capacities that permit to solve classification (pattern recognition) and clusterization tasks in conditions of different classes form that can be overlapped and possibility to learn (or self-learn) using experimental data that may be distorted by disturbances and gaps at the same time. A peculiarity of medical diagnostics is in situation when skilled physician for any group of patients can definitely diagnose any disease and for other one emerge situation when patients can have a few diagnoses (so physician can attribute for each diagnosis some subjective membership level).

And situations when physician cannot to diagnose patient disease due to experience deficiency or due to atypical or distorted patient features also may happens.

From computational intelligence position this situation is subject of active machine learning [15, 16], when diagnostic system during of its training processes observations with known reference signal, which connects each feature vector with specific diagnosis and any nonclassified observations (in self-learning mode). This task is complicated when classes-clusters are overlapped – situation of uncertainty (fuzzy situation) [17]. In this case problem of medical diagnostics transforms in an active learning task of some neuro-fuzzy system, which must tune its parameters in online-mode during the accumulation of experimental data.

Considering these requests this system may be based on popular neural networks as Self-Organizing Map (SOM) and Learning Vector Quantization network (LVQ), introduced by T. Kohonen [18]. These networks have single-layer architecture and permit to train their synaptic weights using simple recurrent self-learning algorithms (SOM) and control learning with teacher (LVQ).Section 2 describes SOM self-learning algorithms in the diagnostics tasks. The Section 3 describes the learning vector quantization learning in the diagnostics tasks. The Section 4 describes diagnostics neuro-fuzzy system with active learning and Section 5 discusses about experimental results.

Cite this paper:

Yevgeniy Bodyanskiy, Iryna Perova, "Medical Online Neuro-Fuzzy Diagnostics System with Active Learning", International Journal of Advances in Computer and Electronics Engineering, Vol. 2, No. 7, pp. 1-10, July 2017.

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

2. SOM SELF-LEARNING IN DIAGNOSTICS TASK

SOM – is single-layer neural network with lateral connections in Kohonen layer between neurons – adaptive linear associators whose synaptic weights describe prototype-centroids of clusters that are formed. When feature vector $x(k)=\left(x_1(k),x_2(k),...x_n(k)\right)^T \in R^n$ (here $k=1,2,...$ is number of this vector in table "object-property" or current discrete time) is fed into input of system, in outputs of each neuron appears the signal

$$y_j(k)=c_j^T(k)x(k),\ j=1,2,...,m \tag{1}$$

where m – quantity of possible classes-clusters that can be change during the training, $c_j(k)$ – weights-centroids vector on $k-$th training step.

We can note, that if feature vector $x(k)$ previously was centered relatively mean as $\sum_{\tau=1}^{k} x(\tau)=0$ and normalized as $\|x(\tau)\|=1$, we can rewrite (1) in the form

$$y_j(k)=\cos\left(c_j(k),x(k)\right)=\cos\theta_j(k), \tag{2}$$

if $\|c_j(k)\|=1$, that is not difficult to realize during training process.

Self-learning process consists of three sequential stages: competition, cooperation and synaptic adaptation, that must be complete by fuzzy inference stage in situation of overlapped diagnoses.

Competition process begins by analysis of current feature vector $x(k)$ that is fed to inputs of all neurons. For each prototype-centroid $c_j(k)$ distance (usually Euclidean distance)

$$d^2\left(c_j(k),x(k)\right)=\|x(k)-c_j(k)\|^2, \tag{3}$$

is calculated, that for standartized vectors

$$\|x(k)\|^2=\|c_j(k)\|^2=1 \tag{4}$$

can be rewritten in the form

$$\begin{aligned} d^2\left(c_j(k),x(k)\right)&=2\left(1-c_j^T(k)x(k)\right)=\\ &=2\left(1-\cos\theta_j(k)\right). \end{aligned} \tag{5}$$

For competition process its possible to use metrics (5) or similarity measure

$$sim\left(c_j(k),x(k)\right)=\cos\theta_j(k). \tag{6}$$

Using (5), (6) neuron-winner $c_j^*(k)$ can be determined such that

$$d^2\left(c_j^*(k),x(k)\right)=\min_p d^2\left(c_p(k),x(k)\right) \tag{7}$$

or

$$\begin{aligned} sim\left(c_j^*(k),x(k)\right)&=\max_p sim\left(c_p(k),x(k)\right)=\\ &=\max_p \cos\theta_p(k). \end{aligned} \tag{8}$$

Then temporally omitting cooperation stage it's possible to introduce synaptic weights training procedure based on WTA-self-learning rule [18]:

$$c_j(k+1)=\begin{cases} c_j^*(k)+\eta(k)\left(x(k)-c_j^*(k)\right), \text{ if i-th} \\ \quad \text{neuron won in the competition} \\ c_j(k), \text{ otherwise.} \end{cases} \tag{9}$$

where $0<\eta(k)<1$ – learning rate parameter, that monotony decreases in time and chosen based on empirical reasons.

For providing condition (4) and simplifying choice of rate parameter $\eta(k)$ instead (9) it's more convenient to use modified procedure

$$c_j(k+1)=\begin{cases} \dfrac{c_j^*(k)+\eta(k)\left(x(k)-c_j^*(k)\right)}{\left\|c_j^*(k)+\eta(k)\left(x(k)-c_j^*(k)\right)\right\|}, \\ \quad \text{if i-th neuron won in the} \\ \quad\quad \text{competition} \\ \eta(k)=r^{-1}(k),\ \ r(k)=\alpha r(k)+1, \\ \quad 0\le\alpha\le 1, \\ c_j(k), \text{ otherwise.} \end{cases} \tag{10}$$

For improvement of data processing quality in SOM training, process cooperation stage and WTM-self-learning rule is introduced. At this stage neuron-winner determines local area of topological neighbourhood and all neurons in this area (except winner) can be tuned but at less measure. So, we can introduce neighbourhood function $\varphi(j,p)$ which depends of distance $d^2\left(c_j^*(k),c_p(k)\right)$ between win-

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

ner and all other neurons in Kohonenlayer $c_p(k)$, $p=1,2,...,m$.

Usually, $\varphi(j,p)$ – is bell-shaped function symmetric relatively $c_j^*(k)$ ($\varphi(j,j)=1$) and decreasing with increasing the distance $d^2\left(c_j^*(k),c_p(k)\right)$. As a rule this function is traditional Gaussian

$$\varphi(j,p)=\exp\left(-\frac{\left\|c_j^*(k)-c_p(k)\right\|^2}{2\sigma^2}\right) \tag{11}$$

(here σ^2 – parameter that defines size of topological neighbourhood), but it's possible to use any other kernel function [19]. In this case self-learning process is connected with gradient minimization of criterion

$$E_p^k=\sum_{\tau=1}^{k}\varphi(j,p)\left\|x(\tau)-c_p\right\|^2 \tag{12}$$

using recurrent procedure

$$\begin{gathered}c_p(k+1)=c_p(k)+\eta(k)\varphi(j,p)\left(x(k)-c_p(k)\right)\\ \forall\ p=1,2,...,m.\end{gathered} \tag{13}$$

Also it's convenient to rewrite WTM-self-learning rule (13) like (10) in the form

$$\begin{cases}c_p(k+1)=c_p(k)+\eta(k)\cdot\\ \cdot\varphi(j,p)\cdot\left(x(k)-c_p(k)\right)\cdot\\ \cdot\left(\left\|c_p(k)+\eta(k)\varphi(j,p)\left(x(k)-c_p(k)\right)\right\|\right)^{-1},\\ \eta(k)=r^{-1}(k),\\ r(k)=\alpha r(k-1)+1,\\ 0\le\alpha\le 1.\end{cases} \tag{14}$$

It's easy to organize WTM-self-learning process using instead of neighbourhood function similarity measure (6), which can be transformed in bell-shaped function that depends of distance (3), (5). So instead (13) we can write

$$\begin{gathered}c_p(k+1)=c_p(k)+\\ +\eta(k)\left[\cos\theta_p(k)\right]_+\left(x(k)-c_p(k)\right)\\ \forall\ p=1,2,...,m\end{gathered} \tag{15}$$

(that structurally coincides with INSTAR-learning rule of S.Grossberg [20]), and instead (14)

$$\begin{cases}c_p(k+1)=c_p(k)+\eta(k)\cdot\\ \cdot\left[\cos\theta_p(k)\right]_+\cdot\left(x(k)-c_p(k)\right)\cdot\\ \cdot\left(\left\|c_p(k)+\eta(k)\left[\cos\theta_p(k)\right]_+\left(x(k)-c_p(k)\right)\right\|\right)^{-1},\\ \eta(k)=r^{-1}(k),\\ r(k)=\alpha r(k-1)+1,\quad 0\le\alpha\le 1.\end{cases} \tag{16}$$

where $\left[\cos\theta_p(k)\right]_+=\max\left\{0,\cos\theta_p(k)\right\}$.

As it was noted, in several real situations clusters can be overlapped and each pattern can belong to several or to all clusters with different membership levels. This situation is a subject of fuzzy clustering analysis [21, 22] and related with minimization of self-learning criterion (goal function)

$$E_k=\sum_{\tau=1}^{k}\sum_{p=1}^{m}u_p^{\beta}(\tau)\left\|x(\tau)-c_p\right\|^2 \tag{17}$$

in presence of constrains

$$\sum_{p=1}^{m}u_p(\tau)=1\ \ \forall\ \ \tau=1,2,...,k, \tag{18}$$

$$0\le\sum_{\tau=1}^{k}u_p(\tau)\le k,\mathrm{p}=1,2,...,m. \tag{19}$$

Here $u_p(\tau)\in[0,1]$ – membership level of vector $x(\tau)$ to $p-$th cluster, β - fuzzifier, that defines boundaries bluring between clusters (usually $\beta=2$).

Minimization (17) with constrains (18), (19) is reduces to solving of standard task of quadratic programming that with $\beta=2$ can be written in the form

$$\begin{cases}u_p(\tau)=\dfrac{\left\|x(\tau)-c_p\right\|^{-2}}{\sum_{l=1}^{m}\left\|x(\tau)-c_l\right\|^{-2}},\\ c_p=\dfrac{\sum_{\tau=1}^{k}u_p^2(\tau)x(\tau)}{\sum_{\tau=1}^{k}u_p^2(\tau)}.\end{cases} \tag{20}$$

Expressions (20) describe batch procedure of data processing and considered SOMs realize sequential online processing. We can note that in [23] fuzzy SOM was introduced, but it works only in batch mode.

That's why we can use in this situation recurrent procedures of fuzzy clusterization [24, 25]. So online algorithm for fuzzy probabilistic clustering in adopting notations can be written in the form

$$\begin{cases} c_p(k+1)=c_p(k)+\eta(k)u_p^2(k)\left(x(k)-c_p(k)\right) \\ \qquad\qquad\qquad\qquad \forall\ p=1,2,\ldots,m. \\ u_p(k)=\dfrac{\left\|x(k)-c_p(k)\right\|^{-2}}{\sum_{l=1}^{m}\left\|x(k)-c_l(k)\right\|^{-2}} \end{cases} \tag{21}$$

where $u_p^2(k)$ in our case has the sense of neighbourhood function.

By considering (4) algorithm (21) can be rewritten in the form

$$\begin{cases} c_p(k+1)=c_p(k)+\eta(k)\cdot \\ \cdot u_p^2(k)\cdot\left(x(k)-c_p(k)\right)\cdot \\ \cdot\left(\left\|c_p(k)+\eta(k)u_p^2(k)\left(x(k)-c_p(k)\right)\right\|\right)^{-1}\ \forall \\ \quad \forall\ p=1,2,\ldots,m, \\ \eta(k)=r^{-1}(k), \\ r(k)=\alpha r(k-1)+1, \\ 0\le\alpha\le1, \\ 0\le u_p(k)=\left\|x(k)-c_p(k)\right\|^{-2}\cdot \\ \cdot\left(\sum_{l=1}^{m}\left\|x(k)-c_l(k)\right\|^{-2}\right)^{-1}\le1. \end{cases} \tag{22}$$

Let's note, that fuzzy clusterization can be realized not only using metrics (3), but also using similarity measure (6). In this case basing on algorithm (16) easy to introduce modification of procedure considered in [26] in the form

$$\begin{cases} c_p(k+1)=c_p(k)+\eta(k)\cdot\left[\cos\theta_p(k)\right]_+\cdot \\ \cdot\left(x(k)-c_p(k)\right)\cdot \\ \cdot\left(\left\|c_p(k)+\eta(k)\left[\cos\theta_p(k)\right]_+\left(x(k)-c_p(k)\right)\right\|\right)^{-1}\ \forall \\ \quad \forall\ p=1,2,\ldots,m, \\ \eta(k)=r^{-1}(k), \\ r(k)=\alpha r(k-1)+1, \\ 0\le\alpha\le1, \\ 0\le u_p(k)=\left[\cos\theta_p(k)\right]_+\left(\sum_{l=1}^{m}\left[\cos\theta_l(k)\right]_+\right)^{-1}\le1. \end{cases} \tag{23}$$

In situation when $c_p(k+1)=c_p(k)$ or when (it's the same) $\left[\cos\theta_p(k)\right]_+=0\ \forall\ p$ it is supposed that $x(k)$ is abnormal observation (outlier) or it does not belong to any possible diagnoses from our consideration.

3. LVQ-LEARNING IN DIAGNOSTICS TASK

Neural networks of learning vector quantization (LVQ) in contradiction to self-organization map can tune its weights in controlled learning mode with teacher, but they have the same architecture as a SOM. The main task, that is solved by LVQ is compact presentation of big data sets (indexes, features, symptoms) by limited centroids (prototypes of diagnoses) set $c_p(k)$, $p=1,2,\ldots,m$, that exactly describes original space R^n. And for each input vector, that was preprocessed like for the SOM, we can introduce neuron-winner $c_j^*(k)$ that corresponds to center of some class-diagnosis. It's clear that winner is chosen using conditions (7) or (8).

Here two possible situations under teacher control can be aroused:

- Input vector $x(k)$ and neuron-winner $c_j^*(k)$ belong to same Voronoy cell [27] or to same diagnosis.
- Input vector $x(k)$ and neuron-winner $c_j^*(k)$ belong to different Voronoy cells.

Then basic LVQ learning algorithm can be written in the form [18, 27]

$$c_j(k+1)=\begin{cases} c_j^*(k)+\eta(k)\left(x(k)-c_j^*(k)\right), \\ \quad \text{if } x(k) \text{ and } c_j^*(k) \text{ belong} \\ \quad \text{to the same Voronoy cell} \\ c_j^*(k)-\eta(k)\left(x(k)-c_j^*(k)\right), \\ \quad \text{if } x(k) \text{ and } c_j^*(k) \text{ belong} \\ \quad \text{to different Voronoy cells} \\ c_j(k), \text{ for neurons, that have} \\ \text{not won in } k\text{-th instant of time.} \end{cases} \tag{24}$$

Algorithm (24) hastransparentphysicalsense: ifneuron-winner $c_j^*(k)$ and vector $x(k)$ belong to the same diagnosis then centroid draws to input feature vector $x(k)$ to distance that determined by learning rate parameter $\eta(k)$, that can be chosen by analogue

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

of SOM. Ifneuron-winner $c_j^*(k)$ and vector $x(k)$ belong to different diagnoses then centroid-winner pushes off from the $x(k)$. All centroids who have not won at k -th instant of time stay constant.

For providing conditions (4) it's easy to introduce modification of (24) in the form

$$c_j(k+1)=\begin{cases}\dfrac{c_j^*(k)+\eta_s(k)\left(x(k)-c_j^*(k)\right)}{\left\|c_j^*(k)+\eta_s(k)\left(x(k)-c_j^*(k)\right)\right\|}, \\ \quad \text{if } x(k) \text{ and } c_j^*(k) \text{ belong} \\ \quad \text{to the same Voronoy cell} \\ \dfrac{c_j^*(k)-\eta_s(k)\left(x(k)-c_j^*(k)\right)}{\left\|c_j^*(k)-\eta_s(k)\left(x(k)-c_j^*(k)\right)\right\|}, \\ \quad \text{if } x(k) \text{ and } c_j^*(k) \text{ belong} \\ \quad \text{to different Voronoy cells} \\ c_j(k), \text{ for neurons, that have} \\ \quad \text{not won at } k\text{-th instant of time.} \\ \eta_s(k)=r^{-1}(k), \\ r(k)=\alpha r(k-1)+1, \quad 0\le\alpha\le 1. \end{cases} \tag{25}$$

Easy to see that first, third and fourth expressions (25) correspond to self-learning procedure (10), second recurrent expression (25) corresponds to repulsion of winner from input vector $x(k)$ with rate $\eta_s(k)$. Chosen of it we have to study more attentively.

Let's rewrite second expression (24) in the form

$$c_j(k+1)=c_j^*(k)-\eta(k)\left(x(k)-c_j^*(k)\right) \tag{26}$$

and introduce elementary transformations:

$$\begin{aligned} x^T(k)c_j(k+1)&=x^T(k)c_j^*(k)- \\ &-\eta(k)\|x(k)\|^2-\eta(k)x^T(k)c_j^*(k), \end{aligned} \tag{27}$$

$$\begin{aligned} \cos\left(c_j(k+1),x(k)\right)&=\cos\left(c_j^*(k),x(k)\right)- \\ &-\eta(k)\left(1+\cos\left(c_j^*(k),x(k)\right)\right) \end{aligned} \tag{28}$$

For providing of condition

$$\cos\left(c_j(k+1),x(k)\right)=\cos\left(c_p(k),x(k)\right) \tag{29}$$

that means that $x(k)$ is located in border of two neighbouring Voronoy cells with centroids $c_j(k+1)$ and $c_p(k)$, learning rate $\eta_L(k)$ has to be chosen using expression [28]

$$\begin{aligned}\eta_L(k)&=\frac{\cos\left(c_j^*(k),x(k)\right)-\cos\left(c_p(k),x(k)\right)}{\cos\left(c_j^*(k),x(k)\right)+1}= \\ &=\frac{\cos\left(c_j^*(k),x(k)\right)-\cos\left(c_p(k),x(k)\right)}{\cos\left(c_j^*(k),x(k)\right)+\cos\left(x(k),x(k)\right)}= \\ &=\frac{\cos\theta_j(k)-\cos\theta_p(k)}{\cos\theta_j(k)+1}.\end{aligned} \tag{30}$$

Then after one step of repulsion $x(k)$ will be belong to clusters with centroids $c_j(k+1)$ and $c_p(k+1)$ with the same membership levels (corresponding with (6) and second expression of (23))

$$u_j(k)=u_p(k)=0{,}5. \tag{31}$$

More complex situation appears using fuzzy classification based on LVQ. We have to note, that in [29] fuzzy vector quantization based on fuzzy c-means algorithm [17, 21] (20) was introduced, but during this procedure opinion of "teacher" (in our case physician) does not considered. Wherein to each patient can be match at once q $(1<q\le m)$ diagnoses with different subjective membership levels $u_p^*(k)$, $p=1,2,...,q$, that correspond to condition (18). If feature vector occurs to one of q indicated Voronoy cells procedure of drawing of q centroids $c_p(k)$ to $x(k)$ is realized using expression

$$\begin{cases} c_p(k+1)=c_p(k)+\eta(k)\cdot \\ \cdot u_p^*(k)\cdot\left(x(k)-c_p(k)\right)\cdot \\ \cdot\left(\left\|c_p(k)+\eta(k)u_p^*(k)\left(x(k)-c_p(k)\right)\right\|\right)^{-1}\ \forall \\ \forall\ p=1,2,...,q, \\ \eta(k)=r^{-1}(k), \\ r(k)=\alpha r(k-1)+1, \quad 0\le\alpha\le 1. \end{cases} \tag{32}$$

In situation when $x(k)$ occurs to Voronoycell with centroid-winner $c_j^*(k)$ that does not related to diagnoses of "teacher", centroid $c_j^*(k)$ pushes off

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

from $x(k)$ using expression

$$\begin{cases} c_j(k+1) = c_j^*(k) - \eta(k) \cdot \\ \cdot \dfrac{\left(1-\cos\theta_j(k)\right)^{-1}}{\sum\limits_{l=1}^{m}\left(1-\cos\theta_l(k)\right)^{-1}} \cdot \left(x(k) - c_j^*(k)\right) \cdot \\ \cdot \left(\left\| c_j^*(k) - \eta(k) \dfrac{\left(1-\cos\theta_j(k)\right)^{-1}}{\sum\limits_{l=1}^{m}\left(1-\cos\theta_l(k)\right)^{-1}} \cdot \right. \right. \\ \left. \left. \cdot \left(x(k) - c_j^*(k)\right) \right\| \right)^{-1} \\ \eta(k) = r^{-1}(k), \\ r(k) = \alpha r(k-1) + 1, \quad 0 \le \alpha \le 1. \end{cases} \tag{33}$$

Multiplier $\left(1-\cos\theta_j(k)\right)^{-1}\left(\sum\limits_{l=1}^{m}\left(1-\cos\theta_l(k)\right)^{-1}\right)^{-1}$ means that if "alien" centroid-winner is closer to input vector than stronger it is pushing off from $x(k)$ according with fuzzy probabilistic membership level [17]:

$$u_j(k) = \frac{\left(1-\cos\theta_j(k)\right)^{-1}}{\sum\limits_{l=1}^{m}\left(1-\cos\theta_j(k)\right)^{-1}}. \tag{34}$$

It's clear that along the repulsion of incorrect centroid using (33), we can draw correct centroids $c_p(k)$ to $x(k)$ using procedure (32).

4. DIAGNOSTICS NEURO-FUZZY SYSTEM WITH ACTIVE LEARNING

Architectures of SOM and LVQ are very close each to other so we can introduce neuro-fuzzy system with active learning for online diagnostics of patients who has several diseases at same time. Architecture of proposed system is presented on Figure 1.

This system consists of two layers of information processing: first hidden Kohonen's layer is formed by adaptive linear assosiators N_p^K $(p = 1, 2, ..., m)$ and second layer – of membership fuzzy levels that are formed by m neurons N_p^M. Input signals in the form of current values vectors $x(k) = \left(x_1(k), x_2(k), ... x_n(k)\right)^T$ are sequentially fed to input of first hidden layer, where centroids $c_p(k)$ of formed classes-clusters are computed.

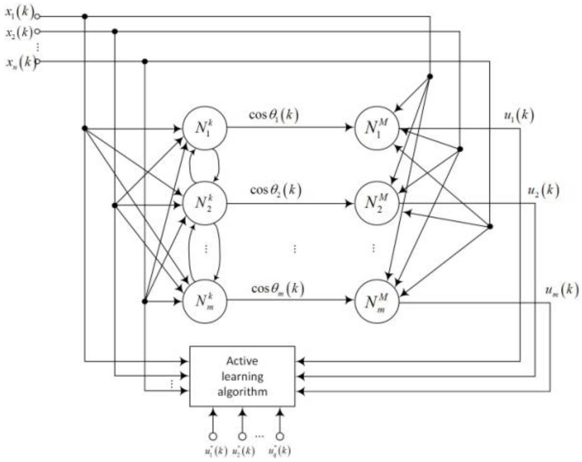

Figure 1 Neuro-fuzzy diagnostics system with active learning

Active learning block can operate in two modes: control learning with teacher and self-learning. In self-learning mode system does not need any additional information excepting input feature vectors and quantity of possible classes. Training process is realized using expression (22) or (23), which structurally coincide and differ only by neighbourhood functions. In control learning mode in system additional information from teacher in the form of subjective membership $u_p^*(k)$ $p = 1, 2, ..., q$ is fed (in the case of crisp diagnostics $q = 1$, $u_1^*(k) = 1$). Training process is realized using expressions (32), (33) which structurally are close to procedure (22), (23).

So, active learning process is realized by united architecture and structurally close learning algorithms, which are gradient procedures of adapted goal function minimization.

5. EXPERIMENTAL RESULTS

Neuro-fuzzy diagnostics system with active learningwas used for breast cancer in Wisconsin data set processing [30-33].

Attributes 2 through 10 have been used to represent instances (Table 1). Each instance has one of 2 possible classes: benign or malignant.

Data set contains 16 patients has gaps in 7-th feature. Before processing we need to delete patients with gaps or to fill ones. Easy to use fuzzy special extrapolation method [34] that permit to fill all gaps with high veracity.

For visualization of this data set we used principal component analysis for extraction of three components. Visualization of breast cancer data set processing those was encoded in interval presented on Figure 2.

ISSN: 2456 - 3935

International Journal of Advances in Computer and Electronics Engineering

Volume 2, Issue 10, July 2017, pp. 1–10

TABLE I ATTRIBUTE INFORMATION

Attribute name in data folder	**Attribute characteristics**
Clump Thickness	id number
Uniformity of Cell Size	1 - 10
Uniformity of Cell Shape	1 - 10
Marginal Adhesion	1 - 10
Single Epithelial Cell Size	1 - 10
Bare Nuclei	1 - 10
Bland Chromatin	1 - 10
Normal Nucleoli	1 - 10
Mitoses	1 - 10
Class	2 for benign, 4 for malignant

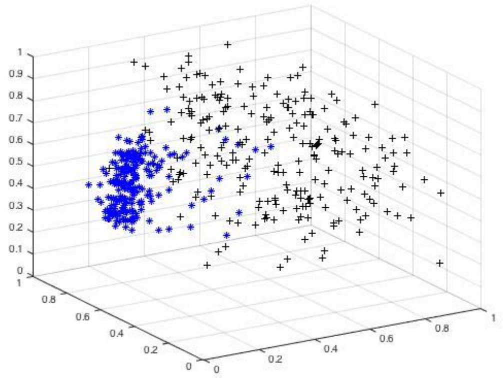

Figure 2 Visualization of breast cancer data set processing those was encoded in interval $[0,1]$

After that vector of these components $R(k)$ was centered relatively mean as $\sum_{\tau=1}^{k} R(\tau)=0$ and normalized as $\|R(\tau)\|=1$. On Figure3 breast cancer data set visualisation is presented and Figure4 – visualisation in three projections. Easy to see that all data are on surface of the sphere.

At next step in data set diagnoses of 295 patients was deleted (marked as NaN). Data set was mixed and divided on training data set (450 patients) and testing data set (249 patients). Initial position of cluster centers was set randomly. Neuro-fuzzy diagnostics system with active learning was trained using algorithm: patients with known diagnoses trained in control learning mode, patients with unknown diagnoses – in self-learning mode (number of possible classes is 2). Then neuro-fuzzy diagnostics system with active learning was tested.

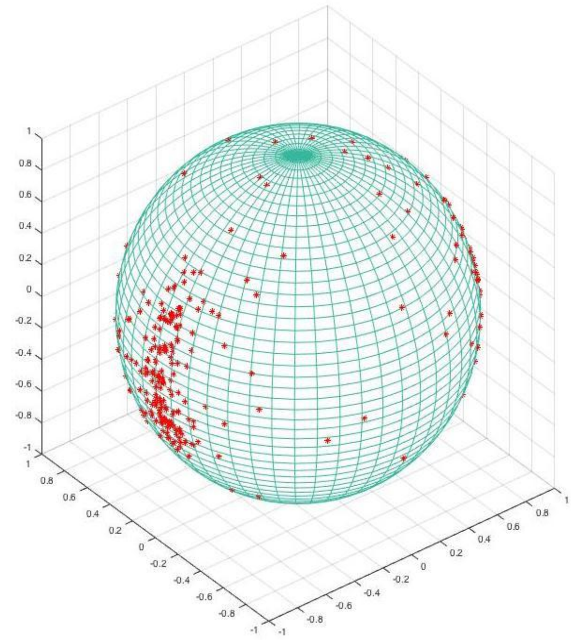

Figure 3 Breast cancer data set visualization

We have to note that all other known computational intelligence methods does not use the fact that a part of data have known diagnosis. Systems of self-learning (like k-means algorithm) for data clusterization use only features and systems of control learning have not enough marked data for its training.

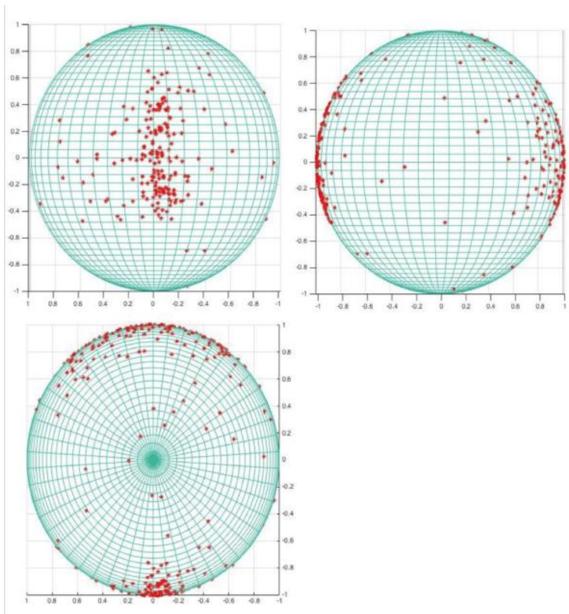

Figure 4 Breast cancer data set visualization

Incorrect classified patterns results presented in Ta-

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

ble 2 and visualization of results presented on Figure 5 and Figure 6.

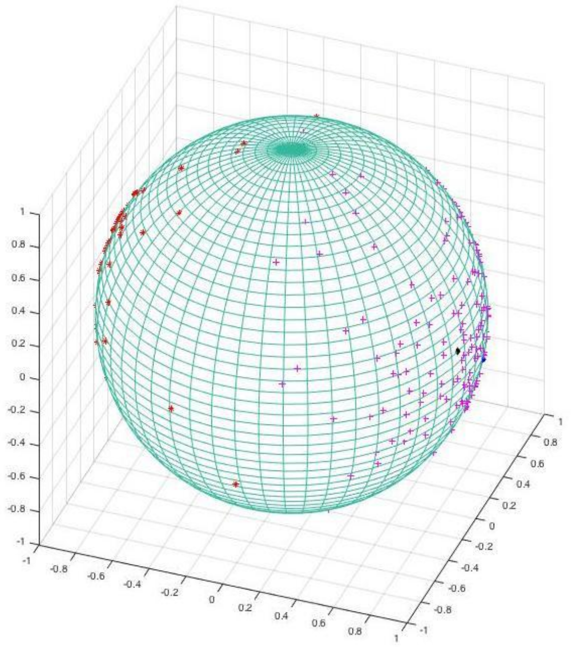

*Figure 5 Visualization of classification results: **+** – Benign, ***** – Malignant, ● – Initial position of clusters centers, ♦ – final position of clusters centers*

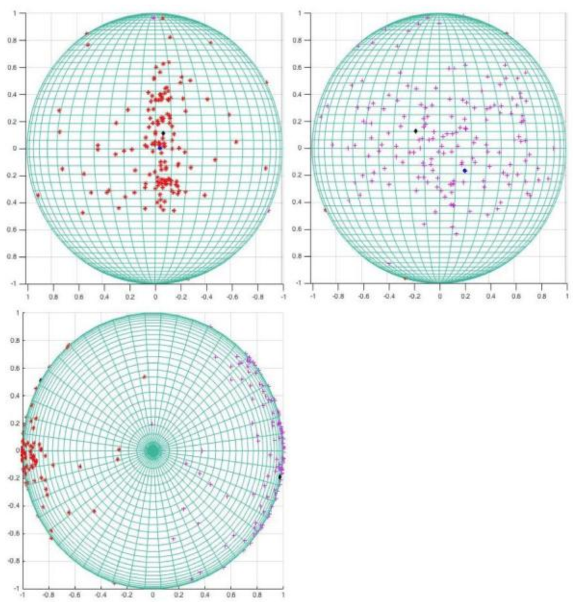

*Figure 6 Visualization of classification results in three projections: **+** – Benign, ***** – Malignant, ● – Initial position of clusters centers, ♦ – final position of clusters centers*

It is easy to that neuro-fuzzy diagnostics system with active learning classified patterns with high veracity.

TABLE II RESULT OF CLASSIFICATION BREAST CANCER DATA SET USING DIFFERENT SYSTEMS OF COMPUTATIONAL INTELLIGENCE

	Incorrect classified patterns			
	Training set		Testing set	
NFDS_AL	11	2,44%	7	2,81%
k-means	21	4,66%	11	4,72%

6. CONCLUSION

At this article neuro-fuzzy system of online medical diagnostics with active learning is proposed. System permits in sequential mode to solve tasks of classification-clusterization in conditions of a priory fuzziness and uncertainty. System under consideration has simple architecture, similar to SOM and LVQ, is effective in numeric realization due using of sequential recurrent procedures, that are gradient optimization algorithms of adopted goal functions.

REFERENCES

[1] Y. Bodyanskiy, A. Dolotov, O. Vynokurova, "Evolving spiking wavelet-neuro-fuzzy self-learning system"*Applied Soft Computing Journal. - 2014. - 14.* - P. 252-258. DOI: 10.1016/j.asoc.2013.05.020

[2] Ye.V. Bodyanskiy, E.A. Vynokurova, A.I. Dolotov, "Self-learning cascade spiking neural network for fuzzy clustering based on group method of data handling"*Journal of Automation and Information Sciences.* – 2013. – 45(3). – P.23-33. DOI: 10.1615/JAutomatInfScien.v45. i3.30

[3] G. Divyabharathi, G. Anupriya "Generation of Question and Answer from Unstructured Document using Gaussian Mixture Neural Topic Model", *International Journal of Advances in Computer and Electronics Engineering*, Volume 1, Issue 1, May 2016, pp. 1–5.

[4] J. Friedman, T. Hastie, R. Tibshirani "The Elements of Statistical Learning. Data Mining, Inference and Prediction", *Berlin: Springer*, 2003. – 552 p.

[5] L. Ruthowski, "Computational Intelligence. Methods and Techniques", *Berlin-Heidelberg: Springer-Verlag*, 2008. – 514 p.

[6] C.L. Mumford, L.C Jain. "Computational Intelligence. Collaboration, Fusion and Emergence"*Berlin: Springer-Verlag*, 2009. – 725p.

[7] R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, P. Held "Computational Intelligence. Methodological Introduction"*Berlin: Springer-Verlag*, 2013. – 488p.

[8] S. MadhuRupa, R.Balamurugan "A Modified Bcbimax Biclustering Algorithm for Market Segmentation", *International Journal of Advances in Computer and Electronics Engineering*, Volume 1, Issue 1, May 2016, pp. 23–27.

[9] K.-L. Du, M.N.S. Swami"Neural Networks and Statistical Learning", *London: Springer-Verlag*, 2014. – 824 p.

[10] I. Izonin, R. Tkachenko, D. Peleshko, T. Rak and D. Batyuk, "Learning-based image super-resolution using weight coefficients of synaptic connections",*Scientific and Technical Conference "Computer Sciences and Information*

Technologies" (CSIT), 2015 Xth International, Lviv, 2015, pp. 25-29. doi: 10.1109/STC-CSIT.2015.7325423

[11] JJ. S. Bhalla, A. Aggarwal "A novel method for medical disease diagnosis using artificial neural networks based on backpropagation algorithm"*Confluence 2013: The Next Generation Information Technology Summit (4th International Conference)*. – 2013. – pp.55-61.

[12] R. Isola , R. Carvalho, M. Iyer, A. K. Tripathy "Automated Differential Diagnosis in Medical Systems Using Neural Networks, kNN and SOM" *Developments in E-systems Engineering (DeSE)*, 2011. – pp.62-67.

[13] F. Jiménez, G. Sánchez, J.M. Juárez "Multi-objective evolutionary algorithms for fuzzy classification in survival prediction"*Artificial Intelligence in Medicine* 60 (2014). – pp. 197-219. DOI: 10.1016/ j.artmed.2013.12.006

[14] I. Perova, I. Pliss, G. Churyumov, Franklin M. Eze, Samer Mohamed Kanaan Mahmoud "Neo-Fuzzy Approach for Medical Diagnostics Tasks in Online-Mode" *2016 IEEE First International Conference on Data Stream Mining &Processing (DSMP)*. – Lviv, Ukraine. – 2016, p. 34-38 DOI: 10.1109/DSMP.2016.7583502.

[15] E. Lughofer "Evolving Fuzzy Systems – Methodologies, Advanced Concept and Applications"*Berlin-Heidelberg: Springer-Verlag*, 2011. – 454p.

[16] E. Lughofer "Single pass active learning with conflict and ignorance" *Evolving systems*. – 2012. – 3. - №4. – P. 251-271.

[17] Bezdek J.C. "Pattern Recognition with Fuzzy Objective Function Algorithms"*New York: Plenum Press*, 1981. – 272p.

[18] Kohonen J. "Self-Organizing Maps",*Berlin: Springer-Verlag*, 1995. – 362p.

[19] Osowski S. "Siecineuronowe do przetwarzaniainformacji "*Warszawa: OficinaWydawnicza Politechniki Warszawskiej*, 2006. – 422s. (in Polish).

[20] Grossberg S. "Classical and instrumental learning by neural networks"*In: Progress in Theoretical Biology. – N.Y.: Academic Press*, 1974. – 3. – P.51-141.

[21] Bezdek J.C., Keller J., Krisnapuram R., Pal N. "Fuzzy Models and Algorithms for Pattern Recognition and Image Processing"*V.4: Springer*, 1999. – 777p.

[22] Hoppner F., Klawonn F., Kruse R., Runkler T. "Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition"*Chichester: John Wiley & Sons*, 1999. – 300 p.

[23] Tsao E.C., Bezdek J.C., Pal N.R. "Fuzzy Kohonen clustering methods"*Pattern Recognition*. – 1994. – 27. – P. 757-764.

[24] Bodyanskiy Ye. "Computational intelligence techniques for data analysis" *Lecture Notes in Informatics*. – P. 72. – Bonn: GI, 2005. – P. 15-36.

[25] Gorshkov Ye., Kolodyazhniy V., Bodyanskiy Ye. "New recursive learning algorithms for fuzzy Kohonen clustering network"*Proc. 17th Int. Workshop on Nonlinear Dynamics of Electronic System*. – Rapperswil, Switzerland, 2009. – P. 58-61.

[26] Bodyanskiy Ye., Vynokurova O., Savvo V., Tverdokhlib T., Mulesa P. "Hybrid clustering-classification neural network in the medical diagnostics of the reactive arthritis"*International Journal of Intelligent Systems and Applications (IJISA)*. – 2016. – 8. – P. 1-9.

[27] S. Haykin, "Neural Networks. A Comprehensive Foundation", *Upper Saddle River, N.J.: Prentice Hall*, Inc., 1999. – 842 p.

[28] Bodyanskiy Ye., Vynokurova O., Mulesa P., Slipchenko O. "Self-organizing map and its learning in the fuzzy clustering-classification tasks"*Bulletin of National University "LvivPolitechnika". Computer Science and Information Technologies*. - 2014. – № 800. – P. 83-91.

[29] Huntsberger T., Ajjimarangsee P. "Parallel self-organizing feature maps for unsupervised pattern recognition"*Int. J. Gen. Systems*. – 1990. – 16. – P. 357-372.

[30] O. L. Mangasarian, W. H. Wolberg: "Cancer diagnosis via linear programming", *SIAM News*, Volume 23, Number 5, September 1990, pp 1 & 18.

[31] William H. Wolberg and O.L. Mangasarian: "Multisurface method of pattern separation for medical diagnosis applied to breast cytology", *Proceedings of the National Academy of Sciences, U.S.A.*, Volume 87, December 1990, pp 9193-9196.

[32] O. L. Mangasarian, R. Setiono, W.H. Wolberg: "Pattern recognition via linear programming: Theory and application to medical diagnosis", in: "Large-scale numerical optimization", Thomas F. Coleman and Yuying Li, editors, *SIAM Publications*, Philadelphia 1990, pp 22-30.

[33] K. P. Bennett & O. L. Mangasarian: "Robust linear programming discrimination of two linearly inseparable sets", *Optimization Methods and Software 1*, 1992, 23-34 (Gordon & Breach Science Publishers).

[34] Mulesa P., Perova I. "Fuzzy Spacial Extrapolation Method Using Manhattan Metrics for Tasks of Medical Data Mining"*Computer Science and Information Technologies CSIT'2015*. – Lviv, Ukraine. – 2015, p. 104-106 DOI: 10.1109/STC-CSIT.2015.7325443.

Authors Biography

Iryna Perova, is an associate professor in Biomedical engineering department at Kharkiv National University of Radio Electronics, Ukraine. She graduated from Kharkiv National University of Radio Electronics in 2004. She got her PhD in 2008. She obtained an academic title of the Senior Researcher in 2015. She obtained an academic title of the Associate Professor in 2016. Her research interests are medical data mining, systems of computational intelligence, neuro- and neo-fuzzy-systems for medical diagnostics tasks, on-line systems that have to do with control, identification, clustering, diagnostics and fault detection.

Yevgeniy Bodyanskiy, is a Professor of Artificial Intelligence Department at Kharkiv National University of Radio Electronics, Scientific Head of Control Systems Research Laboratory at Kharkiv National University of Radio Electronics, Ukraine. In 1971 he graduated with honour from Kharkiv National University of Radio Electronics. In 1980 he defended the Doctoral Thesis. In 1984 he was awarded the academic title of Senior Researcher. In 1990 he was awarded Dr. habil. sc. ing. degree. In 1994 he was

ISSN: 2456 - 3935

International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, July 2017, pp. 1–10

awarded the academic title of Professor. His research interests are evolving hybrid systems of computational intelligence, data stream mining, data science, and big data.

Cite this paper:

Yevgeniy Bodyanskiy, Iryna Perova, "Medical Online Neuro-Fuzzy Diagnostics System with Active Learning", International Journal of Advances in Computer and Electronics Engineering, Vol. 2, No. 7, pp. 1-10, July 2017.