Computer Methods and Programs in Biomedicine 163 (2018) 87–100

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Fetal health status prediction based on maternal clinical history using machine learning techniques

Akhan Akbulut[a,b,*], Egemen Ertugrul[b], Varol Topcu[b]

[a] Department of Computer Science, North Carolina State University, Raleigh, NC 27606, USA
[b] Department of Computer Engineering, Istanbul Kultur University, Atakoy Campus Bakirkoy, Istanbul 34156, Turkey

ARTICLE INFO

Article history:
Received 18 February 2018
Revised 19 May 2018
Accepted 8 June 2018

Keywords:
Machine learning
Medical diagnosis
Risk prediction
Pregnancy
Fetal health
Prognosis
m-Health

ABSTRACT

Background and Objective: Congenital anomalies are seen at 1–3% of the population, probabilities of which are tried to be found out primarily through double, triple and quad tests during pregnancy. Also, ultrasonographical evaluations of fetuses enhance detecting and defining these abnormalities. About 60–70% of the anomalies can be diagnosed via ultrasonography, while the remaining 30–40% can be diagnosed after childbirth. Medical diagnosis and prediction is a topic that is closely related with e-Health and machine learning. e-Health applications are critically important especially for the patients unable to see a doctor or any health professional. Our objective is to help clinicians and families to better predict fetal congenital anomalies besides the traditional pregnancy tests using machine learning techniques and e-Health applications.

Methods: In this work, we developed a prediction system with assistive e-Health applications which both the pregnant women and practitioners can make use of. A performance comparison (considering Accuracy, F1-Score, AUC measures) was made between 9 binary classification models (Averaged Perceptron, Boosted Decision Tree, Bayes Point Machine, Decision Forest, Decision Jungle, Locally-Deep Support Vector Machine, Logistic Regression, Neural Network, Support Vector Machine) which were trained with the clinical dataset of 96 pregnant women and used to process data to predict fetal anomaly status based on the maternal and clinical data. The dataset was obtained through maternal questionnaire and detailed evaluations of 3 clinicians from RadyoEmar radiodiagnostics center in Istanbul, Turkey. Our e-Health applications are used to get pregnant women's health status and clinical history parameters as inputs, recommend them physical activities to perform during pregnancy, and inform the practitioners and finally the patients about possible risks of fetal anomalies as the output.

Results: In this paper, the highest accuracy of prediction was displayed as 89.5% during the development tests with Decision Forest model. In real life testing with 16 users, the performance was 87.5%. This estimate is sufficient to give an idea of fetal health before the patient visits the physician.

Conclusions: The proposed work aims to provide assistive services to pregnant women and clinicians via an online system consisting of a mobile side for the patients, a web application side for their clinicians and a prediction system. In addition, we showed the impact of certain clinical data parameters of pregnant on the fetal health status, statistically correlated the parameters with the existence of fetal anomalies and showed guidelines for future researches.

1. Introduction

The stage of gestation or pregnancy of a woman is a phase that may carry complications for both the mother and the fetus. The fetal health may be affected by the maternal adaptive changes during this period, as well as by the medical history of the mother and maternal/familial attributes. Thus, the observation of fetal health and antenatal care during this phase is crucially important for maternal and fetal health [1].

Clinicians (especially gynaecologists) wish to inform the parents about well-being of their unborn infants and they do, based on already studied cases and past experiences. Hence, defining the pos-

* Corresponding author at: Department of Computer Science, North Carolina State University, Raleigh, NC 27606, USA.
E-mail addresses: aakbulu@ncsu.edu (A. Akbulut), 1401020030@stu.iku.edu.tr (E. Ertugrul), 1600006916@stu.iku.edu.tr (V. Topcu).

https://doi.org/10.1016/j.cmpb.2018.06.010
0169-2607/

sible familial (predominantly maternal) clinical data which is prone to influence fetal health would help antenatal maternal care.

However, patient-oriented health care might not be fully carried out due to some difficulties, i.e. insufficient number of hospitals near the patient's whereabouts (villages outlying in the rural areas) or overcrowded hospitals in cities. In Turkey, m-Health programs that provide access to health services and health information through mobile services are established on a national level [2], where mobile-cellular subscriptions cover 96.02% and Internet users cover 53.74% of the population of Turkey [3]. These establishments and mobile services show the supportive and critical role m-Health plays in inadequate conditions.

Our main objective is to offer both clinicians and pregnant women an assistive technology where they can remotely predict the status of the fetal health with the help of mobile services and ICT. The prediction is achieved by a machine learning system that learns a dataset of pregnant patients including their clinical history and data, and outputs a result considering fetal health as the main criteria, based on the trained dataset. This way, the output of the system helps both sides to take precautions and gives them awareness before the birth is given. The proposed machine learning system uses decision tree learning method. This method will be explained in detail in the methodology section.

Day by day, telehealth [4] is becoming a crucial part of the health care system. In this context, we aim to design and develop a telehealth or m-Health system -since it supports mobile usage- to prognosis the fetal health of pregnant. We offer an alternative and supportive system for remote patients to obtain clinical services.

The contributions of this paper are two-fold:

1. A user-friendly and assistive web and mobile application with a machine learning system employed to predict fetal health status.
2. A dataset, obtained from the pregnant women who were accepted at a radiodiagnostics center (RadyoEmar, Bakirkoy, Istanbul), is used to train the prediction (machine learning) system in order to predict the fetal health, considering the clinical data and history.

In the remainder of this paper, previous works done on this topic are mentioned, methodology and technical details about the proposed model are given, feature importances and weights in the dataset are analyzed, test results of a group of volunteer patients are examined and finally the paper is concluded.

2. Background

The wide popularity of machine learning is increasing in almost every domain. Those methods allow researchers to solve diagnostic and prognostic problems in a variety of medical domains [5]. Medical signal analysis, long-term patient tracking and analysis, medical image and sound processing, drug discovery, assistive robotic surgery, and automatic treatment systems or recommendation services for patients are emerging trends of new generation healthcare applications.

Probably the most promising area of the machine learning usage in the healthcare domain is medical diagnosis. Development of decision support systems for analyzing the new cases using trained data sets and dealing with inappropriate medical data such as having noise or missing data are the most-known implementations. Although our proposed system is not exactly a decision support system, it is considered to be used for pre-diagnostic purposes for the end user.

Rest of this section covers some of the most prominent research projects which were conducted on predicting pregnancy risks and fetal health using machine learning techniques. As an important indicator of fetal health, low birth weight (LBW) risk in pregnancies was predicted via machine learning methods by Yarlapati et al. [6]. With using Bayes minimum error rate classifier on Indian health care data, they managed to predict the fetal status as LBW or NOT-LBW with the accuracy of 96.77%. In other research [7], ultrasound images obtained from lumbar spine of pregnant patients are used to determine proper needle insertion site. Identification of interspinous region is realized using a Support Vector Machine (SVM) with Gaussian kernel. After training this model with 800 images from 20 pregnant subjects, it was tested with 640 images from a separate set of 16 pregnant patients and the proposed model achieved a success rate of 95.00%.

Cheng et al. [8] inspected the effects of risk factors such as age and hematocrit over the gestational age of pregnant. The primary contribution of this study is that the known risk factors are affecting differently on gestational hypertension group and pre-eclampsia group from 412 pregnant women including 1874 clinical follow-up records. Predicting the preterm birth risk was studied by Woolery and Grzymala-Busse [9]. They implemented an expert system to be used in determining the preterm birth risk with using 18,890 subjects and 214 variables. Their proposed system was 53–88% accurate in predicting preterm delivery for 9419 patients. In the study of predicting fetal well-being [10], antepartum cardiotocography (CTG) data was used with 8 different machine learning approaches such as ANN, SVM, k-NN, RF, CART, Logistic Regression, C4.5, and RBFN. Baby's heart rate measured from the mothers abdomen was utilized to extract uterine contraction (UC) and fetal heart rate (FHR) to be used in classifiers. The highest accuracy was achieved by random forest with 99.2%.

The previous researches done on machine learning in medical diagnosis and prediction include C4.5 decision tree classification algorithm and Naive Bayes Kernel algorithm, while the applications of these algorithms can be about predicting and presenting gestational risks [11,12] and normal or abnormal stages of pregnancy [13].

Kenny et al. [14] aimed to make a prognosis for pre-eclampsia disorder for pregnant. They used genetic programming to confirm the patterns of metabolites that distinguish plasma from patients with pre-eclampsia using 97 plasma samples. Smyser et al. [15] conducted a study on predicting brain maturity and neurodevelopmental outcome in infants using multivariate pattern analysis on neonatal functional MRI data. SVM estimated the birth gestational age of individual infants with 84% accuracy (p b 0.0001).

Czabanski et al. [16] focused on cardiotocography (CTG) for biophysical assessment of fetal state. As the first step, the classification of fetal heart rate (FHR) signals with a fuzzy method is carried out. Then in the second phase, using Lagrangian Support Vector Machines (LSVM) a success rate (The highest CC = 92.0% and QI = 88.2%) is achieved with the proposed method. The study in which the fetal brain development is modeled [17], whole-brain functional connectivity was inspected on 105 preterm infants. Connections were estimated with 80% accuracy using the SVM method. In this regard, results were generated for areas that could not be fully interpreted by magnetic resonance image (MRI) analysis. Krupa et al. [18] used empirical mode decomposition and SVM in CTG analysis of newborns. Prediction of normal or at risk classes on 90 randomly selected records was achieved with 86% accuracy.

Ocak [19] performed a CTG analysis with a UCI dataset [20] consists of fetal heart rate and uterine contractions and proposed a fetal heath estimation model with SVM method. The classification performance of the SVM is enhanced by enabling the genetic algorithm to eliminate the irrelevant features from the dataset. His proposed method achieved 99.3% and 100% accuracy in predicting fetal state as normal or pathological, respectively. Research on a different UCI dataset with 2126 fetal CTGs recordings perform a full prediction of pathologic cases using modular neural