MLDA Lifeline 2025 Academic Paper

Cardiotocography (CTG) interpretation is time-consuming and inconsistent among doctors, studies proposed automating foetal health classification using traditional machine-learning models on the public feature sets such as baseline, variability, accelerations/decelerations and histogram stats to provide an extra pair of eyes during labor. Logistic Regression, SVM and Random Forest, Decision Trees, KNN, and simple voting ensembles were trained on standard CTG tabular dataset and are reported as baselines. We utilised Multinomial Logistic Regression, SVM-RBF, Random Forest and XGBoost with the same feature families and labels (NSP 1/2/3) - Normal, Suspect or Pathologic.

The UCI Cardiotocography dataset, including more than 2,000 CTG recordings with medical labels, was used. Medical characteristics such as baseline foetal heart rate (LB), accelerations (AC), decelerations (DL, DS, DP), and variability markers (ASTV, ALTV, MSTV) are included in every record. Short duration measurements (less than 10 minutes) and incomplete or duplicate rows were eliminated as part of the data cleaning process. We numerically encoded the output labels (NSP) as 0 = Normal, 1 = Suspect and 2 = Pathologic after standardising feature scales with StandardScaler. We also included weighted training to address class imbalance and ensured minority cases were fairly represented.

To maintain class proportions, an 80/20 stratified split was used to train all machine-learning models. With the highest balanced accuracy and macro-F1 score across all classes, XGBoost produced the most robust and consistent results. Additionally, it provided interpretable feature importance ratings and was effective in training, both of which are critical in a medical setting.

The results of feature analysis showed that DS, LB and ASTV were important indicators of foetal distress. Neural networks were able to record more complex temporal correlations, however, judgments were harder to understand. Due to its outstanding performance and result, we decided on XGBoost as our model. With further validation and visualization tools, this framework could one day help healthcare professionals detect fetal distress earlier and improve delivery outcomes.