Abstract

Maternal mortality and difficulties during childbirth are important delivery issues in the majority of underdeveloped nations, including Bangladesh. Our objective is to figure out what's behind the rise in caesarean deliveries, build a link between critical components and caesarean section, and seek the key factors are related with caesarean births. We combined 2014 and 2017-18 BDHS data and found that key factors. We employed the logistic regression model to identify and quantify the effects of variables on birth mode and proposed five supervised machine learning approaches to find out the best performing model to predict the birth mode. The performance of these algorithms is evaluated by accuracy, precision, recall and F1 score. This study shows that Division, Highest education level, Wealth index, Total number of children born, BMI and Age of respondent at first birth, Husband’s education level and Respondent's current working status have been statistically linked to caesarean delivery, and using the logistic regression approach all of the categories of the variables are also statistically significant (p value- < 0.05). Since the rich and the upper class are more likely to have a caesarean section than the poor and middle class, they need to be made aware of the disadvantages of this procedure. At the same time, they need to be encouraged to adopt a common approach. As a result, our work has a substantial influence and plays a role in developing a preventative strategy for caesarean situations, particularly in Bangladesh.

Keywords: Maternal mortality, birth mode, machine learning, evaluation, significant

Introduction

Maternal mortality and childbirth complications are major delivery challenges in the majority of developing countries like Bangladesh. According to the World Health Organization, early identification of risks associated with the delivery of a pregnant woman can considerably reduce the death rate. One of the most common causes for elective Caesarean Section (CS) without a medical indication is maternal request. Fear of discomfort, well-being, and convenience are also prominent reasons for opting for CS. Cesarean delivery (C-section) is a surgical technique that is frequently done or advised when the mother's or child's life is in danger (Khan, 2017).
We want to make sure that both the mother and the child are safe during the delivery process. During a baby delivery, the physician's ability to make decisions in a short period of time becomes extremely difficult. A poor judgment puts the mother's life at jeopardy and can have serious consequences (Souza, 2010). A bad decision puts the mother's life at danger and can be hazardous to the newborn baby's health. As a result, the global rate of cesarean birth has been growing in recent years. A variety of variables contribute to the rising trend of cesarean births. The use of cesarean sections (C-sections) is increasing globally, including in Bangladesh, and this is having a detrimental impact.

However, the global rate of caesarean section (CS) is gradually rising. Bangladesh is also seeing an increase in the number of CS cases. However, needless CS has a negative impact on the health of both the mother and the kid. Identifying the risk factors for cesarean birth is critical to avoiding the problem (Symonds, 2003). Pattern extraction is made easier by using Machine Learning (ML) approaches. The use of machine learning (ML) for prediction and classification in health care is growing rapidly (Molina, 2015). Our goal is to discover the causes of the rising rate of cesarean delivery and to develop a prediction model to categorize people as either cesarean section. To predict the birth mode, we used machine learning algorithms like Decision Tree (DT), Random Forest and Support Vector Machine (SVM), LightGBM and XGBoost, also compare their performance with each other.

**Motivation of the study**
- CS is a complicated issue and varies from area to area, country to country and obviously patients to patients.
- We need to identify associated factors related to CS which can help to identify the complications of the patients.
- The relevant issues associated with caesarean delivery need to be identified, a relationship between major elements and caesarean section need to be established.
- Key variables associated with caesarean deliveries need to be investigated.

**Reasons for using logistic regression and machine learning**

**Logistic regression**
- This method assists in identifying important factors.
- Find out the relationship between each of these factors and the dependent variable.

**Machine learning**
- Estimate birth mode.
- Identify potential maternal risks during pregnancy.
- Predict the best mode of birth and detect particular problems during childbirth.

**Objectives**
The aim of this research is to predict the birth mode for caesarean using ML. The specific objectives of this research are as below:
- Established a connection between essential components and caesarean section.
- Use Chi square and logistic regression to find the association and to find the significant factors which are associated in the birth mode for caesarean.
- Use some machine learning methods for predicting Birth Mode and compare the performance of those algorithms.
Related Work

The majority of the time, vaginal birth is the best and most commonly used option, whereas C-sections are only used when a natural delivery is not possible or when the mother's or infant's health is compromised.

A study reported that for both women and children, there are various hazards linked with C-section. Babies born through C-section are more likely to acquire asthma, obesity, diabetes and other diseases. Furthermore, C-section newborns are more likely to have developmental issues such as attention deficit hyperactivity disorder, autism spectrum disorder, learning difficulties, and so on (Albes, 2018). From another study we observed, a caesarean section causes many problems for both the infant and the mother. Allergies, asthma, and diabetes are substantially more common in children born via caesarean section. Another factor contributing to the higher CS rate is a shift in midwives' and obstetricians' opinions toward (Sana, 2012). In 2021, a researcher reported that in Bangladesh, the c-section rate jumped by 51% between 2017 and 2018, with 77 percent of those c-sections being unnecessary (Chen, 2016).

Another researcher reported that according to a review study conducted on 137 nations in 2010, the rate of C-section was less than 10% in 54 countries, greater than 15% in 69 countries, and between 10-15% in 41 countries. Iran was placed second with 41.9 percent, trailing only Brazil with 45.9 percent (Molina, 2015). A research paper intended that one of the most common causes for elective CS with no medical indication is maternal desire. The CS rate varies due to perceptions of dread of childbirth, concern about fetal safety and well-being, and convenience (Hasan, 2019).

Another paper indicated that the largest prevalence of C-sections occurred among females with a secondary or higher level of education. Because education is closely related to women's autonomy, they are more monetarily solvent and, because they mainly live in cities, can choose to give birth through C-section. Some research. However, demonstrate that women's decision of C-Section is no obvious connection to their educational level CS (Patwary, 2019).

Another study involving 404 women was conducted at Al-Yarmouk Teaching Hospital in Baghdad, Iraq, a cross-sectional study. CS terminated the current pregnancy in 154 (38.2%) of the 404 women. There was a significant link between the mother's education and the type of delivery (p < 0.036). Women with less than a primary school education had a substantially greater rate of CS (41.6%) than those with a secondary school or higher education (30.8%) (p <0.036). Women who had previously had an abortion had a considerably lower rate of CS than women who had never had an abortion (p < 0.04). This research, however, revealed that the method of delivery has nothing to do with the neither the women's age nor the husband's occupation (Al-Kubaisy, 2014).

Materials and Methods

We went through four major stages to develop our proposed model: Dataset generation, data preprocessing, model training, and the model performance analysis (Souza, 2010). Data cleaning and feature selection were performed after the dataset was collected (Khan, 2017). After that, prediction models were developed using machine learning methods, and their performance was assessed (Khan, 2017). In addition, we looked into the relationship between factors and caesarean section (Khan, 2017) and also perform binary logistic regression model.

Dataset

In this analysis we used Bangladesh Demographic and Health Survey (BDHS-2014) and Bangladesh Demographic and Health Survey (BDHS-2017-18) datasets. We merged both the dataset and a total of 9593 sample were being used for our analysis.

Data preprocessing

If there is a missing value, it does not have a good impact on our analysis. So, collecting the required dataset then we had to remove the missing data from the final dataset and the sample size for our final dataset was 9593. Based on previous research, 10 characteristics linked with delivery mode were chosen (Khan, 2017).
After that we performed association in regards of birth mode and calculate p-values at 95% CI for checking the significance features. Finally, 10 variables were chosen for this investigation based on the p-value. Division, place of residence, respondent’s education level, wealth index, total number of children born, age of respondent at 1st birth, husband’s educational level, body mass index (BMI), respondent currently working and age of the respondent were selected features for our study.

**Logistic regression**

After finding the associated features, we performed the binary logistic for birth mode of caesarean with those associated features and found eight significant features associated with the birth mode of caesarean section.

**Methods used for classification**

**Decision Tree (DT)**

The most commonly used algorithms in machine learning. At first it works with root nodes where it makes branches with some rules or criteria derived from input data set. Simply works with Yes/No decision. When desired feature fulfilled it makes a result otherwise makes branches which we called sub trees (decision node) and the result is called leaf nodes.

**Support Vector Machine (SVM)**

Support Vector Machine is a common supervised learning technique for Classification and Regression. However, it is most commonly employed in Machine Learning for Classification challenges. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes.

**Random Forest**

The hyper parameters of random forest are virtually identical to those of a decision tree or a bagging classifier. Fortunately, you may utilize the random forest classifier-class instead of combining a decision tree and a bagging classifier. The main difference between decision tree and random forest is, decision tree makes some rules with the training data set and make result (sometimes it can be over fitted) but random forest makes different decision trees and collect their result to create a final result.

**LightGBM**

Microsoft proposed LightGBM, a data model based on GBDT, in 2017. GBDT combines weak learners to generate a powerful one, similar to other boosting techniques. However, because each tree in the GBDT method learns the conclusions and residuals of all previous trees, the decision tree employed in the process can only be a regression tree. A current residual regression tree is constructed by using the residual of each anticipated result and target value as the aim of next learning. The final predicted output is made up of the findings of many decision trees (Friedman, 2001). The LightGBM algorithm employs a leaf-wise generation technique to reduce training data. When compared to the traditional way of growing the same leaf level (depth)-wise, the leaf-wise method can reduce more losses. Additionally, the extra parameter is utilized to limit the decision tree’s depth and avoid overfitting.

**XGBoost**

In 2016 we were introduced to the XGBoost optimization model, which has properties of both linear and tree models. The model prediction values of all decision trees were aggregated into the final one by constructing numerous trees to suit the residuals, and the model was trained using the gradient boosting decision tree (GBDT) algorithm (Chen, 2016). Decision trees are created sequentially in this algorithm. In XGBoost, weights are very significant. All of the independent variables are given weights, which are subsequently fed into the decision tree, which predicts outcomes. The weight of factors that the tree predicted incorrectly is
increased, and these variables are fed into the second decision tree. Individual classifiers/predictors are then combined to form a more powerful and precise model. It can solve problems including regression, classification, rank, and user-defined prediction.

Results

From Table 1, we found that Dhaka region has the highest percentage (27.6%) of the respondents than any other divisions in Bangladesh, where Barisal shows 9.1% of Caesarian and respondents from Sylhet region has 10% of caesarian and due to p-value is less than 0.01, these variables are significant and they are associated to each other. The percentage of the delivery by caesarian section in rural area is 53.8% and in urban area it is 46.2% which is associated with the caesarian section with the p-value is less than 0.01. Here we observed that the percentage of the secondary educational level of the mother is the highest percentage of the delivery by caesarian section which is 50.8%, where highest educational level of the mother is 31.0%, no educational level of the mother is 3.4% and the p-value is less than 0.01. So that there is an association between the respondent’s education level and the delivery of the caesarian section. The respondents from the richest family have the highest percentage of delivery 40.2%, where the percentage of the poorest family is lower 7.2% than the richest family and p-value is less than 0.01. Here the percentage of delivery by caesarian section is 81.4%, where the children is less than 2 and the p-value is less than 0.01. The percentage of the caesarian section is higher (66.3%) where the age of respondent at 1st birth is less than and equal to 20 years than the percentage of the age of respondent at 1st birth is greater than 20 years and here the p-value is <0.01. In this analysis, 36.3% of caesarean delivery occurs when the husbands are secondary educated and the p-value is <0.01. We observed that the respondents who are currently not working have the highest percentage (75.9%) than the respondents who are working, where the p-value is <0.01. The percentage caesarean section of the normal weight women is higher (54.4%) than the percentage of the underweight (11.9%), and the p-value is <0.01. The respondent’s age between 20-24 years has the highest percentage of caesarian section. Where the respondent’s age between 45-49 years has the lowest percentage (0.1%), and the p-value is 0.000, which is less than 0.01. So, the association between the response and exploratory variables is statistically significant.

The following graph shows that the respondents who had non-Caesarean delivery has higher percentage than the Caesarean delivery respondents, where non-Caesarean has 71.30% and Caesarean has 28.70%.
Table 1. An Association between variables and birth mode for caesarean Section (CS)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Group</th>
<th>Over all Sample</th>
<th>Birth mode for CS (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Division</td>
<td>Barisal</td>
<td>1074</td>
<td>252 (9.1 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Chittagong</td>
<td>1698</td>
<td>405 (14.7 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dhaka</td>
<td>2150</td>
<td>760 (27.6 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Khulna</td>
<td>1056</td>
<td>420 (15.2 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rajshahi</td>
<td>1084</td>
<td>358 (13.0 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rangpur</td>
<td>1119</td>
<td>287 (10.4 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sylhet</td>
<td>1412</td>
<td>275 (10.0 %)</td>
<td></td>
</tr>
<tr>
<td>Place of residence</td>
<td>Urban</td>
<td>3175</td>
<td>1275 (33.1 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>6418</td>
<td>1482 (66.9 %)</td>
<td></td>
</tr>
<tr>
<td>Respondent’s education level</td>
<td>No education</td>
<td>932</td>
<td>95 (3.4 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Primary</td>
<td>2663</td>
<td>405 (14.7 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secondary</td>
<td>4571</td>
<td>1401 (50.8 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Higher</td>
<td>1427</td>
<td>856 (31.0 %)</td>
<td></td>
</tr>
<tr>
<td>Wealth Index</td>
<td>Poorest</td>
<td>2055</td>
<td>199 (7.2 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Poorer</td>
<td>1903</td>
<td>332 (12.0 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>1785</td>
<td>457 (16.6 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Richer</td>
<td>1960</td>
<td>661 (24.0 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Richest</td>
<td>1890</td>
<td>1108 (40.2 %)</td>
<td></td>
</tr>
<tr>
<td>Total No. of children born</td>
<td>1-2 children</td>
<td>6715</td>
<td>2245 (81.4 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>&gt; 2 children</td>
<td>2878</td>
<td>512 (18.6 %)</td>
<td></td>
</tr>
<tr>
<td>Age of respondent at 1st birth</td>
<td>&lt;=20 years</td>
<td>7654</td>
<td>1827 (66.3 %)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>&gt;20 years</td>
<td>1939</td>
<td>930 (33.7 %)</td>
<td></td>
</tr>
<tr>
<td>Husband’s education level</td>
<td>No education</td>
<td>1735</td>
<td>208 (7.5 %)</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>
From Table 2 we get logistic regression results and it shows us that all of the division except Chittagong give us the significant results (P value < .05). Barisal, Dhaka, Khulna, Rajshahi, Rangpur are 1.249 times, .1.776 times, 2.332 times, 1.932 times and 1.447 times respectively compared to Sylhet for caesarean delivery.

For respondent's highest education level, respondent with no education is less likely to have caesarean delivery compared to highly educated respondent. Respondent with no education, primary and secondary are .352 times, .447 times and .659 times likelihood respectively compared to higher educated respondent for caesarean delivery with 95% level of significance (P value < .05).

The probability of wealth index is small for all classes of people compared to richest respondent though it gives significance value (P value < .05) for all classes for caesarean delivery. The probability of having caesarean delivery for respondent with 1-2 children is 1.659 times (P value < .05) compared to respondent with more than 2 children. The probability of having caesarean delivery for respondent aged less than and equal to 20 years is .667 times (P value < .05) than respondent aged more than 20 years. If the respondent’s husband is educated then they are more likely to have caesarean delivery and with secondary education it is .703 times, with primary it is .586 times and with no education it is .493 times compared to higher educated husband of respondent with 95% level of significance (P value < .05).

The probability of having caesarean delivery is .655 times, .396 times and .293 times for overweight respondent, normal weight respondent and underweight respondent compared to obese respondent with 95% level of significance (P value < .05). Respondent who are currently not working have 1.320 times (P value < .05) likelihood to have caesarean delivery compared to respondent who are currently working.

Table 1&2 show the results of our investigation into the relationship between factors and caesarean section (CS). In addition, to predict Birth Mode, we applied various machine learning techniques. We trained ten classifiers on a sample of 80% of the individuals in each group (training dataset) and verified them on the remaining 20%. (Test dataset). We compared the results of all of the algorithms listed in Table 3.
Table 2. A binary logistic regression analysis considering birth mode for caesarean

<table>
<thead>
<tr>
<th>Variables</th>
<th>Group</th>
<th>B</th>
<th>S. E</th>
<th>P value</th>
<th>OR</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Division</td>
<td>Barisal</td>
<td>.222</td>
<td>.112</td>
<td>.048</td>
<td>1.249</td>
<td>1.002 1.557</td>
</tr>
<tr>
<td></td>
<td>Chittagong</td>
<td>-.057</td>
<td>.099</td>
<td>.567</td>
<td>.945</td>
<td>.778 1.148</td>
</tr>
<tr>
<td></td>
<td>Dhaka</td>
<td>.574</td>
<td>.093</td>
<td>&lt;0.01</td>
<td>1.776</td>
<td>1.481 2.130</td>
</tr>
<tr>
<td></td>
<td>Khulna</td>
<td>.847</td>
<td>.106</td>
<td>&lt;0.01</td>
<td>2.332</td>
<td>1.896 2.869</td>
</tr>
<tr>
<td></td>
<td>Rajshahi</td>
<td>.658</td>
<td>.107</td>
<td>&lt;0.01</td>
<td>1.932</td>
<td>1.565 2.385</td>
</tr>
<tr>
<td></td>
<td>Rangpur</td>
<td>.369</td>
<td>.112</td>
<td>&lt;0.01</td>
<td>1.447</td>
<td>1.161 1.803</td>
</tr>
<tr>
<td></td>
<td>Sylhet*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of residence</td>
<td>Urban</td>
<td>.093</td>
<td>.059</td>
<td>.118</td>
<td>1.097</td>
<td>.977 1.232</td>
</tr>
<tr>
<td></td>
<td>Rural*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent’s education level</td>
<td>No education</td>
<td>-1.044</td>
<td>.148</td>
<td>&lt;0.01</td>
<td>.352</td>
<td>.264 .470</td>
</tr>
<tr>
<td></td>
<td>Primary</td>
<td>-.805</td>
<td>.101</td>
<td>&lt;0.01</td>
<td>.447</td>
<td>.367 .545</td>
</tr>
<tr>
<td></td>
<td>Secondary</td>
<td>-.417</td>
<td>.079</td>
<td>&lt;0.01</td>
<td>.659</td>
<td>.564 .769</td>
</tr>
<tr>
<td></td>
<td>Higher*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wealth Index</td>
<td>Poorest</td>
<td>-1.437</td>
<td>.109</td>
<td>&lt;0.01</td>
<td>.238</td>
<td>.192 .294</td>
</tr>
<tr>
<td></td>
<td>Poorer</td>
<td>-1.029</td>
<td>.096</td>
<td>&lt;0.01</td>
<td>.357</td>
<td>.296 .431</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>-.789</td>
<td>.086</td>
<td>&lt;0.01</td>
<td>.454</td>
<td>.384 .537</td>
</tr>
<tr>
<td></td>
<td>Richer</td>
<td>-.585</td>
<td>.076</td>
<td>&lt;0.01</td>
<td>.557</td>
<td>.480 .646</td>
</tr>
<tr>
<td></td>
<td>Richest*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total No. of children born</td>
<td>1-2 children</td>
<td>.506</td>
<td>.082</td>
<td>&lt;0.01</td>
<td>1.659</td>
<td>1.411 .950</td>
</tr>
<tr>
<td></td>
<td>&gt;2 children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age respondent at 1st birth</td>
<td>&lt;=20 years</td>
<td>-.405</td>
<td>.070</td>
<td>&lt;0.01</td>
<td>.667</td>
<td>.581 .766</td>
</tr>
<tr>
<td></td>
<td>&gt;20 years*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Husband's education level</td>
<td>No education</td>
<td>-.706</td>
<td>.114</td>
<td>&lt;0.01</td>
<td>.493</td>
<td>.395 .617</td>
</tr>
<tr>
<td></td>
<td>Primary</td>
<td>-.534</td>
<td>.088</td>
<td>&lt;0.01</td>
<td>.586</td>
<td>.493 .697</td>
</tr>
<tr>
<td></td>
<td>Secondary</td>
<td>-.353</td>
<td>.077</td>
<td>&lt;0.01</td>
<td>.703</td>
<td>.605 .816</td>
</tr>
<tr>
<td></td>
<td>Higher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>Underweight (BMI&lt;=18.5)</td>
<td>-.1227</td>
<td>.147</td>
<td>&lt;0.01</td>
<td>.293</td>
<td>.220 .391</td>
</tr>
<tr>
<td></td>
<td>Normal weight</td>
<td>-.926</td>
<td>.133</td>
<td>&lt;0.01</td>
<td>.396</td>
<td>.305 .514</td>
</tr>
</tbody>
</table>
We used a variety of evaluation metrics in our research, including Accuracy, Precision, Recall, and F1 Score.

The performance of the classifiers is shown in Table 3. SVM has the highest accuracy of 0.77 with a F1 score of 0.74, while DT has an accuracy of 0.71 with a F1 score of 0.70. The accuracy of the other algorithms RF, LightGBM, and XGBoost are 0.74, 0.76, 0.75 with F1 values of 0.73, 0.75, and 0.73 respectively.

Table 3. Performance of the algorithms

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT</td>
<td>0.71</td>
<td>0.70</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>RF</td>
<td>0.74</td>
<td>0.73</td>
<td>0.74</td>
<td>0.73</td>
</tr>
<tr>
<td>SVM</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>LightGBM</td>
<td>0.76</td>
<td>0.75</td>
<td>0.76</td>
<td>0.75</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.75</td>
<td>0.73</td>
<td>0.75</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Discussion

C-section is an effective intervention for saving the lives of mothers and newborns during complications during childbirth. Using representative data, this study sought to identify risk factors for CS delivery in Bangladesh. Our analysis demonstrates that Division, Highest education level, Wealth index, Total No. of children born, Age of respondent at 1st birth, Husband’s education level, Respondent currently working and BMI (Body mass index) have statistically significant associations with birth mode for caesarean.

Here in this study, it is noted that the percentages of CS delivery in Bangladesh is 28.7 % which is comparatively very low than many other countries. According to a global survey, China has the highest prevalence of CS delivery (46.2%) (Lumbiganon, 2010). The other countries’ rates ranged from 1.62% (Angola) to 42.0 % (Paraguay) (World Health Organization, 2009; Shah, 2009). Over the last two decades, the
incidence of CS has been reported as 26 % among the Iranian community, and as high as 60 percent in one case study (Vaziran, 2000; Aram, 2002).

According to our research, BMI (Body mass index) is highly and statistically significant and more likely associated in birth for CS and here mothers with a normal weight (54.4%), over weight (26.1%) and Obese (7.7%) are highly correlated in a birth mode for CS. According to our findings, mothers with a higher BMI are more likely to have a cesarean birth. Weight loss can increase the likelihood of vaginal birth after a cesarean section in obese women (Callegari, 2014). Obesity has been linked to poor obstetric outcomes and higher CS (Weiss, 2004; Al-Kubaisy, 2014).

Women with the highest wealth quintile were more likely to be overweight or obese. Six times more likely to be overweight or obese than with women in the lowest wealth quintile (NIPORT, 2013). Overweight or obese women are more common in middleclass and wealthy households, and as the findings indicate, they are more likely to have C-section deliveries than poor households (Press Information Bureau, 2018; Bhartia, 2020; Sandall, 2018; Bahl, 2010). we discovered that older women did not more likely than younger women to have CS delivery. This discovery was made supported by a study from China (Hou, 2017). In this study by the wealth index result people from richest family (40.2 %) and richer family (24%) are highly associated and correlated in a birth mode for CS.

Furthermore, C-section deliveries are most common in India's southern states (Telangana, Kerala, and Andhra Pradesh) (Mishra, 2019; IIPS, 2021; Press Information Bureau, 2018). In this study, Khulna (15.2 %) and Dhaka (27.6%) have the highest significant effect than the other division in Bangladesh. Nassar and Sullivan claimed that most demographic shifts are caused solely by age and parity (birth order) because the primary caesarean rate for first-time mothers is high birth to mothers above the age of 30. Mothers who have low birth order women who have a C-section, explained. The decision was made mostly due to their superiority and danger of pregnancy and complications during birth (Khawaja, 2004; Stanton, 2006; Villar, 2006).

Conclusion
Analysis shows that respondent from Khulna division is more likely to have caesarean deliveries, Khulna residents should be interested in normal delivery. Respondents who are currently not working need to be more aware of cesarean delivery. Higher educated people should be more interested in normal delivery. In machine learning analysis, SVM performed best to predict the birth mode with highest accuracy with 77%. Also, DT gave the lowest accuracy at 71% to predict the birth mode. We recommend other researchers to perform other machine learning algorithms as well as feature importance and others performance measures to evaluate the analysis.

Our research has found many factors for which caesarean delivery is constantly increasing. So, everyone needs to pay more attention to these factors and try to avoid unnecessary caesarean delivery.

Acknowledgement
Type or paste your Acknowledgement here as prescribed by the journal’s instructions for authors. This section is not added if the author does not have anyone to acknowledge.

References
Medicine, 8(1).


Bahl, R., Qazi, S., Darmstadt, G. L., Martins, J. (2010). Why is continuum of care from home to health...


