UDC 004.93.14

FUZZY PROBABILISTIC NEURAL NETWORK IN DOCUMENT CLASSIFICATION TASKS

A. L. Yerokhin, O. V. Zolotukhin

Kharkiv National University of Radio Electronics

E-mail: andriy.yerokhin@nure.ua, oleg.zolotukhin@nure.ua

The modification of a probabilistic neural network based on adding fuzziness is investigated. This article discusses the architecture and learning algorithm of fuzzy probabilistic neural network that can classity on-line text documents.

Keywords: *fuzzy neural network, text document classification.*

НЕЧІТКА ІМОВІРНІСНА НЕЙРОННА МЕРЕЖА В ЗАДАЧАХ КЛАСИФІКАЦІЇ ДОКУМЕНТІВ

А. Л. Єрохін, О. В. Золотухін

Харківський національний університет радіоелектроніки

Досліджено модифікацію ймовірнісної нейронної мережі на основі введення нечіткості. Розглянуто архітектуру та алгоритм навчання нечіткої ймовірнісної нейронної мережі, яка може класифікувати текстові документи у режимі реального часу.

Ключові слова: *нечітка нейронна мережа, текстовий документ, класифікація.*

Introduction. The rapid growth of Internet technology makes possible a broad user access to various documents, including an on-line mode. However, there are new problems, the most urgent is information overload and, consequently, the need for classification of documents that consistently are received in real time.

This task is very important, for example, for news agencies, various online publishers, who must constantly classify data streams, such as news reports, analytical reviews, digests, articles, reports, etc. These usually classifiable web-documents are characterized by multidisciplinary, that affect several topics both very different and very close.

On-line classification of this type of text documents is not a trivial task, because a small piece of the text can contain very valuable information and reference to the relevant class cannot be ignored and closely spaced classes may overlap and / or merge. Therefore, it is desirable to consider the test document belonging to each of the potentially interesting for user classes. At the same time most of the known methods include document classification to one of the classes. In this regard it is urgently needed to develop methods of fuzzy probabilistic classification of text documents.

To address the problem of classification of documents quite effectively the probabilistic neural networks are introduced in paper [1] which is conducted on a “neurons data points”, making it extremely easy and quick.

In [2–5] modified probabilistic neural networks for processing text and the presence of different elements of competition in the learning process and the ability to correct fields of nuclear receptor activation functions are proposed.

However, the use of probabilistic neural networks in problems of word processing is complicated when the volume of information to be analyzed is rather large, and feature vectors (images) are sufficiently high in dimension. This is a difficult situation because of both probabilistic neural networks and other neural networks.

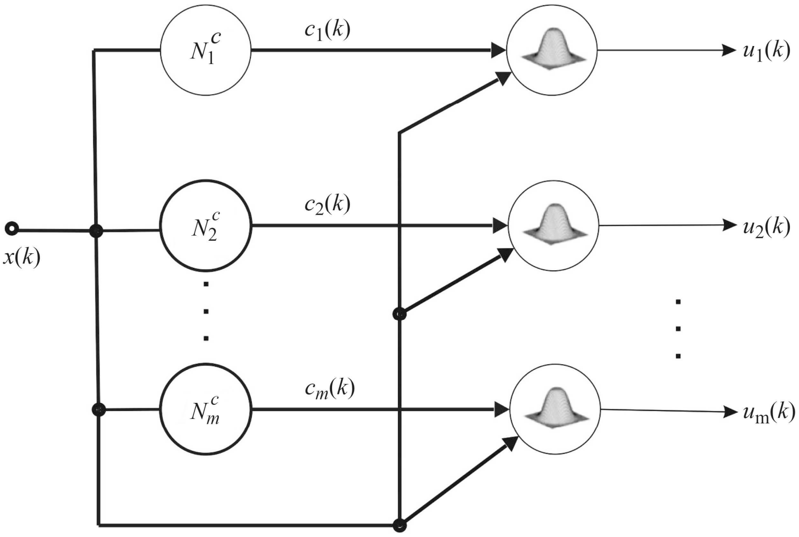
To overcome this drawback in [6] the improved probabilistic neural network (PNN), where the first layer is formed not hidden images and prototypes class, calculated using a conventional C-average in batch mode is proposed.

Since the problem of classification number of possible classes m is usually substantially less then the volume of training set N, the PNN are much better suited to solve applied problems than the standard probabilistic neural network.

However, we note the following main shortcomings of the PNN as an opportunity to study only in batches when training sample is set beforehand, as well as a clear result of classification (classification image is presented to only one class), while in the processing of text documents rather often a situation arises when the analyzed text with different levels of membership can simultaneously refer to several, maybe just ordinary classes.

In this work the approach to the synthesis of neuro-fuzzy network is proposed that learns in an on-line mode and is designed for fuzzy classification of the text documents, presented in the form of images, vectors and successively received for processing.

Fuzzy probabilistic neural network. The architecture proposed neuro-fuzzy network is shown in the Figure.



Architecture of fuzzy probabilistic neural network.

This neuro-fuzzy network contains two layers of information processing: the first hidden layer prototypes (instead of the first hidden layer images in the usual probabilistic neural network) and the output layer calculating the accessories.

The background for learning is classified sequence of vector-images $x(1), x(2),\ldots,$ $x(k),\ldots,x(N),$ where $x(k)=(x_1(k),x_2(k),\ldots,x_n(k))^T \in R^n$.

It is assumed that N may change over time, the number of classes m can also vary prototypes (centroids) classes of described vectors $c_j=(c_{j_1},c_{j_2},\ldots,c_{j_n})^T$ to be measured. Marking $x(k, j)$ image classification means $x(k)$ to j-th class $j=1,2,\ldots,m$. Each class contains N_j classified images $\sum_{j=1}^{m} N_j = N$.

The prototypes in [6] are calculated using ordinary arithmetic mean score

$$c_j = \frac{1}{N_j}\sum_{k=1}^{N_j} x(k,j),$$

which is easy to add to the recurrent form

$$c_j(k) = c_j(k-1) + \frac{1}{k}(x(k, j) - c_j(k-1)) \ , \tag{1}$$

that, as it stands, is responsible T. Kohonen learning rule [7] setting step

$$\eta(k) = \frac{1}{k} \ ,$$

relevant provisions of stochastic approximation.

Because the real problems prototypes classes can drift over time, instead of (1) can be used or exponentially weighted average

$$c_j(k) = \alpha c_j(k-1) + (1-\alpha)(x(k, j) - c_j(k-1)), \ 0 < \alpha < 1,$$

or adaptive procedure [8]

$$\begin{cases} c_j(k) = c_j(k-1) + \eta(k)(x(k, j) - c_j(k-1)) \ , \\ \eta(k) = r^{-1}(k), \ \ r(k) = \alpha r(k-1) + \left\| x(k, j) \right\|^2, \ \ 0 \le \alpha \le 1 \ , \end{cases}$$

under the condition α =1 it satisfies the stochastic approximation of A. Dvoretsky.

The input layer of the network evaluates the membership degree of unclassified observations $x(k)$ $(k > N)$ in existing classes of prototypes $c_j(N)$ using the expression

$$u_j(k) = \frac{\left\| x(k) - c_j(N) \right\|^{-2}}{\sum_{l=1}^{m} \left\| x(k) - c_l(N) \right\|^{-2}}, \tag{2}$$

underlying probabilistic procedures of the fuzzy classification method known as Fuzzy C-means [9].

Thus, in learning networks using both clear and unclear procedures are used.

If rewrite (2) as

$$u_j(k) = \frac{1}{1 + \left\| x(k) - c_j(N) \right\|^2 \sum_{\substack{l=1 \\ l \ne j}}^{m} \left\| x(k) - c_l(N) \right\|^{-2}} \ , \tag{3}$$

we can see that (3) is not like a bellship (nuclear) activation function

$$u_j(k) = \frac{1}{1 + \frac{\left\| x(k) - c_j(N) \right\|^2}{\sigma_j^2}},$$

width of the parameter of receptive field

$$0 \le \sigma_j^2 = \left(\sum_{\substack{l=1 \\ l \ne j}}^{m} \left\| x(k) - c_l(N) \right\|^{-2} \right)^{-1} \le \frac{4}{m-1} \ .$$

This parameter is set automatically during classification.

Expressions (2) and (3) are probabilistic fuzzy classification [10], that is, the following condition is

$$\sum_{j=1}^{m} u_j(k) = 1 \ ,$$

then the situation arises in which $u_j(k) = m^{-1} \forall j,$ means surveillance $x(k)$ or applies to all classes equally which is unlikely or any of them.

In this situation, it is possible to increase the number of classes $m + 1$ putting $x(k)$ as the initial prototype of a new class. If it is found that p classes $p < m$ standard accessories $u_j(k)$ are less than m^{-1}.

This means that $x(k)$ cannot belong to this class, and the level of affiliation should be counted using expression (2) making the summing superscript in the denominator $m–p$.

To avoid possible p classes that do not include $x(k)$ the procedure can be used which is based on the V-criteria [11] (Vigilance criterion) and testing conditions

$$e^{u_j(k)} \left\| x(k) - c_j(N) \right\| \le \varepsilon ,$$

where the threshold value is determined empirically. It is clear that $p = m - 1$ and we get a clear result of the classification [12].

CONCLUSIONS

In the paper the problem of solving on-line classification of text documents entering the processing sequence in real time is presented. The modification fuzzy probabilistic neural network is proposed, giving numerous extremely simple implementation and low volume necessary for the memory implementation.

1. *Specht D. F.* Probabilistic neural networks // Neural Networks. – 1990. – P. 109–118.
2. *Бодянский Е. В., Шубкина О. В.* Семантическое аннотирование текстовых документов с использованием модифицированной вероятностной нейроной сети // Системные технологии. – Днепропетровск, 2011. – Вып. 4 (75). – С. 48–55.
3. *Bodyanskiy Ye., Shubkina O.* Semantic annotation of text documents using modified probabilistic neural network // Proc. 6th IEEE Int. Conf. Intelligent Data Acquisition and Advanced Computing Systems: Techn. and Appl. – 15–17 Sept. 2011. – Czech Republic: Prague, 2011. – P. 328–331.
4. *Bodyanskiy Ye., Shubkina O.* Semantic annotation of text documents using evolving neural network based on principle “Neurons at Data Points”// Proc. 4th Int. Workshop on Inductive Modelling “IWIM 2011”. – Kyiv, 2011. – P. 31–37.
5. *Pattern* recognition using radial basis function network / D. R. Zahirnak, R. Chapman, S. K. Rogers, B. W. Suter, M. Kabriski, V. Pyatti // Proc. 6th Ann. Aerospace Application of AI Conf., Dayton, OH. – 1990. – P. 249–260.
6. *Ciarelli P. M., Oliveira E.* An enhanced probabilistic neural network approach applied to text classification // Lecture Notes on Comp. Sci. – Berlin-Heidelberg: Springer-Verlag, 2009. – **Vol. 5856**. – P. 661–668.
7. *Kohonen T.* Self-Organizing Maps. – Berlin: Springer, 1995. – 362 p.
8. *Бодянский Е. В., Руденко О. Г.* Искусственные нейронные сети: архитектура, обучение, применение. – Харьков: ТЕЛЕТЕХ, 2004. – 372 с.
9. *Convergence* theory for fuzzy c-means: Counterexamples and repairs / J. C. Bezdek, R. J. Hathaway, M. J. Sabin, W. T. Tucker // IEEE Transactions on Systems, Man, and Cybernetics. – 1987. – **Vol. SMC-17**, № 5. – P. 873–877.
10. *Fuzzy* Models and Algorithms for Pattern Recognition and Image Processing / J. C. Bezdek, J. Keller, R. Krishnapuram, N. R. Pal. – N.Y.: Springer Science + Business Media, Inc., 2005. – 776 p.
11. *Cannady J., Garcia R. G.* The Application of Fuzzy ARTMAP in the Detection of Computer Network Attacks // Proc. Int. Conf. Artificial Neural Networks. – ICANN 2001. – Austria: Vienna, 2001. – P. 225–230.
12. *Zolotukhin, O., Kudryavtseva, M.* Authentication Method in Contactless Payment Systems // Int. Scientific and Practical Conf. “Problems of Infocommunications. Science and Technology”, 9–12 October, 2018. – Ukraine: Kharkiv, 2018. – P. 397–400.

Received 19.10.2018