Prediction of Child Birth Delivery Mode Using Hybrid-Boosting Ensemble Machine Learning Model

Mamatha Sekireddy and Sudha Thatimakula
Research Scholar, Department of Computer Science, Sri Padmavati Mahila Visvavidyalayam, Padmavathi Nagar, Tirupati, Andhra Pradesh, India
mamatha.sekireddy@gmail.com, *thatimakula_sudha@yahoo.com*

Keywords: Ensemble Methods, Machine Learning, Artificial Intelligence, Voting.

Abstract: Maternal health is a critical aspect of public health that affects the well-being of both mother and infant. Despite the medical breakthroughs, maternal mortality rates remain high, particularly in developing countries. AI-based models provide new ways to analyze and interpret medical data, which can ultimately improve maternal and fetal health outcomes. The present study proposes a hybrid model for the maternal mode of delivery classification in pregnancy, which utilizes the strength of Ada Boost, Gradient Boost, and XG Boost algorithms. The proposed model culminates the three algorithms to improve the accuracy and efficiency of childbirth delivery classification method in pregnant women. The dataset used in this study consists of features such as age, systolic and diastolic blood pressure, blood sugar, body temperature, heart rate, placenta previa, and gravida. The dataset is divided into training and testing sets, where 70% of the data is used for training and the rest 30% for testing. The output of the Ada Boost, Gradient Boost, and XG Boost classifier is considered, and a maximum probability voting system selects the output with the highest probability as the most correct one. Performance is evaluated using various metrics, such as accuracy, precision, recall, and F1 score as an outcome where the XG Boost showed an accuracy of 97.07%, the Ada Boost showed 85.02%, and the Gradient Boost showed 92.51% respectively. Results also conspicuously showed that the proposed model achieves the highest accuracy of t98.05%, with 97.9% precision, 97.9% recall, and an F1 score of 97.9% on the testing dataset. The hybrid model proposed in this study has the potentiality to improve the accuracy and efficiency of the maternal mode of delivery classification in pregnancy, leading to better health outcomes for pregnant women and their babies.

1 INTRODUCTION

In recent years, artificial intelligence (AI) has emerged as a powerful tool in health care, offering new ways to analyze and interpret the complex medical data. AI-based models have shown promising results in various clinical applications, including disease diagnosis, treatment planning, and patient monitoring. AI-based modelscan analyze vast amounts of health data, both structured and unstructured, to identify patientswho may be at high risk for adverse outcomes. These models can help health care providers makemore accurate and timely decisions; and when it refers to maternal health care in pregnancy; it can ultimately improve maternal and fetal health outcomes. AI offers novel approaches to prediction modelling, diagnosis as well as early detection of mode of delivery in pregnancy.

In general, there are two main types of baby delivery: vaginal delivery, often known as normal delivery and caesarean delivery. Each delivery has advantages and disadvantages. Vaginal delivery is the most common method of childbirth because it is associated with little risk. Pregnant women should take care of their health after caesarean deliveries because they can result in blood loss, infections, and difficulties with subsequent pregnancies [1].

A machine learning-based maternal health care system will anticipate the baby's mode of birth in advance through the analysis of the pregnant woman's medical features generated during the pregnancy period.

2 LITERATURE SURVEY

This section briefly summarizes prior research on the prediction of baby delivery. The quadratic discriminant analysis approach demonstrated the maximum accuracy of 0.979992 with the F1 score of 0.979962 in the study by the author [2] on predicting the right mode of birthing using a machine learning algorithm. The authors in [3] proposed a paper on a comparative analysis of supervised machine learning techniques for diagnosing mode of delivery in the medical sciences, in which the Random Forest algorithm gets the greatest accuracy of 99%. [4] Syed Ali Abbas proposed a study on caesarean analysis and the application of Machine Learning methods for birth data classification. The primary causes of caesarean delivery are of being determined to be old age, high blood pressure, and diabetes. A study on the comparative analysis of Naive Bayes and Decision Tree C4.5 for CS prediction was proposed by Gusti Ayu Suciningsih. Memory utilisation, programme execution time, and accuracy metrics are employed in a comparison analysis of the Nave Bayes & Decision Tree given in [5].

3 METHODOLOGY

2.1 Data Description and Operating Environment

The maternal mode of delivery dataset utilized in this study is acquired from the maternal health care center. The dataset contained eight predictive variables in pregnancy viz., Age (in years), Systolic Blood Pressure (in mmHg), Diastolic Blood Pressure (in mmHg), BloodSugar (in mmol/L), Body Temperature (in degrees Fahrenheit), and Heart Rate (in beats per minute), Gravida (in 1,2,3,4) and placenta previa (in yes/no) all of which were numerical variables except placenta previa. A categorical predict or variable, Placenta previa, is present in two possible cases Yes and No. A categorical target variable, Delivery Method, is present in two possible classes; Caesarean, and Normal. The second sequent totaled 9 variables in the whole dataset. There were 1,021 instances in the dataset; 529 samples of which were in the Caesarean class, and 492 samples are in the normal class, as shown in Figure 1. Attributes and their data types in the dataset are specified in Table 1. There are no missing values in the dataset. All samples with all of their features are used in the model building and evaluation. All data manipulations, model building and evaluations are performed in the python programming language (version 3.8.3). Packages utilized include pycaret (2.3.10), numpy (1.20.3), pandas and (1.5.2), sklearn (0.23.2).

Table 1: Dataset specification.

S.no	Attributes	Datatype
1	Age (Input)	Numerical
2	Systolic Blood Pressure (Input)	Numerical
3	Diastolic Blood Pressure (Input)	Numerical
4	Blood Sugar (Input)	Numerical
5	Body Temperature (Input)	Numerical
6	Heart rate (Input)	Numerical
7	Gravida (Input)	Numerical
8	Placenta Previa (Input)	Categorical
9	Delivery Method (Output)	Categorical

Further feature statistics are shown in Table 2 that are based on count, mean, standard deviation, min (minimum values), 25%, 50%, 75%, and max(maximum values). The minimum value shows the lowest limit and the maximum value shows the highest limit for all features.

Table 2: Description of statistical distribution of Dataset Features.

Attribute	Count	Mean	Std	Min	25%	50%	75%	Max
Age	1021	28.6	8.7	18	21	28	35	45
Systolic BP	1021	114.8	20.4	70	95	120	130	160
Diastloic BP	1021	78.6	15.2	49	65	80	90	100
Blood Sugar	1021	9.8	3.9	6	7	7.7	11	19
Body Tempe-rature	1021	98.6	1.4	98	98	98	98	103
Heart rate	1021	74.6	8.0	60	76	76	80	90
Gravida	1021	1.6	0.7	1	2	2	2	4
Placenta Previa	1021	0.1	0.3	0	0	0	0	1
Delivery Method	1021	0.5	0.4	0	1	1	1	1

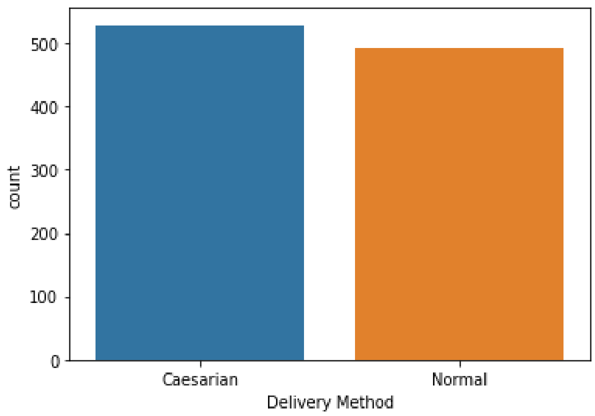


Figure 1: Distribution of dataset based on Delivery Method.

3.2 Data Preprocessing and Analysis

The data file is saved in the comma-separated value (CSV) format, and this facilitated easy reading into the Python script. The major pre-processing step carried out is in the way of encoding the target (categorical) variables and predictor variable placenta previa into numerical variables to facilitate computation. Placenta previa categorical variables yes and no are coded as 1 and 0, and Normal and Caesarean classes were coded as 0 and 1 respectively. Subsequently, the correlation [6] between variables is examined, as depicted in Figure 2. The variable with the highest correlation to the delivery method is found to be blood sugar, while the variable with the lowest correlation is body temperature and heart rate.

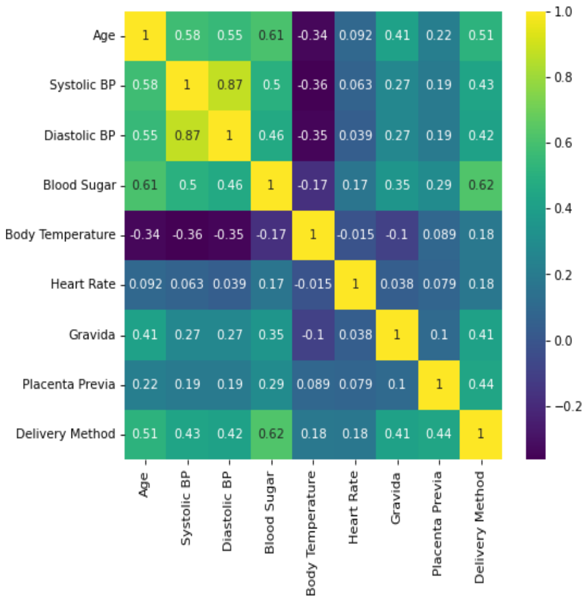


Figure 2: Correlation between the Dataset Variables.

3.3 Machine Learning Ensemble Methods

The data was randomly divided into training and testing sets, with 70% and 30% of the data, respectively. To facilitate data analysis, the Standard Scaler is applied to scale the data. The predictive abilities of various commonly used machine learning ensemble algorithms are initially examined-including, Gradient Boosting Classifier, Ada Boost Classifier, and XG Boost Classifier [7]. A comparison is made among these different ML algorithms using the accuracy, precision, recall, and F1score. The results are summarized in Table 3. The algorithm that performed best according to these metrics is the XG Boost Classifier.

3.4 Training of the Hybrid Model

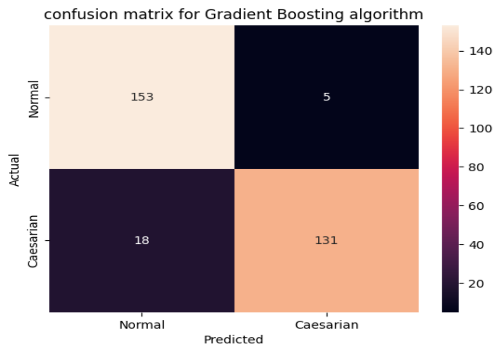
Afterward, these three classifiers are merged into one final hybrid model. The hybrid model utilized a maximum probability voting technique, this is by looking at the predictions (classes and probability values) of Ada boost, Gradient Boosting, and the XG Boost classifier, it subsequently selects and outputs whichever class had the highest probability value regardless of the difference in class predictions of both classifiers. The preference for this probabilistic approach stems from its demonstrated ability to offer simplicity and good interpretability, as similar maximum likelihood ensembles have shown in the past [8]. Additionally, since only 3 classifier units are employed in this binary-class prediction ensemble, the maximum probability voting techniquecapitalizes on their distinct natures, to the advantage that they will not commit the same errors [9]. Thus, the result from the more confident classifier is selected as the most correct one.

4 RESULTS

The performance of the proposed hybrid model has been evaluated based on the following key metrics: Accuracy, Recall, Precision, and F1 Score [10]. An analysis of the detailed comparison of metrics between Ada boost, Gradient Boost, XG Boost and Hybrid ensemble algorithm provided in Table 3 which reveals that the overall hybrid model performed significantly better with an accuracy score of 0.98. This increase in accuracy could be attributed to the fact that both models in the ensemble differ in the nature of their misclassifications, consequently drawing on each’s distinctive nature; to the advantage of one another, to provide the most correct classification. The precision, recall and F1-score from our hybrid model ranged to 0.9700.The confusion matrix, illustrated in Figure 6 that provides valuable insights into hybrid model performance regarding classification of instances into normal and cesarean delivery methods. The analysis reveals that the model accurately predicted 155 instances as Normal, 146 instances as Caesarean. Of particular significance is hybrid algorithm has an ability to avoid misclassifying any Normal instance as Caesarean. This outcome is crucial, especially in real-world medical applications, as misclassifying a Normal instance as Cesarean could have severe consequences. The model's ability to avoid such errors indicates its reliability and suitability for medical decision-making.

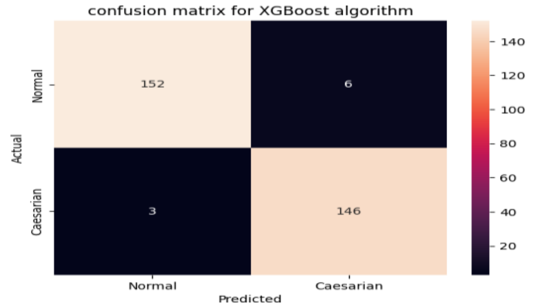
Table 3: Performance comparison of Ada Boost, XG Boost, Gradient Boosting, and Hybrid ensemble model.

Model	Classes	Precision	Recall	F1 Score	Accu-racy
AdaBoost	Normal	0.90	0.80	0.85	0.85
	Cesarean	0.81	0.91	0.85	
Gradient Boost	Normal	0.89	0.97	0.93	0.93
	Cesarean	0.96	0.88	0.92	
XG Boost	Normal	0.98	0.96	0.97	0.97
	Cesarean	0.96	0.98	0.97	
Hybrid boosting ensemble model	Normal	0.98	0.98	0.98	0.98
	Cesarean	0.98	0.98	0.98	



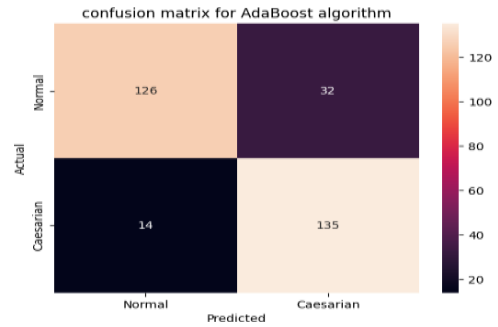
	Precision	Recall	F1-Score	Support
0	0.89	0.97	0.93	158
1	0.96	0.88	0.92	149
accuracy			0.93	307
macro avg	0.93	0.92	0.92	307
weighted avg	0.93	0.93	0.92	307

Figure 3: Gradient Boosting: confusion matrix and classification report.



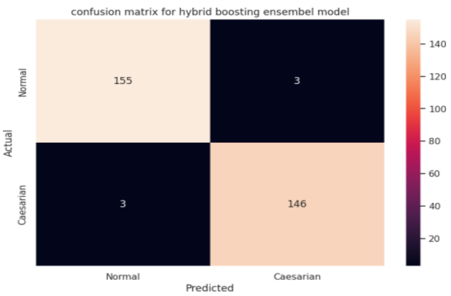
	Precision	Recall	F1-Score	Support
0	0.89	0.97	0.93	158
1	0.96	0.88	0.92	149
accuracy			0.93	307
macro avg	0.93	0.92	0.92	307
weighted avg	0.93	0.93	0.92	307

Figure 4: XG Boost: confusion matrix and classification report.



	Precision	Recall	F1-Score	Support
0	0.90	0.80	0.85	158
1	0.81	0.91	0.85	149
accuracy			0.85	307
macro avg	0.85	0.85	0.85	307
weighted avg	0.86	0.85	0.85	307

Figure 5: Ada Boost : confusion matrix and classification report.



0	0.98	0.98	0.98	158
1	0.98	0.98	0.98	149
accuracy			0.98	307
macro avg	0.98	0.98	0.98	307
weighted avg	0.98	0.98	0.98	307

Figure 6: Hybrid Boosting ensemble model: confusion matrix and classification report.

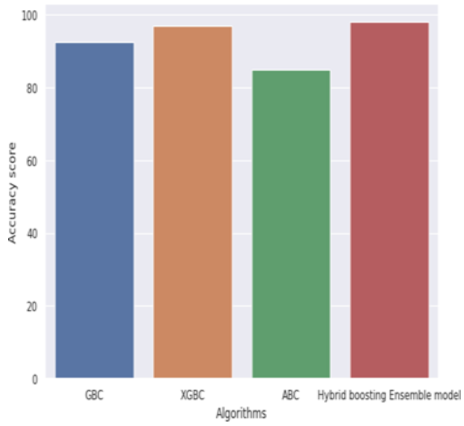


Figure 7: Comparison of Proposed Hybrid model with Ensemble Machine Learning methods.

4 CONCLUSIONS

In conclusion, the present study developed a novel hybrid ensemble model for the maternal mode of baby delivery classification that achieved high accuracy and performed comparably to the previous models. The model has important implications for clinical practice, and the findings suggest that the ensemble technique of machine learning models can be the useful tools for maternal baby delivery classification. Future research is needed to confirm the generalizability of the present findings and to optimize the use of structured and unstructured data in these models. By accurately identifying Cesarean delivery in advance, health care providers can take preventative measures to reduce the likelihood of complications and improve maternal and fetal outcomes. The high accuracy of the model suggests that an ensemble of machine learning models has the potentiality to be the useful tools in the maternal mode of baby delivery classification.

REFERENCES

[1] M. Najrul Islam, "Exploring Machine learning algorithms to find the best features for predicting mode of child birth," Digital Object Identifier 10.1109/ACCESS.2020.3045469.

[2] Md. Kowsher, "Predicting the Appropriate Mode of Childbirth using Machine Learning Algorithms," IJACSA, vol. 12, no. 5, 2021.

[3] S. S. Hussain, "A Comparative Study of Supervised Machine Learning Techniques for Diagnosing Mode of Delivery in Medical Sciences," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 10, no. 12, 2019.

[4] S. A. Abbas, "Cause Analysis of Caesarean Sections and Applications of Machine Learning Methods for Classification of Birth Data," Digital Object Identifier 10.1109/ACCESS. 2018.2879115, IEEE Access, 2018.

[5] G. A. Suciningsih, "Comparative analysis of Naïve Bayes and Decision Tree C4.5 for Cesarean Section Prediction," Journal of Soft Computing Exploration, ISSN: 2746-0991.

[6] A. Crossman, "Correlation Analysis in Research," from www.thoughtco.com: https://www.thoughtco.com/what-is-correlation-analysis-3026696.

[7] Rocca, "Ensemble-methods-bagging-boosting-and-stacking," from towardsdatascience.com: https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205.

[8] K. Dembczynski, W. Kotłowski, and S. Słowi ´ nski, "Maximum Likelihood Rule Ensembles," in Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland: Association for Computing Machinery, p. 224–31. doi: 10.1145/1390156.1390185.

[9] F. Leon, S. A. Floria, and C. Baˇdicaˇ, "Evaluating the Effect of Voting Methods on Ensemble-based Classification," in 2017 IEEE International Conference on Innovations in Intelligent Systems and Applications (INISTA), Gdynia: Institute of Electrical and Electronics Engineers (IEEE). doi: 10.1109/INISTA.2017.8001122.

[10] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," arXiv.