ISSN 1060-992X, Optical Memory and Neural Networks (Information Optics), 2009, Vol. 18, No. 4, pp. 304–311. © Allerton Press, Inc., 2009.

Application of Adaptive Classification of Tensotremorograms for Revealing the Pathological States of Human Motor Control System[1]

Y. A. Kuperin[a], A. S. Minin[b], A. A. Mekler[c], and S. P. Romanov[d]

[a]*Saint-Petersburg State University*
e-mail: yuri.kuperin@gmail.com
[b]*OOO Siemens, Monitoring and Preventive Control*
e-mail: alexey.minin@siemens.com
[c]*Institute of Human Brain RAS*
e-mail: mekler@yandex.ru
[d]*Pavlov Institute of Physiology RAS*
e-mail: romanovsp@mail.ru

Received on June 13, 2009; in final form, September 17, 2009

Abstract—In this paper the adaptive binary classifier is applied for the classification of the tensotremorogramm (TTG) time series. The idea is to reveal pathological states of human motor control system. Adaptive binary classifier being a new type of trained classifiers can be trained on the data for healthy subjects. Then the trained classifier can be used for the examinees division into healthy and sick patients. It is shown, that the trained adaptive binary classifier is able to classify the patients with acceptable accuracy. Other method of classification—Neural Clouds—has also been used. The comparison both methods has been done.

Key words: neural network, tremor, Parkinson disease, classification, isometric muscle force, adaptive binary classifier method, Neural Clouds binary classifier method.

DOI: 10.3103/S1060992X09040092

1. INTRODUCTION

The disorders in activity of human motor control system, in particular Parkinson's disease, relate to a class of socially significant illnesses. In elderly age, this kind of neurological syndromes is most widespread, following the epilepsy and dementia or cerebral-vascular injuries. The Parkinson's disease develops rather slowly. Its revealing is difficult at the early disease stages, while early diagnostics improves the neurological prognosis of treatment. As far as the risk of the Parkinson's syndrome affects the certain groups of the population (elderly men, as well as the groups of risk, for example, which are exposed by chronic bacterial poisoning or occupational intoxication), the task of early monitoring and prophylactic medical examination among these groups is important today. This will make it possible the duly diagnostics of disease.

One of the diagnostic methods is the analysis of tremor by assessment of parameters of isometric muscle force, sustained by hands [1]. The essence of a method is that the examinee applies the certain force to the force gauge, which output electric potential is proportional to the applied pressure. The received signal—tensotremorogram (TTG) is being converted into the digital form and saved for the further processing. Earlier we showed, that in cases of the motor control system dysfunction, the characteristics of TTGs from patients differ from ones from healthy subjects [2–4]. It is necessary to note, that the method described in [1], differs principally from the frequently used methods of tremor registration as fingers ballistics by the accelerometers.

Adaptive Binary Classifier (further ABC) approach to the TTG classification gives an opportunity to develop a diagnostic system that will assess functional state of human motor control system in an automatic mode. Such system would allow monitoring and carrying out clinical examination among the wide segments of population in short terms and without recruiting of the highly skilled medical personnel.

[1] The article is published in the original.

Absolute values of correlations between all the possible pairs of the inputs of the ABC for the examinees

MEAN VALUE	Confidence interval	MAX VALUE	MIN VALUE
0.126	0.037	0.63	0.017

In the present work the trained binary classifier—ABC [5]—is applied for TTG classification. Previously ABC was applied to the classification of other kinds of time series [6–8]. The purpose of the present work is the adaptation of ABC method for classification of TTG signals and its implementing in diagnostics.

2. DATA PREPARATION

We used TTGs which were recorded in six different testing conditions:

(1) Pressing the gauge by the fingers, dorsal carpal side is resting on a table, palm up; subject presses the force gauge from bottom-up;

(a) with minimal force applied to the force gauge; subjects could control visually force value and correct it, if necessary (file 1);

(b) with maximum possible force; subjects could control visually force value and correct it, if necessary (file 2);

(c) with maximum possible force as above, but without visual control (file 3);

(2) Pressing the gauge by the fingers, arms straightened in front of subject, who presses the force gauge from above downwards:

(a) with minimal force applied to the force gauge; subjects could control visually force value and correct it, if necessary (file 4);

(b) with maximum possible force; subjects could control visually force value and correct it, if necessary (file 5);

(c) with maximum possible force as above, but without visual control (file 6);

All registered TTGs were filtered by the low-pass filter (48 Hz). Time series in every test have been recorded in a two text files, containing two columns in each. First file contains involuntary components, from left and right hands respectively. They were obtained by subtraction of the low-frequency components of the signal from the entire time series. Second file contains entire time series or voluntary component from left and right hands as well. In this study only involuntary components were used. Thus, we used 24 TTG time series i.e. twelve for each hand (Fig. 1).

3. DATA NORMALIZATION

We performed the data normalization into the [–1, 1] interval in order to avoid clustering problems, corresponding to the possible significant difference in data distribution. The Min-Max normalization procedure was chosen, where the minimal value corresponds to zero and maximum value corresponds to plus one respectively. In this way we equalize the classifier sensitivity to different input channels. Without such normalization method is more sensitive to inputs, where signal has higher amplitude.

The database was organized according to the Fig. 1, and contained the TTG signals of 10 sick and 10 healthy examinees. The inputs of the ABC were constructed in accordance with this database. Then the correlation coefficients between pairs of inputs were calculated. The results of the calculations are presented in the table.

Table shows low correlations between the inputs in the database. This means that none of the input time series are excess and gives us high quality of the ABC training.

4. THE ADAPTIVE BINARY CLASSIFIER METHOD DESCRIPTION

Let us describe ABC method in details. The first step is data preprocessing, as it is described above. Then preprocessed data should be processed with the advanced K—means method (AKM). Advanced K-means is a K-means [6] which is being performed iteratively, increasing by one the initial number of centers with each iteration. After some iterations one can see the saturation of the resulting number of clusters, (see Fig. 2).

The number of clusters from which saturation is started should be considered as optimal. After the AKM has performed, the centers coordinates, some adjustable parameters should be presented to the Gaussian Mixture Model Method (GMM) [2].

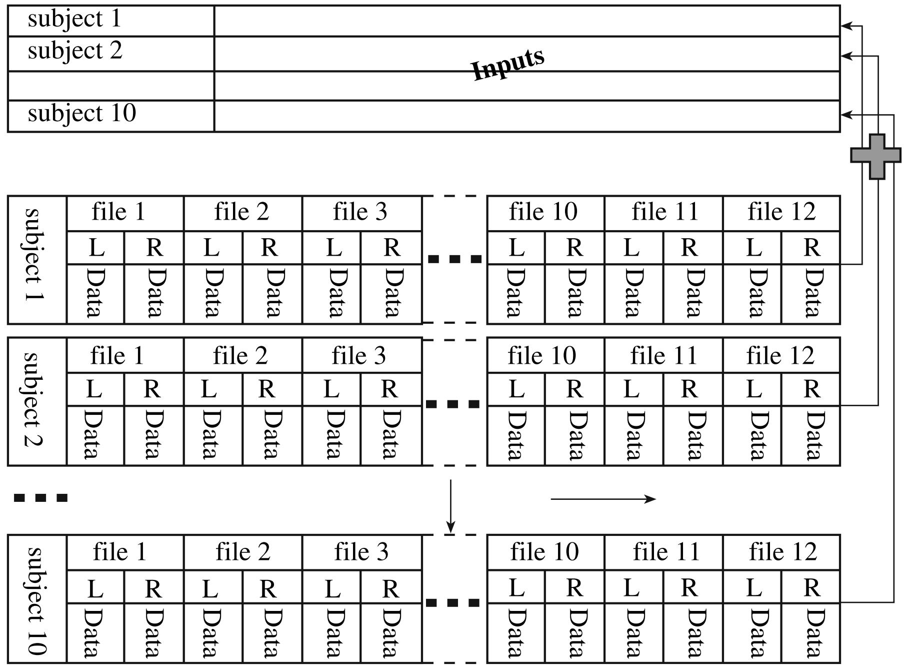

Fig. 1. The data training set for ABC training. Some data lines from this set are shown; arrows show the conformity to the patients.

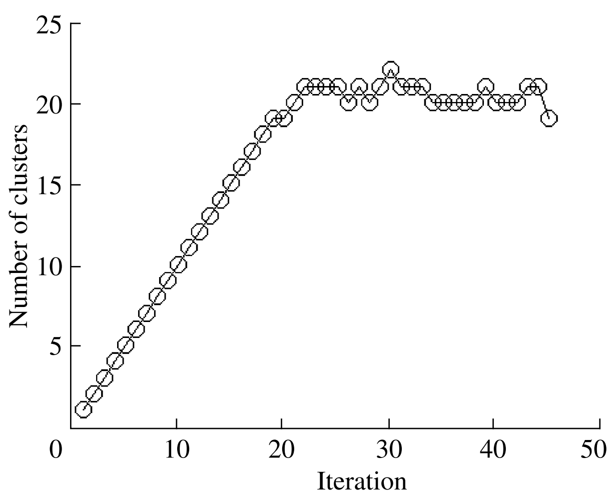

Fig. 2. Saturation of the number of clusters in the AKM algorithm.

Gaussian Mixture Model [3] is a type of density model, comprising a number of functions (components), namely Gaussians. These component functions are combined to provide a density function. Gaussian mixture models can also be considered as a form of generalized radial basis function network in which each Gaussian component is a basis function or "hidden" unit (Bayesian networks). The component priors can be viewed as weights in an output layer of the Bayesian network.

Let the conditional density for a sample data ξ belonging to a data Q be a mixture with M component densities:

$$p(\xi|Q) = \sum_{j=1}^{M} p(\xi|j)P(j), \tag{1}$$

where a mixing parameter $P(j)$ corresponds to the prior probability that sample data ξ was initiated by component j and $\sum_{j=1}^{M} P(j) = 1$, $p(\xi|Q)$ is a conditional density for the data Q, and since Q consist of M mixture components. Then $p(\xi|j)$ is the conditional density for each mixture components. Each mixture component is a Gaussian with the mean μ and the covariance matrix Σ. In the case of a 2D data we have:

$$p(\xi|j) = \frac{1}{\sqrt{(2\pi)|\Sigma_j|^d}} \exp\left(-\frac{1}{2}(\xi-\mu_j)^T \Sigma_j^{-1}(\xi-\mu_j)\right), \tag{2}$$

Here Σ_j is the covariance matrix for j-th mixture component and Σ_j^{-1} is the inverse covariance matrix for the j-th mixture components.

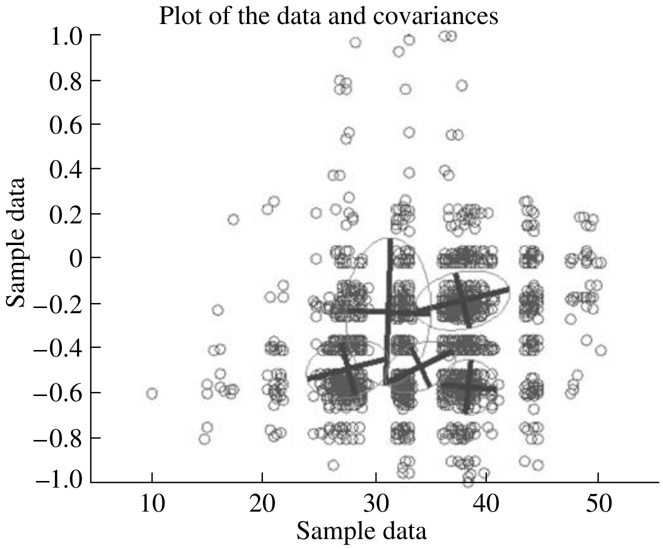

Fig. 3. Visualization of the data as small circles in 2D space. One can see normalized left vs. right hand TTGs in the file. Ellipses show GMM and the centers of the ellipses are obtained implementing the AKM.

The result of the GMM method application to the data is the surface (sum of Gaussian bells).as it is shown on Fig. 3.

The surface should be normalized to the [0; 1] interval, then its height can be interpreted as the probability of the input data to belong the ABC. The visualization of the ABC is presented in the Fig. 4 for 2D data and the 3D ABC itself. In the present study ABC was applied to the 24D real world data.

Under the expression "apply the ABC to the data" authors mean the following. ABC is already trained by the reference data, acquired from the healthy subjects, and then test data should be projected onto the N-dimensional normalized Gaussian Mixture Model (as a part of ABC), where N corresponds to the data dimensionality.

The number of clusters for the TTG data under study was selected by AKM to be 20 (see Fig. 3); the amount of AKM iterations was 40–50. Non exact number of iterations means that there is no strongly pronounced saturation and there is always some instability near the optimal number of clusters.

The data base for ABC analysis included 10 datasets from the sick subjects and 10 datasets from healthy subjects. Inputs sets were formed as it was described above (see Fig. 2). The ABC was trained by TTGs from healthy subjects.

After training is done the whole system presents its response as it is shown on the Fig. 5.

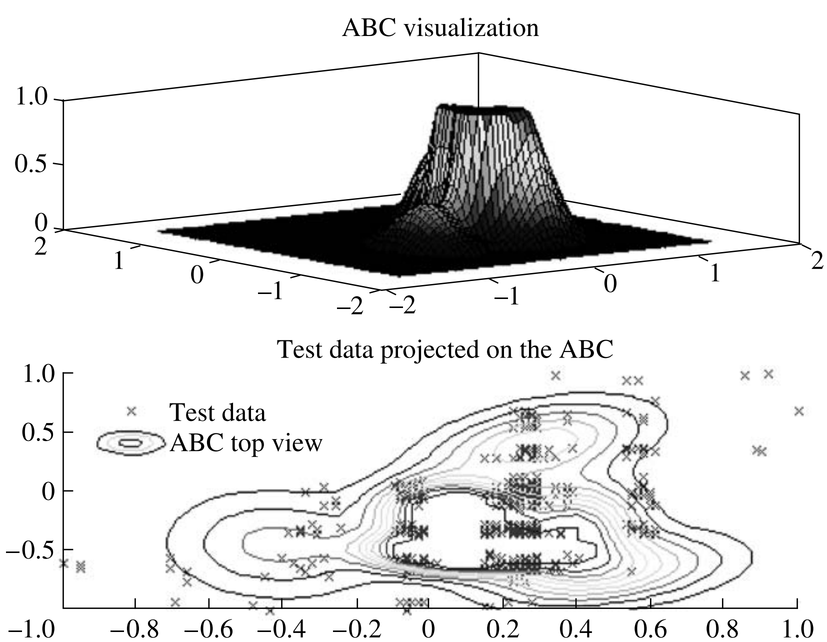

Fig. 4. Visualization of the ABC in 3D space. X axis is TTG measurements for left hand. Y axis is TTG measurements for right hand. Z axis shows the probability of the test TTG data [X1 X2] to belong the ABC. Projection of the data on to the ABC leads us to the monitoring diagram.

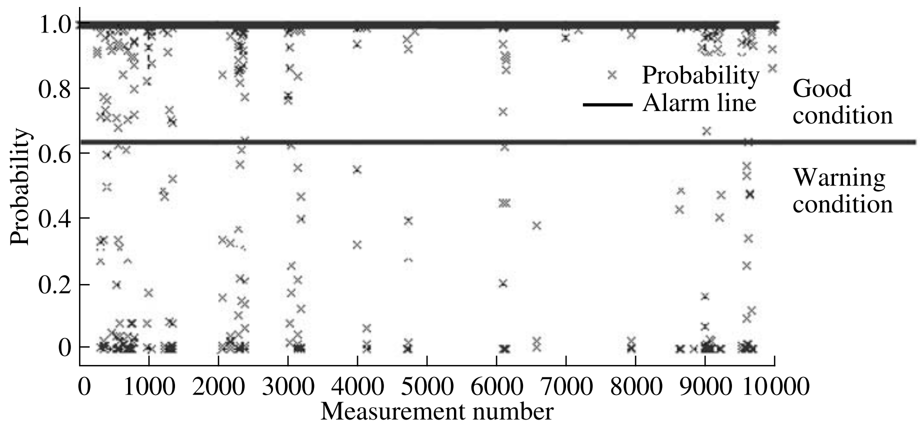

Fig. 5. Confidence along measurements for the detection of the abnormal behavior.

Here the probability values for the warning and the alarm could be set in advance by user according to his preferences. These levels are considered as criteria, according to which the subject should be classified as healthy, or sick.

This picture makes it possible to evaluate the subject and to make a decision, based on the following rule: if average probability is greater then some threshold (say, 0.5), then the subject is healthy. Moreover, it can be performed in a short time after the TTG has been recorded. It takes just a few seconds to create a plot for one subject.

5. INTERMEDIATE RESULTS

It would be reasonable to make some intermediate conclusions based on statistical TTG data manipulation and the results of the computer experiments with ABC trained on the data. Namely, the following was revealed.

According to [5, 6], the optimum number of clusters is determined as a point of saturation of the dependence of the clusters number on the number of algorithm iterations. The point, at which the saturation occurs (see Fig. 2), allows us to estimate optimum number of clusters and corresponding number of the algorithm iterations. For two-dimensional data the optimum number of clusters used by the ABC was approximately 20, which was reached by 40–50 iterations. The quantity uncertainty in the number of iterations is caused by the fact that the saturation is not stable and varies if the dimension of the data is increased up to 24.

6. NUMERICAL RESULTS AND DISCUSSION

Two numerical experiments were carried out. In the first ABC was trained by the TTG data from the healthy subjects. Then the TTG recordings both from healthy subjects, and from the patients, were processed by trained ABC with the purpose of classification, i.e. division of all set of the examinees onto the healthy subjects and patients. The second numerical experiment was carried out in the inverse order. ABC was trained by TTG data for the patients, and the classification was carried out both for the patients and for the healthy examinees.

The results of the study for the two described above experiments are given below. In the first numerical experiment the classifier was constructed and trained by data from 10 healthy subjects, and then, was tested by data from 10 patients. The result of research for one patient is given in Figs. 6 and 7. As one can see, the discrimination is possible with very high probability.

One should also take into account the other method, which can be used instead of the ABC method, namely Neural Clouds (NC) [6]. Let us consider this method and compare it with the ABC method. The normalized Neural Cloud is defined as:

$$P_c(\mathbf{x}) = \frac{\sum_{i=1}^{n} R_i(\mathbf{x})}{\sum_{i=1}^{n} R_i(\mathbf{x}) + g_0}, \tag{3}$$

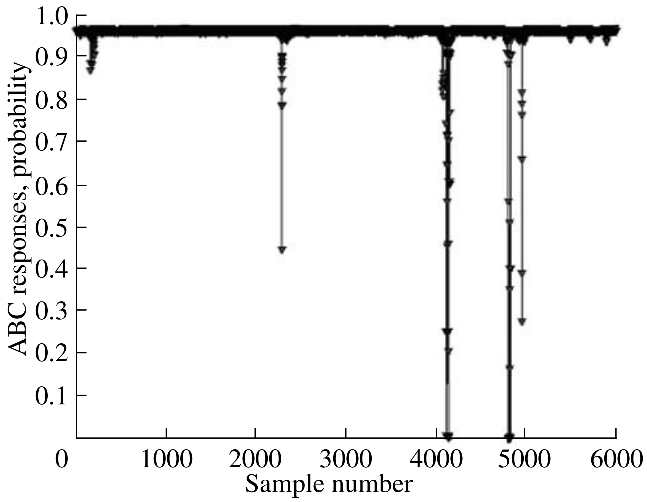

Fig. 6. The Result of the classifier work. On an axis of ordinates, the probability of belonging to ABC is indicated. The axis of abscissa showed the numbers of readings (quantization steps) in TTG recording. One can see that the probability of this patient's TTGs to belong to the set of the patients is close to 1.

where $R_i(\mathbf{x})$ are

$$R_i(x) = e^{-\frac{|x - m_i^2|}{2\sigma^2}} \tag{4}$$

the Gaussian bells, as defined by (4), and g_0 is some tuning parameter. Given input x as a matrix of inputs (patterns) and the output will be a vector of confidence levels, calculated for each element.

Adjustable parameters g_0, σ and the number of AKM clusters n are responsible for the adaptation of the NC to the given data set and determining different confidence levels. As a possible extension of the presented approach authors consider the automation of the mentioned above values selection. However, not all information for these parameters could be extracted out of the given data set. Sensitivity per dimension, the noise level in the data and the evidence of outliers in the data is critical information, usually provided by experts.

Here, it should be mentioned, that the constructed mechanism could be described in a form of Radial Basis Function network [10], as it is shown in Fig. 8. After such a network is manually

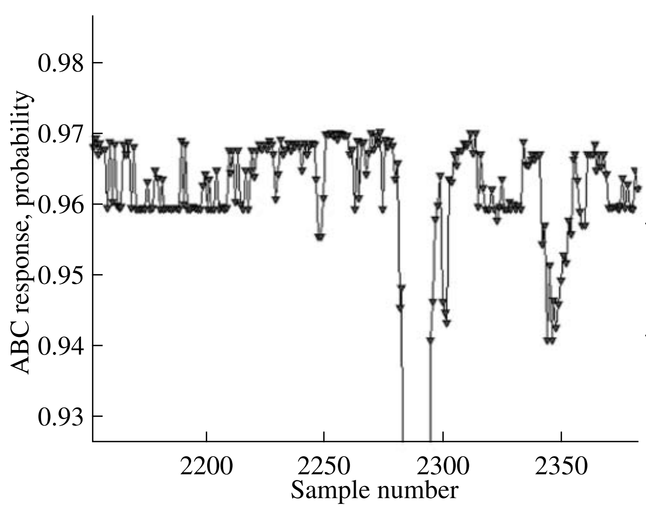

Fig. 7. The detailed fragment of Fig. 6 on which one may see that the probability of belonging of TTG from patient to the set of patients is higher than 94%.

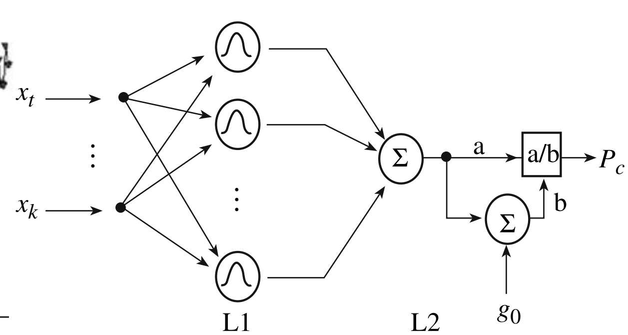

Fig. 8. Radial Basis Function network representation. L1 and L2 are Layer one and Layer two respectively. $X_{1..k}$ are input patterns and P_c is a confidence value for the concrete pattern $X_{1..k}$.

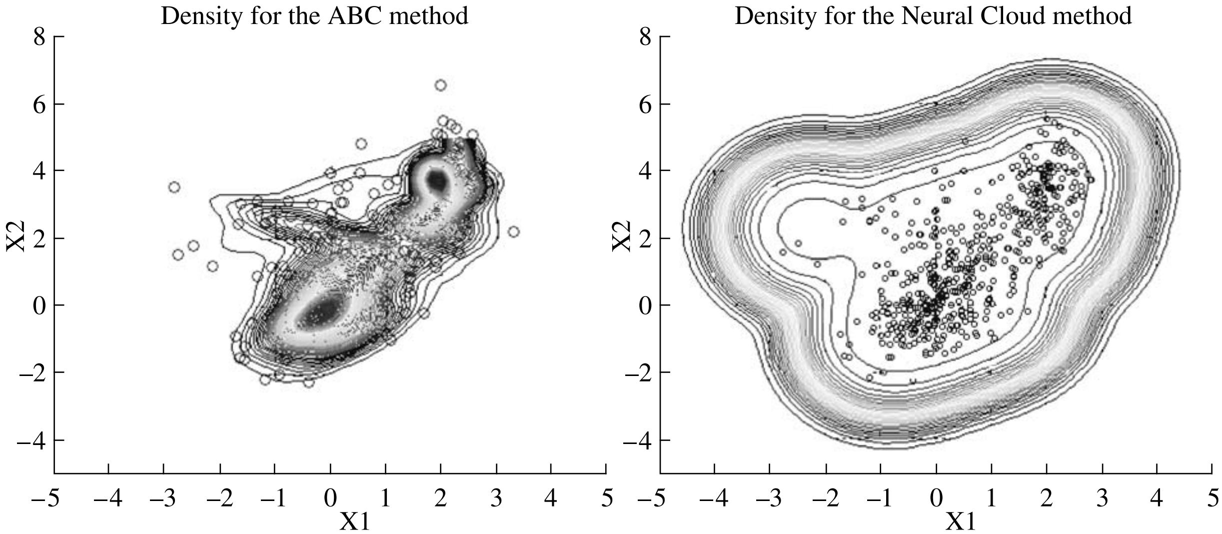

Fig. 9. Comparison of the ABC and Neural Cloud methods. One can see the difference in the sensitivity to the density of the data points.

trained by tuning mentioned parameters by the measured data, it could be used for the confidence level estimation for the new measurements.

The output vector, generated by the NC algorithm, is in a form of a confidence value between 0 and 1 that can be interpreted as a measure of the failure probability.

The characteristic features of the Neural Clouds method are:

(1) the method is not sensitive to the density of the data points;

(2) the method is the combination of the Bayesian neural network and advanced normalization technique;

(3) the method is very good in the binary classification while the amount of training data is very small.

The characteristic features of the ABC method are:

(1) the method is sensitive to the data densities;

(2) the method is a combination of the Gaussian mixture model and the advanced normalization technique;

(3) the method is very good in binary classification when the amount of training data is quiet big.

The comparison of these methods is presented below in the Fig. 9

These methods are both quite capable in binary data classification and moreover do not intersect in their application area. Both methods can be considered as complimentary parts of the binary classification complex solution. ABC method is much more sensitive to data space densities while Neural Clouds is more sensitive to the data space distribution.

7. CONCLUSIONS

Implementing the trained ABC, examinees can be discriminated into the patients and healthy subjects. It was possible to show an opportunity of the data discrimination for 10 patients in the following case.

In the present section two numerical experiments are described. In the first experiment ABC was trained using the 10 datasets, which corresponds to the healthy people. Then 10 datasets which correspond to the sick people were projected on to the ABC. Second experiment was done vise versa. The ABC was trained by 10 datasets which correspond to sick people and then 10 datasets which belong to the healthy people were projected on the ABC.

This diagram can be obtained online if the ABC is trained. One of the monitoring diagrams is presented below (see Figs. 10, 11). From the Figs. 10 and 11 it is clear that in general it is possible to distinguish different people according to the following criteria. If more then 50% of samples from the data under study

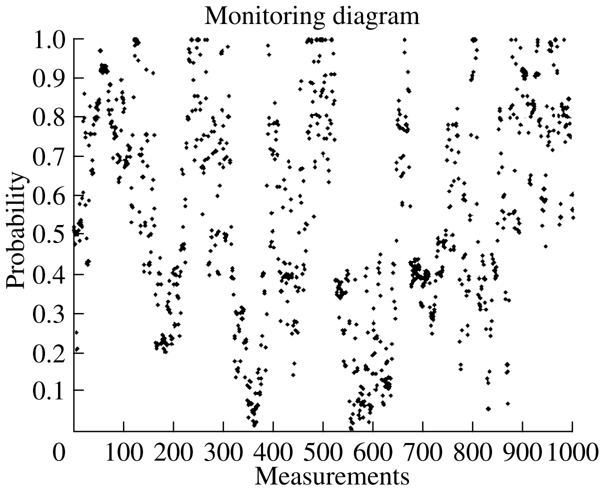

Fig. 10. The monitoring diagram obtained from the ABC. Y axis shows the probability of the sample to belong the ABC. X axis shows TTG samples. Here the dataset which belongs to the healthy man projected on to the ABC trained with the healthy man. Here more then 50% of the test data have the probability above 0.5 to belong to the "healthy group".

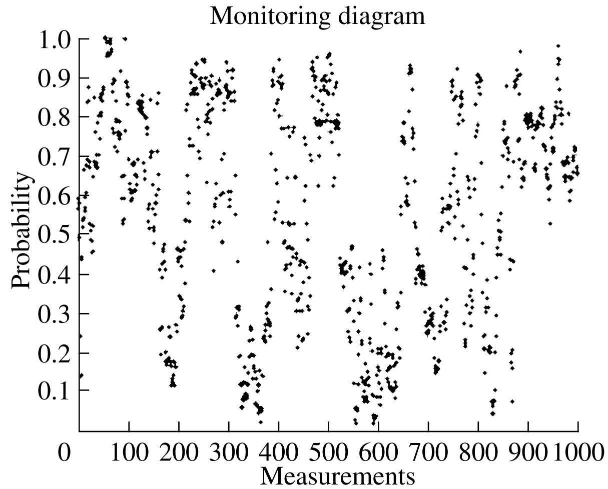

Fig. 11. Monitoring diagram for the sick people. ABC was trained with the data which corresponds to the healthy people Here more than 50% of points have the probability to belong to ABC less then 0.5.

has the probability more then 50% to belong to the "healthy" ABC (i.e. ABC that was trained by datasets from healthy subjects), then this dataset corresponds to the healthy subject.

Sometime this can not be checked visually, like on presented monitoring diagrams. But it is easy to count proportion of points, belonging to "healthy" ABC.

In order to check the ABC, it was applied to the dataset from the patient. The result is presented in the Fig. 11.

The results of the numerical experiments are the following. ABC was trained by the 10 datasets from the healthy subjects. 8 of 10 sick were classified as sick, 7 of 10 healthy were classified as healthy.

ACKNOWLEDGMENTS

The present work was supported by the grant of St. Petersburg government in the sphere of scientific and technical activity, by the grant RFBR no. 08-07-12052-ofi and the grant of Human Capital Foundation no. 185.

REFERENCES

1. Romanov, S.P. and Manoylov, V.V., RF Patent no. 2195869, 2003.
2. Romanov, S.P. and Pchelin, M.G., The Motor Control Output Forming in Healthy Subjects and Parkinsonian Disorder Patients, in *Basal Ganglia and Thalamus in Health and Movement Disorders*, Kultas-Ilinsky, K. and Ilinsky, I.A., Eds., N.Y.: Kluwer, Academic, Plenum Press, 2001, pp. 293–305.
3. Romanov, S.P., Aleksanyan, Z.A., and Manoilov, V.V., Characteristics of Tremor in Normal Subjects and in the Diagnosis and Treatment of Parkinsonism, *Neurosci. Behav. Phys.*, 2004, vol. 34, no. 4, pp. 389–398 [*Rossiiskii Fiziologicheskii Zhurnal imeni I.M. Sechenova*, 2002, vol. 88, no. 10, pp. 1356–1368].
4. Romanov, S.P., Aleksanyan, Z.A., Lyskov, E.B., Merkulova, N.A., and Romanova, L.I., Correlates of Measures of Voluntary Force with the Functional State of the Motor System, *Neurosci. Behav. Phys.*, 2006, vol. 36, no. 4, pp. 391–401. [*Rossiiskii Fiziologicheskii Zhumal imeni I.M. Sechenova*, 2005, vol. 91, no. 5, pp. 488–501].
5. Lang, B., Poppe, T., and Runkler, T., *Application of Artificial Intelligence in Steel Processing*, Automatisierung in der Metallurgie, Heft 89 der Schriftenreihe der GDMB, 2001.
6. Lang, B., Poppe, T., Minin, A., Mokhov, I., Kuperin, Y., Mekler, A., and Liapakina, I., Neural Clouds for Monitoring of Complex Systems, *Opt. Memory Neural Networks (Information Optics)*, 2008, vol. 17, no. 3, pp. 183–192.
7. Lang, B., Mokhov, I., and Minin, A., Neural Clouds for Monitoring of Complex Plant Conditions, Neuroinfomnatics-2009, *X All Russian Scientific Conference, Proceedings* (Moscow, MEPHI, 2008), part 1, p. 125–132.