Neural Networks in Ischemic Strokes Diagnostics

Henadzi Vaitsekhovich, *Brest State Technical University*
Supervisor: Vladimir Golovko, *Brest State Technical University.*

Abstract

In this paper the neural network model for transient ischemic attacks recognition have been addressed. The proposed approach is based on integration of the NPCA neural network and multilayer perceptron. The dataset from clinic have been used for experiments performing. Combining two different neural networks (NPCA and MLP) it is possible to produce efficient performance in terms of transient ischemic attacks detection and recognition. The main advantages of using neural network techniques are quickness and ability to assist a doctor in making decision.

1. Introduction

Nowadays using of Artificial Intelligence has become broadly applied in medicine. Every year in medical journals are issued over 500 academic publications concerning artificial neural networks in medical applications [1]. In accordance with published literature the artificial neural networks are powerful tools for automatic diagnostics of disease with potential to support clinical decision making.

Medical diagnostics is a complicated task that needs to be executed accurately and efficiently. It consists on detection of a disease from many factors or symptoms. Unfortunately medical specialists very often may not enough experience to deal with certain high-risk diseases. Insufficient of medical specialists has increase the mortality of patients suffered from various diseases. So, for instance, the diagnosis of transient ischemic attacks (TIA) by primary care doctors was correct only in 30% cases [2]. Therefore automatic medical decision support system is of great importance for increasing quality of medical diagnostics.

There are different techniques to initial diagnostics of TIA: neuropsychological testing, statistical approach, artificial intelligence approach [3-8]. The main disadvantage of neuropsychological testing is correlation on a doctor qualification and low accuracy. The statistical approach demands large database. The artificial intelligence techniques use neural networks, genetic algorithm, fuzzy logic or combinations of abovementioned approaches. Such a technique is characterized with high accuracy and demands not too big data set in comparison with statistical approaches. Therefore artificial intelligence techniques can be appropriate tools for TIA detection and recognition. In this paper we propose neural network model for TIA detection and recognition. Neural networks technique is used to reduce a diagnostic time and a number of misdiagnosis, as well as to assist a doctor in making decision.

The efficiency of the proposed neural network model in detection and recognition of transient ischemic attack is illustrated by the experimental results.

The paper is organized as follows. A brief review of related works is given in Section 2. The basic information about transient ischemic attacks is presented in Section 3. Section 4 describes patient's data. The proposed neural network model for TIA recognition is detailed in sections 5, 6 and 7. In Section 8 the experimental results are described. Finally, concluding remarks are made in the last section.

2. Related Researches

At present time there exist different approaches to preliminary diagnostics of TIA. Neuropsychological testing is used for initial evaluation of TIA very often. Subsets of neurologic examination consist of cranial nerve testing, speech testing, somatic sensory testing, cerebellar system etc. [3]. However neuropsychological testing greatly depends on a doctor qualification and therefore is very subjective.

The recognition tool for transient ischemic attack is described in [4]. The authors applied multivariate logistic regression using ROC (receiver operating characteristic curves) analysis to develop clinical scoring system. The data base from West Glasgow Stroke Registry was used, which contains approximately 225 000 people. The nine variables (age, headache, diplopia, LOC/pre-syncope, seizure, speech abnormalities, unilateral limb weakness, facial weakness, history of TIA) with corresponding regression coefficients were obtained after regression analysis. To calculate the total score all coefficients should be summed. Using ROC curves was identified TIA, if total score >6.1. Such a system correctly identified 85% of patients with a cerebrovascular diagnosis and 54% with a non-cerebrovascular diagnosis. However such an approach demands a big database.

The neural networks for Ischemic Stroke are presented in [6]. The proposed models were developed for rapid classification into the following outputs: no event, TIA or stroke in left carotid, right carotid, verterobasilar. The questionnaire from asymptotic carotid atherosclerosis study (ACAS) was used as the input data. The 6 sections of the questionnaire contain the following data: loss or change of speech, loss of vision, double vision, numbness or tingling, paralysis or weakness, spells of dizziness or loss of balance. Each input pattern consists of numbers 0 or 1 from questionnaire. The multilayer perceptron was used for each section of the questionnaire. As an example, the network for double vision consists of 4 inputs and 5 outputs. Three outputs correspond to case - no event and the other two - to stroke or TIA.

In [7-8] backpropagation neural networks for the prediction of thrombo-embolic stroke are described. The architecture of neural network consists of 20 input units, 10 hidden and 10 output units. The following parameters were used as the input: hypertensive, diabetes, myocardial, blood cholesterol, left arm and left leg, etc. The output units shows the various categories of stroke diseases: TIA, left hemiplegia, right hemeplegia, dysphasia, monoplegia, left hemianopia, aphasia, right hemianesthesia, dysphasia, and quadriplegia. The prediction accuracy is 78,52% using training set and 90,61% using testing data set.

The mentioned above neural networks approaches differ from each other with input and output data, as well as database of patients. Therefore it is very difficult to compare different approaches.

3. Transient Ischemic Attacks

A transient ischemic attack (TIA) is a transient episode of neurological dysfunction caused by focal brain, spinal-cord or retinal ischemia without acute infarction [3]. It is a result of temporary reduction or cessation of cerebral blood flow in a specific neurovascular distribution due to low flow through a partially occluded vessel, an acute thromboembolic event, or stenosis of a small penetrating vessel. Transient ischemic attacks are named also thrombo-embolic stroke. After TIA a risk of early acute stroke is increased. So, for instance, the early risk of stroke following TIA is approximately 4-5% at 2 days and as high as 11% at 7 days [3]. The patients after TIA should undergo detailed diagnostic at the nearest time (within 24 hours), namely, computed tomography or magnetic resonance imaging evaluation. It can prevent further negative development of disease, which can lead to acute stroke. Therefore the preliminary diagnosis of TIA is of a great importance for prevention of acute stroke. However research has shown a high rate of misdiagnosis of TIA. So, for instance, the diagnosis of TIA by primary care doctors was correct only in 30% cases [2]. Improvement of diagnostic accuracy of TIA from primary doctors would reduce the number of acute stroke, clinic waiting times and facilitate rapid assessment.

4. Patient Data

The data about 114 patients who have symptoms of different TIA diseases have been collected. 38 parameters have been selected for each patient, such as age, sex, residence, education, trade, conflicts on job, residence change for last 10 years, trade change for last 10 year, features of night dream, sleeplessness, heredity on pathology brain vessels, heredity on other diseases, arterial hypertensia, diastolic pressure, auscultation hearts, heart borders, changes on electrocardiogram, heart pain, cardiac arrhythmia, chronic bronchitis, chronic hepatocholecystitis, chronic gastritis, nephrolithiasis, osteochondrosis, meteodependence, the alcohol use, smoking (amount), smoking (age), working capacity, irritability, memory decline (degree), memory decline (occurrence time), vision acuity decrease (degree), vision acuity decrease (occurrence time), vision disorders, headache (nature), headache (occurrence time), dizziness. The TIA can be classified into 3 classes: TIA1, TIA2 and TIA3. The data set contains 28 patterns of TIA1, 25 patterns of TIA2, 27 patterns of TIA3 and 34 patterns of normal state without TIA.

Collected from clinic data about patients with different TIA diseases create training data set for neural network model.

5. Proposed Model

Let's examine the neural network model for detection and recognition of transient ischemic attacks. The proposed model is based on two different neural networks. The 38 features mentioned above are used as input vector, which contains information about patient. The output data of neural network model represent the 4-dimensional vector, where 4 is a number of TIA classes plus normal state. The data processing consists of two stages. The first stage of data processing is feature selection. The important question concerning input data is the following: which input parameters are really useful and contribute significantly to the performance of neural networks? The backward stepwise method or genetic approach for feature selection is used as a rule in papers devoted TIA diagnostics [4,7]. In this work nonlinear principal component analysis (NPCA neural network) for significant information extraction and dimensionality reduction is proposed. It transforms 38-dimensional input vectors into 12-dimensional target vectors.

The second stage of data processing is to detect and to recognize transient ischemic attacks. Compressed on the previous step data contain the useful information from input data and are used as inputs on the second stage of data processing. The multilayer perceptron (MLP) is applied for transient ischemic attacks recognition. MLP processes compressed data to define classes of transient ischemic attack or normal state. Output layer includes four units: three for every class of pathology and one for normal state.

Thereby the neural network model consists of two neural networks: NPCA and MLP (Fig. 1).

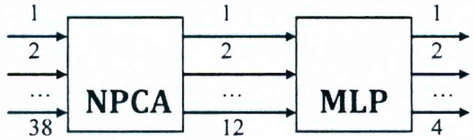

Fig.1. Architecture of the system.

Before entering to NPCA the data should be transformed by the following way:

$$x_i^k = \frac{x_i^k - \mu(x_i)}{\sigma(x_i^k)}, \qquad (1)$$

$$\mu(x_i) = \frac{1}{L}\sum_{k=1}^{L} x_i^k, \qquad (2)$$

$$\sigma(x_i^k) = \frac{1}{L}\sum_{k=1}^{L}(x_i^k - \mu(x_i))^2, \qquad (3)$$

where L is training data set dimension.

After training the neural network model have ability to transient ischemic attacks recognition.

6. NPCA Neural Network

In this section is presented neural network based on nonlinear principal component analysis technique, namely, NPCA neural network.

Let's consider an autoencoder, which is also called recirculation or replicator neural network as it is shown in Fig. 2. It is represented by multilayer perceptron, which performs the nonlinear compression of the dataset through a bottleneck in the hidden layer. As we can see the nodes are partitioned in three layers. The bottleneck layer performs the compression of the input dataset.

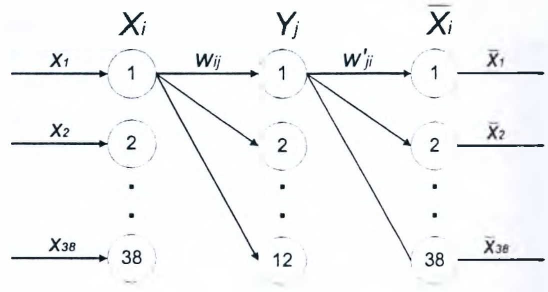

Fig.2. NPCA neural network.

There is a problem, when the principal component analysis is used: which minimum number of main components should we choose for getting maximum information from the input data. The empirical method based on the criteria of informativity can be used to determine necessary number of the main components:

$$I = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_p}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} = \frac{\sum_{i=1}^{p}\lambda_i}{\sum_{i=1}^{n}\lambda_i}, \qquad (4)$$

where n is the dimension of the input vector,

λi – eigen values of the main components, which are equal to their dispersion.

Having curve of the informativity from the main components number dependence, we can determine proper number of the main components (fig. 3).

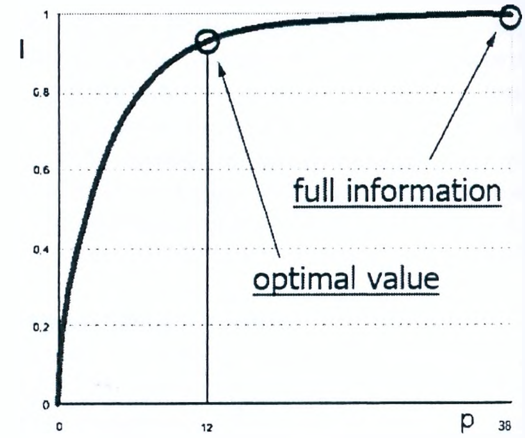

Fig.3. Dependence of the informativity from the main components number.

According to the fig. 3 the number of the main components that is necessary for the sufficient informativity of the compressed data is equal to 12.

Let's consider the dependence of the output data from the inputs for the NPCA neural network.

The j-th hidden unit output is defined as

$$y_j = F(S_j), \qquad (4)$$

$$S_j = \sum_{i=1}^{38} w_{ij} \cdot x_i, \qquad (5)$$

where F is activation function; Sj is weighted sum of the j-th neuron; wij is the weight from the i-th unit to the hidden j-th unit; xi – i-th unit input.

The i-th output unit is given by

$$\overline{x}_i = F(S_i), \qquad (6)$$

$$S_i = \sum_{i=1}^{12} w'_{ji} \cdot y_j . \qquad (7)$$

Let's consider the NPCA neural network training. The backpropagation technique for training NPCA is used. The weights are updated iteratively in accordance with the following rule:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha \cdot \gamma_j \cdot F'(S_j) \cdot x_i, \qquad (8)$$

$$w'_{ji}(t+1) = w'_{ji}(t) - \\ -\alpha \cdot (\overline{x}_i - x_i) \cdot F'(S_i) \cdot y_j \qquad (9)$$

where γ_j is the error of j-th hidden unit:

$$\gamma_j = \sum_{i=1}^{38} (\overline{x}_i - x_i) \cdot F'(S_i) \cdot w'_{ji}, \qquad (10)$$

and $F'(S_j)$ – a derivative of nonlinear activation function on the weighed sum.

The weights data in the hidden layer must be reorthonormalized by using the Gram-Schmidt procedure, as follows:

1) The first vector of the orthonormal frame is chosen as:

$$w'_1 = [\frac{w_{11}}{|w_1|}, \frac{w_{21}}{|w_1|}, \ldots, \frac{w_{n1}}{|w_1|}], \qquad (11)$$

where

$$|w_1| = \sqrt{w_{11}^2 + w_{21}^2 + \ldots + w_{n1}^2} \; . \qquad (12)$$

2) The subsequent weight vector is defined by the following recurrent formulas:

$$w_i = w_i - \sum_{j=1}^{i-1} (w_i^T \cdot w'_j) \cdot w'_j, \qquad (13)$$

$$|w_i| = \sqrt{w_{1i}^2 + w_{2i}^2 + \ldots + w_{ni}^2}, \qquad (14)$$

$$w'_i = [\frac{w_{1i}}{|w_i|}, \frac{w_{2i}}{|w_i|}, \ldots, \frac{w_{ni}}{|w_i|}], \qquad (15)$$

where i=2..12.

After training the NPCA neural network can perform orthogonal compression of the input data set.

7. Multilayer Neural Network

As it has been mentioned before the architecture of the neural network for TIA recognition used in this paper is the multilayered feed-forward network with 12 input units, 5 hidden units and 4 output units.

The activation function for each unit of hidden and output layers is sigmoid function. The number of hidden units corresponds to dimension of compressed data and the number of output units corresponds to number of classes TIA and normal state. The number of hidden units was defined by experimental way.

The backpropagation algorithm is used for training multilayer perceptron. Output value of a neural network is the number in a range from 0 up to 1 which characterizes probability of diagnostics for corresponding class of TIA.

8. Results and Discussion

For training and testing proposed neural network model the clinical observations of 114 patients with 38 parameters have been used. At the beginning the experiments with NPCA neural network have been performed using backpropagation algorithm together with the Gram-Schmidt procedure.

Let's consider the mapping of input space data for normal state and TIA classes of attack on the plane of two principal components (Fig.4).

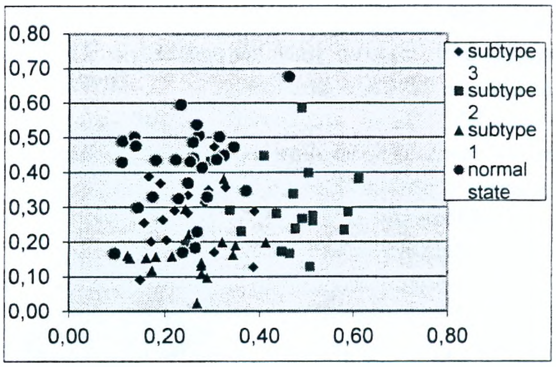

Fig.4. Data visualization on plane of two principal components.

As can be seen from the Fig. 4 the data, which belong different types of attacks are located in compact areas.

The mapping of input space data for normal state and transient ischemic attacks on the plane of three principal components is shown in the Fig.5.

The all data set have been divided into 2 groups: the training data set and testing data set. The recognition accuracy is 100% using training data set and 78% using testing data set (Table 1).

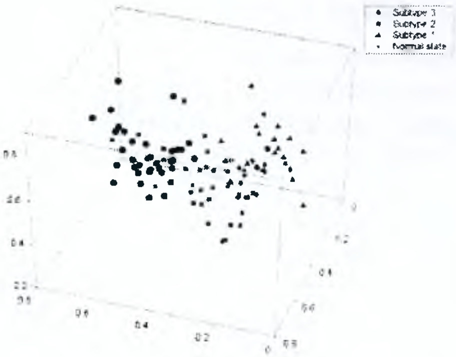

Fig.5. Data visualization on plane of three principal components.

Table 1.
Recognition Accuracy

Number of patterns in training data set	90
Number of patterns in testing data set	24
Recognition accuracy on training data set	100%
Recognition accuracy on testing data set	78%

For estimation of TIA recognition quality can be used the follow characteristics:

- FP (false positives) – the number of patterns with normal state recognized as TIA. They characterize false recognitions;
- FN (false negatives) – the number of patterns with TIA recognized as normal state. They characterize undetection of TIA;
- TP (true positives) – the number of right recognized patterns with TIA;
- TN (true negatives) – the number of right recognized patterns with normal state.

Then the sensitivity that characterizes the probability of right classification of patterns with TIA is determined as

$$TPR = Se = \frac{TP}{TP+FN} = \frac{TP}{P}, \qquad (17)$$

The specificity – probability of right classification normal (without TIA) patterns – is determined as:

$$Sp = \frac{TN}{TN+FP} \qquad (18)$$

Accuracy of classification, characterized quality of the system in general, can be calculated as:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} = \frac{TP+TN}{P+N} \qquad (19)$$

Values of the statistical parameters, that characterize classification quality, calculated on the recognition results of patterns, which are not used in the training sample, are shown in the table 2.

Table 2.
Values of the statistical parameters

Statistical parameters	Values
Specificity	89%
Sensitivity	73%

Significant difference of the accuracy on two samples (training and testing) is because of small number of patterns in the training sample. That's why there is a problem: how to train the system for good classification having small learning sample. This problem can be solved with generating of new patterns for the training sample.

Additional patterns for increasing of dimension of the training sample were generated under the following scheme:

- NPCA network was learned on the input data;
- Random error in the range of 30% was inserted in the least informative component in the hidden layer (12th component);
- Restoring of the patterns was made after error adding;
- New patterns are compared with one of the TIA classes with the help of the trained on the input data multilayer perceptron. Generated patterns which don't have distinct belonging to any class were removed;
- The generated patterns were located in the training sample, and the initial data in the test;
- Neural networks were trained again on the generated patterns and checked on the initial set.

This scheme is shown on the fig. 6.

With the help of this procedure 1000 new patterns which belong to different types of TIA were generated. Results of the system testing using generated patterns are shown in the table 3.

According to the table 3 recognition accuracy of TIA increased significantly.

Values of sensitivity and specificity for the case with generated patterns are shown in the table 4.

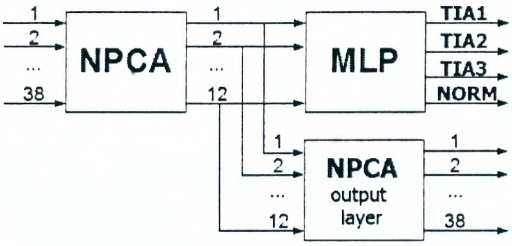

Fig.6. New patterns generating.

Table 3.
Recognition results with generated patterns

Number of patterns in the training sample	1000
Number of patterns in the test sample	114
Recognition accuracy on the training sample	98%
Recognition accuracy on the test sample	92,2%

Table 4.
Values of the statistical parameters

Statistical parameters	Values
Specificity	90%
Sensitivity	97,1%

A model with high sensitivity gives often true result in case of TIA presence (detects disease correctly). On the contrary a model with high specificity more often gives true result in the absence of TIA. Concerning a medicine sensitive diagnostic test is shown in overdiagnosis – prevention of sick patients skipping. Diagnostic test with high specificity defines sick patients very well. It is important in the cases when treatment of a patient is connected with serious side effects and overdiagnosis is very undesirable.

9. Conclusion

In this paper the neural network model for transient ischemic attacks recognition have been addressed. The proposed approach is based on integration of the NPCA neural network and multilayer perceptron. The dataset from clinic have been used for experiments performing. Combining two different neural networks (NPCA and MLP) it is possible to produce efficient performance in terms of detection and recognition of transient ischemic attacks. Neural networks technique permits to reduce the diagnostic time and the number of misdiagnosis, as well as to assist a doctor in making decision.

Bibliography

[1] Paulo J. Lisba, Azzam Taktak. *The use of artificial neural networks in decision support in cancer: A systematic review*, Neural networks 19 (2006), p. 408-415.

[2] A. McNeill. *How Accurate Are Primary Care Referral Letters For Presumed Acute Stroke?* Scottish Medical Journal, V. 53, № 4, 2008. p. 11-12.

[3] J. Easton et all. *Definition and evaluation of transient ischemic attack. Stroke* (Journal of the American Heart associations), 2009.

[4] J. Dawson et all. *A recognition tool for transient ischemic attack*, O J Med, 192, p. 43-49, 2009.

[5] P. Lisba. *A review of evidence of health benefit from artificial neural networks in medical intervention*, Neural networks 15 (2002), p. 11-39.

[6]. R. Barnes, J. Toole, J. Nelson, V, Howard. *Neural networks for ischemic stroke*, Journal of stroke and cerebrovascular diseases, Vol.15, No.5, p. 223-227, 2006.

[7] D. Shanthi, Dr.G. Sahoo, Dr.N. Saravanan. *Input feature selection using hybrid neuro-genetic approach in the diagnosis of stroke disease*, IJCSNS International Journal of Computer Science and Network Security, Vol.8 No.12, December 2008, p.99-107.

[8] D. Shanthi, Dr.G. Sahoo, Dr.N. Saravanan. *Designing an artificial neural network model for the prediction of thrombo-embolic stroke*, International journal of biometric and bioinformatics, Vol.8, No.1,2008, p. 10-18.

Authors:

Henadzi Vaitsekhovich,
PhD student
Brest State Technical University,
email: *napster@tut.by*

Supervisor:
Prof. Vladimir Golovko,
Doctor of Technical Sciences
Brest State Technical University,
email: gva@bstu.by