IJCSN - International Journal of Computer Science and Network, Volume 6, Issue 4, August 2017
ISSN (Online): 2277-5420
www.IJCSN.org
Impact Factor: 1.5

Using Data Mining Technique to Analyze Crime of Bangladesh

[1]Md. Abdul Awal; [2]Jakaria Rabbi; [3]Imran Rana

[1]Department of Computer Science and Engineering,
Khulna University of Engineering & Technology (KUET), Khulna 9203, Bangladesh.

Abstract — Crime is classically unforeseeable and a social nuisance. It is not necessarily random, but neither does it take place consistently in space or time. In the recent past, there has been an enormous increase in the rate of crime, hence the significance of task to predict, prevent or solve the crimes. In this case, machine learning and data mining techniques can play an important role to discover future trends and patterns of crime. In this paper, linear regression model is used to forecast future crime trends of Bangladesh. The real dataset of crime is collected from the website of Bangladesh police. The dataset contains aggregated counts of different types of crime. Then the linear regression model is trained on this dataset. After training the model, crime forecasting is done for dacoit, robbery, murder, women & child repression, kidnapping, burglary and theft for different region of Bangladesh.

Keywords — *Data Mining, Crime Forecasting, Linear Regression, Gradient Descent.*

1. Introduction

Crime , in all its facets, is a well-known social problem. Crime is an action that is regarded deleterious to the public welfare and is legally forbidden. It is also an offense against the society which is punishable by laws. It thwarts not only the quality way of life but also the socio economic development of a country. For this reason, it is ineluctable to analyze crime data to inform the police department and law enforcement agencies about specific and general trends and patterns of crime regularly.

Due to the quick improvement in computerization and digitalization techniques, a vast amount of data are now available in science, business, medicine, education, marketing, banking, airlines and many other areas. Such data may provide a great resource for knowledge discovery and decision support system. Fast computing systems provide an opportunity to peruse this huge amount of data in various ways which were not convenient in earlier time. In order to understand, analyze and gradually make sense of this vast amount of data, a multidisciplinary approach, data mining is introduced to accept the challenge. Data mining techniques can play an important role to analyze this data and discover knowledge from them [1]. One of this data are crime data that highly affect the society. Therefore, to focus on the scientific study of crime and criminal behavior this crime data should be analyzed. The high volume of crime dataset and also the complexity of relationships between this kinds of data have made criminology an appropriate field for applying data mining techniques. Identifying crime characteristics is the first step for developing further analysis. The main objective of crime data analysis is to identify the crime characteristics and their relationships with the criminals [2,23]. The volume of dataset and also the complexity of relationships between this kinds of data make the analysis procedure difficult [3,24]. At present manual investigation of crime data by analysts and law enforcement agencies is limited. In order to make easy the analysis procedure, data mining techniques have become the important tools. Data mining techniques accelerate crime analytics, provide better analysis and generate effective solutions to save important resources and time [4]. The tasks of data mining are classification, clustering, evaluation, association, prediction and trend analysis [5]. The knowledge that is gained from data mining approaches is a very useful tool which may assist the police department and law enforcement agencies to predict, prevent or solve future crime of Bangladesh.

Today, a large number of crimes are causing a lot of problems in many different countries. In fact, scientists are spending time studying crime and criminal behaviors in order to understand the characteristics of crime and to discover crime patterns. It is known that criminals follow repetitive behavioral patterns, so analyzing their behaviors can help to capture relations among events from past crimes [6].

Crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis. There are several characteristics that have great influence on crime data analysis. They includes: different races in a society, poverty, income groups, age groups, family structure (single, divorced, married), level of education, the locality where people live, number of police officers allocated to a locality, number of employed and unemployed people among others [6,7,22].

The Uniform Crime Reporting (UCR) program of Federal Bureau of Investigation (FBI) categorizes crime into four categories: murder, forcible rape, robbery, and aggravated assault [8]. In this paper, different types of crime such as dacoit, robbery, murder, women & child repression, kidnapping, burglary and theft are considered. For the purpose of this research, the dataset is collected from the website of Bangladesh police [1,2]. The dataset contains accumulated counts of different types of crime categorized by the police department and are considered to be forecasted. In this paper, the linear regression model is trained on this dataset. After training the model, basically crime forecasting is done for different types of crime for metropolitan and divisional region of Bangladesh.

The rest of the paper is organized as follows: section 2 describes the existing works, section 3 describes the research methodology and section 4 presents the results of the experimental analysis. Finally, section 5 concludes the paper.

2. Existing Works

Data mining is a powerful technique with great potential to help criminal investigators focus on the most important information in their crime data. Some examples of data mining techniques usage to analyze crime data are classification and machine learning algorithms. Based on existing research, it has been observed that data mining techniques assist the procedure of crime patterns detection. Yu et al. [9], employ an ensemble of data mining classification techniques to perform the crime forecasting. A variety of classification methods such as: One Nearest Neighbor (1NN), Decision Tree (J48), Support Vector Machine (SVM), Neural Network (Neural) with 2 layer network, and Naïve Bayesian (Bayes) were used to predict the crime "hotspot". Finally the best forecasting approach was proposed to achieve the most stable outcomes.

In [4], fuzzy association rule mining approach was used for community crime pattern discovery. The discovered rules were presented and discussed at regional and national levels for crime pattern investigation. Levine [2], developed a spatial statistical program for the analysis of crime incidents or other point locations. It was designed to operate with large crime –incident dataset collected by metropolitan police departments. In [11], different hotspot mapping techniques such as: point mapping, thematic mapping of geographic areas (e.g. Census areas), spatial ellipses, grid thematic mapping and kernel density estimation (KDE) were used for identifying hotspot of crimes. Finally, hotspot mapping accuracy was compared in relation to the mapping technique that was used to identify concentrations of crime events and by crime type – four crime types were compared (burglary, street crime, theft from vehicles and theft of vehicles).

In [1], a comparative study was conducted between some free available data mining and knowledge discovery tools and software packages. The result showed that, the performance of the tools for the classification task is affected by the kind of dataset used and by the way the classification algorithms were implemented within the toolkits. In [3], a clustering based model was used to anticipate crime trends. Performance of clustering technique was analyzed in forming accurate clusters, speed of creating clusters, efficiency in identifying crime trend, identifying crime zones, crime density of a state and efficiency of a state in controlling crime rate. In [6], an experiment was conducted to obtain better supervised classification learning algorithms (Naïve Bayesian (0.898), k Nearest Neighbor (k-NN) (0.895) and Neural Networks (0.892), Decision Tree (J48) (0.727), and Support Vector Machine (SVM) (0.678)) to predict crime status. Two different feature selection methods were tested on real dataset for prediction. Chi-square feature selection technique was used to improve the performance of mining results.

Malathi et al. [7], applied anomalies detection and clustering algorithms to predict the crime patterns and speed up the process of solving crime. MV algorithm, DBScan and PAM outlier detection algorithm were used to assist in the process of filling the missing value and investigation of crime patterns. In [8], a comparative study was conducted between the violent crime patterns from the communities and crime unnormalized dataset and actual crime statistical data for the state of Mississippi. Linear regression, additive regression, and decision stump algorithms were implemented on the communities and crime dataset. The linear regression algorithm performed the best among the three selected algorithms. By considering the geographical approach, Nath et al. [10], applied a combination of k-means and weighting algorithm which show regional crimes on a map and cluster crimes according to their types.

Bruin et al. [12], created digital profiles for all offenders by extracting the factors: crime nature, frequency, duration and severity from the crime database. Comparison of all individuals were performed on these profiles by a new distance measure and cluster them accordingly. Thongtae et al. [13], delivered a comprehensive survey of efficient and effective methods on data mining for crime data analysis. They explored the illegal activities of professional identity fraudsters based on knowledge discovered from their own histories. Bagui et al. [14], used WEKA for mining association rules, developing a decision tree, and clustering to retrieve meaningful information about crime from a U.S. state database.

Abraham et al. [15], employed the association rule data mining technique to generate profiles from log data to

realize the criminal's behavior. In [16], a density tracing based approach was used to discover patterns among datasets by finding those crime and spatial features that exhibit similar spatial distributions by measuring the dissimilarity of their density traces.

3. Methodology

Data mining is a part of the interdisciplinary field of knowledge discovery in database [17]. Data mining consists of collecting raw data and, extracting information that can be applied to make predictions in many real world situations such as the stock market or tracking spending habits at the local Wal-Mart [8]. It is the application of techniques that are used to conduct productive analytics. Successful data mining methodology depends heavily on the particular choice of techniques used by analyst. In this paper, data mining is used to forecast future trends in crime of Bangladesh. The methodology is divided into two parts, these are: 1) dataset description and 2) model selection

3.1 Dataset Description:

To conduct this research, the real dataset of crime of previous years is collected from the website of Bangladesh police. The dataset contains aggregated counts of different types of crime categorized by the police department of Bangladesh. The dataset is divided into two groups according to the region of Bangladesh: metropolitan region data and divisional region data. The dataset consists of 840 instances or crimes that had been collected from across the country. The dataset contains three predictive features and one goal feature. The three predictive features are: region, month and year. The goal feature is the predicted value of different types of crime. The data in each instance belong to different regions of Bangladesh. The regions are represented in the form of a number, every number represents its perspective region of Bangladesh.

3.2 Model Selection:

In order to quantitatively forecast the status of crime, different data mining techniques can be used. The associated task for the dataset used in this paper is regression. Therefore, liner regression model is used to forecast the status of crime. The linear regression model is simple and provides enough description of how the input affects the output. It predicts a variable Y (target variable) as a liner function of another variable X (input variable/features), given m training examples of the form $(x_1,y_1), (x_2,y_2), \ldots, (x_n,y_n)$, where $x_i \varepsilon X$ and $y_i \varepsilon Y$.
The form of hypothesis of linear regression can be expressed as

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$
$$h_\theta(x) = \theta^T x \quad (1)$$

Where $\theta_0, \theta_1, \theta_2, \ldots, \theta_n$ are regression parameters. To find the regression parameters the following cost function is used:

$$J(\theta_0, \theta_1, \theta_2, \ldots, \theta_n) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 \quad (2)$$

The objective of linear regression is to minimize the cost function $J(\theta_0, \theta_1, \theta_2, \ldots, \theta_n)$ by finding $\theta_0, \theta_1, \theta_2, \ldots, \theta_n$, so that $h_\theta(x)$ is close to y for every training examples (x,y). One way to minimize the cost function is to use the batch gradient descent algorithm. The algorithm is below:

Repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) x_j^{(i)} \text{ (for every j)}$$

}

Where α is learning rate. The parameters $\theta_0, \theta_1, \theta_2, \ldots, \theta_n$ must be updated simultaneously. With each step of gradient descent, the parameters $\theta_0, \theta_1, \theta_2, \ldots, \theta_n$ come closer to the optimal values that will achieve the lowest cost $J(\theta_0, \theta_1, \theta_2, \ldots, \theta_n)$. To speed up the minimization process using gradient descent, every feature is normalized in the range [-1,1]. To normalize the feature the following equation [18] is used:

$$x' = \frac{x - \bar{x}}{\sigma} \quad (3)$$

Where x is the original feature vector, $\bar{x}$ is the mean of that feature vector, and σ is its standard deviation.

4. Results and Discussions

All of the results from the implementation of linear regression is provided in this section. The algorithm was run to forecast each of the crime: dacoit, robbery, murder, women & child repression, kidnapping, burglary and theft. Here dacoit, robbery, burglary and theft is considered into a single group to forecast. The result is divided into two parts. One is metropolitan region result and other is divisional region result. Table 1 shows that, how accurate the liner regression is to forecast future crime trends of Bangladesh.

Table 1: Total no of murders in 2014 at metropolitan and divisional region of Bangladesh

Region	Actual no. of Murder	Predicted no. of Murder
DMP	262	255
CMP	120	107
KMP	22	39
RMP	22	37
BMP	15	28
SMP	44	46
Dhaka Range	1395	1243
Chittagong Range	792	757
Sylhet Range	277	366
Khulna Range	520	486
Barisal Range	209	256
Rajshahi Range	463	406
Rangpur Range	349	416

For training purpose, learning parameter α is set to the value 0.01 and gradient descent is iterated 400 times. Fig. 1 shows the convergence of gradient descent with an appropriate learning rate.

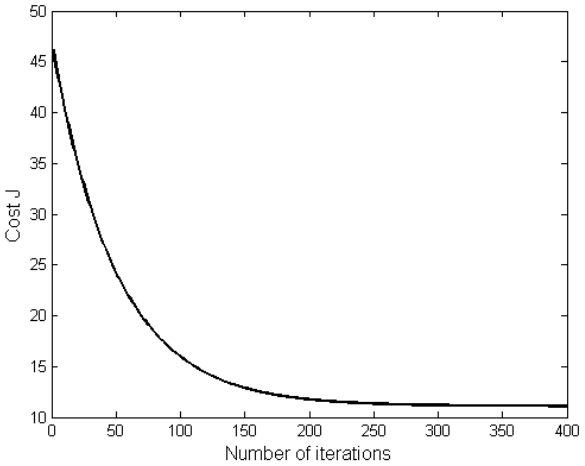

Fig. 1 Convergence of gradient descent with an appropriate learning rate.

After training linear regression, the regression parameters are: $\theta_0 = 1870.38$, $\theta_1 = 2422.64$ and $\theta_2 = -12.831$. These parameters are used to forecast murder in 2014 as follows:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$
$$h_\theta(x) = 1870.38 + 2422.64 * region + (-12.831) * year$$
$$h_\theta(x) = 1870.38 + 2422.64 * 10 + (-12.831) * 2014$$
$$h_\theta(x) = 255$$

The above calculation is done for Dhaka metropolitan police (DMP) region. Before calculation all the features are normalized. Different numeric values are used for different regions which was discussed in the subsection: dataset description. Here numeric value 10 is used for DMP region.

Table 2: Murder forecasting of 2018 at metropolitan and divisional region of Bangladesh.

	Number of crimes versus month for different regions												
Number of murders per month		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	DMP	77	78	80	84	88	94	99	108	116	126	137	149
	CMP	59	60	62	65	70	75	82	89	98	108	119	131
	KMP	40	41	44	47	51	57	63	71	80	89	100	112
	RMP	22	23	25	28	33	38	45	52	61	71	82	94
	BMP	3	5	7	10	14	20	26	34	43	53	63	75
	SMP	0	0	0	0	0	1	8	16	24	34	45	57
	DR	347	352	361	375	394	417	445	477	515	556	603	654
	CR	308	313	322	336	355	378	406	439	476	518	564	615
	SR	270	274	284	298	316	339	367	400	437	479	526	577
	KR	231	236	245	259	278	301	329	361	398	440	487	538
	BR	192	197	206	220	239	262	290	322	360	402	448	499
	BAR	154	158	168	181	200	223	251	284	321	363	409	460
	RAR	115	119	129	143	161	185	212	245	282	324	371	422

Crime forecasting at metropolitan region: In order to forecast different crimes in different months of 2018 at metropolitan region the following regression parameters are found after training linear regression model: $\theta_0 = -212.69$, $\theta_1 = -3.48$, $\theta_2 = 0.24$ and $\theta_3 = 0.13$

$$h_\theta(x) = -212.69 + (-3.48) * region + 0.24 * month + 0.13 * year$$
$$h_\theta(x) = -212.69 + (-3.48) * 10 + 0.24 * 1 + 0.13 * 2018 = 15$$

The above calculation is done for crime forecasting of murder in January of 2018 at DMP region. Similar calculation is done for other types of crime in different months of 2018 at metropolitan region. All the results of crime forecasting for metropolitan regions are shown through the table 2 to table 5. These tables show the change of different types of crime such as: dacoit, robbery, murder, women & child repression, kidnapping, burglary and theft with respect to month.

Table 3: Theft/dacoit forecasting of 2018 at metropolitan and divisional region of Bangladesh.

	Number of crimes versus month for different regions												
Number of theft/dacoit per month		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	DMP	80	80	80	80	81	82	83	84	85	86	88	89
	CMP	57	57	58	58	59	60	60	61	63	64	65	67
	KMP	35	35	36	36	37	37	38	39	40	42	43	45
	RMP	13	13	13	14	14	15	16	17	18	19	21	22
	BMP	0	0	0	0	0	0	0	0	0	0	0	0
	SMP	0	0	0	0	0	0	0	0	0	0	0	0
	DR	34	34	35	36	37	39	41	43	46	49	52	56
	CR	29	29	30	31	32	34	36	38	41	44	47	51
	SR	24	24	25	26	27	29	31	33	36	39	42	46
	KR	19	19	20	21	22	24	26	28	31	34	37	41
	BR	14	14	15	16	17	19	21	23	26	29	32	36
	BAR	9	9	10	11	12	14	16	18	21	24	27	31
	RAR	4	4	5	6	7	9	11	13	16	19	22	26

Table 4: Kidnapping forecasting of 2018 at metropolitan and divisional region of Bangladesh.

	Number of crimes versus month for different regions												
Number of kidnaping per month		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	DMP	7	7	8	8	8	9	10	10	11	12	13	14
	CMP	5	5	6	6	6	7	8	8	9	10	11	12
	KMP	3	4	4	4	5	5	6	6	7	8	9	11
	RMP	2	2	2	2	3	3	4	5	5	6	7	9
	BMP	0	0	0	0	1	1	2	3	3	4	6	7
	SMP	0	0	0	0	0	0	0	1	2	3	4	5
	DR	12	12	12	13	13	14	15	16	17	18	19	20
	CR	10	11	11	11	12	12	13	14	15	16	17	19
	SR	9	9	9	10	10	11	12	12	13	15	16	17
	KR	7	7	8	8	9	9	10	11	12	13	14	16
	BR	6	6	6	6	7	8	8	9	10	11	13	14
	BAR	4	4	4	5	5	6	7	8	9	10	11	12
	RAR	2	3	3	3	4	4	5	6	7	8	9	11

Crime forecasting at divisional region: In order to forecast different crimes in different months of 2018 at divisional region the following regression parameters are found after training linear regression model:
$\theta_0 = 2396.3$, $\theta_1 = -14.46$ and $\theta_2 = -0.67$ and $\theta_3 = -1.08$.

$$h_\theta(x) = 2396.3 + (-14.46) * \text{region} + (-0.67) * month + (-1.08) * year$$

$$h_\theta(x) = 2396.3 + (-14.46) * 10 + (-0.67) * 1 + (-1.08) * 2018 = 73$$

Table 5: Women repression forecasting of 2018 at metropolitan and divisional region of Bangladesh.

	Number of crimes versus month for different regions												
Number of Women repression per month		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	DMP	77	78	80	84	88	94	99	108	116	126	137	149
	CMP	59	60	62	65	70	75	82	89	98	108	119	131
	KMP	40	41	44	47	51	57	63	71	80	89	100	112
	RMP	22	23	25	28	33	38	45	52	61	71	82	94
	BMP	3	5	7	10	14	20	26	34	43	53	63	75
	SMP	0	0	0	0	0	1	8	16	24	34	45	57
	DR	347	352	361	375	394	417	445	477	515	556	603	654
	CR	308	313	322	336	355	378	406	439	476	518	564	615
	SR	270	274	284	298	316	339	367	400	437	479	526	577
	KR	231	236	245	259	278	301	329	361	398	440	487	538
	BR	192	197	206	220	239	262	290	322	360	402	448	499
	BAR	154	158	168	181	200	223	251	284	321	363	409	460
	RAR	115	119	129	143	161	185	212	245	282	324	371	422

The above calculation is done for crime forecasting of murder in January of 2018 at Dhaka range. Similar calculation is done for other types of crime in different months of 2018 at divisional region. All the results of crime forecasting for divisional regions are shown through the table 2 to table 5. These tables show the change of different types of crime such as: dacoit, robbery, murder, women & child repression, kidnapping, burglary and theft with respect to month.

Fig. 2 Population growth in Dhaka city of Bangladesh.

As depicted in fig. 2, the rate of population is increasing day by day. The comparison between fig. 2 and fig. 3 shows that, most of the cases the crime rate is also increasing with the growth of population.

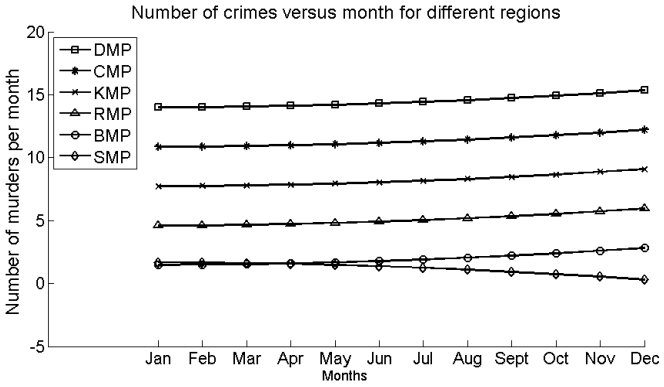

Fig. 3 Murder forecasting of 2018 at metropolitan region.

5. Conclusion

At present, Data mining has become a vital part of crime detection and prevention in many countries. There are other applications of data mining in the domain of law enforcement such as determining criminal "hot spots", creating criminal profiles, and learning crime trends to hold the pledge of abating crime related problem. In this paper, data mining technique is used to forecast future crime trends of Bangladesh. For this purpose, linear regression model is trained by crime data of previous years. After training linear regression, different types of crime are forecasted for the year of 2018. All the results are shown through the table 2 to table 5. Table 1 shows that, how accurate the liner regression is to forecast future crime trends of Bangladesh. From the experimental result it is also seen that, most of the crimes are increasing with the growth of population. The results of this data mining could potentially be used to lessen and even prevent crime for the forthcoming years. The extension of this research work is to forecast the location of crime occurrence, so that prior actions can be taken to prevent crime.

References

[1] http://www.dmp.gov.bd/application/index/page/crime-data

[2] http://www.police.gov.bd/Crime-Statistics-yearly.php?id=337

[3] A. H. Wahbeh, Q. A. Al-Radaideh, M. N. Al-Kabi, E. M. Al-Shawakfa, "A Comparison Study between Data Mining Tools over some Classification Methods", International Journal of Advanced Computer Science and Applications, The SAI Organization, Special Issue on Artificial Intelligence, pp. 18-26, 2011.

[4] N. Levine, "CrimeStat: A Spatial Statistic Program for the Analysis of Crime Incident Locations (v 2.0)", Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC, May 2002.

[5] S. Yamuna, N. S. Bhuvaneswari, "Data mining Techniques to Analyze and Predict Crimes", The International Journal of Engineering and Science, Vol.1, Issue.2, pp.243-247.

[6] A. L. Buczak, C. M. Gifford, "Fuzzy Association Rule Mining for Community Crime Pattern Discovery", In

ACM SIGKDD Workshop on Intelligence and Security Informatics (ISIKDD' 10), 2012.

[7] J. Han, M. Kamber, J. Pei, "Data Mining: Concepts and Techniques", vol. 5. Morgan Kaufmann Publishers, USA, 2012.

[8] S. Shojaee, A. Mustapha, F. Sidi, M. A. Jabar, "A study on classification learning algorithms to predict crime status", In: International Journal of Digital Content Technology and its Applications (JDCTA) 7.9 (May 2013), pp. 361–369, issn: 1975-9339.

[9] A. Malathi, S. S. Baboo, "Enhanced Algorithms to Identify Change in Crime Patterns", International Journal of Combinatorial Optimization Problems and Informatics, Aztec Dragon Academic Publishing, vol. 2, no.3, pp. 32-38, 2011.

[10] L. McClendon, N. Meghanathan, "Using Machine Learning Algorithms to Analyze Crime Data", Machine Learning and Applications: An International Journal (MLAIJ) vol. 2, no. 1, March 2015.

[11] C. H. Yu, M. W. Ward, M. Morabito, W. Ding, "Crime Forecasting Using Data Mining Techniques", In Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW' 11), pp. 779-786, 2011.

[12] S. V. Nath, "Crime pattern detection using data mining", in Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, pp. 41-44, 2006.

[13] S. Chainey, L. Tompson, S. Uhlig, "The utility of hotspot mapping for predicting spatial patterns of crime", Security Journal, 21, 4–28, [104].

[14] J. S. D. Bruin, T. K. Cocx, W. A. Kosters, J. Laros, J. N. Kok, "Data mining approaches to criminal career analysis", in Proceedings of the Sixth International Conference on Data Mining (ICDM'06), pp. 171-177, 2006.

[15] P. Thongtae, S. Srisuk, "An Analysis of Data Mining Applications in Crime Domain", In Proceedings of the IEEE International Conference on Computer and Information Technology Workshops, pp. 122-126, 2006.

[16] S. Bagui, "An Approach to Mining Crime Patterns", International Journal of Data Warehousing and Mining, 2, 1, pp. 50–80, 2006.

[17] T. Abraham, O. D. Vel, "Investigative Profiling with Computer Forensic Log Data and Association Rules", In Proceedings of the IEEE International Conference on Data Mining (ICDM'02), pp. 11 – 18, 2002.

[18] P. Phillips, I. Lee, "Crime Analysis through Spatial Areal Aggregated Density Patterns", GeoInformatica, Springer, vol. 15, no. 1, pp. 49-74, 2011.

[19] S. M. Nirkhi, R. V. Dharaskar, V. M. Thakre. "Data Mining : A Prospective Approach for Digital Forensics", International Journal of Data Mining & Knowledge Management Process, vol. 2, no. 6 pp. 41-48, 2012.

[20] https://en.wikipedia.org/wiki/Feature_scaling.

[21] R. Wortley, L. Mazerolle, "Environmental Criminology and Crime Analysis", Willan Publishing, UK, 2008.

[22] A. Bogomolov, B. Lepri, J. Staiano, N. Oliver, F. Pianesi, and A. Pentland, "Once upon a crime: Towards crime prediction from demographics and mobile data", In Proceedings of the ACM ICMI, to appear, (2014).

[23] E. Ferrara, P. D. Meo, S. Catanese, and G. Fiumara, "Detecting criminal organizations in mobile phone networks", Expert Systems with Applications 41, 5733–5750 (2014).

[24] J. R. Zipkin, M. B. Short, and A. L. Bertozzi, "Cops on the dots in a mathematical model of urban crime and police response", Disc Cont Dyn Syst, B 19 (2014).

[25] Md. Abdul Awal, Jakaria Rabbi, Sk. Imran Hossain, and M. M. A. Hashem, " Using Linear Regression to Forecast Future Trends in Crime of Bangladesh ", Proceedings of International Conference on Informatics, Electronics & Vision (ICIEV16), Dhaka, Bangladesh, IEEE, May 2016.

Authors -

Md. Abdul Awal

Department of Computer Science and Engineering, Khulna University of Engineering & Technology (KUET), Khulna 9203, Bangladesh.

Jakaria Rabbi

Department of Computer Science and Engineering, Khulna University of Engineering & Technology (KUET), Khulna 9203, Bangladesh.

Imran Rana

Department of Computer Science and Engineering, Khulna University of Engineering & Technology (KUET), Khulna 9203, Bangladesh.