fetal monitoring,
neural networks,
pattern classification

Michał JEŻEWSKI[*], Norbert HENZEL[**], Janusz WRÓBEL[***], Paweł ŁABAJ[*]
Krzysztof HOROBA[***], Adam MATONIA[***]

APPLICATION OF NEURAL NETWORKS FOR PREDICTION OF FETAL OUTCOME

In present-day obstetrics cardiotocographic monitoring (CTG) is a basic biophysical method of the fetal state assessment. The reproducible and objective assessment of CTG traces is ensured by computer-aided fetal monitoring systems. However, in clinical centers, there are great needs for automated conclusion generation methods. The properties of neural networks (NN) for handling complex data and capabilities for generalization make them attractive for prediction of fetal outcome. The research material included 685 traces from 189 patients. The final set of 7 parameters of quantitative description of CTG trace were used as input variables. From newborn description four features were used to the NN output classification. Several experiments with data representation and interpretation were done. For the final neural networks obtained sensitivity and specifity indices exceeded 80%. During the experiments some answers to significant problems were received. These solutions being universal can help the application of neural networks in different areas of medical diagnostics.

1. INTRODUCTION

Cardiotocographic monitoring (CTG) is a basic, biophysical method of assessment of the fetal state. It consists in simultaneous acquisition of the three signals: fetal heart rate (FHR), uterine construction activity and fetal movements. Classical visual analysis of CTG traces is rather difficult because of the complex shape of recorded signals. The interpretation of the records is subjective and it depends on experience and emotional state of the clinician. There are high interobserver and intraobserver disagreement of trace assessment.

Analysis of fetal heart rate variability is the most important. the FHR variability is described by two main components: basal level a variability with acceleration/deceleration patterns and instantaneous variability. The instantaneous variability contains the important diagnostic information which is hidden for naked eye and can be quantitatively described with a help of dedicated computer-aided systems. The reproducible and objective assessment of the fetal state based on antenatal CTG records is of particular importance in the so called high risk pregnancy, when thanks to the early diagnosis the appropriate decision can be made. Computerized fetal monitoring systems have become a standard in clinical centers nowadays. They offer automated quantitative analysis of acquired signals, but effective methods, which will enable conclusion

[*] Student of Institute of Computer Science, Silesian University of Technology, Gliwice, Poland
[**] Division of Biomedical Electronics, Institute of Electronics, Silesian University of Technology, Gliwice, Poland
[***] Department of Biomedical Informatics, Institute of Medical Technology and Equipment, Zabrze, Poland

generation, are still being searched. The promising approach seems to be the application of the artificial neural networks (NN). The properties of handling complex data sets, the capability of learning and generalization, and the representation of pattern recognition properties in a distributed fashion, make NNs particularly attractive for prediction of fetal outcome [1,2,4,5]. The goal of this work is to answer some important questions, which occur during neuro-computing for prediction of the fetal state, when the input variables are parameters of quantitative description of CTG traces, computed by fetal monitoring system. The main emphasis will be has been put on input data preprocessing and discussion of obtained results.

2. MATERIAL AND METHODOLOGY

The research material consisted of archive data recorded during 10 months in one of clinics using computer-aided fetal surveillance system – MONAKO (ITAM Zabrze). The material comprised 1274 traces from 341 patients (an average was 3.74 traces for one patient). After several steps of data cleaning, 685 traces from 189 patients were the final research material. These corresponded to 189 newborns, and the traces were acquired in various weeks of pregnancy. The archive included fetal CTG traces as well as medical data referring to patients and their newborns after delivery. It is assumed in obstetrics, that fetal outcome can be assigned only retrospectively to the fetal state during the course of pregnancy, i.e. when the traces were recorded. Additionally, the fetal state can not change during the pregnancy rapidly, neither for worse nor for better. The set of 21 parameters of the quantitative description of the CTG trace was chosen as input variables for NN. These parameters describe all kinds of the fetal heart rate variability, and additionally fetal movements and uterine contractions.

The one woman could have had several cardiotocographic traces if she was in high-risk pregnancy. That caused there were patients with several CTG traces. Taking account this, two main groups of traces called Multiple-Traces and Single-Traces were created. Multiple-Traces group is the group of all CTG traces, where each trace, nevermind from which patient, is treated as a separate case. Single-Traces group is the group of CTG traces fulfilling the relation that a one trace, registered as closest to the delivery as possible, is assigned to one patient. In this group the number of traces is reduced from 685 to 189. This reduction has negative influence on the learning process of network, but it prevents from too much influence of individual patient's features, e.g. when CTG traces from a given patient has specific (very high or very low) FHR variability.

The fetal outcome is described just after the delivery, by clinical staff in the newborn description form. The four main criteria together with ranges describing the normal or abnormal fetal outcome were defined (Table I). The BE and pH parameters describing acid-base fetal status as well as the birth weight are measurable data. The Apgar score is observable data and can be subjective. The following notation was used: FOE(Ki) – fetal outcome assessed by clinical expert according to Ki criterion and FON(Ki) – fetal outcome assessed by neural network according to Ki criterion. Additionally, we defined one complex criterion K5: for which the FOE(K5) = FOE(K1) OR FOE (K2) OR FOE (K3) OR FOE (K4) and the FON(K5) is fetal outcome predicted by NN. For comparison we introduced the OR function FON(OR) of results of individual networks designed for the prediction of fetal outcome according to each simple criteria: FON(OR) = FON(K1) OR FON (K2) OR FON (K3) OR FON (K4). The group of traces for which FOE (Ki)=0

is called class 0, whereas the group of traces for which FOE (Ki)=1 is called class 1. There was a small number of CTG traces in our database, for which fetal outcome was assigned as abnormal.

Table I. Simple criteria for the fetal outcome description.

Ki	Definition	Abnormal fetal outcome
K1	Neonatal birth weight, just after the delivery	Birth weight < 10 centiles
K2	Five minute Apgar score (visually assessed newborn state in range 1-10 points)	Apgar score < 7
K3	Umbilical artery pH at birth	pH < 7,20
K4	Umbilical artery base excess (BE) at birth	BE > 12 mEq/l

This distribution represents the real situation in high-developed countries, where the fetal outcome is normal in most of the cases.

Situation, when sizes of different classes in the learning subset extremely differ, is considered as unfavorable and making the classes equal is recommended. On the other hand, it is stated, that class size corresponding to the real proportions can give more correct classification by neural network. Therefore data were divided in two ways called as: Real-Proportion and Equal-Proportion. Using the Real-Proportion we did not change class sizes in sets Multiple-Traces and Single-Traces. Class sizes are different, and the difference depends on the Ki criterion. Real-Proportion (Multiple-Traces) and Equal-Proportion (Multiple-Traces) sets better represent the real situation. Using the Equal-Proportion, we obtained equal numbers of CTG traces in the class 0 and 1, by removing some traces from the 0 class. This was doned for all criteria: FOE(K1)..FOE(K5) for the Multiple-Traces set. The disadvatage of Equal-Proportion (Multiple-Traces) is the reduction of CTG traces which takes part in neural network learning process.

The Statistica Neural Networks 7.1 (StatSoft, Inc.) software was used for the development of artificial neural networks. The research material was divided into three subsets: training, validation and testing. In the case of Equal-Proportion the class sizes are equal in a given subset. For Natural-Proportion the class proportion is constant in each subset and its value corresponds to the case distribution in the research material. A number of various experiments were done to reach final results. They supported the selection of the most suitable algorithm for task realization. In each step the choice was done from many possible data representations, interpretations of inputs and outputs and structures of neural networks. Results of experiments gave answers to many significant problems. Some of them are so universal, that can be helpful during designing classification systems, using the NN, in various disciplines of medical diagnostics. Major questions were:

- Whether to design NN based on sets, where class sizes extremely differ, according to clinical practice, or sizes should be made equal?
- Whether to take all available parameters describing the normal/abnormal outcome or to make their reduction as regards the clinical value?
- Having several input data sets referring to the same patient (several different data sets referring to the same health state), is it necessary to use all of them as quite separate cases?
- What kind of network should be designed in situation, when several independent and equally significantly components compose the general description of the health state? Should it be the

network designed for the prediction of the general health state, or the logical sum of outputs of networks for predicting only one individual component will be better?

3. RESULTS

The quality of the classification is expressed using: sensitivity, specifity, positive predictive value (PVP) and negative predictive value (NVP). These indices are commonly used for description of prognostic value of a given diagnostic method underwent the testing. Comparison of results of network designed to predict the fetal outcome according to complex criterion K5, i.e. FON(K5), with the FON(OR) – function OR of results of individual networks designed for the prediction of fetal outcome according to each simple criteria, is shown by Figure 1. The difference FON(OR) – FON(K5), indicates that the better is the network designed for prediction of fetal outcome according to complex criterion K5.

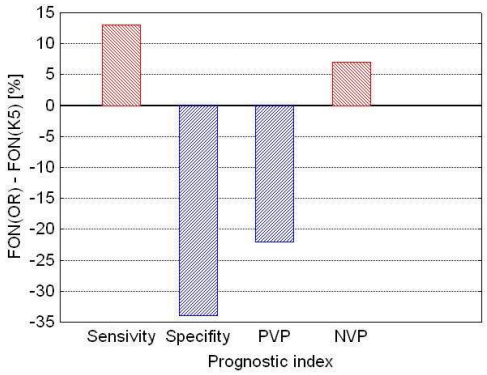

Fig. 1. Comparison of prognostic indices for FON(OR) and for FON(K5).

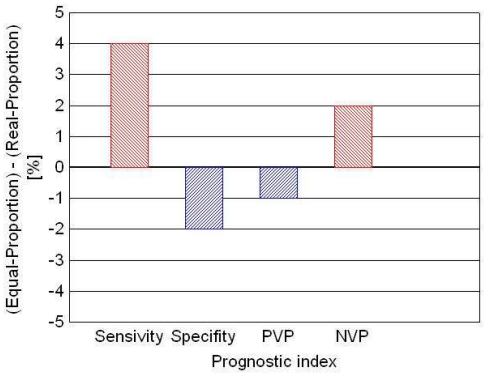

Fig. 2. Comparison of prognostic indices for the data organized as Equal-Proportion and Real-Proportion.

The influence of proportion type (Equal or Real) on the accuracy of the ANN classification, based on the set Multiple-Traces, is presented by Figure 2. Prognostic indices for Equal-Proportion were higher, but not significantly. Therefore, and from the fact, that Real-Proportion better represents reality, we decided to design further networks only basing on this data set.

A large number of various parameters for quantitative description of CTG trace caused attempt to optimize them [3]. For the network designed for prediction of fetal outcome according to the complex criterion K5, a sensitivity ratio for each of input variables was computed. It represents ratio of the network error with the variable being rejected, to the overall network error. The new optimized set of seven input variables, for whose the sensitivity indices were the highest, was created within the particular categories of FHR variability (Table II). The influence of multiple CTG traces from one patient on classification accuracy was examined in the next step. We compared the results of NN based on the Multiple-Traces and Single-Traces sets with both of the type of Real-Proportion and with the optimized set of input variables. It was proved (Figure 3) that better parameters were obtained for the network designed for the set Real-Proportion (Multiple-Traces). This corresponds to clinical practice, where usually several CTG traces are recorded for one patient.

Radial basis function (RBF) and multilayer perceptron (MLP) type networks were designed parallelly in all experiments. Better results were obtained usually for RBF networks. In the last step,

the MLP type NN was designed for the prediction of the fetal outcome according to the complex criterion K5, based on the Real-Proportion (Multiple-Traces). Table III presents results for both final networks of RBF and MLP type, for Multiple-Traces and Single-Traces sets.

Table II. The optimal set of input variables for the network designed for the prediction FON(K5) basing on the set Real-Proportion (Multiple-Traces)

Input variable name and meaning	Sensitivity index
II, Yeh index of longterm FHR variability	1,16
S/L, proportion of shortterm to longterm FHR variability	1,05
LTV, average value of one-minute differences FHR_{max} - FHR_{min}	1,03
DI-BB, Yeh index describing variability on beat-to-beat level	1,03
STI-BB, de Haan index for variability on beat-to-beat level	1,02
ACC, number of FHR acceleration patterns per one hour	1,09
BasalFHR, value describing the fluctuations of FHR baseline	1,00

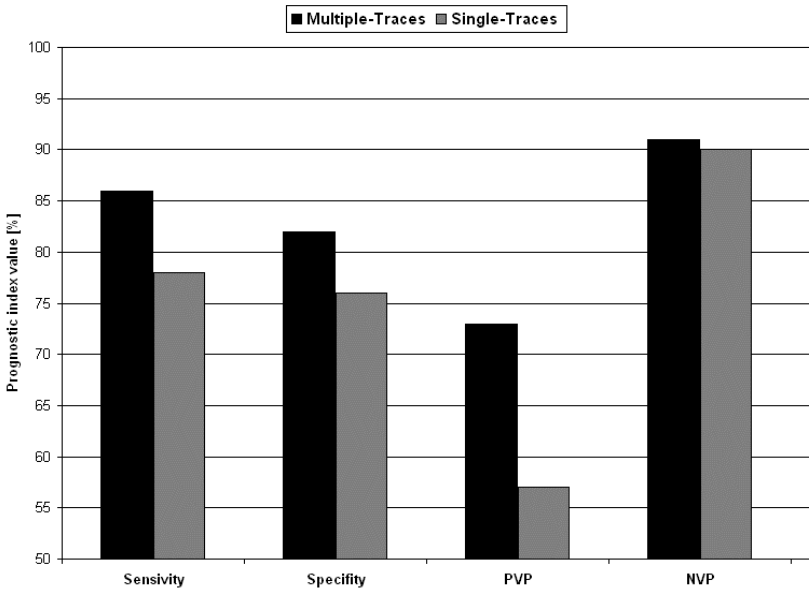

Fig. 3. Comparison of prognostic indices for networks based on Multiple-Traces and Single-Traces sets using the Real-Proportion.

Finally, the artificial neural networks with the following attributes were proposed:

- Type of the network: Radial Basis Function (RBF)
- Input data: set of CTG traces including groups of traces connected with one patient (Multiple-Traces)
- Class proportion: data set representing the reality, i.e. with proportion between the number of traces corresponding to the normal and abnormal fetal outcomes (Real-Proportion)
- Network based on optimized set of input variables, chosen as a result of sensitivity analysis of the NN on a given input

Table III. Comparison of prognostic indices (in %) for final RBF and MLP type networks.

Prognostic indices	Multiple-Traces		Single-Traces	
	RBF	MLP	RBF	MLP
Sensitivity	86	78	78	80
Specifity	82	80	76	79
PVP	73	69	57	60
NVP	91	86	90	91

4. DISCUSSION AND CONCLUSION

The results quite similar as obtained in this work results are presented in [1], where authors took an attempt to prediction of normal or abnormal fetal outcome. Large thirty-element set of parameters of quantitative description of CTG traces was used as input variables. One patient had only one trace. Abnormal fetal outcome was defined as a logical sum of occurrences of four kinds of pathological symptoms, i.e. similarly as at this work. The classification results for the testing set were as follows: sensitivity (73 %), specifity (94 %), PVP (72 %), NVP (94 %). For a comparison, for the MLP type network designed on Real-Proportion (Multiple-Traces) set, we obtained: sensitivity (80 %), specifity (79 %), PVP (60 %), NVP (91 %), see Table III. Indices values in [1], except the PVP one, are quite similar.

The artificial neural network approach for the prediction of fetal outcome based on parameters of quantitative description of CTG traces seems to be very attractive. Details analysis of research material had a strong influence on the research procedure. During various experiments the way of the data representation, interpretation functions describing of inputs and outputs of network were constantly changing. The main emphasis was put on a critical analysis of obtained results in a given step. They gave answers to many questions, which occur during development of conclusion generation systems based on neural networks for medical areas of application.

BIBLIGRAPHY

[1] Arduini D., Giannini F., Magenes G.: Intrapartum serveillance: computer cardiotocography, Proc. of 5th World Congress of Perinatal Medicine, 2001, 1217-1223.
[2] Dalton K.J.: Neuro-computing: a medical overwiew of neural networks, with particular reference to obsterics and gynecology in “A critical appraisal of fetal surveillance”, ed: H. Geijn, Elsevier, 1994, 625-632.
[3] Jeżewski J., Wróbel J., Horoba K.: Computerized perinatal database for retrospective qualitative assessment of CTG traces in „Current Perspectives in Healthcare Computing”, ed: B. Richards, BTHC, 1996, 187-196.
[4] Magenes G., Signorini M.G., Arduini D.: Classification of cardiotocographic records by neural networks, Proc.of the IEEE International Joint Conference on Neural Networks, 2000, Vol. 3, 637-641.
[5] Magenes G., Signorini M.G., Sassi R.: Multiparametric analysis of fetal heart rate: comparison of neural and statistical classifiers, Proc. of IX Conference on Medical Engineering and Computing, 2001, 360-363.

Scientific work financed from the State Committee for Scientific Research resources in 2004–2006 years as a research project No. 3 T11E 017 26.