Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 164 (2022) 1-16 EISSN 2392-2192

An Illustration of ML Models to Determine the Prevalence and Predicting Factors of the First-Day Neonatal Mortality in Bangladesh

Rumana Rois

Department of Statistics, Jahangirnagar University, Savar, Dhaka - 1342, Bangladesh

E-mail address: rois@juniv.edu

ABSTRACT

Neonatal mortality remains unacceptably high in developing countries and the risk is greatest on the first day of life. A better perception of the causes responsible behinds the first-day neonatal mortality is a key to lessening this problem. This study assessed to predict and detect predicting factors of the FNM through different machine learning (ML) algorithms. The study data was based on FNM of 26145 children from the 2017-18 Bangladesh Demographic and Health Survey (BDHS). The Support Vector Machine (SVM) algorithm and chi-square test were used to extract predicting factors of the FNM. Prediction of IM was done using different ML models, for instance, decision tree (DT), random forest (RF), SVM, and logistic regression (LR). The performance of these techniques was evaluated via different parameters of confusion matrix, receiver operating characteristics (ROC) curve, and k-fold cross-validation. The study revealed that the prevalence of FNM was 3% (792 newborns out of 26145 children). Mother's age at first birth, birth interval, region, religion, wealth index, child's gender, birth order, and total children ever born were observed as significant predicting factors of the FNM in Bangladesh using the chi-square test. However, total children ever born, birth order number, father's education, type of cooking fuel, exposure of media, wealth index, gender of the child, mother's education, mother's body mass index (BMI), and religion were the significant predicting factors of the FNM using the SVM method. To predict FNM in Bangladesh, though the LR model was performed better among all four ML algorithms based on the highest accuracy scores and the minimum standard error for the selected predicting factors using the SVM and chi-square test, the LR model failed to correctly predict the positive cases of FNM for sensitivity and precision. Needless to say, to predict the first-day neonatal mortality in Bangladesh for BDHS 2017-18 dataset, the SVM model was

(Received 04 November 2021; Accepted 25 November 2021; Date of Publication 27 November 2021)

recommended (Accuracy = 0.9449, Sensitivity = 0. 0325, Specificity = 0.9745, Precision = 0.0396, area under the ROC curve (AUC) = 0.5227, k-fold accuracy = 0.9487) when the predicting factors will be identified using the SVM method, and the RF model (Accuracy = 0.9675, Sensitivity = 0. 0081, Specificity = 0.9986, Precision = 0.1538, area under the ROC curve (AUC) = 0.6461, k-fold accuracy = 0.9686) was recommended when the associated factors will be identified using the chi-square test. ML framework can be identified the significant predicting factors of the FNM, therefore may help the health-policymakers, stakeholders, and families to understand and prevent this severe public health problem.

***Keywords*:** Newborn's health, decision tree, random forest, support vector machine, features selection, logistic regression, confusion matrix, ROC, k-fold cross-validation

1. INTRODUCTION

According to Save the Children's 14th annual State of the World's Mothers report [1], more than 1 million per year babies die on their day of birth worldwide. Nearly all newborn and maternal deaths, 98 and 99 percent, respectively, occur in developing countries where pregnant women and newborn babies lack access to basic health care services – before, during, and after delivery [1]. Certainly, neonatal and child mortality is a vital indicator of the child health and overall national development of a country [2]. According to World Health Organization (WHO) [3], substantial global progress has been made in reducing child deaths, from 12.7 million in 1990 to 5.9 million in 2015. Between 1990 and 2015, 62 of the 195 countries with available estimates met the Millenium Development Goal 4 (MDG 4) target of a two-thirds reduction in the under-5 mortality rate. Among them, 24 are low- and lower-middle-income countries. Currently, 79 countries have an under-5 mortality rate higher than 25 deaths per 1000 live births [3, 4]. Neonatal deaths alone are responsible for more than two-thirds of all deaths of infant mortality and about fifty percent of all deaths in under-five children [5]. Moreover, a little under half of all neonatal deaths occurred within 24 hours of birth, and around one-third occurred within 6 hours [6].

Though Bangladesh has achieved the Millennium Development Goal (MDG) 4 (to reduce child mortality) and is on track to MDG 5a (to reduce maternal mortality) [4] by experiencing a significant reduction of child mortality over the past decades, neonatal mortality is still reasonably high. Neonatal deaths comprise 61 percent of all under-5 deaths in Bangladesh [7], and it needs a considerable effort to achieve the Sustainable Development Goal (SDG) target 3. The FNM, therefore, plays a crucial role in reducing neonatal as well as child mortality in Bangladesh. According to Bangladesh Maternal Mortality Survey 2010, the main reason behind neonatal deaths, which are determined by the verbal autopsy, are low birth weight and premature delivery (11%), birth asphyxia (21%), sepsis (34%), and acute respiratory infections (10%) [8], with the assistance of some high impact cost-effective, evidence-based interventions and expanded healthcare systems, many of these deaths are preventable [1, 9]. Well-trained and equipped health care worker during the time of delivery is an effective solution [1].

There is much work in the literature concerning the prediction of neonatal mortality using machine learning (ML) models [10-19]. ML models can accurately predict neonatal death, and Artificial Intelligence (AI) is the most frequently used predictor and metrics for neonatal mortality [20]. There is also an adequate number of researches regarding infant mortality, early child mortality, and low birth weight in Bangladesh [12, 21-24]. Therefore, we are motivated to identify the risk factors associated with the FNM and to predict the FNM in Bangladesh based on the latest country-wise representative Bangladesh Demographic and Health Survey (BDHS) 2017-18 [25] using different ML models. Measuring and understanding the true risk factors of the FNM can help the families and health policymakers to take treatments, interventions, and initiatives against the FNM in Bangladesh.

2. MATERIALS AND METHODS

2. 1. Data sources and study design

In detecting the significant factors and in predicting the FNM, this study used survey data from the Bangladesh Demographic and Health Survey (BDHS) 2017-18 [25], which comprises Bangladesh's districts and administrative divisions. This survey collects information from 20,250 households and 20,100 reproductive mothers aged 15–49 years. A total of 26,145 infants were included in this study after removing all missing cases. The detailed information about the survey data is available athttps://dhsprogram.com/data/available-datasets.cfm. To predict the FNM in Bangladesh, we considered the first-day death of newborns as a binary outcome variable. Different maternal, socio-economic, demographic, and environmental factors were considered as exposure variables such as maternal age at first marriage, maternal age at first birth, mother's body mass index, the birth interval between two subsequent pregnancies, antenatal care service during pregnancy, receiving a tetanus toxoid (TT) injection during pregnancy, administrative regions, place of residence, religion, educational attainment of both mother's and father's, occupational status of mother's, women empowerment, exposure of media, total children ever born, child sex, birth order number, sources of drinking water, type of toilet facilities and type of cooking fuel.

2. 2. Statistical Analyses

The study aimed to detect the potential factors and to predict neonatal mortality on the first-day using different machine learning models (DT, RF, SVM, LR). Our methodology involves data collection and processing, feature selection using the chi-square test and SVM. The evaluation process involved splitting the entire data set into training data sets and test data sets, applying ML models in the training data set, and evaluating the performance of these models on the test data set. Therefore, predicting child mortality on the first-day based on the entire data set using the best-performed model. The performances were evaluated using three performance parameters from the confusion matrix such as sensitivity, specificity, and accuracy, the area under the receiver operating characteristics (ROC) curve (AUC), and the K-fold cross-validation. All ML models were performed using the scikit-learn module in Python programming language version 3.7.3.

2. 3. Decision Tree (DT)

A decision tree is the most commonly used ML technique that develops prediction algorithms for a target variable [26, 27]. This method classifies a population into branch-like segments that construct an inverted tree with roots, internal, and leaf nodes [27] Without imposing a complicated parametric structure, the algorithm can deal with a vast quantity of data [27].

2. 4. Random Forest (RF)

Random forest is a tree-structured classifier with each tree depending on a collection of the random variable [28]. The goal is to find a predictor function that minimizes the expected loss value by determining the loss function [29]. We used 100 decision trees and Gini for impurity index to implement the random forests algorithm in Python.

2. 5. Support Vector Machine (SVM)

Support vector machines are a set of related supervised learning methods [30]. The technique uses machine learning theory to maximize predictive accuracy; doing such, it over-fit the data automatically [31]. Structural Risk Minimization (SRM) principle is used in this superior formulation, to the traditional Empirical Risk Minimization (ERM) principle, used by conventional neural networks [32].

2. 6. Logistic Regression (LR)

LR analysis is the study of the association between a categorical dependent variable and a set of independent (explanatory) variables [33]. The response variable's outcome is divided into "failure," which is represented by 1, and "success," which is represented by 0 [34]. Unlike discriminant analysis, logistic regression does not assume that the independent variables are normally distributed [33].

2. 7. Confusion Matrix Performance Parameters

A confusion matrix is a representation of actual and predicted classifications done by a classification system [35]. It compares the predicted classification against the actual classification in the form of false-positive (FP), true positive (TP), false negative (FN), and true negative (TN) information while evaluating the performance [33, 36].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP}, \tag{2.1}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \tag{2.2}$$

$$\text{Specificity} = \frac{TN}{TN + FP}, \tag{2.3}$$

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{2.4}$$

Accuracy indicates the total number of correct predictions, sensitivity indicates how well a classification algorithm classifies data points in the positive class, specificity indicates how well a classification algorithm classifies data points in the negative class, and finally, precision indicates the number of data points correctly classified from the positive class [33, 36].

2. 8. Receiver Operating Characteristic Curve

A receiver operating characteristics (ROC) curve is a two-dimensional plot that visualizes, organizes, and selects classifiers based on their performance [37]. In this curve, the True Positive rate (TP = Sensitivity) is plotted as a function of the False Positive rate (FP = 1 - Specificity) [38]. The area under the ROC curve (AUC) measures how well a parameter can distinguish between two diagnostic groups [37].

2. 9. K-fold Cross-Validation

In *k*-fold cross validation subsampling, the data set is randomly split *k* times. A model is built based on the *k*-1 parts of the data set called training data set. The accuracy of this estimated model is then evaluated on a test set. We choose the model which has the smallest cross-validation score [38, 39]. Any pair of training and test set is disjoint as the sets don't have any common case [39].

3. RESULT AND ANALYSIS

3. 1. Features Selection

A total of 26145 infants of different backgrounds were included in this study from the 2017-18 BDHS dataset. Among 26145 infants, 82.6 % had mothers married before their legal recommended age (at least 18 years in Bangladesh) of the first marriage, 88.0% had mothers the first birth at 20 years or below, 32.8% had overweight or obese mothers, 76.1% of live births were above 2 years interval between two subsequent pregnancies, during mother's pregnancy period 0.4% were received ANC services and 0.3% were received TT-injection, 70.0% of children were selected from rural, 92.1% were Muslim, 27.0% had mothers with no education, 32.5% had illiterate fathers, 58.1% had employed mothers, 45.9% were from low-income families, 82.4% had not empowered mothers, 44.9% had mothers with mass-media exposure, 46.7% of the children were the first child, 80.6% had mothers with 3 or more number of children, 42.4% of families were used unhygienic toilets, only 13.5% of families were from households with less polluted cooking fuel and only 7.1% had no facilities of safe drinking water.

Table 1 also exhibits the association between socio-demographic characteristics and the FNM of children in Bangladesh. Based on the p-value of the chi-square (χ^2) test, significant associations were observed with the FNM in mother's age at first birth (χ^2 = 3.019, p-value < 0.05), birth interval (χ^2 = 354.096, p-value < 0.05), region (χ^2 = 49.088, p-value < 0.05), religion (χ^2 = 8.765, p-value < 0.05), wealth index (χ^2 = 6.397, p-value < 0.05), child's gender (χ^2 = 34.440, p-value < 0.05), birth order (χ^2 = 21.350, p-value < 0.05), total children ever born (χ^2 = 51.472, p-value < 0.05) (Table 1).

Table 1. Socio-demographic Characteristics of the First-day Neonatal Mortality in Bangladesh based on BDHS 2017-18.

Variables	Total n=26145 (%)	First-day Neonatal death n=792 (3%) Yes (%)	χ^2 value	*p*-value
Age at first marriage			0.594	0.235
<18	21586 (82.6)	662 (3.1)		
18 and above	4559 (17.4)	130 (2.9)		
Age at first birth			3.019	0.046*
≤20	22998 (88.0)	681 (3.0)		
>20	3147 (12.0)	111 (3.5)		
Mother's BMI			1.702	0.427
Normal	14471 (55.3)	447 (3.1)		
Underweight	3101 (11.9)	101 (3.3)		
Overweight & Obesity	8573 (32.8)	244 (2.8)		
Birth interval			354.096	<0.001*
≤2 years	6256 (23.9)	412 (6.6)		
> 2 years	19889 (76.1)	380 (1.9)		
Antenatal care during pregnancy			0.008	0.545
No	26051 (99.6)	789 (3.0)		
Yes	94 (0.4)	3 (3.2)		
TT-Injection during pregnancy			4.188	0.058
No	26076 (99.7)	787 (3.0)		
Yes	69 (0.3)	5 (7.2)		
Region			49.088	<0.001*
Dhaka	3276 (12.5)	100 (3.1)		
Barisal	2936 (11.2)	69 (2.4)		
Chittagong	4244 (16.2)	95 (2.2)		
Khulna	2826 (10.8)	97 (3.4)		
Mymensingh	3122 (11.9)	105 (3.4)		

Rajshahi	2859 (10.9)	133 (4.7)		
Rangpur	3273 (12.5)	109 (3.3)		
Sylhet	3609 (13.8)	84 (2.3)		
Place of residence			0.097	0.393
Urban	7855 (30.0)	234 (3.0)		
Rural	18290 (70.0)	558 (3.1)		
Religion			8.765	0.003*
Non-Muslim	2074 (7.9)	58 (4.1)		
Muslim	24071 (92.1)	707 (2.9)		
Maternal education			0.384	0.944
No education	7055 (27.0)	218 (3.1)		
Primary	10534 (40.3)	319 (3.0)		
Secondary	7292 (27.9)	220 (3.0)		
Higher	1264 (4.8)	35 (2.8)		
Paternal education			3.676	0.299
No education	8496 (32.5)	256 (3.0)		
Primary	9162 (35.0)	300 (3.3)		
Secondary	6031 (23.1)	167 (2.8)		
Higher	2456 (9.4)	69 (2.8)		
Mother’s occupation			0.539	0.243
Not working	10961 (41.9)	322 (2.9)		
Working	15184 (58.1)	470 (3.1)		
Wealth quintile			6.397	0.041*
Poor	11994 (45.9)	361 (3.0)		
Middle	5226 (20.0)	184 (3.5)		
Rich	8925 (34.1)	247 (2.8)		
Women empowerment			0.015	0.470
No	21533 (82.4)	651 (3.0)		
Yes	4612 (17.6)	141 (3.1)		
Exposure of media			0.102	0.389
Non-exposure	14406 (55.1)	432 (3.0)		

Exposure	11739 (44.9)	360 (3.1)		
Child sex			34.440	<0.001*
Female	13083 (50.0)	315 (2.4)		
Male	13062 (50.0)	477 (3.7)		
Birth order			21.350	<0.001*
One	12215 (46.7)	428 (3.5)		
Two	7094 (27.1)	205 (2.9)		
Three and more	6836 (26.1)	159 (2.3)		
Total children ever born			51.472	<0.001*
1 or 2	5071 (19.4)	75 (1.5)		
3 and more	21074 (80.6)	717 (3.4)		
Toilet facility			0.115	0.742
Hygienic	15065 (57.6)	461 (3.1)		
Unhygienic	11080 (42.4)	331 (3.0)		
Cooking fuel			2.129	0.156
Less polluted	3526 (13.5)	93 (2.6)		
Polluted	22619 (86.5)	699 (3.1)		
Drinking water			0.730	0.437
Safe water	24294 (92.9)	742 (3.1)		
Unsafe water	1851 (7.1)	50 (2.7)		

*****Statistically significant at the 0.05 level.

Figure 1 illustrates the result of the SVM algorithm with linear kernel, where the important risk predictors of the first-day mortality were explored. Once having the fitted SVM with linear kernel, then the important features can be determined by comparing the size of the classifier coefficients using the .coef_ argument value in the SVC function of the scikit-learn module in Python.

Figure 1 reveals those selected risk predictors with the blue bars and insignificant ones (which hold less variance) with the green bars. Hereafter, the main identified features (risk predictors) were total children ever born, birth order number, father's education, type of cooking fuel, exposure of media, wealth index, gender of the child, mother's education, mother's BMI, and religion. Whereas, mother's age at first birth, birth interval, region, religion, wealth index, child's gender, birth order, and total children ever born were observed significant using the conventional chi-square test. These eight variables were used to evaluate different ML models to predict the first-day neonatal mortality in Bangladesh.

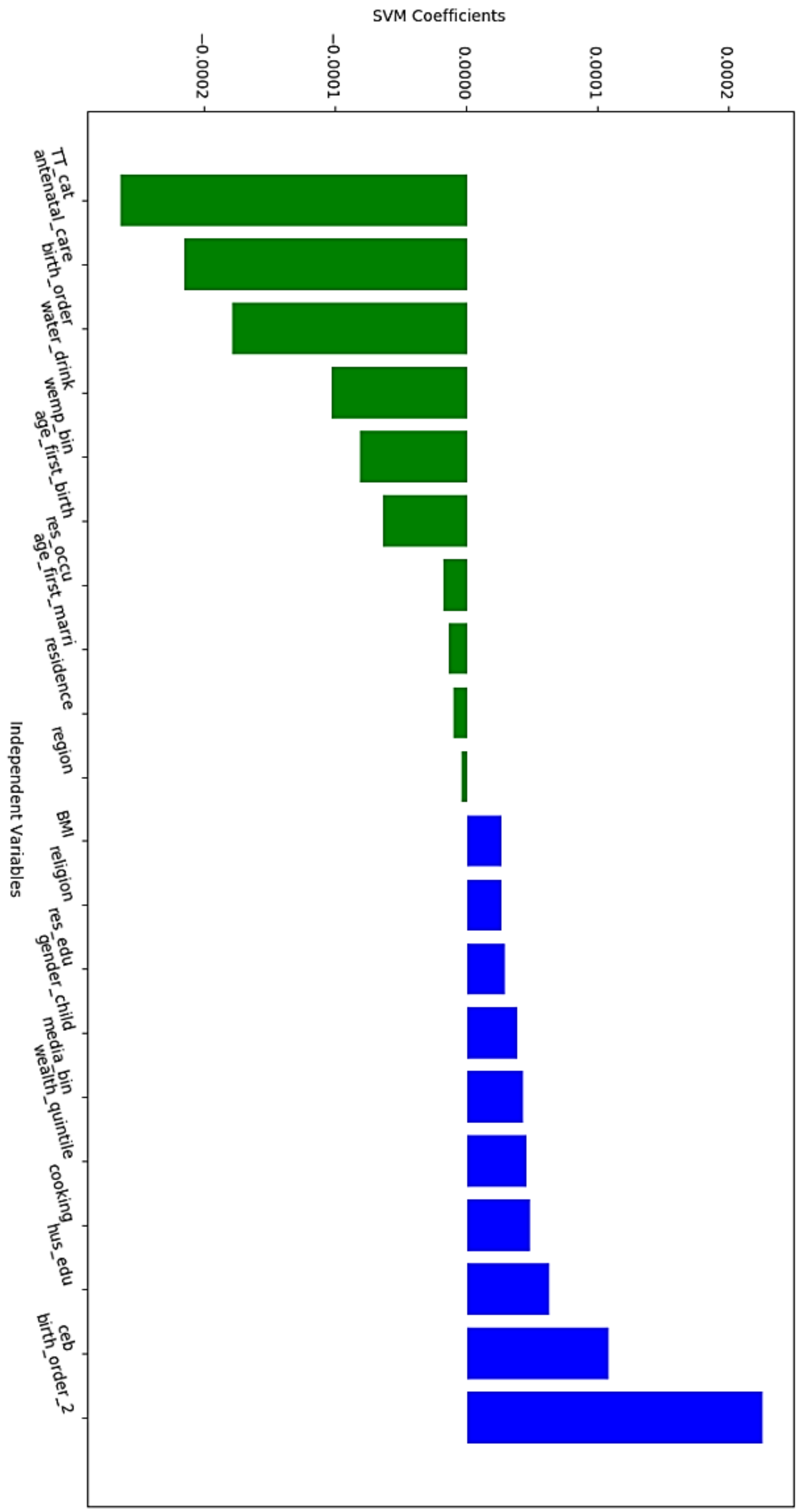


Figure 1. Features selection using SVM.

3. 2. Evaluation of Machine Learning Models

We evaluated the performance of different ML models based on different performance parameters of confusion-matrix (Table 2), ROC curve for different ML models (Figure 3), and k-fold cross-validation (Table 3) using the selected risk factors based on the SVM and the chi-square test. Table 3 reveals accuracy scores, sensitivity, specificity, and precisions of all mentioned machine learning algorithms by considering 70% observations as the training data and 30% observation as the test data with the random seeds 4589 using the scikit-learn module in Python. We found that LR was performed better among four ML algorithms based on accuracy scores for the selected risk factors using both of the methods (SVM and chi-square test). However, based on the sensitivity, specificity, and precisions, the SVM with sigmoid kernel performed better for the selected risk factors using SVM, for instance, 94.5% of accurate predictions (i.e., accuracy = 0.9449), 3.3% of positive cases that were predicted as positive (i.e., sensitivity = 0.0325), 97.5% of negative cases that were predicted as negative (i.e., specificity = 0.9745), 4% of positive predictions that were correct (i.e., precision = 0.0396). Conversely, when potential factors were identified using the chi-square test, RF model performed better, for instance, 96.8% of accurate predictions (i.e., accuracy = 0.9675), 0.8% of positive cases that were predicted as positive (i.e., sensitivity = 0.0081), 99.9% of negative cases that were predicted as negative (i.e., specificity = 0.9986), 15.4% of positive predictions that were correct (i.e., precision = 0.1538) (Table 2).

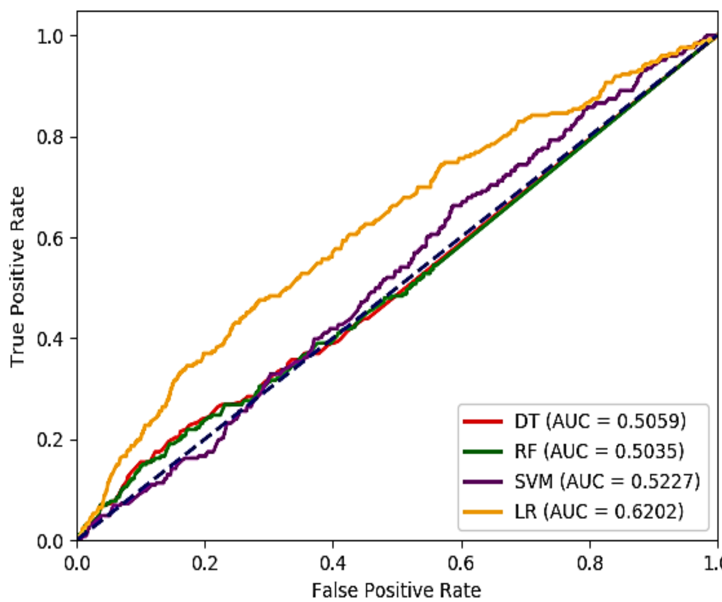
Table 2. Accuracy, sensitivity, specificity and precision of different SVM models.

Features Selection Method	Models	Accuracy	Sensitivity	Specificity	Precision
SVM	DT	0.9663	0.0	0.9976	0.0
	RF	0.9673	0.0	0.9986	0.0
	SVM	0.9449	**0.0325**	0.9745	**0.0396**
	LR	**0.9686**	0.0	**1.0**	N/A
χ^2 Test	DT	0.9671	0.0081	0.9982	0.1250
	RF	0.9675	**0.0081**	0.9986	**0.1538**
	SVM	0.9647	0.0040	0.9957	0.0303
	LR	**0.9686**	0.0	**1.0**	N/A

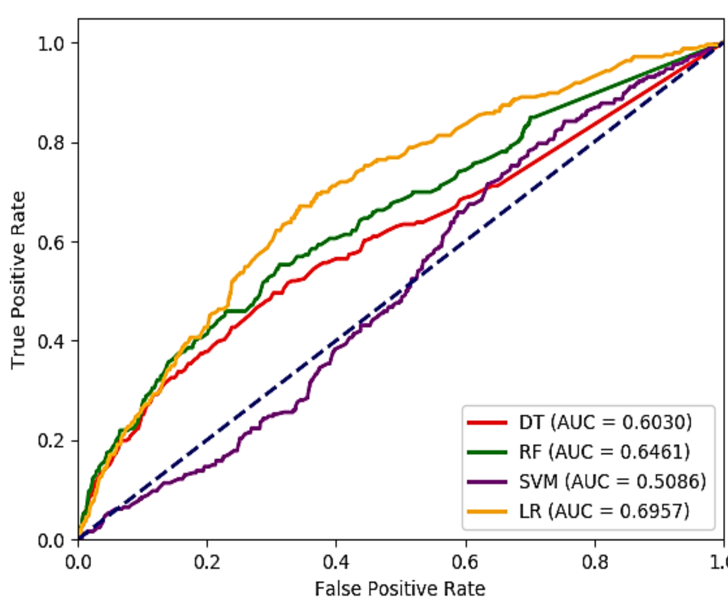
- Highest values are indicated in bold, N/A means not applicable.

Figure 2 illustrates the estimated AUC of SVM models with different kernels, which were run using the scikit-learn module in Python 3.7.3 by considering 70% observations as training data and 30% observation as test data with the random seed 4589. To predict the FNM in Bangladesh based on BDHS 2017-18, the estimated AUC was 0.5059, 0.5035, 0.5227, and

0.6202 using DT, RF, SVM, and LR, respectively, for the selected features using the SVM. However, AUC was 0.6030, 0.6461, 0.5086, and 0.6957 using DT, RF, SVM, and LR, respectively, for the selected features using the chi-square test. Therefore, the LR model performed better among all examined ML models as LR gives the highest estimated AUC.



(a)



(b)

Figure 2. ROC curves to predict first-day mortality of newborns using DT, RF, SVM and LR using selected risk factors based on the SVM in **(a)** and the chi-square Test in **(b)**.

Table 3. Result of K-Fold cross-validation of ML Models.

Features Selection Method	Models	5-Fold		10-Fold		30-Fold	
		MAcc	SE	MAcc	SE	MAcc	SE
SVM	DT	0.9684	0.0022	0.9684	0.0026	0.9688	**0.0045**
	RF	0.9686	0.0019	0.9686	0.0026	0.9687	0.0046
	SVM	0.9483	0.0022	0.9499	0.0050	0.9487	0.0063
	LR	**0.9697**	**0.0016**	**0.9697**	**0.0022**	**0.9697**	0.0046
χ^2 Test	DT	0.9675	0.0018	0.9679	**0.0015**	0.9680	0.0047
	RF	0.9684	0.0019	0.9684	0.0020	0.9686	0.0046
	SVM	0.9501	0.0031	0.9501	0.0040	0.9500	0.0087
	LR	**0.9697**	**0.0016**	**0.9697**	0.0022	**0.9697**	**0.0046**

MAcc=Mean of Accuracy scores from each fold, **SE**=Standard Error of Accuracy scores, the best values are indicated in **bold**.

Table 3 presents the K-fold cross-validation of different ML models, which was performed for 5-Fold, 10-Fold, and 30-Fold repeatedly. The results revealed that the LR model was performed better among four ML algorithms based on the highest accuracy scores and the minimum standard error for the selected risk factors using both of the methods (SVM and chi-square test). However, the LR model failed to correctly predict the positive cases (i.e., the first-day neonatal mortality) as the score of sensitivity was 0 and precision was not applicable to calculate. Consequently, to predict the first-day neonatal mortality in Bangladesh for BDHS 2017-18, the SVM algorithm was recommended when the associated factors will be identified using the SVM method, and the RF model was recommended when the associated factors will be identified using the chi-square test to predict the first-day neonatal mortality in Bangladesh.

4. DISCUSSION AND CONCLUSIONS

Neonatal deaths comprise 61 percent of all under-5 deaths in Bangladesh [7], therefore, the FNM plays a crucial role in reducing neonatal as well as child mortality in Bangladesh. A better perception of the causes responsible behinds the first-day neonatal death is a key to lessening this problem. Motivated by this significant health problem and findings of [24], our study utilized different ML algorithms to find the factors responsible for the first-day mortality and to estimate the FNM in Bangladesh.

The study findings reveal that the proportion of FNM was 3% (792 newborns out of 26145 children). Mother's age at first birth, birth interval, region, religion, wealth index, child's gender, birth order, and total children ever born were observed as significant predicting factors of the FNM in Bangladesh using the chi-square test. However, total children ever born, birth order number, father's education, type of cooking fuel, exposure of media, wealth index, gender of the child, mother's education, mother's body mass index, and religion were the significant predicting factors of the FNM using the SVM method. Though the LR model was performed better among all four ML algorithms based on the highest accuracy scores and the minimum standard error for the selected risk factors using both of the feature selection methods, i.e., SVM and chi-square test, the LR model failed to correctly predict the positive cases (i.e., the first-day neonatal mortality) as the score of sensitivity was 0 and precision was found N/A (not applicable). Needless to say, to predict the first-day neonatal mortality in Bangladesh for BDHS 2017-18 dataset, the SVM algorithm is recommended when the predicting factors will be identified using the SVM method, and the RF model is recommended when the associated factors will be identified using the chi-square test.

Furthermore, no assumptions are required for the DT and RF models, and the implementation of these ML methods is quite easier through any standard software. Whereas, the popular LR model requires to fulfill all the underlying assumptions before estimating the model, among them predictors have to be independent of each other and having a significant association with the outcome variable [24]. Therefore, the proper implementation of the LR model is difficult. In this study, we failed to estimate the precision and sensitivity for the LR model, as improper estimation of the LR model may result in some misleading information. Assumptions confined LR model allows a few predictors to estimate the FNM in Bangladesh, therefore, the fitted model will be correct but less informative. Conversely, SVM and RF models will be more informative and flexible to predict the FNM in Bangladesh.

References

[1] Save the Children. (2013). State of the world's mothers: surviving the first day. London, UK. Available from: http://www.savethechildren.org/atf/cf/%7B9def2ebe-10ae-432c-9bd0-df91d2eba74a%7D/SOWM-FULL-REPORT_2013.PDF

[2] McGuire JW. Basic health care provision and under-5 mortality: a cross-national study of developing countries. *World Dev*. 2006; 34(3): 405–25

[3] WHO. Children: reducing mortality. 2016. http://www.who.int/mediacentre/factsheets/fs178/en/. Accessed 26 Mar 2017.

[4] Khan, J.R., Awan, N. A comprehensive analysis on child mortality and its determinants in Bangladesh using frailty models. *Arch Public Health*, 2017; 75(58): 1-10

[5] Lawn JE, Cousens S, Zupan J, Lancet Neonatal Survival Steering Team. 4 million neonatal deaths: when? Where? Why? *The Lancet* 2005; 365(9462): 891-900

[6] Baqui, A. H., Mitra, D. K., Begum, N., Hurt, L., Soremekun, S., Edmond, K., ... & Nisar, M. I. (2016). Neonatal mortality within 24 hours of birth in six low-and lower-middle-income countries. *Bulletin of the World Health Organization*, 94(10), p. 752.

[7] National Institute of Population Research and Training (NIPORT), Mitra and Associates, & ICF International. (2016). Bangladesh demographic and health survey 2014. p. 99.

[8] Bryce J, Boschi-Pinto C, Shibuya K, Black RE, WHO Child Health Epidemiology Reference Group. WHO estimates of the causes of death in children. *The Lancet* 2005; 365(9465): 1147-1152

[9] Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, De Bernis L, Lancet Neonatal Survival Steering Team. Evidence-based, cost-effective interventions: how many newborn babies can we save? *The Lancet* 2005; 365(9463): 977-988

[10] Batista AF, Diniz CS, Bonilha EA, Kawachi I, Chiavegatto Filho AD. Neonatal mortality prediction with routinely collected data: a machine learning approach. *BMC Pediatrics* 2021; 21(1): 1-6

[11] Ogallo W, Speakman S, Akinwande V, Varshney KR, Walcott-Bryant A, Wayua C, Weldemariam K, Mershon CH, Orobaton N. Identifying Factors Associated with Neonatal Mortality in Sub-Saharan Africa using Machine Learning. *In American Medical Informatics Association (AMIA) Annual Symposium Proceedings*, 2020; 2020 (963).

[12] Sheikhtaheri A, Zarkesh MR, Moradi R, Kermani F. Prediction of neonatal deaths in NICUs: development and validation of machine learning models. *BMC Medical Informatics and Decision Making* 2021; 21(1): 1-4

[13] Jaskari J, Myllärinen J, Leskinen M, Rad AB, Hollmén J, Andersson S, Särkkä S. Machine learning methods for neonatal mortality and morbidity classification. *IEEE Access*. 2020; 8: 123347-58

[14] Greenbury SF, Ougham K, Wu J, Battersby C, Gale C, Modi N, Angelini ED. Identification of variation in nutritional practice in neonatal units in England and association with clinical outcomes using agnostic machine learning. *Scientific Reports*. 2021; 11(1): 1-5

[15] Raja R, Mukherjee I, Sarkar BK. A Machine Learning-Based Prediction Model for Preterm Birth in Rural India. *Journal of Healthcare Engineering*. 2021; 2021: 6665573

[16] Masino AJ, Harris MC, Forsyth D, Ostapenko S, Srinivasan L, Bonafide CP, Balamuth F, Schmatz M, Grundmeier RW. Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. *PLOS ONE* 2019; 14(2): e0212665

[17] Açıkoğlu M, Tuncer SA. Incorporating feature selection methods into a machine learning-based neonatal seizure diagnosis. *Medical Hypotheses* 2020; 135: 109464

[18] Beluzo CE, Alves LC, Silva E, Bresan RC, Arruda NM, de Carvalho TJ. Machine Learning to Predict Neonatal Mortality Using Public Health Data from São Paulo-Brazil. *medRxiv*. 2020 Jan 1.

[19] Lawoyin TO, Onadeko MO, Asekun-Olarinmoye EO. Neonatal mortality and perinatal risk factors in rural southwestern Nigeria: a community-based prospective study. *West African Journal of Medicine*. 2010; 29(1).

[20] Mangold C, Zoretic S, Thallapureddy K, Moreira A, Chorath K, Moreira A. Machine Learning Models for Predicting Neonatal Mortality: A Systematic Review. *Neonatology*. 2021; 118(4): 394-405

[21] Rahman A, Hossain Z, Kabir E, Rois R. Machine Learning Algorithm for Analysing Infant Mortality in Bangladesh. In: Siuly S., Wang H., Chen L., Guo Y., Xing C. (eds) Health Information Science. HIS 2021. *Lecture Notes in Computer Science*, 2021; 13079: 205-219

[22] Galib AH, Nahar N, Hossain BM. The Influences of Pre-birth Factors in Early assessment of Child Mortality using Machine Learning Techniques. *arXiv preprint*: 2011.09536

[23] Borson, NS, M. R. Kabir, Z. Zamal and R. M. Rahman. Correlation analysis of demographic factors on low birth weight and prediction modeling using machine learning techniques. *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 2020, pp. 169-173, doi: 10.1109/WorldS450073.2020.9210338

[24] Rahman SJ, Ahmed NF, Abedin MM, Ahammed B, Ali M, Rahman MJ, Maniruzzaman M. Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. *Plos one*. 2021; 16(6): e0253172

[25] National institute of population research and training (NIPROT), Bangladesh demographic and health survey 2017–2018. Mitra and Associates, Dhaka, Bangladesh and ICF International, Calverton, Maryland, USA (2019).

[26] Shalev S and David SB. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, USA, 2014.

[27] Song YY, Ying LU. Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry* 2015; 27(2): 130

[28] Breiman L. Random Forests. *Machine Learning* 2001; 45(1): 5-32

[29] Cutler A, Cutler DR, Stevens JR. Random forests. In Ensemble machine learning 2012; 157-175

[30] Mantri JK. Comparison between SVM and MLP in predicting stock index trends. *International Journal of Science and Modern Engineering* 2013; 1(9): 81-82

[31] Jakkula V. Tutorial on support vector machine (svm). School of EECS, Washington State University. 2006; 37.

[32] Burges CJ. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 1998; 2(2): 121-67

[33] Igual L, Seguí S. *Introduction to Data Science*. 2017, Springer, Cham.

[34] Sperandei S. Understanding logistic regression analysis. *Biochemia Medica* 2014; 24(1): 12-8

[35] Provost F, Kohavi R. Glossary of terms. *Journal of Machine Learning* 1998; 30(2-3): 271-4

[36] Awad M, Khanna R. Efficient Learning Machines, 2015. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4302-5990-9_1

[37] Yang S, Berdine G. The receiver operating characteristic (ROC) curve. *The Southwest Respiratory and Critical Care Chronicles*. 2017; 5(19): 34-6

[38] Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006; 27(8): 861-74

[39] Bhatt S, Leiman PG, Taylor NMI. Tail Structure and Dynamics. *In Reference Module in Life Sciences*. Elsevier. 2019. https://doi.org/10.1016/B978-0-12-809633-8.20965-5