Universal Journal of Public Health 12(5): 1015-1027, 2024 http://www.hrpub.org
DOI: 10.13189/ujph.2024.120525

Public Health Nurse Perspectives on Predicting Preterm Labor Using Risk Factors and Simple Machine Learning Algorithms

Seeta Devi[1], Barkha Devi[2,*], Sonopant G Joshi[1], Dipali Dumbre[1], Surekha Sakore[1], Lily Podder[3]

[1]Symbiosis College of Nursing (SCON), Symbiosis International Deemed University (SIU), Pune, India
[2]Sikkim Manipal College of Nursing (SMCON), Sikkim Manipal University, India
[3]AIIMS College of Nursing, Bhopal, India

Received May 16, 2024; Revised August 26, 2024; Accepted September 23, 2024

Cite This Paper in the Following Citation Styles

***(a):** [1] Seeta Devi, Barkha Devi, Sonopant G Joshi, Dipali Dumbre, Surekha Sakore, Lily Podder , "Public Health Nurse Perspectives on Predicting Preterm Labor Using Risk Factors and Simple Machine Learning Algorithms," Universal Journal of Public Health, Vol. 12, No. 5, pp. 1015 - 1027, 2024. DOI: 10.13189/ujh.2024.120525.*

***(b):** Seeta Devi, Barkha Devi, Sonopant G Joshi, Dipali Dumbre, Surekha Sakore, Lily Podder (2024). Public Health Nurse Perspectives on Predicting Preterm Labor Using Risk Factors and Simple Machine Learning Algorithms. Universal Journal of Public Health, 12(5), 1015 - 1027. DOI: 10.13189/ujph.2024.120525.*

Abstract Preterm birth (PTB) is one of the most precarious obstetrical conditions, which is one of the leading causes of infant mortality. The capacity to predict PTB during both the first and second trimesters provides great promise for improving pregnancy outcomes. The objective of this study was to predict the preterm labor using various risk factors and machine learning models. Executing a novel methodology, researchers used risk factors of preterm labour and machine learning algorithms to predict premature labour. Our data was normalized using the continuous-discrete variables technique, resulting in a single feature value. Our investigation used various prediction models, such as Naïve Bayes, Neural Network (NN), Stochastic Gradient Descent (SGD), AdaBoost, Gradient Boosting (GB), CN2 rule inducer, and k-nearest neighbors (KNN), to predict PTB. This study includes a total of 300 samples for complete analysis. The results showed that GB, AdaBoost, and the CN2 rule inducer have better accuracy ratings of 0.950, 0.947, and 0.930, respectively, along with outstanding Area Under Curve (AUC) values of 0.996, 0.996, and 0.983. Furthermore, with Precision scores of 0.950, 0.947, and 0.930, these models showed strong performance in predicting the probability of PTB. To conclude, GB, adaboost, and CN2 rule inducer models accurately predict PTB, with high AUC and precision scores, making them useful tools for predicting PTB.

Keywords Preterm Labour, Prediction, Risk Factors, Machine Learning Algorithms

1. Introduction

Preterm birth affects over millions of newborns globally each year and is a major public health concern, occurring within 37 weeks of gestation [1]. In addition to chronic health issues such as neurodevelopmental defects and chronic diseases, it is a major source of both neonatal mortality and morbidity [2]. PTB is caused by a complex and multifaceted combination of fetal and environmental factors, in addition to maternal factors including chronic illnesses, stress, and infections [3]. Mothers with PTB require continuous medical care, which negatively affects their mental and general well-being and places additional financial strain on families and the healthcare system [4]. Thus, it is essential to detect PTB early in order to enable medical professionals to predict it at an early stage to perform prompt therapies that enhance positive pregnancy outcomes. As per the obstetric care guidelines, majority of the countries highly recommend that pregnant women undergo regular antenatal check-ups to detect the high risk pregnancy and to protect their safety. This makes it possible to record information about pregnant mothers in electronic health records (EHRs) more frequently, uniformly, and thoroughly than information about other patients [5]. Early screening has been shown to lower the incidence of PTB in pregnant mothers at risk of premature delivery [6]. Therefore, a prediction model is needed to forecast PTB. Because machine learning, a branch of artificial intelligence, may predict a patient's outcome in the future using data extraction, it has seen substantial use in healthcare settings in recent years [7]. Currently, several studies have shown that machine learning can predict PTB [8].

This study's main goal is to forecast preterm labor by utilizing machine learning (ML) techniques to evaluate multiple risk factors. This study also developed a predictive model using real-time data from electronic medical records belonging to mothers who gave birth prematurely. The model considers various risk factors, including past preterm deliveries, numerous pregnancies, cervical length, smoking, substance abuse, infections, psychological stress, and other relevant conditions. Using ML methods including AdaBoost, Gradient Boosting, and the CN2 rule inducer, the objective is to accurately and consistently predict the likelihood of premature labor.

1.1. Related Work

The authors evaluated relevant research papers from several databases under the requirements provided in the literature section criteria (LSC). The present research reviewed different databases, comprising of SCOPUS, PubMed, and IEEE Xplore. Articles of previous research findings are associated with the current study. The survey period under consideration for inclusion extended from 2015 to August 2023 to gain a more in-depth knowledge of past study findings.

Hallingström et al. [9] studied prenatal prediction by analyzing metabolic patterns in amniotic fluid and identifying indicators associated with poor pregnancy outcomes. While this strategy shows potential for early intervention, its actual application in clinical settings remains undeveloped. Similarly, Chen L. et al. [10] used electro-hysterogram (EHG) imaging to monitor uterine activity, revealing the possibility for non-invasive preterm labour prediction. Despite the innovative nature of these studies, the translation of these discoveries into adaptable therapeutic models for routine clinical application remains limited. Further study is required to improve and integrate these predictive technologies into prenatal care practices.

In 2020, Abraham et al. [11] conducted a study highlighting the potential of algorithms to enhance PTB risk prediction by accurately utilizing diverse data from electronic health records, thereby improving medical care during pregnancy. A similar study conducted by Cheng et al. [12] discovered and evaluated recurrent neural network-based deep learning models that predict extreme premature delivery using electronic health records, identify associated risk factors, and achieve early prediction up to 8 weeks in advance. If an accurate PTB prediction is made, healthcare institutions might be able to distribute resources more effectively and ensure that patients receive the proper care.

In an investigation, Qi et al. [13] found that a prediction model built on the Random Forest (RF) algorithm showed promise for anticipating preterm birth at an early gestational age. The primary PTB influencing factors were also identified by the RF model, indicating that early pregnancy intervention can reduce the chance of preterm delivery. Uri et al. [14] in their work constructed a deep learning model that uses clinical data and EHG measurements to predict preterm deliveries. Despotovic D et al. [15] explored the early prediction of preterm delivery utilizing unique factors and EHG recordings, and the RF classifier performed well.

In this study, pertinent data was gathered retrospectively from electronic records of mothers who had premature deliveries in order to develop a preterm delivery prediction model. To predict premature labor, this model consisted 7 factors, including a history of PTB, pregnancy with multiple babies, a short cervix, smoking habits, drug usage, urinary tract infections, stress, depression, genital tract infections, polyhydramnios, etc. This model's main objective is to forecast preterm labor by using a range of attributes and machine learning techniques.

2. Methods

Design of the Study and Dataset

Retrospective study approach was adopted in this research. The data used in this model was real-time data collected from electronic records of mothers who had premature deliveries in order to develop a preterm delivery prediction technique. The samples included in the study were recruited from selected maternity of hospitals of Pune. The following are the features used in the study for the prediction analysis; Previous history of preterm labour, pregnancy with multiple foetuses, short cervix uterine / placental problems, habit of smoking and use of drugs, chorio amnionitis, urinary tract infections, genital tract infections, chronic conditions such as blood pressure / diabetes/ thyroid diseases, depression, stress, polyhydramnios, per vaginal bleeding, fetal defects, short interval between two pregnancies, age and preterm labour at week. The data was normalized using the continuous-discrete variables technique with a single feature value.

Figure 1 shows the workflow of the classification model. It represents the workflow of seven different algorithms used for the prediction model. These algorithms are hyper-tuned for the better results. Due to class imbalance in the dataset, researchers adopted the hyper-tuning machine learning techniques to obtain a high accuracy value.

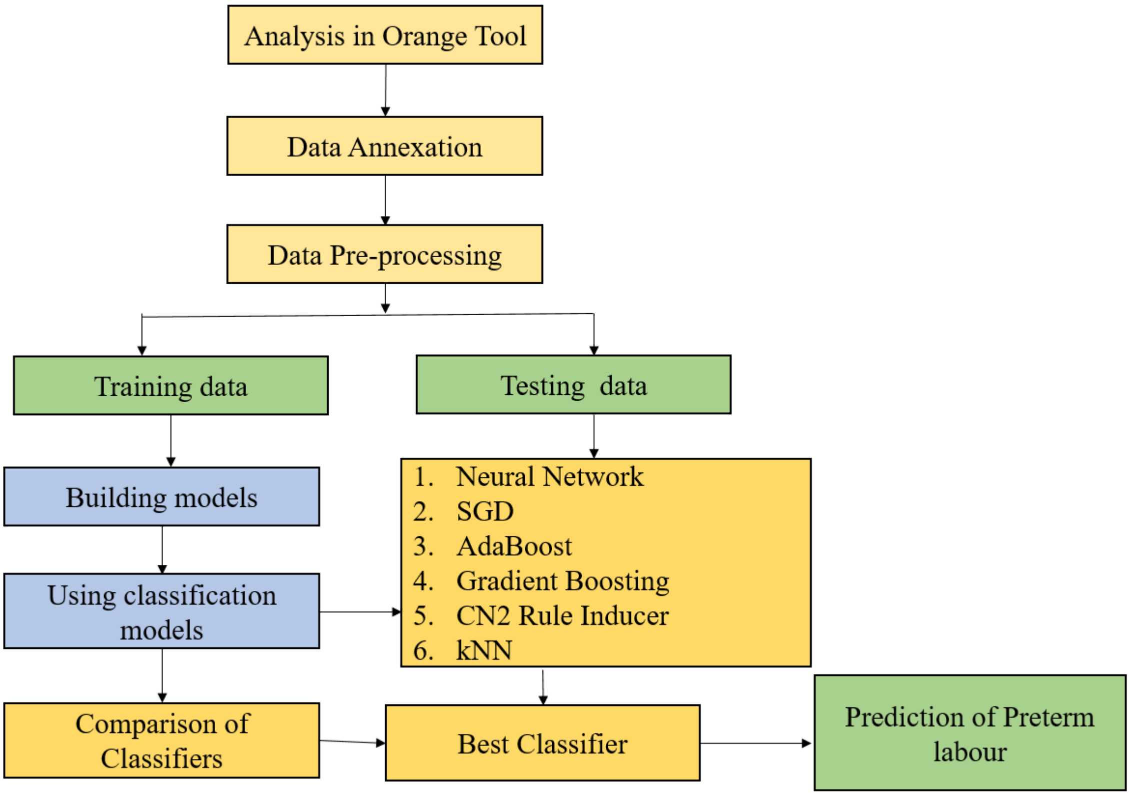

Figure 1. Workflow of classification prediction model for PTB

Following Are the Steps of Model

2.1. Analysis in Orange Tool

The process begins with utilizing the Orange Tool, a versatile and intuitive platform for data analysis, machine learning, and visualization. This tool is used to carry out the complete analysis pipeline, starting from data preparation to model evaluation.

2.2. Data Annexation

This process entails gathering and combining pertinent data from many sources to build a comprehensive dataset ready for analysis. It guarantees that there is enough data to train and assess machine learning models.

2.3. Data Collection and Pre-processing

The data used in the proposed model is real-time data collected from electronic records of the mothers with PTB admitted to the hospital. Mainly, these questionnaires included risk factors that have a direct impact on causing PTB. These are data of risk factors for the prediction of PTB. The data was normalized using a label encoding technique, under which data was converted into numbers, and each factor was scored with 0 or 1 for every true and false value. Due to this data normalization, the current model could perform accurate predictions without overfeeding and overfitting data. None of the columns had a large influence during predictions. The real-data set contained 300 samples and 16 columns with no missing values, of which 14 had categorical values. This data easily fit the classification model, while the classification model needed a categorical target variable. The categories of groups were 0 and 1, making these groups a categorical target variable and skipping the PTB prediction. After this, data was fit for the classification model prediction.

Data pre-processing is an important step that cleans and prepares data for analysis. This may include resolving missing values, normalizing or standardizing the data, encoding categorical variables, and performing other tasks to improve data quality and prepare it for machine learning algorithms.

2.4. Training Data

The dataset is divided into training and testing subgroups following pre-processing. Machine learning models are constructed and trained using training data, which enables the models to discover patterns and relationships in the data. Using an 8:2 ratios, the 300 samples in this investigation were divided into two sets: 240 for training and 60 for testing. The models were trained using the training data, and internal validation was performed using the test set. Moreover, the results of the experiment were confirmed through 10-fold cross-validation.

2.5. Building Model

Using the training data, a variety of machine learning models are built. Various algorithms are used to build models that can classify or predict outcomes depending on the input data. Metrics including AUC, or the receiver operating characteristic (ROC), sensitivity, specificity, and F1 score were used to assess the model's performance. After examining the seven models, the best-performing model was selected for external validation. Information pertaining to pregnant women who underwent childbirth at the hospital during the period spanning January 2022 to December 2023 was compiled. The model's accuracy was determined by comparing its predictions to real clinical outcomes for preterm babies. The statistical analysis was done using the orange tool.

The trained models were evaluated against the chosen test dataset. All experiments were carried out using Orange tool. The evaluation criteria included accuracy, precision, recall, F1 score, and specificity [16]. For multi-class classification, these measures can be averaged in two ways: micro-averaging and macro-averaging. Micro averaging adds up the counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) over all classes before calculating the necessary metrics from them. This method treats all cases equally, regardless of class, to assess the classifier's overall performance across all classes. Macro averaging, on the other hand, computes metrics for each class independently before averaging them. This technique treats each class equally, regardless of class imbalance, and hence evaluates the classifier's performance for each class individually. A confusion matrix was created to show the number of predictions generated by the classifier versus the actual classifications in the test dataset. In the matrix, each row represents the actual instances, and each column represents the anticipated instances.

2.6. Testing the Data

The model's performance is evaluated using the testing data, which was not used in the training phase. This ensures that the models are evaluated using new, previously unseen data, resulting in a meaningful estimate of their effectiveness.

2.7. Using Various Algorithms

In the following subsections, we have presented the theoretical meaning of the algorithms (Figure 1).

Navebayes: NB is a basic supervised machine learning classifier that leverages Bayes' theorem. It computes the posterior probability of each class given the features, assuming independence between features. Mathematically, it estimates the probability P(y=c|x) for each class c using the Equation

$$P(y=c \mid x)=(P(x \mid y=c)P(y=c))/P(x)$$

where P(x|y=c) represents the likelihood of the features given class c, P(y=c) is the prior probability of class c, and P(x) is the marginal probability of the features. For instance, the predicted class is determined by selecting the class with the highest posterior probability. Naïve Bayes's simplicity and efficiency make it well-suited for multiclass classification tasks, providing robust predictions of preterm labour [17].

Neural Network: NN is a custom neural network structured with two hidden layers, each comprising 64 neurons and utilizing Rectified Linear Units (ReLU) for introducing nonlinearity. Softmax activation is used in the output layer for multi-class classification. The Adam optimizer adjusts weights and biases iteratively throughout training. Using accuracy metrics, the sparse categorical cross-entropy loss function calculates prediction accuracy. Early stopping is employed to prevent overfitting, optimizing training by halting when validation performance drops off [18].

Stochastic Gradient Descent (SGD): Finding the optimal value of an objective function—either the lowest or maximum—requires an iterative optimization process. It is a popular method for changing model parameters in machine learning applications in order to lower a cost function. Finding the model parameters that yield the highest accuracy on training and test datasets is the main objective of SGD. The gradient in this method is a vector that, at a given point in time, indicates which way the function is going to climb the steepest. The algorithm can progressively fall to lower function values until it reaches the function's lowest by moving in the gradient's opposite direction [19].

AdaBoost: It serves as a powerful boosting ensemble learning technique. It iteratively trains a sequence of weak learners, such as decision stumps, by assigning higher weights to misclassified instances in each iteration. The final prediction is obtained by aggregating the weighted predictions of all weak learners. Mathematically, it combines the predictions of multiple weak learners, M, as given in the Equation:

$$y_{\text{AdaBoost}}(x) = \text{sign}(\textstyle\sum_{m=1}^{M} \alpha_m h_m(x)),$$

where y_"AdaBoost" (x) represents the predicted class for instance x, α_m denotes the weight assigned to weak learner m, and h_m (x) represents the prediction of weak learner m [20].

Gradient boosting: It is a powerful approach for combining weak learners to enhance prediction. Each consecutive model is trained using gradient descent to reduce the loss inherited from the prior model, such as the mean squared error. The algorithm calculates the gradient of the loss based on current predictions and trains new weak models to minimise this gradient. Predictions from these models are iteratively added to the ensemble until a stop condition is satisfied [21].

CN2 rule induction: This machine learning technique generates explicit and exact categorization rules from datasets. It generates if-then rules, evaluates them using measures like accuracy, selects the most effective rules, prunes to prevent overfitting, adds these rules to the set, and removes cases that have already been covered in the training data. This process continues until a predetermined stopping requirement is fulfilled. CN2 is well-known for its efficiency and capacity to handle both categorical and numerical data, making it useful in areas such as medical diagnosis, marketing, and false detection [22].

K-Nearest Neighbors (KNN): A popular and adaptable machine learning method that may be used for both regression and classification applications is K-Nearest Neighbors (KNN). Since this technique is instance-based and non-parametric, it does not make any assumptions about the distribution of the data. Instead, it bases its prediction on finding cohesions with nearby data points. It operates by locating the 'k' nearest data points in the feature space to a newly discovered data point. To anticipate the result for the new data point, it then uses a majority vote in classification or an average of the values in regression for those 'k' neighbors [23].

2.8. Comparison of Classifiers

To determine which model predicts premature labor the best, the performance of the several classifiers is compared.

2.9. Best Classifier

The classifier that performs the best during testing is selected as the optimal model for prediction [24].

2.10. Prediction of Preterm Labor

The selected model is then employed to predict the likelihood of preterm labor using new data. This step is the final objective, where the model provides reliable predictions that can aid in clinical decision-making or further studies.

3. Results

Figure 2 depicts the distribution of preterm births among people with a short cervix, organized by age group: 26-30, 31-35, and 36-40. Each dot represents a case, with different colors reflecting the various age groups. The graph shows that preterm births occur at all age groups, with a significant cluster in the 31-35 age range (red). The 26-30 group (blue) also has a high number of cases, whereas the 36-40 group (green) has fewer. This implies a possible age-related tendency, with women aged 31 to 35 being more likely to have preterm births due to a short cervix. Figure 3 depicts the distribution of premature labor cases based on gestational weeks and smoking and drug use behaviors. The data points are further divided into three age groups: 26-30, 31-35, and 36-40 years, with each represented by a different color. The scatter plot shows that premature labor occurs in all age categories, with a significant concentration in the 31-35 age group (red), which corresponds to the pattern seen in the preceding image. The distribution shows no evident differential depending on smoking or drug use habits, indicating that these factors may not have a substantial influence on the occurrence of PTB in this dataset. Figure 4 shows the ROC curves for various classification models used to predict PTB between 26 and 30 weeks of gestation. At various threshold settings, the graph contrasts the true positive rate (sensitivity) with the false positive rate (specificity). A large area under the curve (AUC) is shown by curves approaching the upper left corner in well-performing models. The curves indicate a range of prediction capacities, with some models approaching optimal performance (AUC close to 1) and others exhibiting modest accuracy. Overall, the figure shows that various models may accurately predict preterm labor within this gestational window. Figure 5 shows the ROC curves for various classification models used to predict PTB between 31 and 37 weeks' gestation. The trade-off between each model's true positive rate (sensitivity) and false positive rate (specificity) at different thresholds is shown on the graph. A higher AUC is shown by curves toward the upper left corner in well-performing models. The curves reveal a wide range of prediction abilities, with some models performing near-optimally (AUC close to 1) and others being less accurate. Overall, the picture demonstrates that certain models are extremely good at predicting preterm labor during this gestational period. Table 1 presents the comparison of prediction models utilizing various ML algorithms reveals considerable disparities in performance. The top performers are GB and AB, with AUC scores of 0.996 and classification accuracy (CA) values close to 0.95, demonstrating their great predictive skills. These models also have strong F1 scores, precision, recall, and matthews correlation coefficient (MCC) values more than 0.9, demonstrating their reliability. However, NB and NN had poorer AUC (0.666), CA, and F1 values, indicating poor performance. The SGD model performs marginally better, while kNN has moderate accuracy but falls short in recall and MCC. Table 2 presents the precise setups for many ML algorithms, such as DT, kNN, SVM, RF, GB, AB, and NN. It defines critical hyperparameters and their values, including tree count, learning rates, and kernel options. These setups handle important issues such as tree depth, neighbor count, regularization, and activation methods. This detailed perspective guarantees that the algorithms are properly calibrated for their particular jobs. Furthermore, replicable training is used regularly, guaranteeing that the models yield stable and repeatable results across multiple runs, increasing the reliability of Ml studies. Table 3 presents the comparison of multiple models demonstrates that AB and GB are highly accurate at predicting age groups, particularly 26-30 and 31-35. kNN works flawlessly for 26-30 but drastically misclassifies the other groups. NN and SGD have reasonable accuracy with some misclassifications, however Naive Bayes has the largest misclassification, particularly overrepresenting the 31-35 age range.

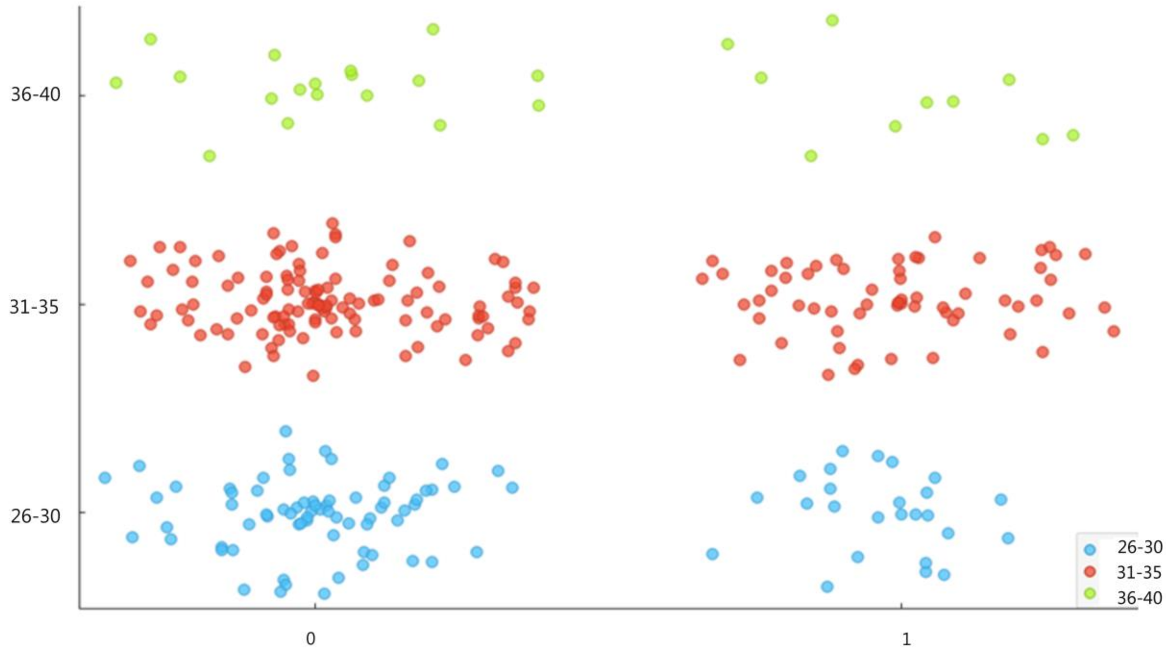

Figure 2. Distribution of the preterm cases with short cervix

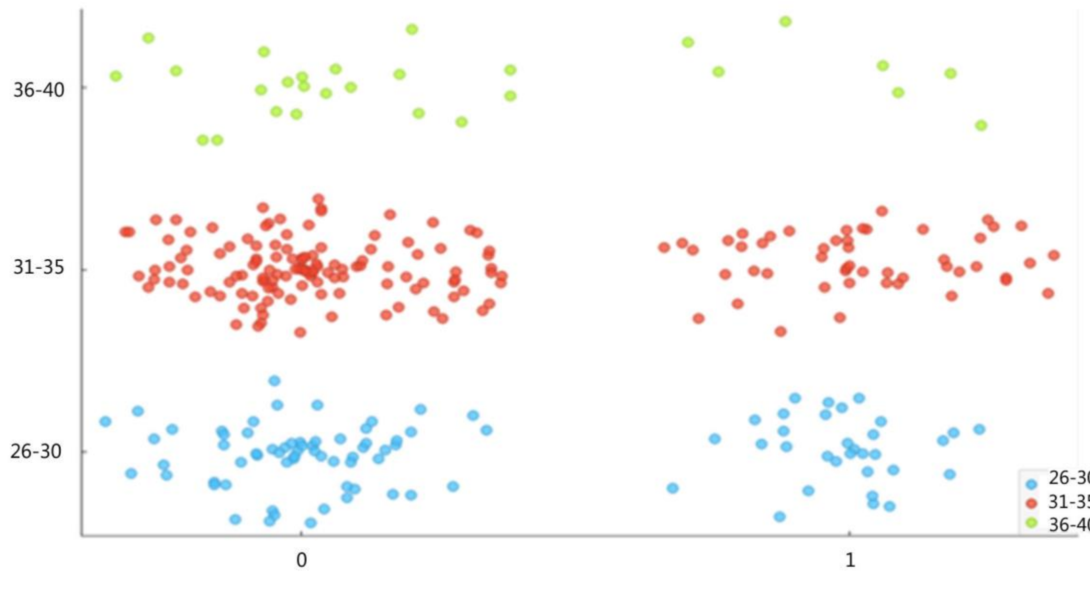

Figure 3. Distribution of the preterm labour in weeks and habit of smoking and use of drugs

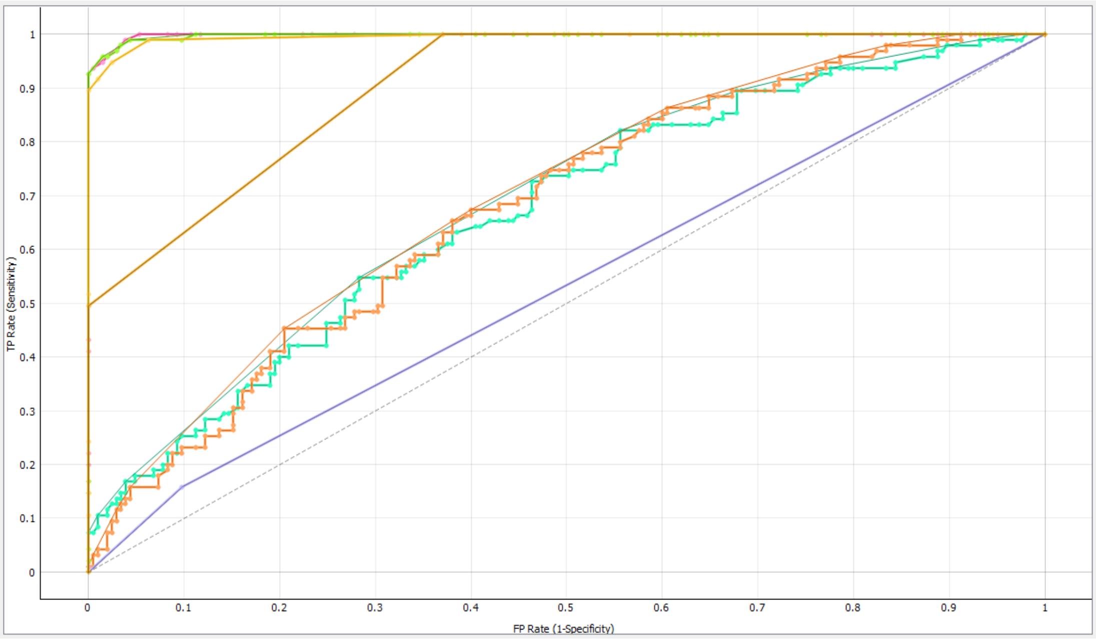

Figure 4. ROC curve for of various classification model for prediction preterm labour (26-30 weeks)

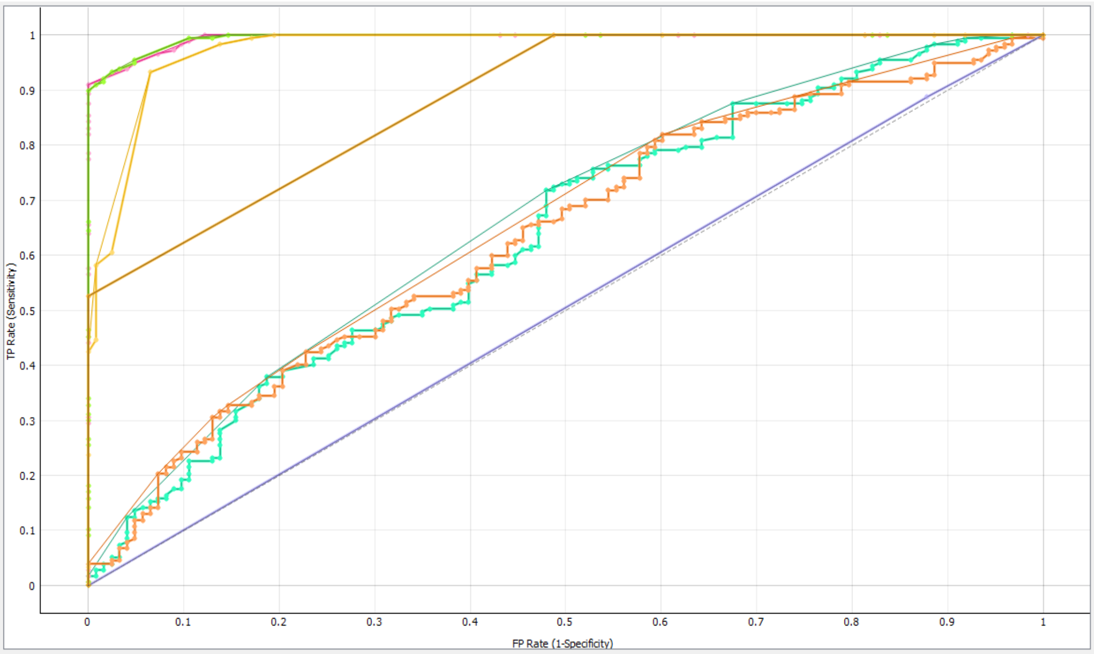

Figure 5. ROC curve for of various classification model for prediction preterm labour (31– 37 weeks)

Table 1. Prediction model deploying various machine learning algorithms

Model	AUC	CA	F1	Precision	Recall	MCC
Naive Bayes	0.666	0.613	0.542	0.573	0.613	0.175
Neural Network	0.666	0.600	0.528	0.524	0.600	0.132
SGD	0.571	0.630	0.562	0.572	0.630	0.216
AdaBoost	0.996	0.947	0.946	0.946	0.947	0.901
GB	0.996	0.950	0.948	0.950	0.950	0.907
CN2 rule inducer	0.983	0.930	0.927	0.929	0.930	0.869
kNN	0.902	0.680	0.654	0.766	0.680	0.480

Table 2. List models and fine-tuned hyperparameters used in classification

Algorithm	Hyperparameter	Value
DT	Create a binary tree	Yes
	Minimum quantity of occurrences in leaves	1
	Never divide subgroups into smaller than	1
	Limit the maximum depth of the tree to	93
	When the majority is reached, stop [%]	98
kNN	The quantity of neighbors	7
	Metric	Euclidean
	Weight	By Distances
SVM	SVM Type	v-SVM
	Regression cost (C)	3.50
	Complexity bound (ν)	0.55
	Kernel Type	Polynomial
	Kernel parameters (g, c, d)	auto, 1.00, 3.5
	Numerical tolerance	0.0011
	Iteration limit	150
Random Forest	Number of trees	25
	Quantity of qualities taken into account at each split	6
	Training that is replicable	Yes
	Equalize the distribution of classes	Yes
	Limit each tree's depth	20
	Never divide subgroups into smaller than	5
GB	Method	Extreme Gradient Boosting (xgboost)
	The quantity of trees	109
	Rate of learning	0.303
	Training that is replicable	Yes
	Regularization (Lambda)	1
	Limit each tree's depth	10
	Percentage of training cases	1.00
	Each tree's feature fraction	1.00
	Percentage of each level's features	1.00
	Partiality of attributes for every division	1.00
AB	Foundation estimator	Tree
	Quantity of estimators	48
	Rate of learning	0.99996
	Random generator with a fixed seed	2
	An algorithm for classification	SAMME.R
	Loss function for regression	Linear
NN	Hidden layers of neurons	200
	Initiation	ReLU
	Solution	L-BFGS-B
	Normalization (α)	0.0001
	Maximum number of cycles	200
	Training that is replicable	Yes

Table 3. Confusion matrix for various machine learning algorithms for prediction of preterm labour

			NB		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	21.1%	77.9%	1.1%	95
Actual	31-35	6.8%	92.1%	1.1%	177
	36-40	10.7%	85.7%	3.6%	28
	Σ	35	261	4	300
			Neural Network		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	22.1%	77.9%	0.0%	95
Actual	31-35	10.2%	89.8%	0.0%	177
	36-40	7.1%	92.9%	0.0%	28
	Σ	41	259	0	300
			SGD		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	11.6%	88.4%	0.0%	95
Actual	31-35	4.0%	96.0%	0.0%	177
	36-40	3.6%	96.4%	0.0%	28
	Σ	19	281	0	300
			AdaBoost		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	95.8%	4.2%	0.0%	95
Actual	31-35	1.7%	96.0%	2.3%	177
	36-40	3.6%	14.3%	82.1%	28
	Σ	95	178	27	300
			GB		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	95.8%	4.2%	0.0%	95
Actual	31-35	1.1%	98.3%	0.6%	177
	36-40	3.6%	25.0%	71.4%	28
	Σ	94	185	21	300
			kNN		
			Predicted		
		26-30	**31-35**	**36-40**	**Σ**
	26-30	100.0%	0.0%	0.0%	95
Actual	31-35	39.5%	60.5%	0.0%	177
	36-40	21.4%	71.4%	7.1%	28
	Σ	171	127	2	300

4. Discussion

The prenatal stage of human life is divided into 3 main stages. The first stage includes the first two weeks after fertilization and is called the germinal period. Then the second stage is called the embryonic stage, where the embryo acquires a cylindrical shape and all the body's systems are formed. This is the critical period as the embryonic development of all human body systems occurs during this period. The last stage, which begins from the third month until birth, is called the fetal stage. During this stage, the fetus begins to acquire a human appearance with more vitality and movement. Moreover, there is a maturation in the body's various systems to be ready to function after birth.

Since preterm birth (PTB) significantly affects the health and survival of newborns, it is a topic of great importance in obstetrics. PTB is characterized by a baby's death and long-term health issues such as chronic illnesses and neurological abnormalities [25]. Accurate prediction and early intervention are critical for improving maternal and newborn health outcomes [26].

Numerous research efforts attempted to accurately forecast preterm birth, enabling timely identification and intervention. Preterm birth can be reliably predicted by early screening, fetal fibronectin testing, or both of these techniques combined. In the context of preterm delivery, Hamilton et al. [26] used decision tree machine learning to identify clusters of pregnancy factors that are most likely to result in PTB and severe neonatal morbidity or death. Similarly, the findings of this AdaBoost trial suggest that fetal hemorrhage, placenta praevia, and preterm membrane rupture all require vigilant monitoring.

It is important to train pregnant mothers to recognize pertinent risk symptoms in order to minimize the chance of unfavourable pregnancy outcomes by ensuring prompt medical attention when early symptoms appear. In order to properly regulate weight and avoid uterine overextension, which may raise the risk of unfavourable pregnancy outcomes, they must adhere to strict nutritional guidelines and engage in modest physical activity [27,28]. In their model to predict preterm deliveries, Khatibi et al. [29] included 112 parameters. Their final model's AUC, however, was 0.68, which is lower than this study's. This demonstrates that the selection of variables that have a strong correlation with preterm birth is more significant than the quantity of variables utilized in the model's construction.

Rawashdeh et al [30] developed a preterm birth prediction model employing 19 input factors and methodologies such as DT and RF, yielding an AUC of 97% that is greater to current study results. However, shortcervix is a major factor in contributing preterm births. Multiple fetuses can be another responsible factor in contributing preterm birth, in line with other research [31]. Research indicates that carrying several fetuses causes the uterus to become overstretched, perhaps resulting in uterine contractions [32]. This could increase the likelihood of issues such intrauterine discomfort, aberrant fetal location, and intrauterine growth restriction (IUGR) [33]. Prenatal bleeding and placenta praevia are also reasonably significant predictors of preterm labour. Numerous conditions, including trauma, high oxytocin, labour obstruction, scarring of the uterus, repeated pregnancies, high amniotic fluid, fetal malposition, etc., can result in prenatal bleeding. However, placental abruption and placenta praevia are the most frequent causes of fetal bleeding in late pregnancy [34]. When the lower uterine wall is expanded in women having anterior placentas, the placenta finds it difficult to adjust and move between the two, reducing the placenta's effective area and limiting the foetus' access to nutrients and oxygen. Moreover, placenta praevia is readily coupled with placental placement, which significantly raises the risk of bleeding and has a significant impact on the foetus' blood circulation, restricting fetal growth and resulting in preterm labour [35].

Predicting PTB requires an understanding of the factors that influence preterm birth. It is difficult to identify the specific mechanisms causing PTB in individuals due to intricate interplay among affecting factors. According to the AdaBoost analysis, early membrane rupture, multiple pregnancies, placenta placement issues, prenatal bleeding, streptococcus agalactiae presence, uterine scarring, abnormal foetal positioning or weight, amniotic fluid characteristics, maternal age, and diabetes all play a role in PTB. Premature membrane rupture is revealed as the strongest predictor. According to studies, only about 8% of women who experience a premature membrane rupture recover [36].

The study also identified key risk factors like early membrane rupture and placenta previa, which have been consistently highlighted in previous research as significant predictors of PTB [37, 38]. Incorporating these factors into predictive models has been shown to significantly improve their accuracy. Menon [38] highlighted that membrane rupture is a critical early sign of PTB, and its inclusion in predictive models can improve early detection rates. Moreover, Romero et al. emphasized the importance of placental abnormalities in the development of PTB, supporting the focus on placenta-related risk factors in this study.

The results of the present investigation align with these advancements, as the Gradient Boosting and AdaBoost models achieved AUC values of 0.996, signifying exceptional precision and dependability. These results suggest that these algorithms are able to forecast pregnancies at risk of preterm birth (PTB) correctly, which enables timely treatments. Conversely, models like Neural Networks and Naïve Bayes showed worse performance, maybe because of their [39].

Incorporating these prediction models into EHRs has the potential to enhance clinical settings by facilitating real-time risk assessment and decision-making, hence leading to better outcomes for mothers and newborns [40]. Postpartum hemorrhage would be less common if medical professionals could recognize high-risk pregnancies early on and take the necessary precautions.

In summary, machine learning techniques offer a great deal of promise for predicting premature labor, especially when applied to ensemble models like Gradient Boosting and AdaBoost. These methods enable early detection and intervention in high-risk pregnancies, which could significantly improve results for the health of the mother and the newborn. Subsequent studies ought to concentrate on confirming these models in larger, more varied populations and exploring the application of deep learning to enhance prediction precision.

5. Limitations and Future Research Directions

Although the study provides helpful information, there are some issues with it. Although 300 is a sufficient sample size for first analysis, it could not offer sufficient statistical power for more extensive generalization. Stronger validation across a wide range of demographics would be possible with a larger and more varied sample. Furthermore, some algorithms' performance may still be impacted, which could result in biased results, even though hyper-tuning techniques have been utilized to lessen the dataset's class imbalance. There's a chance that other important variables, such as genetic predispositions or more nuanced socioeconomic aspects, were overlooked because the study concentrated on just 17 risk factors.

The study's use of retrospectively gathered data may induce biases in data recording, emphasizing the significance of prospective studies for more trustworthy data collecting. Furthermore, the study's single-center methodology, which used data from selected maternity facilities in Pune, limits its applicability to other regions. Future research should include multicenter trials to validate the model in other healthcare settings. Incorporating more complex machine learning methods, such as deep learning models, may also increase predicted accuracy. Finally, conducting external validation tests and incorporating predictive models into EHRs would improve the model's usability and generalizability in clinical settings.

The limitations of this study should be addressed in future research to increase the predicted accuracy and practicality of preterm labor models. Increasing the sample size and merging information from many sites would offer more thorough validation for a wider range of demographics. Moreover, adding factors like socioeconomic status and genetics could increase the accuracy of the model. Investigating cutting-edge machine learning technologies like deep learning will help to increase forecast accuracy. The biases present in retrospective data collecting may be mitigated with the help of prospective investigations. Moreover, integrating these prediction models into electronic health records (EHRs) could facilitate prompt decision-making, leading to better results for the health of mothers and fetuses. To verify the models' relevance and usefulness in clinical applications, external validation in diverse healthcare settings is required, as is constant updating with new data.

6. Conclusions

Simple machine learning techniques can be used to predict premature labour. These algorithms have the potential to be useful tools for diagnosing high-risk pregnancies and improving outcomes for both the mother and the baby, as evidenced by their high precision and AUC-ROC scores. Future research should focus on replicating these findings in bigger and more diverse cohorts, as well as investigating the viability of utilizing more advanced machine learning algorithms to predict premature labour. This technique can help us better understand preterm labour and design more effective preventative and intervention strategies.

Abbreviations

The following abbreviations are used in this manuscript:

AB	:	AdaBoost
AUC	:	Area Under Curve
EHG	:	Electro-hysterogram
GB	:	Gradient Boosting
FN	:	False Negatives
FP	:	False Positives
KNN	:	k-nearest neighbors
NB	:	Naïve Bayes
NN	:	Neural Network
PTB	:	Preterm birth
RF	:	Random Forest
ROC	:	the receiver operating characteristic
SGD	:	Stochastic Gradient Descent
TN	:	True Negatives
TP	:	True positives

REFERENCES

[1] Cao G., Liu J., Liu M, "Global, Regional, and National Incidence and Mortality of Neonatal Preterm Birth, 1990-2019," *JAMA Pediatrics*, vol. 176, no. 8, pp. 787–96, 2022. DOI: 10.1001/jamapediatrics.2022.1622.

[2] Walani S. R, "Global Burden of Preterm Birth," *International Journal of Gynaecology and Obstetrics*, vol. 150, no. 1, pp. 31–3, 2020. DOI: 10.1002/ijgo.13195.

[3] Vogel J. P., Chawanpaiboon S., Moller A-B., Watananirun K., Bonet M., Lumbiganon P,"The Global Epidemiology of Preterm Birth," *Best Practice & Research Clinical*

Obstetrics & Gynaecology, vol. 52, pp. 3–12, 2018. DOI: 10.1016/j.bpobgyn.2018.04.003.

[4] Baumann N., Bartmann P., Wolke D, "Health-related quality of life into adulthood after very preterm birth," Pediatrics, vol. 137, no. 4, 2016. DOI: 10.1542/peds.2015-3148.

[5] Paquette A. G., Hoo L., Price N. D., Sadovsky Y, "Deep phenotyping during pregnancy for predictive and preventive medicine," Science Translational Medicine, vol. 12, no. 527, 2020. DOI: 10.1126/scitranslmed. aay1059.

[6] Ville Y., Rozenberg P, "Predictors of preterm birth," Best Practice & Research Clinical Obstetrics & Gynaecology, vol. 52, pp. 23–32, 2018. DOI: 10.1016/j.bpobgyn.2018.05 .002.

[7] Obermeyer Z., Emanuel E. J, "Predicting the future-big data, machine learning, and clinical medicine," The New England Journal of Medicine, vol. 375, no. 13, pp. 1216-19, 2009. DOI: 10.1056/NEJMp1606181

[8] Iftikhar P., Kuijpers M. V., Khayyat A., Iftikhar A., DeGouvia De Sa M, "Artificial intelligence: A new paradigm in obstetrics and gynecology research and clinical practice," Cureus, vol. 12, no. 2, 2020. DOI: 10.7759/cureus.7124.

[9] Hallingström M., Barman M., Savolainen O., Viklund F., Kacerovsky M., Brunius C, "Metabolomic profiles of mid-trimester amniotic fluid are not associated with subsequent spontaneous preterm delivery or gestational duration at delivery," Journal of Maternal-Fetal & Neonatal Medicine, vol. 35, no. 11, pp. 2054–62, 2022. DOI: 10.1080/14767058.2020.1777271.

[10] Chen L., Hao Y, "Feature extraction and classification of EHG between pregnancy and labour group using Hilbert-Huang transform and extreme learning machine," Computational and Mathematical Methods in Medicine, vol. 2017, 2017. DOI: 10.1155/2017/7949507

[11] Abraham A., Le B., Kosti I., Straub P., Velez-Edwards D. R., Davis L. K., Newton J. M., Mugli, L. J., Rokas A., Bejan C. A., Sirota M., Capra J. A, "Dense phenotyping from electronic health records enables machine learning-based prediction of preterm birth," BMC Medicine, vol. 20, no. 1, p. 333, 2022. DOI: 10.1186/s12916-022-02522-x.

[12] Gao C., Osmundson S., Velez Edwards D R., Jackson G. P., Malin B A., Chen Y, "Deep learning predicts extreme preterm birth from electronic health records," Journal of Biomedical Informatics, vol. 100, no. 103334, pp. 103334, 2019. DOI: 10.1016/j.jbi.2019.103334.

[13] Sun Q., Zou X., Yan Y., Zhang H., Wang S., Gao Y., Liu H., Liu S., Lu J., Yang Y., Ma X, "Machine learning-based prediction model of preterm birth using electronic health record," Journal of Healthcare Engineering, vol. 2022, pp. 1–12, 2022. DOI: 10.1155/2022/9635526.

[14] Goldsztejn U., Nehorai A, "Predicting preterm births from electrohysterogram recordings via deep learning," PLOS ONE, vol. 18, no. 5, 2023. DOI: 10.1371/journal.pone.028 5219.

[15] Despotovic D., Zec A., Mladenovic K., Radin N., Turukalo T L,"A machine learning approach for an early prediction of preterm delivery," 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), pp. 000265-000270 2018. DOI: 10.1109/SISY.2018.85248 18

[16] Hossin M., Sulaiman M N, "A review on evaluation metrics for data classification evaluations," International Journal of Data Mining & Knowledge Management Process, vol. 5, no. 2, pp. 1-11, 2015. DOI: 10.5121/ijdkp.2015.5201

[17] Muhammad A., Kamran Ali M., Abdul Aleem J., Saleemullah M., Anees A, "Multinomial Naive Bayes classification model for sentiment analysis," IJCSNS International Journal of Computer Science and Network Security, vol. 19, no. 3, pp. 62-67, 2019. DOI: 10.13140/RG.2.2.30021.40169

[18] Féraud R., Fabrice C, "A methodology to explain neural network classification," Neural Networks, vol. 15, no. 2, pp. 237-246, 2002. DOI: 10.1016/s0893-6080(01)00127-7.

[19] R. J. P. Princy, S. Parthasarathy, S. T. George, and M. S. P. Subathra, "Predicting the gestational period using machine learning algorithms," in Data Intelligence and Cognitive Informatics, Singapore: Springer Nature Singapore, 2023, pp. 545–560.

[20] S. Mohammadi Far, M. Beiramvand, M. Shahbakhti, and P. Augustyniak, "Prediction of preterm labor from the electrohysterogram signals based on different gestational weeks," Sensors (Basel), vol. 23, no. 13, p. 5965, 2023.

[21] Y. Zhang et al., "Establishment of a model for predicting preterm birth based on the machine learning algorithm," BMC Pregnancy Childbirth, vol. 23, no. 1, 2023.

[22] M. Saha, S. Nayak, N. Mohanty, V. Baral, and I. Rout, "Preterm delivery prediction using gradient boosting algorithms," in Lecture Notes in Networks and Systems, Singapore: Springer Singapore, 2021, pp. 59–68.

[23] J. W. Grzymalabusse and L. K. Woolery, "Improving Prediction of Preterm Birth Using a New Classification Scheme and Rule Induction," JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION, pp. 730–734, 1994.

[24] Li J., Li G., Hai C., Guo M, "Transformer fault diagnosis based on multi-class AdaBoost algorithm," IEEE Access: Practical Innovations, Open Solutions, vol. 10, pp. 1522-1532, 2021. DOI: 10.1109/access.2021.3135467

[25] Zhang Y., Lu S., Wu Y., Hu W., Yuan Z, "The prediction of preterm birth using time-series technology-based machine learning: Retrospective cohort study," JMIR Medical Informatics, vol. 10, no. 6, pp. 1-14, 2022. DOI: 10.2196/33835.

[26] Hamilton F., Dyachenko A., Ciampi A., Maurel K., Warrick A., Garite J, "Estimating risk of severe neonatal morbidity in preterm births under 32 weeks of gestation," Journal of Maternal-Fetal & Neonatal Medicine, vol. 33, no. 1, pp. 73–80, 2020. DOI: 10.1080/14767058.2018.148739 5.

[27] Crowson M. G., Moukheiber D., Arévalo R., Lam D., Mantena S., Rana A., Goss D., Bates W., Celi A, "A systematic review of federated learning applications for biomedical data," PLOS Digital Health, vol. 1, no. 5, article e0000033, 2022. DOI: 10.1371/journal.pdig.0000033.

[28] Akazawa M., Hashimoto, K. "Prediction of preterm birth using artificial intelligence: a systematic review," Journal of

Obstetrics and Gynaecology, vol. 42, no. 6, pp. 1662–8, 2022. DOI: 10.1080/01443615.2022.2056828.

[29] Khatibi T., Kheyrikoochaksarayee N., Sepehri M. M,"Analysis of big data for prediction of provider-initiated preterm birth and spontaneous premature deliveries and ranking the predictive features," Archives of Gynecology and Obstetrics, vol. 300, no. 6, pp. 1565–82, 2019. DOI: 10.1007/s00404-019-05325-3.

[30] Rawashdeh H., Awawdeh S., Shannag F., Henawi E., Faris H., Obeid N., Hyett, J, "Intelligent system based on data mining techniques for prediction of preterm birth for women with cervical cerclage," Computational and Structural Biotechnology Journal, vol. 85, article 107233, 2020. DOI: 10.1016/j.compbiolchem.

[31] Zhang Y-J., Shen J., Lin S. B., Lu C., Jiang H., Sun Y., Cheng X., Wang H., Cui S., Liu X., Huang L., Lin X., Zhao G., Yang L., Chen C, "The risk factors of preterm birth: A multicentre case–control survey in China in 2018," Journal of Paediatrics and Child Health, vol. 58, no. 8, pp. 1396–406, 2022. DOI: 10.1111/jpc.16002.

[32] Norwitz E. R., Edusa V., Park J. S, "Maternal physiology and complications of multiple pregnancy," Seminars in Perinatology, vol. 29, no. 5, pp. 338–48, 2005. DOI: 10.1053/j.semperi.2005.08.002.

[33] D’Alton M., Breslin N, "Management of multiple gestations," International Journal of Gynaecology and Obstetrics, vol. 150, no. 1, pp. 3–9, 2020. DOI: 10.1002/ijgo.13168

[34] Silver R. M, "Abnormal placentation: Placenta previa, Vasa previa, and placenta accrete," Obstetrics & Gynecology, vol. 126, no. 3, pp. 654–68, 2015. DOI: 10.1097/aog.00000000 00001005.

[35] Ananth, C. V., Demissie K., Hanley M. L, "Birth weight discordancy and adverse perinatal outcomes among twin gestations in the United States: The effect of placental abruption," American Journal of Obstetrics & Gynecology, vol. 188, no. 4, pp. 954–60, 2003. DOI: 10.1067/mob.2003.210.

[36] Lee S. M., Park K. H., Hong S., Kim Y. M., Park Y. H., Lee Y. E., Jeon S.J, "Identification of cultivable bacteria in amniotic fluid using cervicovaginal fluid protein microarray in preterm premature rupture of membranes," Reproductive Sciences, vol. 27, no. 4, pp. 1008–17, 2020. DOI: 10.1007/s43032-020-00143-4.

[37] Menon, R., "Spontaneous preterm birth, a clinical dilemma: etiologic, pathophysiologic and genetic heterogeneities and racial disparity," Acta Obstetricia et Gynecologica Scandinavica, vol. 87, no. 6, pp. 590-600, 2008. DOI: 10.1080/00016340802005126.

[38] Romero, R., Dey, S. K., & Fisher, S. J., "Preterm labor: one syndrome, many causes," Science, vol. 345, no. 6198, 760-765, 2014. DOI: 10.1126/science.1251816.

[39] Zhu, Y., Zhang, Y., Liu, X., & Fan, W., "A comparison of feature selection methods for machine learning-based predictive modeling of preterm birth in women undergoing IVF," Reproductive Biology and Endocrinology, vol. 17, no. 1, pp. 1-9, 2019.

[40] Obermeyer, Z., & Emanuel, E. J., "Predicting the future—big data, machine learning, and clinical medicine," The New England Journal of Medicine, vol. 375, no. 13, pp. 1216-1219, 2016. DOI: 10.1056/NEJMp1606181.