An Empirical Comparison of Machine and Deep Learning Algorithms for Predicting Maternal Health Risk

Srinivasa Gupta N[1], B. Valarmathi[2,2](✉), A. Kavitha[3], Sai Bhuvan Gopini[2], and Hari Kumar Seela[2]

[1] Department of Manufacturing Engineering, School of Mechanical Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

[2] Department of Software and Systems Engineering, School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, Tamil Nadu, India

valarmathi.b@vit.ac.in, {gopinisai.bhuvan2020, seelahari.kumar2020}@vitstudent.ac.in

[3] Department of Mathematics, School of Advanced of Sciences, Vellore Institute of Technology, Vellore, Tamil Nadu, India

kavitha@vit.ac.in

Abstract. Maternal health is among the greatest challenges in the world, especially in rural areas as there lack medical practitioners, they do not have easily accessible publics clinics and transport is difficult. Therefore, high rates of maternal as well as infant morbidity and mortalities are recorded. This research utilizes Artificial Intelligence (AI) with machine learning algorithms to forecast and address maternal health hazards right at their onset stage. The current research utilizes the concept of AI along with many Machine Learning (ML) methods like the Ensemble Learning Model (ELM), Random Forest (RF), K-Nearest Neighbour (KNN), Decision-Tree (DT), XG-Boost (XGB), Cat Boost (CB), and Gradient Boosting (GB), along with Synthetic Minority Over-sampling Technique (SMOTE) algorithm used for dealing with the problem class imbalance within the data set. SMOTE algorithm is utilized for the dataset balancing process. The handling system involves refining data preprocessing with the help of feature engineering and robust data cleaning which makes sure that anomalies do not erode the reliability of the predictive model. The existing methods [1] used RF (90%), DT (87%), XGB (85%), CB (86%), and GB (81%) algorithms and were compared with the accuracies of the proposed models like Logistic Regression (LR), Ensemble Learning Bagging (ELB), Ensemble Learning Stacking (ELS), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). The existing methods used only imbalance dataset. The accuracies of the proposed models with using SMOTE algorithm (balanced dataset) are LR (61.33%), KNN (81%), ELB (92.33%), ELS (90.66%) CNN (40.67%), RNN (59.67%), LSTM (54%), GRU (56%) respectively. Among these methods, ELB achieved 92.33% of accuracy with using SMOTE algorithm using imbalanced dataset. Whereas the accuracies of the proposed models without using SMOTE algorithm (imbalanced dataset) are LR (66.09%), KNN (68.47%), ELB (79.31%), ELS (82.26%), CNN (36.95%),

P. Pareek et al. (Eds.): IC4S 2024, LNICST 598, pp. 232–251, 2025.
https://doi.org/10.1007/978-3-031-77078-4_20

RNN (54.68%), LSTM (50.74%), and GRU (52.22%) respectively. Among these methods, ELS achieved 82.26% of accuracy using imbalanced dataset. The proposed model accuracies outperformed well using balanced dataset comparatively with existing models.

Keywords: Blood Sugar · SMOTE · XG Boost Classifier · Random Forest Classifier · Gradient Boost Classifier · Ensemble Learning · Bagging · Stacking · Logistic Regression · Convolutional Neural Network · Recurrent Neural Network · Long Short-Term Memory · Gated Recurrent Unit · Diabetes

1 Introduction

Globally, maternal health still poses as a notable issue of concern, with disparities manifested in maternal and infant outcomes among various areas. Accessibility of sufficient healthcare services is a challenge among many rural areas. Such challenges contribute to high cases of maternal morbidity and deaths. This challenge calls for creative ways of forecasting and managing maternal health risks, particularly in financially constrained regions.

There have been significant strides made towards enhanced maternal care practices through the application of some advanced techniques such as the Artificial Intelligence and Machine Learning Algorithms. These Machine Learning algorithms have capacity of analysing huge databases containing medical records that have been linked to maternal health concerns. With these observations, it is possible to create predictive models that will help in identifying high-risk pregnancies at first instance for better maternal welfare.

This research aims to investigate the potential of ML algorithms in predicting maternal health risks, focusing on Eight classification algorithms: A variety of models, namely Ensemble Model, Random-Forest, K-Nearest-Neighbors (KNN), Decision-Tree, Cat-Boosting, Gradient-Boosting, and Logistic-Regression Using along with SMOTE Algorithm. The SMOTE approach is used to tackle the problem of class imbalance in the dataset. Additionally, six characteristics are taken to categorize maternal health risk into three levels: High Risk, Low Risk, Mid Risk.

Health of the mothers is an imperative aspect under the area of public health, as it determines the state of mothers and their babies in general. Identifying and predicting maternal health risk is vital for improving maternity outcome. The last few years have seen tremendous advancements in computational methods for deep learning and traditional machine learning on health care.

The study involves extensive data on maternal health record, which cover varying ethnic groups, previous pregnancy conditions, and pregnancy-based aspects. Pre-processing of the dataset ensures that there is uniformity with regards to data quality. The SMOTE algorithm receives the dataset as input. Following that, there are more instances of the minority class. It helps to balance datasets, which enhances algorithm performance and avoids overfitting issues. To find a new point on the line segment by adding the random number to the feature vector, SMOTE usually entails finding a feature vector and its closest neighbour, measuring the difference between the two, multiplying it by a random number between 0 and 1, and repeating the process for all located feature vectors. Compared to copies that marginally deviate from the original data points, SMOTE offers the advantage of creating synthetic data points. The technique of SMOTE is applied to balance the class imbalance in which the amount to high-risk pregnancies is way too fewer compared to low-risk pregnancies. SMOTE creates additional minority classes samples thereby balancing and enhancing performances of the models that make prediction.

The main contribution of the proposed work are as follows:

1. The evaluation of eight classification algorithms for predicting maternal health risks using a pre-processed and balanced data set.
2. To evaluate the efficiency of SMOTE on balancing the data set and enhancing the prediction models' accuracy.
3. To apply feature Ensemble Machine Learning techniques for identifying the most significant characteristics related to maternal health risk.
4. To develop a predictive model that can accurately classify maternal health risk into three levels: utilizing their selected features and the best possible classification algorithm for High Risk, Low Risk, and Mid Risk models.

It is anticipated that study results will have far-reaching effects on the discipline of maternal health and low resourced setups in particular. Using an extensive data set to develop correct and trusted predictive models will enable healthcare providers to identify high risk pregnancies early enough thus allowing for the right interventions, that would in turn lead to improved maternal and neonatal outcomes.

2 Literature Survey

MD Assaduzzaman et. al. [1] addressed abnormalities in the data values, this research suggested enhancing data pretreatment methods involving feature engineering and data cleaning. Several ML methods, including CB, RF, XG-Boost, Decision-Tree, and Gradient-Boost, were employed to determine the maternal health risk factor. Hursit Burak MUTLU et. al. [2] suggested system six separate machine learning techniques were employed to assess the health risk of the mother. When the outcomes of different techniques were evaluated, it became clear that decision tree was the most effective approach for predicting the health risk of mothers.

Anika Rahman and Md. Golam Rabiul Alam [3] did various machine learning and deep learning algorithms were assessed, revealing that the Gradient Boosting algorithm demonstrated the highest accuracy in predicting risk levels. Taofeeq Oluwatosin Togunwa et. al. [4] suggested a vital component of community health that impacted the welfare of maternal and neonatal was Maternal-health. A deep-hybrid-model for predicting maternal health threat during conception was presented in this work. The two algorithms were combined in the suggested model to increase the precision and effectiveness of risk classification in expectant mothers.

Kausthav Pratim Kalita et. al. [5] aimed to forecast the risk level of pregnancy issues by combining blockchain technology with predictive data analytics through the use of an ensemble machine learning algorithm for data-collection, analysis, and storage related to maternal health in a secure and reliable manner. Additionally, the suggested system′s efficacy was assessed.

Aditi Ravi et. al. [6] investigated early diagnosis of the changes contributed to a lower death rate and a safer pregnancy period. Their primary goal was to create a primal prediction model by applying traditional machine learning techniques. Lokesh Pawar et. al. [7] suggested approach predicted the risk to the health of mothers using traditional machine learning algorithms. A strong model that functioned in the worst, average, and best cases was suggested.

Tizita Dereje et. al. [8] considered the best option for a home monitoring device because it was lightweight and extremely sensitive, even for slight movements. Continuous monitoring was not advised while using an ultrasound scan. They computed the normal and abnormal rates using the GSM module. They suggested a smartphone application to automatically track the health and activities of the fetus. B.Priyanka et. al. [9] considered forecast neonatal fatality using parental data and investigate risk variables. They used AdaBoost and bagging decision tree algorithms. Better results were obtained using a decision tree using the AdaBoost ensemble approach.

Marzia Ahmed and Mohammod Abul Kashem [10] considered a system for efficiently tracking and estimating a maternal risk assessment level was established, in the context of Bangladesh. Following comparisons between a few groups of the machine learning algorithm, the risk level for prediction and classification was determined. Lena Davidson and Mary Regina Boland [11] discussed the ways in which artificial intelligence (AI) had been used to treat preconception, prenatal, perinatal, and postnatal health issues for mothers. Nevertheless, there was a dearth of studies using AI techniques to investigate the effects of pharmaceutical therapies on pregnancy. The biggest gap in the literature that they found related to the use of AI techniques to maximize translational research, including drug exposures related to pregnancy in both humans and animals.

S.Shiny Amala and S.Mythili [12] used a wireless accelerometer sensor to measure fetal movement and vital parameters like temperature, heart rate, and blood pressure for rural pregnant women. The device used IoT and mobile phones for personalized care. Akhan Akbulut et al. [13] progressed a machine learning-based forecasting system for fetal congenital anomalies.

Aparna Gorthi et. al. [14] suggested using machine learning to identify the pregnancy risk group early on by looking for trends in profiles of well-established clinical markers. In particular, considering the decision-making process and the significance of individual characteristics were clearly represented in the tree, they showed the application of decision-trees for addressing complex issues in antenatal care. Christina Catley et. al. [15] aimed to develop a revamped strategy for initial stages preterm birth forecasting using artificial neural networks, with physician support, on heterogeneous maternal populations. The trained ANN classified new patient cases, identifying mothers at high risk of delivering premature infants, using an antenatal health decision-aid.

Maternal mortality in Nigeria was a significant issue due to long travel distances and a shortage of primary healthcare providers. A review of risk prediction models suggested providing a framework for healthcare providers to reduce maternal mortality rates. A parametric-based approach and risk assessment system, particularly for rural areas, were proposed by Imeh Jonah Umoren and Abasiama Silas [16].

Marzia Ahmed et. al. [17] used wearable sensing technology to monitor pregnant women in rural developing countries and it was showcased, aiming to reduce maternal and fetal mortality rates. The research used machine-learning algorithms to identify risk-levels based on conception risk-factors, aiming to achieve zero tolerance for maternal and fetal mortality. Sowmya K.N [18] monitored physiological characteristics and signs of the human body was essential for medical diagnosis and the health-care business. Many invasive and non-invasive methods were used nowadays to gather important data.

J. Schnarr and F. Smaill [19] used symptomatic and asymptomatic bacteriuria and they were common in pregnant women, with antibiotic treatment reducing low birth weight. There was no consensus on antibiotic choice or therapy duration. With increasing antibiotic resistance, local resistance rates were crucial. Rabiatu Sageer et. al. [20] used most maternal mortality investigations in Nigeria were individual accounts originating from a sole healthcare institution. Preventive measures and adequate care could have avoided many contributory factors of maternal mortality. MPDSR provided critical evidence and information on prevention strategies. Implementation was underway in Ogun State, but it should have been institutionalized in all Nigerian states. Commitment to MPDSR findings was crucial for success.

Naciye Nur ARSLAN et. al. [21] used a Temporal Convolutional Network (TCN) based Bidirectional Long Short Term Memory network (BiLSTM) network and machine learning algorithms to classify low, medium, and high-risk conditions during pregnancy. The dataset included maternal risk factors like age, blood-pressure, blood-sugar, and heart rate. Ali Raza et. al. [22] studied a deep neural network-based system called DT-BiLTCN to predict pregnancy-related health hazard using health data records. The system used Decision-Trees, a dual-directional LSTM-network, alongside a temporal convolution network. The model identified diastolic and systolic blood-pressure, heart-rate and maternal-age age as the strongest indicators of pregnancy health risks.

Mst Irin Sultana et. al. [23] used a separate data repository, their goal was to find additional risk variables and detect characteristics associated to the maternal death rate in Bangladesh that have been found in numerous research. Akhan Akbulut et al. [13] aimed to improve fetal congenital anomalies prediction using machine-learning and e-Health applications. They created a forecasting system with supportive e-Health applications for expectant mothers and healthcare professionals. The planned project aimed to provide assistive services to expectant mothers and medical practitioners through an online system with mobile and web applications.

João Alexandre Lobo Marques et. al. [24] proposed the construction of an integrated solution based on Data-Analytics for feature extraction, smart diagnostic support system utilizing a 1-D convolutional neural network (CNN) classifier, and IoT sensors for monitoring maternal and fetal markers in at-risk pregnancies. Taking into account six potential outcomes, a classification technique was suggested as a forecasting system for simultaneous maternal and fetal-health position categorization.

Md. Alif Sheakh et. al. [25] provided a risk factor analysis utilizing machine learning techniques to lower infant and mother mortality. This work assessed Eight machine learning algorithms as part of the investigation. Three metrics were used to evaluate the effectiveness of various classification algorithms: recall, accuracy, and precision. The main goal of the current study was to use clinical data to estimate the risk factor for maternal and infant death.

Erly Krisnanik et. al. [26] used Panimbang Health Center risk criteria and they identified four criteria for pregnancy risk: underlying health conditions, complex pregnancies, unfavorable obstetric history, and compromised maternal condition. A system was developed using descriptive, forecastive, and prescriptive data analysis to provide recommendations based on symptoms experienced by pregnant women. The system aimed to help midwives anticipate risk levels and reduce mortality rates for pregnant women and fetuses. Z. Comert and A.F. Kocamaz [27] identified the best Machine-Learning method for categorizing fetal-heart-rate signals. As a result, the study concentrated on popular and useful machine learning approaches, including random forest, radial basis function network, Support-Vector Machine, Artificial-Neural-Network, and Extreme-Learning-Machine.

Y. P. Sinha et al. [28] offered a useful and feasible contextual care protocol approach because this was also how doctors diagnose in real life. Monitoring the patient's changing symptoms might lead to a faster and more precise diagnosis. It was also self-correcting because it gave medical practitioners a way to provide comments and suggest changes). Chenai Mlandu et. al. [29] identified the most important variables in three countries—the Democratic Republic of the Congo, Kenya, and Tanzania—and to forecast the probability of a mother or child leaving the MNCH continuum by using dependable machine learning predictive models.

3 Dataset Description and Sample Data

Uniform Resource Locator (URL) for data-set is given below.

https://archive.ics.uci.edu/dataset/863/maternal+health+risk.

Many hospitals and community clinics as well as a number of maternal health care providers located in various parts of Bangladesh's rural areas have been connected through an IOT based risk monitoring system. The attributes that are measured include age (in years), body temperature (body-temp in Fahrenheit), heart rate (in beats per minute), Blood-Sugar (BS in mmol/L), Diastolic-Blood-Pressure (BP in mmHg), and Risk-Level. These will be considered as the major risk factors that are responsible for maternal mortality, which is one of the UN′s Sustainable Development Goals. The details are given below.

- Number of Attributes – 7 (6 Independent & 1 Dependent)
- Number of classes in Dependent attribute column – 3 classes
- Before Applying SMOTE algorithm (Total Number of Records – 1014)
- After Applying SMOTE algorithm (Total Number of Records – 1497)

The original dataset has an uneven distribution of class labels, meaning one class has significantly more data points than the other. The SMOTE algorithm is used to balance this unevenness. This methodology increases the number of records belonging to minority category using fake samples produced from real minority class information. Using oversampling technique to balance the dataset helps to enhance the efficiency of machine learning models.

The method used in SMOTE is locating the KNN of any minority class point and synthesizing new points across the line that connects those k nearest points. The filling of these vacancies assists is equalizing the minority group's representation and ultimately makes the data set balanced. Now, after using SMOTE Algorithm each class or each type of Risk level is either having 499 records which combines to 1497 Records in total, where before SMOTE the total records used to be 1014 with varying number of records for each Risk level.

In the Table 1, the SMOTE algorithm was effective in balancing the dataset. After applying the algorithm, there are now 499 records in each class. This should make it easier to train machine learning models that are accurate across all classes.

Table 1. Size of dataset with and without applying SMOTE Algorithm

Risk Level	Imbalanced Dataset (Before applying SMOTE algorithm)	Balanced Dataset (After applying SMOTE algorithm)
Low Risk	272	499
Mid Risk	406	499
High Risk	336	499
Total	1014	1497

In the Table 2, it is about the different variables that are used to predict the Risk-Level of a pregnant women.

Table 2. Sample dataset

Age	Systolic-BP	Diastolic-BP	BS	Body Temp	Heart Rate	Risk Level
25	130	80	15	98	86	high risk
35	140	90	13	98	70	high risk
29	90	70	8	100	80	high risk
30	140	85	7	98	70	high risk
35	120	60	6.1	98	76	low risk
23	140	80	7.01	98	70	high risk
23	130	70	7.01	98	78	mid risk
35	85	60	11	102	86	high risk
32	120	90	6.9	98	70	mid risk
42	130	80	18	98	70	high risk

The correlation matrix in the diagram shows the correlation between the seven variables in the dataset you provided: include age, body temperature (body-temp), heart rate, Blood-Sugar (BS), Diastolic-Blood-Pressure (BP), and Risk-Level. The correlation coefficients are shown in the boxes at the intersections of the rows and columns.

A correlation matrix is a square matrix that shows the correlation between each pair of variables in a data-set. The correlation between two variables is a measure of how strongly related they are. Heatmap representation of dataset attributes (before and after applying SMOTE algorithm) are shown in the Fig. 1 and Fig. 2.

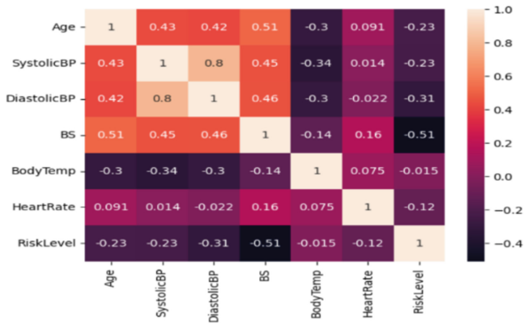

Fig. 1. Heatmap Representation of Dataset Attributes (Before SMOTE algorithm)

Correlation matrix show that there are strong positive correlations between the following pairs of variables:

- Age and Risk Level
- Systolic BP and Risk Level
- Diastolic BP and Risk Level
- BS and Risk Level

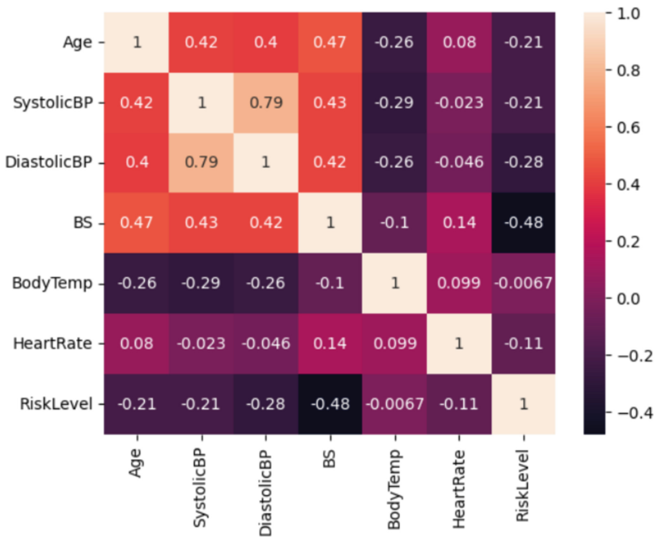

Fig. 2. Heatmap Representation of Dataset Attributes (After SMOTE algorithm)

4 Methodology

The work flow for predicting maternal health risk is shown in the Fig. 3. The pipeline begins with preparing and collecting the data. This information is subsequently partitioned into Two portions; one for Training, and another for Testing. A number of machine learning algorithms are trained with the available training data. The trained models are finally put through testing against the test-data in order to choose a model for the specific objective.

There are ten steps involved in the proposed work and they are given below.

Step 1: Load data-set	Step 6: Data splitting
Step 2: Perform data pre-processing	Step 7: Model building
Step 3: Identify and handle imbalanced classes	Step 8: Model testing
Step 4: Data visualization	Step 9: Performance measures calculation for every model
Step 5: Scaling dataset values	Step 10: Accuracy comparison and best model declaration

Step 1 - Load Data-set: Insert the dataset into a Pandas Data Frame, which is an excellent construct of a framework meant for data manipulation or even scrutiny.

Step 2 - Perform Data Pre-processing: The next step is to clean the data if any possibility is needed which makes it suitable for machine learning. These could comprise of using label encoder to label the outcomes obtain from the model.

Step 3 - Identify and Handle Imbalanced Classes: The balance class is identified in these data and necessary measures are taken. A typical strategy involves using methodologies like oversampling (SMOTE) to raise the number of data instances within the minor group.

Step 4 - Data Visualization: This work uses Matplot.lib library to visualize the relationship existing between dependent variables and understand how they related with each other. The pairwise relationships between variables within a dataset is shown in the Fig. 4.

Step 5 - Scaling Dataset Values: In this process, we will scale the data (using minmax scaler) so that all the features have a similar scale. Machine learning algorithms, however, can be enhanced by this.

Step 6 - Data Splitting: This way, they divide data into training and test samples. The Machine Learning Models are trained using the training Data and tested using the test Data. The data is usually Partitioned into 80% for training, and 20% for testing.

Step 7 - Model Building: It involves choosing a machine learning algorithm, and training it using a training set. Many algorithms of machine learning exist, each having their advantages and disadvantages. Examples of such algorithms are Ensemble Learning Model, Random-Forest, KNN, Decision-Tree, CB, gradient boosting, and neural network.

Step 8 – Model Testing: The next step entails testing the trained Machine Learning model on a separate set of Data. Here, it makes forecasts on test set which are compared with the true class label.

Step 9 - Performance Measures Calculation for Every Model: During this stage, the trained Machine Learning model's accuracy in the case of test set is measured. Accuracy, Precision, F-Measure, Recall the measure of how well the model estimates or predicts the target output associated with a particular data point.

The performance measures used for the proposed works are accuracy, precision, recall and F-Measure respectively. The accuracy rate is the number of correct predictions out of all the predictions made by model and it is shown in the Eq. 1.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{1}$$

The precision is the percentage of all instances predicted as positive by the model that were actually correct in being classified as so and it is shown in the Eq. 2.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2}$$

The recall is the percentage of correctly predicted positive instances over all true positives and it is shown in the Eq. 3.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{3}$$

The F-Measure is the geometric mean of precision and recall, a balanced measure of the two metrics and it is shown in the Eq. 4.

$$\text{Precision} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} \pm \text{Recall}} \tag{4}$$

Step 10 - Accuracy Comparison and Best Model Declaration: The last step entails ranking of all the acquired trained machine learning models for their precision and choosing the most appropriate model. The best model is the one with the highest accuracy in the test set.

5 Results and Discussion

Various machine learning algorithms and deep learning algorithms were used to classify maternal health risk factors including XG-Boost, G-Boost, Random-Forest, Decision-Tree, KNN, also the dataset is used in two ways: one imbalanced version and one balanced version using SMOTE. Graphs are also displayed, and it shows that the models with a balanced data set have higher accuracies. Train and test data split are 80% and 20%. Random state = 40 is used in train-test split.

Parameters used in the implementation of building the proposal models are shown in the Table 3.

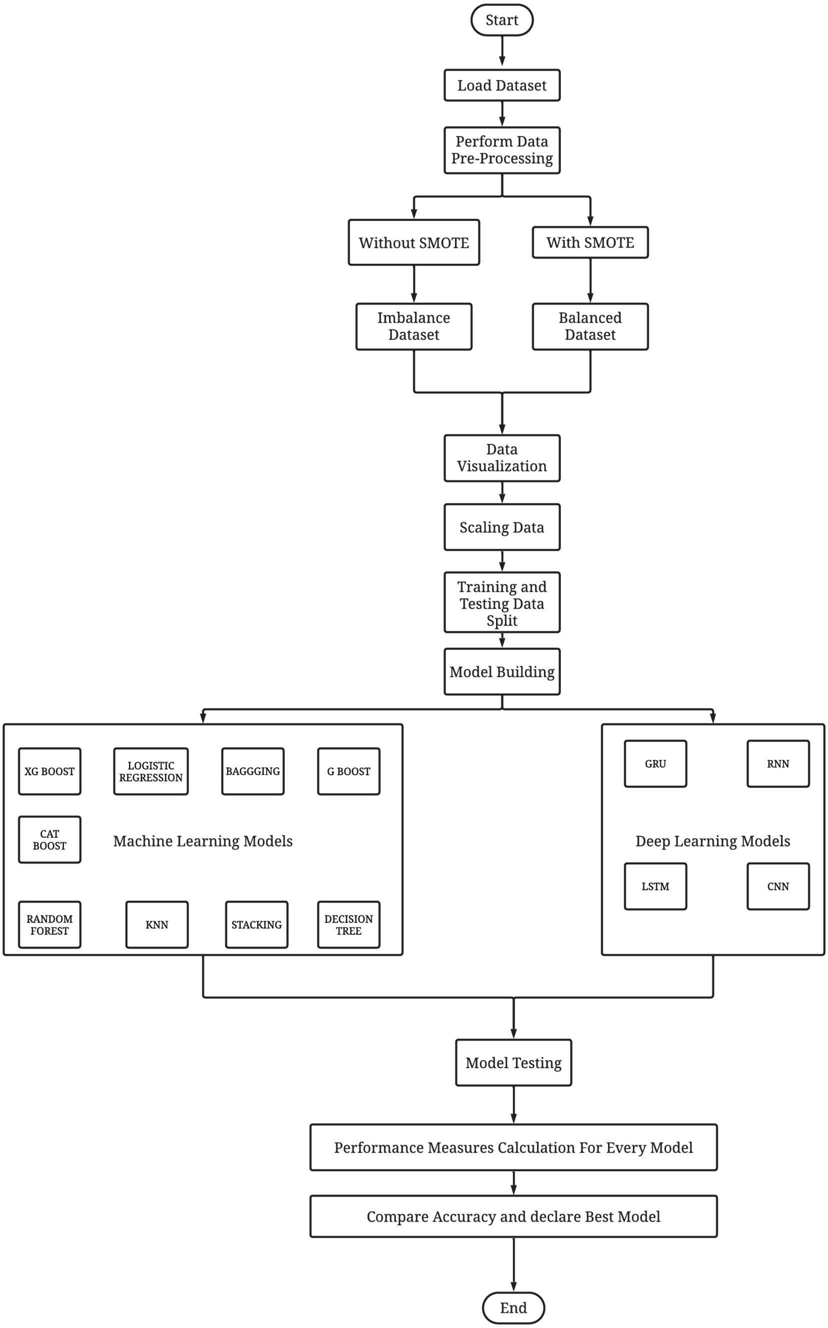

Fig. 3. The work flow for predicting maternal health risk

Fig. 4. The pair wise relationship between variables

There is variation in accuracy when dealing with an imbalanced data-set without using SMOTE algorithm for Maternal Health Risk. Various machine learning techniques and deep learning techniques are employed and their accuracies are XGB (81.77%), GB (73.9%), RT (78.81%), DT (79.8%), KNN (68.47%), LR (66.09%), CB (79.80%), ELB (79.31%), ELS (82.26%), CNN (36.95%), RNN (54.68%), LSTM (50.74%), GRU (52.22%) respectively. Among these models, ELS algorithm gives the highest accuracy of 82.26%. The performance measures like accuracy, precision, recall and F1-score of the various classifiers are compared when dealing with an imbalanced dataset without using the SMOTE algorithm for Maternal Health Risk are shown in the Table 4.

The performance measures like accuracy, precision, recall and F1-score of the various classifiers are compared when dealing with a balanced dataset using the SMOTE

Table 3. Parameters applied to various models

Models	PARAMETERS FOLLOWED
XGB	Random state = 40
GB, KNN & LR	Nil
RF	Random state = 40, estimators = 100, criterion = entropy
DT	Random state = 40, criterion = entropy
ELB	final_estimator = RandomForestRegressor(n_estimators = 100, random_state = 40) voting ='soft'
ELS	final_estimator = RandomForestClassifier(n_estimators = 100, random_state = 40)
CNN, GRU, LSTM, RNN	loss = "binary_crossentropy", optimizer = "adam"
CB	Learning_rate = 0.3, depth = 6, l2_leaf_reg = 3, iterations = 100

Table 4. Accuracy results comparison for imbalanced dataset

Models	Accuracy	Precision	Recall	F1-score
XGB	81.77%	82%	82%	82%
GB	73.9%	74%	74%	74%
RF	78.81%	79%	79%	79%
DT	79.8%	80%	80%	80%
KNN	68.47%	69%	68%	68%
LR	66.09%	66%	65%	64%
CB	79.80%	80%	80%	80%
ELB	79.31%	80%	79%	79%
ELS	82.26%	82%	82%	82%
GRU	52.22%	43%	50%	42%
RNN	54.68%	47%	53%	45%
LSTM	50.74%	45%	49%	41%
CNN	36.95%	12%	33%	18%

algorithm for Maternal Health Risk are shown in the Table 5. Various machine Learning techniques and deep learning techniques, are employed and their accuracies are GB (84.67%), XGB (90.33%), RT (91.33%), DT (87.67%), LR (61.33%), CB (89%), KNN (81%), ELB (92.33%), ELS (90.66%) CNN (40.67%), RNN (59.67%), LSTM (54%), GRU (56 respectively. Among these models, ELB algorithm gives the highest accuracy of 92.33%.

The Table 6 compares the accuracy levels of the various models before and after applying SMOTE algorithm. Out of 13 models used, 10 model accuracies are increased,

Table 5. Accuracy results comparison for balanced data-set using SMOTE Algorithm

Models	Accuracy	Precision	Recall	F1-score
XBG	90.33%	90%	89%	90%
GB	84.67%	83%	83%	83%
RF	91.33%	91%	91%	91%
DT	87.67%	87%	87%	87%
KNN	81%	80%	80%	80%
LR	61.33%	60%	60%	60%
CB	89%	88%	88%	88%
ELB	92.33%	92%	92%	92%
ELS	90.66%	90%	90%	90%
GRU	56%	46%	54%	45%
LSTM	54%	46%	52%	43%
RNN	59.67%	47%	56%	48%
CNN	40.67%	44%	41%	30%

Table 6. Accuracy comparison of various classifiers using imbalanced and balanced dataset

MODELS	Accuracy of IMBALANCED Data-Set	Accuracy of BALANCED Data-Set
XBG	81.77%	90.33%
GB	73.90%	84.67%
RF	78.81%	91.33%
DT	79.80%	87.67%
KNN	68.47%	81%
LR	66.09%	61.33%
CB	79.80%	89%
ELB	79.31%	92.33%
ELS	82.26%	90.66%
GRU	52.22%	56%
LSTM	54.68%	54%
RNN	50.74%	59.67%
CNN	36.95%	40.67%

2 model accuracies are decreased, and 1 model get same value after applying SMOTE algorithm to the workflow. So, the purpose of SMOTE Algorithm**,** which is accustomed to

balance dataset, significantly enhanced the proposed model's quality and the accuracies are considered to be final and accepted for this system.

In the Fig. 5, it represents the bar plot for all the model accuracies and gives a comparative result of the proposed system using imbalanced and balanced dataset and the knowledge to be inferred is that 8 out of 9 models accuracies have been increased after applying SMOTE algorithm.

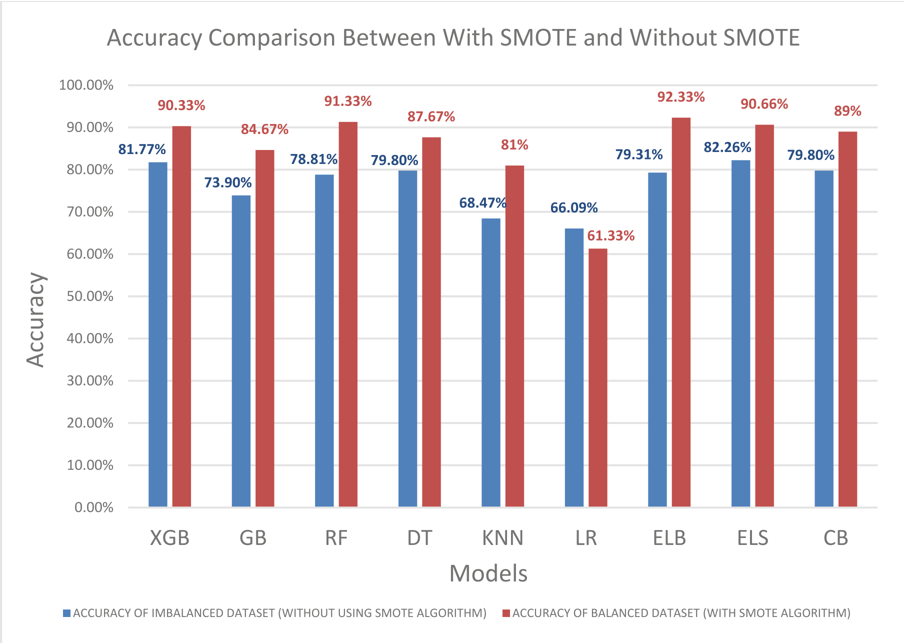

Fig. 5. Comparison of accuracy values for the proposed system using imbalanced and balanced dataset

The Table 7 compares the accuracies of existing and proposed models. The proposed models have outperformed their accuracy results using balanced dataset. They are GB (84.67%), XGB (90.33%), RF (91.33%), DT (87.67%), LR (61.33%), CB (89%), KNN (81%), ELB (92.33%), ELS (90.66%), CNN (40.67%), RNN (59.67%), LSTM (54%), and GRU (56%).

In the existing model, the maximum accuracy is gained by using RF with 90%, but in the proposed model, the maximum accuracy is gained by using Ensemble Learning Bagging (RF and XGB) with 92.33% outperformed that benchmarked accuracy of existing model.

In the Fig. 6, it represents the bar plot for all the model accuracies (XGB, GB, RF, DT, CB) and gives a comparative result of the existing and proposed systems using imbalanced and balanced dataset. In the Fig. 7, it represents confusion matrices for the proposed systems like LR, KNN, ELB, ELS, CNN, RNN, LSTM and GRU models using balanced dataset.

Table 7. Accuracy comparison between existing and proposed model accuracies

MODELS	Existing Model Accuracy	Proposed Model Accuracy (using balanced dataset)
XBG	85%	90.33%
GB	81%	84.67%
RF	90%	91.33%
DT	87%	87.67%
KNN	86%	81%
LR	-	61.33%
CB	-	89%
ELB	-	92.33%
ELS	-	90.66%
GRU	-	56%
LSTM	-	54%
RNN	-	59.67%
CNN	-	40.67%

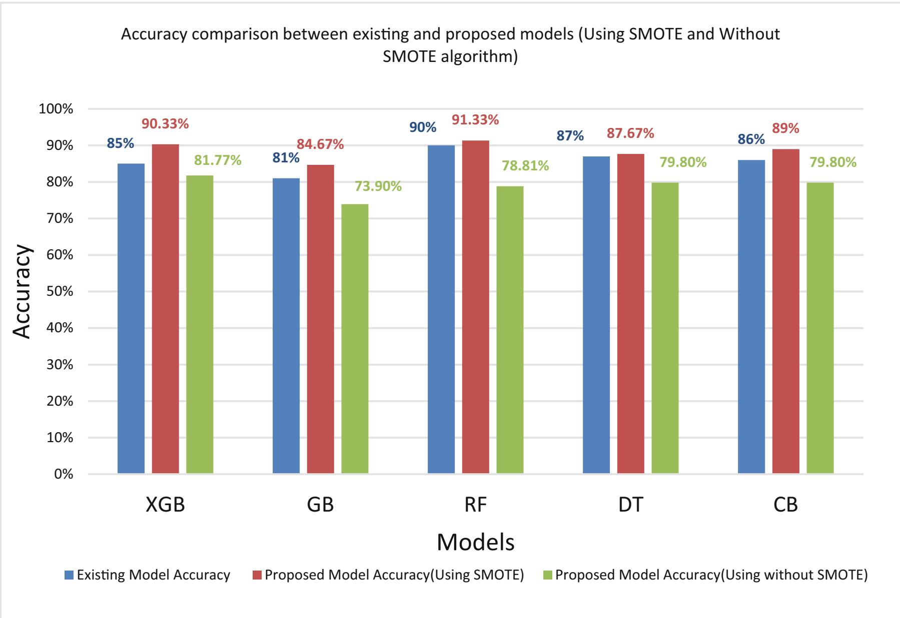

Fig. 6. Accuracy comparison between existing and proposed models (Using SMOTE and Without SMOTE algorithm)

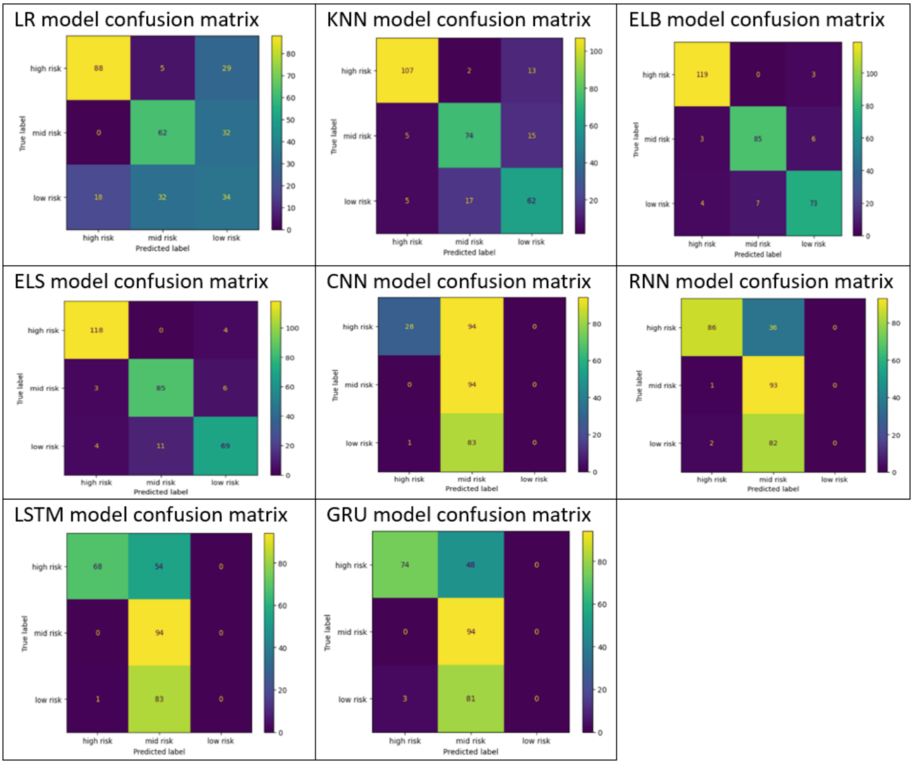

Fig. 7. The confusion matrices for the proposed systems using balanced dataset

6 Conclusion and Future Work

The present investigation assessed the efficacy of SMOTE algorithm in conjunction with various machine learning and deep learning algorithms in forecasting maternal peril. The accuracies of the proposed models with using SMOTE algorithm are LR (61.33%), KNN (81%), ELB (92.33%), ELS (90.66%) CNN (40.67%), RNN (59.67%), LSTM (54%), and GRU (56%) respectively. Whereas the accuracies of the proposed models without using SMOTE algorithm are LR (66.09%), KNN (68.47%), ELB (79.32%), ELS (82.26%), CNN (36.95%), RNN (54%), LSTM (50%), and GRU (52%) respectively. The proposed model accuracies outperformed well using balanced dataset comparatively with existing models. The combination of SMOTE algorithm and ELB (92.33% of accuracy), proved itself as the most accurate one. It scored better by compared to ML models and DL models applied on imbalanced (without using SMOTE algorithm) and balanced (using SMOTE algorithm) dataset involved in predicting maternal health risks.

The future works are given below.

- **Explore different data balancing techniques:** Henceforth, another data balancing technique like ADASYN and Near-miss need to be studied to decide on the best data balancing strategy for maternal health risk prediction.

- **Incorporate additional risk factors:** Adding more confounders like socio-economic status, environmental exposure and genetics can enhance prediction for the enlarged data set.
- **Explore deep learning Ensemble approaches:** In addition, there is high level of accuracy among other complex deep learning techniques such as Ensemble neural networks which might be applied in maternal health risk prediction.
- **Incorporate explainable AI methods:** Develop explainable design AI approaches that would be able to extract the predictions of the ML models to help clinicians under the conditions increasing maternal risk.

References

1. Assaduzzaman, M., Mamun, A.A, Hasan, M.Z.: Early prediction of maternal health risk factors using machine learning techniques. In: 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, pp. 1–6. IEEE (2023)
2. Mutlu, H.B., Durmaz, F., Yücel, N., Cengil, E., Yıldırım, M.: Prediction of maternal health risk with traditional machine learning methods. MTU J. Eng. Nat. Sci. **4**(1), 16–23 (2023)
3. Rahman, A., Alam, M.G.: Explainable AI based maternal health risk prediction using machine learning and deep learning. In: 2023 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, pp. 0013–0018. IEEE (2023)
4. Togunwa, T.O., Babatunde, A.O., Abdullah, K.U.: Deep hybrid model for maternal health risk classification in pregnancy: synergy of ANN and random forest. Front. Artif. Intell. **6**, 1213436 (2023)
5. Kalita, K.P., Chettri, S.K., Deka, R.K.: A blockchain-based model for maternal health information exchange and prediction of health risks using machine learning. In: 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, pp. 1184–1189. IEEE (2023)
6. Ravi, A., Joshi, S.P., Kodipalli, A., Kamal, S.: Analysis of maternal health risk using computational models. In: International Conference on Smart and Sustainable Technologies in Energy and Power Sectors (SSTEPS). Mahendragarh, India, pp. 311–315. IEEE (2022)
7. Pawar, L., Malhotra, J., Sharma, A., Arora, D., Vaidya, D.: A robust machine learning predictive model for maternal health risk. In: International Conference on Electronics and Sustainable Communication Systems (ICESC 2022). Coimbatore, India, pp. 882–888. IEEE (2022)
8. Dereje, T., Abuhay, T.M., Letta, A., Alelign, M.: Investigate risk factors and predict neonatal and infant mortality based on maternal determinants using homogenous ensemble methods. In: International Conference on Information and Communication Technology for Development for Africa (ICT4DA) 2022. Bahir Dar, Ehiopia. pp. 18–23. IEEE (2021)
9. Priyanka, B., Kalaivanan, V.M., Pavish, R.A., Kanageshwaran, M.: IOT based pregnancy women health monitoring system for prenatal care. In: 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 1264–1269. IEEE (2021)
10. Ahmed, M., Kashem, M.A.: IoT based risk level prediction model for maternal health care in the context of Bangladesh. In: 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2020, pp. 1–6. IEEE (2020)
11. Davidson, L., Boland, M.R.: Enabling pregnant women and their physicians to make informed medication decisions using artificial intelligence. J. Pharmacokinet. Pharmacodyn. **47**, 305–318 (2020)

12. Amala, S., Mythili, S.: IoT based health care monitoring system for rural pregnant women. Int. J. Pure Appl. Math. **119**(15), 837–843 (2018)
13. Akbulut, A., Ertugrul, E., Topcu, V.: Fetal health status prediction based on maternal clinical history using machine learning techniques. Comput. Methods Programs Biomed. **163**, 87–100 (2018)
14. Gorthi, A., Firtion, C., Vepa, J.: Automated risk assessment tool for pregnancy care. In: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, pp. 6222–6225. IEEE (2009)
15. Catley, C., Frize, M., Walker, R.C., Petriu, D.C.: Predicting high-risk preterm birth using artificial neural networks. In: IEEE Transactions on Information Technology in Biomedicine **10**(3), 540–549 (2006)
16. Umoren, I., Chigozirim, F., Silas, A., Ekong, B.: Modeling and prediction of pregnancy risk for efficient birth outcomes using decision tree classification and regression model 12–26 (2021)
17. Ahmed, M., Kashem, M.A., Rahman, M., Khatun, S.: Review and analysis of risk factor of maternal health in remote area using the internet of things (IoT). In: Kasruddin Nasir, A.N., et al. In ECCE2019. Lecture Notes in Electrical Engineering, vol. 632. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-2317-5_30
18. Rai, S.K., Sowmya, K.: A review on use of machine learning techniques in diagnostic health-care. CiiT Int. J. Artif. Intell. Syst. Mach. Learn. **10**(4), 102–107 (2018)
19. Schnarr, J., Smaill, F.: Asymptomatic bacteriuria and symptomatic urinary tract infections in pregnancy. Eur. J. Clin. Invest. **38**(S2), 50–57 (2008)
20. Sageer, R., Kongnyuy, E., Adebimpe, W.O., Omosehin, O., Ogunsola, E.A., Sanni, B.: Causes and contributory factors of maternal mortality: evidence from maternal and perinatal death surveillance and response in Ogun state. Southwest Niger. BMC Pregnancy Childbirth **19**(63), 1–8 (2019)
21. Arslan, N., Arslan, M., Sahin, E.: Automated Classification of maternal risks in pregnancy: Anal. Using Mach. Learn. Artif. Neural Netw. (2023)
22. Raza, A., Siddiqui, H.U., Munir, K., Almutairi, M., Rustam, F., Ashraf, I.: Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction. PLoS ONE **17**(11), 1–29 (2022)
23. Sultana, M.I., Lovely, M.L.S., Hasan, M.M.: Building prediction models for maternal mortality rate in Bangladesh. In: 2019 5th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, pp. 375–380. IEEE (2019)
24. Marques, J.A.L., et al.: IoT-based smart health system for ambulatory maternal and fetal monitoring. IEEE Internet Things J. **8**(23), 16814–16824 (2021)
25. Sheakh, M.A., Tahosin, M.S., Hasan, M.M., Islam, T., Islam, O., Rana, M.M.: Child and maternal mortality risk factor analysis using machine learning approaches. In: 11th International Symposium on Digital Forensics and Security (ISDFS), Chattanooga, TN, USA, pp. 1–6. IEEE (2023)
26. Krisnanik, E., Tambunan, K., Irmanda, H.N.: Analysis of pregnancy risk factors for pregnant women using analysis data based on expert system. In: International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, pp. 151–156. IEEE (2019)
27. Cömert, Z., Kocamaz, A.F.: Comparison of machine learning techniques for fetal heart rate classification. Acta Phys. Pol. A **132**(3), 451–454 (2017)

28. Sinha, Y.P., Malviya, P., Panda, M., Ali, S.M.: Contextual care protocol using neural networks and decision trees. In: Second International Conference on Advances in Electronics, Computers and Communications (ICAECC), Bangalore, India, pp. 1–4. IEEE (2018)
29. Mlandu, C., Matsena-Zingoni, Z., Musenge, E.: Predicting the drop out from the maternal, newborn and child healthcare continuum in three East African community countries: application of machine learning models. BMC Med. Inf. Decis. Making **23**(191), 1–10 (2023)