MODIFIED PROBABILISTIC NEURO-FUZZY NETWORK FOR TEXT DOCUMENTS PROCESSING

Yevgeniy Bodyanskiy, Irina Pliss, Valentyna Volkova

Dept. of Artificial Intelligence, Kharkiv National University of Radio Electronics
Lenina av. 14, 255, Kharkiv, Ukraine, 61166
e-mail: bodya@kture.kharkov.ua, pliss@kture.kharkov.ua, volkova@kture.kharkov.ua
www.kture.kharkov.ua

Abstract: *We consider the problem of text processing (classification problem) using the modification of the probabilistic neural network proposed by D. Specht. Since classes often overlap during texts processing, there were introduced the modification that implements a procedure of fuzzy inference. For this into the network were introduced two additional layers. The results of the outputs of the network are levels of belonging each text document to each of the possible classes.*

Keywords: *Text document processing, probabilistic neuro-fuzzy network, multilayer architecture.*

1. INTRODUCTION

Data classification is key problem of data mining, and solving this problem is important for effective knowledge accumulation on the basis of observational analysis. One of the popular techniques of data classification is the probabilistic neural network (PNN) proposed by D.F. Specht [1,2]. This network is effectively applied to various problems of classification, diagnosis, pattern recognition etc. [3-5]. However, the conventional PNN solves the classification problem according to the traditional "crisp" approach when every observation is assigned to only one class. But in practice the problems with overlapping classes are common, when the classified observations can belong to several classes simultaneously with certain degrees of membership. To solve such problems the methods based on the fuzzy approach to clustering, classification and pattern recognition are used [6,7].

The fuzzy situations often arise in text document processing task. Such a process is called semantic metadata processing [9] and consists of three main components: ontology, text corpus, and the method of constructing a classifier for text mining.

There is a set of standard solutions that are proposed to describe the metadata and to generate semantic annotations, such as the standard Dublin Core [10]. However, given a sets of tags for describing text documents do not reflect the information that may be relevant to the current ontology and often contain only general information. It should be noted that the creation of semantic annotations manually is quite time consuming and requires large costs. This has led to the development of methods for semi-automatic text processing, which in turn has several disadvantages, such as using templates or a priori defined rules. Therefore, the actual task is to develop models and methods for on-line semantic annotation of text documents.

2. PROBLEM STATEMENT

The process of text processing can be regarded as a classification problem, and the automation can be achieved by applying different data mining technique. The main challenge in this case is to build a classifier based on an ontology O and the reference data sample X for the text object projection on the domain ontology. In this case "projection" should understand as the assignment of a text object to the ontological class as an instance of this concept.

It should be observed that artificial neural networks (ANN) have good reputation in the present modern developments using intelligent technologies, so it seems appropriate to use the neural network approach for semantic annotation of text documents [11-14]. In addition, the using of ANN makes it possible to build a classifier, regardless of the semantic annotation level (the level of words, sentences, paragraphs, documents) for the formation of the text documents' metadata. In this case, it is allowed to use different ways of presenting the

original text data in vector space of attributes that are not reflected in the quality of data processing.

The formal representation of semantic annotation of text documents can be obtained from the standpoint of the classification problem as follows. For the given domain ontology O the set of the concepts (classes) are defined as $C = (c(1), c(2), ..., c(N_1))$, where $c(i)$ – i-th concept from O. The set of text objects for the concerned text corpus are represented as $X = (x(1), x(2), ..., x(j), ..., x(N_2))$, where $x(j)$ – j-th text object presented in the vector form as a set of relevant features, N_1 and N_2 – number of ontological concepts (classes) and the power of the original text objects sample respectively.

Then the semantic annotation – labeling or the set of metadata for considered text document – on the basis of the given ontology will be defined as $L = \{l_i \mid \exists c_j \in C \wedge l_i = c_j\}$ where L – the unique set for each text document, consisting of concepts (classes) of the domain ontology, obtained by projections of textual objects belonging to each document to a given ontology using techniques based on ANN. It should be pointed out that under this approach does not exclude membership of one text object $x(k)$ to several classes (fuzzy situation), namely, the output of the classifier is determined by the membership to each given class of ontology O for input pattern. It is possible to include an object to a number of potentially possible concepts of the ontology by setting a threshold that is determined based on the membership difference. The received labeling can then be presented in tabular form for the formation of RDF-descriptions and its further using by various software tools.

3. MODIFIED PROBABILISTIC NEURO-FUZZY NETWORK

In the general case the probabilistic neural networks which solve the problem of Bayesian classification [15] via the recovery of unknown probability distributions by means of the Parzen kernels [15,16], belong to the feed forward architectures and are closely related to the radial-basis function networks [17] and generalized regression neural networks [18].

The modified probabilistic neuro-fuzzy network (MPNFN) has three-layer structure shown in Fig. 1 and is the modification of neuro-fuzzy systems introduced in [19,20].

The source information for the synthesis of network is the training set of patterns formed as a "packet" of n-dimensional vectors

$x(1), ..., x(i), ..., x(N), x(N+1), ..., x(N+M)$,

$x(j) = (x_1(j), ..., x_i(j), ..., x_n(j))^T$. The order of patterns in the packet is not significant. It is assumed that the crisp classification (the membership to one of m clusters) is known for each of the training patterns $x(j)$, and the representatives of all the possible clusters must be present in the dataset. That is, if N_l observations from the training set belong to the l-th class, then

$$\sum_{l=1}^{m} N_l = N. \qquad (1)$$

The input layer (number 0) receives an n-dimensional vector patterns $x(k), k > N+M$ that are transferred to the hidden layers for further processing.

The first hidden layer (the prototype layer) contains N neurons with the bell-shaped (usually Gaussian) activation functions and their synaptic weights are determined by the components of the training patterns, i.e.

$$w_{ji} = x_i(j), \qquad (2)$$

$i = 1, 2, ..., n, \ j = 1, 2, ..., N$, or in a vector form

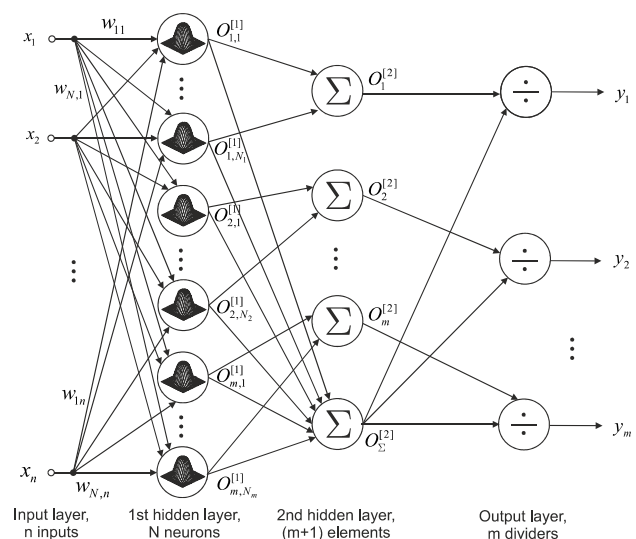$$w_j = x(j) = (x_1(j), ..., x_i(j), ..., x_n(j))^T.$$

Fig. 1 – Modified probabilistic neuro-fuzzy network

For the convenience of notation, all the neurons in the prototype layer are divided into m groups with N_l nodes in each, corresponding to its class.

The vector of weights of the p-th neuron in the l-th group will be denoted as $w_{l,p} = w_{l+p-1}$, $l = 1, 2, ..., m, p = 1, 2, ..., N_l$.

It is obvious that the training of the synaptic weights in this case is reduced to their one-time setting that is very simple.

When the vector $x(k)$ is fed to the 0-th layer the neurons of the 1-st layer produce the signals

$$o_{l,p}^{[1]}(k) = \Phi_{l,p}(x(k), \sigma_l) =$$
$$= \exp\left(-\frac{\|x(k) - w_{l,p}\|^2}{2\sigma_l^2}\right), \quad (4)$$

$l = 1, ..., m; p = 1, ..., N_l; k > N + M$, (where σ_l^2 is the width parameter of the activation function).

If all the input vectors are pre-normalized so that $\|x\| = 1$, the expression (4) assumes the following simple form

$$o_{l,p}^{[1]}(k) = \exp\left(-\frac{1 - x^T(k)w_{l,p}}{\sigma_l^2}\right). \quad (5)$$

It is also obvious that $\|w_{l,p}\| = 1$.

Since it holds

$$-1 \leq x^T(k)w_{l,p} \leq 1, \quad (6)$$

the outputs of the neurons of the first layer can vary only within the interval

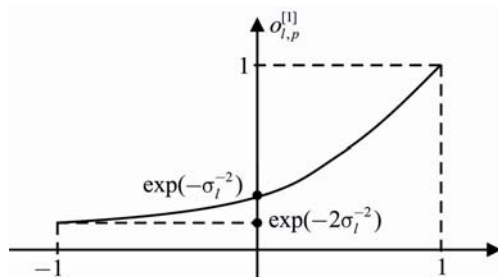$$\exp(-2\sigma_l^{-2}) \leq o_{l,p}^{[1]} \leq 1 \quad (7)$$

as is shown in Fig. 2.

Fig. 2 – Activation function of neurons in the first layer

At the level of the first layer it is also possible to determine the diameters of the classes formed by the training data as

$$\begin{cases} D_l^2 = \max_{p,q=1,...,N_l} \|w_{l,p} - w_{l,q}\|^2 = \\ = \max_{p,q=1,...,N_l} 2(1 - w_{l,p}^T w_{l,q}), \quad (8) \\ 0 \leq D_l = \sqrt{\max 2(1 - w_{l,p}^T w_{l,q})} \leq 2, \end{cases}$$

and to roughly estimate how much the overlap.

The second hidden layer is formed by $m+1$ elementary summing units with first m of them receiving the outputs of the prototype layer so that

$$o_l^{[2]}(k) = \sum_{p=1}^{N_l} o_{l,p}^{[1]}(k), \quad (9)$$

And the output of the $(m+1)$-th summing unit calculates the total sum

$$o_\Sigma^{[2]}(k) = \sum_{l=1}^{m} \sum_{p=1}^{N_l} o_{l,p}^{[1]}(k). \quad (10)$$

It can be readily seen that the sums (9) are Parzen approximations [5,6] of the unknown data distributions in the classes.

Finally the output of normalization layer formed by m dividers calculates the vector of degrees of membership $y(k) = (y_1(k), ..., y_l(k), ..., y_m(k))^T$ of the processed observation as

$$0 \leq y_l(k) = \frac{o_l^{[2]}}{o_\Sigma^{[2]}} \leq 1, \sum_{l=1}^{m} y_l(k) = 1. \quad (11)$$

It is easy to see that the described network is a combination of the probabilistic and generalized regression neural networks and is capable of data classification on the basis of fuzzy decision on membership of a particular observation to a certain class.

It is also evident that the first two layers of this network are essentially a standard radial basis function architecture with fixed synaptic weights $w_{lj}^{RBF} \equiv 1$ and centers $c_j^{RBF} \equiv w_j$.

Undetermined are only the cluster width parameters σ_l^2 which significantly affect the classification accuracy and are usually chosen empirically [5].

Therefore it is advisable to introduce the learning capabilities into the proposed network.

The learning algorithm adjusts the width parameters σ_l^2 to provide more accurate approximation of the data belonging to the l-th class

from the observations $x(u)$, $u = N+1,...,N+M$ in the training set and their crisp classification $d(u)$.

4. LEARNING ALGORITHM

Let introduce the one-step learning criterion (error function)

$$E(u) = \sum_{h=1}^{m} E_h(u) = \frac{1}{2}\|e(u)\|^2, \qquad (12)$$

(where $e(u) = (e_1(u),...,e_h(u),...,e_m(u))^T$, $E_h(u) = \frac{1}{2}e_h^2(u) = \frac{1}{2}(d_h(u) - y_h(u))^2$, $d_h(u)$ is the training signal equal to 1 if the vector $x(u)$, $u - N+1,...,N+M$ belongs to the h-th class and 0 otherwise). The derivative of the error function with respect to the tunable parameters will be

$$\begin{cases} \dfrac{\partial E_l(u)}{\partial \sigma_l^2} = -e_l(u)\dfrac{1-y_l(u)}{o_\Sigma^{[2]}(u)}\sum_{p=1}^{N_l}\dfrac{\partial o_{l,p}^{[1]}(u)}{\partial \sigma_l^2}, \\[4mm] \dfrac{\partial E_h(u)}{\partial \sigma_l^2} = e_h(u)\dfrac{y_h(u)}{o_\Sigma^{[2]}(u)}\sum_{p=1}^{N_l}\dfrac{\partial o_{l,p}^{[1]}(u)}{\partial \sigma_l^2}, h \neq l, \end{cases} \quad (13)$$

or

$$\frac{\partial E_h(u)}{\partial \sigma_l^2} = e_h(u)\frac{y_h(u) - \delta_{hl}}{o_\Sigma^{[2]}(u)}\sum_{p=1}^{N_l}\frac{\partial o_{l,p}^{[1]}(u)}{\partial \sigma_l^2}, \qquad (14)$$

$$h = 1,2,...,m,$$

where $\delta_{hl} = \begin{cases} 1, & if\ h = l, \\ 0, & otherwise. \end{cases}$

It can be readily seen that

$$\frac{\partial E(u)}{\partial \sigma_l^2} = \frac{1}{o_\Sigma^{[2]}(u)}\sum_{p=1}^{N_l}\frac{\partial o_{l,p}^{[1]}(u)}{\partial \sigma_l^2}\cdot$$
$$\cdot\sum_{h=1}^{m}(y_h(u) - \delta_{hl})e_h(u), \qquad (15)$$

and since for the Gaussian membership function

$$\frac{\partial o_{l,p}^{[1]}(u)}{\partial \sigma_l^2} = \frac{\|x(u) - w_{l,p}\|^2}{\sigma_l^4}\cdot$$
$$\cdot\exp\left(-\frac{\|x(u) - w_{l,p}\|^2}{2\sigma_l^2}\right), \qquad (16)$$

the expression (15) can be rewritten in the form

$$\frac{\partial E(u)}{\partial \sigma_l^2} = \frac{1}{2\sigma_l^4 o_\Sigma^{[2]}(u)}\sum_{h=1}^{m}(y_h(u) - \delta_{hl})e_h(u)\cdot$$
$$\cdot\sum_{p=1}^{N_l}\|x(u) - w_{l,p}\|^2\exp\left(-\frac{\|x(u) - w_{l,p}\|^2}{2\sigma_l^2}\right). \qquad (17)$$

Minimizing (12) with the gradient-based procedure

$$\sigma_l^2(q+1) = \sigma_l^2(u) - \eta(u)\frac{\partial E(u)}{\partial \sigma_l^2}, \quad (18)$$

we can finally write the learning algorithm of the probabilistic neuro-fuzzy network

$$\sigma_l^2(q+1) = \sigma_l^2(u) - \eta(u)\frac{1}{2\sigma_l^4(u)o_\Sigma^{[2]}(u)}\cdot$$
$$\cdot\sum_{p=1}^{N_l}\|x(u) - w_{l,p}\|^2\exp\left(-\frac{\|x(u) - w_{l,p}\|^2}{2\sigma_l^2(u)}\right)\cdot \quad (19)$$
$$\cdot\sum_{h=1}^{m}(y_h(u) - \delta_{hl})e_h(u),$$

or taking into account (5)

$$\sigma_l^2(q+1) = \sigma_l^2(u) - \eta(u)\frac{1}{\sigma_l^4(u)o_\Sigma^{[2]}(u)}\cdot$$
$$\cdot\sum_{p=1}^{N_l}\|1 - x^T(u)w_{l,p}\|^2\exp\left(-\frac{1 - x^T(u)w_{l,p}}{\sigma_l^2(u)}\right). \quad (20)$$
$$\cdot\sum_{h=1}^{m}(y_h(u) - \delta_{hl})e_h(u),$$

where $\eta(u)$ is the scalar learning rate parameter.

When the learning is finished the values $\sigma_l^2(N+M+1)$, $l = 1,2,...,m$ are used as estimates of the activation function parameters (4), (5) in the classification of objects $x(k), k > N+M$ with unknown membership.

5. THE RESULTS OF EXPERIMENT

The processing semantic annotation procedure for text documents, which is based on MPNFN, was tested on sample texts belonging to different ontological classes (50 features, 100 objects). The initial data was considered as a corpus "20

Newsgroups DataSet" (comp.graphics, comp.os.ms-windows.misc, comp.sys.ibm.pc.hardware classes), and text information is treated from highly specialized subject area. Initial taxonomy (ontology) for "20 Newsgroups DataSet" is considered on Fig. 3.

comp.graphics
comp.os.ms-windows.misc
comp.sys.ibm.pc.hardware
comp.sys.mac.hardware
comp.windows.xrec.autos
rec.motorcycles
rec.sport.baseball
rec.sport.hockeysci.crypt
sci.electronics
sci.med
sci.spacemisc.forsaletalk.politics.misc
talk.politics.guns
talk.politics.mideasttalk.religion.misc
alt.atheism
soc.religion.christian

Fig. 3 – Initial taxonomy for "20 Newsgroups DataSet"

Converting text corpus in the vector space based on statistical evaluation of TF-IDF using WordNet 2.0 (for the attributes formation was considered only nouns, stop words were removed). The experiment was considered, first of all, to estimate quality of MPNFN processing [13]. Table 1 shows the results of the processing. It is indicated for an input text object that the operation generates a set of the membership values to several classes, which are considered as classes of domain ontology.

Table 1. Examples of the program work

Input text object number	Membership values to each classes		
	comp.graph ics	*comp.os.ms windows.m isc*	*comp.sys.ib m.pc. hardware*
10157	0,99964	0,00010032	0,00025828
10158	0,67091	0,094663	0,23443
10159	0,02427	0,95577	0,019964
10160	0,092507	0,88085	0,02664
10161	0,32385	0,62807	0,048079
10162	0,13043	0,84376	0,025803
10163	0,061536	0,91668	0,021786
10164	0,25684	0,37909	0,36406
10165	0,10571	0,033676	0,86061
10166	0,0044423	0,00086429	0,99469

Thus it becomes possible to get semantic annotations on the basis of proposed method. Example of the semantic annotation for the text document № 10159 from the "20 Newsgroups DataSet" is shown in Fig. 4. In this case semantic annotation includes some tags from Dublin Core (title, author) and the other obtained by modified probabilistic neural network.

```
<!-- http://www.semanticweb.org/ontologies/2011/4/Ontology1304595614765.owl#Document10159-->
  <owl:NamedIndividual rdf:about="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#Document10159">
    <rdf:type rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#Document"/>
    <dcterms:title xml:lang="en-US">WfWorkgroups server goes deaf</dcterms:title>
    <dcterms:publisher rdf:resource="http://people.csail.mit.edu/~jrennie/20Newsgroups/"/>
    <dcterms:modified>2010-05-15</dcterms:modified>
    <creation_date rdf:datatype="&xsd;dateTime">2003-03-18</creation_date>
    <from rdf:datatype="&xsd;string">taso@munnari.OZ.AU (Taso Hatzi)</from>
    <organization rdf:datatype="&xsd;string">Comp Sci, Univ Melbourne, AU</organization>
    <X-Newsreader rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#TIN"/>
    <belongs_to rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#CP10159_1"/>
    <belongs_to rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#CP10159_2"/>
    <belongs_to rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#CP10159_3"/>
    <X-Newsreader rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#TIN"/>
  </owl:NamedIndividual>
  <owl:NamedIndividual rdf:about="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#CP10159_1">
    <rdf:type rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#ComplexProbability"/>
    <probability rdf:datatype="&xsd;double">0,02427</probability>
    <class_label rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#comp.graphics"/>
  </owl:NamedIndividual>
  <owl:NamedIndividual rdf:about="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#CP10159_2">
    <rdf:type rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#ComplexProbability"/>
    <probability rdf:datatype="&xsd;double">0,95577</probability>
    <class_label rdf:resource="http://www.semanticweb.org/ontologies/2011/4/
Ontology1304595614765.owl#comp.os.ms-windows.misc"/>
  </owl:NamedIndividual>
```

Fig. 4 – Example of text document semantic annotation

6. SUMMARY

In this paper a method for semantic annotation based on modified probabilistic neuro-fuzzy network, which is a hybrid of the probabilistic neural network, a general regression neural network and neuro-fuzzy systems is described. Due to this, it becomes possible to determine the memberships for an incoming text object to each of the potentially possible classes of ontology. This method provides the ability to process information as it becomes available in sequential on-line mode, characterized by simplicity of implementation and speed of information processing.

6. REFERENCES

[1] D.F. Specht, Probabilistic neural networks, Neural Networks, (1) 3 (1990), pp. 109-118.

[2] D.F. Specht, Probabilistic neural networks and polynomial adaline as complementary techniques for classification, *IEEE Trans. on Neural Networks*, (1) 1 (1990), pp. 111-121.

[3] C.M. Bishop, *Neural Networks for Pattern Recognition,* Clarendon Press, Oxford, 1995, 482 p.

[4] C. Looney, *Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists,* Oxford University Press, N.Y., 1997, 480 p.

[5] R. Callan, *The Essence of Neural Networks,* Prentice Hall Europe, London, 1999, 288 p.

[6] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms,* Plenum Press, N.Y., 1981, 272 p.

[7] F. Höppner, F. Klawonn, R. Kruse, *Fuzzy Cluster-Analyse,* Vieweg, Braunschweig, 1999, 289 p.

[8] J.-S. R. Jang, C.-T. Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing – Computational Approach to Learning and Machine Intelligence,* Prentice Hall, Upper Saddle River, 1997, 640 p.

[9] V. Uren, Ph. Cimiano, J. Iria, S. Handschuh, M. Vargas-Vera, E. Motta, F. Ciravegna, Semantic annotation for knowledge management: Requirements and a survey of the state of the art, *Web Semantics: Science, Services and Agents on the World Wide Web*, (1) 4 (2006), pp. 14-28.

[10] Dublin Core: Metadata Initiative (http://www.dublincore.org [last accessed 22/02/2011]).

[11] Ye. Bodyanskiy, O. Shubkina, Semantic annotation of text documents using hierarchical radial basis function neural network, *Eastern-European Journal of Enterprise Technologies*, (48) 6/3 (2010), pp. 72-77.

[12] Ye. Bodyanskiy, O. Shubkina, Semantic annotation of text documents using evolving neural network based on principle "Neurons at Data Points", *4th Int. Workshop on Inductive Modeling "IWIM 2011"*, Kyiv, Ukraine (2011), pp. 31-37.

[13] Ye. Bodyanskiy, O. Shubkina, Semantic annotation of text documents using modified probabilistic neural network, *6th IEEE Int. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Prague, Czech Republic (2011), pp. 328-331.

[14] O. Shubkina, I. Pliss, Ye. Bodyanskiy, Using of competitive probabilistic network in the text information processing tasks, *Bulletin of NU "Lvivska Politechnika", Computer Science and Information Technologies*, 710 (2011), pp. 175-181.

[15] K. Fukunaga, *Intruduction to Statical Pattern Recognition.* Academic Press, New York and London, 1972, 368 p.

[16] E. Parzen, On the estimation of a probability density function and the mode, *Ann. Math. Statist.*, 38 (1962), pp. 1065-1076.

[17] J. Moody, C.J. Darken, Fast learning in networks of locally-tuned processing units. *Neural Computing*, 1 (1989), pp. 281-299.

[18] D.F. Specht, A general regression neural network, *IEEE Transactions on Neural Networks*, (2) 6 (1991), pp. 568-576.

[19] Ye. Bodyanskiy, Ye. Gorshkov, V. Kolodyazhniy, J. Wernstedt, A learning probabilistic neural network with fuzzy inference, Proc. *6th Int. Conf. on Artificial Neural Nets and Genetic Algorithms "ICANNGA 2003"*, Roanne, France, Springer-Verlag, Wien (2003), pp. 13-17.

[20] Ye. Bodyanskiy, Ye. Gorshkov, V. Kolodyazhniy, J. Wernstedt, *Probabilistic neuro-fuzzy network with non-conventional activation functions*, in: Lecture Notes in Artificial Intelligence, v. 2773, Springer, Berlin-Heidelberg-New York, 2003, pp. 973-979.

Yevgeniy Bodyanskiy, in 1971 graduated with honour Kharkiv National University of Radio Electronics. In 1980 has defended the candidate thesis. In 1984 has taken academic title Senior Reseacher. In 1990 defended the doctor thesis (Dr.habil.sci.ing.). Since 1974 he has been working at Kharkiv National University of Radio Electronics. Since 1992 professor of Artificial Intelligence Department KhNURE, scientific head of Control Systems Research Laboratory KhNURE. He has more then 500 scientific publications including 40 inventions and 10 monographs. Research interests include hybrid systems of computational intelligence: adaptive, neuro-, wavelet-, neo-fuzzy-, real-time systems including problems connected with control, identification, forecasting, clustering, diagnostics.

Iryna Pliss, In 1970 received her Engineer-electrician degree in "Automatic and Remote Control" from Kharkiv National University of Radio Electronics, Ukraine. She became Candidate of technical Sciences in 1979 on "Technical Cybernetics and Information Theory", and Senior Researcher in 1984 on "Technical Cybernetics and Information Theory". She is currently a Leading Researcher of the Control Systems Research Laboratory. She is a member of the IEEE Signal Processing Society and the Neural Network Society. She has more than 150 scientific publications and five inventions. Her current research interests are neuro-fuzzy hybrid systems for forecasting, identification and diagnostics.

Valentyna Volkova, in 2005 graduated from Kharkiv National University of Radio Electronics by specialty "Intelligent Decision Making Systems", 2005-2006 rese-archer, 2006-2009 post-graduate study in Artificial Intelligence Department, has defended the PhD thesis in 2011. Her major field of research is fuzzy clustering of multi-topic text documents. After post-graduate study works in Artificial Intelligence Department at Kharkiv National University of Radio Electronics. She has more than 20 publications. Research interests are computational intelligence, data mining: fuzzy clustering algorithms based on neuro-fuzzy models.