JOURNAL OF DATA SCIENCE | Vol.2024:17
eISSN:2805-5160

Prediction of Fetal Health Status Using Machine Learning

Naidile S Saragodu, [1]Shreedhara N Hegde[1], Harprith Kaur[2]

[1]Dayananda Sagar Academy of Technology & Management, Karnataka, India
[2]Faculty of Data Science and Information Technology, INTI International University, Malaysia

Emails: naidilessaragodu99@gmail.com, shreedharadsatm@gmail.com

Abstract

The goal of this promising area of study is to enhance prenatal care and lower fetal morbidity and mortality by utilizing machine learning to anticipate fetal disease. In this study, we present a machine learning-based strategy for predicting fetal diseases from clinical data. First, we gathered a sizable collection of clinical information from expectant mothers with various fetal disorders. Using clinical guidelines, we pre-processed the data and retrieved pertinent features. We integrated a range of machine learning algorithms, including logistic regression, support vector machines, decision trees, and random forests, to train and test our model. We evaluated the performance of our model using several factors, such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). The results of this study demonstrate how machine learning algorithms can accurately forecast fetal health status. The developed models achieve good accuracy and AUC-ROC ratings to distinguish between healthy and at-risk fetuses. The interpretability study identifies key clinical characteristics that have a significant impact on the prediction, providing medical practitioners with useful information when making decisions about prenatal care. Through the provision of more unbiased and precise assessments of fetal health status, machine learning techniques incorporated into prenatal care have the potential to transform the industry. By providing accurate and early projections, this technology can assist healthcare professionals in identifying high-risk pregnancies and carrying out the necessary procedures, improving mother and fetal outcomes. Future research should concentrate on verifying and improving predictive models on larger and more varied datasets to ensure real-world applicability and reliability

Keywords

SVM, LR, Random Forest Classification, Fetal Health, Machine Learning.

Introduction

Fetal health is the state of the unborn child. In the United States, the fetal death rate in 2019 was roughly 5.6, or roughly 6 newborns per 1000 (Vimala, K., & Usha, D.,2020). Based on the sources, infants. The rate of fetal mortality can be reduced by being aware of fetal health issues and taking the appropriate safety measures (Arif, M. Z., 2020). The health of the fetus can also be inferred from other continuous aspects like acceleration, baseline value, histogram mean, kicks, etc., although there are several ways to do so depending on how these values are calculated. Therefore, our technique comprises utilizing and applying machine learning

Submission: 9 May 2024; **Acceptance:** 20 July 2024

principles to the dataset we have to forecast fetal health with a focus on accuracy. The phrase "dataset" describes a group of records (Ramla, M.,et;al., 2018). These records include the values that give the data some context and are essential in determining the health of the infant. For this prediction, the Kaggle dataset is being used. The dataset contains about 21 features and 21 records, but we only take into account 5 characteristics in our forecast. This process is known as preprocessing. If we offer these five features as its input, the Random- Forest-Classifier algorithm will provide the result as Normal, pathologic, and suspect (Bhowmik, P.,2021).

Pre-processing, also known as data pre-processing, is the process of choosing the data and excluding all extraneous aspects. Pre-processing is the process of choosing features based on how much of a contribution they make to health prediction. Pre-processing is the process of choosing features based on how much of a contribution they make to health prediction. We additionally clean our datasets during this process by deleting data with null feature values in order to obtain accurate and unambiguous data (Muhammad Hussain, N, et.al. 2022). The Random Forest Classifier is one machine learning approach for categorization and prediction. Random forest classifiers employ a large number of decision trees that are built from different subsets of the dataset and take into account the average in order to improve the dataset's predictive accuracy (Subasi, A., 2020) The Decision Tree Classifier and Random Forest Classifier differ primarily in two ways. The first one talks about the training set's data. The Decision Tree Classifier only needs one training set, whereas the RandomForest-Classifier needs many training datasets (Huang, Q. A, 2016). Another distinction is the number of trees required for prediction.

Based on the number of training datasets, Random Forest Classifiers require more than one tree for prediction, but Decision Tree Classifiers only require one tree. Because there is only one tree in the Decision Tree Classifier, there is only one output. Although there are more outputs since the Random-forest-Classifier employs more trees, the majority of these outputs will be used to produce the final result. When we use the majority, we can anticipate the result more precisely and with less overfitting. Overfitting is the phrase used in data science to describe the situation where a built-in model perfectly matches its training data. MongoDB is the database we employed for the implementation.

Methodology

Creating a report and examining the findings of tests performed on expectant women to identify and assess the condition of the fetus. The answer will be provided through a website created using a random forest classifier prediction algorithm based on datasets and the outcomes of countless prior tests. by determining the health, abnormality, or pathology of the fetus using the findings of tests conducted on pregnant women as input (Iraji, M. S. ,2019).

- Getting a clean dataset and obtaining the dataset.
- trying different things with the dataset to increase accuracy.
- to design an intuitive user interface.
- To use the provided input to create a PDF of a package's results.
- To send the user's email with the PDF.

Figure 1 depicts the learning model's flowchart, while Figures 2 depict the suggested model and system architecture. Using any random subset of the training dataset, the random forest classifier creates a series of decision trees. To choose the final class for the test objects, it integrates the votes from several decision trees.

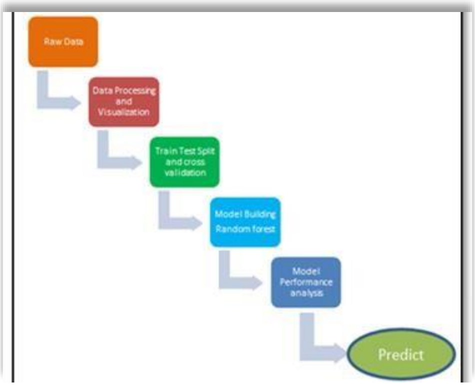

Fig1. Learning model flow

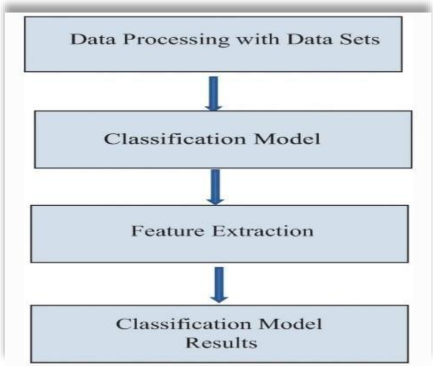

Fig2. Proposed learning model flow

State of fetus	*Abbreviations*	*FHR recordings*
Normal	N	1655
Suspect	S	295
Pathologic	P	176
Total		2126

Table1. Class distribution of CTGs

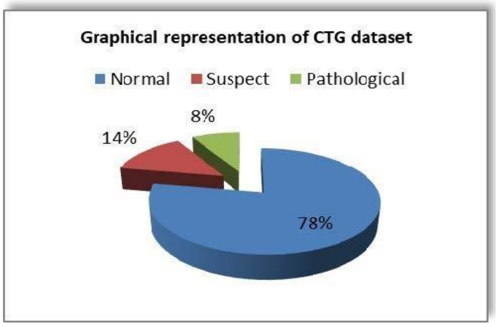

Fig3. Graphical representation of fetal dataset

Results and Discussion

While solving predictions involving unstructured data. Our data were modeled, and the accuracy was 98%. We've had good accuracy thus far, as shown in Figure 7.

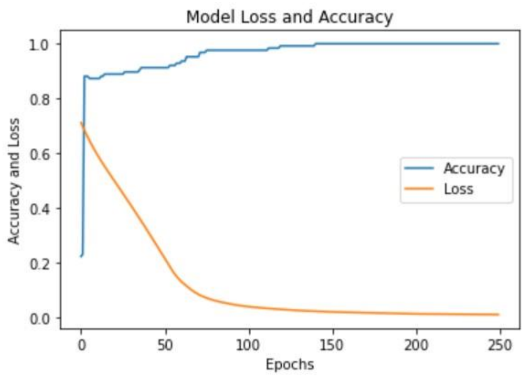

Fig7.Accuracy & loss graph for Random forest

Fetal heart rate, acceleration, mobility, uterine contractions, and general fetal health are tracked after applying ML algorithms to various datasets. Figure 8 depicts the fetal heart rate at rest; less than 110 bpm is deemed "abnormal" or "slow heart rate," and the range of 110 to 160 bpm is regarded "normal."

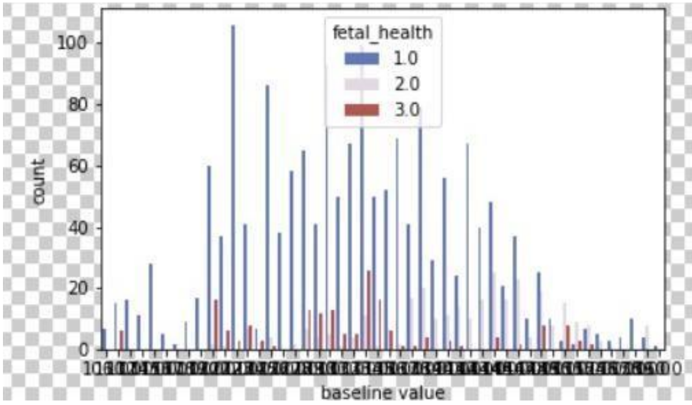

Fig8. Baseline Value

The fetal heart rate acceleration per second is depicted in Figure 9. Short-term increases in heart rate that continue at least 15 seconds or are at least 15 beats per minute in the "x-axis" are regarded "normal," whereas other instances are deemed "abnormal." It can be found by looking at the oxygen supply relative to mm Hg in the 'y-axis'. Figure 10 shows how many fetal movements occur every second, with 10 or fewer considered "normal." The movement is deemed "abnormal" if it is not recorded at the designated time. The 'normal' condition for the uterine contraction in Fig. 11 is 0.005 milliseconds. It is 0.001 milliseconds for the suspect and 0.003 milliseconds for 'pathological' or 'abnormal' conditions. The overall impact of the factors on the fetal health status, with 1 denoting a "normal" heart rate and anything above being seen as "suspicious" or "pathological."

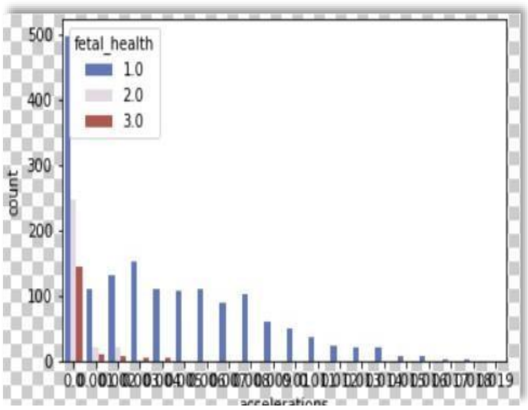

Fig9. Fetal heart rate acceleration

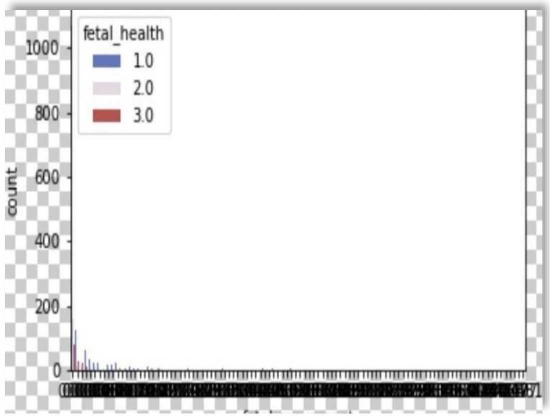

Fig10. Fetal movements

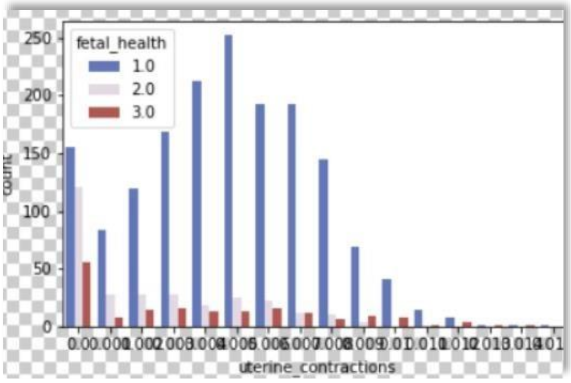

Fig11. Uterine Contraction

Based on the input of the data, our approach predicts that 92–98% of the parameters will be normal and that 2-4% would be ambiguous or aberrant, as illustrated in Fig. 12.

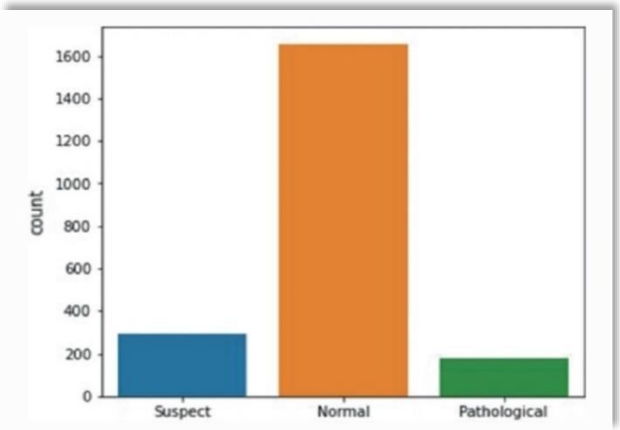

Fig12. Fetal health with results

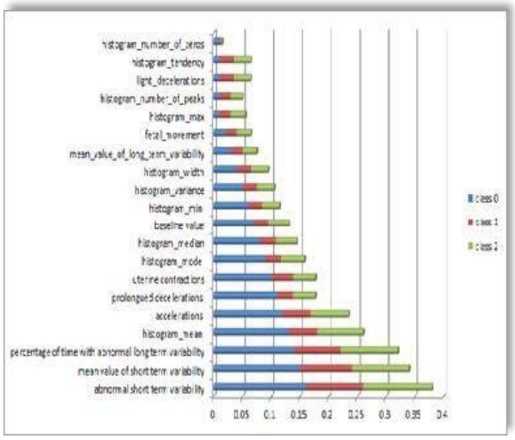

Fig13.Proof how we select 5 parameters

1. Short-Term Abnormal Variability: The fetal heart's short-term variability is the range of 3 to 5 beats per minute (BPM) that it experiences from one second to the next.
2. The average of the absolute variances between subsequent intervals is the "short-term variability mean value" of the ECG.
3. Period with an abnormal long-term variation: The fetal heart rate "acceleration" is a considerable long-term variation.
4. They are frequently triggered by fetal movement, last 10 to 20 seconds or more, and are at least 15 beats per minute quicker than the baseline.
5. Histogram mean: The histogram's mean displays the fetus's heart rate throughout various time periods.

Conclusion

In summary, predicting fetal health status using machine learning has shown considerable potential. Researchers and medical experts have been able to create precise models for determining the health of unborn children by utilizing a variety of algorithms and vast datasets.

These models can aid in risk assessment and early intervention, which will eventually improve outcomes for both the mother and the fetus.

Machine learning algorithms can efficiently analyze complicated patterns and create predictions about fetal health status by using characteristics like maternal health records, fetal heart rate patterns, and other pertinent information. These models may be able to spot irregularities and difficulties, such as intrauterine growth restriction, fetal distress, and congenital defects, which can have a big impact on the fetus's health.

More individualized and proactive methods of treating pregnancies may result from the use of machine learning in prenatal care. Informed decisions on interventions, such as medical treatments, closer monitoring, or even early delivery if necessary, may be made by healthcare professionals with the correct and timely knowledge about the fetal health status.

While machine learning has shown potential in this field, it is crucial to remember that technology has its limits. The caliber and variety of the training data have a significant impact on the models' accuracy. When applying machine learning in healthcare, it's also important to carefully evaluate ethical issues, privacy issues, and the requirement for interpretability in the decision-making process.

Acknowledgment

The authors would like to express our heartfelt gratitude to Dayananda Sagar Academy of Technology and Management (DSATM) for providing us with the necessary resources and facilities to conduct this research project Prediction of Fetal Health Status using machine learning.

Reference

Vimala, K., & Usha, D. (2020). An efficient classification of congenital fetal heart disorder using improved random forest algorithm. *International Journal of Engineering Trends and Technology*, *68*(12), 182-186.

Arif, M. Z., Ahmed, R., Sadia, U. H., Tultul, M. S. I., & Chakma, R. (2020). Decision tree method using for fetal state classification from cardiotography data. *Journal of Advanced Engineering and Computation*, *4*(1), 64-73.

Ramla, M., Sangeetha, S., & Nickolas, S. (2018, June). Fetal health state monitoring using decision tree classifier from cardiotocography measurements. In *2018 second international conference on intelligent computing and control systems (ICICCS)* (pp. 1799-1803). IEEE.s

Bhowmik, P., Bhowmik, P. C., Ali, U. M. E., & Sohrawordi, M. (2021). Cardiotocography data analysis to predict fetal health risks with tree-based ensemble learning. *Inf. Technol. Comput. Sci*, *5*, 30-40.

Muhammad Hussain, N., Rehman, A. U., Othman, M. T. B., Zafar, J., Zafar, H., & Hamam, H. (2022). Accessing artificial intelligence for fetus health status using hybrid deep learning algorithm (AlexNet-SVM) on cardiotocographic data. *Sensors*, *22*(14), 5103.

Subasi, A., Kadasa, B., & Kremic, E. (2020). Classification of the cardiotocogram data for anticipation of fetal risks using bagging ensemble classifier. *Procedia Computer Science*, *168*, 34-39.

Huang, Q. A., Dong, L., & Wang, L. F. (2016). Cardiotocography analysis for fetal state classification using machine learning algorithms. *J. Micro Electromech. Syst*, *25*.

Iraji, M. S. (2019). Prediction of fetal state from the cardiotocogram recordings using neural network models. *Artificial intelligence in medicine*, *96*, 33-44.

Bibliography

Chandana, C., Neha, P. N., Nisarga, S. M., Thanvi, P., & Balarengadurai, C. (2021). Fetal health prediction using machine learning approach. *SpringerLink*. https://doi.org/10.1007/s00542-021-05927-8

Mehbodniya, A., Choudhary, A., & Mehbodniya, M. (2022). Optimizing fetal health prediction: Ensemble modeling with fusion of cardiotocogram data. *ScienceDirect*. https://doi.org/10.1016/j.compbiomed.2022.105637

Lee, L. H., & Noble, J. A. (2020). Automatic determination of the fetal cardiac cycle in ultrasound using spatio-temporal neural networks. *IEEE International Symposium on Biomedical Imaging (ISBI)*. https://doi.org/10.1109/ISBI45749.2020.9098371

Liu, G. F., & Gao, P. C. (2019). Predicting fetal health using cardiotocograms: A machine learning approach. *TensorGate Research*. https://doi.org/10.5281/zenodo.3820823

Christopher, W. G., & Deressa, T. D. (2019). Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. *IEEE Access*, *7*, 112026-112036. https://doi.org/10.1109/ACCESS.2019.2934432