Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Predicting Birth Mode Using Ensemble Machine Learning: Evidence from Bangladesh Demographic and Health Survey

Md. Ziaul Alam
Jahangirnagar University
Farhana Afrin Duty
Jahangirnagar University
Tasneem Binte Reza
Jahangirnagar University
Nahid Salma
nahidsalma@juniv.edu

Jahangirnagar University
Md. Musharraf Hossain
Jamalpur Medical College Hospital

Article

Keywords: Caesarean Delivery, Boruta Algorithm, Ensemble method, Maternal Mortality, Feature Selection

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9726982/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background and Aim:

Maternal mortality and childbirth complications remain significant challenges in many developing countries, particularly in rural areas. Early identification of childbirth risks is crucial for reducing mortality rates. This study employs a range of machine learning techniques to identify key features related to Birth Mode and determine the most effective predictive model. The analysis is primarily based on data from the BDHS 2017–2018, aiming to improve risk prediction and ultimately enhance maternal health outcomes.

Methods

To identify key factors influencing birth mode, the Boruta algorithm was applied. The optimal predictive framework was then determined by evaluating several machine learning algorithms, including Random Forest (RF), Bagging, GradientBoost (GB), XGBoost (eXtreme Gradient Boosting), and AdaBoost (Adaptive Boosting). The best model was selected based on performance metrics such as precision, sensitivity, specificity, F-score, accuracy, and AUC value.

Results

The chi-square test revealed a significant association between birth mode and factors such as division, place of residence, age at first birth, number of living children, husband's education, respondent's employment status, wealth index, iron intake during pregnancy, education level, and BMI, all at a 1% significance level. The Boruta algorithm identified 13 key features, including those from the chi-square test, plus two additional ones. Ensemble models (RF, Bagging, GB, XGBoost, AdaBoost) were trained using these features. Bagging exhibited the highest accuracy (0.7611), precision (0.7746), F-score (0.8278), and sensitivity (0.8887). XGBoost achieved the highest specificity (0.8962), while AdaBoost had the best AUC (0.7833). Overall, Bagging outperformed the other models across four out of six metrics.

Conclusions

The study demonstrates the effectiveness of the Bagging algorithm for predicting birth mode, providing important insights into the factors associated with caesarean deliveries. The findings indicate a higher likelihood of caesarean sections among women from wealthier socioeconomic groups compared to those from lower-income backgrounds. This observation underscores the importance of increasing awareness regarding the potential risks and overuse of caesarean procedures, particularly among higher-income populations. The results suggest that incorporating these identified risk factors into

clinical decision-making and public health strategies may support healthcare providers and policymakers in Bangladesh in reducing the prevalence of unnecessary caesarean deliveries.

1. Introduction

Birth Mode (BM), or the method of delivery, plays a pivotal role in determining maternal and neonatal outcomes (Hasan et al., 2023). Childbirth is a complex and transformative process influenced by a variety of factors, with delivery occurring either through vaginal birth or cesarean section (C-section). In recent years, the global healthcare community has seen a significant rise in the number of C-sections, raising concerns about their appropriate use, associated health risks, and the economic burden they impose on healthcare systems (Islam & Noor, 2021). While C-sections are a crucial medical intervention for managing complications and safeguarding the health of both mothers and infants, their overuse can lead to adverse consequences. C-sections are linked to increased risks for both mothers (such as infections, prolonged recovery, and complications in future pregnancies) and infants (including higher rates of allergies, asthma, and diabetes) (Sana et al., 2012; Gy et al., 2017; Lee et al., 2023). Thus, a balanced approach to their use is necessary to mitigate long-term health impacts and optimize healthcare resources.

Several factors contribute to the rising global C-section rates, including socioeconomic status, maternal age, and educational background (Irfan & Nisar, 2018). However, making rapid and accurate decisions during childbirth can be difficult for healthcare professionals, especially when the consequences of incorrect choices can be life-threatening for both the mother and child (Souza et al., 2010). Understanding the drivers of increasing C-section rates is therefore essential to inform evidence-based decision-making, improve healthcare policies, and reduce unnecessary medical interventions.

In Bangladesh, C-section prevalence has dramatically increased, with the national rate now around 33%, far surpassing the WHO's recommended 10–15%. Urban areas have even higher rates, reaching 36.9%, while regions like Noakhali report a striking 53.75% prevalence (Ahmmed et al., 2021; Alam et al., 2021; S. S. Alam et al., 2021). The rise is particularly notable in private healthcare facilities, where C-section rates can reach 83% (Nahar et al., 2022). Factors contributing to this increase include maternal age, education, wealth, antenatal care visits, and clinical issues like prolonged labor and previous C-sections (Ahmmed et al., 2021; Kumar & Sharma, 2023; Hasan et al., 2019; Ahamed & Jahan, 2022; Begum et al., 2019). Profit-driven private hospitals also favor C-sections, exacerbating the trend (Hossain et al., 2022; Nahar et al., 2022). This rise in C-sections is concerning due to its association with higher maternal and neonatal morbidity, including complications like back pain, urinary incontinence, and postnatal infections (Hossain et al., 2022). Despite significant progress in reducing maternal mortality in Bangladesh, the country still faces substantial challenges in meeting its health targets. According to Bhutta et al. (2004), maternal death rates in South Asia in 2015 ranged from 29 per 100,000 live births in Sri Lanka to 400 per 100,000 live births in Afghanistan. To meet Millennium Development Goal 5, Bangladesh was required to reduce its maternal mortality ratio from 569 to 143 per 100,000 live births between 1990 and 2015. Although the country made considerable progress and achieved a maternal mortality ratio of 176 per

100,000 live births by 2015, it still falls short of meeting the goal. Further research is needed to identify the factors contributing to maternal mortality and to inform policy decisions aimed at improving maternal health outcomes.

Machine learning (ML) techniques have become increasingly popular in Bangladesh as a tool for modeling and predicting birth modes, with the aim of reducing maternal mortality and childbirth complications. Several studies have employed various ML algorithms to identify the key factors influencing the mode of delivery and to develop predictive models. For example, logistic regression, decision trees, and other classifiers have been used to predict C-sections, with statistically significant predictors including factors such as division, education level, wealth index, BMI, and age at first birth (Khatun et al., 2022). Additionally, ensemble methods like bagging classifiers, based on traditional algorithms such as Decision Tree, K-Nearest Neighbor, Naive Bayes, and Support Vector Machines, have shown promising results in improving the prediction of birth modes. These approaches highlight the potential of ML techniques in creating decision support systems to help reduce unnecessary C-sections (M. S. B. Alam et al., 2021).

However, to the best of our knowledge, no study has focused exclusively on ensemble-based ML techniques for forecasting birth modes in Bangladesh. While traditional clinical methods remain essential, machine learning offers a more precise way to estimate BM risks and guide healthcare professionals in making informed decisions. The goal of this study is to identify key characteristics influencing birth mode through a variety of ensemble-based machine learning techniques and to determine the optimal predictive model for birth mode in Bangladesh. By doing so, this research aims to contribute to evidence-based decision-making and ultimately improve maternal health outcomes.

2. Related Works

Several studies have explored factors influencing birth mode, highlighting key determinants and the role of machine learning in predicting delivery methods. Campillo-Artero (2018) found that older mothers, those with higher body mass index (BMI) and weight, and those with a history of previous C-sections were more likely to undergo a repeat cesarean section. Additionally, women facing higher obstetric risks were more prone to emergency C-sections. Khatun (2022) suggested that women in the Khulna division should be particularly encouraged to consider natural delivery options. Begum (2017) identified several factors increasing the likelihood of a C-section, including higher socioeconomic status, education, older age, lower birth order, more antenatal care visits, and a history of pregnancy complications. Dashek et al. (1993) found that when the anticipated birth weight exceeded 3600 grams and the mother was over 35, the likelihood of a C-section increased.

Machine learning algorithms have also been used to predict birth mode. Studies by Kowsher (2020), M. Hasan (2023), and Campillo-Artero (2018) investigated both ensemble and non-ensemble methods, showing promise in improving predictive accuracy for C-sections. Karlström (2011) analyzed data from 150 countries, revealing that 18.6% of births now involve C-sections, with significant regional variations.

A subsequent study of 121 countries indicated a 12.4% global increase in C-section rates from 1990 to 2014, with varying annual increases across regions. Hasan (2019) found that C-section rates in Bangladesh had risen to 23.94%, surpassing the World Health Organization's recommended threshold. Furthermore, Karlström (2011) noted that some women who opted for C-sections were anxious about childbirth and dissatisfied with their pre-delivery care. This dissatisfaction often led to uncertainty about future pregnancies. These findings emphasize the importance of informed decision-making in childbirth and the need for a balanced, evidence-based approach to the use of C-sections. A summary of these related studies is presented in Table 1.

Table 1
Reviewing the current Birth Mode prediction models

Authors	Country	Data Types	Approach	Models/Analysis	Remarks
Khatun et al., 2022	Bangladesh	Secondary data collected from BDHS 2017-18	Classification	Decision Tree (DT), XGBoost, Random Forest, Support Vector Machine (SVM), and lightGBM	The Support Vector Machine (SVM) exhibited the highest accuracy
Hasan et al., 2019	Bangladesh	Secondary data from BDHS 2014	Secondary Data Analysis	Population Attributable Fractions	Identified multiple predictors for CS deliveries; highlighted need for targeted interventions to align CS rates with global recommendation
Begum et al., 2017	Bangladesh	HDSS data from Matlab Upazila under Chandpur District in BD	Retrospective Study	Binary Logistic Regression	Found higher socioeconomic status, education, and older age increased likelihood of C-section
Adashek et al., 1993	USA	Prospective data collection	Estimation	Multiple Logistic Regression, t-tests, and chi-square analyses	Established connection between higher birth weight and cesarean sections, particularly for women under 35
Sandall et al., 2018	Global	Variety of sources, data from 169 countries	Comprehensive Review	None	Examined short-term and long-term impacts of CS on maternal and infant health; promoted evidence-based decision-making
Kowsher et al., 2020	Bangladesh	Dataset from Tarail Upazilla Health Complex	Classification	Nine machine learning classifiers, including "impact learning"	"Impact learning" approach provided the highest accuracy (0.89) and F1 score (0.88)

Authors	Country	Data Types	Approach	Models/Analysis	Remarks
Karlström et al., 2011	Sweden	Survey data	Comparative Study	Multivariate Logistic Regression	Compared women's experiences of CS vs. vaginal births; highlighted dissatisfaction among women who chose CS
Betrán et al., 2016b	Global	Data from 150 countries	Trend Analysis	Descriptive Analysis	Analyzed global, regional, and national trends in CS rates from 1990 to 2014; emphasized need for balanced approach to maternal health and neonatal health
M. Hasan et al., 2023	Bangladesh	from another study in Spain: Campillo-Artero et al., 2018	Classification	SVM, Decision Tree, Random Forest, Gradient Boosting, etc.	Proposed an ensemble model SVXGBRF; achieved 95.52% accuracy and high performance metrics; laid groundwork for decision-support system
Campillo-Artero et al., 2018	Spain	Data from 4 public hospitals	Classification	Logistic Regression, Classification Tree, Random Forest	Random Forest model achieved high accuracy (AUC 0.94)

3. Methods and Materials

To achieve our goal, we developed a structured study approach, carefully executing each step from data collection to result presentation, as illustrated in Fig. 1. The study utilized the BDHS 2017–18 dataset, starting with descriptive statistics to summarize key respondent information. After data processing, the dataset was randomly split into training and test sets in a 70:30 ratio. To identify significant features, we employed the Boruta algorithm and chi-square test. We then applied a variety of machine learning techniques, including Random Forest (RF), Bagging, Gradient Boosting (GB), XGBoost, and AdaBoost. The performance of each model was assessed based on precision, sensitivity, specificity, F-score,

accuracy, and AUC value. Ultimately, the best predictive model for Birth Mode (BM) was selected, and the most important BM characteristics were identified. All analyses were conducted using SPSS V.23.

3.1 Data Source and Study Variables

In this study, all the variables are extracted from the dataset BDHS 2017-18. The outcome variable is Delivery by caesarean section (Caesarean, which) was expressed as a binary outcome. The binary nature of this variable implies two distinct possibilities: a value of "1" represents cases where the mode of delivery was Cesarean, while a value of "0" corresponds to instances of vaginal delivery.

To investigate factors associated with caesarean section deliveries, this study incorporates a range of independent variables at various levels. At the individual level, The geographic division where the respondent is located (Barisal, Chittagong, Dhaka, Khulna, Sylhet, etc.), Type of residence(Urban, Rural), age of the respondent at the time of their first birth, number of living children the respondent has, education level of the respondent's husband/partner (No education, Primary, Secondary, Higher education), respondent currently working (No, Yes), a child is alive (No, Yes), Wealth index combined (Poorest, Poorer, Middle, Richer, Richest), the respondent was given or bought iron tablets/syrup during pregnancy (No, Yes, Don't know), Respondent's highest educational level (No education, Primary incomplete, Primary complete, Secondary incomplete, Secondary complete), BMI cate: Body mass index category (Underweight, Normal weight, Overweight, Obese), Religion (Islam. Hinduism, Buddhism), Sex of Household Head (Male or Female), Number of Household member.

3.2 Feature Selection Method

To identify the most relevant predictors, both statistical and machine learning-based feature selection methods were employed.

3.5.1 Chi-square Test

A chi-square test of independence was conducted to examine the association between categorical predictors and the outcome variable. Variables with statistically significant associations ($p < 0.01$) were considered relevant for further analysis.

3.2.2 Boruta Algorithm

The Boruta algorithm serves as a wrapper for the Random Forest classification technique, aiming to pinpoint the most relevant input features from a potentially large set of features in the dataset collection for a target variable.

The steps of the Boruta algorithm are as follows:

- Add more variables to the information system, including at least 5 shadow attributes.
- Rearrange attributes to remove correlations with the outcome.

- Use the extended information system to run a random forest classifier and collect the calculated Z scores.
- Award a hit to attributes scoring over the highest Z score of the shadow attribute (MZSA).
- Use MZSA to run a two-sided test of equality for each attribute with unknown relevance.
- Permanently delete unimportant attributes with significantly lower importance than MZSA.
- Identify attributes as "important" if their importance significantly exceeds MZSA.
- Remove any shadow properties.
- Continue until all attributes are prioritized or until all available random forest runs are used.(Kursa & Rudnicki, 2010).

3.3 Ensemble Learning Models

Ensemble learning combines multiple models to achieve better performance than individual models. It leverages diverse models to mitigate overfitting, enhance generalization, and address complex datasets. Techniques include Bagging, Boosting, Stacking, and Random Forest. Ensemble learning is a cornerstone in modern machine learning, boosting accuracy and reliability across various applications.

3.3.1 Bagging

Bagging is a popular technique in machine learning that aims to improve the accuracy and stability of models. It works by creating multiple copies of the same model, each trained on different subsets of the original training data. These subsets are generated by randomly sampling the training data with replacement. When making predictions, bagging combines the predictions of all these models through a majority vote (for classification) or averaging (for regression), reducing variance and preventing overfitting.

3.3.2 AdaBoost

AdaBoost is an ensemble method that combines weak classifiers, often shallow decision trees, to create a strong classifier. It adjusts sample weights, giving more emphasis to misclassified instances, and trains classifiers in parallel. While effective for imbalanced data and preventing overfitting, it can be sensitive to noise. Hyperparameter tuning is crucial but can be less complex than other algorithms like random forests (Dong et al.,2020).

3.3.3 Gradient Boosting

Gradient Boosting is a useful method in machine learning that combines several simple models, usually decision trees, to create a strong predictive model. It focuses on reducing errors from previous models by adjusting each new model to the difference from the group. It uses a process called gradient descent to find the best settings for each simple model, gradually improving the overall performance. Gradient Boosting is very effective for many different tasks and often works better than other methods for making

predictions. However, it needs careful adjustment of settings like learning rate, tree depth, and number of iterations to avoid focusing too much on the training data (Dong et al.,2020).

3.3.4 XGBoost

XGBoost is a powerful ensemble machine learning algorithm. It combines multiple decision trees and offers various features such as automatic missing data handling, tree pruning, feature importance evaluation, early stopping, and cross-validation. XGBoost is known for its exceptional performance and is widely used for structured data problems in academic research and practical applications.

3.3.5 Random Forest

Random forest, also known as decision tree forest, is an ensemble-based learning method that focuses on decision tree ensembles. It uses a bagging approach to create multiple decision trees with a random subset of data, which are then combined to make a final decision. Random forests can be used for both classification and regression problems in machine learning, by combining various classifiers to address complex issues and improve model performance.

4. Results

After carefully selecting and processing relevant data, we examined the variables displayed in Table 2. The study involved 4868 participants, with the majority from Chittagong at 16.6%. Most participants were from rural areas. Among them, 23.3% had completed secondary education or higher, while 6.1% had received no education. In contrast, a larger proportion of their partners had completed primary education, at 33.7%. An observation of the wealth index revealed that 21.7% of participants belonged to the poorest category and 19.9% to the richest category. Additionally, the majority of families were male-dominated, with 80.3% of family heads being male.

Table 2
Demographic Characteristics of Independent Variables

Factors	Category	Frequency	Percent	Cumulative Percent
Division	Barisal	521	10.7	10.7
	Chittagong	808	16.6	27.3
	Dhaka	706	14.5	41.8
	Khulna	506	10.4	52.2
	Mymensingh	594	12.2	64.4
	Rajshahi	516	10.6	75
	Rangpur	550	11.3	86.3
	Sylhet	667	13.7	100
Type of Residence	Urban	1563	32.1	32.1
	Rural	3305	67.9	100
Highest education level	No education	299	6.1	6.1
	Primary incomplete	840	17.3	23.4
	Primary complete	517	10.6	34
	Secondary incomplete	2078	42.7	76.7
	Secondary incomplete or higher	1134	23.3	100
BMI	Underweight	780	16	16
	Normal weight	2993	61.5	77.5
	Overweight	886	18.2	95.7
	Obese	209	4.3	100
Husband/partner's education level	No education	672	13.8	13.8
	Primary	1639	33.7	47.5
	Secondary	1610	33.1	80.5
	Higher	934	19.2	99.7
	Don't know	13	0.3	100.0
Wealth Index Combined	Poorest	1055	21.7	21.7

Factors	Category	Frequency	Percent	Cumulative Percent
	Poorer	988	20.3	42.0
	Middle	885	18.2	60.1
	Richer	971	19.9	80.1
	Richest	969	19.9	100.0
Age at First Birth	Below 20	3819	78.5	78.5
	20–30	1012	20.7	99.2
	30+	37	0.8	100.0
Working Status	No	3044	62.5	62.5
	Yes	1824	37.5	100.0
Number of living children	1–2	3622	74.4	74.4
	> 2	1246	25.6	100.0
Child is Alive	No	109	2.2	2.2
	Yes	4759	97.8	100.0
During pregnancy, given or bought iron tablets/syrup	Yes	3720	76.4	76.4
	No	1138	23.4	99.8
	Don't know	10	0.2	100.0
Sex of Household Head	Male	3910	80.3	80.3
	Female	958	19.7	100.0

Then, this study used the Boruta algorithm to select features, with 100 iterations performed. The analysis revealed that out of the predictor variables, 13 attributes were confirmed to be important and significantly contribute to the predictive power of the model. These key variables include Division, Type of place of residence, Age at First Birth, Number of living children, Husband/partner's education level, Respondent currently working, Wealth index combined, during pregnancy, given or bought iron tablets/syrup, Highest educational level, Body mass index, Religion, Sex of household head, and Number of household members (listed). Conversely, the Boruta algorithm identified one attribute as confirmed unimportant, indicating it has little impact on the model's predictive performance. One of the unimportant attributes identified is the number of children alive, which can be safely excluded from the model-building process to streamline complexity. These important attributes are presented in Table 3 and the graphical representation of them is provided in Fig. 2.

Table 3
An overview of the Boruta algorithm's findings

Features	Mean Importance	Decision
Division	16.49	Confirmed
Residence	22.57	Confirmed
Age_first_birth	24.04	Confirmed
Living_child	15.47	Confirmed
Partners_Edu	36.19	Confirmed
Working_status	13.88	Rejected
Alive_child	1.89	Confirmed
Wealth_index	61.96	Confirmed
Iron_tab	13.62	Confirmed
Education	41.70	Confirmed
BMI_cate	31.19	Confirmed
Religion	11.95	Confirmed
Head_mem_sex	6.43	Confirmed
Family_mem_cate	5.41	Confirmed

Table 4 illustrates the associations between the categorical variables Division, Residential Area, and Wealth with the delivery status. To evaluate the importance of these correlations, the p-values were used.

Table 4
Association between Caesarean Status as Different Characteristics

Variable	Category	Caesarean Status (%)		Total	P value
		Yes	No		
Division	Barisal	142 (27.36%)	377 (72.64%)	519	< 0.001*
	Chittagong	224 (28%)	576 (72%)	800	
	Dhaka	311 (44.17%)	393 (55.83%)	704	
	Khulna	227 (44.86%)	279 (45.14%)	506	
	Mymensingh	166 (28.04%)	426 (71.96%)	592	
	Rajshahi	198 (38.29%)	319 (61.71%)	517	
	Rangpur	174 (31.52%)	378 (68.48%)	552	
	Sylhet	177 (26.10%)	501 (73.90%)	678	
Residence	Urban	591 (37.76%)	974 (72.24%)	1565	< 0.001*
	Rural	1028(31.12%)	2275 (68.88%)	3303	
Highest Educational Level	No education	47 (15.71%)	252 (84.29%)	299	< 0.001*
	Primary incomplete	143 (17.02%)	697 (82.98%)	840	
	Primary complete	99 (19.14%)	418 (80.86%)	517	
	Secondary incomplete	667 (32.09%)	1411 (67.91%)	2078	
	Secondary complete or higher	663 (58.46%)	471 (61.54%)	1134	
Husband/partner's Education Level	No education	109 (16.22%)	563 (83.78%)	672	< 0.001*

***. Significant at the level of 1%**

	Primary	363 (22.14%)	1276 (77.86%)	1639	
	Secondary	569 (35.34%)	1041 (64.66%)	1610	
	Higher	574 (61.45%)	360 (38.55%)	934	
	Don't know	4 (30%)	9 (70%)	13	
Wealth Index Combined	Poorest	135 (12.79%)	920 (87.21%)	1055	< 0.001*
	Poorer	220 (22.26%)	768 (77.74%)	988	
	Middle	279 (31.52%)	606 (98.48%)	885	
	Richer	370 (38.10%)	601 (61.90%)	971	
	Richest	615 (63.46%)	354 (36.54%)	969	
Respondent Currently Working	No	1156(37.97%)	1888 (62.03%)	3044	< 0.001*
	Yes	463 (25.38%)	1361 (74.62%)	1824	
Age at First Birth	below 20	877 (26.07%)	2486 (73.73%)	3363	< 0.001*
	20–30	716 (48.50%)	760 (51.50%)	1476	
	30+	26 (89.65%)	3 (10.35%)	29	
Number of living children	1–2	1364 (37.65%)	2258 (62.35%)	3622	< 0.001*
	> 2	800 (8.75%)	1137 (91.25%)	1246	
During pregnancy, given or bought iron tablets/syrup	No	248 (21.79%)	890 (78.21%)	1138	< 0.001*
	Yes	1366 (36.72%)	2354 (63.28%)	3720	
	Don't know	5 (50%)	5 (50%)	10	
***. Significant at the level of 1%**					

Child is Alive	No	25 (22.93%)	84 (77.07%)	109	0.02
	Yes	1594 (33.49%)	3165 (66.51%)	4759	
Body Mass Index	Underweight	204 (26.15%)	576 (73.85%)	780	< 0.001*
	Normal	962 (32.14%)	2031 (67.86%)	2993	
	Overweight	334 (37.69%)	552 (62.31%)	886	
	Obese	119 (56.93%)	90 (43.07%)	209	
Religion	Islam	1498 (33.40%)	2987 (66.60%)	4485	< 0.001*
	Hinduism	111 (36.75%)	191 (43.25%)	302	
	Buddhism	10 (12.34%)	71 (87.66%)	81	
Sex of Household head	Male	1282 (32.78%)	2628 (67.22%)	3910	0.159
	Female	337 (35.17%)	621 (64.83%)	958	
Number of Household Members	<=5	925 (33.42%)	1842 (66.58%)	2767	0.163
	6–10	587 (32.27%)	1232 (67.73%)	1819	
	> 10	107 (0.37%)	175 (99.63%)	282	
***. Significant at the level of 1%**					

Based on the data presented in the table, we have identified several significant associations. Firstly, there is a strong link between the division and the likelihood of Caesarean delivery, with the highest percentage found in the Khulna division at 44.84%. Furthermore, we observed a significant association between residence and Caesarean delivery, with urban areas showing a higher percentage (37.76%) compared to rural areas (31.12%). The Wealth Index Combined Status also demonstrates a strong association with Caesarean deliveries, indicating potential financial and healthcare access factors. Body Mass Index (BMI) is significantly associated with Caesarean deliveries, with higher BMI values (Obese) linked to a higher likelihood of Caesarean deliveries. Additionally, the survival status of the child is significantly associated with Caesarean deliveries, with mothers of living children more likely to undergo Caesarean

deliveries. Lastly, the respondent's education level is significantly associated with Caesarean status, with higher education levels linked to a higher likelihood of Caesarean deliveries (p-value = 0.00).

The statistical analysis shows that there is no significant association between the sex of the household head and Caesarean status (p-value = 0.159). In simple terms, whether the household is headed by a male or a female does not seem to affect the likelihood of a Caesarean delivery. This finding suggests that the gender of the household head does not play a significant role in maternal healthcare decision-making regarding the mode of childbirth. Additionally, the number of household members (p-value = 0.163) and the sex of the household head (p-value = 0.159) do not exhibit a statistically significant association with Caesarean status. This means that both the gender of the household head and the size of the household do not appear to influence the likelihood of a Caesarean delivery.

The study evaluated the performance of various Ensemble Models, including Random Forest, Bagging, AdaBoost, GradientBoost, and XGBoost. The assessment involved using key performance metrics derived from the confusion matrix, which included Accuracy, Sensitivity, Specificity, and Precision, in addition to the evaluation of the area under the ROC curve. In the initial phase of the evaluation process, the dataset was split into two distinct subsets in a random manner: a training set and a testing set. The training set comprised 70% of the data, while the remaining 30% constituted the testing set. This partitioning strategy allowed the machine learning models to be trained on a substantial portion of the data while a separate dataset was preserved for unbiased evaluation purposes.

Table 5
Performance evaluation of different ensemble ML methods

Model	Accuracy	Precision	f1 score	Sensitivity	Specificity	AUC Value
RF	0.7371	0.7422	0.8171	**0.9088**	0.4235	0.6662
Bagging	**0.7611**	**0.7746**	**0.8278**	**0.8887**	0.5280	0.7084
AdaBoost	0.7365	0.7471	0.8144	0.8951	0.4468	**0.7833**
GB	0.7187	0.4043	0.5042	0.6699	0.7319	0.7623
XGBoost	0.7234	0.6828	0.5109	0.4081	**0.8962**	0.7727

In this performance evaluation, we evaluated the Random Forest, Bagging, AdaBoost, GradientBoost, and XGBoost ensemble models. Important performance metrics, such as precision, sensitivity, specificity, and AUC Value were used to evaluate the models. From the above Table 5, we can see that the Random Forest model achieved 73.71% accuracy and 74.22% precision. However, it had low specificity at 42.35%, resulting in a higher rate of false positives. The AUC value was 0.6662, indicating moderate discriminative ability. And the F-score is 0.8171.

The Bagging model achieved 76.11% accuracy, meaning it correctly classified most cases. It has a high precision of 77.46%, showing it’s good at identifying positive cases. The F-score of 0.8278 shows a strong balance between precision and recall, while its sensitivity is very high at 88.87%, but its specificity

is 52.80%. The AUC of 0.7084 suggests it's reasonably good at distinguishing between classes. On the other hand, the AdaBoost model reached 73.65% accuracy with a precision of 74.71%, also good at identifying positive cases. Its F-score is 0.8144. But it has a lower specificity of 44.68%, meaning it struggles more with negative cases. However, its AUC of 0.7833 shows it's still good at separating classes.

The GB model has an accuracy of 71.87%, but its precision is relatively low at 40.43%. The F-score of 0.5042 suggests there's room for improvement in balancing precision and recall. With a sensitivity of 67%, it does a decent job identifying positive cases, while its specificity of 73.19% indicates good performance at recognizing negative cases. The AUC of 0.7623 shows it can reasonably distinguish between classes. The XGBoost model has a slightly better accuracy at 72.34% and a precision of 68.28%, meaning it identifies positive cases fairly well. However, the F-score of 0.5109 suggests a need to improve the balance between precision and recall. Its sensitivity is lower at 40.81%, meaning it struggles to identify positive cases, but its specificity is high at 89.62%, making it excellent at detecting negative cases. The AUC of 0.7727 indicates strong class discrimination.

Using 70% of the data for training and 30% for testing, an ROC curve was generated. To identify the best model for predicting birth mode, the area under the ROC curve (AUC) was calculated and visualized in the figure below:

In Fig. 3, the Area under the ROC Curve (AUC) is used to predict birth mode using different ensemble models. The AUC values for each model are as follows: random forest (0.666), Bagging (0.708), AdaBoost (0.785), Gradient Boost (0.762), and XGBoost (0.773). Based on the ROC curve analysis, it is evident that the AdaBoost model has the highest AUC value (0.785), indicating better performance compared to the other ensemble models.

Based on the comparison of models using accuracy, precision, F-score, sensitivity, specificity, and AUC value, it is clear that Bagging outperformed all other models. It achieved the highest scores in accuracy, precision, F-score, and sensitivity. AdaBoost only achieved the highest score for the AUC value, while XgBoost achieved the highest score for the specificity value.

From Fig. 4, we found that wealth Index, number of family member are most important then Division, Partners Education, Respondents education and BMI are moderate and age at first birth, sex of household head, working status, taking iron tablet or syrup, Residence, Number of living children and Religion are the less important variables. Therefore, number of living children and Religion are least important variables.

5. Discussion

The global rate of Caesarean section (CS), often referred to as C-section, is indeed increasing steadily in many countries, including Bangladesh. Cesarean sections now comprise around 18.6% of global births, with Latin America and the Caribbean showing the highest rates. Additionally, there has been a 12.4%

average global increase in cesarean rates from 1990 to 2014, varying annually across regions(Betrán et al., 2016b). However, some women who chose cesarean sections were anxious and dissatisfied with the decision-making process and pre-birth care, leading to a poor overall experience and uncertainty about having more children.(Karlström et al., 2011).

In this study, we seek to identify the causes of the rising caesarean delivery rate and develop a model for forecasting to categorize people as having a caesarean section or a vaginal delivery. In this regard, several ensemble methods were applied in a dataset where caesarean is a target variable and 70% of observations were considered as Train data and 30% of observations were considered as test data.

5.1 Demographic Characteristics

The demographic analysis examined that from the eight divisions Chittagong was the highest data collection area (16.10%) and Khulna was the lowest area (10.40%). Besides 32.10% of observations came from urban and the rest observation is from rural. We discovered the association between different variables and cesarean status which was presented in Table 2. The data highlights that a significant portion of participants, 23.3%, had achieved at least a secondary education, while 6.1% had no education at all. A greater proportion of their partners had completed primary education, at 33.7%. The wealth index revealed a disparity, with 21.7% falling into the poorest category and 19.9% in the richest. Additionally, the majority of families were male-headed, with 80.3% of households led by men, underscoring the prevalent gender dynamics in family leadership.

From the Table 4, this study discovered that the urban women (37.76%) were more interested than rural (31.12%) in CS and Dhaka had a cesarean delivery rate of 44.17%, whereas Khulna had the highest rate at 44.86%. A researcher in Bangladesh found the similar results analyzing BDHS-14 data and recommended that individuals residing in the Khulna division should be particularly encouraged to consider natural delivery options (Khatun et al., 2022). We found that highly educated mothers and husbands showed greater interest in CS. According to our research, working women had more cesarean deliveries than non-working women. The poorest women had just 12.79 percent of their deliveries by caesarean section, while the richest women had 63.46 percent were most interested in Cesarean delivery. The analysis examines those mothers with higher BMI values (Obese 56.93%) are markedly more likely to undergo Caesarean deliveries. The women from Hindu families were more interested than other religions. Also, the women whose family member is less than 6 are more interested in CS.

5.2 Identifying Factors and Model Selection

The factors associated to the birth mode were retrieved using the Boruta algorithm and an association matrix [Table 3]. In Boruta, 13 attributes were confirmed to be important and significantly contribute to the predictive power of the model. These key variables include Division, Type of place of residence, Age at First Birth, Number of living children, Husband/partner's education level, Respondent currently working, Wealth index combined, during pregnancy, given or bought iron tablets/syrup, Highest educational level, Body mass index, Religion, Sex of household head, and Number of household

members (listed). These findings are in line with previous studies in Bangladesh(Khatun et al., 2022, Hasan et al., 2019, Begum et al., 2017).

Several studies employed machine learning algorithms to predict the birth mode and found assortment in the model's performances (Khatun et al., 2022, Kowsher et al., 2020, M. Hasan et al., 2023,Campillo-Artero et al., 2018). To predict the birth mode, this study analyzed various ensemble models including RF, Bagging, GB, AdaBoost, and XGBoost. The study used multiple metrics for comparison, such as accuracy, precision, F-score, sensitivity, specificity, and AUC value. The RF model achieved an accuracy of 0.7371, Bagging achieved 0.7611, AdaBoost achieved 0.7365, GB achieved 0.7187, and XGBoost achieved 0.7234, indicating that Bagging achieved the highest accuracy. Regarding precision, Bagging achieved the highest score at 0.7746, followed by AdaBoost at 0.7471. Bagging also achieved the highest F-score at 0.8270, while GB scored the lowest at 0.5042. Regarding sensitivity, RF scored the highest at 0.9088, followed by Bagging at 0.8887, AdaBoost at 0.8951, GB at 0.6699, and XGBoost at 0.4081. In terms of specificity, XGBoost scored the highest at 0.8962, with RF scoring the lowest at 0.4235. AdaBoost achieved the highest AUC value at 0.7833, while RF scored 0.6662, Bagging 0.7084, GB 0.7623, and XGBoost 0.7727. Overall, Bagging outperformed all the models, achieving the highest scores in four out of six metrics.

In contrary,Khatun (2022) found support vector machine (SVM) with 77% accuracy,Kowsher (2020) found the "impact learning" approach (89% accuracy),M. Hasan (2023) proposed a novel ensemble algorithm named SVXGBRF, combining SVC, XGB, and RF (95.5% accuracy),Campillo-Artero (2018) found random forest model with AUC of 0.94 perform better in predicting birth mode.

6. Conclusion

The worldwide rate of Caesarean section (CS), commonly known as C-section, is steadily increasing in many countries, including Bangladesh. This study aims to identify the reasons for the increasing rate of caesarean deliveries and to create a model for predicting whether individuals will have a caesarean section or a vaginal delivery. To sum up with, this study identified risk factors associated with cesarean section deliveries, such as Division, Type of place of residence, Age at First Birth, Number of living children, Husband/partner's education level, Respondent currently working, Wealth index combined, during pregnancy, given or bought iron tablets/syrup, Highest educational level, and Body mass index. Also, among the considered ensemble method, Bagging is the top-performing model for predicting the birth mode.

Declarations

Ethics approval and consent to participate

A secondary dataset utilized in this investigation was obtained from the Demographic and Health Surveys (DHS) Programme website (https://dhsprogram.com/data/). Accreditation for ethics is not essential for this dataset.

Consent for publication

Each author has given their complete approval for the manuscript to be published.

Competing interests

The writers asserted that they had no conflicting agendas.

Funding

For this research, no specific money from a governmental, private, or nonprofit organization was provided.

Author Contribution

MZA: Conceptualization; Methodology; Data Preparation; formal analysis; Writing-original draft. FAD: Original draft writing, formatting; TBR: Methodology; Writing-original draft. NS: Conceptualization; Methodology; Writing and Editing original draft; Main supervision, MMH: Original draft writing, methodology, formatting.

Data Availability

Data Availability StatementThe datasets analyzed in this study are available in the [dhsprogram] repository, [https://dhsprogram.com/data/].

References

1. Alam, M. S. B., Patwary, M. J. A. & Hassan, M. *Birth Mode Prediction Using Bagging Ensemble Classifier: A Case Study of Bangladesh*. (2021). https://doi.org/10.1109/ICICT4SD50815.2021.9396909
2. Alam, S. S. et al. *The Prevalence and Associated Factors of Caesarean Section at Noakhali Sadar, Bangladesh*. 1–12. (2021). https://doi.org/10.9734/JOCAMR/2021/V13I230219
3. Adashek, J. A., Peaceman, A. M., Lopez-Zeno, J. A., Minogue, J. P. & Socol, M. L. Factors contributing to the increased cesarean birth rate in older parturient women. *Am. J. Obstet. Gynecol.* **169** (4), 936–940 (1993).
4. Ahamed, D. T. & Jahan, D. K. Caesarean Delivery in a Tertiary Care Hospital; A One-Year Survey. *Scholars J. Appl. Med. Sci.* **10** (11), 2036–2041. https://doi.org/10.36347/sjams.2022.v10i11.039 (2022).

5. Ahmmed, F., Manik, M. M. R. & Hossain, M. J. Caesarian section (CS) delivery in Bangladesh: A nationally representative cross-sectional study. *PLOS ONE*. **16** (7). https://doi.org/10.1371/JOURNAL.PONE.0254777 (2021).
6. Begum, T. et al. Indications and determinants of caesarean section delivery: Evidence from a population-based study in Matlab, Bangladesh. *PLoS ONE*, **12**(11). (2017).
7. Begum, T. et al. Monitoring caesarean births using the Robson ten group classification system: A cross-sectional survey of private for-profit facilities in urban Bangladesh. *PLOS ONE*. **14** (8). https://doi.org/10.1371/JOURNAL.PONE.0220693 (2019).
8. Betrán, A. P. et al. The increasing trend in caesarean section rates: Global, regional and national estimates: 1990–2014. *PLoS ONE*, *11*(2). (2016a).
9. Campillo-Artero, C., Serra-Burriel, M. & Calvo-Pérez, A. Predictive modeling of emergency cesarean delivery. *PLoS ONE*, **13**(1). (2018).
10. Dong, X., Yu, Z., Cao, W., Shi, Y. & Ma, Q. A survey on ensemble learning. In *Frontiers of Computer Science* (Vol. 14, Issue 2, pp. 241–258). Higher Education Press. (2020). https://doi.org/10.1007/s11704-019-8208-z
11. Gibbons, L. et al. The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage. *World Health Rep.* **30** (1), 1–31 (2010).
12. Gy, K. N. N. et al. (n.d.). Cesarean section birth: Background, Indications, Epidemiology. Retrieved from https://emedicine.medscape.com/article/1609430-overview
13. Hasan, F., Alam, M. M. & Hossain, M. G. Associated factors and their individual contributions to caesarean delivery among married women in Bangladesh: Analysis of Bangladesh demographic and health survey data. In *BMC Pregnancy and Childbirth* (Vol. 19, Issue 1). BioMed Central. (2019). https://doi.org/10.1186/s12884-019-2588-9
14. Hasan, M. et al. Ensemble Based Machine Learning Model for Early Detection of Mother's Delivery Mode. *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 1–6. (2023).
15. Hasan, F., Alam, M. M. & Hossain, M. G. Associated factors and their individual contributions to caesarean delivery among married women in Bangladesh: Analysis of Bangladesh demographic and health survey data. *BMC Pregnancy Childbirth*. **19** (1), 433–433. https://doi.org/10.1186/S12884-019-2588-9 (2019).
16. Hossain, M. A. et al. Rising Trends of Cesarean Section in Bangladesh: Associated Factors and Long-Term Complications on Health of Mother and Children. *J. Maternal Child. Health*. **7** (5), 532–542. https://doi.org/10.26911/thejmch.2022.07.05.04 (2022).
17. Irfan, M. M. & Nisar, N. Factors associated with caesarean delivery in district Badin, Sindh, Pakistan: a cross-sectional analysis. *Pakistan J. Med. Sci.* **34** (6), 1411–1416 (2018).
18. Islam, M. M. & Noor, F. M. Prevalence and factors of cesarean delivery among Bangladeshi reproductive aged women: Evidence from multiple Indicator cluster survey 2019 data. *J. Public*

Health. 1–10. https://doi.org/10.1007/S10389-021-01551-9 (2021).

19. Karlström, A., Nystedt, A. & Hildingsson, I. A comparative study of the experience of childbirth between women who preferred and had a caesarean section and women who preferred and had a vaginal birth. *Sex. Reproductive Healthc.* **2** (3), 93–99. https://doi.org/10.1016/j.srhc.2011.03.002 (2011).
20. Khatun, F., Rahman, R., Das, S. & Talukdar, A. Identification of significant factors for birth mode in bangladesh: Evidence from cross-sectional study. *Khulna Univ. Stud.* 664–675. https://doi.org/10.53808/kus.2022.icstem4ir.0121-se (2022).
21. Kowsher, M. et al. Predicting the Appropriate Mode of Childbirth using Machine Learning Algorithm. *Int. J. Adv. Comput. Sci. Appl.* **12** (5). https://doi.org/10.14569/IJACSA.2021.0120582 (2021).
22. Kowsher, M., Prottasha, N. J., Tahabilder, A. & Islam, M. B. *Machine Learning Based Recommendation Systems for the Mode of Childbirth* (pp. 295–306). (2020). https://doi.org/10.1007/978-3-030-52856-0_23
23. Kumar, P. & Sharma, H. Prevalence and determinants of socioeconomic inequality in caesarean section deliveries in Bangladesh: An analysis of cross-sectional data from Bangladesh Demographic Health Survey, 2017-18. *BMC Pregnancy Childbirth*. **23** (1). https://doi.org/10.1186/s12884-023-05782-4 (2023).
24. Kursa, M. B. & Rudnicki, W. R. Feature Selection with the Boruta Package. In *JSS Journal of Statistical Software* (Vol. 36). (2010). http://www.jstatsoft.org/
25. Lee, S., Cha, D. H., Park, C. W. & Kim, E. H. Maternal and Neonatal Outcomes of Elective Induction of Labor at 39 or More Weeks: A Prospective, Observational Study. *Diagnostics* **13** (1). https://doi.org/10.3390/diagnostics13010038 (2023).
26. Liu, Y., Wang, Y. & Zhang, J. New Machine Learning Algorithm: Random Forest. In *LNCS* (Vol. 7473). (2012).
27. Sana, A., Razzaq, S. & Ferzund, J. Automated Diagnosis and Cause Analysis of Cesarean Section Using Machine Learning Techniques. *Int. J. Mach. Learn. Comput.* 677–680. https://doi.org/10.7763/ijmlc.2012.v2.213 (2012).
28. Sandall, J. et al. Short-term and long-term effects of caesarean section on the health of women and children. In *The Lancet* (Vol. 392, Issue 10155, pp. 1349–1357). Lancet Publishing Group. (2018). https://doi.org/10.1016/S0140-6736(18)31930-5
29. Souza, J. P. et al. Caesarean section without medical indications is associated with an increased risk of adverse short-term maternal outcomes: the 2004–2008 WHO Global Survey on Maternal and Perinatal Health. *BMC Med.* **8** (1), 1–10 (2010).
30. Minsart, A. F. et al. Maternal and neonatal outcomes of elective induction of labor and elective cesarean section at 39 weeks in uncomplicated singleton pregnancies: aretrospectiveanalysis. *BMC Pregnancy Childbirth*. **18** (1), 57 (2018).
31. Nahar, Z., Sohan, M., Hossain, M. J. & Islam, M. R. *Unnecessary Cesarean Section Delivery Causes Risk to Both Mother and Baby: A Commentary on Pregnancy Complications and Women's Health* (A

Journal of Medical Care, 2022).

Figures

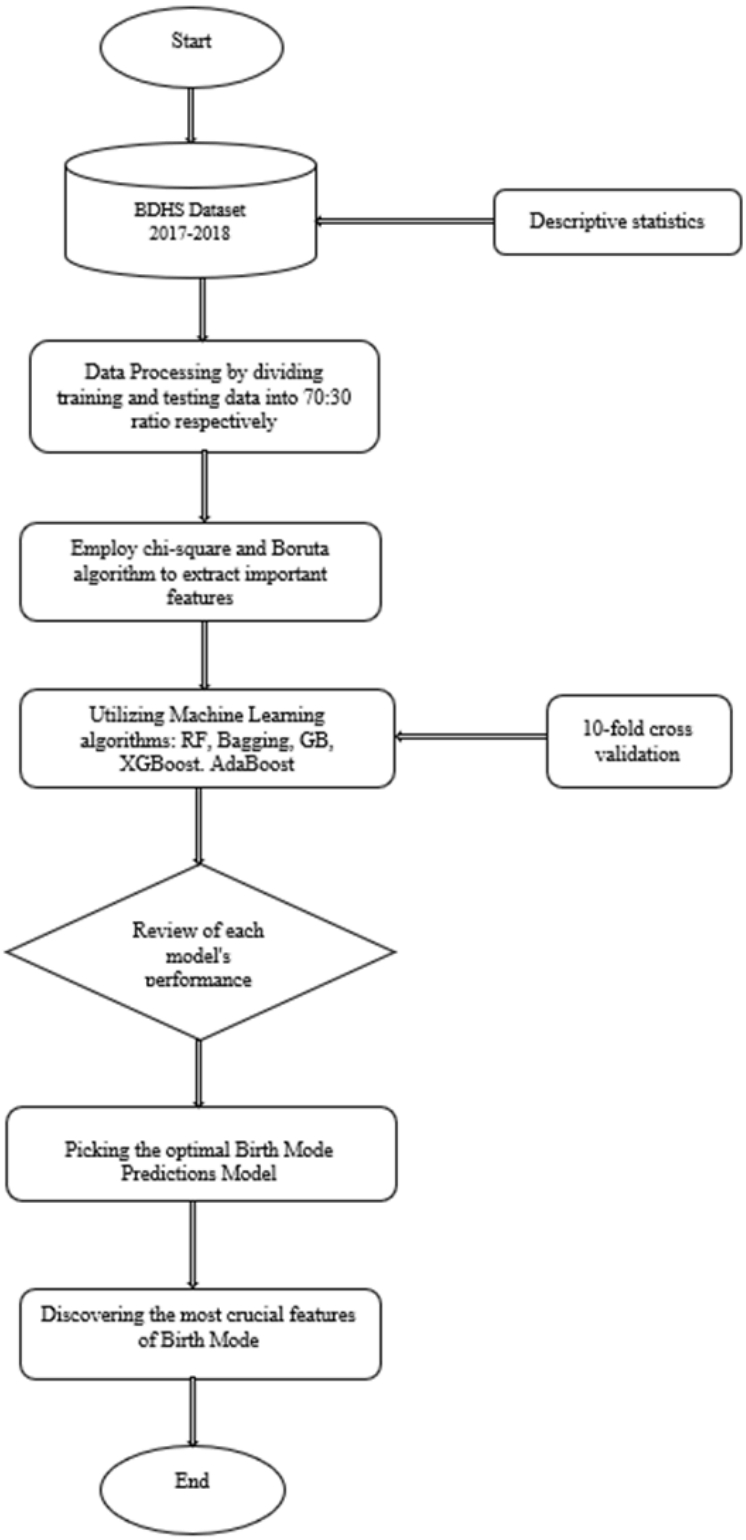


Figure 1

Comprehensive research approach

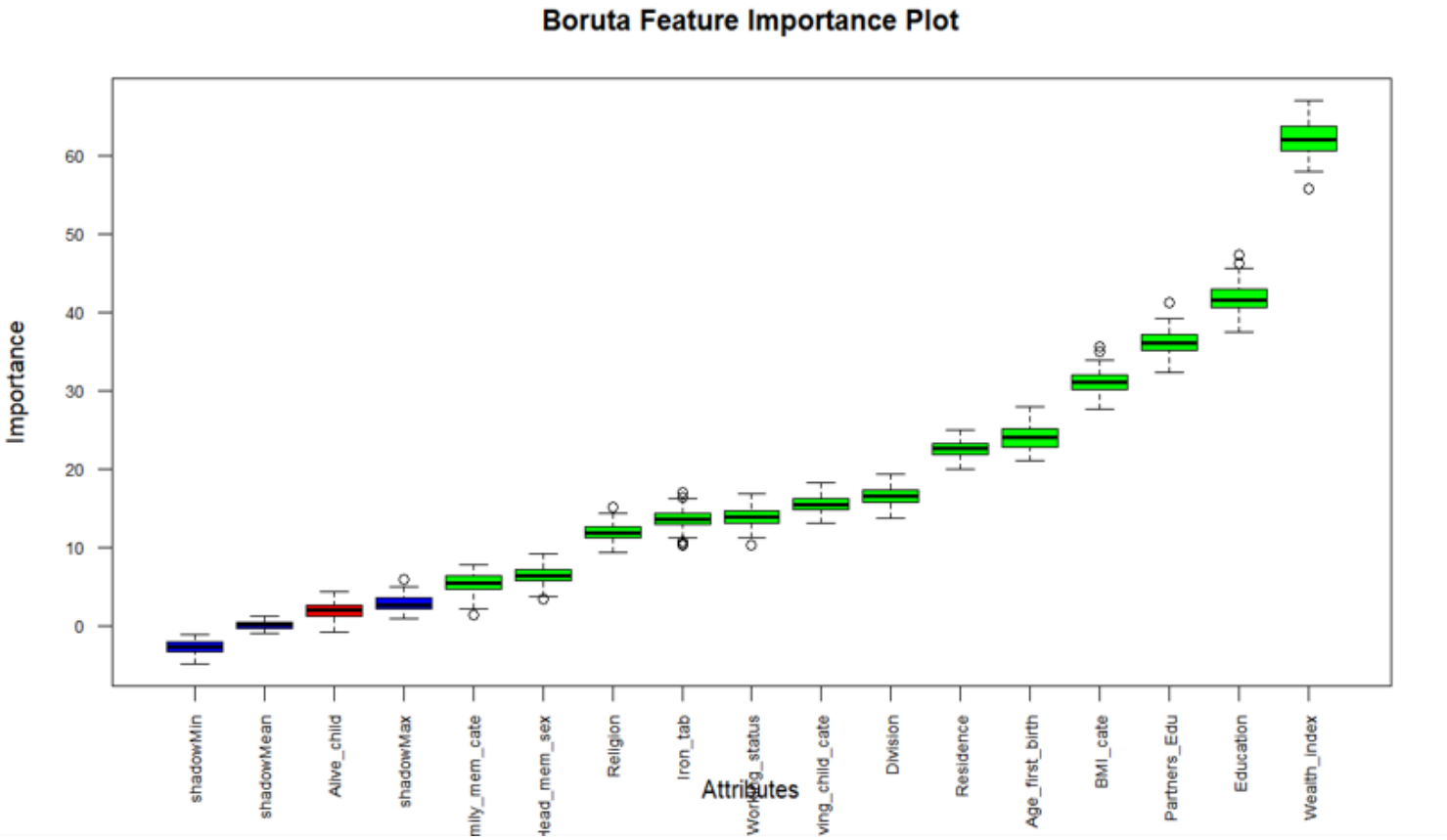


Figure 2

Feature Selection Using Boruta Algorithm

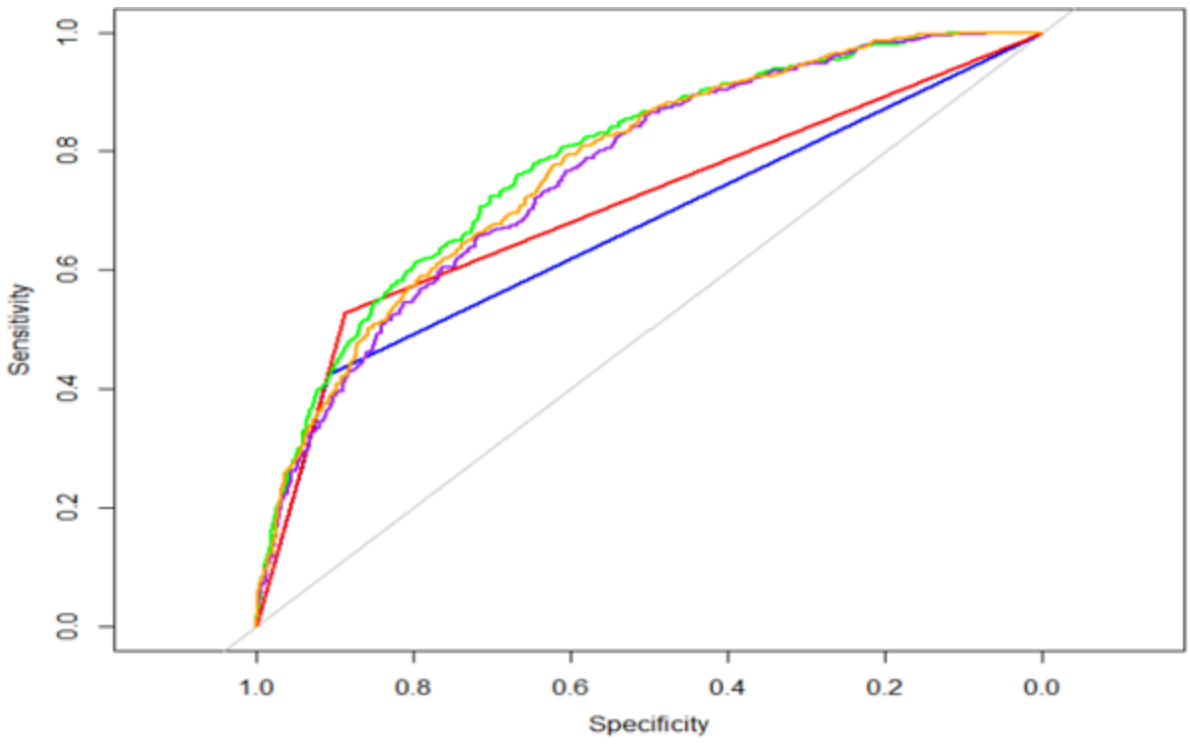


Figure 3

ROC Curve for predicting Birth Mode using different Ensemble methods

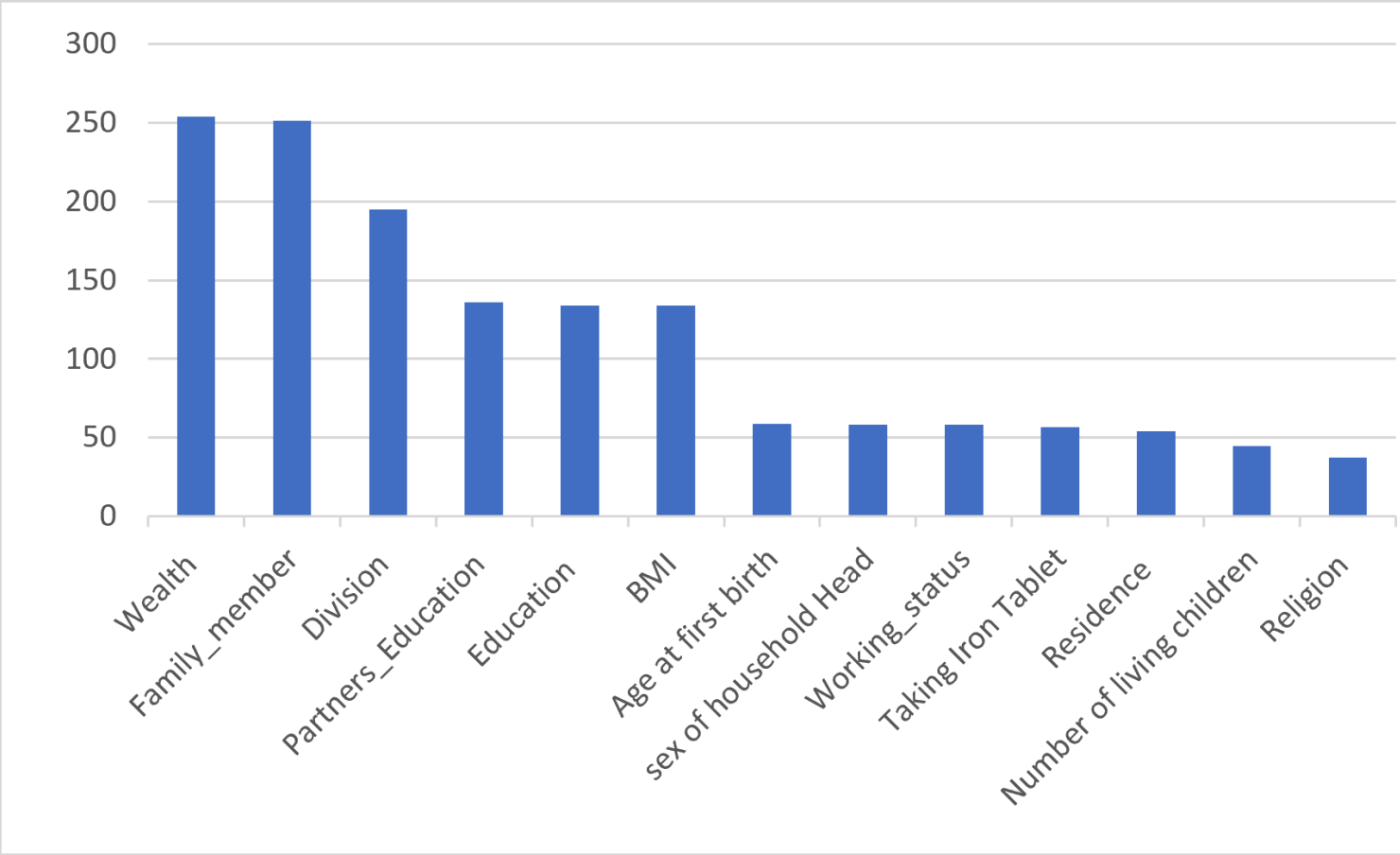


Figure 4

Variable Importance Measures for Bagging Model