Ensemble Based Machine Learning Model for Early Detection of Mother's Delivery Mode

Conference Paper · February 2023

DOI: 10.1109/ECCE57851.2023.10101558

6 authors, including:

Mahmudul Hasan
Deakin University

Md Jakaria Zobair
Daffodil International University

Sumya Akter
Lecturer at Hajee Mohammad Danesh Science and Technology University

Mahir Ashef

Ensemble Based Machine Learning Model for Early Detection of Mother's Delivery Mode

Mahmudul Hasan
Department of CSE
HSTU, Dinajpur-5200, Bangladesh
mahmudulmoon123@gmail.com

Md Jakaria Zobair
Department of CSE
DIU, Dhaka-1341, Bangladesh
jakariacse16@gmail.com

Sumya Akter
Department of CSE
HSTU, Dinajpur-5200, Bangladesh
sumya.hstu@gmail.com

Mahir Ashef
Department of CSE
BUBT, Dhaka-1216, Bangladesh
mahirasef30@gmail.com

Nazrin Akter
PGD in Data Analytics
Purdue University, USA
nazrinuddin@gmail.com

Nahid Binte Sadia
Department of ICT
Ministry of Post, Telecommunications and IT, Bangladesh
nahidtee10@gmail.com

2023 International Conference on Electrical, Computer and Communication Engineering (ECCE) | 979-8-3503-4536-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ECCE57851.2023.10101558

***Abstract*—The mother's mode of delivery greatly impacts the relationship between the newborn baby and the mother, as well as the mother's and baby's health. Currently, the cesarean rate is increasing at an alarming rate. The inability to predict the mother's health status and mode of delivery are mainly responsible for this situation. Support Vector Machine (SVM), Decision Tree, Random Forest (RF), Gradient Boosting Classifier(GBC), Logistic Regression, Gaussian Naive Bayes, Stochastic Gradient Descent, CatBoost (CB), Adaptive Boosting (AB), Gaussian Na ıve Bayes, Extreme Gradient Boosting(XGB) are used to predict the mother's mode of delivery. This study also proposed an ensemble machine learning algorithm that stacked the SVC, XGB, and RF together and named the ensemble SVXGBRF. To preprocess the dataset, we use a pipeline that basic preprocessing techniques, data balancing and feature selection. Our proposed SVXGBRF classifiers show 95.52% accuracy, 96% precision, recall, f1 score, and 99% AUC score. SVXGBRF shows its superiority, where most models show an accuracy of less than 90% except RF, GBC, CB, and AB. Eventually, this research could be utilized to develop a decision-support system for reducing the number of cesarean sections by trying to extract insights from complex data patterns.**

***Index Terms*—Mother's Mode of Delivery, machine learning, ensemble learning**

I. INTRODUCTION

Mother's delivery rules are changing day by day due to the evaluation of modern technologies. Now-a-days there are a lots of delivery modes such as Vaginal delivery, Natural childbirth, Caesarean Section (CS), Forceps delivery, Vacuum Extraction, Vaginal birth after caesarean (VBAC) [1] and so on. Types of delivery modes depend on mother's characteristics as well as child. Natural childbirth is more popular among the others. On the other hand, in Cesarean section, also known as C-Section, delivery is done by opening up mother's abdomen and surgically removes the baby by opening the uterus. The rates of CS are increasing day by day commensurately in Bangladesh as well as around the world. In one study, they showed that the unnecessary C-Section increases about 51% in the year 2016 to 2018 where the rate was only 4% in the year 2004 [2]. According to the WHO (World Health Organization), uses of C-Section for delivering of babies were almost double from the year 2000 to the year 2015 which was about 12% to 21% of all deliveries [3]. This unnecessary C-Section increases the rate of maternal death. Again according to the WHO, CS increases the chances of maternal mortality. It also affects to the future pregnancies and mothers can have some problems such as infection, bleeding, and surgical injury to the bladder and so on [4]. It also creates complications to a child. Children who are delivered by C-Section may have the berating issues, lung problems. In this respect, CS plays important role as a lifesaver method. The decisions come from the medical professionals whether C-Section delivery is needed or not. Sometimes they confused to take decisions. Apart from these, the family suffers from a lot by the cost of this C-Section. The clinics changes about 22,085 BD-TK or $ 276 in the year 2010 whereas the cost of a normal delivery was 3,565 BD-TK or $ 45 [5]. But, sometimes complications arise and natural childbirth is not possible.

Machine learning (ML) works with hidden pattern of some data, understand the pattern and discovers valuable information from the data. Acutely it creates the relationship between dependent and independent variables. ML is successfully applied in many sectors such as Health care, Retail, Manufacturing, Banking and finance, Transportations and so on [6]. Especially in Health care sector ML is a blessing one. With the help of ML, diverse problems are solving and increasing the success of experiments. ML is applied in diseases predictions [7] [8], forecasting epidemic [9], predicting autism [10] and so on. With the simulation of ML our life becomes easier. In this respect, Machine Learning (ML) can play a significant role. Many researchers have proposed different techniques to predict the modes of delivery as well as the risks of those modes which is discussed in the literature review section.

In this research, we propose a preprocessing pipeline that accommodates missing values, corrects erroneous data, corrects wrongly formatted data, normalizes the data, selects relevant features, and deals with data imbalance problems. We train some machine learning classifiers by this preprocessed dataset. In addition, we proposed an ensemble machine learning model to enhance the classification performance of identifying the mother's mode of delivery at an early stage.

II. Review of Literature

A lot of research is being done using machine learning techniques for biomedical decision-making systems which use key information from the different hideous patterns in historical data. Here we will discuss some models and methods which are relevant to our work. M. Sakib et al. [11] implemented a Bagging Ensemble Classifier based on traditional machine learning algorithms, which is a novel approach for predicting the mood of childbirth. They examine the performance of four ML algorithms, with bagging ensemble classifiers, and used 4493 cross-sectional data. They found bagging classifiers work well for identifying the key factors which influence Caesarean Section (CS). In another paper, Annes et al. [12] implemented three prognostic models for complications of planned c-section which identified, and predicted blood transfusion, spinal hypotension, and postpartum hemorrhage. Michal et al. [13] examine the feasibility of using machine learning methods for predicting a successful normal birth after cesarean delivery. They used two models on 9888 patient's data, one model for providing a personalized risk score for normal birth after cesarean delivery and another for reassessing that score after adding available features. Again S. Ali et al. [14] use bagging and boosting classification algorithms on birth data. BagFda is the very classification model for their birth data and their training and testing accuracy is 94.53% and 93.44% respectively. Nafiz et al. [15] develop three different ensemble prediction models (XGBoost, AdaBoost, and Catboost) to predict whether the cesarean section is necessary or not. XGBoost gave the highest accuracy 88.91% whereas AdaBoost and Catboost showed 88.69% and 87.66% accuracy respectlively. In another study C. Campillo et al. [16] fitted a classification tree (CTREE or conditionally unbiased inference classification tree) and analyzed a random forest model to increase the discriminatory accuracy (DA) using 157 birth data. In [17], they built a supervised machine learning-based decision-making model to predict the most significant labor mode which will reduce the maternal and infant health risk. The analysis of quadratic discriminant showed the highest accuracy of 0.979992 with the F1 score of 0.979962. S. Maroufizadeh et al. [18] used Classification methods to identify factors that were related to Cesarean Section. They used 2120 primiparas who gave singleton birth, such as logistic regression (LR), random forest (RF), and artificial neural network (ANN).The ANN method showed the best performance which classified CS delivery compared to the other two methods. P. Fergus et al. [19] discuss about the importance of machine learning in medical sectors such as, cesarean section(CS) is acceptable or not and help to avoid the perinatal deaths.

III. Materials and Methods

A. Overview of proposed methodology

Including data preprocessing, model development, and performance measurement is a top-down process of our proposed methodology. Details of the steps are described one by one. Fig. 1 shows individual parts separately and according to the flow. To use the inter dataset for training and testing purposes, we use 10-fold cross-validation.

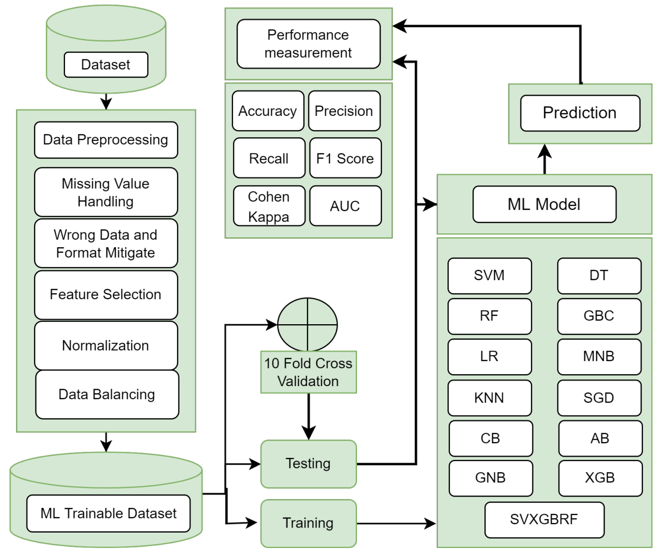

Fig. 1. Overview of proposed methodology

The dataset is highly noise and need a long preprocessing on it. We handle the missing value, wrong data and wrong format. Also, we perform feature selection, data normalization and data balancing. Then the dataset converted raw dataset to ML trainable dataset. We use the machine learning classifiers to predict the mother's mode of delivery. We also propose an ensemble techniques that shows better performance than the benchmark ML algorithms. To use the inter dataset as training and testing, we perform 10-fold cross validation. Different performance measure technique are used to measure the performance of the classifiers and tabulated in result section. Different charts are introduced in the result and analysis section to show the full experimental result in details.

B. Dataset description

This study relies on data collected and compiled by Campillo-Artero et al. [20]. It contains 6,157 patient records from four public hospitals in Spain in 2014. There are a total of 161 features, 142 of which are categorical and 19 numerical. The collection included information organized into six distinct categories. Except for the C-section group, all of the groups have been switched to the "regular delivery" group. There are 692 cesarean records and 5465 non-cesarean records, indicating an imbalance in the data. There are a lot of features in this dataset that aren't necessary for training a machine learning model. This section will provide a high-level overview of the procedure that follows.

C. Data prepossessing techniques

In the prepossessing section a pipe-lined based processing has been applied. At first for the missing values, we have used the mean. In this respect we have calculated the mean values for some attribute and put the values into the missing values in that attribute. We use synthetic minority oversampling technique (SMOTE) technique for handling the imbalance data. This technique adds more value to our original dataset without losing any information by creating synthetic value [21].

D. Feature selection

For the feature selection process we have applied three methods namely chi-square test, univariate feature selection and forward feature selection.

Chi-square Test: In chi-square test, the important feature is selected by observing the observed value and the expected value. If the value of chi-square test is small between two attributes then it can be said that the attributes are independent. The formula for calculating the Chi-square test is:

$$\chi^2 = \frac{\sum(O-E)^2}{E}$$

Where, O represents the observed frequency, E represents the expected frequency under the null hypothesis.

Univariate feature selection: In univariate feature selection method valuable features are selected by find the relationship between the output variable and each input value. Actually it works based on selecting the best features from all the features. Suppose we have some attributes and the values of that attributes can be measured by univarient statistical test and the highest valued attribute is the most important feature of all features.

Forward feature selection: On the other hand, in forward feature selection method, we calculate the metric value on cross validation dataset and if get better value then add the feature to useful feature list.

E. Description of the algorithms

Support Vector Machine (SVM): SVM is one of the mostly used supervised machine learning algorithm that can be used both regression and classification purpose [22]. In the Support Vector Machine, each piece of data is plotted as a point in n-dimensional space (where n is the number of features). Data objects are classified with the use of hyperplane, also known as decision boundaries. The main goal of svm is to maximize the margin or distance of the support vectors. This algorithm use the Kernel method to convert low dimensional input space to a higher dimensional space. The hyperplane is define as $w.x + b = 0$ where w is a vector normal to the hyperplane and b is offset. The following equations define the svm clearly:

$$\vec{X}.\vec{w} - c \geq 0$$

$$putting - c\ as\ b, we\ get$$

$$\vec{X}.\vec{w} + b \geq 0\ hence$$

$$f(x) = \begin{cases} +1, & \text{if } \vec{X}.\vec{w} + b \geq 0 \\ -1, & \text{if } \vec{X}.\vec{w} + b < 0 \end{cases}$$

If $\vec{X}.\vec{w} - c \geq 0$ is active then the result being positive, otherwise it show negative point. w and b is responsible to maximize the margin. When the dataset is huge and has more noise, there are some classification challenges that SVC struggles with. Additionally, SVM will perform poorly if the target cases overlap and there are more features per data point than there are training data.

Decision Tree (DT): A decision tree is a tree structure, hierarchical and non-parametric supervised machine learning method which can do both classification and regression tasks [23]. It actually works on divide and conquer strategy and finds an optimal split point within the tree by using greedy search. More important attributes can be found by seeing the hierarchical nature of this tree which is missing in other algorithms. But complex decision tree can lead over fitting. It gives a way to represent the algorithm with the conditional statement and every branch gives the output of the attributes.

Random Forest (RF): A random forest is an extension version of bagging (one kinds of ensemble method) and constructed by multiple decision trees [24]. The over fitting problem is solved by this random forest by selecting the subset of all possible features where a decision tree consider all of those features. From the training data set different decision tree is made up and then summarized this tree to get the final result. But this algorithm requires more space and it is time consuming process. For classification random forest gives a vote and it selects the classification with the majority of votes. For regression it does average all the outputs of the tree.

Gradient Boosting Classifier (GBC): Gradient boosting classifier is one kinds of boosting method which works on the basic of weak predictors by combining them to make more accurate predictor [25]. This classifier is used for both regression and classification problem. In each iteration, a new learner is introduced to the errors that the previous learner did. So the new learner is concerned about the errors and try to resolve or updates the errors and performs well than the previous learner. By following this process the classifier finally provides better performance.

Logistic Regression (LR): Logistic regression is a statistical model which tells us the probability of occurring an event by investigating some independent variables and it deals to the issues of classification. Logistic regression has only two outcomes either it is happen or not happen. It provides accurate result for those which have yes or no outcomes. As it is a probability based outcome model, the dependent variable is bounded 0 to 1. This model provides good accuracy in terms of binary class classification. The probability is calculated by,

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Multinomial Naïve Bayes (MNB): Multinomial Naïve Bayes is one kinds of Naïve Bayes classifier algorithm which is very useful when data set is multinomially distributed. This classifier algorithm is a master in those types of tasks which are related to natural language processing. It is a popular classifier to classify news related to global, political, regional and so no.

K Nearest Neighbor (KNN): K nearest neighbor is the simplest machine learning classifier which classifies the data point by measuring the similarities of previously classified data. For example if you have two types of fruits like mango and apple (consider color and shape), when a new object or fruit come KNN compare this with its earlier stored fruit data (color and Shape). A new data choose its nearest neighbor by calculating the Euclidean distance. KNN is very powerful algorithm for this it becomes the most popular machine learning classifier for small types of data set.

Stochastic Gradient Descent (SGD): Stochastic gradient descent is the improved version of gradient descent which has better performance on large scale data set. In gradient descent algorithm every item of the dataset is used to find the pattern and predict the output. But the problem is if there are millions of item in the dataset then the computational complexity will be increased. But using the stochastic gradient descent this problem can be solved. In this respect, SGD takes some random items for each iteration and try to find the pattern and so there is no need to consider all the values of the dataset. So it is also called optimizer algorithm and it minimize the cost function and thus fit the model in a better way.

CatBoost (CB): CatBoost or categorical boosting is the one of the latest boosting classifier algorithms which is very useful when there are some categorical data. It improves the accuracy by reducing over fitting and it performs well and fast in Heterogeneous data. This classifier can handle categorical features automatically by using one hot encoding. This classifier can be used in regression, classification as well as for casting, ranking, recommendation systems and sometimes in personal assistants by using greedy method.

Adaptive Boosting (AB): Adaboost classifier is a strong ensemble classifier and its used to increase the performance of machine learning algorithms. It combines all classifier which does not provide proper result individually but after combination of those classifiers, set weights for each classifier and training the data in every iteration so that it gives good accuracy for overall classifier. Adaboost uses decision tree with one level means only 1 split.

Gaussian Naïve Bayes (GNB): Gaussian Naïve Bayes is one kind of naïve bayes which works according to Gaussian distribution, also called normal distribution. It produces a bell shaped curve to calculate the mean and standard deviation of the training data. It is a probabilistic method which determines an output with the help of conditional probability. GNB exchanges the parameter with the new assign value of the variable and makes a result .It is a quick classification method which gives good accuracy without extra effort.

Extreme Gradient Boosting (XGB): Extreme gradient boosting is the latest extension version of gradient boosted decision tree which provide better speed and performance with the help of linear model solver and tree learning algorithms. It conducts parallel computation in a single machine to solve data science problem faster than any other machine learning algorithms. XGB is a decision tree based models which develops a graph for examining the input under several 'IF' statements. When the If condition is true then it goes to another if statement and eventual prediction .XGB adds more and more IF statement to the decision tree to build a strong model.

Support Vector-Extreme Gradient Boosting-Random Forest Classifier (SVXGBRF): This classifier proposed by us is built with SVC, XGB, and RF by stacking the ensemble procedure. In order to construct a new model, stacking takes the use of predictions for several nodes. The final model is then set to use predicting values from the test set. The SVXGBRF model contains XGB as a level1 classifier for the final_estimator, and SVC, XGB, and RF are used as remaining estimators as the level0 with the default parameter of SKLearn. It shows a better performance than the benchmark machine learning models.

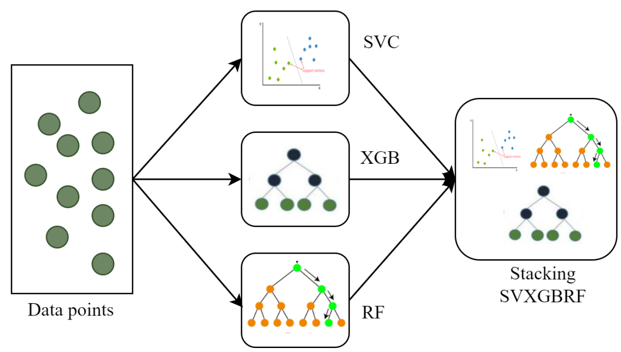

Fig. 2. Building blocks of stacking SVXGBRF

F. Performance measure techniques

In this section, we will discuss the matrices which are used to evaluate the performance of the models.

Precision: It tells us how many positive predictions are made correctly by the model.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall: It tells how many actual positive cases are there from total predicted case.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F1-score (F1): It tells us the harmonic mean of the precision and recall score.

$$F1score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Cohn Kappa: It is a statistical measurement that is used to measure the inter and intra rater reliability to classify items.

$$CohenKappa = \frac{P - P1}{1 - P1}$$

Where P = observed argument and P1 = hypothetical probability.

AUC: The level of separation can be measured by calculating the Area Under the Curve (AUC) score. A stronger predictor of True Positives and True Negatives is a model with a higher AUC. Area under the ROC curve (AUC) is the sum of the ROC areas.

ROC: The ROC curve illustrates the compromise that must be made between sensitivity (sometimes referred to as TPR) and specificity (1 − FPR). A higher performance can be inferred from classifiers that produce curves that are situated closer to the upper left corner. A random classifier is expected to give points that are located along the diagonal (FPR = TPR), and thus serves as a baseline expectation.

IV. Result and Discussion

Using the machine learning classifiers, we predict the mothers mode of delivery. Also, to check the full dataset performance, we use 10 fold cross validation. The evaluation matrics result are tabulated in the Table I. Also, the K Fold cross validation result are show in Fig. 4 as a scatter plot. We apply SMOTE to balance the dataset. Before balancing class 1 contains 5465 and class 0 contains 692 instance. After balancing both class contain 5465 instances.

TABLE I
Performance of the classifiers

Algorithm	Accuracy	Precision	Recall	F1 Score	Cohen Kappa	AUC
SVM	86.52	0.87	0.87	0.87	0.73	0.94
DT	87.55	0.88	0.88	0.88	0.75	0.93
RF	95.13	0.95	0.95	0.95	0.90	0.99
GBC	92.49	0.93	0.92	0.92	0.84	0.98
LR	85.98	0.86	0.86	0.86	0.72	0.94
MNB	79.53	0.80	0.80	0.80	0.59	0.88
KNN	86.81	0.88	0.87	0.87	0.74	0.94
SGD	85.34	0.85	0.85	0.85	0.71	0.94
CB	95.00	0.95	0.95	0.95	0.89	0.99
AB	90.00	0.90	0.90	0.90	0.80	0.96
GNB	74.08	0.81	0.74	0.73	0.48	0.88
XGB	94.76	0.94	0.94	0.94	0.89	0.99
SVXGBRF	**95.52**	**0.96**	**0.96**	**0.96**	**0.91**	**0.99**

From table 1 it can be observed that, in terms of accuracy, multinomial naive bayes and gaussian naive bayes give lower accuracy than any other models which are below 80%. On the other hand, random forest, catboost and extreme gradient boosting perform well which are 95.13%, 95% and 94.76% respectively. The accuracy of other models are in the range of 80% to 92%. But, here our model SVXGBRF achieves higher accuracy than any other models which is 95.52%. For the precision section, our model experienced higher score than the others which is 96%, whereas multimonial naive bayes has lower performance which is only 80%. Again, for the recall and f1-score, random forest and catboost model have good score but our model gains 96% score and outperforms from any other models. In terms of cohen kappa, most of the models have low scores except random forest algorithm. And here our model gives better performance which is above 90%. For AUC curve, it can be also seen that our proposed ensemble model achieves higher score which is similar to random forest, catboost and extreme gradient boosting algorithms.

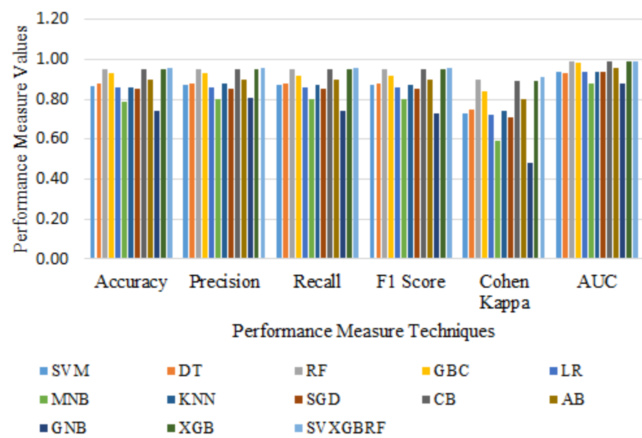

Fig. 3. Performance of the classifiers

The Fig 3 shows the performance of the classifiers that are tabulated in the table 1. Last bar of the every segment shows the performance of the proposed ensemble model. The last bar is superior to all.

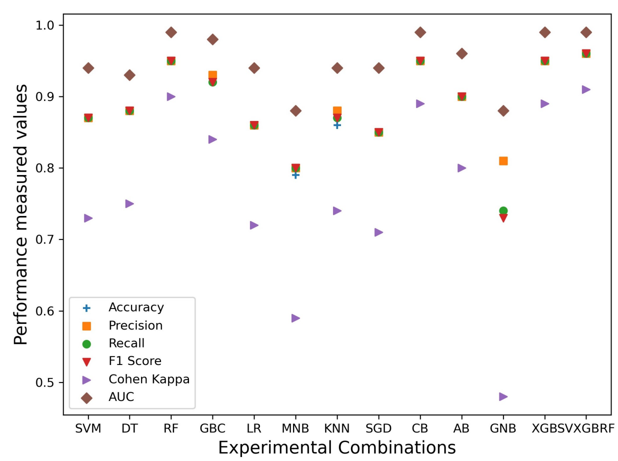

Fig. 4. 10 fold cross validation result of the classifiers

Fig. 4 shows us the performances after applying the 10 fold dataset for cross validation. And again our model experienced higher accuracy than any other models which is almost 96%. Some models have very low score especially in terms of cohen kappa but here our model handles this in an impressive way.

Fig. 5 provides us the information about the true positive and false positive rate which is also known as ROC curve where different colors represents different models performance. It is clearly seen that our proposed ensemble model has the highest value among all others algorithms as our model is

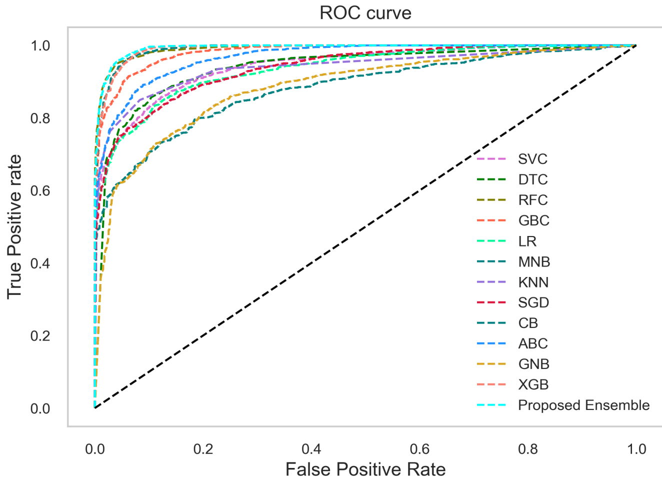

Fig. 5. ROC curve to show the performance of the models

the combination of support vector machine, extreme gradient boosting and random forest and all of these models have good mechanism to perfectly handle any dataset.

After evaluating we get some top features, Episotomy, Fetal Intrapartum pH, Age, Height, Anesthesia get more priority than other features.

V. Conclusion and Future work

In order to ensure the safety of the mother and the newborn, it is essential to choose the mode of delivery carefully. It is still necessary, however, to investigate the most useful combinations of features for computing this decision. Therefore, we are attempting to utilize AI to determine the safest and most effective modes of delivering a baby. Among all the binary classifiers, our proposed stacking SVXGBRF ensemble model shows superiority using the data preprocessing, feature selection, and data balancing pipeline. How our proposed method shows a better result is not shown in this study, and which features are getting more priority will investigate in further study.

References

[1] M. Arora. (2019) 6 different types of delivery methods you must know. [Online]. Available: https://parenting.firstcry.com/articles/different-childbirth-methods-you-must-know/

[2] (2021) Bangladesh: 51 per cent increase in "unnecessary" c-sections in two years. [Online]. Available: https://www.savethechildren.net/news/bangladesh-51-cent-increase-

[3] T. Boerma, C. Ronsmans, D. Y. Melesse, A. J. Barros, F. C. Barros, L. Juan, A.-B. Moller, L. Say, A. R. Hosseinpoor, M. Yi *et al.*, "Global epidemiology of use of and disparities in caesarean sections," *The Lancet*, vol. 392, no. 10155, pp. 1341–1348, 2018.

[4] T. C. Johnson. (2020) What are the risks of a c-section? [Online]. Available: https://www.webmd.com/baby/risks-of-a-c-section

[5] M. R. Haider, M. M. Rahman, M. Moinuddin, A. E. Rahman, S. Ahmed, and M. M. Khan, "Ever-increasing caesarean section and its economic burden in bangladesh," *PloS one*, vol. 13, no. 12, p. e0208623, 2018.

[6] C. Global. (2021) Why is machine learning important? [Online]. Available: https://csuglobal.edu/blog/why-machine-learning-important

[7] N. S. Khan, M. H. Muaz, A. Kabir, and M. N. Islam, "Diabetes predicting mhealth application using machine learning," in *2017 IEEE international WIE conference on electrical and computer engineering (WIECON-ECE)*. IEEE, 2017, pp. 237–240.

[8] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE access*, vol. 7, pp. 81 542–81 554, 2019.

[9] M. N. Islam and A. N. Islam, "A systematic review of the digital interventions for fighting covid-19: the bangladesh perspective," *Ieee Access*, vol. 8, pp. 114 078–114 087, 2020.

[10] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi, and M. N. Islam, "A machine learning approach to predict autism spectrum disorder," in *2019 International conference on electrical, computer and communication engineering (ECCE)*. IEEE, 2019, pp. 1–6.

[11] M. S. B. Alam, M. J. Patwary, and M. Hassan, "Birth mode prediction using bagging ensemble classifier: A case study of bangladesh," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*. IEEE, 2021, pp. 95–99.

[12] A. Ahmeidat, W. J. Kotts, J. Wong, D. J. McLernon, and M. Black, "Predictive models of individual risk of elective caesarean section complications: a systematic review," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 262, pp. 248–255, 2021.

[13] M. Lipschuetz et al., "Prediction of vaginal birth after cesarean deliveries using machine learning," *American journal of obstetrics and gynecology*, vol. 222, no. 6, pp. 613–e1, 2020.

[14] S. A. Abbas, A. U. Rehman, F. Majeed, A. Majid, M. S. A. Malik, Z. H. Kazmi, and S. Zafar, "Performance analysis of classification algorithms on birth dataset," *IEEE Access*, vol. 8, pp. 102 146–102 154, 2020.

[15] N. I. Khan, T. Mahmud, M. N. Islam, and S. N. Mustafina, "Prediction of cesarean childbirth using ensemble machine learning methods," in *Proceedings of the 22nd international conference on information integration and web-based applications & services*, 2020, pp. 331–339.

[16] C. Campillo-Artero, M. Serra-Burriel, and A. Calvo-Pérez, "Predictive modeling of emergency cesarean delivery," *PloS one*, vol. 13, no. 1, p. e0191248, 2018.

[17] M. Kowsher, N. J. Prottasha, A. Tahabilder, H. Kaiser, M. Abdur-Rakib, and M. S. Alam, "Predicting the appropriate mode of childbirth using machine learning algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, 2021.

[18] S. Maroufizadeh, P. Amini, M. Hosseini, A. Almasi-Hashiani, M. Mohammadi, B. Navid, and R. Omani-Samani, "Determinants of cesarean section among primiparas: a comparison of classification methods," *Iranian Journal of Public Health*, vol. 47, no. 12, p. 1913, 2018.

[19] P. Fergus, A. Hussain, D. Al-Jumeily, D.-S. Huang, and N. Bouguila, "Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms," *Biomedical engineering online*, vol. 16, no. 1, pp. 1–26, 2017.

[20] C. Campillo-Artero, M. Serra-Burriel, and A. Calvo-Pérez, "Predictive modeling of emergency cesarean delivery," *PloS one*, vol. 13, no. 1, p. e0191248, 2018.

[21] M. A. Sahid, M. Hasan, N. Akter, and M. M. R. Tareq, "Effect of imbalance data handling techniques to improve the accuracy of heart disease prediction using machine learning and deep learning," in *2022 IEEE Region 10 Symposium (TENSYMP)*, 2022, pp. 1–6.

[22] S. Suthaharan, "Support vector machine," in *Machine learning models and algorithms for big data classification*. Springer, 2016, pp. 207–235.

[23] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 18, no. 6, pp. 275–285, 2004.

[24] Y. Qi, "Random forest for bioinformatics," in *Ensemble machine learning*. Springer, 2012, pp. 307–323.

[25] P. Bahad and P. Saxena, "Study of adaboost and gradient boosting algorithms for predictive analytics," in *International Conference on Intelligent Computing and Smart Communication 2019*. Springer, 2020, pp. 235–244.