# Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques

Hakan Sahin*, Abdulhamit Subasi

*International Burch University, Faculty of Engineering and Information Technologies, Francuske Revolucije b.b., Ilidza, Sarajevo 71000, Bosnia and Herzegovina*

ABSTRACT

The aim of the research is evaluating the classification performances of eight different machine-learning methods on the antepartum cardiotocography (CTG) data. The classification is necessary to predict new-born health, especially for the critical cases. Cardiotocography is used for assisting the obstetricians' to obtain detailed information during the pregnancy as a technique of measuring fetal well-being, essentially in pregnant women having potential complications. The obstetricians describe CTG shortly as a continuous electronic record of the baby's heart rate took from the mother's abdomen. The acquired information is necessary to visualize unhealthiness of the embryo and gives an opportunity for early intervention prior to happening a permanent impairment to the embryo. The aim of the machine learning methods is by using attributes of data obtained from the uterine contraction (UC) and fetal heart rate (FHR) signals to classify as pathological or normal. The dataset contains 1831 instances with 21 attributes, examined by applying the methods. In the paper, the highest accuracy displayed as 99.2%.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

During pregnancy, it is not easy to acquire direct information about the fetus. Therefore, obstetricians should use indirect information related to a fetal state. Most important one is measuring the fetal heart rate (FHR) [1].

The term 'electronic fetal monitoring' is the other alternative name of CTG, however, it is conceivable as less specific appellation since CTG contains observation about the contractions of the mother and different kind of fetal monitoring too. One type of fetal assessments in pregnancy can be the antenatal CTG and it checks the fetal heart rate regarding as a biological indicator of baby health [2]. CTG is usually applied during the last trimester it means after the 28th week of pregnancy. The idea behind the usage of CTG is a monitoring test for the diagnosis of babies having either acute or chronic fetal hypoxia and getting information about the baby under the risk of developing hypoxia [2].

The accepted parameters for the fetus are stated as follows (they are all in beats per minute). Baseline fetal heart rate between 110 and 160. Baseline variance is larger than 5 [3].

Nowadays, cardiotocography is the most preferable oblique, investigative method, in practical usage, to observe fetal well-being. Cardiotocogram (CTG) comprises of two different signals, its uninterrupted recording of instant fetal heart rate (FHR) and uterine activity (UC). The information, which is obtained from CTG, is used for early identification of a pathological state (i.e. inherited heart deficiency, fetal suffering or hypoxia, etc.) and may assist the obstetrician to anticipate future complications and interfere before there is a permanent harm to the fetus. During the delivery the baby who is exposed to hypoxia may cause of death or temporary disablement. Due to improper diagnosis of the FHR pattern recordings and improper treatments applied to the fetus can cause more than half of these deaths [4].

While its practicality, there has been some discrepancy as to the utility and the success of CTG monitoring, particularly in low-risk pregnancies. If there is an incorrectly analyzed fetal pain, then, it may be referred to needless treatments or if there is an improper analysis of fetal welfare then it may be rejected required treatments [5].

The classification uses a performance evaluation measure, but it is not enough to decide for the vital case especially in medical diagnosis. Therefore, it is also suggested another type of performance evaluation tools such as the ROC (receiver operation characteristics) [6] and $F_1$-measure [7].

* Corresponding author. Tel.: +387 33 944 513; fax: +387 33 782 131.
*E-mail addresses:* hsahin@ibu.edu.ba (H. Sahin), asubasi@ibu.edu.ba (A. Subasi).

**Table 1**
Summary of all CTG features used in classification [11].

| Symbol | Attribute information |
| --- | --- |
| LB | FHR baseline (beats per minute) |
| AC | # of accelerations per second |
| FM | # of fetal movements per second |
| UC | # of uterine contractions per second |
| DL | # of light decelerations per second |
| DS | # of severe decelerations per second |
| DP | # of prolonged decelerations per second |
| ASTV | Percentage of time with abnormal short-term variability |
| MSTV | Mean value of short-term variability |
| ALTV | Percentage of time with abnormal long-term variability |
| MLTV | Mean value of long-term variability |
| Width | Width of FHR histogram |
| Min | Minimum of FHR histogram |
| Max | Maximum of FHR histogram |
| Nmax | # of histogram peaks |
| Nzeros | # of histogram zeros |
| Mode | Histogram mode |
| Mean | Histogram mean |
| Median | Histogram median |
| Variance | Histogram variance |
| Tendency | Histogram tendency |
| NSP | Fetal state class (code (N = normal; P = pathological)) |

In this study, it is used eight different machine-learning methods to examine the CTG data categorized as healthy or unhealthy according to the decisions of three obstetricians. This study presents a comparison of the performances of the machine learning methods using the open source software WEKA [8] in terms of accuracy, specificity, sensitivity, *F*-measure and ROC curve.

## 2. Materials and methods

### 2.1. Dataset

The analysis shows the performance of the methods over the datasets from UCI [9] including CTG data with some indicative features. Three expert obstetricians decided to condition of the CTG data, whether normal or pathological checking the status of the embryo. The UCI cardiotocography data [9] was obtained by the automatic SISPORTO 2.0 [10] software. It is isolated from the suspicious entries and normal and pathologic class added to the NP feature. The Table 1 gives an explanation for each property of the respective features in the data. The CTG data has 21 features, 8 of them are continuous and 13 are discrete. The classification of the data with respect to condition of the fetus either normal or pathological. After giving a short explanation for each of the algorithms which are ANN, SVM, SL, RBF, C4.5, CART and RF decision trees. WEKA classification algorithms available at [11] were used for the algorithms. Each algorithm performance was tested using 10-fold cross validation during the examination for CTG dataset that tested around 1831 entries accurately. The CTG data entered in the WEKA classification module and trained several times by varying the respective parameters of the each algorithm to maximize the classification performance. The performances of the algorithms recorded and then tabulated.

### 2.2. Logistic regressions

Logistic regression analyzes the effect of several dynamics in a two-grouped outcome via approximating the probability of the experimental event [12]. The outcome of logistic regression is a binary event, like spam versus not spam, or malignant versus benign. The mathematical background of logistic regression is the concept logit that is described as "the natural logarithm of the odds ratio". In the literature, you can come up against the researchers who were calling the logistic regression as the logit model or logistic model formulated as follows,

$$log \frac{p(x)}{1 - p(x)} = \beta_0 + x \cdot \beta \qquad (1)$$

Why? Because, to achieve a goal, such as predicting the dichotomous outcome, instead of using a complicated equation, it opts linear one like its cousin linear regression. Solving for p, gives us a new formula depends on the variable *x*

$$p(x; w, n) = \frac{e^{\beta_0 + x \cdot \beta}}{1 + e^{\beta_0 + x \cdot \beta}} = \frac{1}{1 + e^{-(\beta_0 + x \cdot \beta)}} \qquad (2)$$

That equation includes all necessary coefficients, which are accepted as the inputs of the model. In the literature, it is encountered with two different types of usage of logistic model: Of course, to categorize the data according to its classes by calculating the ratio of probabilities success and failure in the form of the odds ratio. Secondly, to find out the relationships between the variables and how effects each other [13].

### 2.3. k-Nearest neighbor

The *k*-nearest-neighbors are an example of a classification method with the independence of the parameters. The method can be classified as by the point of implication as simple, but efficient in many datasets [14].

The *k*-NN algorithm is defined by three terms (*S*, *k*, *T*) where *S* represents a resemblance measure which links to each pair of data in an appropriate *N*-dimensional space at (real or integer) number, *k* represents the number of nearest data that are trained to carry out the classification and *T* represents the vector of *M* training data applied by the classifier to actually carry out the classification [15].

*k*-NN uses distance metrics usually Euclidean distance to perform the similarity measure formulated by

$$d_E = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \qquad (3)$$

If *t* is a data sample, which classification implemented and its *k* nearest neighbors turn up, then this makes a *neighborhood of t*. While using the method to classify the data sample in the neighborhood about *t*, the distance is whether considered or not. Of course, to use *k*-NN choosing the right value for *k* is critical; due to the rate of success of classification directly appertain to this value. The *k*-NN method can pass for biased by *k*. It is possible to choose the *k*-value in so many different ways. Of course, the smooth and efficient one is applying the algorithm several times changing the value of *k* to get the highest accuracy. To decrease the dependency level of choice of *k* for *k*-NN method, Wang [16] recommended checking multiple sets of nearest neighbors instead of one set of nearest neighbors.

*k*-NN can be applied by using the following procedure,

| | |
| --- | --- |
| 1. | Keep all the outputs of the *P* nearest neighbors to examine a part *r* in the vector set $m = \left\{ m^1, \ldots, m^p \right\}$ by applying the steps *P* times: |
| a. | Take the next part $r^i$ in the data, where *i* represents the current iteration in the given set {1, ..., Q} |
| b. | If *r* is not assigned or $r < d(r, m^i) : r \leftarrow d(r, m^i), t \leftarrow s^i$ |
| c. | Repeat until the last element of the data (i.e. *i* = *Q*) |
| d. | Keep *r* as the vector *c* and *t* as the vector *m* |
| 2. | Compute the average output for the vector *m* by the formula as follows: $\bar{m} = \frac{1}{P} \sum_{i=1}^{P} m_i$ |
| 3. | Assign $\bar{m}$ as the output value of the examination data part *r*. |