JOURNAL OF MEDICAL INFORMATICS & TECHNOLOGIES Vol. 12/2008, ISSN 1642-6037

fetal monitoring, cardiotocography, classification, neural networks

Michał JEŻEWSKI*, Robert CZABAŃSKI*, Krzysztof HOROBA**, Janusz WRÓBEL**, Jacek ŁĘSKI*, Janusz JEŻEWSKI**

INFLUENCE OF GESTATIONAL AGE ON NEURAL NETWORKS INTERPRETATION OF FETAL MONITORING SIGNALS

Cardiotocographic monitoring (CTG) is a primary biophysical monitoring method for assessment of the fetal state and is based on analysis of fetal heart rate, uterine contraction activity and fetal movement signals. Visual analysis of CTG traces is very difficult so computer-aided fetal monitoring systems have become a standard in clinical centres. We proposed the application of neural networks for the prediction of fetal outcome using the parameters of quantitative description of acquired signals as inputs. We focused on the influence of the gestational age (during trace recording) on the fetal outcome classification quality. We designed MLP and RBF neural networks with changing the number of neurons in the hidden layer to find the best structure. Networks were trained and tested fifty times, with random cases assignment to training, validating and testing subset. We obtained the value of sensitivity index above 0.7, what may be regarded as good result. However additional trace grouping within similar gestational age, increased classification quality in the case of MLP networks.

1. INTRODUCTION

Cardiotocographic fetal monitoring (CTG) is the most common biophysical method for assessment of the fetal state. It consists in simultaneous registration and analysis of three signals: fetal heart rate (FHR), uterine contraction activity and fetal movements. Classical, visual analysis of CTG traces is rather difficult because of the complex shape of the signals. The interpretation is subjective and characterized by high interobserver and intraobserver disagreement. The repeatable and objective assessment of the fetal state is especially important in the case of high risk pregnancy, when thanks to the early diagnosis, appropriate medical management can be carried out. The fetal heart rate variability contains the most important diagnostic information, which is hidden for a naked eye, but can be quantitatively evaluated with the help of dedicated computer-aided systems. Therefore, computerized fetal monitoring systems quickly have become a standard in clinical centres [8, 10]. However efficient methods, which enable fetal state assessment based on quantitative parameters of CTG traces, are still being searched [2, 9, 11]. Evident advantages of artificial neural networks [3, 7] like properties of handling complex data sets, capability of learning and generalization and a distributed pattern recognition process, make them particularly attractive tool for these purpose. In the proposed work we tried to apply the neural networks for the prediction of fetal outcome using the set of parameters of quantitative description of CTG traces as the input data.

We focused on the influence of the gestational age (in weeks), at which the fetal signals were registered, on the fetal outcome classification quality. The reason was that the effects of gestational age on the FHR variability were reported [4, 5, 12]. Additionally, the proposed experiments will help to answer some important questions, which may occur during application of any computational intelligence methods as tools for efficient classification of the cardiotocographic signals.

2. METHODOLOGY

The research material was obtained from the archive of the computerized fetal surveillance system MONAKO [10]. It included parameters of quantitative description of recorded CTG traces as well as data referring to the patient's medical history and description of the newborns. At the beginning, the data set included 2431 traces registered from 293 patients. The data were preprocessed in two steps. In the first step, the traces of duration less than 20 minutes and with signal loss higher than 20% were rejected. Signal loss parameter represents the percentage of duration of episodes with lack of signal in the whole trace. Additionally, according to clinical expert suggestions, we also rejected traces for which the percentage of undetermined minute values of FHR variability was higher than 20%. The final rejection concerned traces from those patients, for which there was no data referring to their newborns. Such situations could happen as a result of uncompleted or even empty newborn's data forms, or when the child was delivered in another hospital than this one, where the CTG monitoring was performed.

Finally the data set included 749 traces registered from 103 patients (one patient could have had several traces), with 210 (28%) corresponding to abnormal and 639 (72%) to normal fetal outcome. The records duration time varies from 21 to 340 minutes (an average of 44 minutes), they were registered between the 28th and 41st week of gestation (an average of 36th week). As input variables for neural networks we used 17 parameters of quantitative description of the recorded cardiotocographic signals. Fifteen parameters were describing the fetal heart rate signal in time domain (the FHR baseline, acceleration /deceleration patterns, long-term variability as well as beat-to-beat variability). Additionally, the number of fetal

* Division of Biomedical Electronics, Institute of Electronics, Silesian University of Technology, Gliwice, Poland (mjezewski@polsl.pl).
** Department of Biomedical Informatics, Institute of Medical Technology and Equipment, Zabrze, Poland.

outcome: normal or abnormal. The fetal outcome was obtained from the newborn description data forms. It was assessed as abnormal, if the value of at least one of four attributes describing the newborn was outside the physiological range.

The Statistica Neural Networks 7.1 (StatSoft, Inc.) software was used for the development of artificial neural networks. We investigated two most common types of neural networks: multilayer perceptron (MLP) and radial basis function (RBF). Using both types of networks, we changed the number of neurons in the hidden layer to find the best network (with the highest generalization ability). In the case of the MLP networks, number of the hidden neurons varied from 2 to 10 (with a step of 1), and the sigmoid activation function was applied. In the case of the RBF networks, we changed the number of hidden neurons from 5 to 50 (with a step of 5). The MLP networks were trained for 500 epochs by the steepest descent gradient algorithm with the constant learning rate of 0.01. The centres of RBF neurons were determined by K–Means algorithm, the radii by K–Nearest Neighbours algorithm with 10 neighbours. Each network was trained and tested 50 times, with random cases assignment to three data subsets: training, validating and testing. The sizes of the subsets (where the training subset was twice size of the other ones) and the number of traces in classes corresponding to normal and abnormal fetal outcome were constant in each trial. The results of the experiments were represented as a set of statistical values for all the 50 trials.

For a given date of CTG recording, the monitoring system computes the gestational age basing on the date of the last menstrual period being entered to the database as the beginning of pregnancy, assuming that the delivery should occur in the 40th week of pregnancy. Typically, the gestational age is expressed in completed weeks of pregnancy, however in the database it is calculated with accuracy of days. Since the gestational week is associated with the monitoring signals, we decided to investigate in detail the influence of the gestational age on the classification quality, because with the development of the fetus, the parameters describing quantitatively the acquired signals are changing. Our investigation consisted of two approaches: decomposition of records into gestational and labour groups and introduction of the additional neural networks input. In the first approach we created four groups of traces registered in similar weeks of gestation as well as one group of traces registered directly on the day of the labour – the earliest labour was in the 30th week of gestation, the latest in 42nd. The gestational groups overlap each other to make their sizes similar to the size of the group of labour traces (286 traces), Table 1.

Table 1. Number of traces corresponding to particular fetal outcomes for gestational and labour groups.

Group	Gestational Age [weeks]	Traces Corresponding to Fetal Outcome	
		Normal	Abnormal
G1	33 ÷ 36 (34.5)*	193	92
G2	34 ÷ 37 (35.5)	195	89
G3	35 ÷ 38 (36.5)	196	83
G4	36 ÷ 41 (38.5)	198	80
Labour	Day of labour	163	123

* Group centre

In the second approach we use all traces in the learning process, but with additional neural networks input describing the gestational age, at which the trace was registered. Two possibilities of the information representation were investigated. In the first one, the gestational age corresponding to a given CTG trace was converted into the number of one of the four groups of antenatal traces established previously. For a given trace with gestational age expressed with an accuracy of days, age distances (absolute differences) to centres of all the four groups were determined. The number of the group, for which the minimal distance was obtained, was assigned to the given CTG trace. If the absolute difference was equal for two groups, the group with the highest number was chosen. The second representation of the information on the gestational age was simply applied as the number of completed weeks of gestation.

3. RESULTS

The efficiency of the designed neural networks is evaluated by a set of prognostic indices: Sensitivity, Specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV). Indices are computed basing on the confusion matrix obtained for the testing data subset. From a clinical point of view, particularly important is the high value of the Sensitivity index, which relates to false negative cases having more serious consequences then the false positive ones. However the Specificity index has also influence on the classification quality, so we introduced the logarithmic Quality Index (QI) [6] according to the formula:

$$QI = -(Sensitivity^N \cdot ln(1 - Sensitivity \cdot Specificity)) \qquad (1)$$

The QI expands toward infinity when both the Sensitivity and Specificity are 1, but it is slightly weighted in favour of Sensitivity [6]. The degree of weighting (bias) toward Sensitivity is controlled by the exponent N [1], which is equal to 0.75 for our experiments. The presented results concern the best neural networks (with the highest QI), and the resulted number of neurons in the hidden layer. The reported values are describing statistical parameters obtained for 50 learning trials.

At first, we analyzed the classification quality for both types of networks, using the basic input dataset – all traces recorded at different gestational age. Better results were obtained for RBF networks (Table 2). Figure 1 presents values of QI for all the 50 trials for these types of neural networks.

Table 2. Values of all prognostic indices obtained for the basic input dataset for both types of neural networks.

Index	MLP network with 6 hidden neurons		RBF network with 35 hidden neurons	
	Mean (SD)	Min÷Max	Mean (SD)	Min÷Max
Sensitivity	0.65 (0.07)	0.44÷0.81	0.71 (0.04)	0.63÷0.80
Specificity	0.67 (0.04)	0.58÷0.74	0.63 (0.02)	0.57÷0.69
PPV	0.44 (0.05)	0.31÷0.54	0.41 (0.02)	0.37÷0.45
NPV	0.82 (0.04)	0.73÷0.91	0.86 (0.02)	0.83÷0.90
QI	0.41 (0.09)	0.18÷0.63	0.46 (0.05)	0.37÷0.61

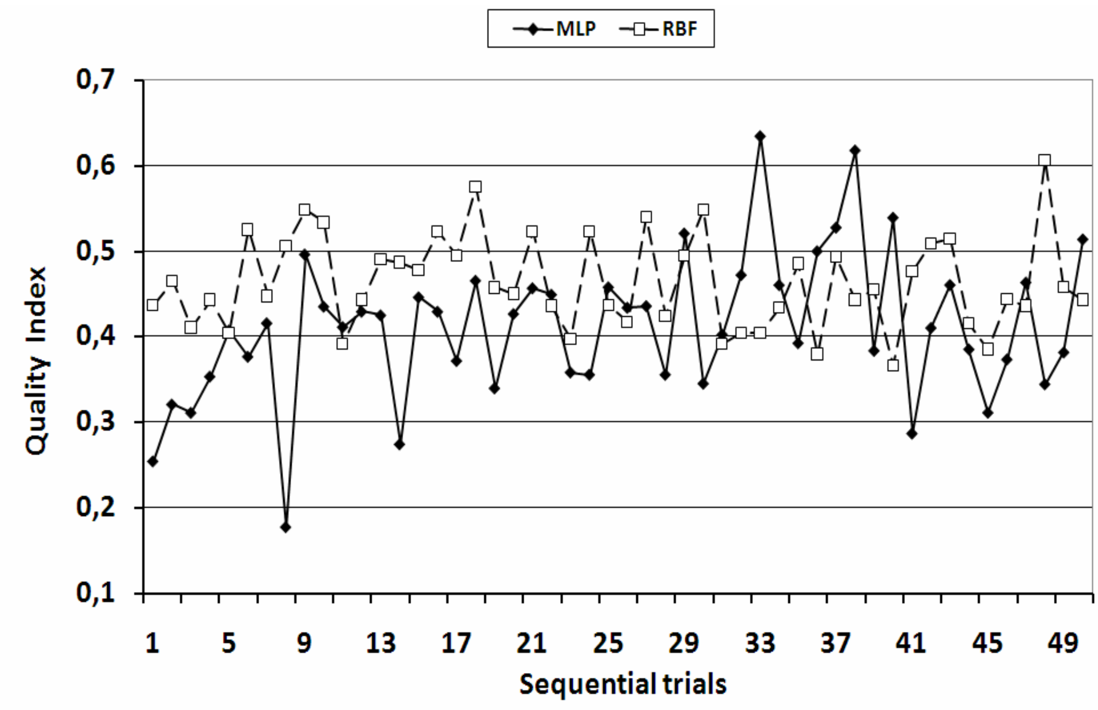

Fig. 1. The Quality Index values for all of the 50 trials, calculated for the neural networks from Table 2.

In the case of the trace grouping according to the gestational age for both of neural networks types, the values of QI decreased with the increase of the gestational age, however a single increase for labour traces was observed (Table 3). This single increase may be caused by the disproportion between the numbers of traces corresponding to normal and abnormal fetal outcomes which was smaller in labour group than in gestational groups. The same relation for some of other prognostic indices may be seen – Figures 2 and 3. In opposite to the basic dataset (Table 2), the MLP networks provided higher values of QI than the RBF networks. The most important is the fact, that grouping the traces within similar gestational age improves the classification quality only in the case of the MLP networks (higher values of QI for groups G1÷G3 in comparison to the basic dataset). For the RBF networks, the best results were obtained for the basic dataset.

Table 3. Values of Quality Index obtained for MLP and RBF neural networks
with trace grouping according to the gestational age.

Group	MLP Neural Network			RBF Neural Network		
	K	Mean (SD)	Min÷Max	K	Mean (SD)	Min÷Max
G1	6	0.59 (0.20)	0.25÷1.03	35	0.46 (0.13)	0.20÷0.85
G2	6	0.48 (0.18)	0.16÷1.07	40	0.42 (0.16)	0.11÷0.74
G3	6	0.48 (0.17)	0.15÷0.93	50	0.36 (0.14)	0.07÷0.83
G4	6	0.40 (0.14)	0.16÷0.74	50	0.27 (0.11)	0.06÷0.52
Labour	6	0.46 (0.15)	0.13÷0.84	50	0.30 (0.11)	0.15÷0.58

K – number of neurons in the hidden layer for the best neural network structure

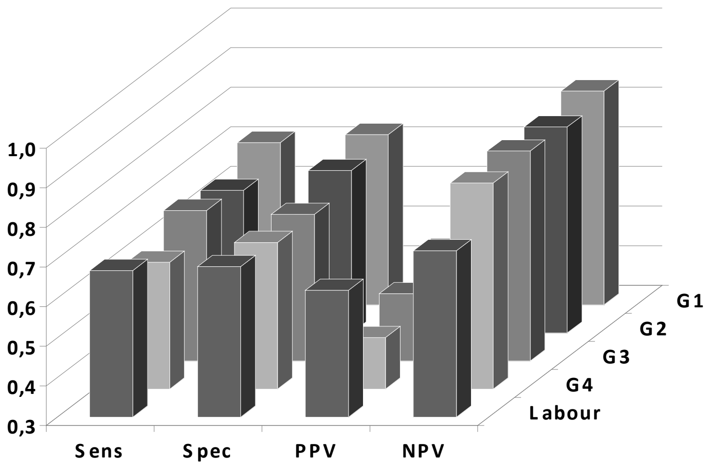

Fig.2. Values of prognostic indices for MLP neural networks using grouping according to the gestational age.

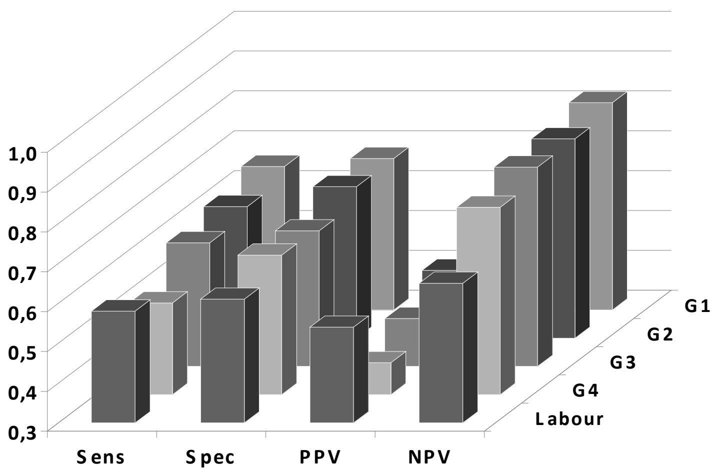

Fig.3. Values of prognostic indices for RBF networks using grouping according to the gestational age.

Adding the information about the gestational age through the additional neural networks input improved (in comparison to the basic dataset) the classification quality only in the case of MLP networks and when the additional input variable was describing the number of the range of the gestational age - Table 4. In other cases the results decreased.

Table 4. The influence of additional networks input describing the gestational age on the classification quality.

Additional Input	Neural Networks	Sensitivity	Specificity	PPV	NPV	QI
Number Representing Gestational Age Group	MLP 6#	0.66 (0.09)*	0.69 (0.06)	0.46 (0.08)	0.83 (0.05)	0.45 (0.13)
	RBF 40	0.57 (0.07)	0.62 (0.04)	0.38 (0.05)	0.78 (0.04)	0.28 (0.07)
Number of Completed Gestational Weeks	MLP 6	0.63 (0.08)	0.64 (0.06)	0.42 (0.08)	0.81 (0.05)	0.38 (0.11)
	RBF 50	0.56 (0.06)	0.62 (0.05)	0.37 (0.06)	0.78 (0.04)	0.29 (0.07)

#Number of neurons in the hidden layer of the best neural networks structure, *Mean (SD)

4. CONCLUSIONS

In the presented work we investigated the possibility of predicting the fetal outcome basing on parameters of quantitative description of cardiotocographic signals using both the MLP and RBF neural networks. We focused on the influence of the gestational age (when the trace was registered) on the prediction quality. At first, four groups of gestational traces and one group of traces registered on the day of labour were created. This input data decomposition improved classification quality only in the case of the MLP networks. In the next step, additional neural networks input was used, directly as a number of completed weeks of gestation or indirectly as a number of appropriate gestational age group. This results similarly to the first step, improved the classification quality of the MLP networks, when additional input was the number of gestational age group. In the case of the RBF networks, the best results were obtained using basic dataset, including all the traces registered at different gestational age. We didn't find any work, where the influence of gestational age on the classification quality was examined, so it is very difficult to compare our results. The authors of logarithmic index [6] focused on explaining its advantage over other criteria for stopping learning process. There are works, which describes effects of gestational age on the FHR variability [4, 5, 12], and our results confirm such relation.

In the experiments, the number of neurons in the hidden layer were changed to find the best structure. It is very important procedure, because different structures were appropriate for different data sets – MLP networks for the separate dataset with traces registered in similar gestational age, RBF networks for dataset with all traces. Each neural network was trained and tested fifty times, with random cases assignment to three subsets: training, validating and testing. Such approach increases the reliability and eliminates accidental coincidences. Because sometimes the classification quality did not change significantly in individual trials, comparison of mean values is more useful. This prevents the situation, when the network with a given structure provides very good results only by chance.

Our investigation showed the usefulness of both types of networks for fetal monitoring signals classification. However, additional improvements of this process are possible, particularly through a new input with information on the gestational age in the learning process and using the appropriate data set and networks.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Sciences and Higher Education resources in 2007-2009 under Research Project R1302802.

BIBLIOGRAPHY

[1] CATLEY C., FRIZE M., WALKER R., PETRIU D.C., Predicting high-risk preterm birth using artificial neural networks, IEEE Trans. Inf. Technol. Biomed., Vol. 10, No 3, 2006, pp. 540-549.

[2] CZABAŃSKI R., JEŻEWSKI M., WRÓBEL J., HOROBA K., JEŻEWSKI J.: A neuro-fuzzy approach to the classification of fetal cardiotocograms, IFMBE Proc., Vol. 20, 2008, pp. 446-449.

[3] CZOGAŁA E., ŁĘSKI J.: Fuzzy and Neuro-Fuzzy Intelligent Systems, Physica-Verlag, A Springer –Verlag Company, 2000.

[4] DAWES N. W., DAWES G. S., MOULDEN M., REDMAN CH. W. G.: Fetal heart rate patterns in term labour vary with sex, gestational age, epidural analgesia, and fetal weight, Am. J. Obstet. Gynaecol., 1999, Vol. 180, pp. 181-187.

[5] DRUZIN M. L., MILTON HUTSON J., EDERSHEIM T.G.: Relationship of baseline fetal heart rate to gestational age and fetal sex, Am. J. Obstet. Gynaecol., 1986, Vol. 154, pp. 1102-1103.

[6] ENNETT C.M., FRIZE M., SCALES N.: Evaluation of the logarithmic-sensitivity index as a neural network stopping criterion for rare outcomes, Proc. of 4th IEEE Conf. Inform. Technol. Applic. Biomed., 2003, pp. 338-341.

[7] HAYKIN S.: Neural networks: A guided tour, in: Nonlinear Biomedical Signal Processing, Vol. 1, Fuzzy Logic, Neural Networks and New Algorithms, Editor: Akay M.

[8] JEZEWSKI J., WROBEL J., HOROBA K.: Computerized perinatal database for retrospective qualitative assessment of CTG traces in „Current Perspectives in Healthcare Computing", Editor: B. Richards, BTHC, 1996, pp. 187-196.

[9] JEŻEWSKI M., WRÓBEL J., HOROBA K., GACEK A., HENZEL N., ŁĘSKI J.: The prediction of fetal outcome by applying neural network for evaluation of CTG records, in: Computer Recognition Systems, Editors: Kurzyński M., Puchała E., et. al., Advances in Soft Computing Series, Springer Verlag, 2007, pp. 532-541.

[10] JEZEWSKI J., WROBEL J., HOROBA K., KUPKA T., MATONIA A.: Centralised fetal monitoring system with hardware-based data flow control, Proc. of 3rd Inter. Conf. MEDSIP, Glasgow, VII 2006, pp. 51–54.

[11] JEŻEWSKI M., WRÓBEL J., ŁABAJ P., ŁĘSKI J., HENZEL N., HOROBA K., JEŻEWSKI J.: Some practical remarks on neural networks approach to fetal cardiotocograms classification, Proc. of 29th Int. Conf. IEEE EMBS, 2007, pp. 5170-5173.

[12] LANGE S., VAN LEEUWEN P., GEUE D., HATZMANN W., GRONEMEYER D.: Influence of gestational age, heart rate, gender and time of day on fetal heart rate variability, Med. Biol. Eng. Comput., Vol. 43, 2005, pp. 481-486.