A machine learning-based domestic violence prediction and Android application-based domestic violence prevention assistance system

Sheikh Rasel and Mahfuzulhoq Chowdhury*

Computer Science and Engineering Department,
Chittagong University of Engineering and Technology,
Chittagong 4349, Bangladesh
Email: sheikhrasel195@gmail.com
Email: mahfuz@cuet.ac.bd
Email: mahfuzulhoq.cse05@gmail.com
*Corresponding author

Abstract: Domestic violence is a widespread issue in today's society. Numerous efforts are currently being made to reduce domestic violence. Due to the number of repeated offenses and the various pattern of behaviour, this is a difficult problem to control. Existing works did not investigate both emergency and non-emergency help for users of domestic violence help using Android applications as well as machine learning-based violence prediction. This paper prepares the dataset through survey questions and compares the performance of various machine learning-based algorithms for violence prediction. This paper develops a mobile application to provide instant and legal help to victims of domestic violence. This paper offers call button features that make automatic calling, an instant location tracking system, and automatic location along with video/image transfer to the nearest police station, phone shaking features for emergency help. The evaluation shows the usefulness of this application.

Keywords: domestic violence; machine learning; prediction; legal help; instant help; Android application.

Reference to this paper should be made as follows: Rasel, S. and Chowdhury, M. (2024) 'A machine learning-based domestic violence prediction and Android application-based domestic violence prevention assistance system', *Int. J. Applied Management Science*, Vol. 16, No. 1, pp.68–88.

Biographical notes: Sheikh Rasel received his BSc in Computer Science and Engineering at the Department of CSE, Chittagong University of Engineering and Technology in 2022. His research interests include machine learning and mobile application development.

Mahfuzulhoq Chowdhury received his PhD in Telecommunications at the University of Quebec, INRS, Montreal, Canada. He is a faculty member at the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology since 2010. His research interests are related to machine learning, optimisation, mobile application development, game theory, resource allocation in future generation networks, cloud computing, IoT, age of information, and blockchain. He has published several research papers at highly cited IEEE journals, transactions, magazines, and conference proceedings as well as chapters of books.

1 Introduction

Domestic violence (DV) is described as violence between partners or spouses. Family members and friends can also be violence victims. Everyone, regardless of age, culture, social status, or degree of education, is susceptible to the global epidemic of domestic abuse. Violence may be present on occasion or regularly. As a result of family violence, suicides, violent incidents, and the torturing of women and children are frequently documented (Campbell, 2016). According to the Global Health Organization, 35% of women have experienced intimate partner violence (Garcıa-Moreno et al., 2013). According to certain national studies, DV affects up to 70% of women nationwide. Physical assaults like striking, kicking, hurling objects, slapping, sexual abuse, abuse of emotions, mental illness, and alcohol use can all be considered forms of violence. During the pandemic, there has been an increase in family violence. There is not a straightforward idea in the research that can give clinicians who treat victims of DV useful direction (Sharma and Borah, 2020). Numerous heartbreaking incidents of DV against women over the past ten years have intensely compelled scholars and administrators to consider the harm it causes to society as a whole and to take effective action to stop it. In terms of DV, 50% to 70% of developing countries' women (i.e., Bangladeshi) report being abused by their male partners, placing the country among the world's worst for violence against women (Khatun and Rahman, 2012). With all of these factors in mind, it is only necessary to take some sensible steps to eradicate DV from Bangladeshi society.

DV has been one of the world's most common problems, and victims are subjected to not just physical but also sexual, psychological, and verbal abuse (Sharma and Borah, 2020). Victims of DV can get both informal and formal assistance to improve their safety, physical, and mental health (Khatun and Rahman, 2012). This idea has the potential to be vital in providing DV victims with resources such as a crisis hotline and an emergency shelter. As per a 2015 poll released by the Bangladesh Bureau of Statistics (BBS) and the United Nations Population Fund (UNFPA), more than 70% of wedded women and girls in Bangladesh had witnessed family violence, with roughly half disclosing physical attacks by their partners. In the first nine months of 2020, a minimum of 235 women were killed by husbands or relatives, according to Bangladeshi Human Rights Organization Ain o Salish Kendra (ASK). Between January 2001 and December 2019, nearly 3,300 females were murdered because of dowry disputes, according to Odhikar, a prominent Bangladeshi human rights group (Ara, 2020). These numbers, however, are based on news reports and are likely a small fraction of the actual level of violence. Violence against girls women and seems to have risen in Bangladesh throughout the COVID-19 epidemic, with NGO helplines having recorded a rise in panicked calls.

However, the prevention of DV has aroused serious concerns all over the world because of the disparity between the increasing number of reported DV cases. Various software solutions address victim violence, but only a few have developed user-friendly systems. Because of the lack of multiple features in one single application, most applications have some limitations. By analysing recent literature, some studies have examined the factors underlying family abuse. However, only a small number of studies included machine learning methods. Moreover, the literature works did not investigate emergency and non-emergency help for DV using Android applications. Further, different real-time DV features and factors were out of their investigations.

To assist developing countries' people, this paper proposes a machine learning approach to predict whether DV had occurred or not and what type of assistance the victim would require. This paper also creates a smartphone application that aids victims of DV. Victims will not only report on the situation but also send evidence for the case to the police station. On the other hand, it will also provide some legal suggestions to the victim on what steps to follow after the violence. Although the Android operating system is used by the vast majority of smartphones in developing countries like Bangladesh, our proposed application is built for both Android and iOS. The remarkable contributions and features of this paper are given as follows:

1 This paper collects DV data that can be used for future research purposes.

2 This paper designs a model that can predict DV with higher accuracy.

3 This paper develops an Android application for preventing DV. An automated message will be sent to the emergency number in case of an emergency if the user shakes the phone three times.

4 In this application, a real-time database has been used, but in previous work, this has not been used. By using our proposed features, the user can attach necessary evidence. By analysing the evidence, the police will provide immediate help and legal help according to the situation.

5 Our proposed mobile application is a combination of more than one module in a single application. Here, the user can notify the police station with evidence and can take legal help. The DV victim can use this application to notify the police, and they can also seek help.

Next, Section 2 briefly summarises some related previous works. Section 3 demonstrates the detailed research of the proposed technique, along with the design of the model that predicts whether or not DV has occurred. It will also illustrate the data flow and development process of the mobile application. Section 3 also explains the operational dataset as well as the analysis of accuracy, precision, recall, F-measure, and other metrics using various algorithms. Section 4 gives the proposed android application discussion with features. Section 5 provides the user review based evaluation results regarding our proposed application. Section 6 provides an overview of this paper as well as some future work suggestions.

2 Related works

Instead of concentrating on women who risk DV indoors, academics have been focusing in recent years on mobile applications addressing outdoor safety for women. Researchers mainly focus on local monitoring and alerting authorities to a tense situation by rocking the phone which sends alerts to the police. As a result, a system with all the features in a single program is important. In recent years, mobile applications as well as other strategies, have been employed in studies to reduce violence against women. Srinivas et al. (2021) presented the Android app for women's safety. When a user has a problem or wants assistance, he can use this program to aid. The HELP button is visible when the user opens this application. A message and three phone numbers can also be stored for him. Chakraborty et al. (2021) designed this NAARI Safety Application, which is

intended to help users learn about vehicles by scanning QR codes by directing the camera toward the code. Siddique et al. (2022) utilised machine learning techniques to predict the depression pattern of undergraduate university students.

Khandoker et al. (2019) developed an Android-based application system for women's safety. It is a smartphone app designed for women's safety, even as men can also use it in an emergency. It can be triggered by voice commands or the SOS button. Every five minutes, until the system is turned off, an SMS alert with the position is sent to the user-defined numbers. Here, this application does not inform the police station and only the location with text is sent to the user-defined number without any evidence. Hossain et al. (2019) released a smartphone app for personal safety. This system uses SOS modules to ensure security, such as top number calling and sending the original position to designated mobile phone numbers. This application was platform dependent and was mainly for the outdoor safety of women. The FEMME system was introduced for safety by Monisha and others in Braithwaite et al. (2016). The FEMME keeps track of the current position, then sends it to a known phone number, where audio is recorded as evidence. Here, just the pre-set contact number receives the actual position. But this application does not send location to the police station, and they do not use any real-time database. Sriranjini et al. (2017) presented a women's safety self-defence system based on the GPS and GSM systems. The proposed method allows tracking the victim's position using Google Maps by displaying the victim's latitude and longitude.

Chand et al. (2015) suggested a safety mobile app. By using this application, the user can report an unpleasant scenario to their relatives. Currently, in this application, the emergency message and call are set automatically by shaking the phone 40 times in a row within about eight seconds. Sharma and More (2016) proposed and built a mobile program for sending alerts to a chosen contact number and dialling a default number automatically. The Women's Security App will only call the first registered number when the volume key is pressed, but it will send an emergency message to every contact placed on the emergency contact list, but it does not send any evidence of the violence. Mane et al. (2016) presented a safety-focused study. This system continuously monitors the current location and communicates the location as well as an image to a phone number that has been assigned. For interior use, the bulk of the systems listed above lacks adequate security, usability, communication, and system interface features (Sharma and Borah, 2020). However, these programs do not provide all of the functionality in one package. As a result, it is necessary to develop a new system that supplies indoor amenities to all required features in one application. This is particularly important for mothers and children who are confined to their homes. Different from the existing study, this paper proposes an Android-based application that gives victims instant and legal assistance while also predicting DV and also uses machine learning for victim prediction.

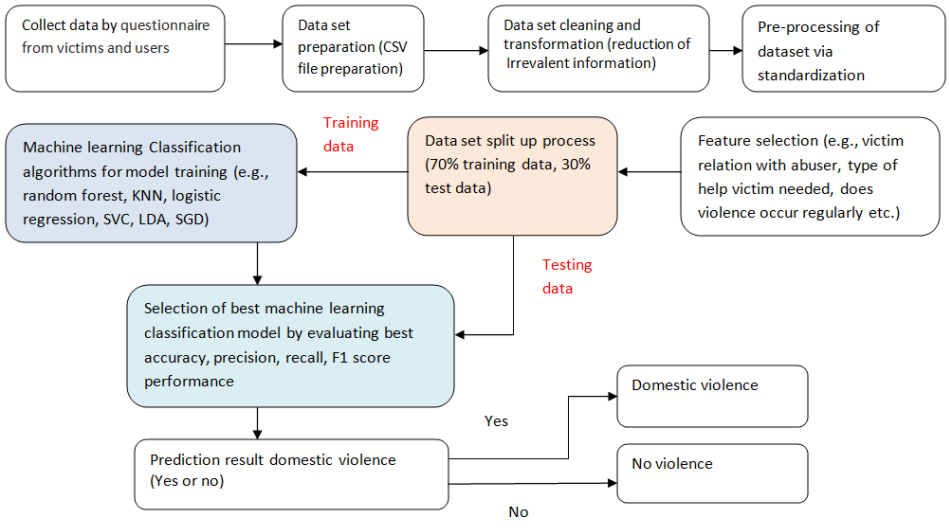
3 Proposed model

This section provides a comprehensive overview of how we created the model that predicts whether DV has occurred or not. Then, data collection procedures are discussed in detail. Then we demonstrated how we designed the mobile application to provide the victim with immediate and legal assistance.

3.1 System design

The entire procedure begins with questionnaires asked of the victim or user. Then user data was collected in an Excel sheet for further use. Data cleaning was done by removing incomplete data and finalising it for data analysis. This entire procedure is depicted in Figure 1. First, we have collected the dataset. Then, we performed the dataset cleaning and pre-processing. Then, we performed the feature selection. After that, we have split the dataset into training and testing data. Then, we have selected the best machine learning classifier model for DV prediction. A comparison of various machine learning algorithms such as logistic regression (LR), K-neighbours classifier, support vector classifier (SVC), random forest (RF), stochastic gradient descent (SGD) classifier, and linear discriminant analysis (LDA) including accuracy, precision, and recall for analysis will be shown in the next section. The main reason for using different algorithms is that different algorithms perform better in different situations. Depending on the algorithms' specialisation, some operate better with little data while others are effective with large data. When determining whether a user has experienced domestic abuse or not, we want our module to provide the highest level of accuracy.

Figure 1 Block diagram for DV prediction (see online version for colours)

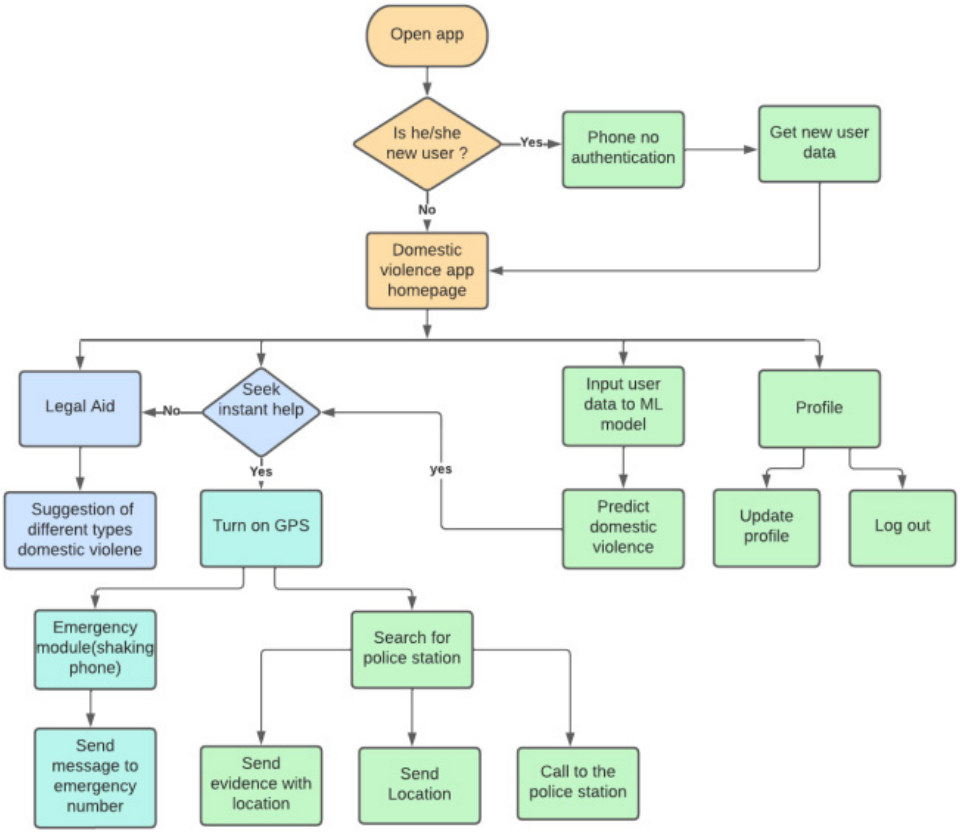


3.2 System design for mobile application

The complete data flow diagram for the mobile application is shown in Figure 2. If the user is new, then first his phone number will be authenticated and he or she will also provide some data along with an emergency number. Then the user will move to the DV homepage. If a user previously used this app and did not log out, then he or she will directly move to the homepage. On this page, the user will find his location, and if the user faces DV and needs emergency help, he or she will select 'contact police' and turn on the GPS module. Here user will select the nearest police station, and then the victim can either call the police station or user can send a message along with a video or image

to the police station number. If the victim is unable to call, he can shake his phone three times, which will send his location to the victim's emergency number. On the legal help page, victims will get suggestions on how they can proceed to get legal justice against the abuser. Usually, the majority of domestic abuse victims are not aware of the law on DV. That is why we think this application will provide relief for the victims of DV.

Figure 2 Data flow diagram of mobile application (see online version for colours)



3.2.1 Dataset preparation and data pre-processing

To collect the dataset from the respondents, we have used information from the police station officers, district-level court judges, and local village authority council members. We have asked any victims to fill up the questionnaire form (i.e., those who want to share their DV experience) via both online and offline basis. The dataset is verified by the local police officers and district-level judges or officers (i.e., those who provide DV support to users). To maintain the confidentiality of users, we did not include their names in the dataset. As the dataset is collected based on real-time occurrences from the victims, we can say that it is valid for any future scientific experiments or research. For future or public use, we have already uploaded our dataset into the online Kaggle platform

(Chowdhury et al., 2023). It can be noted that, before our work, there is no internet dataset for DV among the Bangladeshi people. As a result, we have gathered information from the victims as well as via surveys using Google Forms. We have collected all police station numbers and location information. We have collected datasets directly from DV victims. To get further information, we conducted an online survey on domestic abuse in Bangladesh. Using the online Google Doc platform and offline form, we initially created several family violence-related questionnaires for this survey. Then, to get the data, we forwarded this link to the respondent through e-mail and messenger. We got 352 datasets in ten days. Our dataset included factors like age, gender, and marital status, place of residence, have you or your family members faced any DV? Then if yes, what kind of violence did you or your family members face?, abuser causes of DV, is the violence you or your family members have been facing life-threatening?, does violence happen regularly?, what type of help did the victim need?, do you think that due to a lack of basic needs, family violence may happen?, do you think that family violence may occur due to dowry?, etc. Figure 3 shows the dataset snapshot for DV. Data is noisy, thus pre-processing is essential before creating a model. Otherwise, the model might not work or produce reliable results. The process of converting unprocessed data into a format that may be used is referred to as 'data processing'. For preparing the data, we used the pandas' library. The initial step should include pre-processing the data to reduce implementation time and improve results. We normalised the data for this purpose. For all of the features in this sample, we applied min-max feature scaling. It is a form of scaling in which points are changed and re-scaled until they fall between 0 and 1.

3.2.2 Data visualisation

The study of how to graphically represent data is known as data visualisation. It effectively communicates inferences from data by graphically plotting the data. Both small and large datasets benefit greatly from the use of data visualisation. Python libraries used for data visualisation include Matplotlib and Seaborn. They have built-in modules for generating various graphs. While Seaborn is mainly used for statistical charts, Matplotlib is utilised to embed graphs into applications. In our application, we have used the Seaborn library for data visualisation. There are various ways to visualise data. As a consequence, we will represent our outcomes with a histogram, as it is a method for representing the distribution of numerical data. The complete list of bars is displayed on the y-axis of a histogram, although each bar represents a specific piece of data on the x-axis. The orange bar represents 'faced DV' whereas the blue bar represents 'not faced DV'. Figure 4 demonstrates that among 338 users, 200 faced DV. On the other hand, 138 users did not face DV.

Figure 5 demonstrates that, among 338 users, 102 regularly faced DV. On the other hand, 96 users did not face DV regularly, and 140 users did not face DV. The orange bar represents 'faced DV regularly', whereas the blue bar represents 'not faced DV regularly' and the green bar represents 'not faced DV'.

Figure 3 Dataset of DV

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Timestamp	Your Ag	Gender*	Marital status	Place of res	Have you or your far	If yes then victim relation	If yes, what kind of violenc	Abuser causes of dom	Is the violence you	Does violence happen	What type of help victim r	Do you think that due t	What Do you think
135	7/27/2022 1:11:47	24	Male	Unmarried	Urban	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Disagree	Neutral
136	7/27/2022 1:15:03	25	Female	Unmarried	Rural	Yes	Father, Mother	Mental, Verbal	Family problem	No	Yes	Legal help	Disagree	Agree
137	7/27/2022 1:17:15	22	Male	Unmarried	Urban	Yes	Mother, Siblings	Verbal	Economic problem, Fa	No	Yes	Not applicable	Agree	Strongly Agree
138	7/27/2022 6:40:39	18	Male	Unmarried	Rural	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Neutral	Strongly Agree
139	7/27/2022 8:50:57	24	Male	Unmarried	Urban	Yes	Wife	Physical, Economical, M	Economic problem, Fa	Yes	Yes	Legal help	Strongly Disagree	Agree
140	7/27/2022 9:47:16	25	Male	Unmarried	Rural	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Neutral	Strongly Disagree
141	7/27/2022 10:19:1	24	Male	Unmarried	Rural	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Strongly Agree	Strongly Agree
142	7/27/2022 12:18:2	24	Female	Unmarried	Urban	Yes	Mother, Others	Physical, Mental, Verbal	Family problem, Other	No	No	Not applicable	Neutral	Agree
143	7/27/2022 14:24:0	32	Female	Married	Urban	Yes	Siblings	Mental, Verbal	Alcoholic, Economic p	Yes	Yes	Instant help, Legal help	Disagree	Strongly Agree
144	7/27/2022 14:26:2	48	Male	Married	Rural	Yes	Son	Verbal	Alcoholic	No	Yes	Legal help	Agree	Strongly Agree
145	7/27/2022 14:30:41	32	Female	Married	Urban	Yes	Husband	Sexual	Alcoholic	Yes	Yes	Instant help	Neutral	Agree
146	7/27/2022 14:37:3	22	Female	Married	Rural	Yes	Husband, Mother in law	Physical, Verbal	Family problem	Yes	Yes	Instant help	Agree	Strongly Agree
147	7/27/2022 14:37:4	40	Female	Married	Rural	Yes	Son	Mental, Verbal	Alcoholic	No	No	Legal help	Disagree	Agree
148	7/27/2022 14:41:4	24	Male	Unmarried	Urban	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Neutral	Agree
149	7/27/2022 14:44:2	35	Female	Married	Urban	Yes	Son	Verbal	Economic problem	No	Yes	Legal help	Disagree	Agree
150	7/27/2022 14:52:1	24	Male	Unmarried	Urban	Yes	Siblings	Mental, Verbal	Economic problem, Fa	No	Yes	Legal help	Disagree	Strongly Agree
151	7/27/2022 15:01:0	12	Male	Unmarried	Rural	Yes	Father	Physical	Others	No	Not applicable	Legal help	Disagree	Neutral
152	7/27/2022 15:15:31	22	Male	Unmarried	Urban	No	Not applicable	Not applicable	Not applicable	No	No	Not applicable	Strongly Agree	Strongly Agree
153	7/27/2022 15:45:4	35	Female	Married	Urban	Yes	Siblings	Economical, Verbal	Alcoholic	Yes	Not applicable	Instant help	Disagree	Agree
154	7/27/2022 15:48:5	23	Male	Unmarried	Rural	No	Not applicable	Not applicable	Not applicable	No	Not applicable	Not applicable	Agree	Agree
155	7/27/2022 16:43:5	28	Female	Married	Rural	Yes	Husband	Sexual	Alcoholic	No	Not applicable	Legal help	Disagree	Agree
156	7/27/2022 17:47:1	21	Male	Unmarried	Urban	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Agree	Agree
157	7/27/2022 19:49:2	24	Female	Unmarried	Urban	Yes	Husband	Economical, Mental, Verb	Mental illness, Family	No	No	Not applicable	Neutral	Neutral
158	7/27/2022 22:36:1	24	Male	Unmarried	Rural	No	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Agree	Agree

Figure 4 Data comparison for have you or your family members facing/faced any domestic violence? (see online version for colours)

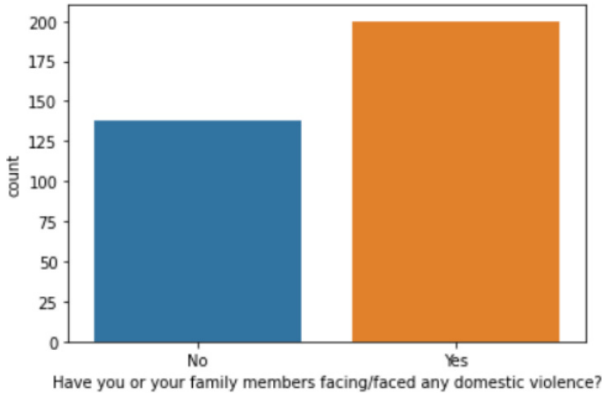


Figure 5 Data comparison for does violence happen regularly? (see online version for colours)

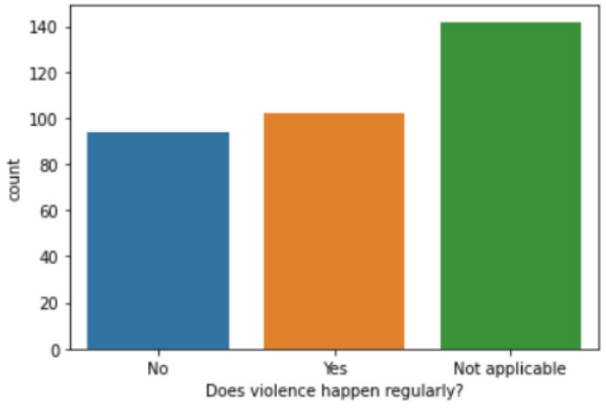


Figure 6 Confusion matrix for LR (see online version for colours)

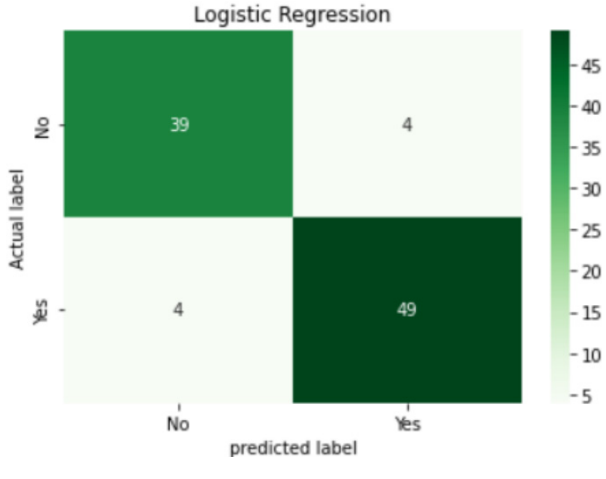


Figure 7 Confusion matrix for SVC (see online version for colours)

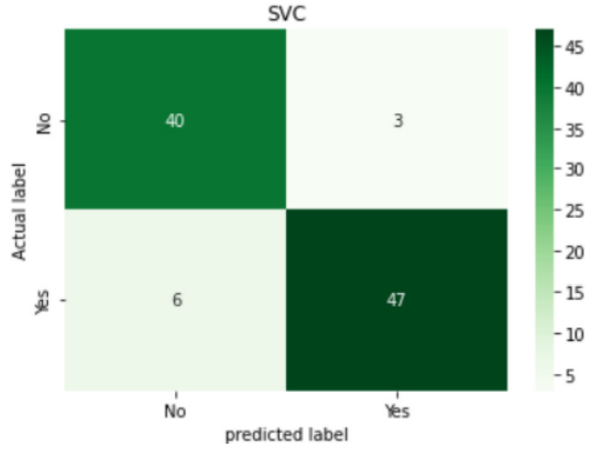


Figure 8 Confusion matrix for KNN (see online version for colours)

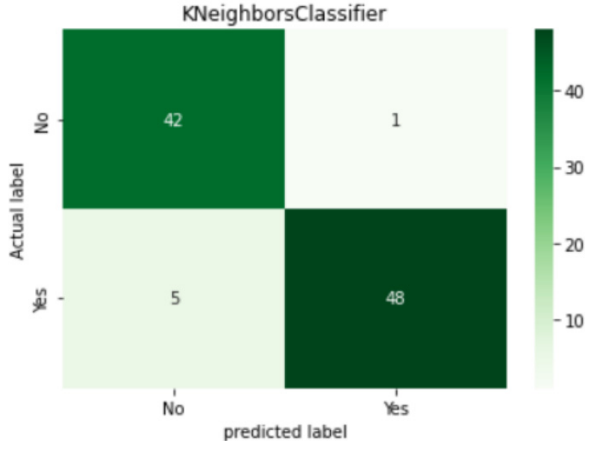


Figure 9 Confusion matrix for SGD (see online version for colours)

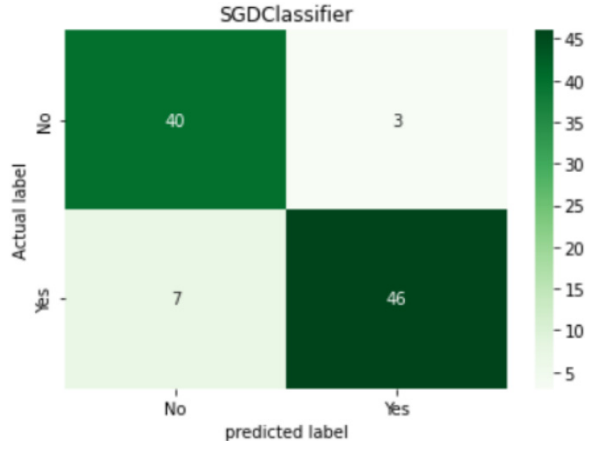


Figure 10 Confusion matrix for LDA (see online version for colours)

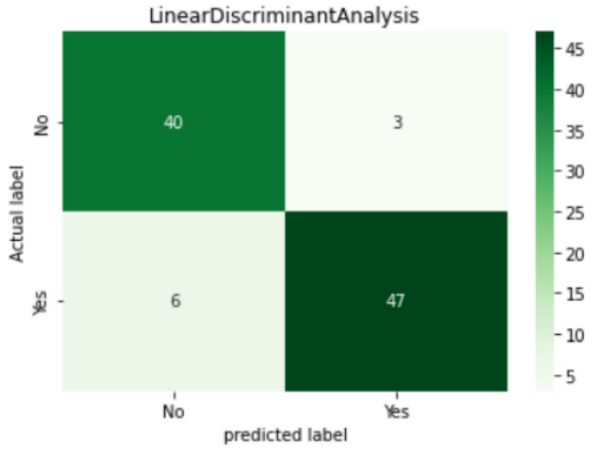


Figure 11 Confusion matrix for RF (see online version for colours)

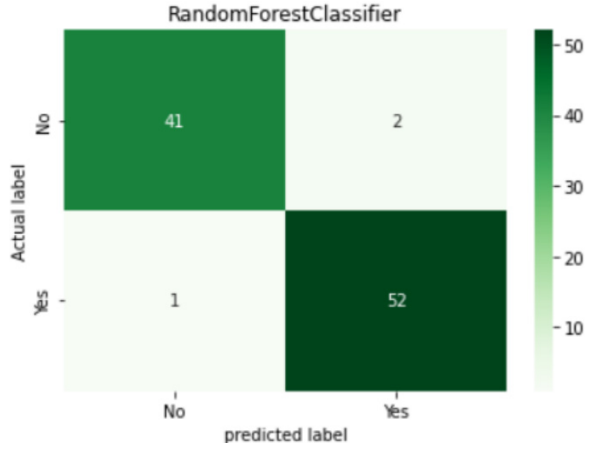
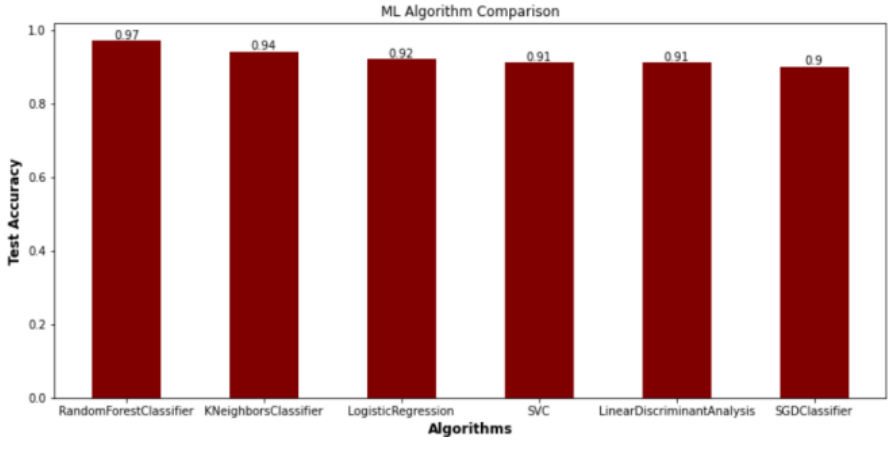


Figure 12 Accuracy comparison of different ML algorithms (see online version for colours)



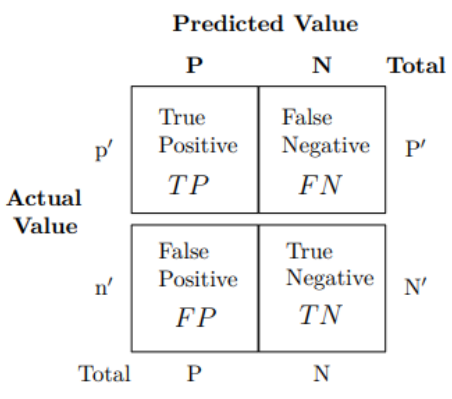
3.2.3 Feature extraction

The feature significance describes which features in the data collection are more valuable or significant than others. By using feature selection, it may help to better understand the problem that has been solved and occasionally result in model enhancements. In essence, feature significance is a method that rates input features according to how well they can predict a target variable. From our dataset, we computed the feature importance and uncorrelated and pointless features will be eliminated during feature extraction, which might aid in fitting the valuable features to machine learning algorithms and increase the model's accuracy.

Table 1 Evaluation result of different metrics

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*
Random forest classifier	97%	96%	98%	97%
KNN classifier	94%	98%	91%	94%
Logistic regression	92%	93%	92%	92%
SVC	91%	94%	89%	91%
Linear discriminant analysis	91%	94%	89%	91%
SGD classifier	90%	94%	87%	90%

Figure 13 Evaluation formula of different metrics



Confusion Matrix

(a)

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

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

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

$$\text{F-score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

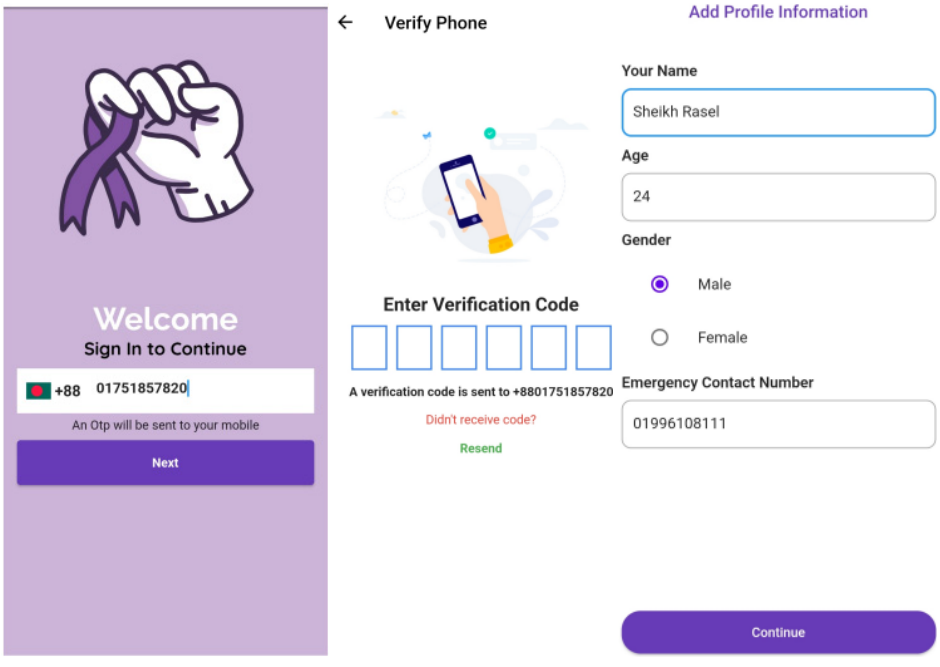
(b)

3.2.4 Evaluation of framework

Our research seeks to determine whether or not a victim experienced DV. Because our dataset is binary, either of the two outcomes is possible. Because our data is binary, we used the most commonly used algorithms. The accuracy, recall, precision, and F-measure of our prediction model were evaluated using six algorithms. If the accuracy and

F-measure were raised, the procedure would be more effective. Precision, measured as the overall number of true positives (TPs) in all yes forecasts, is another crucial element. Another critical component of the prediction system is the number of TPs in actual yes data. Next, we will compare the accuracy and F-measure of each algorithm to see which one best fits our model. False negatives (FN) and false positives (FPs) are two more important aspects of our scheme. The model would be improved if the FN and FP rates were lower. If both are lower and there is a decent balance between them, a model is said to be good. Because our system predicts DV, it is vital to have a low FN, because FN means the system incorrectly identifies a DV incidence as no DV, whereas FP means the system incorrectly identifies no DV as DV. To train the model, we used some Python libraries. These include pandas, sci-kit-learn, and others. For model prediction, we used the SVC, KNN, LR, RF, SGD, and LDA algorithms. We used the Python library to import the dataset. One portion of the dataset is for training data, while the other is for test data. Overall, 70% of the data has been divided into train data and 30% has been divided into test data. Now, we will discuss some performance metrics.

Figure 14 User authentication with phone number (see online version for colours)



(a) (b) (c)

Confusion matrices are used to show important predictive statistics including recall, specificity, accuracy, and precision. A binary classification confusion matrix is a two-by-two table that is formed by counting the number of binary classifier results including TP, FP, true negative (TN), and FN. The accuracy (TP + TN/TP + TN + FP + FN) is calculated by dividing the number of correct predictions by the total number of

predictions in the dataset. Precision (TP/TP + FP) is calculated by dividing the number of correct positive predictions by the total number of positive predictions. Sensitivity is calculated by dividing the number of correct positive predictions by the total number of positives. It is often referred to as the recall rate or the genuine positive rate (TP/TP + FN). The F-score is the harmonic mean of precision and recall. A confusion matrix is shown in Figure 13(a). The formulae of the metrics (e.g., accuracy, precision, recall, and F-score) are given in Figure 13(b).

Figure 15 User information stored in firebase (see online version for colours)

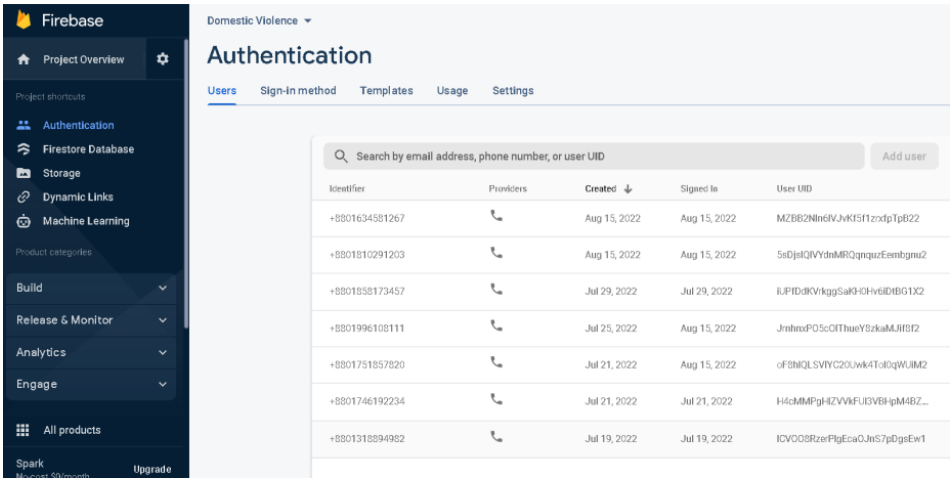


Figure 16 Homepage of DV (see online version for colours)

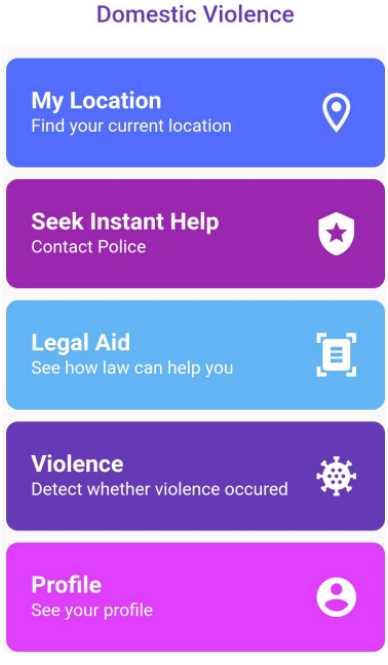
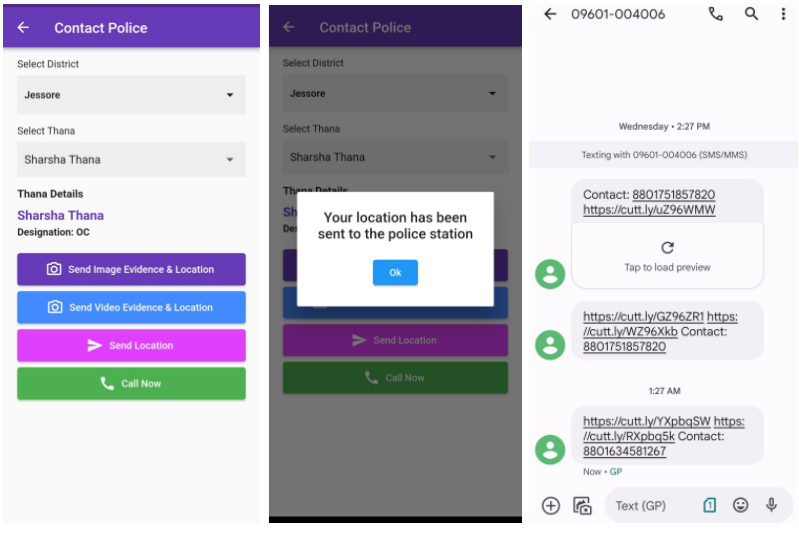
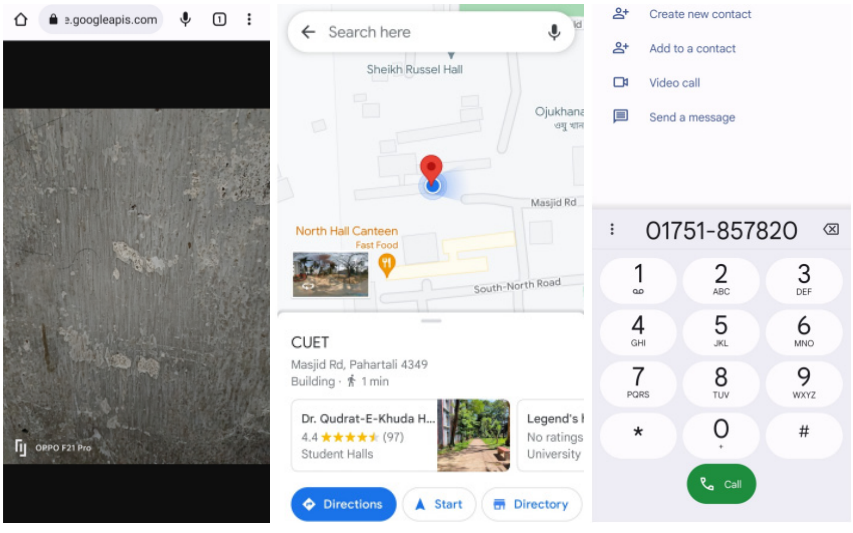


Figure 17 Instant help features for DV victim using contact police (see online version for colours)



(a) (b) (c)

Figure 18 Instant help features for DV victim using shaking, tracking location, and evidence attachment (see online version for colours)

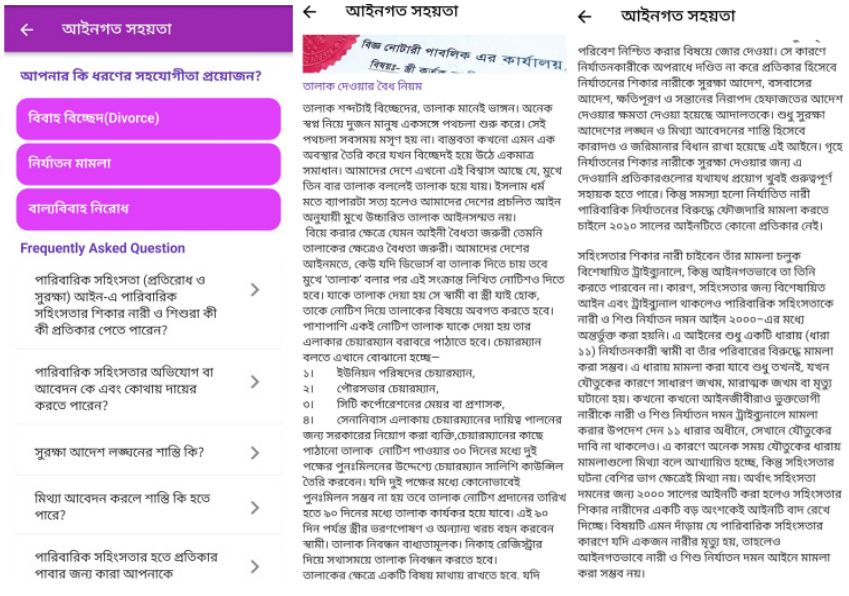


(a) (b) (c)

The confusion matrix of LR, which includes 49 TPs, 39 TNs, four false positives, and four FNs, is shown in Figure 6. We can determine our accuracy from this, which is around 92%. The confusion matrix of SVC, which includes 47 TPs, 40 TNs, three false positives, and six FNs, is shown in Figure 7. We can determine our accuracy from this, which is around 92%. Figure 8 depicts a confusion matrix for K-nearest neighbour

(KNN) with 48 TPs, 42 TNs, one FP, and five FNs. Figure 9 illustrates the confusion matrix of SGD, which has 46 TPs, 40 TNs, three FPs, and seven FNs. This allows us to determine our accuracy, which is nearly 90%. Figure 10 shows the confusion matrix of LDA, which has 47 TPs, 40 TNs, three FPs, and six FNs. This allows us to determine our accuracy, which is around 91%. The confusion matrix of the RF classifier, which includes 52 TPs, 41 TNs, two false positives, and one FN, is shown in Figure 11. Now, we will show a comparative analysis of the performance of different machine learning algorithms. Figure 12 shows the accuracy comparison between algorithms. Table 1 compares different ML algorithms based on accuracy, precision, recall, and F1-score. According to the figure, RF provides the highest accuracy, which is 97%. In terms of precision, KNN gives the best precision of 98%. In the case of a recall, RF gives the best recall, which is 98%. In the case of the F-score, RF gives the best F-measure, which is 97%.

Figure 19 Legal help (some legal rules for police action in Bengali) (see online version for colours)



(a) (b) (c)

4 Android application features and user evaluation

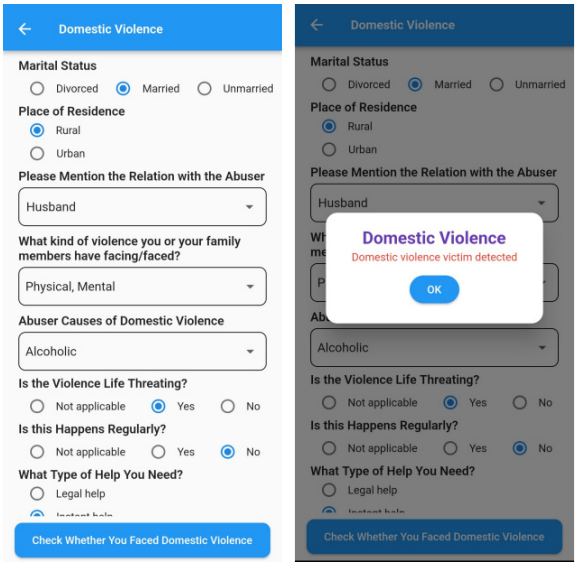
In this section, the features and interface of our Android-based DV application will be shown and described. After successful authentication, the application has different features. We will go through each of them sequentially.

4.1 User sign in and authentication

During sign-in, if the user is new, then he/she will provide a mobile number and a verification code will be sent to that particular number. The user can also add some information along with the emergency contact number of his relatives. After successful

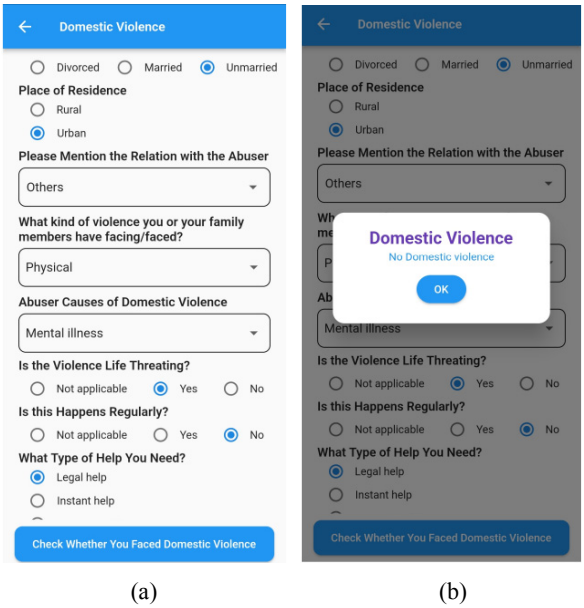
authentication, the user will be moved to the homepage. Here, users will stay signed in until they intentionally sign out of the app. Figure 14 shows the user authentication and sign-in feature using a phone number. But after signing out, if a user again signs in to the system, he/she does not need to provide profile information because user information will be stored in the firebase (see Figure 15) after the first time they sign in.

Figure 20 DV prediction (see online version for colours)



(a) (b)

Figure 21 Evaluation of non-DV prediction (see online version for colours)



(a) (b)

Figure 22 Profile information of user (see online version for colours)

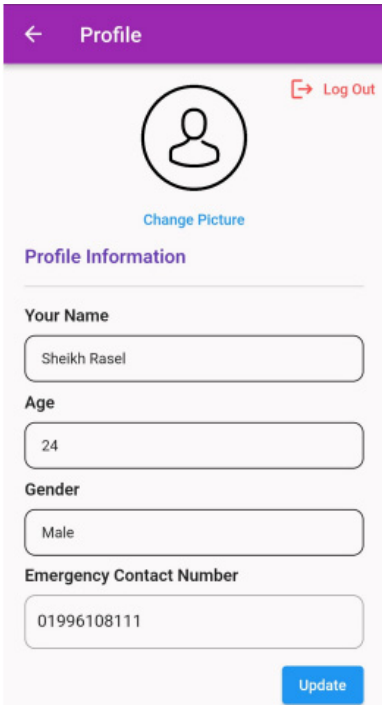
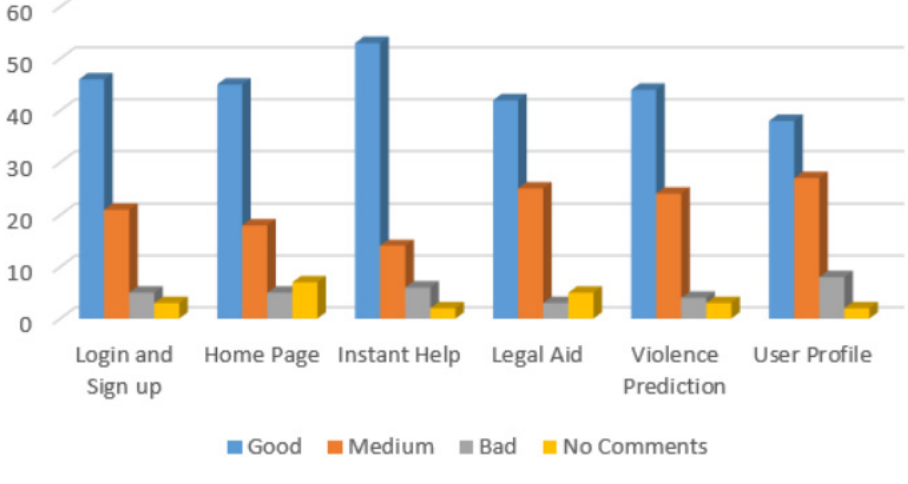


Figure 23 User rating (see online version for colours)



4.2 Homepage and instant help feature

After signing in to the system, the system will automatically direct users to the homepage. Figure 16 shows the homepage. All the features the system will offer can be seen on this page. The user will know their location from the my location widget. There

are also some other widgets such as seek instant help, legal aid, violence, etc. A detailed explanation of each widget will be explained later in the next figures. If a DV victim needs instant help, the user will press the seek instant help widget. Then the user will select a district and thana (i.e., sub-district area name) based on the violence location. After that, the user will send the victim's location along with a video/image to the nearest police station (see Figure 17). The 'call now' widget allows the user to call that specific police station number (see Figure 17). So, this way, victims can get instant help from the nearest police station. Similarly, users can shake their phone and their location can be traced by the police in case of emergency. The users can also attach their evidence by using the proposed mobile application (see Figure 18).

4.3 Legal help feature

Figure 19 shows the legal help feature (in Bengali language). The user can read the associated government rules in terms of each violence type. In a particular case, if the victim's violence is not life-threatening but he/she needs legal help, then the user will get suggestions about different types of DV occurrences. What should a victim do for a divorce? What is the punishment for early marriage? And how should the victim legally proceed against DV? All types of suggestions will be informed by this legal aid widget.

4.4 DV prediction and User profile information

Next, this application widget will predict whether DV occurred or not. For that, it will take input from the user via a questionnaire that we have used in our dataset, and, based on these questions, it will give a result. We have built our model and used six different algorithms for this prediction part. The main reason for using different algorithms is that different algorithms perform better in different situations. Several machine learning algorithms, including LDA, K-neighbours classifier, SVC, RF, and LR, have been compared for accuracy, precision, and recall. From that, we have chosen our best model to convert it into native code. For this, we have used m2cgen. m2cgen is a simple Python library that converts a trained machine-learning model into different programming languages. Figures 20 and 21 show machine learning-based DV and non-violence prediction for users, respectively. This is user profile information that users must fill in when they first sign in to the app. After that, the user can change profile information according to his needs. From this widget, the user can also log out and change their emergency number. Figure 22 shows the profile information of users.

5 User evaluation

Next, Figure 23 highlights the user rating curve for our proposed application features. We have collected a total of 75 user ratings. It can be seen that most of the reviews gave good remarks. Whereas, the medium, bad, and no comments users are secured in second, third, and fourth position respectively. Thus, from the user review results it can be said that the proposed application is suitable and useful for DV prediction assistance for the Bangladeshi people.

6 Conclusions

Domestic abuse is a major societal issue all around the world. But in our country, most people are not aware of the law on DV and they do not want to disclose the situation due to the lack of an effective solution and some other cases. This paper builds a model that can predict whether DV has occurred or not. This paper shows a comparative analysis of the algorithms and selects the best one for the prediction model. This paper develops an application that provides multiple relevant features in one single application. This paper also builds a mobile application that will provide instant help along with legal help to the victim. If people are familiar with this system, then victims will easily get help through our app. Because victims may be unable to call in times of danger because of a variety of factors, our suggested solution includes phone shaking characteristics. In this case, by shaking the phone three times, the victim can send their position to an emergency number. Thus, our system improves overall performance by eliminating the limitations of the existing system. The main limitation of this work is our dataset participants are not huge. Our model could be more efficient if the number of participants was increased. Our developed mobile application could also be improved. In the case of a police station, we have only used the police station number. That is why we have to choose the nearest police station by searching. If we used the longitude and latitude of the police station in the dataset, then the nearest police station should be automatically selected based on user location.

References

Ara, M. (2022) 'The forbidden word – the life during COVID-19', in *COVID-19 Assemblages*, pp.45–51, Taylor and Francis Group, Routledge, India.

Braithwaite, S.R. et al. (2016) 'Women safety device and application-femme', *Indian Journal of Science and Technology*, Vol. 9, No. 10, pp.1–6.

Campbell, A.M. (2016) 'An increasing risk of family violence during the COVID-19 pandemic: strengthening community collaborations to save lives', *Forensic Science International: Reports*, Vol. 2, No. 100089, pp.1–2, Elsevier Public Health Emergency Collection.

Chakraborty, S. et al. (2021) 'NAARI: an intelligent Android app for women safety', in *Applications of Artificial Intelligence in Engineering, Part of Algorithms for Intelligent Systems Book Series (AIS), First Global Conference on Artificial Intelligence and Applications (GCAIA 2020)*, Springer, Jaipur, India, pp.625–637.

Chand, D. et al. (2015) 'A mobile application for women's safety: Wosapp', in *TENCON 2015 IEEE Region 10 Conference*, pp.1–5.

Chowdhury, M. et al. (2023) *Domestic Violence Prediction Dataset*. Available online at: https://www.kaggle.com/datasets/mahfuzulhoqchowdhury/domestic-violence-prediction-dataset (accessed on 27 June 2023).

Garcıa-Moreno, C. et al. (2013) *Global and Regional Estimates of Violence against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-Partner Sexual Violence*, pp.1–50, Department of Reproductive Health and Research, World Health Organization.

Hossain, M.E. et al. (2019) 'Manifesting a mobile application on safety which ascertains women salus in Bangladesh', *International Journal of Electrical and Computer Engineering*, Vol. 9, No. 5, p.4355.

Khandoker, R.R. et al. (2019) 'Lifecraft: an Android based application system for women safety', *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp.1–6.

Khatun, M.T. and Rahman, K.F. (2012) 'Domestic violence against women in Bangladesh: analysis from a socio-legal perspective, *Bangladesh E-Journal of Sociology*, Vol. 9, No. 2, pp.19–29.

Mane, I.A. et al. (2016) 'Stay safe application', *International Research Journal of Engineering and Technology*, Vol. 3, No. 5, pp.2157–2160.

Sharma, A. and Borah, S.B. (2020) 'COVID-19 and domestic violence: an indirect path to social and economic crisis', in *Journal of Family Violence*, Vol. 37, pp.759–765 https://doi.org/10.1007/s10896-020-00188-8.

Sharma, K. and More, A. (2016) 'Advance woman security system based on Android', *International Journal for Innovative Research in Science and Technology*, Vol. 2, No. 12, pp.2349–6010.

Siddique, A.B. et al. (2022) 'A machine learning-based approach to predict university students' depression pattern and mental healthcare assistance scheme using Android application', *International Journal of Data Analysis Techniques and Strategies*, Vol. 14, No. 2, pp.122–139.

Srinivas, D. et al. (2021) 'Android app for women safety', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, Vol. 7, No. 3, pp.378–386.

Sriranjini, R. et al. (2017) 'GPS and GSM based self defense system for women safety', *Journal of Electrical and Electronic Systems*, Vol. 6, No. 2, pp.2332–2796.