Cardiotocogram Data Analysis for **Fetal Health Classification** Using Machine Learning

Presented By

Ashraf Ul Alam

Table of contents

Introduction

- Monitoring fetal health is essential for ensuring the well-being of unborn infants and successful delivery outcomes.

- Machine learning analysis of cardiotocogram (CTG) data enables healthcare providers to conduct more accurate, objective, and timely assessments.

- This application of machine learning significantly enhances prenatal care by facilitating better decision-making and intervention strategies.

Introduction (to CTG) Cntd'

- CTG is a crucial tool for monitoring fetal heart rate (FHR) and uterine contractions during pregnancy.
- Provides essential data for evaluating fetal health.
- Thus, helps in early detection of potential complications.

Research Proposal

- Develop machine learning algorithms tailored for fetal health classification using cardiotocogram (CTG) data.

- Evaluate the performance of different machine learning models, such as Random Forest, K Nearest Neighbors, and Gradient Boosting, in accurately categorizing fetal health status.

- Investigate the impact of preprocessing techniques, feature selection, and algorithm tuning on classification accuracy and efficiency.

Problem Statement

Despite advancements in healthcare, perinatal mortality rates remain high, especially in regions like South Asia and Sub-Saharan Africa.

There is a need for more efficient and accurate methods to classify fetal health, allowing for timely interventions and personalized care.

Traditional methods of assessing fetal health, such as manual analysis of CTG) are time-consuming and subject to human error.

Aim of Research

To leverage machine learning algorithms for analyzing CTG data to accurately classify fetal health and detect potential issues.

To evaluate the performance of various machine learning models, including Random Forest, K Nearest Neighbors, and Gradient Boosting, in fetal health classification.

To improve prediction accuracy and provide insights into the scalability and applicability of machine learning models for assessing fetal health.

Context, Boundaries, and Assumptions

Context

This research focuses on utilizing machine learning techniques to enhance fetal health classification based on CTG data, aiming to address the persisting issue of perinatal mortality rates.

Boundaries

The study will primarily focus on analyzing CTG data and evaluating machine learning algorithms' performance. Other factors influencing fetal health, such as maternal health conditions, will be considered but not directly analyzed in this research.

Assumptions

The research assumes the availability of sufficient and reliable CTG data for analysis. Additionally, it assumes that machine learning algorithms can effectively learn patterns from CTG data to classify fetal health accurately.

Literature Review

2017

J. Spilka et al. explore identifying acidosis during childbirth by analyzing intrapartum FHR variability. They use scales and wavelet leader analysis, apply Sparse SVM for classification, and monitor feature space trajectories. Their dataset comprises intrapartum FHR records from a public hospital in France.

2018

M. Ramla et al. introduce a predictive approach for high-risk pregnancies based on fetal health using the CART algorithm. Results indicate CART's promise with an accuracy rate of 88.87% using entropy calculation and 90.12% using the Gini index method.

2019

K. Agrawal et al. focus on classifying health based on CTG data using machine learning techniques like Decision Tree, SVM, and Naïve Bayes. Their study utilizes a dataset from the UCI Machine Learning Repository, revealing high accuracies: 93.17% for Decision Tree, 92.84% for SVM, and 83.65% for Naïve Bayes.

2021

N. F. M. Noor et al. employ supervised machine learning methods on CTG data for health monitoring, achieving high accuracy. Ada Boost combined with a submodel of Random Forest stands out, reaching 94.7% accuracy when k=10. Additionally, J. Li and colleagues test twelve machine learning models on CTG data and propose a voting integration technique to merge top-performing models into a Blender Model, outperforming conventional stacking methods.

2023

K. Singh et al. a method combining ultrasound images and clinical factors to determine fetal health status, aiming to equip obstetricians with tools for improved care. They plan to gather a dataset containing ultrasound images and clinical information to train and test their model. Particularly IBM Watson and Python-based models, to forecast health outcomes from cardiotocograph recordings, achieving a success rate of 95.3% with CatBoostClassifier.

Research Workflow

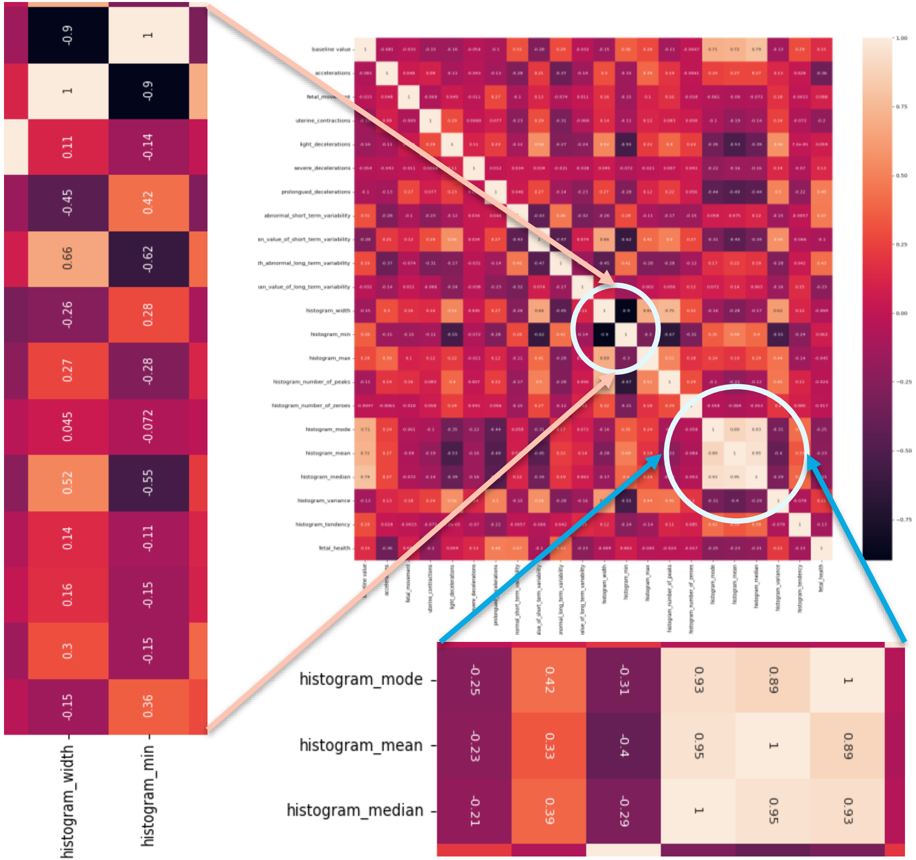
1. Data Collection
2. Data Preprocessing
3. Model Training
4. Result Analysis

Dataset Description

The CTG dataset comprises 2126 records, where each record consists of 21 features extracted from CTG exams and one feature used for classification, annotated by three expert obstetricians into one of three distinct classes:

- Normal
- Suspect
- Pathological

Data Preprocessing



Assess Correlation

Use a correlation matrix or heatmap to identify and address highly correlated features that may indicate redundancy and affect model performance.

Remove Highly Correlated Features

Drop highly correlated or unnecessary features. For example, 'histogram_median' and 'histogram_min' were removed based on correlation analysis.

Data Preprocessing

Feature Scaling

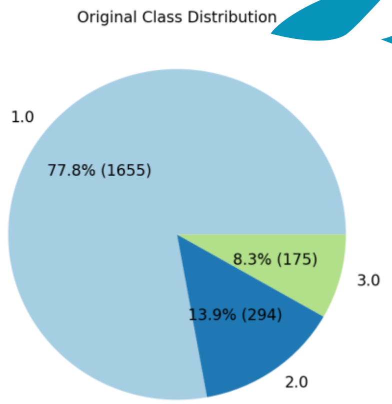
Standardize numerical features to have a mean of 0 and a standard deviation of 1 to prevent certain features from dominating the model training.

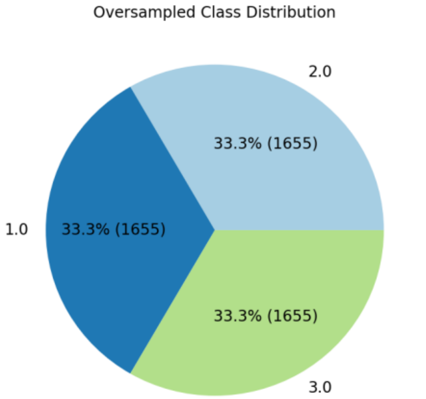
Handling Imbalanced Data

To address the imbalance in the dataset, Synthetic Minority Over-sampling Technique (SMOTE) was employed. SMOTE generates synthetic samples for the minority class, effectively balancing the distribution of classes and enhancing the model's ability to learn from the data.

Data Partitioning

The dataset was partitioned into training and testing sets using the train_test_split function from scikit-learn. The split was performed with a ratio of 75% for training data and 25% for testing data.





Model Implementation

01

Random Forest

No. of Estimators = 100

02

K-Nearest Neighbors

No. of Neighbors = 5

03

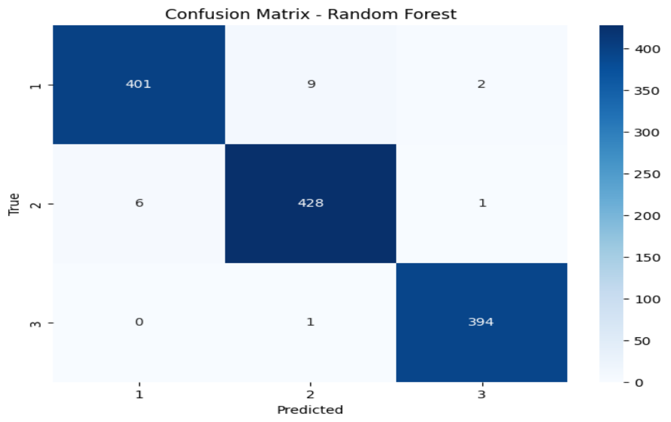
Gradient Boosting

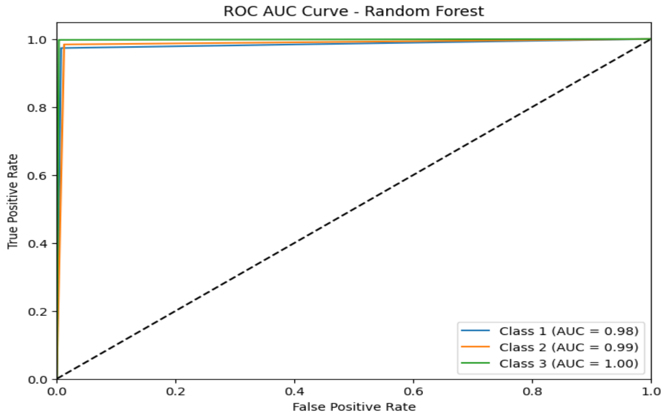
No. of Estimators = 100

Result Analysis

1. Random Forest

- Accuracy: 98.47%
- Precision: 98.50%
- Recall: 98.49%
- F1-score: 98.49%

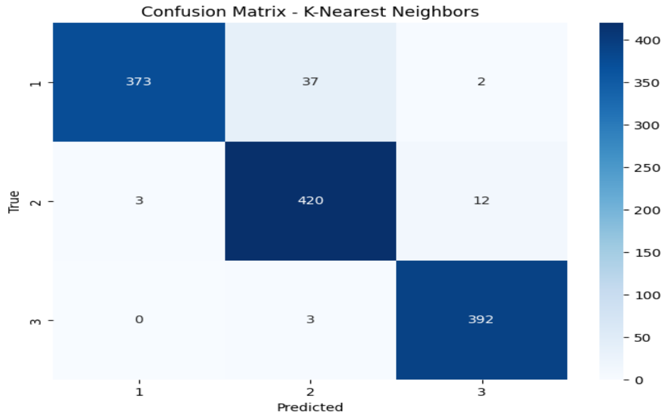


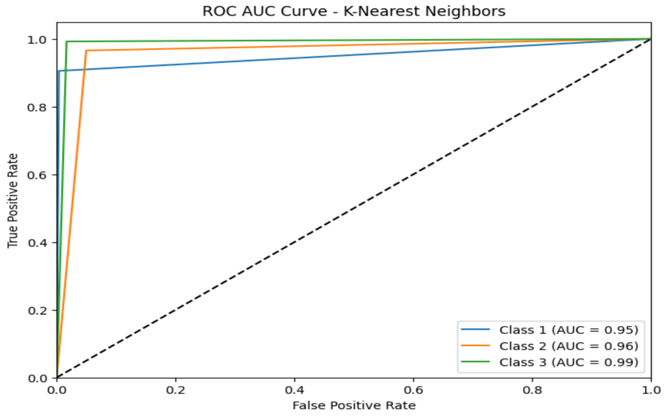


Result Analysis

2. K-Nearest Neighbors

- Accuracy: 95.41%
- Precision: 95.69%
- Recall: 95.44%
- F1-score: 95.47%

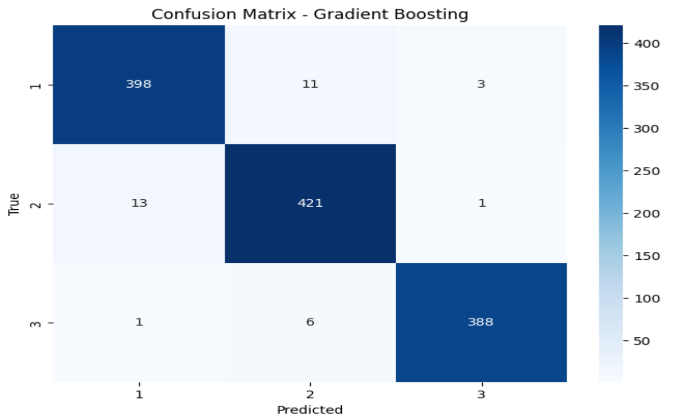


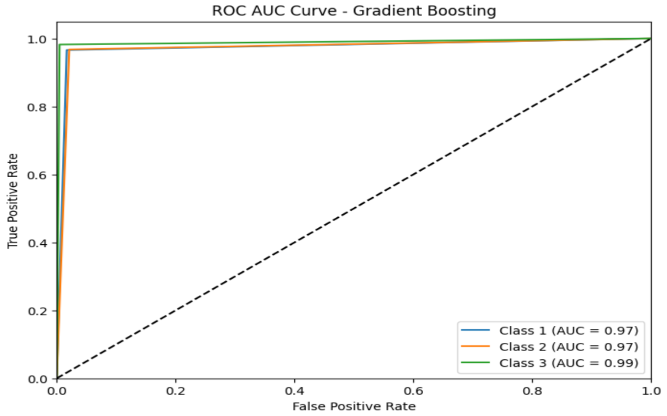


Result Analysis

3. Gradient Boosting

- Accuracy: 97.18%
- Precision: 97.23%
- Recall: 97.20%
- F1-score: 97.22%





Conclusion

- Developed an effective system for fetal health classification using machine learning.
- Algorithms like Random Forest, K-Nearest Neighbors, and Gradient Boosting were employed.
- Particularly Random Forest showed significant potential with an accuracy of 98.47%.
- Strides towards personalized and data-driven prenatal care for better health outcomes.

References

- J. Spilka, J. Frecon, R. Leonarduzzi, N. Pustelnik, P. Abry and M. Doret, "Sparse Support Vector Machine for Intrapartum Fetal Heart Rate Classification," in *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 664-671, May 2017, doi: 10.1109/JBHI.2016.2546312.
- M. Ramla, S. Sangeetha and S. Nickolas, "Fetal Health State Monitoring Using Decision Tree Classifier from Cardiotocography Measurements," *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2018, pp. 1799-1803, doi: 10.1109/ICCONS.2018.8663047.
- K. Agrawal and H. Mohan, "Cardiotocography Analysis for Fetal State Classification Using Machine Learning Algorithms," *2019 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2019, pp. 1-6, doi: 10.1109/ICCCI.2019.8822218.
- N. F. M. Noor, N. Ahmad and N. M. Noor, "Fetal Health Classification Using Supervised Learning Approach," *2021 IEEE National Biomedical Engineering Conference (NBEC)*, Kuala Lumpur, Malaysia, 2021, pp. 36-41, doi: 10.1109/NBEC53282.2021.9618748.
- K. Singh, P. Shyry and R. G. Franklin, "Efficient Fetal Health Monitoring and Classification with Machine Learning," *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2023, pp. 355-359, doi: 10.1109/ICICCS56967.2023.10142312.

Thank You