Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

Article

Prediction on Domestic Violence in Bangladesh during the COVID-19 Outbreak using Machine Learning Methods

Md. Murad Hossain [1,2,§], Md. Asadullah [2,§], Abidur Rahaman [3,§], Md. Sipon Miah [4,§], Md. Zahid Hasan [2,§], Tonmay Paul [2,§] and Mohammad Amzad Hossain [3,5,*,§]

1 Department of Computer Science, University of Turin, Italy; md.hossain50@edu.unito.it (M.M.H)
2 Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh; asadullahstat@gmail.com (M.A.), zahid1680@gmail.com (M.Z.H), tonmaypaul006@gmail.com (T.P.)
3 Department of Information and Communication Engineering, Noakhali Science and Technology University, Noakhali, Bangladesh; abidur@nstu.edu.bd (A.R.)
4 Department of Information and Communication Technology (ICT), Islamic University, Kushtia-7003, Bangladesh; m.miah1@nuigalway.ie (M.S.M)
5 School of Computer Science, National University of Ireland Galway, Galway, Republic of Ireland; m.hossain3@nuigalway.ie (M.A.H)
* Correspondence: m.hossain3@nuigalway.ie (M.A.H); Tel.: +353-892-447-669 (M.A.H)
§ These authors contributed equally to this work.

Abstract: In Southern Asia, Bangladesh is a well-known developing country. Because of COVID-19, we continuously face challenges. Not only can these issues occur beyond economic or health concerns, but they also generate dangerous social problems, such as family abuse. Since the inception of this epidemic, multiple social crimes are looming. Remaining home during the lockout period enhances divorce rates. This research presents a customized forecast of family violence during the COVID-19 outbreak by using machine learning methods. In this paper, we have applied Random Forest, Logistic Regression, and Naive Bayes machine learning classifiers to predict family violence and discovered the feature importance. The performance of the classifiers is evaluated based on accuracy, precision, recall, and F-score. We have employed an oversampling strategy named synthetic minority oversampling technique (SMOTE) to solve the imbalance problem of our data. Even, we have tried to compare three machine learning model performances before and after balancing of normalization data. Finally, ROC analyses and confusion matrices were developed and analyzed by using data augmentation. Our proposed system with the random forest classifier performed better with 77% accuracy in comparison with other two machine learning classifiers.

Keywords: Family violence; Machine Learning; Classification; ROC; Accuracy; COVID-19

1. Introduction

COVID-19 is now the most devastating global epidemic recorded in recent times, resulting in the deaths of up to 75-200 million people in Eurasia and North Africa, reaching a peak around 1347 to 1351 in Europe. But if we equate COVID-19 to the Black Death, it won't be quite different. During December 2019, COVID-19 spread from Wuhan, China, and accidents and mortality grew dramatically around the world. Not only economic, emotional, medical, but also social problems such as family violence are faced by Bangladesh. Due to family violence, records of too many suicide cases, violent events, female and child torture are reported. The graph of family violence is assumed to increase during pandemic circumstances. There is no simple concept in the academic literature available to provide valuable guidance for family violence survivor clinicians by exploring natural occurrences. With the extension of this report, the bridge between human welfare and animal welfare [1]. As during COVID-19 epidemic, various recent research literature's on tactics and recommendations was reviewed with the provision of integrated separation in reaction to violence towards women. In this review, a consolidated overview was carried out among December 2019 and June 2020 [2]. Economic deficiencies, work loss, disturbance or social alienation plays a major role in family caregivers to the understanding of child violence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

The consumption of abusive content and cyber bullying was responsible for more time spent online. This study gathered information by social media users on family violence [3]. Many limitations are placed to protect the welfare of children and teenagers, which causes the possibility of family abuse. The consequences of social isolation are economic uncertainty, stress levels, work losses and altered types of relationships. In this study, the significant side effects associated with the good growth of children and teens with their rights were listed [4].

The key aim of this research is to investigate the evidence and instances of gender abuse during the ongoing COVID-19 global pandemic. Both quantitative-qualitative approaches are clarified through bibliographic analysis, taking into account accurate and recent sources [5]. The study explained negative effect of COVID-19 on family in China thorough discussion and recommendations for the future researcher for the policy implications of assistance required during this pandemic about family abuse [6]. Considering the decreasing time of social contact, entertainment opportunities, job losses and financial deficiency, the study focused on New Zealand's psychological well-being. In an online panel study, the Kessler cognitive anxiety measure, the GAD-7 and the Well-being chart are all included. The purpose of this analysis was to examine the effect and effects of lockdown on a certain group of societies [7]. The research was performed following a recent article to demonstrate the association between remaining at home and family abuse. It plays an essential part in reducing the impact of the virus on politicians and other public adverse outcomes [8].

In order to determine the performance of clinicians in forecasting aggression in men, patterns of conflict related to age, race and sex are tracked. In order to predict aggression when the status of variation admission was followed up in the group over a 6-month period [9]. By implementing machine learning methods to actual panel survey results from Liberia, this study adopts an alternative approach to predicting community, interpersonal, and extra judicial killings violence two decades into the future. In specific, it also focused on three areas: Random Forest, Lasso, and Neural Networks. Besides comparative purposes, it contained findings from a portfolio context [10]. In particular, researchers considered whether it was possible to obtain usefully reliable predictions of domestic abuse. They introduced algorithms to evidence on more than 28,000 court date cases from a major metro area in which a perpetrator is facing sexual violence charges. As part of a broader initiative to improve pre-trial practices and findings in a wide major city, the data used in this study was compiled [11]. This research project contrasted predictions from a traditional risk management protocol-based framework with those founded on an equipment analysis. It proposed using criminal record forecasts to prioritize incoming service calls, and creating a more sensitive tool to differentiate true from false positives arising from this initial evaluation [12]. Initially, in this project, in order to assess GBV, a database of features relevant to the importance of GBV for over a decade is collected and constructed from government sources. Second, an approach is presented that includes evaluating various feature selection technologies such that four predictive rules are used and illustrated with each of the generated subcategories [13]. However, some of the articles listed above define family violence through various statistical analyses, such as core patterns, dispersion, logistic regression, etc. By analyzing recent posts, a couple of papers describe the factors behind family abuse. Even some study papers projected family violence using distinct verification and authentication methods, such as children, elderly people and couples facing their family. Some papers have clarified the forms in which COVID-19 increased family violence during lockdown period. But a rare number of papers used the techniques of machine learning. There is no such research available that has attempted to identify and forecast family violence through the use of machine learning software. In this research, using machine learning algorithms, we concentrate on predicting how family violence happens and what the key factor behind family violence. Throughout this article, from a description, validation and accuracy standpoint, we also seek to assess the utility of various advanced analysis. Finally, we aim to prove that our suggested model represents the used

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

data set correctly with an outstanding quality of precision, accuracy, sensitivity, specificity, F1 Score, and AUC. This paper therefore aims to unlock the primary and successful variables and explanations for the increased breach of the family during the deadlock. We hope it helps to remedy these heinous incidents, and people in our motherland come to stable life along with whole nations.

The rest of the chapter is organized as follows: The related work is summarized in Section 2. In Section 3, we describe the materials and methods which consists of the data set collection and processing as well as machine learning technique. In Section 4, we explain the proposed model architecture. In Section 5, we describe the performance evaluation criteria of our proposed model. The simulation results are given in Section 6. Finally, our conclusion is addressed in Section 7.

2. Related Work

Several researchers have been worked to analysis the domestic violence activity during the COVID-19 outbreak. In [14], the authors determined the frequency, rates and seriousness of injuries in intimate partner abuse relative to the previous 3 years at the time of the pandemic during the year 2020. This study was undertaken to determine the severity of domestic violence in northern Ethiopia towards pregnant women. For the causal relationship variables, binary and simultaneous logistic regressions were used to forecast. This paper has used central tendency and descriptive analysis based on a cross sectional community based study [15]. The research was conducted with a survey of 166 Victorian Practitioners to keep practitioners' voices and perspectives in the context of abuse experienced by women during the lockdown of COVID-19 in Victoria, Australia [16]. The main aim of this research is to determine the link between COVID-19 and domestic internal abuse. It also reveals the reasons for increased cases of violence due to COVID-19. The vacancies, loss of earnings, extended residential stays, and vulnerability to actions due to stay-at-home orders were deemed responsible for the increased incidence of family violence in this analysis [17]. The purpose of this research was to determine the effects of the COVID-19 lockdown on the mental health and gender-based violence of Tunisian women. An online survey was performed using the depression, anxiety and stress scale and the Facebook Bergen scale through sampling method of networking. Various statistical techniques such as frequencies, mean, standard deviation, Chi-square tests, odds ratio, ANOVA and correlation were used in this research [18]. The goal of this review was to examine the disparity in records of cases of domestic abuse from police statistics in Atlanta, Georgia. This research examined by compiling the residential felony counts mainstreamed to the metropolitan area. They were able to analyze the fluctuations and severity of do-residents and cross the rows with these studies. These studies helped them to evaluate the fluctuations and severity of domestic crimes [19]. A summary of substance abuse and behavioral condition issues has been identified to examine the hidden viewpoints of this paper on an emerging economic downturn. This study discusses particular ideas that foster veracity and accountability; security; respect for compatriots; solidarity and mutuality; autonomy and choice; environmental, historic and sexuality politics [20]. This research assessed the impact of the COVID-19 disease outbreak in relation to paternal excitement and child endangerment. Corona risk factors, risk factors for mental well being, potential perceived conflict and child abuse were examined using hierarchical multiple regression analysis and group difference tests to assess the associations between demographic characteristics [21]. This research analyzed public debate and sentiment regarding elderly people including pandemic on social networks and assessed the extent of age discrimination in civil debate. They used a mixture of qualitative thematic analysis of data science approaches and traditional statistics [22]. This study ignored how other natural disasters are not identical to a pandemic. During COVID-19, it discussed the shifts and times of intimate partner abuse. This paper also attempted to find the same violence scenario rather than a pandemic era [23]. Teenagers face multiple types of bullying, ignorance, and domestic violence, which is a U.S. public health problem of concern. It has

 doi:10.20944/preprints202104.0343.v1

a greater effect on low-income communities and race. In the recent period of the public health crisis, this paper described adverse childhood experiences and avoided health and social issues [24]. The main health issues most often identified with IPV are addressed at the start of this story. This article outlined the current problems faced by health care practitioners and offers future guidance on steps to be taken to avoid such cases during and after the COVID-19 pandemic [25]. The purpose of this research was to outline the Corona virus documentation and juvenile psychological problems associated with shutdown. This research identified psychiatric problems such as post-traumatic stress, psychological and severe anxiety, as well as signs relating to sadness that are so harmful for teens, when the worldwide disease outbreak crisis was extended [26]. The purposes of loneliness across various cultures were described in this article and the expression of loneliness was determined by Twitter during the pandemic period. This study identified key features of machine learning tools through Twitter feeds [27]. The key purpose of this analysis is to investigate pandemic-related debates, fears and feelings shared by Twitter users. Throughout the collected tweets, machine learning tools are used to identify common embedding and pos-tagging, popular topics and trends, and thoughts [28]. The objective of this essay is to include a vast overview of national conversation on family abuse and the contagion of corona virus on Social media. This essay post also used the Latent Dirichlet Allocation machine learning method and defined salient trends, issues and representative tweets [29]. The purpose of this research is to define the prejudice among women in Jordan and to assess the potential correlation of violence among women during the eruption of COVID-19 [30]. The main objective of this paper is to address the growing academic literature by monitoring framework connected with the hazard of IPV exposure. The research aimed for machine learning approaches that understand concealed and dynamic data trends and regularities [31]. Through current review thus aims to establish a predictive method that is clinically applicable. Cross-sectional evaluations were carried out in this article to classify demographic, clinical, and socio-cultural variables. There have also been several computational methods being used forecast physical aggression in previous events, including 28 predictors [32]. In order to test patients for IPV and injury, this study introduced machine learning models. It educated the advanced model on diagnostic files with IPV tags centered on entry to a violence reduction initiative and accident marks by emergency pediatrics congregation clinicians. Random forest, logistic regression, gradient enhanced forests, waistcoat frame neural network, and neural network clinical BERT42 formulation were being used [33].

3. Materials and Methods

3.1. Data Description

To collect data, we surveyed an Internet questionnaire on "Domestic Violence in Bangladesh During the COVID-19 Outbreak". For this survey, we first developed a series of family violence-related questionnaires using the Google Doc online platform. Thereafter, we forward this link to the respondent via email, messenger, and Facebook to collect the data. Within 10 days, we have received 511 replies. Our data set consists of some distinct variables, such as Age, Gender, Marital Status, Respondent Education, Profession, Family Type, Number of Family Members, Number of Earning Person, Head of Family, Religion, Residence Location, Wealth Status, Income before Corona, Income after Corona, and Lost Job during Corona. All variables with corresponding definitions are detailed in Table 1. Based on the values of these variables, we predicted family violence in the Bangladesh during the COVID-19 outbreak. For data processing and analysis, we use R package version 4.03.

Table 1: Series of questions in the Study dealing with family violence

Items	Corresponding Definitions
Age	The Participant's Age
Gender	The Participant's Gender
Marital Status	Marital Status of the Participant's
Respondent Education	Educational qualification of the Participant's
Profession	Occupation of the Respondent
Family Type	Family type of the Respondent
Number of Family Members	Number of Family member of the Respondent
Number of Earning Person	Number of earning person in the family of the Respondent
Head of Family	Head of family of the Respondent
Religion	Religion of the Respondent
Residence Location	Place of residence of the Respondent
Wealth Status	Wealth status of the Participant's
Income before Corona	The Participant's Monthly income for families before corona
Income after Corona	The Participant's Monthly income for families after corona
Lost Job during Corona	Respondent or any family member lose job during pandemic situation

The descriptive characteristics of the study respondents are given in the Table 2. The survey included a total of 511 respondents, reflecting a participation rate of 69.67% male and 30.33% female. The 438 respondents (85.71%) were Muslims with respect to their religious inclinations, while 73 (14.28%) were Hindus. Unmarried respondents made up 80.43% of the population in the survey and 19.57% were married. From Table 2, we found that most of the participants were between the ages of 15 and 25 (77.49%). With regard to the educational qualification feature, the highest respondent (63.99%) is an undergraduate student. In relation to profession feature, the highest respondent (81.02%) is a student. In our data set, 70.45% of the respondents are members of a joint family, 59.30% of the respondents live in rural areas, 82.20% of the respondents belong to middle-class wealth status, and 89.04% of household earners are 1-2 members. On the other hand, 20.35% of respondents or any of their family members have lost their job due to the COVID-19 outbreak.

3.2. Data Preprocessing: Data Normalization

In the first step, the data should be preprocessed to reduce the implementation time and improve the results. For this purpose, we normalized the data so that the attributes are normalized as follows. In this work, we have performed min-max feature scaling (normalization) for all the features. It is a scaling methodology where esteems are moved and re-scaled so they wind up going somewhere in the range of 0 and 1. The principle for applying normalization is given as follows:

$$X_{normalized} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \tag{1}$$

3.3. Feature Importance Plot

The feature importance describes which features are more helpful or important than other features in the data set. It can help with better understanding of the solved problem and sometimes lead to model improvements by employing the feature selection. Basically, feature importance is a technique that assign a score to input features based on

how useful they are at predicting a target variable. We computed the feature importance from our data set. This has been represented in Figure 1.

According to Figure 1, the feature Income after Corona, Income before Corona, and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

Table 2: Descriptive Characteristics of Study Respondents

Indicator's		N=511, n (%)
Age	15-25	396(77.49)
	26-35	100 (19.57)
	36-45	11 (2.15)
	46-55	3 (0.59)
	56-65	1 (0.19)
Gender	Male	356 (69.67)
	Female	155 (30.33)
Marital Status	Unmarried	411(80.43)
	Married	100 (19.57)
Respondent Education	HSC	34 (6.65)
	Undergraduate	327 (63.99)
	Graduate	46 (9.00)
	Post Graduate	93 (18.20)
	PhD	11 (2.15)
Profession	Student	414(81.02)
	Private Employee	51 (9.98)
	Government Employee	46 (9.00)
Family Type	Joint	360 (70.45)
	Single	151 (29.55)
Number of Family Members	1-4	82 (16.05)
	5-8	348 (68.10)
	8+	81 (15.85)
Number of Earning Person	1-2	455 (89.04)
	3-4	48 (9.40)
	4+	8 (1.56)
Head of Family	Father	405 (79.28)
	Husband	43 (8.41)
	Mother	34 (6.65)
	Others	29 (5.67)
Religion	Muslim	438 (85.71)
	Hinduism	73 (14.28)
Residence Location	Rural	303 (59.30)
	Urban	208 (40.70)
Wealth Status	Middle	420 (82.20)
	Poor	68 (13.71)
	Poorest	14 (2.74)
	Rich	9 (1.76)
Income before Corona	5000	31 (6.07)
	5000-15000	147 (28.77)
	16000-25000	107 (20.94)
	26000-35000	66 (12.91)
	36000-50000	66 (12.91)
	50000+	94 (18.40)
Income after Corona	5000	115 (22.50)
	5000-15000	125 (24.46)
	16000-25000	80 (16.66)
	26000-35000	60 (11.74)
	36000-50000	57 (11.15)
	50000+	74 (14.48)
Lost Job in Corona	No	407 (79.65)
	Yes	104 (20.35)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

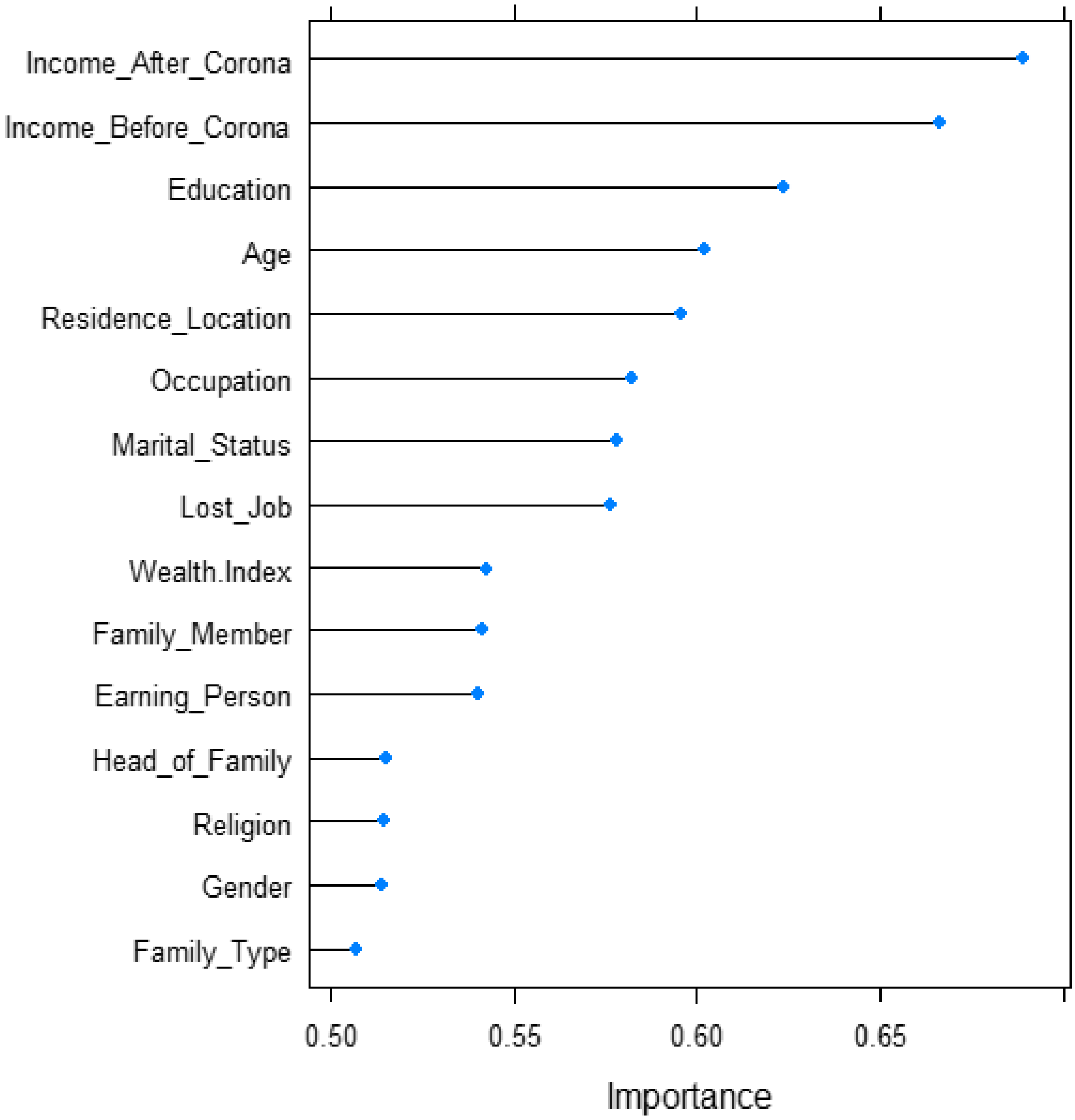


Figure 1. Feature importance score.

Education are the top three significant features in the data set. Whereas Religion, Gender, Family Type are the less important features in our data set

3.4. Machine Learning Technique

In this sub-section, we can briefly explain the three machine learning (ML) algorithms with mathematical expressions.

3.4.1. Random Forest

The random forest (RF) method is an easy way to include a classifier that, even without calibrating the hyper-parameter, induces a great outcome most of the time . It is also one of the most used frameworks because of its elegance and usability [34]. A significant advantage of RF is that they can be utilized supervised learning questions, which make up the bulk of cognitive computing programs. Like a decision tree or bagging classifier in random trees, there are almost the same hyper parameters. We can also do regression tasks for random forests, using the regression algorithm. The RF algorithm forms of a multitude of tree classifiers during which a randomized vector computed individually of the input

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

vector is used to construct each classification model, and each tree imposes a unit vote to assign the input vector according to the most common category.

3.4.2. Logistic Regression

Logistic regression (LR) is a kind of parametric classification model that have a certain fixed number of parameters that depend on the number of input features, and they output categorical prediction. It is binary classification model. The LR model is performed based on a logistical function that is defined as follows [35]:

$$f(X) = \frac{1}{1 + e^{-X}} \tag{2}$$

where X is a weighted sum of the input feature which is defined as $(X = w_1 x_1 + w_2 x_2 + ... + w_n x_n)$, and n is a number of input features.

Now, the logit form of the logistic model can be obtained by the following formula [36]:

$$\text{y=logit(p)=ln}\left(\frac{p}{1-p}\right) = w^T x + b$$

(3)

where logistic $(logit)$ is the ratio of class probabilities, x is the data feature vector, w is the weight, and b is the bias of the model.

Therefore, the benefits of using LR include its flexibility, reliability, and the ability to resist over-fitting without any hyper-parameter tuning in small-scaled data sets.

3.4.3. Naive Bayes

The Naive Bayes (NB) classifier framework is easy to construct for very large volumes of data, and particularly helpful. It is a mathematical model based on Bayes' rule, with the premise that determinants are distinct. In simple terms, a NB learning algorithm from a certain feature in a class is irrelevant to any other functionality being included. The key problem of the naive Bayes approach is the calculation of class conditional density [37]. The class conditional density is typically calculated depending on the data points. We may therefore know the class conditional density from unknown data objects identified by probability distributions for unknown classification problems. The equation provided a mechanism to calculate likelihood function for $P(c)$, $P(x|c)$ and $P(x)$. Look beneath the equation:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{4}$$

where $P(.)$ and $P(|)$ denotes the probability and the conditional probability, respectively, $P(c|x)$ seems to be the posterior probability of group (target) includes integrated (attribute), $P(c)$ is the reflection coefficient of class, $P(x|c)$ is the probability of class received indicator, and $P(x)$ is the likelihood of determinant.

4. Proposed Model Architecture

The pictorial demonstration of the proposed model architecture is illustrated in Figure 2. The architecture comprised of several blocks like data collection, data preprocessing, data slitter, SMOTE (Synthetic Minority Oversampling Technique), three ML algorithm for model training and testing. Normalization and scaling techniques are used to preprocess the collected data. Preprocess data is divided into two groups using data splitter to training and testing the SMOTE. The SMOTE is a data processing technique, which used to solve the imbalance problem of the data set. Thereafter, this balanced data set is divided into training data set (80%) and testing data set (20%) for ML algorithm. Each ML algorithm is trained and tested individually using training data set and test data set, respectively. Finally, we calculate the performance of the proposed model using the ML algorithms for our data set.

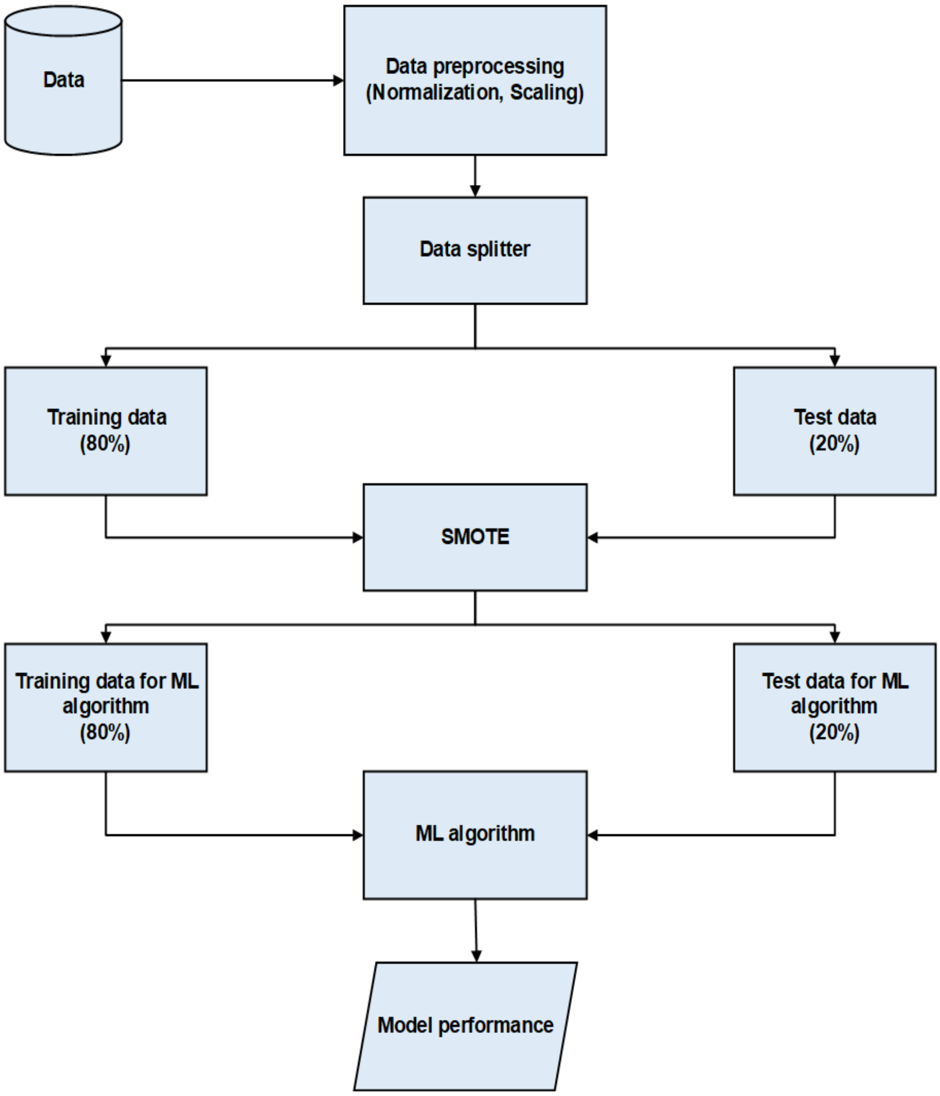


Figure 2. A block diagram of the proposed Model Architecture.

5. Performance Evaluation Criteria

The results of applying proposed model architecture on the data set are evaluated through accuracy, precision, recall/sensitivity, and F-Measure criteria calculated from confusion matrix values. A confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. A confusion matrix for a typical two-value classification problem is presented in Figure 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

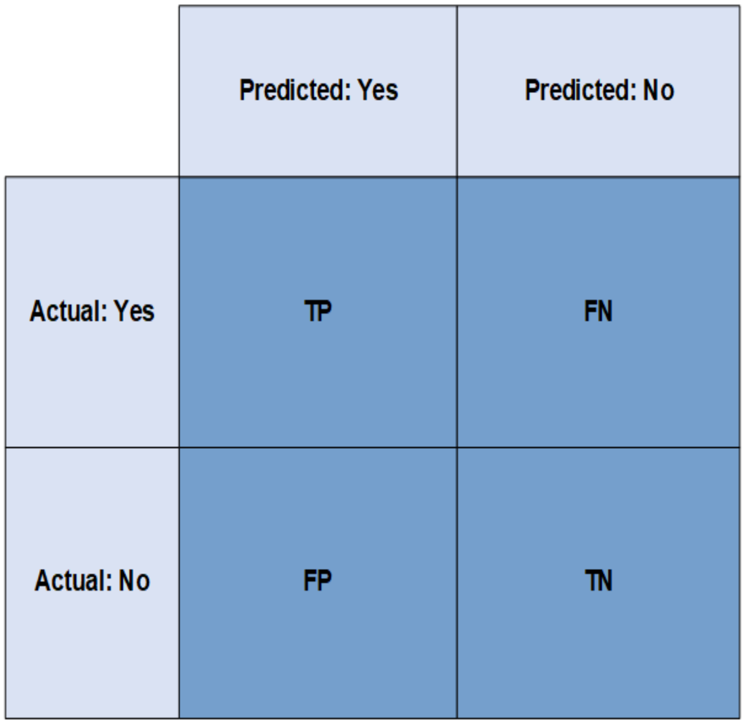


Figure 3. A Confusion matrix of the proposed model Architecture.

The values of accuracy, precision, recall/sensitivity, specificity, F1-score, and specificity are determined based on the true positive value, true negative value, false positive value and false negative value. The definition of true positive value, true negative value, false positive value and false negative value for ML algorithm is given as follows:

- TP (True Positive): The actual observation indicates that domestic violence has occurred and the ML algorithm detects domestic violence from the given data (i. e., the detection result is true positive).
- TN (True Negative): The actual observation indicates that domestic violence has occurred whereas the ML algorithm can not detect domestic violence from the given data (i. e., the detection result is true negative).
- FP (False Positive): The actual observation indicates that no domestic violence occurred and the ML algorithm indicates that no domestic violence is detected from the given data (i. e., the detection result is false positive).
- FN (False Negative): The actual observation indicates that no domestic violence occurred whereas the ML algorithm detects domestic violence from the given data (i.e., the detection result is false negative).

Accuracy is one of the important criteria for evaluating classification models. For binary classification, accuracy can be calculated by using following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

precision is a another important criteria for evaluating classification models. Precision is the ratio between the True Positives and all the Positives (i.e., true positive and true negative). The precision is defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

Recall/sensitivity is the true positive rate. It measures how frequently the experiment detects the domestic violence from the given data when the actually domestic violence has occurred. The recall is defined as follows:

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

F1-Score is a measure of a model's accuracy on a data set. It is used to evaluate binary classification systems, which classify examples into 'positive' or 'negative'. The F-score is a way of combining the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall. The F1-score is defined as follows:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

Specificity is the ability of a experiment to correctly classify the non-domestic violence case. It is the ability to test correctly classify the non-domestic violence case. The specificity is calculated as given:

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

6. Simulation Results and Discussion

In this section, to evaluate the family violence prediction accuracy of three ML algorithms like RF, LR, and NB for our data set. We conduct comprehensive experiments using our data set. For our study, 511 responses are considered from the data collection. In our data set, we consider that family violence has occurred in 229 families and family violence has not occurred in 282 families during the COVID-19 outbreak in Bangladesh. We break our array of data into two sections, where 80% is included in the training phase and the corresponding 20% is also included in the testing set. We have used 10-fold cross validation approach to assess the prediction performance of the three ML algorithms.

We create confusion matrix for test data set to evaluate the predication performance of the three ML algorithms. Firstly, we use imbalance data set to create confusion matrix for three ML algorithms. Secondly, we use balance data set to create confusion matrix for three ML algorithms. The confusion matrix of the three ML algorithms for imbalance data is shown in Figure 4. The confusion matrix of the three ML algorithms for balance data is shown in Figure 5. .

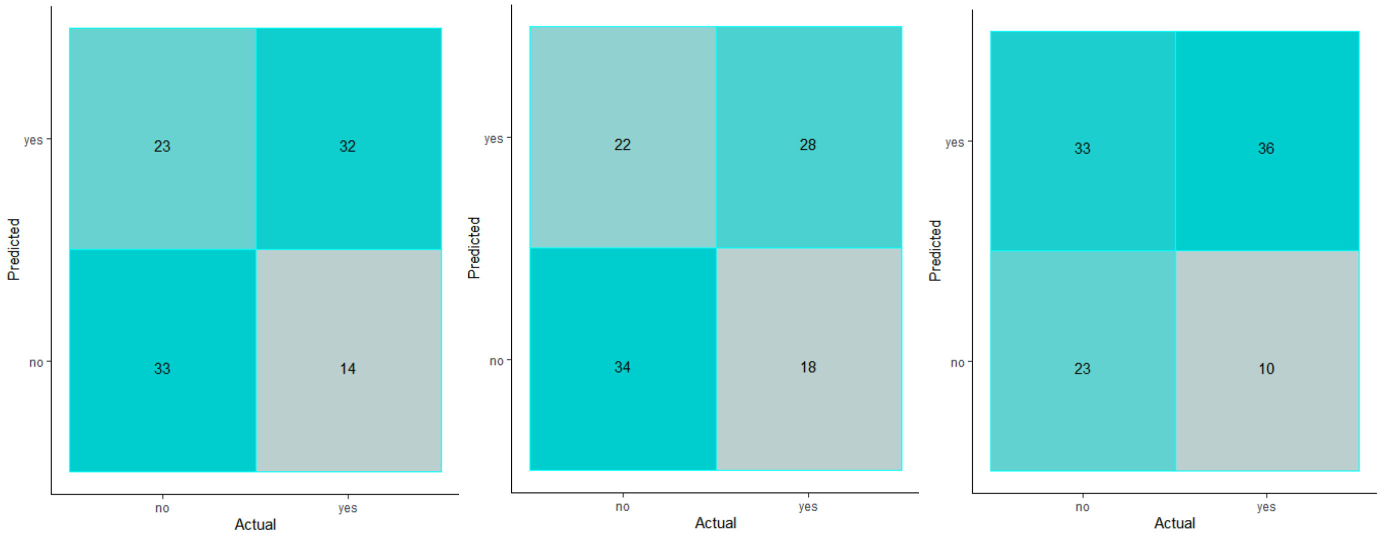
Figure 4 (a), 4 (b), and 4 (c) show the confusion matrix for imbalance data of RF, LR, and NB algorithms, respectively. From these figures, we can see that the true positive and true negative values of these confusion matrices are not better compared to false positive and false negative values. Therefore, we need to balancing our data because our data is not balance. There are many algorithm to handling imbalance data set for machine learning algorithm like under sampling, over sampling, SMOTE. We used SMOTE technique for balancing our data to get better true positive and true negative values.

Figure 5 (a), 5 (b), and 5 (c) show the confusion matrix after data balance and normalization for RF, LR, and NB algorithms, respectively. From these figures, we can see that the true positive and true negative values of these confusion matrices are better compared to false positive and false negative values. Therefore, the values of accuracy, precision, recall/sensitivity, specificity, F1-score, and specificity are enhanced after balancing our data.

From Figure 4 and Figure 5, we can be summarized that the prediction performance of RF, LR, and NB algorithms is better for balance data set compared to imbalance data set.

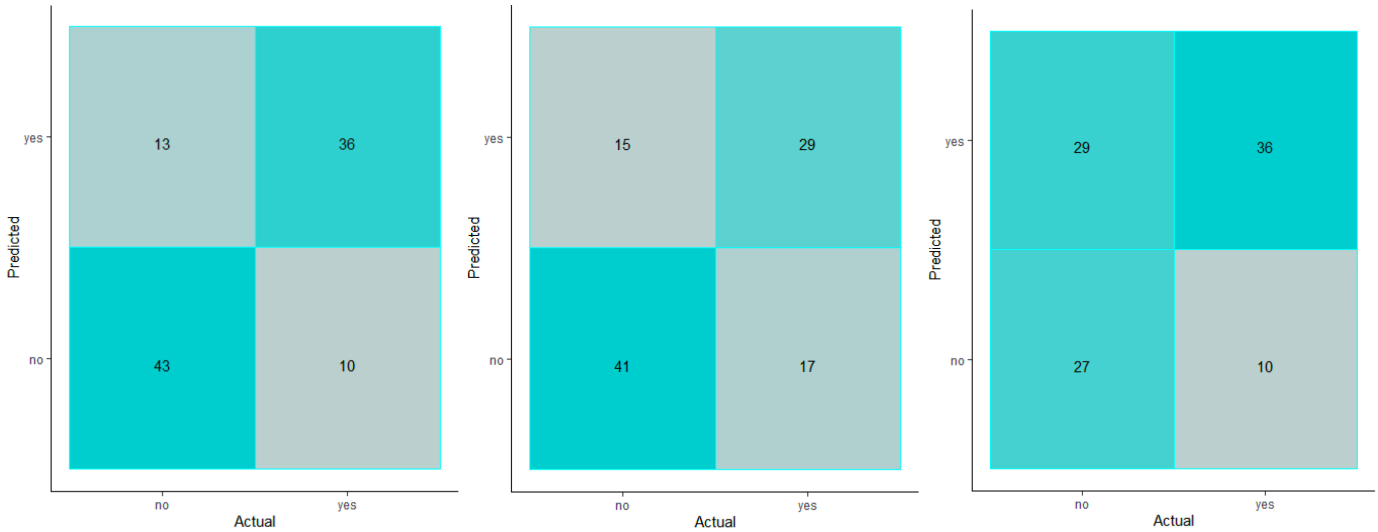
Now we want to check the four measurement criteria of three ML algorithms for our data set. From the Figure 6, we observe that the prediction accuracy of RF, LR, and NB algorithms for imbalance data set is 64%, 61%, and 58%, respectively. In this case, the RF

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1



(a) Confusion matrix for RF **(b)** Confusion matrix for LR **(c)** Confusion matrix for NB

Figure 4. Confusion matrix before data balance and normalization for RF, LR, and NB algorithms.



(a) Confusion matrix for RF **(b)** Confusion matrix for LR **(c)** Confusion matrix for NB

Figure 5. Confusion matrix after balance data set and normalization for RF, LR, and NB algorithms.

algorithm provided highest accuracy, precision, recall and F1-score values whereas NB algorithm is provided the lowest accuracy, precision, recall, and F1-score values.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

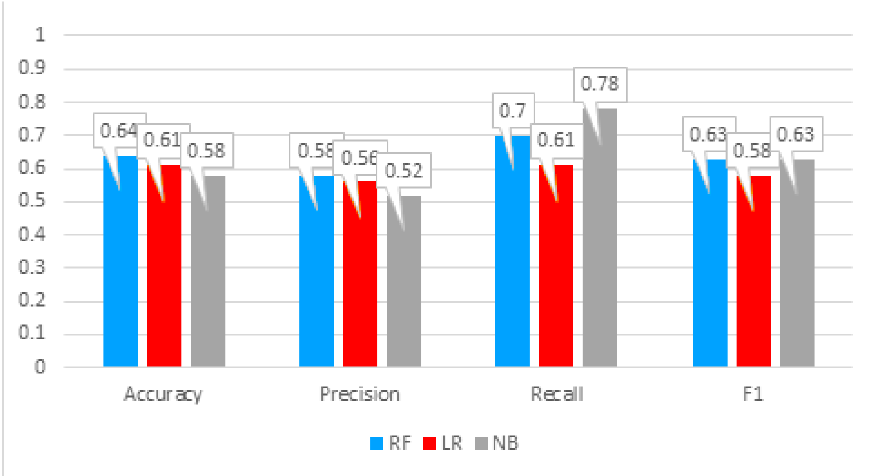


Figure 6. Percentage of classification results with imbalance data set.

We use SMOTE technique to handle the imbalance data set. For balance data set, the prediction performance of RF, LR, and NB algorithms is improved when compared to the imbalance data set. From the Figure 7, we observe that the prediction accuracy of RF, LR, and NB algorithms for balance data set is 77%, 69%, and 62%, respectively. In this case, the RF algorithm is provided the highest accuracy, precision, recall, and F1-score values whilst the NB algorithm is provided the lowest accuracy, precision, recall, and F1-score values.

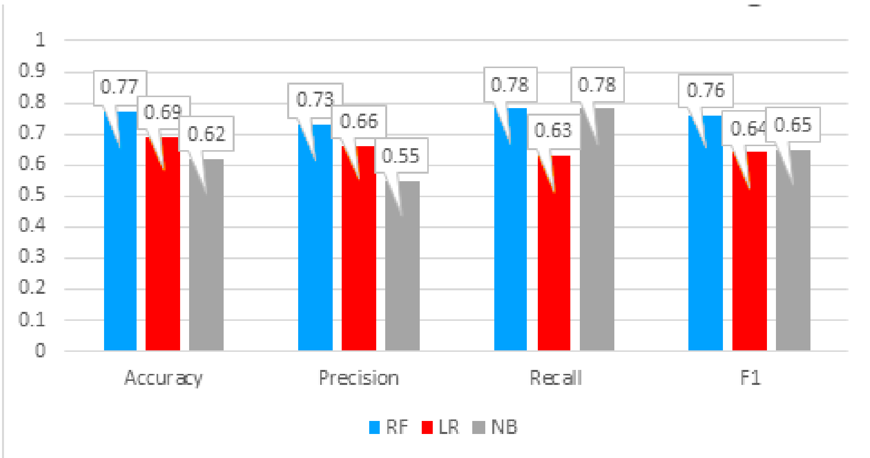


Figure 7. Percentage of classification results with balance data set.

From Figures 6 and Figure 7, we can be summarized that the prediction performance of RF, LR, and NB algorithms is better for balance data set when compared to imbalance data set.

6.1. ROC Curve

The receiver operator characteristic (ROC) curve is a graphical plot used to show the diagnostic ability of binary classifiers. In this paper our main objective to identify the family violence and compare the result for balance data set and imbalance data set. From Figure 8, we can observe that the diagnostic ability for three ML algorithms is very low when we considered the imbalance data set. This means that when fit ML algorithm with imbalance data set then decreases the diagnostic ability whilst increases classification error. Although with imbalance data set the RF algorithm is best performer compared to other ML algorithms like LR, and NB.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2021 doi:10.20944/preprints202104.0343.v1

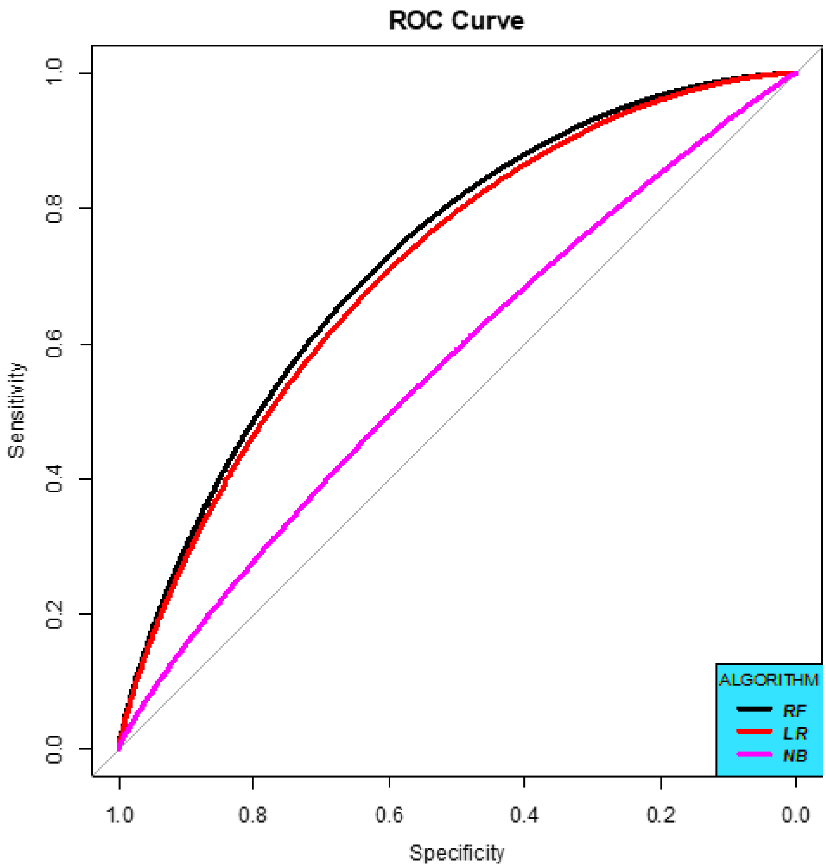


Figure 8. Diagnostic result with imbalance data.

After using SMOTE for balancing our imbalance data. From Figure 9, we can observe that the diagnostic ability for three ML algorithms is increased with balance data when compared to the imbalance data. As a result, the classification error is decreased. Here also RF algorithm is provided better result for balance data set when compared to both LR and NB algorithms.

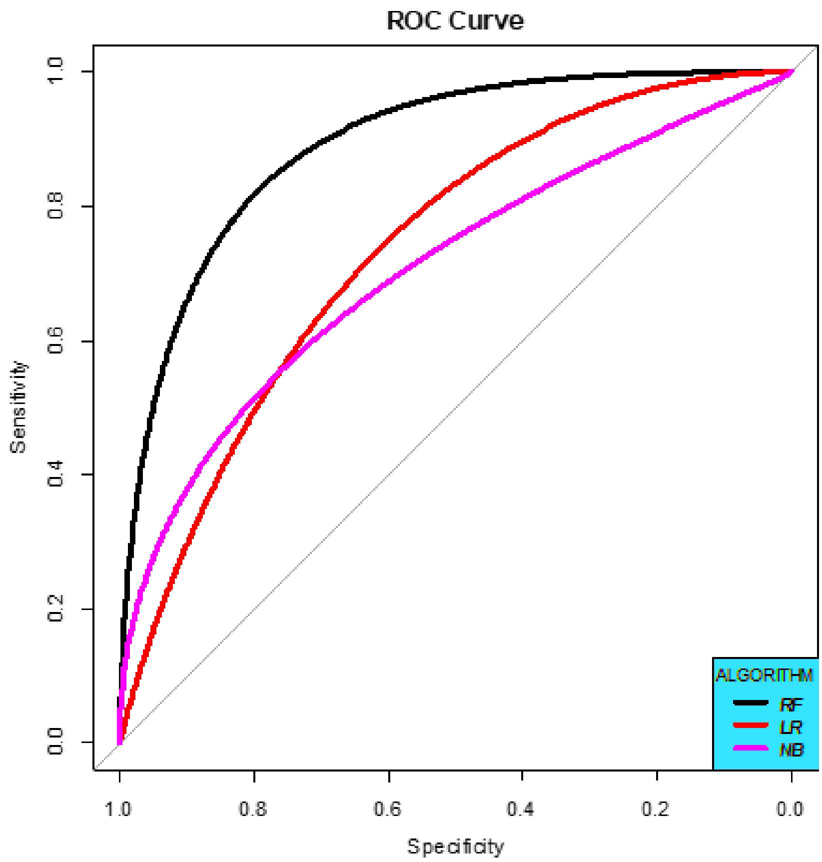


Figure 9. Diagnostic result with balance data set (using SMOTE technique).

7. Conclusion

In this paper we propose the ML algorithms based model to predict domestic violence in Bangladesh during the COVID-19 pandemic. We have monitored the effectiveness of the three ML classification models and assess their performance for our data set. In this study, we have applied the SMOTE technique for data balancing to enhance classifier performance. We achieved a significant improvement in different classifier performance metrics with balance data set. The RF classifier provides better performance compared to the other two ML classifiers. From experimental results, we observed that the accuracy of domestic violence prediction of the RF, LR, and NB classifiers for imbalance data is 64%, 63%, and 58%, respectively. For balance data, the accuracy of domestic violence prediction of the RF, LR, and NB classifiers is 77%, 69%, and 62%, respectively. Therefore, the maximum prediction accuracy is achieved by RF classifier and lowest prediction accuracy achieved by NB for both data sets. With regard to classification precision or the region under the ROC curve, RF classifier showed the best results compared to other algorithms for our data set.

In our future work, we will use other oversampling techniques with ML algorithms to improve the results more reliably.

Author Contributions: Conceptualization: M.M.H., M.A.; Formal Analysis: M.M.H., M.A., M.Z.H., T.P., and M.A.H.; Methodology and Validation: M.M.H., M.A., and M.A.H.; Software: M.A.; Writing - Original Draft Preparation: M.M.H.; Writing - Review & Editing; M.S.M, A.R., and M.A.H.

The authors declare no conflict of interest.

Acknowledgments: This research was supported in part by the Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh.

References

1. Campbell, A.M. An increasing risk of family violence during the Covid-19 pandemic: Strengthening community collaborations to save lives. *Forensic Science International: Reports* **2020**, *2*, 100089.
2. Sánchez, O.R.; Vale, D.B.; Rodrigues, L.; Surita, F.G. Violence against women during the COVID-19 pandemic: An integrative review. *International Journal of Gynecology & Obstetrics* **2020**, *151*, 180–187.
3. Babvey, P.; Capela, F.; Cappa, C.; Lipizzi, C.; Petrowski, N.; Ramirez-Marquez, J. Using social media data for assessing children's exposure to violence during the COVID-19 pandemic. *Child Abuse & Neglect* **2020**, p. 104747.
4. Pereda, N.; Díaz-Faes, D.A. Family violence against children in the wake of COVID-19 pandemic: a review of current perspectives and risk factors. *Child and adolescent psychiatry and mental health* **2020**, *14*, 1–7.
5. Solórzano, D.A.N.; Gamez, M.R.; Corcho, O.d. Gender violence on pandemic of COVID-19. *International Journal of Health Sciences* **2020**, *4*, 10–18.
6. Zhang, H. The influence of the ongoing COVID-19 pandemic on family violence in China. *Journal of family violence* **2020**, pp. 1–11.
7. Every-Palmer, S.; Jenkins, M.; Gendall, P.; Hoek, J.; Beaglehole, B.; Bell, C.; Williman, J.; Rapsey, C.; Stanley, J. Psychological distress, anxiety, family violence, suicidality, and wellbeing in New Zealand during the COVID-19 lockdown: A cross-sectional study. *PLoS one* **2020**, *15*, e0241658.
8. Gonzalez, J.M.R.; Molsberry, R.; Maskaly, J.; Jetelina, K.K. Trends in Family Violence Are Not Causally Associated with COVID-19 Stay-at-Home Orders: a Commentary on Piquero et al. *American journal of criminal justice* **2020**, *45*, 1100–1110.
9. Lidz, C.W.; Mulvey, E.P.; Gardner, W. The accuracy of predictions of violence to others. *Jama* **1993**, *269*, 1007–1011.
10. Blair, R.A.; Blattman, C.; Hartman, A. Predicting local violence: Evidence from a panel survey in Liberia. *Journal of Peace Research* **2017**, *54*, 298–312.
11. Berk, R.A.; Sorenson, S.B.; Barnes, G. Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies* **2016**, *13*, 94–115.
12. Grogger, J.; Gupta, S.; Ivandic, R.; Kirchmaier, T. Comparing conventional and machine-learning approaches to risk assessment in domestic abuse cases. Technical report, National Bureau of Economic Research, 2021.
13. Rodríguez-Rodríguez, I.; Rodríguez, J.V.; Pardo-Quiles, D.J.; Heras-González, P.; Chatzigiannakis, I. Modeling and Forecasting Gender-Based Violence through Machine Learning Techniques. *Applied Sciences* **2020**, *10*, 8244.
14. Gosangi, B.; Park, H.; Thomas, R.; Gujrathi, R.; Bay, C.P.; Raja, A.S.; Seltzer, S.E.; Balcom, M.C.; McDonald, M.L.; Orgill, D.P.; others. Exacerbation of Physical Intimate Partner Violence during COVID-19 Pandemic. *Radiology* **2021**, *298*, E38–E45.
15. Gebrewahd, G.T.; Gebremeskel, G.G.; Tadesse, D.B. Intimate partner violence against reproductive age women during COVID-19 pandemic in northern Ethiopia 2020: a community-based cross-sectional study. *Reproductive health* **2020**, *17*, 1–8.
16. Pfitzner, N.; Fitz-Gibbon, K.; True, J. Responding to the 'shadow pandemic': practitioner views on the nature of and responses to violence against women in Victoria, Australia during the COVID-19 restrictions **2020**.
17. Sharma, A.; Borah, S.B. Covid-19 and domestic violence: an indirect path to social and economic crisis. *Journal of family violence* **2020**, pp. 1–7.
18. Sediri, S.; Zgueb, Y.; Ouanes, S.; Ouali, U.; Bourgou, S.; Jomli, R.; Nacef, F. Women's mental health: acute impact of COVID-19 pandemic on domestic violence. *Archives of women's mental health* **2020**, pp. 1–8.
19. Evans, D.P.; Hawk, S.R.; Ripkey, C.E. Domestic Violence in Atlanta, Georgia Before and During COVID-19. *Violence and Gender* **2020**.
20. Collin-Vézina, D.; Brend, D.; Beeman, I. When it counts the most: Trauma-informed care and the COVID-19 global pandemic. *Developmental Child Welfare* **2020**, *2*, 172–179.
21. Brown, S.M.; Doom, J.R.; Lechuga-Peña, S.; Watamura, S.E.; Koppels, T. Stress and parenting during the global COVID-19 pandemic. *Child abuse & neglect* **2020**, *110*, 104699.

22. Xiang, X.; Lu, X.; Halavanau, A.; Xue, J.; Sun, Y.; Lai, P.H.L.; Wu, Z. Modern senicide in the face of a pandemic: an examination of public discourse and sentiment about older adults and COVID-19 using machine learning. *The Journals of Gerontology: Series B* **2020**.
23. Buttell, F.; Ferreira, R.J. The hidden disaster of COVID-19: Intimate partner violence. *Psychological trauma: theory, research, practice, and policy* **2020**, *12*, S197.
24. Bryant, D.J.; Oo, M.; Damian, A.J. The rise of adverse childhood experiences during the COVID-19 pandemic. *Psychological Trauma: Theory, Research, Practice, and Policy* **2020**.
25. Moreira, D.N.; da Costa, M.P. The impact of the Covid-19 pandemic in the precipitation of intimate partner violence. *International journal of law and psychiatry* **2020**, *71*, 101606.
26. Guessoum, S.B.; Lachal, J.; Radjack, R.; Carretier, E.; Minassian, S.; Benoit, L.; Moro, M.R. Adolescent psychiatric disorders during the COVID-19 pandemic and lockdown. *Psychiatry research* **2020**, p. 113264.
27. Koh, J.X.; Liew, T.M. How loneliness is talked about in social media during COVID-19 pandemic: text mining of 4,492 Twitter feeds. *Journal of psychiatric research* **2020**.
28. Xue, J.; Chen, J.; Hu, R.; Chen, C.; Zheng, C.; Su, Y.; Zhu, T. Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. *Journal of medical Internet research* **2020**, *22*, e20550.
29. Xue, J.; Chen, J.; Chen, C.; Hu, R.; Zhu, T. The Hidden Pandemic of Family Violence During COVID-19: Unsupervised Learning of Tweets. *Journal of medical Internet research* **2020**, *22*, e24361.
30. Abuhammad, S. Violence against Jordanian Women during COVID-19 Outbreak. *International journal of clinical practice* **2021**, *75*, e13824.
31. Amusa, L.B.; Bengesai, A.V.; Khan, H.T. Predicting the vulnerability of women to intimate partner violence in South Africa: evidence from tree-based machine learning techniques. *Journal of interpersonal violence* **2020**, p. 0886260520960110.
32. Wang, K.Z.; Bani-Fatemi, A.; Adanty, C.; Harripaul, R.; Griffiths, J.; Kolla, N.; Gerretsen, P.; Graff, A.; De Luca, V. Prediction of physical violence in schizophrenia with machine learning algorithms. *Psychiatry research* **2020**, *289*, 112960.
33. Chen, I.Y.; Alsentzer, E.; Park, H.; Thomas, R.; Gosangi, B.; Gujrathi, R.; Khurana, B. Intimate Partner Violence and Injury Prediction From Radiology Reports. *arXiv preprint arXiv:2009.09084* **2020**.
34. Sadorsky, P. A Random Forests Approach to Predicting Clean Energy Stock Prices. *Journal of Risk and Financial Management* **2021**, *14*, 48.
35. Miah, M.S.; Hossain, M.A.; Ahmed, K.M.; Rahman, M.; Calhan, A.; Cicioglu, M.; others. Machine Learning-Based Malicious User Detection in Energy Harvested Cognitive Radio-Internet of Things **2021**.
36. Borucka, A.; Grzelak, M. Application of logistic regression for production machinery efficiency evaluation. *Applied Sciences* **2019**, *9*, 4770.
37. Abubakar, A. Comparative Analysis of Classification Algorithms Using CNN Transferable Features: A Case Study Using Burn Datasets from Black Africans. *Applied System Innovation* **2020**, *3*, 43.