Fetal Health Prediction Using Machine Learning

H.K. Gayani V.L. Wickramarathna
Department of Computer Science and Engineering
University of Moratuwa
Colombo, Sri Lanka
gayani.20@cse.mrt.ac.lk

Dewmini C. Bulegodaaarachchi
Department of Computer Engineering
University of Ruhuna
Galle, Sri Lanka
dewmini.bulegodaarachchi@gmail.com

Keywords—Fetal health, Machine learning, K-Nearest Neighbour, Random Forest, Classification.

I. Introduction

Fetal health assessment plays a crucial role in prenatal care, as early detection of fetal distress can help prevent complications and ensure better outcomes for both the mother and baby. Traditionally, this assessment relies on cardiotocography (CTG) data, which medical professionals analyze to determine fetal well-being. However, this process can be subjective and influenced by individual expertise. Machine learning presents a valuable opportunity to enhance this assessment by automating classification, reducing variability, and improving diagnostic accuracy. By recognizing subtle patterns in CTG data that may be difficult for the human eye to detect, these models can provide more consistent and reliable evaluations. This paper explores the application of various machine learning models for fetal health classification, comparing their effectiveness in identifying normal, suspected, and pathological cases. Additionally, we discuss key aspects such as feature selection, model interpretability, and the potential for integrating these approaches into clinical practice to support healthcare professionals in making informed decisions..

II. Literature Review

Machine learning has significantly improved fetal health prediction by automating the analysis of fetal heart rate (FHR) and cardiotocographic (CTG) data. Spilka et al. [1] utilized neural networks to analyze FHR signals, demonstrating their potential in detecting abnormalities. Ensemble learning techniques, as explored by Kavitha et al. [2], further enhanced classification accuracy by combining multiple models. Recent studies, such as Patel et al. [3], introduced reinforcement learning for optimizing fetal health classification, improving adaptability. Deep learning approaches have also been employed for fetal health monitoring, with studies focusing on ultrasound-based gestational age estimation [4] and computational intelligence models for intrauterine growth restriction prediction [5]. Additionally, Kumar and Singh [6] demonstrated that ensemble techniques improve classification performance. These studies highlight the effectiveness of machine learning in fetal health assessment, supporting the integration of hybrid and deep learning models for enhanced predictive accuracy.

III. Methodology

The methodology followed in this study consists of several key steps, including data preprocessing, feature selection, model training, and evaluation.

A. Dataset Description

The dataset used in this study comprises 21 clinical features obtained from cardiotocographic (CTG) measurements. These features include fetal heart rate (FHR) baseline, accelerations, decelerations, uterine contractions, and other signal-derived characteristics. The dataset labels classify each case into three categories: Normal, Suspect, and Pathological.

B. Data Preprocessing

Preprocessing is an essential step to improve model performance and includes:

- **Outlier Handling:** Identified via box plots and Zscore analysis; addressed through removal or winsorization.
- **Normalization:** Min-Max scaling applied for feature uniformity and model convergence.
- **Feature Engineering:** Derived relevant features and removed redundancies to improve efficiency.
- **Class Balancing:** Addressed imbalance using SMOTE to prevent majority class bias.
- **Data Preparation:** Split into training/testing sets; categorical features encoded via One-Hot Encoding.Encoding to make them suitable for machine learning models.

C. Feature Selection

To improve efficiency and prevent overfitting, feature selection techniques were employed. A correlation threshold of 0.7 was set to identify and remove highly correlated features, mitigating multicollinearity. The correlation matrix was analyzed to detect feature pairs exceeding this threshold, leading to the removal of redundant features such as histogram_width, baseline value, and histogram_mean. Additionally, Recursive Feature Elimination (RFE) was applied to further refine the selection, ensuring that only the most relevant 14 features were retained. This process enhances the model's interpretability and performance by reducing redundancy while preserving essential predictive information.

https://doi.org/10.31705/ADScAI.2025.50

D. Model Selection and Training

Several machine learning models were implemented and tested:

- **Logistic Regression (LR):** A baseline model used for initial evaluation.
- **Support Vector Machines (SVM): Applied** to separate the classes with a hyperplane in a highdimensional space.
- **Random Forest (RF):** A decision tree-based ensemble model used for improved accuracy and robustness.
- **Gradient Boosting (GB):** A boosting technique combining weak classifiers to enhance performance.
- **Neural Networks (NN)**: A deep learning approach for complex pattern recognition in CTG data.

Each model was trained using a stratified 10-fold cross-validation approach to ensure generalization. Hyperparameter tuning was performed using grid search and random search techniques to optimize model performance.

E. Performance Evaluation

Model performance was assessed using multiple evaluation metrics:

- **Accuracy:** Measures overall correctness of the model.
- **Precision, Recall, and F1-score:** Provide insights into class-specific performance.

IV. Results and Discussion

A. Model Performance

Two machine learning models were evaluated: Random Forest Classifier and K-Nearest Neighbors (KNN). The performance comparison is summarized in Table I.

TABLE I. OVERALL PERFORMANCE COMPARISON

Metric	Random Forest	KNN
Test Accuracy	0.95	0.93
Precision	0.95	0.93
Recall	0.95	0.93
F1 Score	0.95	0.93
Cross-validation Accuracy	0.94(±0.01)	0.89

B. Confusion Matrix Analysis

The classification performance for each fetal health category is detailed in Table II.

TABLE II. CONFUSION MATRIX PERFORMANCE

CLASS	RANDOM FOREST		KNN	
	PRECISION	RECALL	PRECISION	RECALL
NORMAL(1.0)	96%	99%	95%	96%
SUSPECT(2.0)	93%	78%	82%	77%
PATHOLOGICAL(3.0)	93%	90%	90%	89%

C. Comparison and Insights

The Random Forest model outperformed KNN, achieving a higher overall accuracy and more balanced recall across all classes. It showed particularly strong performance in correctly classifying normal and pathological cases, though recall for suspect cases (78%) was slightly lower. The Random Forest model performed well but had still lower recall for suspect cases.

D. Novelty and Contribution

This research introduces a comparative analysis of machine learning models for fetal health classification, emphasizing the effectiveness of ensemble-based approaches over traditional distance-based classifiers. The study highlights the importance of recall for suspect cases, suggesting potential improvements using hybrid models or additional feature engineering to enhance classification performance in future work.

V. Conclusion

Machine learning significantly enhances fetal health prediction accuracy. Our study highlights the importance of feature selection and model tuning in achieving optimal results. Future research should explore hybrid models integrating multiple approaches for better interpretability and reliability. Additionally, integrating these models into clinical workflows remains a crucial step toward real-world applicability. This includes developing automated decision support systems that seamlessly assist obstetricians in risk assessment. Further studies should also focus on real-time implementation and validation in diverse clinical settings to ensure robustness and generalizability.

Acknowledgement

We extend our gratitude to our mentor, Ms.Yugani Gamlath for her guidance, support, and insights throughout this project. We also express our appreciation to all who contributed directly or indirectly to this work. This research was carried out independently without the support of external funding.

References

[1] J. Spilka et al., “Neural network analysis of fetal heart rate signals,” 2012.

[2] K. Kavitha et al., “Ensemble learning approaches for fetal health classification,” 2021.

[3] R. Patel et al., “Reinforcement learning optimization in fetal health classification,” 2023.

[4] L. H. Lee et al., “Machine learning for accurate estimation of fetal gestational age based on ultrasound images,” 2023.

[5] M. Aslam et al., “Explainable computational intelligence model for predicting intrauterine growth restriction,” 2022.

[6] P. Kumar and M. Singh, “Predicting fetal health status using ensemble machine learning techniques,” 2020.

[7] Ayres de Campos et al. SisPorto 2.0 A Program for Automated Analysis of Cardiotocograms. J Matern Fetal Med 5:311-318, 2000.