JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

FETAL HEALTH CLASSIFICATION USING AI & ML

N. GOWTHAM ABHISHEK- Student, Department of CSE (Data Science), Geethanjali College of Engineering and Technology

V. A. TANMAI- Student, Department of CSE (Data Science), Geethanjali College of Engineering and Technology

KONDA KARTHIK SRI DATTA- Student, Department of CSE (Data Science), Geethanjali College of Engineering and Technology

Mr. B. VENKATESHWARLU - Assistant Professor, Department of Computer Science, Geethanjali College of Engineering and Technology

Abstract

The United Nations estimates around 140 million childbirths every year and around 2.4 - 3 million deaths. However, this number can seem very small in comparison; all these deaths could have been avoided with prior information and proper care during gestation. An unborn baby is called a fetus; most humans' gestation lasts around nine months. The mother and the child's health have to be monitored regularly to keep track of various characteristics that affect the fetus's health and development. Using the traditional methods, we have specialized procedures to determine if the fetus is normal or not, which requires medical personnel to be trained extensively, which is highly expensive as we have to train the professional, which is time-consuming. The development of Machine Learning Models based on features extracted from CTG is an alternative approach to the conventional methodology, as domain expertise is required for the conventional method. The aim is to develop a machine learning model that is efficient and accurate at predicting the fetal health state based on the Cardiotocogram (CTG) data. The models will be trained on the data that has been carefully classified according to the fetal health state by the doctors. After the model has been trained, we can make predictions on the new data using these models.

1. INTRODUCTION

The world has made significant progress in reducing child and maternal deaths in the past few decades. The child and maternal mortality rates have almost halved in the past 30 years. However, they are still dying in unacceptably large numbers and most of these deaths could have been preventable. Cardiotocography plays a significant role in understanding the health of unborn babies and how it impacts the mother's health. Cardiotocography (CTG) is defined as the graphic recording of fetal heart rate and uterine contractions by using electronic devices indicated for assessing the fetal condition (Fig. 1.1). A cardiotocograph records the fetal heart rate (cardio-tach) and alterations in the uterine muscle tone curve (toco-dynamo) simultaneously. If a pregnancy is considered high-risk, a precise diagnosis with respect to fetal hypoxia or acidosis can be established by using CTG. Fetal hypoxia (FH), or intrauterine hypoxia (IH), occurs when the fetus is deprived of an adequate oxygen supply. Hypoxia in pregnancy is rare; however, the anatomical and physiological changes associated with pregnancy may exacerbate hypoxia and hypoxemia arising from pathological processes. Similarly, Maternal acidosis, due to acute starvation, is a medical emergency in which both mother and child are at risk for significant morbidity and mortality. This issue is especially prominent in low-income societies. Acute starvation in the third trimester of pregnancy causing maternal ketoacidosis should be identified rapidly, followed by the proper treatment immediately. The fetal Heart Rate can be recorded by applying a Doppler Transducer to the maternal abdominal wall at the fetal cardiac projection. Using a Doppler transducer during pregnancy or labor is convenient. Still, some disadvantages should be mentioned: recorded changes in fetal Heart Rate are not precise and, in some cases, (especially when the fetus is dead), the sensor may capture not the fetal Heart Rate but the maternal abdominal aortic pulsation instead. For those reasons mentioned above, direct fetal electrocardiography characterized by the application of an electrode into the fetal scalp skin may be required during labor. The placement of fetal scalp electrode (FSE) is rarely indicated unless conventional CTG recording is difficult to perform or the precision of CTG results is suspect. The fetal scalp electrode can only be placed once fetal membranes are ruptured and cervix dilation is at least 2-3 cm. Alterations in the uterine muscle tone are recorded by a sensitive pressure transducer, which is placed at the uterine fundus. It also records various fetal movements simultaneously. The regular CTG paper speed is 1 cm/min. CTG is one of the most essential documents of labor supervision; thus, evaluation of CTG changes and interpretation must be present in every labor history–in particular, where adjustments in birth care tactics based on CTG data are made. The CTG Data can be studied and made sense of by an obstetrician - a physician specializing in delivering babies and caring for people during pregnancy and after birth. The data we have chosen for our Project is from the University of California Irvine Machine Learning Repository. Over 2000+ instances of such fetal case studies have been performed, and the most prominent features from the CTG scan have been recorded. The requirement of such highly trained professionals is a scarce resource in low-income and underdeveloped countries; using machine learning in such medical cases where we can predict the outcome from such reports is a significant necessity and has been at the forefront of medical development in the recent years, as such technology removes the barrier for crucial and quality healthcare.

1.2 Scope of the Project:

The project aims to develop a machine learning model using Cardiotocogram (CTG) data to predict fetal health conditions, particularly focusing on detecting fetal hypoxia or acidosis. By leveraging over 2000 instances of fetal case studies from the University of California Irvine Machine Learning Repository, the model will be trained to analyze CTG features and accurately classify fetal health states. This approach addresses the limitations of traditional methods, which require extensive training of medical personnel, and aims to improve access to quality healthcare, especially in low-income and underdeveloped regions, where the scarcity of highly trained professionals poses significant challenges to maternal and fetal health.

2. LITERATURE SURVEY

Fetal health classification is an important task in obstetrics and gynecology. Several studies have been conducted to develop machine learning algorithms that can accurately classify fetal health using cardiotocography (CTG) signals. University of California, Irvine, has created a dataset comprising of 22 different cardiotocogram parameters [1]. Rosalie M Grivell researched the effectiveness of antenatal cardiotocography, both traditional and computerized assessments, in improving outcomes and for mothers and babies during and after pregnancy. For this, the team has gone through the Cochrane Pregnancy and Childbirth Group's Trials Register (26 June 2015) and reference lists of retrieved studies, considering the selection criteria of randomized and quasi-randomized trials that compared traditional antenatal CTG with no CTG or CTG results concealed; computerized CTG with no CTG or CTG results concealed; and computerized CTG with traditional CTG [2]. V. Gintautas, G. Ramonienė, D. Simanavičiūtė discussed the concept of cardiotocography (CTG), its recording method, indications of performances during pregnancy and labor as well as assessment of results [3]. In another study, Vaclav Chudasek researched the various methods of automatic signal processing and evaluation of CTG data and its minimal progress in the past 20 years. They emphasized on the lack of common grounds for clinicians and technicians and stated that their open-access database aims to change that [4]. Jun Ogasawara, Satoru Ikenoue, and Hiroko Yamamoto created a newly-constructed deep neural network model (CTG-net) to detect compromised fetal status, classified into normal or abnormal cases. They evaluated the performance of the CTG-net with the F1 score and compared it with conventional algorithms, namely, support vector machine and k-means clustering, and another deep neural network model, long short-term memory [5]. Thomas G. Day researched the potential uses of Artificial Intelligence (AI) in the field of fetal cardiology including the provision of more accurate prognoses for individuals, and the automatic quantification of various metrics including cardiac function [6]. Edoardo Spairani created a hybrid approach based on a neural architecture that receives heterogeneous data in input, a set of quantitative parameters and images, for classifying healthy and pathological fetuses.

The neural architecture was trained on a huge and balanced set of clinical data (14,000 CTG tracings, 7000 healthy, and 7000 pathological) recorded during ambulatory non-stress tests at the University Hospital Federico II, Napoli, Italy. After hyperparameters tuning and training, the neural network proposed has reached an overall accuracy of 80.1%, which is a promising result, as it has been obtained on a huge dataset [7]. Shomona Gracia Jacob and R. Geetha Ramani developed a classification model for cardiotocography data using data mining methods and techniques. This research work emphasizes three major phases of Data mining viz, Clustering, Outlier Detection, and Classification [8]. V. Subha, Dr. D. Murugan, Jency Rani, and Dr. K. Rajalakshmi have done a comparative analysis of classification techniques in AI/ML. By using CTG data, they compared the classification methods, Naive Bayes, Decision Tree, Multi-Layer Perceptron, and Radial Basis Function. The metrics (Accuracy, Sensitivity, Specificity, F-Score) are used to evaluate them. In order to improve the accuracy, feature selection methods were also applied and the performance of classifiers was compared [9]. Masahiro Akiba compared machine-learning classification models like neural network (NN), naïve Bayes (NB), support vector machine (SVM), and gradient-boosted decision trees (GBDT) for glaucoma management. A comparison of the performance of the three machine-learning classification models showed that the NN had the best classification performance with a validated accuracy of 87.8% using only nine ocular parameters [10]. Sai Prasad Potharaju and M. Sreedevi suggested the data mining approach to increase the classification accuracy of cardiotocography by applying preprocessing techniques. They suggested a technique to handle imbalanced data to increase the classification performance of various lazy learners, rule-based induction models, and tree-based models. They used Synthetic Minority Over Sampling Technique (SMOTE) on a real dataset to accelerate the performance of various classifiers [11]. Sharma and Gupta comprehensively reviewed deep-learning techniques for fetal health classification. They evaluated various deep learning models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention-based models. The study found that CNN-based models perform better than RNN-based models for fetal health classification [12]. Another study by Heredia-Jimenez et al. proposed a machine learning-based approach for fetal health classification using a small dataset. The study compared several machines learning algorithms, including logistic regression, support vector machines, random forests, and neural networks. The study found that neural networks perform better than other algorithms for fetal health classification [13]. Rueda-Ruiz et al. (2020) proposed a fetal health classification method using CNNs. The study used ultrasound images to classify fetal health into three categories: normal, suspicious, and pathological. The study achieved an accuracy of 93% for fetal health classification [14]. Ayres et al. developed a machine learning-based approach for fetal health classification using CTG signals. The study compared several machine learning algorithms, including decision trees, random forests, and gradient boosting [15]. Finally, Patil and Sonavane proposed a wavelet-based approach for fetal health classification using a random forest classifier. The study used CTG signals to classify fetal health into three categories: normal, suspicious, and pathological. The study achieved an accuracy of 95% for fetal health classification [16].

3. OVERVIEW OF THE SYSTEM

3.1 Existing System

The existing system for fetal health classification primarily utilizes traditional methods, particularly cardiotocography (CTG). CTG involves the manual interpretation of fetal heart rate and uterine contractions by obstetricians to assess fetal well-being during pregnancy and labor. This process has been a standard practice in obstetrics for decades and remains essential for monitoring fetal health. However, its reliance on human interpretation can lead to variability in diagnoses due to differences in expertise among healthcare providers. Despite these challenges, the existing system serves as a foundational tool in fetal health assessment.

3.1.1 Disadvantages of Existing System

Subjectivity: Manual interpretation of CTG by obstetricians can lead to variability in diagnoses due to differences in expertise among healthcare providers.
Inconsistencies: Variability in interpretation can result in inconsistent fetal health assessments, potentially impacting patient outcomes.
Labor-Intensive: The reliance on human interpretation makes the process labor-intensive and time-consuming, requiring significant effort from healthcare providers.

3.2 Proposed System

The proposed system aims to revolutionize fetal health assessment by integrating machine learning techniques with cardiotocography (CTG) data analysis. Through the development of advanced algorithms trained on CTG data, the system seeks to automate and enhance the accuracy of fetal health classification. By analyzing patterns in fetal heart rate and uterine contractions, the system can predict conditions such as hypoxia or acidosis with greater efficiency and consistency compared to traditional methods. This innovative approach has the potential to significantly improve prenatal care by providing more timely and reliable assessments of fetal well-being.

3.2.1 Advantages of Proposed System

Automation: Integration of machine learning techniques automates the analysis of CTG data, reducing dependency on individual expertise and minimizing variability in diagnoses.

Enhanced Accuracy: Advanced algorithms trained on CTG data can analyze patterns in fetal heart rate and uterine contractions more efficiently and consistently than traditional methods, leading to more accurate assessments of fetal well-being.

Timely Assessments: The proposed system has the potential to provide more timely assessments of fetal health, enabling prompt interventions and improving maternal and fetal outcomes.

3.3 Proposed System Design

In this project work, there are SIX modules and each module has specific functions, they are:

1. Data Collection
2. Data Preparation
3. Data Analysis
4. Model Selection
5. Model Training
6. Model Evaluation

3.3.1 Data Collection

This module acts as a bridge between user and classifier modules. It performs loading of dataset and then does the preprocessing of the data. It requires training and testing data in order to train the classifier. This module also provides higher security by authenticating the user credentials. It then generates results according to the classification and suggests the necessary medication or action to be taken.

3.3.2 Data Preparation

Data cleanup is an essential step in machine learning that involves identifying and correcting or removing errors, inconsistencies, and inaccuracies in the dataset to improve the accuracy and reliability of the machine learning model.

3.3.3 Data Analysis

The fetal health dataset on which the model is trained has highly imbalanced data; due to the imbalance in the data, the model will be trained with a bias towards the data that has a majority in the number of occurrences; to eliminate this, we have to balance the dataset, this can be done in two ways by manipulation of the dataset which is

1. Undersampling
2. Oversampling

the latter option has been chosen, i.e., oversampling because by undersampling the dataset to balance the minority data, many data points would be lost, which affects the model training process as fewer data is available to work with; this can result in an inferior model.

3.3.4 Model Selection

There are many types of machine learning models, and selection of the model needs to be done after performance evaluation of different models on the train and the test data to come to a conclusion; in this process, we are using machine learning algorithms and conducting a performance analysis to find the best suitable model.

1. Linear Regression
2. K Nearest Neighbours
3. Decision Tree Classifier
4. Gradient Boosting Classifier
5. Support Vector Classifier
6. Random Forest Classifier

3.3.5 Model Training

Model training is the most crucial step that affects the final outcome of the predicted values. To begin the model training, the existing dataset is distributed into a 7:3 ratio for train and test data; we then oversample the training data such that the minority classes are oversampled, and the data imbalance problem is resolved, after which have trained the models on the Machine Learning Algorithms mentioned in the previous section to perform the model evaluation. After comparing the accuracy values from the trained models, it was found that Random Forest Classifier yielded the best and the optimal results. To train the model better, GridSearchCV was employed to find the best hyper-tuning parameters for this algorithm; after this process, the best hyperparameters were found and were employed during the training of the hyper-tuned model, which yielded the best results

3.3.6 Model Evaluation

Once the models are trained, performance evaluation is done using existing data; this is called cross-validation, where the test data is further split into subsets, and the model is evaluated against its 'n' number of times, with a different subset each time. This helps to reduce the impact of randomness and variability in the dataset. Various performance metrics such as Accuracy, Precision, and F1 Score are used to evaluate and decide the model as the best model.

3.4 Architecture

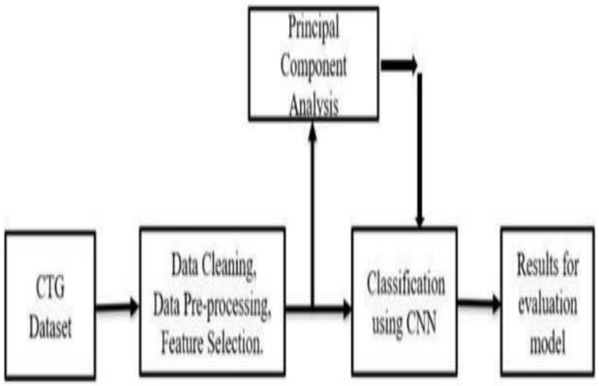
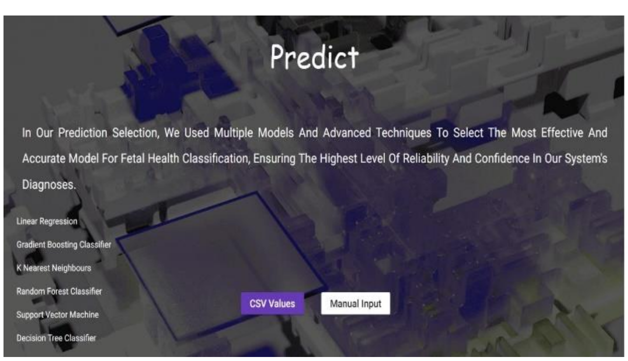
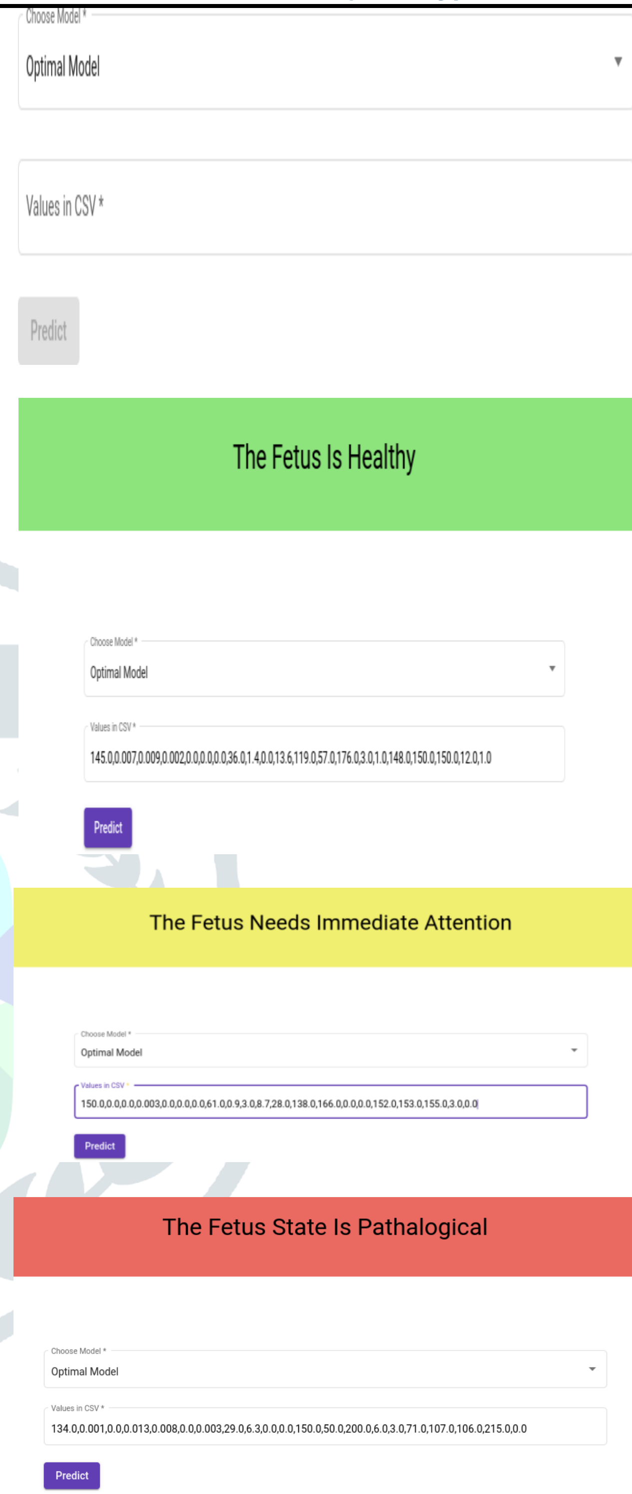


Fig 1: System Architecture

4. RESULT SCREEN SHOTS





5. CONCLUSION

In conclusion, our Paper provides a novel and innovative solution to the challenges of fetal health classification from CTG data, particularly in remote and rural areas where access to trained medical professionals and resources is limited. By leveraging the power of machine learning, a highly accurate model was deployed that can predict fetal health with a 97% success rate. This approach shows a significant improvement over traditional methods that rely on costly and time-consuming manual analysis by medical professionals. The model can provide accurate results quickly and at a fraction of the cost and time, which makes it a valuable tool while addressing the urgent need for maternal and child healthcare in developing countries. Overall, this project highlights the potential of machine learning to revolutionize healthcare delivery, particularly in resource-limited settings. By democratizing access to medical expertise, we can improve health outcomes and save lives. One of the most exciting aspects of our project is its potential for scalability and widespread adoption. The use of machine learning for fetal health classification has the potential to revolutionize the field of obstetrics and gynecology, providing accurate and timely diagnoses to women around the world. Our model can be easily deployed on a range of platforms, from desktop computers to mobile devices, as the application can be used on any device connected to the internet, making it highly accessible and adaptable to a range of healthcare settings.

6. REFERENCES

[1] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[2] Grivell, Rosalie M et al. "Antenatal cardiotocography for fetal assessment." The Cochrane database of systematic reviews vol. 2015,9 CD007863. 12 Sep. 2015, doi:10.1002/14651858.CD007863.pub4

[3] Gintautas, V., Ramonienė, G., Simanavičiūtė, D., "Cardiotocography," Cardiotocography. http://eknygos.lsmuni.lt/akuserijaen/Obstetrics.html

[4] Chudáček, V., Spilka, J., Burša, M. et al. Open access intrapartum CTG database. BMC Pregnancy Childbirth 14, 16 (2014). https://doi.org/10.1186/1471-2393-14-16

[5] Ogasawara, J., Ikenoue, S., Yamamoto, H. et al. Deep neural network-based classification of cardiotocograms outperformed conventional algorithms. Sci Rep 11, 13367 (2021). https://doi.org/10.1038/s41598-021-92805-9

[6] Khan M, Khurshid M, Vatsa M, Singh R, Duggal M and Singh K (2022) On AI Approaches for Promoting Maternal and Neonatal Health in Low Resource Settings: A Review. Front. Public Health 10:880034. doi: 10.3389/fpubh.2022.880034

[7] Spairani, Edoardo et al. "A deep learning mixed-data type approach for the classification of FHR signals." Frontiers in bioengineering and biotechnology vol. 10 887549. 8 Aug. 2022, doi:10.3389/fbioe.2022.887549

[8] Jacob, Shomona & Ramani, Geetha. (2012). Evolving Efficient Classification Rules from Cardiotocography Data through Data Mining Methods and Techniques. 78.

[9] [1R. Afridi, Z. Iqbal, M. Khan, A. Ahmad, and R. Naseem, "Fetal Heart Rate Classification and Comparative Analysis Using Cardiotocography Data and Known Classifiers," International Journal of Grid and Distributed Computing, vol. 12, no. 1, pp. 31–42, Jun. 2019, doi: 10.33832/ijgdc.2019.12.1.03. [Online]. Available: http://dx.doi.org/10.33832/ijgdc.2019.12.1.03

[10] Guangzhou An, Kazuko Omodaka, Satoru Tsuda, Yukihiro Shiga, Naoko Takada, Tsutomu Kikawa, Toru Nakazawa, Hideo Yokota, Masahiro Akiba, "Comparison of Machine-Learning Classification Models for Glaucoma Management", Journal of Healthcare Engineering, vol. 2018, Article ID 6874765, 8 pages, 2018. https://doi.org/10.1155/2018/6874765

[11] Sai Prasad Potharaju, M. Sreedevi, Vinay Kumar Ande, Ravi Kumar Tirandasu, "Data mining approach for accelerating the classification accuracy of cardiotocography", Clinical Epidemiology and Global Health, Volume 7, Issue 2, 2019, Pages 160-164, ISSN2213-3984, https://doi.org/10.1016/j.cegh.2018.03.004. (https://www.sciencedirect.com/science/article/pii/S2213398418300277)

[12] "Fetal Health Classification Using Deep Learning Techniques: A Comprehensive Review," by S. Sharma and R. K. Gupta (2021)

[13] "Fetal Health Classification Using Machine Learning Techniques and a Small Dataset," by L. Heredia-Jimenez et al. (2021)

[14] "Fetal Health Classification from Ultrasound Images Using Convolutional Neural Networks," by M. A. Rueda-Ruiz et al. (2020)

[15] "Fetal Health Classification from Cardiotocography Signals Using Deep Learning Techniques," by R. M. Ayres et al. (2020)

[16] "Fetal Health Classification Using Wavelet-Based Features and Random Forest Classifier," by S. S. Patil and S. S. Sonavane (2019).