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

fetal monitoring, cardiotocography, signal classification, neuro-fuzzy systems

Robert CZABAŃSKI*, Michał JEŻEWSKI*, Janusz WRÓBEL**, Tomasz KUPKA**, Jacek ŁĘSKI*, Janusz JEŻEWSKI**

THE PREDICTION OF THE LOW FETAL BIRTH WEIGHT BASED ON QUANTITATIVE DESCRIPTION OF CARDIOTOCOGRAPHIC SIGNALS

Cardiotocography (CTG) is a routine method of fetal condition assessment used in modern obstetrics. It is a biophysical method based on simultaneous recording and analysis of activity of fetal heart, fetal movements and maternal uterine contractions. The fetal condition is diagnosed on the basis of printed CTG trace evaluation. The correct interpretation of CTG traces from a bedside monitor is very difficult even for experienced clinicians. Therefore, computerized fetal monitoring systems are used to yield the quantitative description of the signal. However, the effective methods, aiming to support the conclusion generation, are still being searched. One of the most important features defining the state of fetal outcome is the weight of the newborn. The presented work describes an application of the Artificial Neural Network Based on Logical Interpretation of fuzzy if-then Rules (ANBLIR) to evaluate the risk of the low birth weight using a set of parameters quantitatively describing the CTG traces. The obtained results confirm that the neuro-fuzzy based CTG classification methods are very efficient for the prediction of the fetal outcome.

1. INTRODUCTION

Cardiotocography (CTG) is a primary biophysical method for monitoring of fetal condition before and during labour. It consists in analysis of the fetal heart rate (FHR) variability and their relationship to fetal movements and maternal uterine contractions. At a present, CTG become a standard clinical technique for identifying a wellbeing of the fetus. However, visual analysis of graphical patterns describing the FHR variability is relatively difficult. Therefore, computerized fetal monitoring systems are used to provide the quantitative description of the CTG signal. Nevertheless, the interpretation is still made by clinicians and remains highly subjective and dependent on the human expert capability and experience. The repeatable and objective assessment of the fetal state is of particular importance. Early detection of fetuses that during pregnancy are at significant risk may help to avoid dangerous situations which are more difficult or even impossible to manage in case of the newborn. Consequantly, effective methods, aiming to support the diagnosis, are still the topic of many research studies [3], [8]-[10].

The CTG monitoring evaluates the actual (at the time of monitoring session) fetal state. But the diagnosis verification is possible only after the delivery. There is no reference information about the fetal health condition during pregnancy. This information will be obtained only after the delivery, and the fetal outcome is retrospectively assigned to the fetal state. Such prediction of fetal outcome during pregnancy is possible, because in perinatology it is assumed that the fetal state can not change rapidly. One of the most important features defining the fetal outcome is the newborn's weight. Babies categorized as low birth weight are at particular risk of health problems, disability, or even death. In obstetrics, the newborns weight is evaluated as a percentile of reference values of birth mass in relation to the completed week of gestation. Any birth weight below the 10th percentile of the reference values is considered as too low .

The presented work describes an application of the Artificial Neural Network Based on Logical Interpretation of fuzzy if-then Rules (ANBLIR) [4] to evaluate the risk of low birth weight using parameters for quantitative description of CTG traces. The ANBLIR is a computationally effective fuzzy reasoning system which connects advantages of neural networks (capability of learning and generalization) and fuzzy systems (capacity to handle inherently imprecise concepts and ability of linguistic interpretation of learning results). It was successfully applied to solve many practical problems leading to significant increase of performance in comparison to the solutions using classical computational intelligence algorithms [4], [3]. To establish values of the neuro-fuzzy system parameters for birth weight assessment we used a preprocessed database obtained from an archive of computerized fetal surveillance system MONAKO [7]. We investigated three different learning algorithms of ANBLIR based on integration of steepest descent method, least squares algorithm and deterministic annealing learning as well as different training data set structures in aim to achieved the best CTG classification accuracy.

2. RESEARCH MATERIAL

The research database used in our experiments contains the results of quantitative analysis of CTG traces recorded from bedside monitors. It includes parameters describing the FHR signal in time domain as well as the number of recognized uterine contractions and fetal movements. The original, raw research material included 1274 CTG traces collected with computerized fetal surveillance system MONAKO [7] from 341 unselected patients of Obstetrical Department of the Silesian Medical

* Silesian University of Technology, Institute of Electronics, ul. Akademicka 16, 44-100 Gliwice, Poland.
** Institute of Medical Technology and Equipment, Biomedical Informatics Department, ul. Roosevelta 118, 41-800 Zabrze, Poland

University in Katowice. After removing all the incomplete data we obtained the database comprised of 685 records from 189 patient.

As input variables we used the set of 19 parameters describing the CTG signals in time domain: statistical parameters concerning the basal fetal heart rate (2), long-term FHR variability indices (7), beat-to-beat FHR variability indices (6), number of identified patterns decelerations, accelerations, uterine contractions and fetal movements (4). The output of ANBLIR system was defined as two-state variable defining the percentile of birth weight as being correct (+1) or incorrect (-1). The percentile value was determined basing on neonatal birth weight in relation to the reference values of birth weight derived form Polish national data charts. Any newborns' weight below the 10th percentile of reference values for a given week of pregnancy was regarded as incorrect. According to our assumptions, 68 (9.93%) records was classified as related to the low birth weight class.

3. NEURO-FUZZY SYSTEM

ANBLIR is a neuro-fuzzy system with parameterized consequents that generates inference results based on fuzzy if-then rules. Fuzzy sets of the rules premises have Gaussian membership function and their firing strength is defined as [4]:

$$\underset{i=1,2,\ldots,I}{\forall} F^{(i)} = \exp\left[-\frac{1}{2}\sum_{j=1}^{t}\left(\frac{x_{0j}-c_j^{(i)}}{s_j^{(i)}}\right)^2\right] \tag{1}$$

where I denotes the number of the rules, t is the number of inputs, x_{0j} is the j-th element of input vector $\mathbf{x}_0 = [x_{01}, x_{02}, \ldots, x_{0t}]^T$, $c_j^{(i)}$ is the centre of Gaussian membership function and $s_j^{(i)}$ is its dispersion.

Consequents fuzzy sets of ANBLIR system have symmetric triangular membership functions. They are defined by the width of the triangle base $w^{(i)}$ and the location of its centre of gravity $y^{(i)}(\mathbf{x}_0)$, which is determined by a linear combinations of inputs [4]:

$$y^{(i)}(\mathbf{x}_0) = p_0^{(i)} + p_1^{(i)}x_{01} + \ldots + p_t^{(i)}x_{0t} = \mathbf{p}^{(i)T}\mathbf{x}_{e0} \tag{2}$$

where $\mathbf{x}_{e0} = [1, \mathbf{x}_0]^T$ is the extended input vector.

The relationship between inputs and output of the fuzzy system is represented by mathematical relation defined on membership functions of fuzzy sets of premises and consequents. In the next considerations we assume the logical interpretation of fuzzy conditional statements using Lukasiewicz fuzzy implication. It satisfies all condition required for fuzzy implication given by Fodor [5]. In the ANBLIR inference procedure a normalized arithmetic mean as a aggregation operator and modified indexed centre of gravity as a defuzzifier are used. As a result we get the crisp output value of the system in the following form [4]:

$$y_0 = \sum_{i=1}^{I}\frac{w^{(i)}\left(1-\frac{1}{2}\left(F^{(i)}(\mathbf{x}_0)\right)^2\right)}{\sum_{j=1}^{I}w^{(j)}\left(1-\frac{1}{2}\left(F^{(i)}(\mathbf{x}_0)\right)^2\right)}y^{(i)}(\mathbf{x}_0) = \sum_{i=1}^{I}G^{(i)}(\mathbf{x}_0)y^{(i)}(\mathbf{x}_0) \tag{3}$$

One of the most important features of fuzzy system with parameterized consequents is its functional equivalence to the radial basis function neural network. Therefore, the unknown neuro-fuzzy system parameters (parameters of membership functions of fuzzy sets defining if-then rules) can be estimated using learning algorithms of artificial neural networks. Several solutions of the ANBLIR learning problem have been already introduced in literature [4], [1], [2]. In this work, we investigate learning procedures based on integration of steepest descent method, least squares algorithm and deterministic annealing procedure.

The learning set used in our numerical experiments is comprised of N=749 input vectors $\mathbf{x}_0(n) \in \Re^t$ with quantitative description of the CTG records and the same number of binary representation of the birth weight $t_0(n) \in \{-1,1\}$. The learning goal is the extraction of the set of fuzzy if-then rules that enables ANBLIR neuro-fuzzy system to classify CTG records. The learning process consists in the estimation of membership function parameters of antecedents as well as consequents $\underset{i=1,2,\ldots,I}{\forall}\ \underset{j=1,2,\ldots,t}{\forall} \alpha = \{c_j^{(i)}, s_j^{(i)}, p_j^{(i)}, w^{(i)}\}$. The number of rules I is pre-set arbitrarily. The original learning algorithm of ANBLIR uses a supervised training procedure that combines the steepest descent (SD) and the least squares (LS) methods [4]. In order to increase the learning efficiency, the deterministic annealing (DA) procedure [11] adapted for the purposes of learning the neuro-fuzzy system with parameterized consequents can be applied [1].

Deterministic annealing consists in the minimization of the squared-error cost E with simultaneous control of the Shannon entropy level S. The constrained deterministic annealing optimization be formulated as an unconstrained minimization of the Lagrangian [11]:

$$L = E - TS \tag{4}$$

where T is the Lagrange multiplier, which is very often called temperature what emphases the analogy between the DA algorithm and the process of the annealing of solids as the quantity L may be interpreted as the Helmholtz free energy of physical system with energy E, entropy S and temperature T. The DA method uses a simulated annealing procedure framework involving a series of minimization steps with gradual reduction of the entropy level. The algorithm starts at a high level of temperature $T_{\max}$ and tracks the solution for lowered values of T. The temperature parameter is reduced according to the so called annealing schedule function. In our experiments the temperature is decreased geometrically with a constant common ratio $q \in (0,1)$. At each level of temperature the Lagrangian is minimized using steepest descent method. The parameters of ANBLIR system are given as:

$$\alpha(k+1) = \alpha(k) - \frac{\eta_{ini}}{\sqrt{\sum_{i=1}^{n_i} \left(\frac{\partial L}{\partial \alpha_i} \right)^2 \Bigg|_{\alpha_i = \alpha_i(k)}}} \frac{\partial L}{\partial \alpha} \Bigg|_{\alpha = \alpha(k)} \tag{5}$$

where k denotes the iteration index, η_{ini} is the initial learning rate value and n_{i} represents the number of optimized parameters.

The DA learning method can be summarized as follows:

1. Set parameters of the algorithm: initial solution α, initial temperature $T_{\max}$, final temperature $T_{\min}$ and annealing schedule function. Set $T = T_{\max}$.
2. Minimize L at given level of temperature.
3. Check the entropy equilibrium $|\Delta S| = |(S(k-1) - S(k))/S(k-1)| > \delta$ or iterations stop condition $k \le k_{\max}$, where δ is pre-set parameter and $k_{\max}$ denotes the maximum number of iteration at given level of temperature, if one of them is fulfilled go to Step 2°.
4. Decrease temperature $T \leftarrow qT$.
5. If $T > T_{\min}$ go to Step 2°.
6. Stop the algorithm.

In the original ANBLIR learning method the parameters of consequents $\mathbf{p}^{(i)}$ were estimated using the least squares method. This solution speeds up the learning convergence [4]. Integration of LS algorithm with DA procedure or SD method leads to the ANBLIR learning method where parameters of antecedents and consequents of fuzzy rules are adjusted separately [1]. Premise parameters $c_j^{(i)}$, $s_j^{(i)}$ as well as triangle base widths $w^{(i)}$ of fuzzy sets in consequents are determined by means of steepest descent or deterministic annealing method, whereas parameters of linear equations from consequents $\mathbf{p}^{(i)}$ on the basis least squares procedure. For decreasing the computational time of the learning procedure the deterministic annealing method with *freezing* phase (DAF) can be also applied [2]. The freezing phase consists in the calculation of $\mathbf{p}^{(i)}$ using LS procedure after every decreasing step of temperature while keeping values of other parameters constant. This solution leads to decrease of the computational complexity of the DA based learning method. Another problem is the estimation of initial values of membership functions for antecedents. However, it can be solved by means of preliminary clustering of the training data [4].

4. EXPERIMENTS

To evaluate the risk of low birth weight we used three learning algorithms of the ANBLIR system: (i) the original training procedure which connects the steepest descent and the least squares methods (SDLS), (ii) a learning method that combines the deterministic annealing with the least squares algorithm (DALS) and (iii) its modified version, with the freezing phase (DAFLS). To asses the quality of the classification we evaluated the percent of correct classifications (ECL), Sensitivity (SE), Specificity (SP), Positive Predictive Value (PPV) and Negative Predictive Value (NPV). The evaluation of the learning performance on the basis of comparison of all prognostic indices values is rather difficult, therefore we applied the overall quality index $\mathrm{OI} = (2 \cdot \mathrm{SE} + \mathrm{NPV}) \cdot (\mathrm{SP} + \mathrm{PPV})/600$ [8]. The influence of the sensitivity index on the OI value is doubled. It is due to the fact that in clinical practice the false negative recognition of the newborns wellbeing may lead to the most serious consequences.

All numerical experiments were conducted in the MATLAB7.0® environment. The learning process was carried out for the number of if-then rules changed from 2 to 10. The initial values of parameters of antecedents were calculated on the basis of fuzzy c-means clustering results. For the deterministic annealing procedure the following parameters values were applied: $\eta_{ini} = 0.01$, $q = 0.95$ and $T_{\min} = 10^{-5} T_{\max}$. As there is no method for automated estimation of initial temperature T_{max} for the deterministic annealing algorithm, at the first stage of our experiments we changed its value in the range from 10^{+5} to 10^{-5} with the common ratio equal to 0.1. In every experiment, the cases were 50 times randomly assigned to three data sets: learning, validating and testing. The whole data set was partitioned in the learning, validating and testing subsets in proportion 50%, 25% and 25% respectively. The ratio of cases with normal birth weight to abnormal one in each set was constant in all trials.

At the first stage of the experiments performed we tried to find the most appropriate structure of learning data. To increase the learning quality we enlarged the number of the cases in the class of low birth from 10% to 30% with a step of 5. It is due to the difficulties with classifying skewed data. The cases with incorrect birth weight were randomly resampled until their representation in the training subset was sufficiently high. Table 1 presents the classification results expressed as mean values of OI index and the average percent of correct classification ECL. The prognostic indices were calculated for all 50 trials. For DA based algorithms they were averaged over the whole range of temperatures.

Table 1. Values of classification quality indices for different class structures of the training subset

Low birth weight [%]	Learning method					
	SDLS		DALS		DAFLS	
	OI	ECL	OI	ECL	OI	ECL
10	$39.6_{(4)}\pm8.5^{*}$	$89.9_{(4)}\pm1.9$	$39.6_{(4)}\pm 7.5$	$90.2_{(4)}\pm1.8$	$39.6_{(4)}\pm7.1$	$90.1_{(4)}\pm1.8$
15	$61.2_{(7)}\pm10.0$	$92.6_{(7)}\pm1.9$	$61.2_{(10)}\pm9.4$	$92.6_{(10)}\pm2.2$	$60.6_{(10)}\pm9.8$	$92.4_{(10)}\pm2.3$
20	$65.0_{(8)}\pm7.6$	$93.3_{(8)}\pm2.0$	$64.9_{(8)}\pm6.7$	$93.4_{(8)}\pm1.7$	$63.9_{(8)}\pm7.4$	$93.3_{(8)}\pm1.7$
25	$61.6_{(5)}\pm9.3$	$92.5_{(5)}\pm2.4$	$65.6_{(10)}\pm7.0$	$92.9_{(10)}\pm1.9$	$65.1_{(10)}\pm7.0$	$92.8_{(10)}\pm1.8$
30	$61.9_{(4)}\pm8.1$	$91.8_{(4)}\pm2.0$	$67.1_{(10)}\pm7.3$	$93.2_{(10)}\pm1.9$	$67.0_{(10)}\pm7.3$	$93.2_{(10)}\pm1.9$

*mean $_{\text{(no. of rules)}}$ ± SD [%]

The best prediction quality was obtained for DALS procedure when the percentage of cases characterized by the low birth weight was artificially increased to 30%. The detailed results of all prognostic indices used are shown in the Table 2.

Table 2. Summary statistics of the best learning results

	Learning method $_{(I,\%)}$		
Index	$\text{SDLS}_{(8,20\%)}$	$\text{DALS}_{(10,30\%)}$	$\text{DAFLS}_{(10,30\%)}$
ECL	$93.30 \pm 1.95^{*}$	93.17 ± 1.88	93.15 ± 1.88
Sensitivity	63.47 ± 9.40	60.67 ± 8.39	60.66 ± 8.42
Specificity	97.54 ± 1.22	98.65 ± 0.93	98.63 ± 0.96
PPV	78.21 ± 10.29	88.35 ± 8.28	88.17 ± 8.43
NPV	94.96 ± 1.93	93.69 ± 1.78	93.15 ± 1.88
OI	65.02 ± 7.56	67.11 ± 7.33	67.02 ± 7.34

*mean ± SD [%]

Table 3. Summary statistics of the best learning results obtained for single values of temperature

	Learning method $_{(I,\%)}$		
Index	$\text{SDLS}_{(8,20\%)}$	$\text{DALS}_{(8,20\%)}$	$\text{DAFLS}_{(8,20\%)}$
ECL	$93.30 \pm 1.95^{*}$	94.15 ± 1.71	94.32 ± 1.57
Sensitivity	63.47 ± 9.40	67.93 ± 8.78	69.46 ± 9.46
Specificity	97.54 ± 1.22	97.78 ± 1.28	97.79 ± 1.20
PPV	78.21 ± 10.29	80.30 ± 11.21	80.32 ± 9.92
NPV	94.96 ± 1.93	95.91 ± 1.61	96.34 ± 1.64
OI	65.02 ± 7.56	68.30 ± 7.21	68.75 ± 7.38

*mean ± SD [%]

The results of our study on finding the most appropriate training subset structure leading to an increase of CTG classification performance will be slightly different when considering the learning efficiency obtained with a single value of temperature, for which we get the best prediction accuracy. In this case, all ANBLIR learning algorithms showed the best performance for the percentage of cases characterized by low birth weight at the level of 20% and eight fuzzy conditional rules (Table 3).

The learning results reveal the crucial influence of the temperature value on the results of the prediction of the low fetal birth weight. The initial temperature should be high enough to ensure entropy maximization at the beginning of the optimization procedure, whereas the final temperature should assure the minimization of the square error at the end. The formula for the calculation of the annealing schedule parameter that guarantees finding a global minimum of the cost for simulated annealing procedure was given in [6]. However, this method leads to a significant increase of the learning time. Even for the considered annealing schedule the calculation time for the maximal number of rules is unacceptable in the real medical applications (Fig. 1).

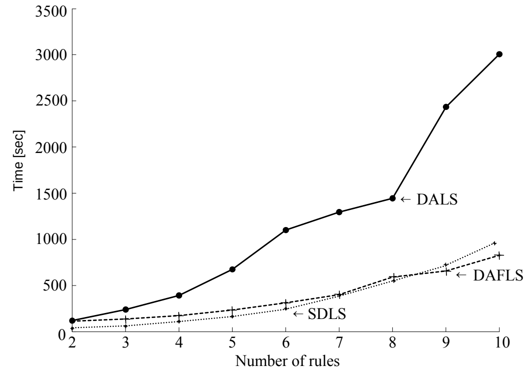

Fig. 1 The ANBLIR learning time as a function of the number of rules when using a PC equipped with Intel® Core™2 Duo processor 6400@2.13 GHz

It can be noticed that the DALS procedure is the slowest because the highest computational complexity. An application of the learning method with freezing phase enables to obtain the classification results in the similar period of time to the original ANBLIR learning procedure however, it provides better learning results. As the finding the optimal value of temperature using trial and error procedure is highly time consuming, we made an attempt to give some guidelines for the T parameter selection. The DA based algorithms lead to the acceptable results if the initial value of temperature is calculated on the basis of the inequality $T_{\max} \geq E_{ini} / S_{ini}$, where E_{ini} and S_{ini} are the values of square error and entropy respectively, calculated for the first iteration of ANBLIR learning.

Table 4. Summary statistics of the low birth weight prediction with constant temperature ($T_{\max}=10^{-2}$, $T_{\min}=10^{-7}$)

	Learning method $_{(r,\,\%)}$	
Index	$\text{DALS}_{(10,30\%)}$	$\text{DAFLS}_{(10,30\%)}$
ECL	93.16 ± 1.88*	93.13 ± 1.87
Sensitivity	60.60 ± 8.82	60.54 ± 8.33
Specificity	98.65 ± 0.95	98.63 ± 0.97
PPV	88.30 ± 8.38	88.19 ± 8.56
NPV	93.16 ± 1.88	93.13 ± 1.87
OI	67.04 ± 7.33	66.95 ± 7.32

*mean ± SD [%]

The value of temperature $T_{\min}$ should assure the $E_{end} >> S_{end} T_{\min}$ relation, where E_{end} and S_{end} are the values of square error and entropy respectively, evaluated for the last ANBLIR training iteration. The learning results obtained with the paradigm given are shown in the Table 4. They are only slightly worse than the low birth weight prediction results averaged over all temperatures within the considered range (see Table 2). Another way for improving the learning performance relies on application of zero-entropy iterations [1], however this solution leads also to the increase of the computational complexity and therefore the learning time.

5. CONCLUSIONS

In the presented work, we investigated the Artificial Neural Network Based on Logical Interpretation of fuzzy if-then Rules for prediction of the low fetal birth weight basing on the classification of cardiotocograms. We applied different learning algorithm, combining steepest descent method, least squares algorithm and deterministic annealing procedure. The selected learning procedure, integrates the deterministic annealing algorithm with freezing phase and least squares method was revealed as the most reasonable compromise between classification quality and the computational complexity.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Sciences and Higher Education resources in 2007-2010 under Research Project N518 014 32/0980.

BIBLIOGRAPHY

[1] CZABAŃSKI R., Neuro-fuzzy modeling based in a deterministic annealing approach. Int. J. Appl. Math. Comput. Sci. Vol.15, No.4, pp.561-576, 2005.
[2] CZABAŃSKI R., Extraction of fuzzy rules using deterministic annealing integrated with ε-insensitive learning. Int. J. Appl. Math. Comput. Sci. Vol.16, No.3, pp.35-372, 2006.
[3] CZABAŃSKI R., JEŻEWSKI M., WRÓBEL J., HOROBA K., JEŻEWSKI J., A Neuro-Fuzzy Approach to the Classification of Fetal Cardiotocograms, Proc. of 14th Int. Conf. NBC2008, Vol.20, pp.446-449, 2008.
[4] CZOGAŁA E., ŁĘSKI J., Fuzzy and Neuro-Fuzzy Intelligent Systems, Physica-Verlag, A Springer –Verlag Company, 2000.
[5] FODOR J.C., On fuzzy implication operators, Fuzzy Sets and Systems, Vol.42, No.3, pp.293-300, 1991.
[6] GERMAN S., GERMAN D., Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images. IEEE Trans. Pattern Anal. Mach. Intell. Vol.6, No.6, pp.721–741, 1984
[7] JEŻEWSKI J., WRÓBEL J., HOROBA K., KUPKA T., MATONIA A., Centralised fetal monitoring system with hardware-based data flow control, Proc. of III Int. Conf. MEDSIP, pp.51–54, Glasgow, 2006.
[8] 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 2, Eds.: Kurzyński M., et. al., ASC 45, Springer Verlag, pp.532-541, 2007.
[9] KOL S., THALER I., PAZ N., SHMUELI O., Interpretation of nonstress test by an artificial neural networks, American Journal of Obstetrics & Gynecology, Vol.172, No.5, pp.1372-1378, 1995.
[10] MAGENES G., SIGNORINI M.G., ARDUINI D., Classification of cardiotocographic records by neural networks, Proc.of the IEEE Int. Joint Conf. on Neural Networks, Vol.3, pp.637–641, 2000.
[11] ROSE K., Deterministic annealing for clustering, compression, classification, regression and related optimization problems. Proc. IEEE Vol.86, No.11, pp.2210–2239, 1998.