**Fetal Health Classification and predicting the Birth-Weight of Fetus**

**ABSTRACT**

This work explores the critical area of fetal health prediction and birth weight estimation using machine learning techniques that focus on the analysis of cardiotocography (CTG) data. Given the growing importance of pregnancy and childbirth, this effort highlights the transformative journey of motherhood in parallel with the development of medical technology. Fetal health prediction and birth weight should be calculated. Using sophisticated algorithms such as CatBoost Regressor, SGD Regressor, LightGBM Regressor and XGBoost Regressor, the model achieves remarkable accuracy in birth weight prediction and provides valuable insights into fetal growth and potential risk factors for delivery. Moreover, in terms of normal, suspicious and pathological working conditions, a robust majority voting prediction algorithm is built using a combination of DNN, ANN, CNN to assign the final labels, and the model can obtain macro-impressive accuracy scores on training and testing datasets. . Millions of babies die within the first 24 hours after birth and with alarming maternal mortality rates, prenatal care is essential. This effort will focus on CTG, the most common method of adoption, so keep an eye on fetal well-being. By analyzing the patterns captured in the data, the model effectively differentiates fetal health conditions, enabling early medical intervention whenever needed. This project represents a major advance in the application of machine learning in prenatal care.

**Keywords:** *Cardiotocography (CTG), Heart Rate of fetal, Gestational age, Low birth weight of fetus, Fetal Health, Machine learning, Prenatal Care.*

**I. INTRODUCTION**

A healthy pregnancy culminates in the joyful arrival of a newborn. However, the primary concern remains ensuring a safe journey for both mother and child during pregnancy. Here, advances in medical technology offer key tools for monitoring fetal health and predicting birth weight.The fetal health classification uses a variety of techniques to assess the well-being of the developing baby. This may include analyzing data from prenatal tests such as ultrasound scans and cardiotocography (CTG), which measures fetal heart rate and uterine contractions. By leveraging machine learning algorithms, healthcare professionals can gain valuable insights into potential risks, allowing for early intervention and better pregnancy outcomes.

Predicting birth weight is another important aspect of prenatal care. An accurate estimate of the baby's size allows doctors to identify potential complications associated with low birth weight or macrosomia (large birth weight). This information allows them to tailor birth plans and ensure a safe birth for both mother and baby.

Fetal health classification involves assessing the condition of the fetus based on various parameters such as fetal heart rate, fetal movements, amniotic fluid volume, and biophysical profile. These assessments help healthcare providers identify any signs of distress or abnormalities that may indicate underlying health issues in the fetus. Timely recognition and classification of fetal health can guide medical interventions and management strategies to optimize outcomes for both the mother and the baby.

One of the key components of fetal health classification is the evaluation of fetal growth and development, which includes monitoring the baby's size and weight throughout pregnancy. Birth weight prediction plays a vital role in determining the likelihood of adverse perinatal outcomes, such as preterm birth, low birth weight, or macrosomia (excessive birth weight). By accurately estimating the baby's weight before birth, healthcare providers can anticipate potential delivery complications and tailor obstetric care accordingly.These non-invasive techniques allow for detailed visualization of fetal anatomy and physiology, enabling healthcare providers to detect abnormalities early and provide timely interventions.

Moreover, the integration of deep learning and machine learning algorithms has revolutionized fetal health assessment by facilitating automated analysis of CTG data and clinical data. These computational tools can analyze complex patterns and biomarkers to provide predictive models for fetal health outcomes, enabling personalized risk assessment and intervention strategies.This survey delves deeper into the methods and benefits of fetal health classification and birth weight prediction. We'll explore the types of data used, the machine learning techniques used, and the impact these improvements can have on ensuring a healthy pregnancy and smooth birth.

**II.LITERATURE SURVEY**

Design a robust ensemble model called ETSE (Ensemble of Tuned Support Vector Machine and Extratrees) to predict fetal health analyze the performance of the proposed ETSE model and compare it with other models The model has not been validated on other datasets, which may affect its robustness and reliability. Tuning hyperparameters using techniques such as grid search can optimize machine learning models and improve their performance The KNN (K- Nearest Neighbors) model achieved the lowest accuracy of 90%. The decision tree model was reported to be 96%The proposed ensemble model, ETSE, which achieved 99.66% accuracy, 100% precision, and 100% recall The study does not mention the specific dataset used for training and testing the models, which could affect the generalizability of the results. The study does not provide information on the size of the data set used, which could affect the performance and reliability of the models.[1]

To develop a model for predicting the health status of fetuses during pregnancy, speci cally to categorize them into three ranges: normal, risk, and abnormal conditions those pregnant women. This model easily predicts the state of the fetus. The absence of this information limits the transparency and generalizability of study results, raising concerns about the reliability of predictive models. Ambiguous Model Evaluation Metrics Uses supervised learning algorithms such as decision tree, gradient boosting, random forest, SVM, k-NN, AdaBoost, and stochastic gradient descent to obtain an accurate decision tree prediction model for fetal health.F1-score-0.918, Accuracy-0.91, Recall - 0.922AUC-0.84, Transition Boost: F1-score-0.935, Accuracy-0.935, Recall - 0.938AUC-0.965, Random forest: F1-score-0.939, Accuracy-0.940, Recall - 0.942AUC - 0.966AdaBoost: F1-score-0.935, Accuracy-0.934, Recall - 0.937AUC-0.930. There is a gap in the explicit comparison of the proposed approach with existing studies on the prediction of pregnancy risk and fetal health, which limits a comprehensive understanding of the novelty and effectiveness of the presented model.[2]

To design an ensemble learning-based predictive method for classifying fetal health using a cardiotocographic ensemble of fetal movements and fetal heart rate accelerations from nonstress tests (NSTs) which can help healthcare professionals take early preventive measures and interventions for maternal and fetal well-being The paper does not provide information on the size and diversity of the data set used, which could affect the generalizability of the resultsThe proposed predictive method based on ensemble learning achieved an accuracy level above 99.5% on the test data set, indicating its effectiveness in classifying fetal health.XGBoost: Accuracy -0.99, Accuracy-0.99Appeal-1F1- Score-0.99Transition Gain (GB): Accuracy -0.99Accuracy-1Appeal-1F1-Score- 0.99.Extreme Learning Machine (ELM) Algorithm Accuracy - 99.29 The article does not provide information on the specific features or variables used in the dataset to classify fetal health. The article does not deal with the interpretability of the ensemble learning model and how the predictions are made.[3]

Classify cardiotocography (CTG) data into three categories: healthy, suspicious, and pathological using ensemble learning techniques. The lack of detailed examination of the interpretability of the model raises concerns about the transparency of decisions in the critical medical context, which hinders the practical adoption of the classi cation system. Small dataset size (3602 CTG readings), potentially affecting the generalizability of the proposed model A soft voting classifier is used to combine the results of individual classi fiers, eliminating the weakness of one classifier and providing better performance of the overall model. Using clustered classifier in a balanced CTG dataset improves the detection accuracy of the Random Forest(RF) Classifier: Accuracy: 0.98Accuracy: 0.96, Recall: 0.94, F1-Score: 0.98XGBoostClassi er: Accuracy:0.99, Accuracy:0.98, Recall:0.98, F1Score:0.99,AdaBoostClassifier: Accuracy: Accuracy:0.84, Recall:0.84, F1Score:0.84, CatBoost Classifier: Accuracy:0.99Accuracy:0.98Recall:0.99, F1- Score:0.98.Lacking is an in-depth examination of the interpretability of the model and a discussion of potential biases in the data set, which are key aspects in the context of medical applications. There is a gap in the absence of validation on external datasets, which limits the generalizability of the proposed classification approach to different populations or clinical settings.[4]

Predict fetal health using machine learning models. Evaluate CTG (continuous cardiotocography) to analyze fetal health in utero and assess heart rate and overall fetal health during pregnancy Lack of detailed information on speci c methodology used for CTG monitoring and machine learning algorithms used to detect fetal abnormalities. Continuous cardiotocography (CTG) monitoring provides valuable information about fetal health in utero,specifically assessing heart rate and overall fetal well-being during pregnancy Random Forest Classi er: Accuracy 91.62, Support Vector Machine (SVM): Accuracy - 85.45%, F1-score-86.66%, Download - 85.45%, Accuracy - 89.67%, Extra tree model: Accuracy - 93.66%, F1-score-93.71%, Download - 93.66%, Accuracy - 93.78%. Exploring the use of additional machine learning algorithms or hybrid approaches to improve classification performance for fetal risk prediction.Investigating the impact of additional features or data sources on the accuracy and reliability of the prediction model.[5]

A deep neural network model called CTG-net for cardiotocogram (CTG) classification to detect fetal distress. It focused only on quantitative and biased algorithms for the evaluation of the cardiotocogram (CTG) but did not take into account other factors that can affect the state of the fetus. CTG-net is a quantitative and automated diagnostic aid system that can enable early intervention in potentially abnormal fetuses to reduce the risk of hypoxic injury. K stands for clustering. Deep Neural Network, F1 score 0.67, Accuracy on Roc- Auc 0.68. The development of a deep neural network (CTG-net) model for classifying abnormal and normal cardiotocogram (CTG) data, however, did not address other potential factors that may in uence fetal status, such as maternal health conditions or other physiological parameters.[6]

To develop a computer-aided diagnosis (CAD) system based on deep learning technology to predict fetal acidemia using fetal heart rate (FHR) signals. The authors propose an eight-layer deep convolutional neural network (CNN) framework that can automatically predict fetal acidemia without the need for complex feature engineering. The study did not address potential biases in the prediction models. The deep learning approach of the CNN model enables self- learning of useful features from the input data, which represents a significant advantage over conventional machine learning approaches. the CNN modelaccuracy:98.34%, sensitivity:98.22%,speci city:94.87%.quality index:96.53%and the area under the curve value97.82% Future research could focus on optimizing the architecture of the CNN model by experimenting with different conjurations of layers, lters, and hyperparameters to achieve even better classification.[7]

The paper aims to reduce fetal mortality by developing a Convolutional Neural Network (CNN)--b-based fetal health classification information system and a hybrid CNN with dimensionality reduction using Principal Component Analysis (PCA). It does not provide information about the size and diversity of the data set used for training and testing the models. Achieved approximately 95% accuracy when integrating CNN with PCA for dimensionality reduction. The use of CNN and PCA helps in accurate fetal heart rate detection and fetal health classification. The CNN model accuracy is 98%, F1-score 0.88, and RF (Random Forest) 77.8%The work does not compare the performance of the proposed method with other existing methods for fetal health Classification Comparative studies should be conducted to evaluate the proposed method against other existing methods for fetal health classification, taking into account various performance metrics such as precision, recall, and F1 score.[8]

CTG support during pregnancy, along with intelligent remote monitoring, can improve fetal health outcomes and healthcare availability amid logistical and pandemic challenges. The potential lack of addressing other factors in uencing access to prenatal care outside of traffic and office hours, as well as the need for further investigation into the effectiveness and feasibility of implementing CTG in different healthcare settings.CTG helps pregnant women overcome barriers to access and ensures fetal monitoring despite barriers. The article highlights the use of chi-square and other feature selection methods to simplify classifier models by identifying key predictions. The paper evaluates the degree of classifi cation accuracy of the proposed method, which is stated as 91.62% random forest (99.2%), SVM (98.42%), and (98.9), this is the highest accuracy The paper does not mention any future directions, or. potential areas of improvement of the proposed method and leave room for further research and development.[9]

The paper aims to compare and analyze the performance of different machine learning methods for fetal health classi cation paper aims to assess the performance of decision trees, logistic regression, and KNN classifiers in fetal health classification. The contribution does not provide information about potential conficts of interest that could affect the objectivity of the research findings. The article lacks detailed information about the method used and the methodology of the experiment, which may limit the reproducibility of the study, and uses performance metrics such as precision, precision-recall trade- off, and AUC to evaluate the performance of models. It is also mentioned that the KNN (K-Nearest Neighbors) model achieved the lowest accuracy of 90%.The accuracy rate of the decision tree model was reported as 96% in this paper. The lack of specific data set details in the paper prevents an assessment of the generalizability of the findings. Appeals for evidence of manipulation of the publication process undermine reliability and integrity.[10]

Research focuses on using advanced machine learning techniques to improve early detection of potential health problems in fetuses, facilitate early interventions, and improve health care outcomes ML models can be biased, leading to inaccurate predictions for certain groups, which is a problem for fetal health. Limited labeled samples and data complexity present challenges in classifying fetal health. The comprehensive feature selection and model evaluation methodology offers a more holistic and reliable evaluation compared to traditional methods. Proposed The machine learning approach achieves an impressive 98.31% accuracy on the test set, demonstrating its effectiveness in LightGBM fetal health classification Accuracy of 98.81% and F1-score 0.9937. Precision and Recall 0.9863 Detailed analysis of the confusion matrix and its implications for model performance are not discussed in the sources.[11]

Research focuses on using UCI cardiotocography data to train and evaluate an association classification model to access fetal health status. The research aims to investigate a classification-based association (CBA) approach for cardiotocographic analysis of fetal assessment, specifically for the classification of fetal health status. The data set used for training the classification model is unbalanced, which can affect the model performance and the accuracy of the results. The association classification model showed high accuracy (83% before, 84% after feature selection) in predicting fetal health, indicating efficacy in risk assessment. Before and after feature selection, the accuracy rates were 83% and 84%. Random Forest achieved 99.18% accuracy in fetal state classification, while Zhang et al. He proposed a hybrid PCA and AdaBoost method with an accuracy of 93% and 98.6% for the total and selected features, respectively. One of the main gaps is the imbalance in the data set used to train the model. Research recognizes that the data set is unbalanced, which can affect the performance of the classification model. It is suggested that in future work, the data set should be balanced first to ensure better accuracy of the results.[12]

To use the soft voting integration method to integrate the top four models and compare it with the stacking method. Lack of information about the specific CTG data set used, such as its size, source, or characteristics. No discussion of potential clinical implications or practical applications of the proposed model. The paper evaluates the performance of twelve machine learning models on the CTG dataset and provides a comprehensive analysis of the various models' accuracy, AUC, recall, precision, F1, and MCC. Precision rate 0.959, AUC 0.988, recall rate 0.916, precision rate 0.959, F1 0.958, and MCC 0.886 The article does not provide information about the specific CTG data set used, such as its size, source, or characteristics. The article does not discuss the limitations or potential biases of the machine learning models used in the study.[13]

The study aimed to analyze antenatal cardiotocography (CTG) data and develop an efficient tree-based ensemble learning (EL) classifier model for fetal health prediction. The CTG data set used in the study is unbalanced, and although tree-based algorithms have shown higher performance on this type of data set, it is unclear how well the model would perform on other types of data sets. The study focused on the analysis of antenatal cardiotocography ( CTG ), which provides sophisticated information by monitoring the fetal heart rate signal, making them a valuable resource for predicting fetal health. fetal health risks. using cardiotocography (CTG) data The study does not discuss the limitations or potential problems of using tree ensemble learning techniques for the analysis of cardiotocography (CTG) data.[14]

Determine the most in uential features in classifying fetal health using the developed model. Develop a model using Support Vector Machine (SVM) and resampling to classify fetal health with high accuracy (99.59%. The model does not take into account other factors that can affect the health of the fetus, such as the mother's lifestyle and demographics. This may limit its effectiveness in diagnosing the risk of fetal mortality and nding appropriate treatment options. Efficiency: Interpretation of cardiotocograms (ctgs) can be time-consuming and inefficient, especially in underdeveloped areas where skilled obstetricians may be scarce. Machine learning models can automate the classification process, saving time and resources. The proposed model achieved an accuracy of 99.59% in fetal health classification using support vector machine (SVM) and resampling techniques. The paper achieves 99.59% fetal health classification accuracy using SVM and resampling, introduces FAB for model explanation, and recommends careful tuning of the number of light estimates to avoid overfitting and excessive computational cost.[15]

Develop a method for automatic determination of fetal health status using cardiotocography (CTG) and machine learning algorithms. Enhancing the diagnostic capability of medical practitioners by automating the process of CTG signal interpretation. The study focused on improving the sensitivity of the pathologic and suspicious classes, even if it meant sacrificing the accuracy of the normal class. This may result in more false positives in the general category. Automating the process of CTG signal interpretation can increase the diagnostic capacity of medical practitioners, potentially leading to earlier detection of fetal heart conditions and prevention of death. The extreme gradient boosting algorithm-based model showed an accuracy of 96.7% and an F1-score of 0.963. In the pathological class, it shows its effectiveness in detecting the health status of the fetus. The study focuses on developing machine learning models to automatically detect fetal health status using CTG data but does not specify the specific algorithms used for feature selection and model training.[16]

Create a machine learning model that can accurately predict fetal health status based on clinical test results. Reducing hospital costs and waiting times for pregnant women by providing a more efficient and convenient way to monitor fetal health. Machine learning algorithms such as random forest classifiers can accurately predict fetal health, potentially saving parents time, stress, and money. Web applications for monitoring and predicting fetal health status may have limitations in terms of accessibility and usability for different users. The proposed hybrid method combining information acquisition features and an opposition-based re y algorithm (OBFA) showed better classification performance than the existing methods and achieved the following accuracy 96.24 percent. This study does not provide information on specific laboratory tests or parameters used to predict fetal health. There is no mention of how accurate the model is predicted, so uncertainty remains about the reliability of the prediction.[17]

Predict low birth weight using machine learning models. Evaluate the diagnostic performance of predictive models for LBW. Identify critical predictors for predicting LBW. Previous studies have had limitations such as small sample size or feature selection. More research is needed to conclude the LBW prediction performance of ML models. LBW prediction, a preventive tool for newborn health risks Machine learning models such as XGBoost show high diagnostic performance. Maternal age, education, and residence factors associated with LBW extreme gradient amplification model: precision 0.79, precision 0.87, recall 0.69. Extreme gradient gain model: best for predicting LBW with high accuracy. Critical predictors: gestational age, and previous history of LBW. Sample sizes: 12,500 participants, SVM accuracy 0.93. Data source: Iranian Maternal and Newborn Network (IMaN Net).[18]

Create a data set of 4212 clinical records to determine fetal weight. Developing an ensemble machine learning model to accurately predict fetal weight. Evaluate the performance of the model using real test data and propose a new benchmark. Because of population differences, regression models are not universal. strict sonographer requirements and equipment standards for ultrasound. obstetric ultrasound screening in less prosperous rural areas. Compared to the ultrasound method, the accuracy of the assessment increased by 12%. With ensemble machine learning, the average relative error is reduced by 3%. Accuracy increased by 12%, the mean error was reduced by 3%, and the novel intercept-unity (loU) estimation index showed the ensemble model loU 0.64 for predicting fetal weight, while ultrasound increased fetal weight by 12% using the ensemble model. Machine learning improves the accuracy of fetal weight estimation.[19]

Review machine learning models for fetal weight prediction accuracy. Evaluate the impact of machine learning on maternal and fetal health. The study did not address potential biases in the prediction models. Machine learning in studies predicts fetal weight accurately above 60%. Machine learning helps pregnant women track their weight effectively on their own. Machine learning models achieved 89-99% accuracy in predicting fetal health. XGB regression predicted baby weight with 42% accuracy. XG Boost, Light GBM, and Random Forest le model improved results. The research rules out non- machine learning approaches in fetal weight prediction. Limited discussion of the challenges sonographers face in accessing ultrasound. Convolutional neural networks may not perform as well as other methods.[20]

Predict low birth weight using machine learning algorithms. Identify critical predictors of low birth weight in Ethiopia. Compare classifiers for LBW prediction accuracy in Ethiopia. Limited to data from Ethiopia Demographic and Health Survey 2016. Predicts low birth weight accurately with 91.60% accuracy. Identifies top predictors: child's gender, marriage interval, and mother's age. Utilizes Random Forest classifier for effective LBW prediction. RF predicts LBW with 91.60% accuracy, 91.60% Recall, and 96.80% ROC-AUC. RF has a 91.60% F1 Score, 1.05% Hamming loss, and 81.86% Jaccard score. The study identified critical predictors of low birth weight in Ethiopia. Extreme Gradient Boosting (XGB) efficiently handles vast amounts of medical data. Random Forest (RF) was the best classifier for predicting LBW. The gender of the child and the mother's occupation are top predictors.[21]

Predict late IUGR using machine learning classification algorithms. Assess late IUGR using CTG findings with machine learning. Machine learning models show high accuracy, precision, recall, and F-score. Eliminating features increases model performance and reduces reassembly risks. The feature importance method identifies the main relevant features for prediction. Accuracy rate 84-85%. Accuracy 79-80%. Recall at 85-89%. Late prediction of IUGR using machine learning models with high accuracy. Models include Logistic Regression and Support Vector Machine. Feature elimination is essential to improve model performance.[22]

To compare four machine learning classifiers (decision tree, random forest, artificial neural network, support vector machine, and logistic regression) in predicting low birth weight (LBW) in Hamadan, Iran. The study was conducted in a specific area (Hamadan, Iran), which may limit the generalizability of the findings to other populations. The study uses machine learning classifiers to predict low birth weight, which can provide more accurate and efficient predictions compared to traditional statistical methods. The average accuracy of all models was 87% or higher. The LR method provided a sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and accuracy of 74%, 89%, 7.04%, 29%, and 88%, respectively. The study was conducted in a specific area based on retrospective data, which limits the generalizability of findings to other populations and potentially introduces bias.[23]

The main objective of this paper is to use machine learning (ML) techniques to predict fetal birth weight during early pregnancy and classify it into three categories: low, normal, and abnormal. Despite the promising results and progress presented in the paper, some limitations should be considered. First, the performance of the proposed ML model can vary depending on the quality and quantity of data used for training. This paper highlights several advantages of using ML methods to predict fetal birth weight. The performance of the proposed ML model is evaluated by metrics such as accuracy, precision, recall, and F1 score. These metrics assess the ability to correctly classify fetal birth weight into certain categories (low, normal, abnormal), and although the paper presents a promising approach to predicting fetal birth weight, there are important gaps that require further research. First, research can bene t from a more comprehensive analysis of factors that affect fetal development and their effects.[24]

The study aims to develop a method for accurate prediction of fetal birth weight using fetal ultrasound video scans and clinical data. The study is aware of the difficulty of obtaining high-quality ultrasound images during advanced pregnancy due to the lack of amniotic uid. The proposed method outperforms several state-of-the-art automated methods and estimates fetal birth weight with an accuracy comparable to human experts. The mean absolute error (MAE) for the head is reported as 218 grams with a standard deviation of 19 grams, while the root mean square error (RMSE) is 262 grams. Similarly, for the femur, the MAE is 204 grams with a standard deviation of 28 grams and the RMSE is 250 grams. The thesis does not explicitly state any research gaps.[25]

The study aimed to analyze the main predictors of adverse birth outcomes in very low birth weight (VLBW) infants, including the concentration of particulate matter (PM10). The study did not analyze possible mediating effects between predictors. The study adopted binary categories of adverse birth outcomes, which could be expanded to more categories for more clinical findings. The study used machine learning and a national prospective cohort database to provide a comprehensive investigation of predictors of adverse birth outcomes in VLBW infants. The random forest model performed best among the six prediction models, with an accuracy of 0.79 and an area under the receiver-operating characteristic curve of 0.72The study did not analyze possible mediating effects between predictors, which could provide additional insight into the relationships between predictors and adverse birth outcomes.[26]

The objective of the study was to construct a hybrid birth weight predictor classifier based on long short-term memory (LSTM) network using multiple electronic medical records and B-ultrasonic examination data of pregnant women. The study only used data from a single hospital in eastern China, which may limit the generalizability of the findings. The proposed hybrid-LSTM model showed a higher accuracy rate and empirical formula for predicting birth weight compared to other machine learning models. Back Propagation Neural Network (BPNN), and proposed hybrid-LSTM at the 40th gestational week for first delivery were 0.498, 0.662, 0.670, 0.680, 0.705, and 0.793, respectively. This paper does not explicitly address any research gaps. However, one potential gap may be the lack of external validation of the hybrid-LSTM model using data from different hospitals or regions.[27]

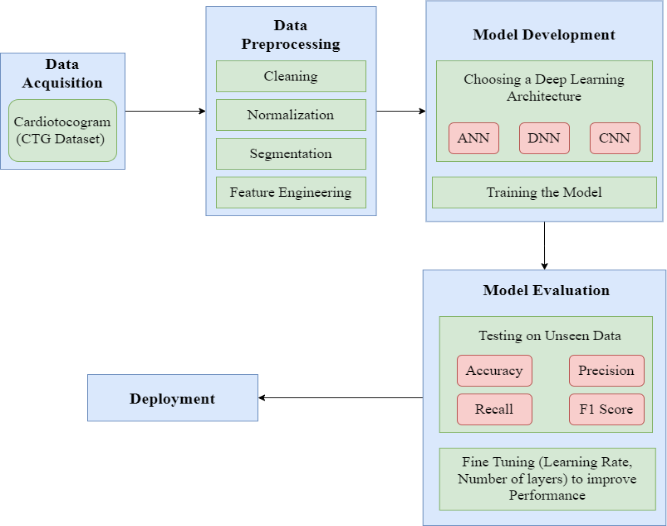
The thesis aims to develop a clinically applicable model for fetal weight estimation based on data from electronic health records (EHR) from pregnant women. The study acknowledges several limitations. First, it relies on retrospective EHR data from a single center, potentially introducing bias. Second, the data set may not be fully representative of pregnant women from other regions Despite these limitations, the study offers several advantages. It presents a fetal weight estimation model that potentially offers more accurate predictions compared to traditional methods such as Hadlock's formula and multiple linear regression (MLR). 16 g) and the mean absolute percentage error was 5.79% (95% CI: 5.70% - 5.81%). Despite its contributions, the article leaves some gaps that could be addressed in future research.[28]

The main aim of this study was to investigate the association between obesity before pregnancy and the risk of low birth weight (LBW), fetal growth restriction (FGR), and macrosomia in a Polish prospective cohort. While the sample size of 912 mothers was suf ficient for the analyses performed, larger sample sizes could provide more robust results, especially for rare outcomes such as fetal growth restriction. The study used a prospective cohort design, which allows data to be collected over time, minimizes recall bias, and provides more reliable information on exposures and outcome conditions and stillbirth prevention. The study used multiple logistic regression to calculate adjusted odds ratios (AORs) with 95% confidence intervals (CIs) for neonatal outcomes about obesity (BMI ≥ 30 kg/m^2) compared with the normal BMI range (18.5–24.9 kg/m^2). Low birth weight (LBW) in the cohort affected 6.6% of pregnancies, fetal growth restriction (FGR) occurred in 2.3% of cases, and macrosomia was observed in 10.6% of neonates. and birth outcomes, some gaps could be addressed in future research.[29]

The primary objective of the research is to evaluate the performance of the Dual Regression Model (DREM) in handling incomplete datasets. A limitation of the study is the inability to perform paired t-tests due to the non-scalar nature of the error range metric used to evaluate the models. Despite the limitations, DREM shows resilience in handling incomplete datasets, with only a marginal decrease in accuracy as the percentage of missing values increases, They proposed regression models for weight estimation based on head size, abdomen size, and femur length, based on its better than On head and body measurements. The error means in their model percentages are 1.3%, 1.5%, 0.4%, 1.4%, 2.3%, and –0.7% while the error standard deviations are 10.1%, 9.8%, 7.7%, 7.3%, 7.4%, and 7.3%. While the study identifies areas where DREM performs well, it also points to some gaps that warrant further investigation.[30]

**III. METHODOLOGY**

**1.Fetal Health Classification Proposed Model:**

****

**1.1:Data Acquisition:**

Data acquisition refers to the process of collecting raw data from various sources for further processing, analysis, or storage. It is a fundamental step in many scientific, engineering, and industrial applications where data-driven insights are needed. Here we use CTG Dataset known as Cardiotocogram. The data should be taken into various medical records. The data is from Kaggle in this dataset consists of 2127 rows and 22 columns, the datasets consist of attributes such as baseline value, accelerations, fetal\_movement, uterine\_contractions and light\_decelerations. Cardiotocography (CTG) is a monitoring technique used during pregnancy and labor to monitor the fetal heartbeat and uterine contractions. It helps assess the well-being of the fetus and detect any signs of distress.CTG data sets typically include recordings of fetal heart rate (FHR) and uterine contractions over a period of time.

**1.2:Datapreprocessing:**  
Data preprocessing is a crucial step in the data analysis pipeline that involves cleaning, transforming, and preparing raw data into a format suitable for further analysis. In the realm of data preprocessing, ensuring data quality and consistency is paramount for meaningful analysis and reliable results. Data cleaning serves as the foundational step, involving the identifcation and rectification of missing values, outliers, and duplicates.

**1.2.1:Data Cleaning**:

**1.2.1.1:Handling missing values:** Identifying and dealing with missing data points, which can involve imputation (replacing missing values with estimated ones), deletion of rows or columns with missing data, or using algorithms that can handle missing values.

**1.2.1.2:Handling outliers:** Detecting and addressing outliers that can skew analysis results. This might involve removing outliers, transforming them, or using robust statistical methods.

**1.2.1.3:Removing duplicates**: Identifying and removing duplicate records to ensure data integrity and avoid redundancy.

**1.2.2:Data Normalization**:

Ensuring that data is in a consistent format and structure, which might involve converting units, standardizing date formats, or resolving inconsistencies in naming conventions.

Its objective is to bring all features to a similar scale or range, which can enhance the performance of machine learning algorithms and improve convergence during training

**1.2.3:Segmentation:**

Segmentation involves dividing a dataset into meaningful subsets or segments based on certain criteria. These criteria could include spatial, temporal, or logical characteristics of the data.

Segmentation involves partitioning datasets into coherent subsets based on specific criteria, tailored to the nature of the data and the analytical objectives.

**1.2.4:Feature engineering:**

Focuses on crafting informative representations from the segmented data to enhance model performance. By creating new features or transforming existing ones, feature engineering equips machine learning algorithms with richer inputs for decision-making.

**1.3:Model Development:**

Model development in deep learning involves a systematic approach to solving a problem. It begins with clearly defining the task at hand, whether it's classification, regression, or another type of problem. Once the problem is defined, the next step is to gather and preprocess the data. This includes cleaning the data, handling missing values, and ensuring it's in a suitable format for input into the model.Choosing the right model architecture is crucial. Depending on the nature of the data and the problem, different types of models may be appropriate.Once the model architecture is selected, it needs to be designed, specifying the number of layers, types of layers, activation functions, and any regularization techniques to prevent overfitting.

**1.3.1:Artifical Neural Network(ANN):**

The ANN model typically consists of an input layer, one or more hidden layers, and an output layer.Each neuron in the input layer represents a feature of the CTG data, such as fetal heart rate or uterine contractions.The hidden layers perform nonlinear transformations on the input data, allowing the model to learn complex patterns and relationships.Activation functions like ReLU (Rectified Linear Unit) or sigmoid are applied to introduce nonlinearity into the model.The output layer produces probabilities or scores for each class of fetal health (e.g., normal, suspicious, pathological) using an appropriate activation function (e.g., softmax for multi-class classification).

**1.3.2:Deep Neural Network(DNN):**

A DNN model consists of multiple layers of neurons, typically organized into an input layer, multiple hidden layers, and an output layer.Each neuron in the input layer represents a feature of the CTG data.Hidden layers perform nonlinear transformations on the input data, allowing the model to learn complex patterns and relationships. Activation functions like ReLU (Rectified Linear Unit) are applied to introduce nonlinearity into the model.The output layer produces probabilities or scores for each class of fetal health using an appropriate activation function.

**1.3.3:Convolutional Neural Networks(CNN):**

A CNN model architecture suitable for CTG data typically includes convolutional layers followed by pooling layers to extract relevant features from the input signals. The convolutional layers capture spatial patterns in the data, while pooling layers reduce dimensionality and enhance computational efficiency.Additional layers such as batch normalization and dropout may be incorporated to improve model generalization and prevent overfitting.The final layers of the CNN model consist of fully connected layers followed by output layers, which classify the fetal health status based on the extracted features.

**1.4:Model Evalutation:Top of Form**

Training a model is the iterative process of adjusting its parameters to minimize the disparity between predicted outputs and actual targets, as quantified by a defined loss function. It begins with initializing the model's parameters, either randomly or through pre-trained weights for transfer learning. During forward propagation, the training data is fed through the network, producing predictions that are compared to the true targets to compute the loss. Backpropagation then calculates the gradients of the loss function with respect to each parameter, enabling parameter updates through optimization algorithms like stochastic gradient descent (SGD). This iterative process continues for multiple epochs, with the model gradually improving its performance by adjusting its parameters based on the training data.

In deep learning, as in other machine learning domains, accuracy, precision, recall, and F1 score are commonly used metrics to evaluate the performance of classification models. Here's a brief overview of each metric:

**1.4.1:Accuracy**: Accuracy measures the overall correctness of the model's predictions and is calculated as the ratio of correctly predicted instances to the total number of instances in the dataset. It provides a general assessment of how well the model performs across all classes.

**Accuracy=True Positives+True Negatives/True Positives+ False Positives+ True Negatives+ False Positives**

**1.4.2:Precision**: Precision measures the proportion of true positive predictions out of all positive predictions made by the model. It focuses on the relevance of the positive predictions and is particularly useful when the cost of false positives is high.

**Precision=True Positives/True Positives+ False Positives**

**1.4.3:Recall (Sensitivity)**: Recall measures the ability of the model to correctly identify all relevant instances, or true positives, out of all actual positive instances in the dataset. It is also known as sensitivity or true positive rate.

**Recall=True Positives/True Positives+False Negatives**

**1.4.4:F1 Score:** The F1 score is the harmonic mean of precision and recall and provides a balanced measure of a model's performance. It takes into account both false positives and false negatives and is particularly useful when the classes are imbalanced.

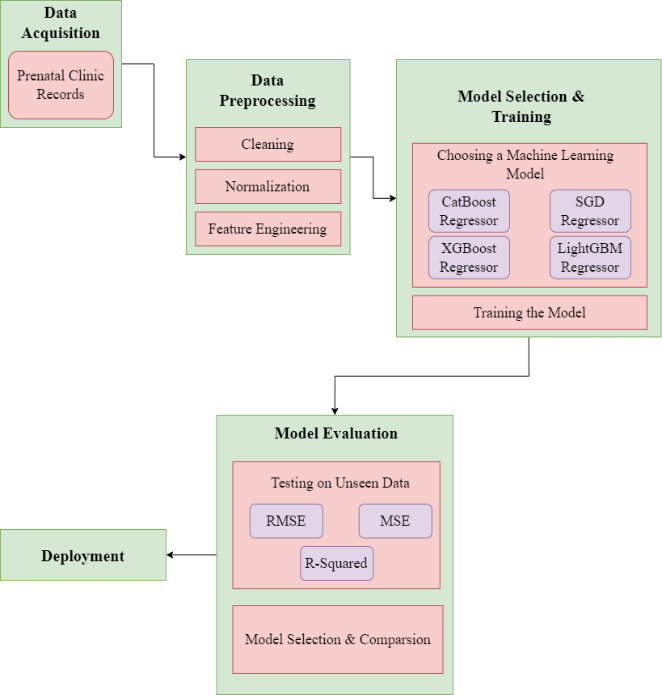
**F1 Score=2×Precision×Recall / Precision+Recall**

These metrics are essential for evaluating the performance of classification models in deep learning tasks, helping practitioners understand the trade-offs between precision and recall and providing insights into the model's overall effectiveness in making correct predictions.

**1.4.5:Fine-tuning:**

Fine-tuning in the context of deep learning typically refers to the process of taking a pre-trained neural network model and further training it on a new dataset or task. This approach leverages the knowledge and representations learned by the pre-trained model on a large dataset adapts it to perform well on a different but related task or dataset.

**2.Birth-Weight of Fetus Prediction Proposed Model:**

****

**2.1:Data Acquisition**:

The data should be taken into various medical records. The data is from Kaggle and in this dataset consists of 7 rows and 1236 columns, the datasets consist of attributes such as birth weight, gestational age, parity, maternal age, maternal weight, and smoking. A healthy newborn usually weighs 2.5 to 4 kilograms, with the ideal weight being 3.5 kilograms. Using these attributes, we try to categorize children as "normal" if their weight falls between 2.5 and 4 kilograms, "suspicious" if below this range, and "pathological" if above. This classification will help predict potential health risks and guide preventive measures accordingly**.**

**2.2:Data Preprocessing:**

After collecting the data, we need to move to the step of pre-processing the data. Data preprocessing is actually a fundamental step in the data analysis and machine learning process. It involves transforming raw data into a clean, organized and structured format that is suitable for analysis, modeling or visualization. The primary goal of data preprocessing is to ensure that the data is consistent, accurate, and relevant to the tasks at hand.

**2.2.1:Data Cleaning:**

In the pre-processing of medical data on fetal health, data cleaning is an essential step aimed at ensuring the quality, accuracy and reliability of the data used for analysis or modelling. Identify and analyze missing values ​​in the data set, especially for important features related to fetal health parameters.Decide on appropriate strategies for handling missing data, such as imputation techniques (eg mean imputation, regression imputation) or removing records with missing values ​​if negligible.

**2.2.2:Data Normalization:**

Data normalization is a crucial step in data preprocessing, especially in the context of medical data such as fetal health records. It involves transforming data to a common scale without distorting differences in value ranges. This process is essential for accurate analysis and modelling.

**2.2.3:Feature Engineering:**

Feature Engineering involves creating new features or modifying existing ones to improve the performance of machine learning models. In the context of fetal health data, functional engineering techniques can be particularly effective in obtaining meaningful information that can aid in accurate predictions or analyses. Here are some technical techniques commonly used in the preprocessing of medical data for fetal health:

**2.2.3.1:Feature Scale:** As mentioned earlier, feature scaling techniques such as minimum and maximum scaling, Z-score standardization, or robust scaling can be used to normalize features within a consistent range. This ensures that features with different scales do not dominate the model training process.

**2.2.3.2: Feature Aggregation:** This involves combining multiple related features to create new, more informative features. For example, in fetal health data, you can aggregate features related to heart rate variability, such as mean heart rate, standard deviation of heart rate, and peak heart rate, into a single feature that captures overall heart rate variability.

**2.2.3.3:Feature Selection:** Not all features may contribute equally to the predictive power of the model. Feature selection techniques such as correlation analysis, recursive feature elimination, or feature importance ratings (eg, using random forests or gradient boosting models) can help identify and retain the most relevant features while discarding redundant or irrelevant ones.

**2.3:Model Selection And Training:**

After preprocessing the data, we can proceed to the next step of model selection and training. In this step we can choose a model and train the models and test them for better and best accuracy. And model selection and training are crucial steps in machine learning (ML) and statistical modeling. And the main motto of model selection involves choosing the best Model from a set of candidate models to solve a particular problem. In this we use Techinque model building and model training Model training refers to the process of using training data to adjust the model's parameters or weights so that it learns to make accurate predictions.

**2.3.1:Catboost Regressor:**

The CatBoost Regressor is trained using the prepared dataset, with the birth weight of the fetus as the target variable.During training, the algorithm automatically handles categorical features without the need for explicit encoding, simplifying the preprocessing step.The model learns to predict the birth weight by minimizing a defined loss function, typically mean squared error (MSE) or root mean squared error (RMSE).

**2.3.2:Sgd REgressor:**

The SGD Regressor is an optimization algorithm commonly used for training linear models. During training, it iteratively updates the model parameters to minimize a defined loss function.It updates the model parameters incrementally by minimizing the loss function using small random batches of data, making it computationally efficient for large datasets.

**2.3.3:Xgboost:**

Xgboost (Extreme Gradient Boosting) is a powerful machine learning algorithm known for its speed, scalability and high performance. It excels in structured data processing, feature importance evaluation, and ensemble learning, making it popular for classification and regression tasks in various domains.Xgboost's capabilities in structured data manipulation, feature importance evaluation, and ensemble learning make it a valuable tool in medical research, diagnostics, treatment planning, and healthcare management.

**2.3.4:Lightgbm Boost:**

LightGBM is a gradient boosting framework that uses tree-based learning algorithms. During training, the model builds decision trees in a leaf-wise manner, optimizing the split at each node to minimize the loss function.It is known for its efficiency, speed and ability to process large data sets.Lightgbm's capabilities in handling large data sets, providing fast and accurate predictions, and supporting various machine learning tasks make it a valuable tool in medical research, diagnosis, treatment planning, and drug development.

**2.4:Model Evalutation:Top of Form**

In the model evaluation step, we typically perform several tasks to assess the performance of our machine learning model. Model evaluation involves key steps Calculate various performance metrics to assess how well the model is performing. Common metrics include: Regression: RMSE, MSE, R-squared, MAE (Mean Absolute Error) Effectively evaluates our machine learning model and makes informed decisions about its performance and potential improvements.

**2.4.1:Root Mean Squared Error (RMSE):** Measures the average size of the errors between predicted and actual values, giving more weight to large errors. Lower RMSE means better model performance**.**

**2.4.2:Mean Squared Error (MSE):**

Similar to RMSE, but without square root, so it is sensitive to outliers. It is calculated as the average of the squares of the differences between the predicted and actual values.

**2.4.3:R-squared (R2):** Also known as the coefficient of determination, R-squared measures the proportion of variance in the dependent variable (target) that is predictable from the independent variables (traits). R-squared ranges from 0 to 1, where higher values ​​indicate a better fit of the model to the data.

**IV.RESULTS**

**Fetal Health Classification**:

**Training set :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| CNN | 91.41% | 92.8% | 90% | 91% |
| DNN | 99.41% | 99.5% | 98.6% | 98% |
| ANN | 98.18% | 98.7% | 97.5% | 97% |

**Testing set :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| CNN | 87.79% | 88% | 89.7% | 89% |
| DNN | 94.13% | 93.54% | 93.1% | 92% |
| ANN | 93.42% | 92.99% | 92% | 93% |

**Birth-Weight of Fetus Prediction:**

**Training set :**

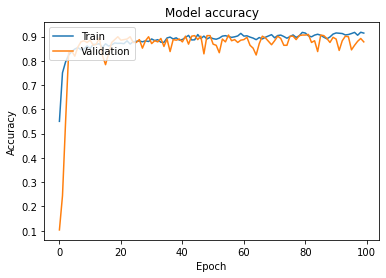
|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **RMSE** | **MSE** | **R- Squared** |
| CatBoost Regressor | 6.42 | 41.29 | 0.78 |
| SGD Regressor | 12.18 | 148.5 | 0.24 |
| XGBoost Regressor | 8.98 | 80.74 | 0.58 |
| LightGBM Regressor | 7.22 | 52.14 | 0.73 |

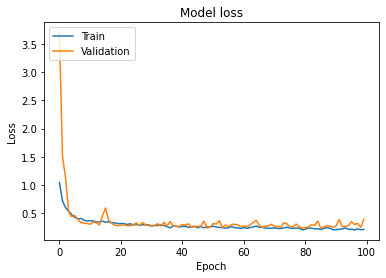
**Testing set :**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **RMSE** | **MSE** | **R- Squared** |
| CatBoost Regressor | 12.84 | 165.0 | 0.08 |
| SGD Regressor | 11.61 | 134.7 | 0.25 |
| XGBoost Regressor | 12.34 | 152.5 | 0.15 |
| LightGBM Regressor | 13.04 | 170.2 | 0.05 |

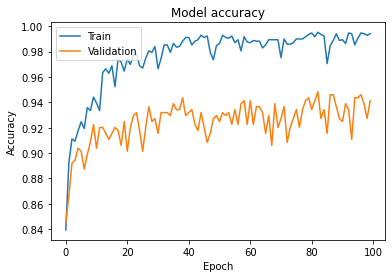
**Charts:**

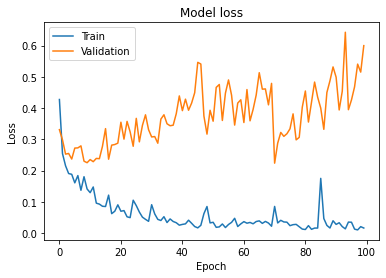
**CNN:**



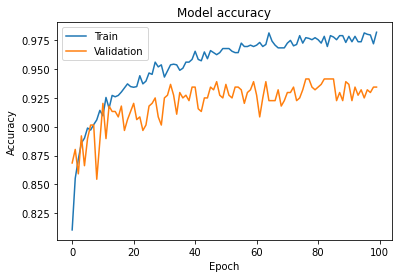


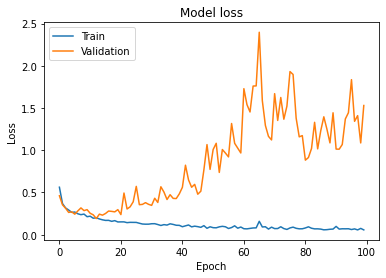
**DNN:**





**ANN:**





**CONCLUSION & FUTURE SCOPE**

With its improved techniques for predicting birth weight and classifying fetal health, machine learning (ML) is transforming prenatal care. To identify fetuses at high risk of complications, machine learning algorithms evaluate a variety of data, including ultrasound scans, maternal health, etc., and fetal heart rate patterns. Pregnancy outcomes are improved by early detection, which allows for rapid intervention. ML models estimate fetal birth weight more accurately than conventional techniques. This helps to detect possible problems such as large for gestational age (LGA) or low birth weight (LBW), allowing the required safety measures to be taken. Although technology should be seen as an aid to doctors rather than a replacement for their knowledge, machine learning (ML) offers remarkable improvements in prenatal care. Pregnancy outcomes could be greatly improved by incorporating machine learning into fetal health monitoring.

To improve the ensemble model's performance and accuracy in forecasting fetal health, more research should be done on cutting-edge machine learning methods and algorithms. To enhance the caliber and applicability of the input data for the ensemble model, look into the possibilities of additional data preparation strategies and feature engineering approaches. To confirm that the suggested ensemble model is effective and generalizable, carry out more tests and analyses on bigger and more varied datasets.

**REFERENCES**

[1]Talukder, M. S. H., & Akter, S. (2023). An Improved Model Ensembled of Different Hyper-parameter Tuned Machine Learning Algorithms for Fetal Health Prediction. arXiv preprint arXiv:2305.17156.

[2]Bahar, E., Agushinta, D., Arimbi, Y. D., & Reksoprodjo, M. (2022). Model Structure of Fetal Health Status Prediction. JUITA: Jurnal Informatika, 10(2), 159-166.

[3]Kuzu, A., & Santur, Y. (2023). Early Diagnosis and Classification of Fetal Health Status from a Fetal Cardiotocography Dataset Using Ensemble Learning. Diagnostics, 13(15), 2471.

[4]Al Duhayyim, M., Abbas, S., Al Hejaili, A., Kryvinska, N., Almadhor, A., & Mughal, H. (2023). Ensemble Learning for Fetal Health Classification. Computer Systems Science & Engineering, 47(1).

[5]Dixit, R. R. (2022). Predicting Fetal Health using Cardiotocograms: A Machine Learning Approach. Journal of Advanced Analytics in Healthcare Management, 6(1), 43-57.

[6]Ogasawara, J., Ikenoue, S., Yamamoto, H., Sato, M., Kasuga, Y., Mitsukura, Y., ... & Ochiai, D. (2021). Deep neural network-based classification of cardiotocograms outperformed conventional algorithms. Scientific reports, 11(1), 13367.

[7]Zhao, Z., Deng, Y., Zhang, Y., Zhang, Y., Zhang, X., & Shao, L. (2019). DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network. BMC medical informatics and decision making, 19, 1-15.

[8]Arun, K., Phaneesh, S., Reddy, S. P., Sreeman, S., & Ghildiyal, Y. (2023). An Information System on Fetal Health Classification based on CNN and Hybrid-CNN with Dimensionality Reduction. In E3S Web of Conferences (Vol. 430, p. 01029). EDP Sciences.

[9]Leruste, S., Delfarguiel, L., Doray, B., Loubaresse, C., Sennsfelder, L., Maillard, T., ... & Spodenkiewicz, M. (2023). The role of general practitioners in Reunion in detecting alcohol use in pregnant women and identifying fetal alcohol spectrum disorder: a qualitative study. Archives of Public Health, 81(1), 210.

[10]Raihen, M. N., & Akter, S. (2024). Comparative Assessment of Several Effective Machine Learning Classification Methods for Maternal Health Risk. Computational Journal of Mathematical and Statistical Sciences, 3(1), 161-176.

[11]Mandala, S. K. (2023). Unveiling the Unborn: Advancing Fetal Health Classification through Machine Learning. arXiv preprint arXiv:2310.00505.

[12]Piri, J., & Mohapatra, P. (2019, December). Exploring fetal health status using an association based classification approach. In 2019 International Conference on Information Technology (ICIT) (pp. 166-171). IEEE.

[13]Li, J., & Liu, X. (2021, March). Fetal health classification based on machine learning. In 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) (pp. 899-902). IEEE.

[14]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.

[15]Yin, Y., & Bingi, Y. (2023). Using machine learning to classify human fetal health and analyze feature importance. BioMedInformatics, 3(2), 280-298.

[16]Rayhan, M. T., Arefin, A. S. M., & Chowdhury, S. A. (2021). Automatic detection of fetal health status from cardiotocography data using machine learning algorithms. Journal of the Bangladesh Academy of Sciences, 45(2).

[17]Mehbodniya, A., Lazar, A. J. P., Webber, J., Sharma, D. K., Jayagopalan, S., K, K., ... & Sengan, S. (2022). Fetal health classification from cardiotocographic data using machine learning. Expert Systems, 39(6), e12899.

[18]Ranjbar, A., Montazeri, F., Farashah, M. V., Mehrnoush, V., Darsareh, F., & Roozbeh, N. (2023). Machine learning-based approach for predicting low birth weight. BMC Pregnancy and Childbirth, 23(1), 803.

[19]Lu, Y., Zhang, X., Fu, X., Chen, F., & Wong, K. K. (2019, July). Ensemble machine learning for estimating fetal weight at varying gestational age. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 9522-9527).

[20]Mane, D. T., Mante, J., Bakare, A. A., Gandhi, Y., Khetani, V., & Mahajan, R. A. (2023). A Systematic Review on The Applications of Machine Learning for Fetal Birth Weight Prediction.

[21]Bekele, W. T. (2022). Machine learning algorithms for predicting low birth weight in Ethiopia. BMC medical informatics and decision making, 22(1), 232.

[22]Magboo, V. P. C., & Magboo, M. S. A. (2022). Prediction of late intrauterine growth restriction using machine learning models. Procedia Computer Science, 207, 1427-1436.

[23]Arayeshgari, M., Najafi-Ghobadi, S., Tarhsaz, H., Parami, S., & Tapak, L. (2023). Machine learning-based classifiers for the prediction of low birth weight. Healthcare Informatics Research, 29(1), 54.

[24]Apoorva, B. A., Nanditha, B. R., Gopi, N., Nidhi, K. S., & Niharika, G. H. PREDICTION OF FETUS WEIGHT IN HIGH RISK PREGANANCIES.

[25]Płotka, S., Grzeszczyk, M. K., Brawura-Biskupski-Samaha, R., Gutaj, P., Lipa, M., Trzciński, T., ... & Sitek, A. (2023). BabyNet++: Fetal birth weight prediction using biometry multimodal data acquired less than 24 hours before delivery. Computers in Biology and Medicine, 167, 107602.

[26]Płotka, S., Grzeszczyk, M. K., Brawura-Biskupski-Samaha, R., Gutaj, P., Lipa, M., Trzciński, T., ... & Sitek, A. (2023). BabyNet++: Fetal birth weight prediction using biometry multimodal data acquired less than 24 hours before delivery. Computers in Biology and Medicine, 167, 107602.

[27]Tao, J., Yuan, Z., Sun, L., Yu, K., & Zhang, Z. (2021). Fetal birthweight prediction with measured data by a temporal machine learning method. BMC Medical Informatics and Decision Making, 21, 1-10.

[28]Wang, Y., Shi, Y., Zhang, C., Su, K., Hu, Y., Chen, L., ... & Huang, H. (2023). Fetal weight estimation based on deep neural network: a retrospective observational study. BMC Pregnancy and Childbirth, 23(1), 560.

[29]Lewandowska, M. (2021). Maternal obesity and risk of low birth weight, fetal growth restriction, and macrosomia: multiple analyses. Nutrients, 13(4), 1213.

[30]Nguyen, L., & Ho, T. H. T. (2018). Early fetal weight estimation with expectation maximization algorithm.