Baghdad Science Journal

Volume 22 | Issue 7 Article 30

7-25-2025

Enhancing Fetal Health Assessment Using Machine Learning and XAI Techniques

Niaz Ashraf Khan
Department of CSE, BRAC University, Dhaka, Bangladesh, niaz.ashraf@bracu.ac.bd

A. M. Saadman Rafat
Department of ECE, North South University, Dhaka, Bangladesh, saadmanhere@gmail.com

How to Cite this Article

Khan, Niaz Ashraf and Rafat, A. M. Saadman (2025) "Enhancing Fetal Health Assessment Using Machine Learning and XAI Techniques," *Baghdad Science Journal*: Vol. 22: Iss. 7, Article 30.
DOI: https://doi.org/10.21123/2411-7986.5010

RESEARCH ARTICLE

Enhancing Fetal Health Assessment Using Machine Learning and XAI Techniques

Niaz Ashraf Khan [1,*], A. M. Saadman Rafat [2]

[1] Department of CSE, BRAC University, Dhaka, Bangladesh
[2] Department of ECE, North South University, Dhaka, Bangladesh

ABSTRACT

Fetal health assessment is crucial for ensuring the well-being of both mother and fetus during pregnancy. Accurate monitoring can reduce child and maternal mortality rates, particularly in low- and middle-income countries where these rates are higher. This research aims to enhance fetal health classification using advanced machine learning techniques and Explainable Artificial Intelligence (XAI) to address the gaps in predictive accuracy and model transparency. A dataset of 2,126 records with 22 features from Cardiotocography (CTG) readings. The dataset was balanced using Synthetic Minority Over-sampling Technique (SMOTE) to handle class imbalances. Various machine learning algorithms, including LightGBM, XGBoost, and Random Forest, and employed Pearson's Correlation for feature selection were implemented. Shapley values were used to ensure model interpretability. LightGBM achieved the highest accuracy at 95.9%, followed by XGBoost at 95.5%. Feature importance and SHAP analysis revealed features that are critical for accurate predictions. Our study also demonstrates that combining machine learning with XAI can drastically improve fetal health monitoring by providing interpretable models. Ultimately, this contributes to more informed decision-making in fetal health monitoring and supports global efforts to reduce maternal and neonatal mortality in line with the Sustainable Development Goals.

Keywords: Cardiotocography, Ensemble model, Fetal health, SMOTE, XAI

Introduction

Fetal health assessment is a key part of prenatal care, helping ensure both the mother and baby stay healthy throughout pregnancy. As the baby grows and develops over the nine months, regular check-ups become crucial to catch any potential issues early. One of the most trusted tools for monitoring the baby's well-being is Cardiotocography (CTG), often known as electronic fetal monitoring. It works by using a small ultrasound sensor placed on the mother's belly to keep track of the baby's heartbeat and the mother's contractions in real time. CTG is especially useful during the final stages of pregnancy and during labor, giving doctors important insights to make timely and informed decisions.

The primary goal of using CTG is to classify fetal health and prevent child and maternal mortality, aligning with the United Nations' Sustainable Development Goals (SDGs)[1] to reduce preventable deaths of newborns and children under the age of five and to lower maternal mortality rates. As of 2017, maternal mortality accounted for approximately 295,000 deaths. According to UNICEF, in 2018, maternal and fetal mortality rates were significantly higher in low- and middle-income countries, with seven deaths per 1000 live births, compared to three in higher-income countries.

The dataset includes 2,126 records, each containing 22 features that capture different aspects of the CTG readings. The analysis found that certain features like heart rate accelerations, prolonged slowdowns

Received 30 May 2024; revised 23 August 2024; accepted 25 August 2024.
Available online 25 July 2025

* Corresponding author.
E-mail addresses: niaz.ashraf@bracu.ac.bd (N. A. Khan), saadmanhere@gmail.com (A. M. S. Rafat).

https://doi.org/10.21123/2411-7986.5010

(decelerations), unusual short-term variability, a high percentage of abnormal long-term variability, and the average value of long-term variability are strongly linked to fetal health. Many factors can affect a baby's development before birth, including genetic conditions, illnesses in the mother, environmental exposures, or complications during pregnancy.

To boost the accuracy of predicting fetal health, we tested a wide range of machine learning models. We also explored other techniques like AdaBoost, Logistic Regression, Stochastic Gradient Descent (SGD), Perceptron, Linear Discriminant Analysis (LDA), Naive Bayes, and Quadratic Discriminant Analysis (QDA). Before implementing these algorithms, the dataset was balanced using SMOTE to address class imbalances. Pearson's Correlation was utilized to identify and remove highly correlated features (correlation coefficient > 0.8) to prevent duplication and improve model performance.

This study aims at leveraging advanced machine learning techniques to classify fetal health accurately. By providing actionable insights into fetal well-being, the work supports the broader goal of achieving the SDGs by 2030.[1] Ensuring that all births are attended by skilled health professionals, coupled with the use of effective monitoring tools and predictive algorithms, can make the difference between life and death for both the mother and the newborn.

This study introduces several novel aspects to the field of fetal health monitoring:

- Firstly, it involves a comprehensive analysis of key features correlated with fetal health. By identifying and analyzing these features, the study provides deeper insights into which parameters are most indicative of potential complications.
- Secondly, it systematically compares a wide range of ML algorithms, including traditional models and advanced techniques like neural networks and ensemble methods. Unlike many earlier studies that only explored a few algorithms, our work takes a more comprehensive approach by evaluating a wider range of models.
- Thirdly, it integrates XAI into the ML models. By ensuring that the models are not only accurate but also transparent and explainable, the research addresses a critical gap in the current application of ML in healthcare, where interpretability is often overlooked.

The primary technical gap this study addresses is the lack of transparent, accurate, and reliable ML models for fetal health monitoring. While many studies have explored the use of ML in this field, few have effectively combined high accuracy with interpretability. This study bridges this gap by using XAI techniques to create models that are both accurate and explainable.

Another key challenge is the imbalance in fetal health data, where abnormal cases are much rarer than normal ones, and this study addresses it by using SMOTE to create a more balanced dataset. In summary, this study leverages the power of machine learning to improve fetal health monitoring by addressing key technical gaps and ensuring that the models are both accurate and explainable.

Related works

Abubaker et al.[2] demonstrate that employing the extra tree feature selection method significantly enhances the accuracy of heart disease prediction models. By prioritizing key features, their research streamlines computational complexity which resulted in an improved accuracy with the extra tree classifier. Jassim et al.[3] concentrate on refining lung cancer prediction through the utilization of deep learning and ensemble techniques. Afridi et al.[4] tackled the important issue of fetal distress during childbirth by analyzing fetal heart rate (FHR) data collected through cardiotocography (CTG). Mandala's[5] study presents a new machine-learning method for classifying fetal health using a LightGBM classifier. Kannan et al.[6] evaluate the calibration of classifiers based on rules, trees, and functions to explore uncertain information in the Cardiotocography (CTG) dataset.

Ramla[7] highlights the importance of Cardiotocography (CTG) as a crucial tool for monitoring fetal health during both pregnancy and after birth. Krupa et al.[8] introduced a new method for analyzing FHR using CTG, a common tool for fetal monitoring. Mendis et al.,[9] showcases the importance of CTG in monitoring fetal distress during labor by analyzing FHR and UC. Warmerdam et al.[10] discusses how analyzing the variability in fetal heart rate through spectral analysis can indicate the baby's health. Zhao et al.,[11] proposes a computer-aided diagnostic tool called Hybrid-FHR to address this issue. Spairani et al.,[12] addresses the challenge of interpreting Cardiotocography (CTG) signals for assessing fetal well-being, a practice prone to significant variability among observers. Esteban et al.[13] studied predicting acidemia in newborns by examining specific fetal heart signal features in a 30-minute window, especially focusing on the last deceleration before birth. Das et al.[14] introduce a soft computing-based method for classifying fetal health using cardiotocograph (CTG) data during labor. Their proposed approach involves preprocessing CTG signals, extracting relevant features, and utilizing soft computing algorithms for classification. The study provided in Ayres-de-Campos et al.,[15] aims to offer assistance in utilizing and interpreting intrapartum cardiotocography (CTG) while managing specific CTG patterns in clinical settings. The study detailed by Sharma et al.,[16] dives into the evaluation of cardiotocography (CTG) in high-risk pregnancies and its association with both maternal and fetal outcomes. Conducted on 200 antenatal patients with high-risk pregnancies, the study correlates CTG analyses, as per RCOG guidelines, with various outcomes. Results indicate that patients with non-reactive CTG faced significantly higher risks of premature rupture of membranes and prolonged labor.

Salini et al.[17] developed machine learning models to improve fetal health classification using cardiotocography data. Their models achieved a high accuracy of 93%. The researchers explored various ML models, including Random Forests, Logistic Regression, Decision Trees, Support Vector Classifiers, Voting Classes, and K-Nearest Neighbors. In Petrozziello et al.,[18] the authors focused to improve how fetal heart rate (FHR) monitoring during childbirth is assessed. The study found that their models performed better than traditional methods in predicting cord acidemia.

Nafea et al.[19] introduce an ensemble model combining decision trees, support vector machines, random forests, and ADA-boost for detecting adverse drug reactions (ADRs). Omotunde and Mouhamed[20] provide a comprehensive overview of how artificial intelligence (AI) systems are revolutionizing healthcare, particularly in the management and analysis of electronic medical records (EMRs). Mohammed et al.[21] address the challenge of skin cancer detection using a hybrid model that combines DenseNet201 and an auto-encoder for feature extraction with a support vector machine (SVM) classifier. Kadhim et al.[22] explored the application of ensemble machine learning algorithms, including SVM, and several ensemble models to detect for prostate cancer using MRI scans.

Unlike earlier studies, this research fills an important gap by combining high prediction accuracy with model interpretability in fetal health monitoring. While many previous works focused on testing different machine learning algorithms, few have successfully included Explainable AI (XAI) to make the models more transparent and trustworthy for clinical use. We also addressed the issue of data imbalance by applying the Synthetic Minority Over-sampling Technique (SMOTE) to ensure fair representation of all health conditions.

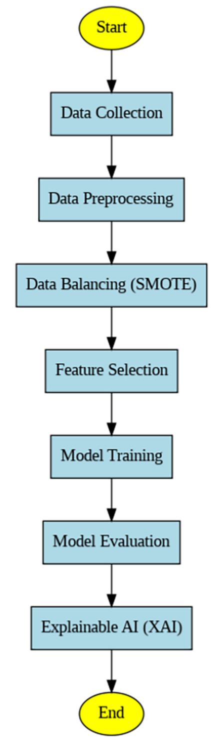

Fig. 1. Flowchart of the methodology.

Materials and methods

The methodology of this study follows a clear, step-by-step approach, as shown in Fig. 1. It begins with data collection and preprocessing, followed by balancing the dataset using SMOTE, selecting the most important features, training various models, evaluating their performance, and finally using XAI to showcase interpretability.

The dataset comprises 2126 records with 22 features capturing various parameters from CTG readings. The features include numerical attributes such as ‘mean value of short-term variability’ and ‘percentage of time with abnormal long-term variability,’ and categorical attributes such as ‘fetal health class.’

The important steps of the methodologies are described in detail below:

1. Data Preprocessing The initial step involves cleaning and preparing the data to ensure its quality and suitability for analysis. This includes handling missing values, normalizing numerical features, and encoding categorical variables into numerical formats compatible with machine learning algorithms. In the data preprocessing step, missing values were handled using multiple imputation methods. Numerical features

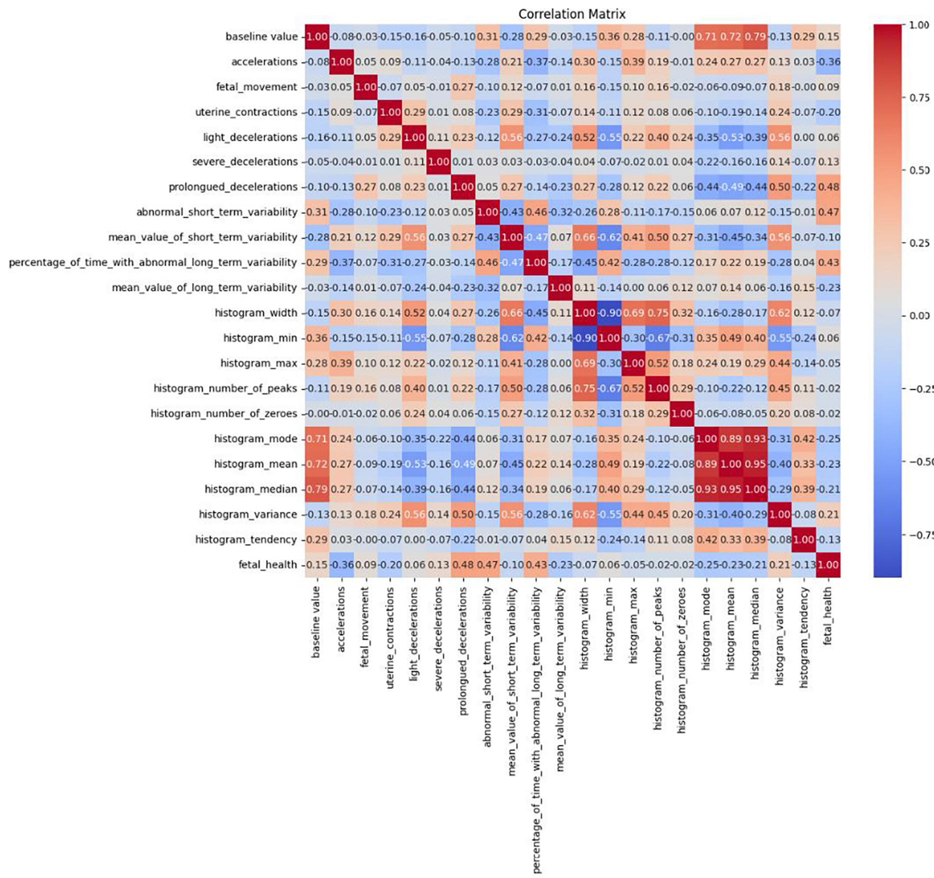

Fig. 2. Correlation map of the features.

were normalized using Min-Max scaling and Z-score normalization to standardize the data. Categorical variables were encoded using one-hot encoding for nominal categories and label encoding for ordinal categories.

2. Data Balancing: To handle class imbalances in the dataset SMOTE was applied: The following steps are involved:
   - For each minority class sample (x_i), find its (k)-nearest neighbors.
   - Randomly select one of the (k)-nearest neighbors $(x_{i,k})$.
   - Generate a new synthetic sample (x_{new}) along the line segment joining (x_i) and $(x_{i,k})$ as defined in Eq. (1)

   $$\left[x_{new} = x_i + \lambda\left(x_{i,k} - x_i\right)\right] \tag{1}$$

   where (λ) is a random number between 0 and 1.
3. Feature Selection using Pearson Correlation Coefficient To reduce dimensionality and remove redundant features; Pearson's correlation coefficient is calculated between pairs of features. Features with a high correlation coefficient (+0.8 and –0.8 in this research) are considered redundant and removed to prevent multicollinearity, thereby enhancing model performance. In total, 4 features were removed based on this threshold Fig. 2 contains the correlation matrix of the features in the dataset. Pearson correlation coefficient is defined in Eq. (2).

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{2}$$

where $(\bar{X})$ and $(\bar{Y})$ are the means of (X) and (Y), respectively.

4. Application of Machine Learning and Deep Learning Methods: A range of ML and DL algorithms to classify fetal health was applied. The description of the best performing algorithm is provided below:
   - XGBoost is based on the gradient boosting framework, which builds an ensemble of decision trees in a sequential manner. Each

tree is constructed to correct the errors of the previous tree, thus improving the overall model accuracy. XGBoost includes regularization terms in its objective function to prevent overfitting. It employs both L1 (Lasso) and L2 (Ridge) regularization to control the complexity of the model, making it more robust and generalizable.

Unlike traditional gradient boosting methods that use first-order gradients, XGBoost incorporates both first-order and second-order gradients in its optimization process. The objective function is shown in Eq. (3).

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} l\left(y_i, \widehat{y_i}\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{3}$$

where:
– $(l(y_i, \widehat{y_i}))$ is the loss function (e.g., mean squared error for regression).
– $(\Omega(f_k))$ is the regularization term.

• LightGBM also uses a gradient boosting framework but optimizes the training process by growing trees leaf-wise (best-first) instead of level-wise.
• Random Forest is an ensemble method which builds on multiple models (decision trees) and combines their outputs.

For classification, the final prediction $(\hat{y})$ as in Eq. (4) is the mode of the predictions from all trees $(\{h_1(x), h_2(x), h_m(x)\})$:

$$\hat{y} = \text{mode}\left(h_1(x), h_2(x), h_m(x)\right) \tag{4}$$

For regression, the final prediction $(\hat{y})$ is the average of the predictions as shown in Eq. (5).

$$\hat{y} = \frac{1}{m} \sum_{j=1}^{m} h_j(x) \tag{5}$$

5. Integration of Explainable Artificial Intelligence (XAI): To make the model more transparent and easier to understand, we used Shapley values to see how much each feature contributes to the model's predictions. Eq. (6) shows how the Shapley values are calculated:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \left[v(S \cup \{i\}) - v(S)\right] \tag{6}$$

Here, $(v(S))$ is the value of the coalition (S).

Table 1. Model accuracy comparison.

ML Algorithm	Accuracy
XGBoost	95.6%
LightGBM	94.8%
Random Forest	94.4%
Gradient Boosting	94.4%
Extra Trees	93.9%
Decision Tree	90.1%
Neural Network	89.5%
SVM	88.7%
K-Nearest Neighbors	88.4%
Adaboost	87.9%
SGD	87.5%
Logistic Regression	86.7%
Perceptron	85.3%
LDA	81.7%
Naïve Bayes	70.4%
QDA	56.6%

Table 2. Precision, recall and F1 score of the ensemble algorithms.

ML Algorithm	Precision	Recall	F1 Score
Random Forest	0.920	0.903	0.910
Decision Tree	0.888	0.913	0.900
Extra Trees	0.900	0.853	0.875
XGBoost	0.928	0.938	0.933
LightGBM	0.950	0.940	0.945

Results and discussion

The results of our analysis on fetal health, using a range of machine learning models, are summarized below. Accuracy of various models is compared, feature importance is evaluated, and SHAP analysis is used to comprehend the influence of individual features on predictions.

Model accuracy comparison

The comparison of accuracies of various machine learning model is shown in Table 1. The precision, recall and the F1-Score of the ensemble algorithms are provided in Table 2. The models are listed on the y-axis, and their accuracy scores are on the x-axis. The models are ranked from highest to lowest accuracy. Here are some key observations:

• LightGBM has the highest accuracy at 0.959.
• XGBoost follows closely with an accuracy of 0.955.
• Gradient Boosting, Random Forest, and Extra Trees also show high accuracies of 0.945, 0.939, and 0.937, respectively.
• Traditional models like Logistic Regression, SGD, Perceptron, and LDA have lower accuracies, ranging from 0.851 to 0.815.
• Naive Bayes and QDA have the lowest accuracies, 0.737 and 0.643, respectively.

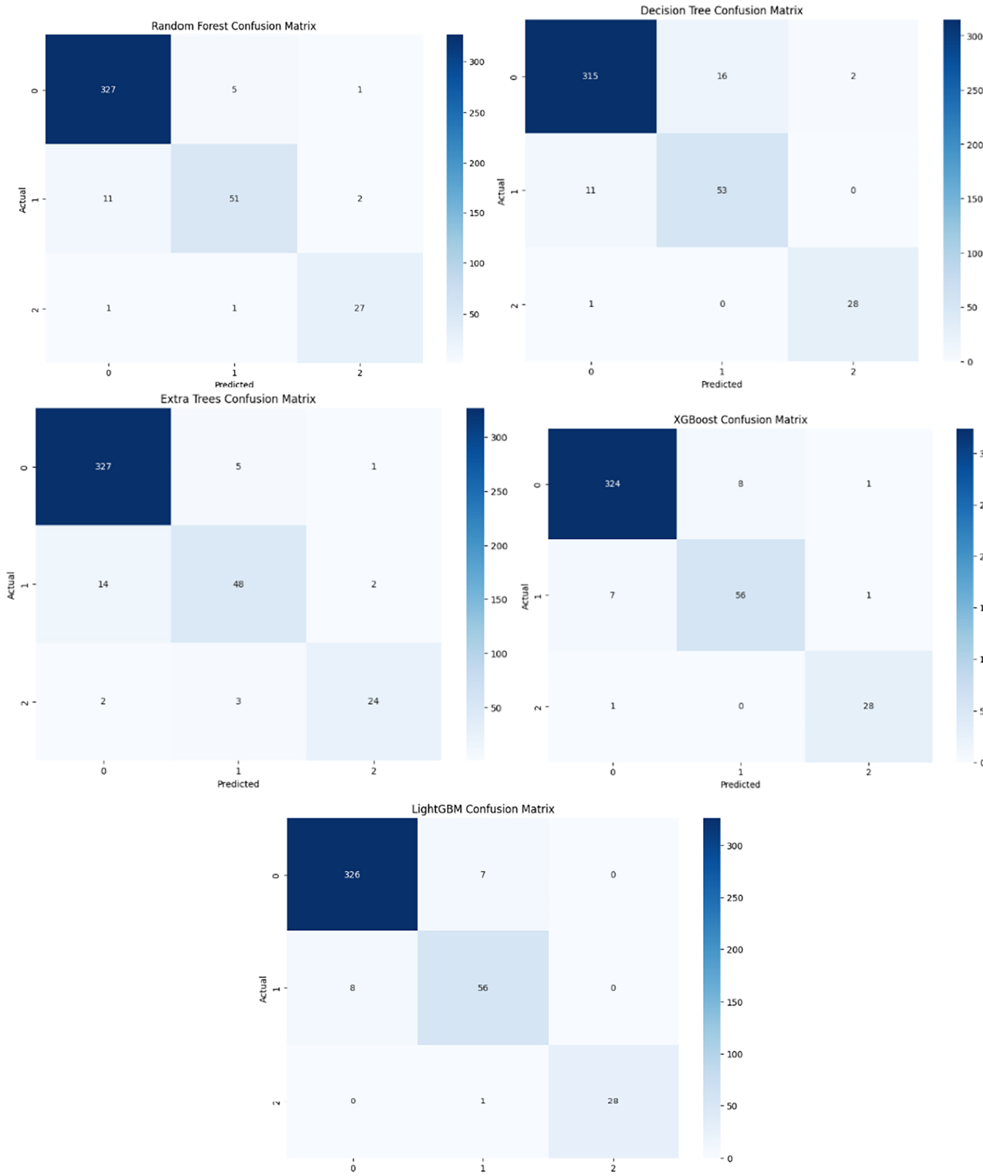

Fig. 3. Confusion matrix for RF, decision tree, extra trees, XGBoost, LightGBM.

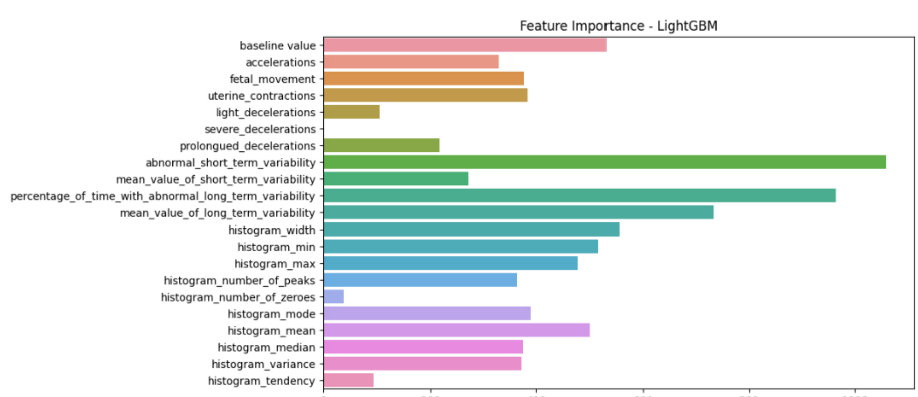

Fig. 4. Important features for LightGBM.

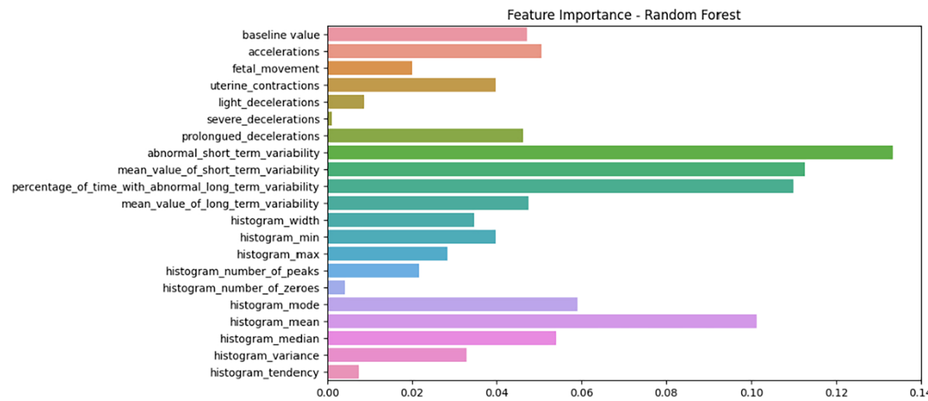

Fig. 5. Important features for random forest.

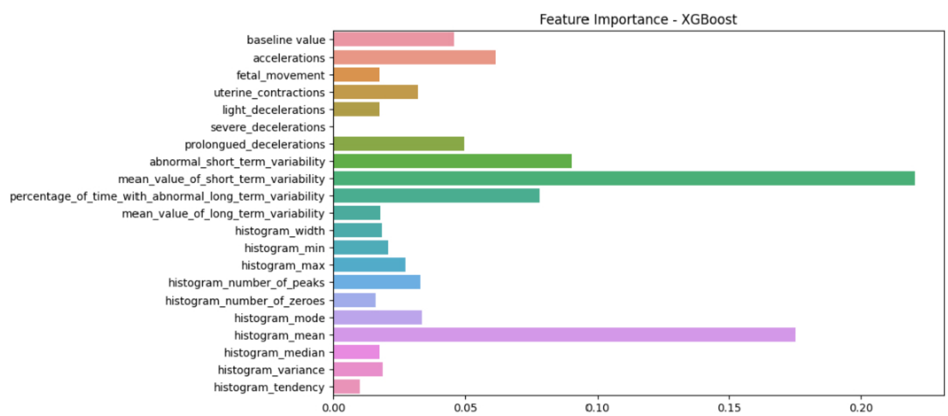

Fig. 6. Important features for XGBoost.

Table 3. Comparison with previous works.

ML Algorithm	ML Model	Accuracy
This Research	XGBoost	95.6%
Chamidah N, et al.[23]	SVM	90.64%
Piri J, et al.[24]	XGBoost	94%

This indicates that ensemble methods and tree-based algorithms generally perform better in this particular context. Table 3 also contains comparison with previous state-of-the-art works.

Confusion matrices

The confusion matrices for the top-performing models, including Random Forest, Decision Tree, Extra Trees, XGBoost, and LightGBM, provide a detailed view of the classification performance in Fig. 3. These matrices illustrate the true positives, false positives, true negatives, and false negatives, which give insight into the models' accuracy.

Feature importance analysis

Feature importance analysis for LightGBM shown in Fig. 4, Random Forest shown in Fig. 5, and XGBoost

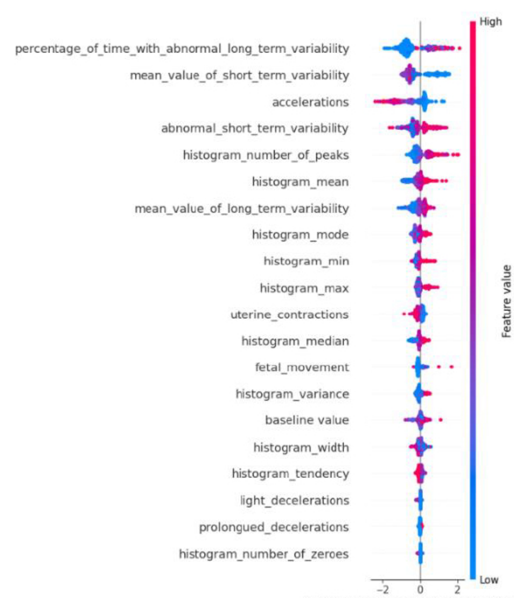

Fig. 7. SHAP analysis on LightGBM.

shown in Fig. 6, models was conducted to identify the key features impacting model predictions. The top features included:

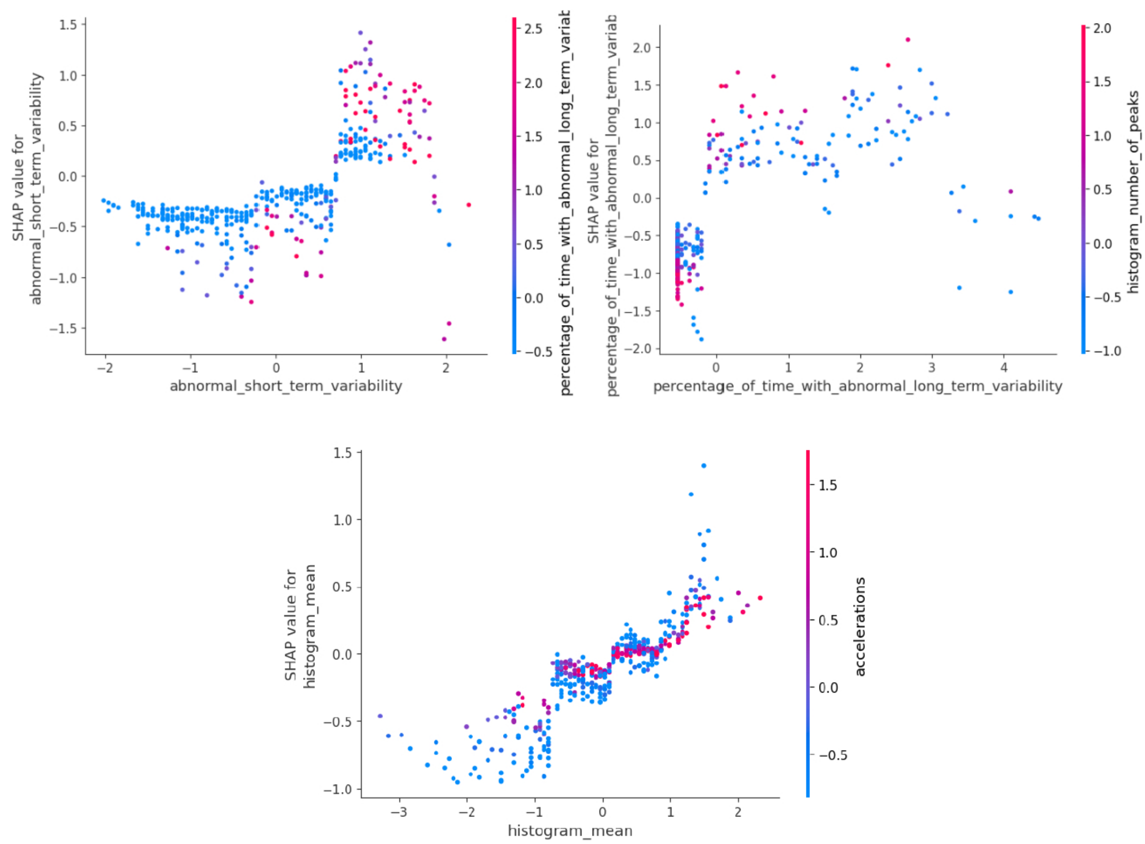

Fig. 8. SHAP plots for the best features of LightGBM.

- Percentage of time with abnormal long-term variability
- Mean value of short-term variability
- Accelerations
- Abnormal short-term variability

SHAP analysis

The SHAP analysis for LightGBM provided a deeper understanding of feature impacts on model predictions as shown in Fig. 7, The SHAP value plot highlighted additional features, such as accelerations and abnormal short-term variability, which were not as prominently featured in traditional feature importance plots. This suggests that SHAP values offer a more nuanced understanding of feature impacts on model predictions. The SHAP plots for the best features of the LightGBM model, as shown in in Fig. 8, demonstrated the contribution of each feature to the model's predictions.

Conclusion

Machine learning has shown significant potential in improving the accuracy and reliability of fetal health monitoring. Integrating these models into clinical practice requires creating user-friendly interfaces, validating the models in clinical settings, and addressing challenges like data privacy, user acceptance, and system compatibility. Ensemble methods, particularly LightGBM and XGBoost, have proven highly effective in classifying fetal health states.To make the models more understandable and transparent, we used SHAP values, which help explain how specific features influence predictions. The model's performance was improved by addressing class imbalance with SMOTE and applying a careful feature selection process. Future work will explore feature engineering techniques, such as polynomial features, interaction terms, and domain-specific transformations, to enrich the feature set and boost model performance.

Authors' declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images, that are not ours, have been included with the necessary permission for republication, which is attached to the manuscript.
- Authors sign on ethical consideration's approval.
- No animal studies are present in the manuscript.

- Ethical Clearance: The project was approved by the local ethical committee at University of Liberal Arts Bangladesh (ULAB), Dhaka, Bangladesh.

Authors' contribution statement

N. A. K. and A. M. S. R. designed the study and drafted the result. N.A.K. wrote the manuscript, performed simulations and analyzed the data. A. M. S. R. read and approved the final manuscript.

References

1. SDG 3: Ensure healthy lives and promote well-being for all at all ages. UN Women. 2023.
2. Abubaker H, Muchtar F, Khairuddin AR, Nuar ANA, Yunos ZM, Salimun C. Exploring important factors in predicting heart disease based on ensemble-extra feature selection approach. Baghdad Sci J. 2024;21(2):0812. https://doi.org/10.21123/bsj.2024.9711.
3. Jassim OA, Abed MJ, Saied ZHS. Deep learning techniques in the cancer-related medical domain: a transfer deep learning ensemble model for lung cancer prediction. Baghdad Sci J. 2024;21(3):1101–1118. https://doi.org/10.21123/bsj.2023.8340.da.
4. Afridi R, Iqbal Z, Khan M, Ahmad A, Naseem R. Fetal heart rate classification and comparative analysis using cardiotocography data and known classifiers. Int J Grid Distrib Comput. 2019;12:31–42. https://doi.org/10.33832/ijgdc.2019.12.1.03.
5. Mandala SK. Unveiling the unborn: advancing fetal health classification through machine learning. Artif Intell Health. 2023;1:2121. https://doi.org/10.36922/aih.2121.
6. Kannan E, Ravikumar S, Anitha A, Kumar S, Vijayasarathy M. Analyzing uncertainty in cardiotocogram data for the prediction of fetal risks based on machine learning techniques using rough set. J Ambient Intell Human Comput. 2021;21:01–13. https://doi.org/10.1007/s12652-020-02803-4.
7. Ramla M. Influence of feature selection methods on cardiotocography data: a quantitative investigation. Int J Eng Adv Technol. 2019;8:2249–8958. https://doi.org/10.35940/ijeat.D1006.0484S219.
8. Krupa N, Mohd Ali M, Zahedi E, Ahmad S, Hassan F. Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. Biomed Eng Online. 2011;10:6. https://doi.org/10.1186/1475-925X-10-6.
9. Mendis L, Palaniswami M, Brownfoot F, Keenan E. Computerised cardiotocography analysis for the automated detection of fetal compromise during labour: A review. Bioengineering. 2023;10:1007. https://doi.org/10.3390/bioengineering10091007.
10. Warmerdam G, Vullings R, Bergmans J, Oei G. Reliability of spectral analysis of fetal heart rate variability. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014. 2014;2817–20. https://doi.org/10.1109/EMBC.2014.6944209.
11. Zhao Z, Zhu J, Jiao P, Wang J, Zhang X, Lu X, *et al.* Hybrid-FHR: a multi-modal AI approach for automated fetal acidosis diagnosis. BMC Med Inform Decis Mak. 2024;24:19. https://doi.org/10.1186/s12911-024-02423-4.
12. Spairani E, Daniele B, Signorini MG, Magenes G. A deep learning mixed-data type approach for the classification of FHR signals. Front Bioeng Biotechnol. 2022;10. https://doi.org/10.3389/fbioe.2022.887549.
13. Esteban L, Castán B, Esteban Escaño J, Sanz-Enguita G, Laliena A, Chóliz-Ezquerro M, *et al.* Machine learning algorithms combining slope deceleration and fetal heart rate features to predict acidemia. Appl Sci. 2023;13:7478. https://doi.org/10.3390/app13137478.
14. Das S, Mukherjee H, Roy K, Saha C. Fetal health classification from cardiotocograph for both stages of labor – A soft computing based approach. 2022. https://doi.org/10.3390/diagnostics13050858Abstract.
15. Ayres-de-Campos D, Spong C, Chandraharan E, Valent S, Alhaidari T. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. Int J Gynecol Obstet. 2015;131:13–24. https://doi.org/10.1016/j.ijgo.2015.06.020.
16. Sharma S, Grover M, Behal M, Jindal M, Goel V, Minhas D. Role of cardiotocography in improving maternal and fetal outcomes in high-risk pregnancy. J Cardiovasc Dis Res. 2024;14:683–7.
17. Y Salini, Mohanty SN, Ramesh JV, Yang M, Chalapathi MMV. Cardiotocography data analysis for fetal health classification using machine learning models. IEEE Access. 2024;12:26005–22. https://doi.org/10.1109/ACCESS.2024.3364755.
18. Petrozziello A, Redman C, Papageorghiou AT, Jordanov I, Georgieva A. Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. IEEE Access. 2019:1–1. https://doi.org/10.1109/ACCESS.2019.2933368.
19. Nafea AA, Ibrahim MS, Mukhlif AA, AL-Ani MM, Omar N. An ensemble model for detection of adverse drug reactions. Aro-The Scientific Journal of Koya University. 2024;12(1):41–47. https://doi.org/10.14500/aro.11403.
20. Omotunde H, Mouhamed MR. The modern impact of artificial intelligence systems in healthcare: A concise analysis. Mesopotamian Journal of Artificial Intelligence in Healthcare. 2023;2023:66–70. https://doi.org/10.58496/MJAIH/2023/013.
21. Mohammed HJ, Nafea AA, Almulla HK, Aliesawi SAS, Al-Ani MM. An effective hybrid model for skin cancer detection using transfer learning. In 2023 16th International Conference on Developments in eSystems Engineering (DeSE). 2023;840–845. https://doi.org/10.1109/DeSE60595.2023.10468994.
22. Kadhim OJ, Nafea AA, Aliesawi SAS, Al-Ani MM. Ensemble model for prostate cancer detection using MRI images. In 2023 16th International Conference on Developments in eSystems Engineering (DeSE). 2023;492–497. https://doi.org/10.1109/DeSE60595.2023.10468866.
23. Chamidah N, Wasito I. Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine. In 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS). 2015;37–41. https://doi.org/10.1109/ICACSIS.2015.7415166.
24. Piri J, Mohapatra P, Dey R. Fetal health status classification using MOGA - CD based feature selection approach. In 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). 2020;1–6. https://doi.org/10.1109/conecct50063.2020.9198377.

تعزيز تقييم صحة الجنين باستخدام تقنيات التعلم الآلي وXAI

نياز أشرف خان[1]، أ. م. سعدمان رافت[2]

[1]قسم علوم الحاسب والهندسة، جامعة الفنون الحرة بنغلاديش، دكا، بنغلاديش.
[2]قسم الهندسة الكهربائية والحاسوبية، جامعة شمال الجنوب، دكا، بنغلاديش.

الخلاصة

تقييم صحة الجنين أمر حاسم لضمان رفاهية الأم والجنين طوال فترة الحمل. يهدف هذا البحث إلى تحسين تصنيف صحة الجنين باستخدام تقنيات التعلم الآلي المتقدمة والذكاء الاصطناعي المفسر (XAI) لمعالجة الفجوات في دقة التنبؤ وشفافية النموذج. استخدمنا مجموعة بيانات تتضمن 2126 سجلًا مع 22 ميزة من قراءات تخطيط القلب الرحمي .(CTG) تم تحقيق التوازن في مجموعة البيانات باستخدام تقنية الزيادة الاصطناعية للأقليات (SMOTE) للتعامل مع التفاوتات في الفئات. طبقنا خوارزميات تعلم آلي مختلفة، بما في ذلك LightGBM و XGBoostو Random Forest، واستخدمنا معامل الارتباط لبيرسون لاختيار الميزات. استخدمنا قيم شابلي لضمان تفسير النموذج. حقق LightGBM أعلى دقة بنسبة 95.9%، يليه XGBoost بنسبة 95.5%. أظهر تحليل أهمية الميزات وتحليل SHAP أن التسارعات، والتباين القصير المدى غير الطبيعي، ونسبة الوقت مع التباين الطويل المدى غير الطبيعي هي ميزات حاسمة للتنبؤ الدقيق. يوضح دراستنا أن دمج التعلم الآلي مع XAI يمكن أن يحسن بشكل كبير مراقبة صحة الجنين من خلال تقديم نماذج دقيقة وقابلة للتفسير. تدعم هذه الطريقة اتخاذ القرارات السريرية الأفضل وتتماشى مع أهداف التنمية المستدامة في تقليل الوفيات الأمومية والطفولية.

الكلمات المفتاحية: تخطيط قلب الجنين، النموذج المتجانس، صحة الجنين، تقنيةSMOTE ، الذكاء الاصطناعي القابل للتفسير(XAI)