Kuperin Yuri**, Alexey Mekler ***, Bernhard Lang*,
Ilya Mokhov **, Alexey Minin***

* OOO Siemens Corporate Technology, ** Saint-Petersburg State University, *** Institute of Human brain RAS
yuri.kuperin@gmail.com, mekler@narod.ru, bernhard.lang@siemens.com,
ilya.mokhov@siemens.com, alexey.minin@siemens.com

CLASSIFICATION OF EEG RECORDINGS WITH NEURAL CLOUDS

Abstract. Classification of EEG recordings is a challenging task for neurophysiologists due to the complexity of the data. EEG recordings help neurophysiologists in monitoring the status of the subject in order to determine and prevent possible diseases, which have an impact on the EEG measurements. In this paper the possible application of a novel machine learning method for the automated one-side classification is presented. The Neural Clouds (NC) is the data encapsulation method, which provides a confidence measure regarding classification of EEG recordings. Main advantage of the NC method is that the algorithm requires only data which corresponds to the normal conditions of the subject.

I. INTRODUCTION

Modern electroencephalography uses constantly expanding range of signal processing algorithms. Along with wide developing of computer systems in last decades this makes it possible to extract more information about brain processes in wide range of studies. Numerous papers on quantitative electroencephalography have been published [1, 2]. These studies are very important for understanding the nature of processes in human brain. Nevertheless, inverse problem meets many difficulties. In situations, when it is necessary to make diagnostics of disease, state or peculiarities of the person examined; qualitative methods are still most effective.

The main disadvantage of these methods is the necessity of qualified neurophysiologist who can make subjective assessment of EEG recording. In addition, subjective assessment cannot be instant and be used in long time monitoring. One of the possible ways to solve this problem is the usage of intellectual systems of classification. Their successful development will lead to possibility to express diagnostics of brain diseases and building of diagnostic, monitoring systems etc.
Modern studies on diagnostic systems based on artificial neural networks show the potential of such methods [3, 4, 5, 6, 7]. In present work, we suggest a new approach in these methods.

II. THE BASIS OF THE NEURAL CLOUDS ALGORITHM

For the classification of the EEG recordings, a novel technique, namely Neural Clouds, has been chosen. The NC concept has been successfully developed and has been applied by the Corporate Technology Department of Siemens for the optimization of steel production [8]. However, this technique has been also transferred, with the necessary modifications, to the EEG analysis field. The details of NC algorithm are presented in [9]. Nevertheless, the main ideas of the method are briefly discussed below. The concept, which stands behind the term NC, consists in creation of the effective mechanism for the data encapsulation and the so-called one-side classification, using the advanced k-means clustering algorithm (further AKM) and the expanded Radial Basis Function approach. The basic idea of the one-side classification usage in the field of the EEG recordings analysis is that most of the data which could be measured corresponds to the normal conditions of the subject, while the data acquisition related to the particular disease is not always possible due to certain reasons. The system has been trained on the given data set, which corresponds to normal operational modes of a brain and supports detection of fluctuations from these conditions. Such approach does not provide any information regarding the nature and reasons of the deviation from the trained normal state, however even the warning which appears immediately after some changes are detected could significantly help the experts.

A simple example of the abovementioned classification is shown in the figure 1. Here, the items of data are shown with circles, and the confidence levels are shown with lines. The plateau on the figure 2 corresponds to the conditions, which assumed to be normal according to the training procedure, with the probability equal to 1. The slopes show how the probability decays to 0. When NC is constructed it could be applied for classification of the new measured state of the subject, and for the estimation of the confidence value. Then the probability of potentially unexpected situation can be calculated with a simple expression: $P_f = |1 - Conf|$, where $P_f$ is probability of failure and $Conf$ is confidence level.
III. REAL WORLD APPLICATION

In order to show how powerful the method can be authors decided to apply this method for the classification of one of the most difficult time series types, namely EEG. Classification of the EEG recordings is very difficult task.
The idea behind the classification is to say whether the EEG belongs to healthy subject or not, or to evaluate the state of the subject during the EEG registration. In present abnormal fragments in EEG can be firmly identified only by subjective evaluation made by neurophysiologist. As the first experience it was decided to try to distinguish the EEG of the man with open eyes from the EEG of the man with closed eyes. EEG’s, recorded in these two states can be easy distinguished subjectively. We used EEGs, recorded from 19 electrodes, placed according to the international 10-20 system [10]. The idea was to build NC (in 19 dimensional space) for the EEG recording of the man with open eyes and then to present the EEG of the same man with closed eyes to the trained NC in order to obtain the answer whether the new points belong to the NC or not without inducing the prior knowledge about the EEG time series into the algorithm. The results presented in the figures 3, 4, 5 and in the table 1.

In the following “training set” (see fig. 3) means that some measurements have been randomly taken from the EEG recorded with open eyes, which was used for the NC creation. Test set (see fig. 4) stands for the projection of the measurements on the NC, which is not used for NC creation, but which corresponds to the EEG recorded with open eyes. Generalization set (see fig. 5) means that the data, which corresponds to the EEG recorded with closed eyes, have been projected on NC constructed using the EEG data, which correspond to the EEG recorded with open eyes. Probability means that with such probability the measurement projected on NC belongs to the EEG recorded with open eyes.

Fig.3: Training set. Ratio of measurements which do not belong to NC (P<0.5) is 3.78%
Fig. 4: Test set. Ratio of measurements which do not belong to NC (P<0.5) is 7.6%

Fig. 5: Generalization set. Ratio of measurements which do not belong to NC (P<0.5) is 57.6%

For more details about the computational experiment see table 1.
<table>
<thead>
<tr>
<th>Parameter description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of points which do not belong to the NC when the subject closes eyes (ideal case 100%)</td>
<td>58%</td>
</tr>
<tr>
<td>Number of AKM centroids</td>
<td>53</td>
</tr>
<tr>
<td>Computation time to configure NC (CPU:1.86Ghz, RAM: 4 GB DDR2)</td>
<td>25 sec</td>
</tr>
<tr>
<td>Evaluation results</td>
<td></td>
</tr>
<tr>
<td>Training set</td>
<td>3.8%</td>
</tr>
<tr>
<td>Test set</td>
<td>7.8%</td>
</tr>
<tr>
<td>Generalization set</td>
<td>57.6%</td>
</tr>
</tbody>
</table>

Tab. 1: EEG with open and closed eyes. Results obtained in computational experiment. Here, the field “Number of AKM centroids” emphasizes the sensitivity of this algorithm with respect to this parameter.

To make the obtained results more evident the above mentioned technique has been applied for another dataset. These are EEGs, recorded before and during epileptic seizure. They correspond to two different states of the brain (in context of epilepsy). These two kinds of EEG recordings also can be easy distinguished visually. Results one can see in figures 6, 7 and 8 and in the table 2 below.

![Training set for the EEG, recorded during seizure](image)

Fig.6: Training set. Ratio of measurements which do not belong to NC (P<0.5) is 0.30% (ideal case would be 0%)
Fig. 7: Test set. Ratio of measurements which do not belong to NC (P<0.5) is 0.93% (ideal case would be 0%)

Fig. 8: Generalization set. Ratio of measurements which do not belong to NC (P<0.5) is 95.7%

For more details about computational experiment see table 2.
<table>
<thead>
<tr>
<th>Parameter description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of points which do not belong to the NC when the subject is in seizure (ideal case 100%)</td>
<td>95.7%</td>
</tr>
<tr>
<td>Number of AKM centroids</td>
<td>60</td>
</tr>
<tr>
<td>Computation time to configure NC</td>
<td>65 sec</td>
</tr>
<tr>
<td>Evaluation results</td>
<td></td>
</tr>
<tr>
<td>Training set</td>
<td>0.3%</td>
</tr>
<tr>
<td>Test set</td>
<td>0.93%</td>
</tr>
<tr>
<td>Generalization set</td>
<td>95.7%</td>
</tr>
</tbody>
</table>

Tab. 2: EEG before and after seizure. Summary of all results obtained in computational experiment.

As one can see from the results above, this method have easily separated the EEG recordings, made before seizure, from the EEG recordings, made during seizure.

IV. CONCLUSION

The application of the efficient one-side classification algorithm is proposed with the EEG examples. The advantage of this algorithm is that it can be trained with data which correspond to the normal conditions. The results obtained in the real world experiments with EEG recordings give the possibility to develop so-called complete human health estimators. If one could collect all possible external parameters, which are easy to measure, it would be possible to construct NC which corresponds to the healthy condition of the patient. Therefore, it would be possible to organize inspection in the hospital at once and as a result, one could obtain the probability of how far is the condition of the health from the normal.

Acknowledgment

Authors are grateful to N. Koroleva for supplying authors with EEG recordings from epilepsy patients.

References


Опубликовано в: