Indonesian Journal of Electrical Engineering and Computer Science
Vol. 22, No. 2, May 2021, pp. 1052~1060
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v22.i2.pp1052-1060

Data mining technique to analyse and predict crime using crime categories and arrest records

Most. Rokeya Khatun[1], Safial Islam Ayon[2], Md Rahat Hossain[3], Md. Jaber Alam[4]
[1,2]Department of CSE, Green University of Bangladesh, Dhaka, Bangladesh
[3]Department of ICT, School of Engineering and Technology, Central Queensland University, Australia
[4]Faculty of Engineering, Multimedia University, Malaysia

Article Info

Article history:

Received Oct 12, 2020
Revised Mar 4, 2021
Accepted Mar 21, 2021

Keywords:

Arrest attribute
Crime type
Crimes
Decision tree
K-nearest neighbours
Random forest

ABSTRACT

Generally, crimes influence organisations as it starts occurring frequently in society. Because of having many dimensions of crime data, it is difficult to mine the available information using off the shelf or statistical data analysis tools. Improving this process will aid the police as well as crime protection agencies to solve the crime rate in a faster period. Also, criminals can often be identified based on crime data. Data mining includes strategies for the convergence of machine learning and database frameworks. Using this concept, we can extract previously unknown useful information and their patterns of occurrence from unstructured data. The sole purpose of this paper is to give an idea of how data mining can be utilised by crime investigation agencies to discover relevant precautionary measures from prediction rates. Data sets are analysed by some supervised classification algorithms, namely decision tree, K-nearest neighbours (KNN), and random forest algorithms. Crime forecasting is done for frequently occurring crimes like robbery, assault, and theft. Specifically, the results indicate the superiority of the random forest algorithm in test accuracy.

Corresponding Author:

Md Rahat Hossain
College of Information & Communication Technology
School of Engineering & Technology | Tertiary Education Division
CQUniversity Australia, Building 30/1.12, Bruce Highway
North Rockhampton, Queensland, 4701
Email: m.hossain@cqu.edu.au

1. INTRODUCTION

Day by day, the crime rate is rising considerably. Crime prediction is not an easy process since it is neither systematic nor random [1]. From crime statistics, some crimes like arson, and burglary, have been decreasing however crimes like murder, gang rape, and sex abuse, have been increasing [2]. Although we cannot predict the crime victims, we can predict the most probable crime locations. Predicting the crime will not completely prevent it from occurring, however, it will offer security to some extent in crime sensitive areas. Usually, people do not take into account the way to be secured from sudden occurrences. Both people who are strangers from outside the area and those already living inside the area should know how and what crime incidents are taking place through that particular area [3]. Varieties of crimes happen in different regions at various times. Life and property of general people can be shattered just because of not having a minimum-security mindset.

The mounting crime level has become one of the pressing challenges for society. Police can use crime databases to inspect criminal incidents and associated factors of previous phenomena to implement and

frame crime prevention strategies. The crime data examination may assist to comprehend the behaviour of the trends over time. Therefore from past observations, future values may be predicted. Society's and country's development cannot increase if we cannot safeguard people's life and property from destruction. It is not so difficult with modern technology to design a system by which people can get assistance for their safety. Increased crime levels reduce house prices and neighbourhood satisfaction and increase the desire to move on to another place [4]. To reduce crimes, it is important to identify the reasons behind crimes, predict the occurrence of crimes and prescribe solutions. Due to large volumes of data, it is unrealistic to do a manual analysis [5]. So, it is necessary to have a platform that is capable of applying any algorithm required to do a descriptive, predictive, and prescriptive analysis on a large volume of crime records.

The crime reports have several information categories as follows: crime location, crime types, date or time, case number, block, arrest, district, ward, community, crime classification code, latitude, and longitude. There is also information about the victim and identified or unidentified suspect. Additionally, there is the description or narrative of the crime that is more often than not in the form of text. The crime detectives or investigators utilise free text to trace most of their interpretations that are difficult to be incorporated in the checkboxes. Crime prediction is an up-and-coming move toward criminological research and criminal justice studies. By reviewing past data justice practitioners, criminologists and related researchers can better comprehend the outline of the historical behaviour of crime parameters and also guess future criminal behaviour more accurately.

The main challenges of this paper are to optimise the storage and analysis of a massive amount of incomplete and inconsistent criminal records, to resolve the limitation of obtaining crime records from Law Enforcement agencies, to achieve an acceptable level of prediction accuracy, and so on. Sorting out the crime patterns is another challenging and critical factor. It takes a huge amount of time for crime analysts to spot a pattern, screening through data to come across if a certain crime fits into a recognized pattern.

It is a difficult task to solve crimes and a lot of experience is needed. To model crime detection problems data mining can be utilised. The thought is to seek to input experiences of humans converted into statistical data into computer models using data mining. Crimes are a societal annoyance that cost us severely in numerous ways. The steps we follow to predict the crimes are data collection, pre-processing, feature extraction, application of classifiers, evaluation of the performance of classifiers, comparison of classifiers, and visualisation. Here, datasets are collected from multiple sources from different countries. Pre-processing of the datasets is done by several techniques and methods. Then important features are extracted by different feature extraction methods like Gini impurity, and cross-entropy. For analysing crime data, we have used three classification algorithms: K-nearest neighbours (KNN) algorithm, decision tree algorithm, and random forest algorithm.

A crime can be defined as an action that violates a law which leads to punishment. To understand the nature of crimes, one has to understand their spatio-temporal dimensions, the victim-offender relationship and the history of similar incidents. Data mining is a strong proven method of analysis to find trends and patterns from a cluster of data. Nowadays data mining has been studied as one of the major forefronts to aid criminal investigators to concentrate on the most significant information in the provided crime data. Few of the popular data mining schemes to analyze crime data relates to machine learning techniques and classification methods [6]. Based on the state of art, it has been noted that data mining methods enhance crime data analysis and predict crime pattern efficiently. Brown [7] states that Richmond city in the USA has approximately 1,00,000 criminal records per year. Data mining analysis of such a rich data set could identify complex crime patterns and assist in solving crimes faster and in an efficient way.

N.A. Rahman and W.A. Khader [8] suggested a method for predicting crime in San Francisco using KNN and Naïve Bayes classification. They compared the classifiers for crime prediction and classification. They used validation and cross-validation methods to test the results. There are some limitations in the paper in that they applied each method directly on a training dataset before any pre-processing. Also, the data set was not evaluated for outlier or entry errors. A. Gupta, *et al,* [9] undertook a comparative study of the classification of algorithms on accidents and crime in Denver city of the USA. They just compared the algorithms with blind measure 'accuracy', with no use of precision or recall. A. Awal, *et al,* [10] also analysed the crime of Bangladesh using a linear regression model. Their analysis was limited to the linear relationship of crime data. T. Almanie, *et al,* [11] predicted crime depended on varieties of crime and utilising spatial and temporary hotspots of crime occurrence. They utilised the Decision tree classifier and Naive Bayes classifier to guess potential crime types. P.Yerpude and V.Gudur [12] proposed predictive modelling of a crime dataset using Datamining. They used the Decision tree, Naïve Bayes, and regression model for predicting the properties accountable for causing crime in an area. R. Kiani, *et al,* [13] analysed and predicted crime types by classification and clustering.

The structure of the rest of the paper is as follows. Section 2 is the proposed methodology. Results and discussions of the suggested scheme are elaborated in Section 3. In the end, the paper is concluded in Section 4.

2. RESEARCH METHOD

The following section discusses the introduced methodology for predicting two criteria: crime types and predicted arrest area. Firstly, crime types are measured based on the area, time, and FBI code. Secondly, the predicted arrest area is measured based on crime type, environment, and FBI code. Moreover, the premise code [11] is measured based on area, district, victim age, and criminal code. All the predictions and findings are done with the KNN, decision tree, and random forest algorithms. In this paper, a method for predicted crime analysis is done in five steps as follows: (1) Data Collection, (2) Pre-processing, (3) Feature Extraction, (4) Apply Classification Algorithm, and (5) Evaluate performance. The following subsections explain all the steps in detail.

2.1. Data collection

Several data sets from different recognized literature are used in this paper to investigate the prediction of crimes. One dataset has been collected from police recorded crime in Northern Ireland [14]. The size of the dataset is 84,800*25. Furthermore, a real-time crime dataset [15] of Chicago from 2012 to 2017(0.5 million*22) has been used. Also, the crime dataset of Chicago [16] from 2001 to the present (1 million*22) have been used. Another real crime data set of summer 2014 in San Francisco has been used in this paper. The size of the dataset is 28,994*13 [15]. Also, a crime dataset of Los Angeles from 2010 to the present has been extracted from their Open Data Portal [17]. The size of this data set is 74767*6. Table 1 shows the general attributes of these datasets.

Table 1. General attributes of data

Attribute	Description
ID	ID Unique identifier
Case no	Unique crime incident id
Date	Date when the incident occurred
Block	redacted address of the incident place
IUCR	Uniform crime reporting code
Primary type	Primary description of IUCR
Description	Secondary description of IUCR
Local desc	Location of incident place
Arrest	Indicates whether the arrest was made
Domestic	tells whether incident domestic rel.
Beat	Smallest police geographic area
District	Police district of incident occurred
Ward	Ward where the incident occurred
Community	The community area of the incident
FBI code	Indicate Crime classification code
X-coordinate	X co-ordinate of the incident location
Y-coordinate	Y co-ordinate of the incident location
Year	Year the incident occurred
Updated on	The date the record was last updated
Latitude	Latitude of the incident location
Longitude	Longitude of the incident location
Location	Location of the incident in map form

2.2. Pre-processing

The standard of the dataset sometimes affects the outcome of any classification problems. The results are affected by missing values. Hence, it is needed to manage the missing parameters of the dataset first. Misplaced values can be controlled in several ways, such as overlook the misplaced values, change the misplaced values with any numeric value, exchange the misplaced values with the maximum value appearing for that trait or change the value with the mean value for that characteristic. In this paper, the misplaced values of numeric data are managed by replacing the values with the mean value of that characteristic. The loss of the data can be negated by this method which yields better results compared to the removal of rows and columns. For the Boolean type data, the number of 1's and 0's are counted first, then missing values are replaced by the highest counTable 1's or 0's.

2.3. Feature extraction

Feature selection is known to be the procedure of decreasing the inputs for analysis and processing, or of sorting out the most significant data. Moreover, feature selection is referred to as attribute selection, variable selection, or variable subset selection and is the procedure of selecting a subset of related properties to utilise in the construction of the model. Feature extraction includes decreasing the quantity of the resource needed to portray a large dataset. Generally, it needs a large quantity of computational power and power to analyse a large number of variables. It may also cause a classification algorithm to overfit the training samples and generalise poorly to new examples. In this paper, we use the decision tree algorithm to extract the features. It inherently estimates the suitability of features for the separation of objects representing different classes using Gini impurity.

2.4. Apply classification algorithm

In this paper, three classification algorithms are utilized to categorize the crime, namely (1) K-Nearest Neighbours algorithm, (2) Decision tree algorithm, and (3) Random forest algorithm.

(1) K-Nearest Neighbours

The K-Nearest neighbours (KNN) algorithm can be utilised for both regression predictive and classification problems. This algorithm suits all considerable parameters [18]. It is usually utilised for its lower calculation time and ease of interpretation. There is no need to make additional assumptions. It works easily on multi-class problems [19]. The KNN algorithm predicts that alike things subsist near each other. We can also say, alike things are nearby [18]. For distance calculation, it uses Euclidean, Manhattan, and Minkowski distance functions. The process is to run the KNN algorithm several times with diverse K-values and select the value of K that decreases the error number. Using a cross-validation method and by measuring accuracy or validation error, we get the optimal value of K.

(2) Decision tree algorithm

A decision tree is a tree structure where an internal node depicts a property, the branch refers to a decision rule, and each leaf node refers to the outcome. The topmost node learns to partition depending on the attribute value. It partitions the tree in a recursive manner called recursive partitioning. This algorithm aids in decision making [20].

The decision tree algorithm works in the following way:

a) Select the best feature utilising the Gini index or cross-entropy to divide the records.
b) Makes that feature decision node and splits the set of data into smaller subsets.
c) Begins building of tree by repeating this procedure again and again for every child until all the tuples dedicated to the same feature value or no more leftover features or no more occurrences. For subset selection, it uses information gain and Gini impurity [20]. We found the information gain using the following equation:

$$\mathrm{I} = \sum_{i=1}^{n} p(i) * \log\left(p(i)\right) \tag{1}$$

The equation of Gini impurity is:

$$G = 1 - \sum_{i=1}^{n} \left(p(i)\right)^2 \tag{2}$$

Here, I= information gain, G= Gini impurity, *n* =number of features, *i* =feature and *p*=probability of *i*.
p(i) is the probability of randomly picking an element of class *i* i.e., the proportion of the dataset made up of class *i*.

(3) Random forest algorithm

Random forest is a supervised learning algorithm. It builds multiple decision trees (makes it somehow random), merges them, builds a forest, and obtains a stable and more accurate prediction. As an alternative to probing for the most crucial characteristic while dividing a node, it seeks the best attribute among a random subset of features [21]. The higher amount of trees in the forest gives more accurate results. The optimal amount of trees depends on the number of predictors. This algorithm selects the attribute's subset randomly. Feature's importance is determined by the reduction of Gini impurity or cross impurity.

$$Subset = \sqrt{no.of feature} \tag{3}$$

2.5. Evaluate performance

The performance of a model is evaluated using the evidence of experimental actual events. During the training of any model, a labelled set of data that involves the real values to be guessed is taken into consideration. This suggested the theories of a confusion matrix. There are four classification performance indices in the confusion matrix. Those are (1) true positives (TP), (2) true negatives (TN), (3) false positives (FP), and (4) false negatives (FN). When the data of a dataset is imbalanced, then accuracy does not give the best result. In this case, the F1 score gives an accurate result. In this paper, most of the dataset is imbalanced so we take the F1 score to check our system accuracy. The following qualities are measured to estimate the performance of the system:

I. Precision = TP / (TP + FP)
II. Recall = TP / (TP + FN)
III. F1score = (2 * Precision * Recall) / (Precision + Recall)
IV. MCC = ((TP * TN) – (FP * FN)) / √ ((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN))

3. RESULTS AND DISCUSSION

Different classification methods such as decision tree, K-Nearest Neighbour (KNN), and random forest algorithm have been utilised to forecast different characteristics of crime data. In this paper, two types of attributes are used to predict crime. One is crime type and the other is the number of arrests. The following section discusses the two types of attributes in detail.

3.1. Prediction of crime types

We have predicted crime types using KNN, decision tree, and random forest algorithms. Different crime types are present such as assault, burglary, theft, robbery, weapons violation, vehicle theft, and public peace violation. We have worked with 9 attributes and 632 instances. A 10-fold cross-validation method has been utilised for result measurement. The benefit of the cross-validation technique is to examine for both testing and training. Besides, every observation was utilised exactly once for the test set. The ratio of splitting the set of data in all the cases was 25% for testing and 75% for training. With the support of the confusion matrix precision, recall, F1 score, and MCC was measured. Results are presented in Table 2, Table 3, and Table 4 for KNN, decision tree, and random forest respectively.

From Table 2, Table 3, and Table 4 we see that most of the cases have precision values between 0.93 to 1.00, where the best value of precision is 1.00. It is also true for recall; most of the recall values are between 0.94 to 1.00. The best value of recall is also 1.00. The F1 score varies from 0.92 to 1.00, where 1.00 is the best value for an F1 score. MCC score varies from 0.92 to 0.99. So, we can say that our experimental results are very close to the best values of precision, recall, F1 score, and MCC.

Table 2. Measurement for crime data using KNN

Class	Precision	Recall	F1 score	MCC
Battery	0.85	1.00	0.92	0.91
Theft	0.91	1.00	0.95	0.93
Robbery	1.00	0.94	0.97	0.92
Vehicle theft	0.93	1.00	0.96	0.96
Assault	0.96	0.89	0.92	0.94

Table 3. Measurement for crime data using decision tree

Class	Precision	Recall	F1 score	MCC
Battery	1.00	1.00	1.00	0.98
Theft	0.99	1.00	0.99	0.96
Robbery	1.00	1.00	1.00	0.97
Vehicle theft	1.00	1.00	1.00	0.98
Assault	0.98	1.00	0.98	0.99

Table 4. Measurement for crime data using random forest algorithm

Class	Precision	Recall	F1 score	MCC
Battery	0.96	0.96	0.96	0.96
Theft	0.86	0.92	0.96	0.97
Robbery	1.00	1.00	1.00	0.99
Vehicle theft	1.00	0.94	0.97	0.94
Assault	0.92	0.96	0.94	0.96

For the KNN algorithm, we use K (number of nearest neighbours) = 7. In the decision tree algorithm, a confidence factor (CF) has been used. CF is used for pruning. Larger CF gives more specific rules to predict the target class. In this paper, we applied CF = 0.45. For the random forest algorithm, batch size and number of iterations have been used. For this paper, both the batch size and the number of iterations are 100. For all the cases, we used different types of numbers in these parameters (K, CF, or batch size). But we chose the number with which we get the best result.

3.2. Prediction of the arrest record

We have predicted the arrest attribute which means whether or not criminals will be arrested using KNN, Decision tree, and Random forest algorithms. Arrest and not arrest are the two target classes here. For this prediction, we used the Chicago crime dataset from 2001 to the present [16]. Primary type, local description, beat, district, domestic, ward, community, and FBI code are the input data. For analysis, the 10-fold cross-validation technique has been utilised, as a result, eliminating the chance of overfitting the data. Performances have been measured by accuracy, precision, recall, F1 score, and MCC. The ratio of dividing the sets of data in all the cases was 25% for testing and 75% for training. Table 5, Table 6, and Table 7 show the precision, recall, F1 score, and MCC for KNN, Decision Tree, and Random forest respectively.

From Table 5, Table 6, and Table 7, we see that most of the cases have precision values between 0.79 to 0.89, where the best value of precision is 1.00. It is also true for recall; most of the recall values are between 0.44 to 0.58. The best value of recall is also 1.00. The F1 score varies from 0.56 to 0.64, where 1.00 is the best value for the F1 score. For MCC, the score is between 0.55 to 0.61. Here most of the cases of our experimental results are not very close to the best values of precision, recall, F1 score, and MCC.

In the KNN algorithm, we applied K (number of nearest neighbours) = 7. Here the Minkowski distance calculation has been used for measuring distances. For the decision tree algorithm, a confidence factor (CF) has been used. The confidence factor represents a threshold of allowed inherent error in data while pruning the decision tree. For attribute selection, we use the Gini index criterion. In this paper, we applied CF = 0.45 for all the experiments. For the random forest algorithm, batch size and number of iterations have been used. For this experiment, 100 is fixed for both the batch size (how many samples were taken at a time) and the number of iterations. As was done with crime types, we also used different types of numbers in these parameters (K, CF, or batch size). But we chose the number with which we get the best result.

Table 5. Measurement for crime data using KNN

Class	Precision	Recall	F1 score	MCC
True	0.79	0.44	0.56	0.55
False	0.86	0.97	0.91	0.92

Table 6. Measurement for crime data using decision tree

Class	Precision	Recall	F1 score	MCC
True	0.89	0.43	0.58	0.56
False	0.86	0.99	0.92	0.93

Table 7. Measurement for crime data using random forest algorithm

Class	Precision	Recall	F1 score	MCC
True	0.73	0.58	0.64	0.61
False	0.89	0.94	0.91	0.92

3.3. Result analysis

Figure 1 and Figure 2 show the experimental analysis undertaken for this paper. In Figure 1, the F1 score of different classes of the crime type is shown. From this result, we clearly show that, in the battery class, the decision tree works well. Also, for theft, robbery, vehicle theft, and assault classes, decision tree work better than KNN and random forest. KNN shows the worst results among the three algorithms.

In Figure 2 we show that, for the measurement of the F1 score for the arrest attribute, random forest shows better results for the TRUE class among KNN, decision tree, and random forest algorithms. For the FALSE class, all three algorithms show a similar result.

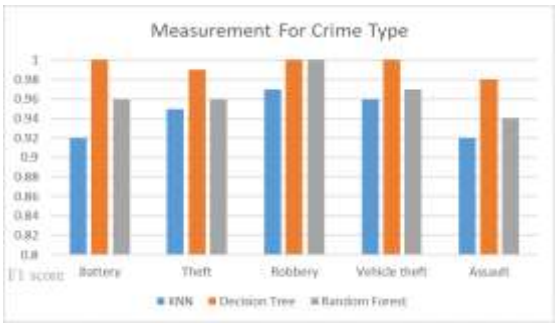

Figure 1. Measurement for crime data using different algorithms

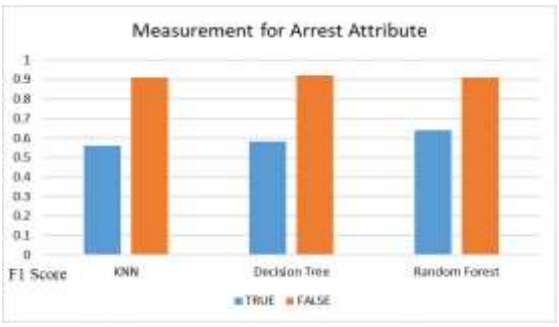

Figure 2. Measurement for arrest attributes using different algorithms

3.4. Comparative analysis

In this section, we compare our experimental results with different existing methods. P. Yerpude and V. Gudur [12] proposed a method where they predict the crime dataset using 4 different data-mining schemes of the decision tree, random forest, naïve bayes, and linear regression. Their F1 scores are 81.22%, 86.54%, 79.58%, and 82.3% respectively for those techniques. H.B.F. David and A. Suruliandi [22] worked on the analysis and prediction of crimes utilising data-mining schemes. They used SIIMCO, CrimeNet Explorer, and Log analysis methods to predict the crime and got 59%, 38%, and 52% accuracy respectively [23]-[25]. Table 8 shows the comparative results of the various studies of the prediction of crime. In Table 8, we take the best result among the different classes we analysed. For all the methods, we indicate the Robbery class results. Analysing all the results, it is observed that our result is better than the other methods. The reasons behind this better performance are the preprocessing of the dataset, the machine learning techniques we use and the parameter settings for those machine learning techniques.

Table 8. Comparative study of prediction codes

Author	Method	F1 score
P. Yerpude and V. Gudur [12]	Decision tree	81.22%
	Random forest	86.54%,
	Naïve Bayes	79.58%
	Linear Regression	82.3%
H.B.F. David and A. Suruliandi [22]	SIMCO	59%
	CrimeNet Explorer	38%,
	Log analysis	52%
In this paper	Decision Tree	100%
	Random Forest	100%
	K-Nearest Neighbour	97%

4. CONCLUSION

In this paper, we have tested the F1 score and various measurements like precision, recall, and MCC of classification and prediction depended on diverse train and test sets of data. Crime patterns change over time. So, we have considered only some limiting factors; for this reason, the full accuracy of the system cannot be attained. To achieve better results, we have to come across more crime features instead of fixing some characteristics. To date, the models are trained to utilise specific characteristics, but more factors are included to improve accuracy. We have applied different data pre-processing techniques in this paper, namely K-Nearest Neighbour, Decision tree, and Random forest. Through the results in this paper, it is seen that the Decision tree gives slightly better performance than the Random forest algorithm for prediction and classification of different crime characteristics, however, the Decision tree creates an overfitting problem. So, we consider the Random forest algorithm as a better model than the KNN and Decision tree algorithms. This paper may help law enforcement agencies to discover precautionary measures for different crimes.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the authority of the Green University of Bangladesh (GUB), Dhaka, Bangladesh for its contribution in sponsoring this article processing charge.

REFERENCES

[1] D. R. Ruiz and A. Sawant, "Quantitative Analysis of Crime Incidents in Chicago Using Data Analysis Techniques", *CMC*, vol. 59, no.2, pp. 389-396, 2019, doi:10.32604/cmc.2019.06433.

[2] H. Saif and H. Dossari, "Detecting and classifying crimes from Arabic Twitter posts using text mining techniques", *International Journal of Advanced Computer Science and Applications (IJACSA*), vol. 9, no. 10, 2018, doi: 10.14569/IJACSA.2018.091046.

[3] A. Bogomolov, B. Lepri, and J. Staiano, "Towards crime prediction from demographics and mobile data," *16th international conference on multimodal interaction*, pp. 427-434, Sept. 2014. Arxiv:1409.2983v1.

[4] C. Kadar, R. Maculan and S. Feuerriegel, "An imbalance aware hyper ensemble for spatiotemporal crime prediction," *International conference on multimodal interaction*, pp. 421-424, Feb. 2019. Arxiv:1902.03237v1.

[5] R. C. Arellano, "Spatial prediction of annual burglaries in Los Angeles", *UCLA Electronic Theses and Dissertations*, University of Los Angeles, 2019.

[6] C. Kadar, J. Iria, and I. Pletikosa, "Exploring derived features for crime prediction in New York City", *ACM* ISBN 978-1-4503-2138-9, Feb2018, doi: 10.1145/1235.

[7] R. Parvez, T. Mosharraf, and M. E. Ali, "A novel approach to identify spatiotemporal crime pattern in Dhaka city," *ACM* ISBN 978-1-4503-4306-0/16/06, June 2016, doi: 10.1145/2909609.2909624.

[8] N. A. Rahman and W. A. Khader, "KNN classifier and Naive Bayes Classifier for crime prediction in SAN FRANCISCO context," *International Journal of Database Management System (IJDMS)*, vol.9, no.4, August 2017, doi: 10.5121/ijdms.2017.9401.

[9] A. Gupta, A. Mohammad, A. Syed, and M. N. Halgamuge, "Classification Algorithms using Data Mining: Crime and Accident in Denver City the USA*", International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 7, no. 7, 2016, doi: 10.14569/IJACSA.2016.070753.

[10] A. Awal, J. Rabbi, and I. Rana, "Using Data Mining Technique to Analyze Crime of Bangladesh", *International Journal of Computer Science and Network(IJCSN)*, vol. 6, no. 4, ISSN:22175420 August 2017.

[11] T. Almanie, R. Mirza, and E. Lor, "Crime Prediction based on crime types and using spatial and temporal criminal hotspots*", International Journal of Data mining & Knowledge Management Process (IJDKP*), vol. 5, no. 4, July 2015, doi: 10.5121/ijdkp.2015.5401.

[12] P. Yerpude and V. Gudur, "Predictive modelling of crime dataset using Data mining", *International Journal of Data mining & Knowledge Management Process (IJDKP*), vol. 7, no. 4, July 2017, doi: 10.5121/ijdkp.2017.7404.

[13] R. Kiani, S. Mahdavi, and A. Keshavarzi, "Analysis and prediction of crimes by clustering and classification", *International Journal of Advanced Research in Artificial Intelligence (IJARAI*), vol. 4, no. 8, 2015.

[14] PSNI Statistics Branch, Crime-Bulletin [Online]. Available: https://www.psni.police.uk/globalassets/inside-thepsniour-statistics/police-recorded-crime-statistics [Accessed Jan.13, 2019].

[15] Kaggle, Chicago Crime Data analysis [online]. Available : https://www.kaggle.com/gnem/Chicago-crime-rate-analysis [Accessed Mar.22,2019].

[16] GitHub, Crimes-in-Chicago/crimes in Chicago 2015- 2016 [online]. Available: https://github.com/Drachna/Crimes-inchicago/blob/master/crimes%20in%20chicago%202015-2016.ipynb [Accessed Mar.20,2019].

[17] GitHub, Crime-prediction [online]. Available : https://github.com/shuck0407/crime-prediction [Accessed April.25, 2019].

[18] Chicago Data Portal, Crimes-2001 to present city of Chicago data portal [online]. Available: https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data. [Accessed Mar. 22, 2019].

[19] GitHub, Crime analytics in San Francisco [online]. Available : https :// github .com/ cmenguy/ crimeanalytics/ blob/ master/ crime analytics .ipynb.[Accessed July. 18, 2019].

[20] Los Angeles Data Portal, Crimes-2010 to present city of Los Angeles data portal [online]. Available: https:// data. city of Los Angeles. org/ Public-Safety/ Crimes-2010-to-present/ ijzp-q8t2/data .[Accessed Mar.22, 2019]

[21] Analytics Vidhya, K-nearest neighbors algorithm [Online] Available: https :// www . analyticsvidhya . com/ blog/ 2017/ 09/ KNN-explain ed/. [Accessed Feb.27, 2019]

[22] H. B. F. David and A. Suruliandi, "Survey On Crime Analysis And Prediction Using Data Mining Techniques," *Ictact Journal On Soft Computing*, vol. 07, no. 03, April 2017, doi: 10.21917/ijsc.2017.0202.

[23] S. Prabakaran and S Mitra, "Survey of Analysis of Crime Detection Techniques using Data Mining and Machine Learning," *Journal of Physics: Conference Series*, 2018.

[24] D. K Tayal and *et al.*, "Crime Detection and Criminal Identification in India using Data Mining Techniques", *AI and society*, 2015.

[25] M Farsi *et al.*, "Crime Data Mining Threat Analysis and Prediction", *Cyber Criminology*, pp. 183-202, 2018, Springer.

BIOGRAPHIES OF AUTHORS

Most. Rokeya Khatun has completed her B.Sc. degree in the Department of Computer Science and Engineering (CSE) from Rajshahi University of Engineering and Technology (RUET), Bangladesh in 2019. She joined as a lecturer of CSE, the Green University of Bangladesh in 2020. Her research interests include data mining and machine learning.

Safial Islam Ayon has completed his B.Sc. degree from the Department of Computer Science and Engineering (CSE) at the Khulna University of Engineering and Technology (KUET), Bangladesh in 2019. He is currently worked as a lecturer of the CSE department at the Green University of Bangladesh, Dhaka. His research interests focus on deep neural networks, machine learning, and swarm intelligence.

Dr Rahat Hossain is a dedicated and articulate lecturer at CQ University, Australia with extensive teaching and learning experience across different areas of Information and Communication Technology (ICT). He moved to Rockhampton, Australia from Bangladesh in March 2010 and completed his PhD in Computational Intelligence in 2013 at CQ University. Before his current academic position, Rahat acquired more than seven years of learning and teaching experience in the Department of Computer Science and Information Technology at the Islamic University of Technology (IUT), Bangladesh.

Md. Jaber Alam has completed his Masters of Engineering Science (M.Eng.Sc.) degree from Multimedia University, Malaysia in 2019. He has done his bachelor's (B.Sc.) degree in Electronics and Telecommunications Engineering in 2016 from IIUC, Bangladesh. He has been working in the field of wireless communication for the last few years. Moreover, his research interest also lies in Electrical, Electronics, and Computer Science topics.