**Crime Analysis and Prediction in Bangladesh using Machine Learning**

**Abstract:**

**Crimes occurrence is an alarming issue in Bangladesh. If crimes and their trends could be predicted from beforehand, it would be beneficial to the law enforcements to distribute the scarce resources and ensure a safer country by attempting to prevent those crimes. This paper conducts an investigation into various machine learning approaches to performing forecasting of crimes in Bangladesh based on the official crimes data available on official Bangladesh Police website. On analysis, it is observed that Linear Regression, Logistic Regression, Regression Tree, Random Forest Regressor, Support Vector Regression, Adaboost, Gradient boost, XGboost and Multilayer Perceptron all achieve above 97% accuracy where Multilayer Perceptron and Logistic Regression enjoys a 100% accuracy when it comes to forecasting crime values of 2018 in Bangladesh based on the data available from 2010 to 2017.**

**Introduction:**

Crime is a social problem which generally refers to harmful and violent act by breach of laws. This is one of the most threatening global issues at present and is increasing throughout the world. The detrimental effects of crime can hinder the socio-economic development of a country. Besides, the end result of criminal acts has various emotional and physical impacts in a society violating human rights. Hence, prevention of crime has become crucial for preserving human rights. In order to prevent crime, it is important to understand the patterns of criminal occurrences in a certain area by analyzing crime data.

With the advance development in Artificial Intelligence and Machine Learning, a wide range of Machine Learning applications is being implemented in the field of Criminology. Criminology is a process that aims to identify crime characteristics. Actually crime analysis includes exploring and detecting crimes and their relationships with criminals. The high volume of crime datasets and also the complexity of relationships between these 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 knowledge that is gained from data mining approaches is a very useful tool which can help and support police forces. An approach based on data mining techniques is discussed in this paper [5]

Companies and cities all over the world are experimenting with using artificial intelligence to reduce and prevent crime, and to more quickly respond to crimes in progress. The ideas behind many of these projects is that crimes are relatively predictable; it just requires being able to sort through a massive volume of data to find patterns that are useful to law enforcement. This kind of data analysis was technologically impossible a few decades ago, but the hope is that recent developments in machine learning are up to the task.

There is a good reason why companies and government are both interested in trying to use AI in this manner. As of 2010, the United States spent over [$80 billion a year](https://www.brookings.edu/wp-content/uploads/2016/06/v8_THP_10CrimeFacts.pdf) on incarations at the state, local, and federal levels. Estimates put the United States’ total spending on law enforcement at over [$100 billion a year](http://www.justicepolicy.org/research/3906). Law enforcement and prisons make up a substantial percentage of local government budgets.[6]

Although this is a worldwide problem, many of the most serious crime problems are now to be found in the developing countries [4]. The growth rate of crime in Bangladesh is rampant. According to Global Peace Index 2019, Bangladesh is 94th among 195 countries of the world having 2.128 as peace index [3]. As social level of morality is one of the factors influencing the crime rate [1], criminal behaviour varies in different societies or areas.

**Literature Review:**

Police and the government are working ceaselessly to ensure that crimes occur minimally. If it was possible to somehow predict where and what crimes are occurring most frequently, it would be possible to target these regions and employ the required police force in that area to deter criminals. Numerous researchers, often with aid from the police and government have dedicated their time and resources to detect criminal activities. Decision Tree is a popular algorithm for classification and the classification of crimes is no exception to this, evident from proceedings in [9,11,16,17,18,28]. [9] produced a dataset using data entry for each crime that occured which contained the type of event, the location in longitude and latitude, and the time and date of the crime. This dataset was organized in a spatial grid pattern and analysed using 1NN using Euclidean distance which was used as a base algorithm with a location constrained variation, along with SVM hypertuned with kernel type of radial basis function, a 2 layer Neural Network, J48 (a C4.5 Decision Tree) and Naive Bayes. Naive Bayes outperforms all the other algorithms for this particular dataset. US State database where SQL was used to retrieve and preprocess the data and was mined used a Decision Tree and Clustering Algorithm in [16]. Similarly, Communities and Crime Unnormalized Dataset provided by the University of California-Irvine repository and actual crime statistical data for the state of Mississippi that has been provided by neighborhoodscout.com was employed by [11] to implement Linear Regression, Additive Regression, and Decision Stump algorithms on the basis of Correlation coefficient, Mean absolute error, Root mean squared error, Relative absolute error, and the Root relative squared error. Amongst the three algorithms, Linear Regression provides the most efficient and accurate in predicting crime rates. [29] established an auto regressive model (ARIMA) on the urban data of Chicago using Crimes - 2001 to present dataset. This regression model was able to forecast crime trends with an accuracy of 84 percent on one year-ahead and 80 percent on two-year-ahead forecasts. [18] compared the performance of Decision Tree and Naive Bayes algorithms also on Communities and Crime Unnormalized Dataset from [11] where the Decision Tree surpassed the classification accuracy of Naive Bayes. Fuzzy association rule mining was applied for community crime pattern discovery using the same Communities and Crime mentioned in [11] for 40 attributes in [10] where a pruning method based on support, confidence, and lift was also suggested to discover the rare rules present in the dataset. [28] evaluated crime data mining on the basis of ANN, decision trees, rule induction, nearest-neighbor method, and genetic algorithm. CrimeStat [7] is a brilliant tool that displays spatial distribution, spatial autocorrelation using monte carlo simulation and distance analysis between crimes and provides analysis about where the most frequent crimes occur (based on category) using a fuzzy logic association rule as well as spatial prediction of crimes separable by various feature variables. It also provides a spatial regression of crimes separable by various feature variables. [14] designed a tool that can extract information such as crime nature, frequency, duration and severity from police database and create digital profiles for the offenders and use distance measures to cluster them accordingly. Another general crime detecting framework was developed by [19] on the four different types of data mining techniques when it came to crime detection, namely - entity extraction, association, prediction, and pattern visualization. A challenge when handling with data from real scenarios is how to handle the missing data, whether due to lack of information or human laziness. [8] focuses on filling in missing values using methods such as a KNN-based imputation method enhanced utilizing LVQ (Learning Vector Quantization) methods combined with generalized relevance learning (which can also be used in classification) and Expectation-Maximization (EM) algorithm and clusters the cleaned dataset using K means algorithm and DBScan algorithm. Anomalies are then detected using PDSAD (Partition and Distance Method based on Shape for Anomaly Detection) model which reduces the time complexity compared to traditional anomaly detection methods. Another KNN based work where LDA is utilized for dimension reduction was presented in [21] where Spatial-temporal characteristics are extracted from time sequence of area-specific heat levels, temporal distance of important holidays, and neighborhood features. To properly determine the neighbourhood features, histogram based statistical methods is used. The analysis shows optimum prediction occurs on weekly based KNN application. KNN, Parzen windows, and Neural Networks, analyzed to compare the performance metrics for predicting the crimes in San Francisco in [33] where KNN displayed superior results. Gradient Boosting Machine Learning algorithms were employed to reveal hidden links in the criminal data in [34]. It was also observed that more decentralized and less clustered network has better performance on finding hidden links as the machine learning model avoids overfitting. It can be observed that researchers from all over the world have used a variety of algorithms to detect and predict crimes. [15] attempted to provide a comprehensive guideline for developing algorithms that would encompass variables such as types of crimes, background knowledge of the analyst performing the research work and so on. Crime was predicted using K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Naïve Bayesian, Decision Tree, and Ensemble Methods in Vancouver [37] where KNN and boosted decision tree resulted in crime-prediction accuracy between 39% to 44%. In [17], crime hotspots for London, UK was predicted using logistic regression, support vector machines, neural networks, decision trees, and different implementations of ensembles of tree classifiers with different parameters from a multitude of geographic data collected from phones and demography. The research resulted in the decision tree classifier based on the Breiman’s Random Forest (RF) algorithm yielding the best performance when compared to all other classifiers. An approach [27] that leverages Bayesian framework with a Gaussian Process prior and MCMC attempts to handle uncertainties in crime prediction by modelling the dependency between offence data and environmental factors such as demographic characteristics and spatial location for New South Wales, Australia. In Brazil, random forest algorithm was used in [31] to evaluate crime prediction where random forest resulted in 97% accuracy. The Dempster-Shafer theory of evidence integrated with the multi-kernel method was used to establish a crime prediction forecast for the region of Santiago, Chile [32]. In [23], drug-related criminal activity in Taiwan was predicted using Random Forest, Naïve Bayes and a proposed deep learning method. This deep learning method is a data-driven method based on “broken windows” theory and spatial analysis which surpassed the performance metrics in comparison to Random Forest and Naïve Bayes. Another broken windows theory-based work [25] uses feature-level data fusion method with environmental context based on a deep neural network (DNN) where training occurs on spatial, temporal, environmental, and joint-feature representation layers. The data for the experiment comes from various online databases of crime statistics, demographic and meteorological data, and images in Chicago, Illinois. DNN is said to give the best results set against other machine learning algorithms. Various researchers aimed to investigate the crime detection and prediction using spatial features. [7, 9, 12, 13, 20, 21, 24, 25, 26, 27,30] are some of the works that presented crime detection and prediction on the basis of spatial features. In [12], K-means clustering with dynamical weights assignment combined with a semi supervised learning technique for knowledge discovery is used in crime detection based on their geo-spatial plots. In [13], Hotspot mapping accuracy is compared in relation to the mapping technique that is used to identify concentrations of crime events (thematic mapping of Census Output Areas, spatial ellipses, grid thematic mapping, and KDE) and by crime type – four crime types are compared (burglary, street crime, theft from vehicles and theft of vehicles). KDE performed best in predicting spatial patterns of crimes and street crimes was the best type of crime to predict where the next crime will occur. In [20], Area specific crime hotspots were identified using Kernel Density Estimation (KDE) and Risk Terrain Modeling (RTM) algorithms. In [24], Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) and reverse-geocoding technique is used to find crime hotspots for different categories of crime on crime data from Royal Canadian Mounted Police of Halifax, NS. In [26] A transfer learning framework was suggested that explored spatio-temporal patterns for crime prediction with cross-domain urban data. A framework - CRIMETRACER [24] utilizes random walk algorithm for spatial crime analysis (a probabilistic model of spatial behavior of known offenders within their activity space) and crime location prediction outside of hotspots. It can be seen that a great deal of exploration has been done on attempting to detect or predict crimes in other countries. However, the area is nearly untouched when it comes to Bangladesh. [22] introduced a novel approach for crime hotspot detection in Bangladesh using Artificial Neural Network (ANN) - CRIMECAST. The time performance of the ANN was improved by utilizing a Gamma test. A type of crime, crib, is analysed using decision tree on data obtained from Dumki police station in Bangladesh, web sources and Facebook [35]. Linear regression trained by gradient descent is employed in [36] to predict the crime rates for Bangladesh using real dataset of crime of previous years.

**Methodology**

References:

1. <https://legalbeagle.com/4572257-do-juveniles-commit-crimes.html>
2. <https://www.numbeo.com/crime/country_result.jsp?country=Bangladesh>
3. <http://worldpopulationreview.com/countries/safest-countries-in-the-world/>
4. <https://link.springer.com/article/10.1186/s40163-016-0056-7>
5. <https://www.sciencedirect.com/science/article/pii/S1877050910005181>
6. <https://emerj.com/ai-sector-overviews/ai-crime-prevention-5-current-applications/>

7. Ned Levine (2015). CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 4.02). Ned Levine & Associates, Houston, Texas, and the National Institute of Justice, Washington, D.C. August.

8. 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.

9. 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.

10. 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.

11. 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.

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. 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," Proc. of the 16th Intl. Conf. on Multimodal Interaction, pp. 427-434, 2014.

18. R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. Shariat Panahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," Indian J. of Sci. and Technol., vol. 6, no. 3, pp. 4219- 4225, Mar. 2013.

19. H. Chen, W. Chung, J. J. Xu, G. Wang, Y. Qin, and M. Chau, "Crime data mining: a general framework and some examples," IEEE Computer, vol. 37, no. 4, pp. 50-56, Apr. 2004

20. M. Al Boni and M. S. Gerber, "Area-specific crime prediction models," 15th IEEE Intl. Conf. on Mach. Learn. and Appl., Anaheim, CA, USA, Dec. 2016.

21. Q. Zhang, P. Yuan, Q. Zhou, and Z. Yang, "Mixed spatial-temporal characteristics based crime hot spots prediction," IEEE 20th Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD), Nanchang, China, May 2016.

22. N. Mahmud, K. Ibn Zinnah, Y. Ar Rahman, and N. Ahmed, "CRIMECAST: a crime prediction and strategy direction service," IEEE 19th Intl. Conf. on Comput. and Inform. Technol., Dhaka, Bangladesh, Dec. 2016.

23. Y. L. Lin, L. C. Yu, and T. Y. Chen, "Using machine learning to assist crime prevention," IEEE 6th Intl. Congr. on Advanced Appl. Inform. (IIAI- AAI), Hamamatsu, Japan, Jul. 2017.

24. F. K. Bappee, A. S. Júnior, and S. Matwin, "Predicting crime using spatial features," Can. AI 2018: Advances in Artificial Intel.-Lecture Notes in Comput. Sci., vol. 10832, pp. 367-373, Springer, Mar. 2018.

25. H. W. Kang, H. B. Kang, "Prediction of crime occurrence from multi-modal data using deep learning," PLoS ONE, vol. 12, no. 4, Apr. 2017.

26. X. Zhao and J. Tang, "Exploring transfer learning for crime prediction," IEEE Intl. Conf. on Data Mining Workshop (ICDMW), New Orleans, LA, USA, Nov. 2017.

27. R. Marchant, S. Haan, G. Clancey, and S. Cripps, "Applying machine learning to criminology: semiparametric spatial demographic Bayesian regression," Security Inform., vol. 7, no. 1, Dec. 2018.

28. S. Prabakaran and S. Mitra, "Survey of analysis of crime detection techniques using data mining and machine learning," Nat. Conf. on Math. Techn. and its Appl. (NCMTA 2018), IOP J. of Physics: Conf. Series, vol. 1000, 2018.

29. E. Cesario, C. Catlett, and D. Talia, "Forecasting crimes using autoregressive models," IEEE 14th Intl. Conf. on Dependable, Auton. And Secure Comput., Auckland, New Zealand, Aug. 2016.

30. M. A. Tayebi, U. Glässer, and P. L. Brantingham, "Learning where to inspect: location learning for crime prediction," IEEE Intl. Conf. on Intel. and Security Inform. (ISI), Baltimore, MD, USA, May 2015,.

31. L. G. A. Alves, H. V. Ribeiro, and F. A. Rodrigues, "Crime prediction through urban metrics and statistical learning," Physica A, vol. 505, pp.435-443, 2018.

32. N. Baloian, E. Bassaletti, M. Fernández, O. Figueroa, P. Fuentes. R. Manasevich, M. Orchard. S. Peñafiel, J. A. Pino, and M. Vergara, "Crime prediction using patterns and context," IEEE 21st Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD), Wellington, New Zealand, Apr. 2017.

33. M. V. Barnadas, Machine learning applied to crime prediction, Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, Sep. 2016.

34. E. Budur, S. Lee, and V. S. Kong, "Structural analysis of criminal network and predicting hidden links using machine learning," arXiv:1507.05739, Sep. 2015.

35. Sarna Majumder et al, “Crib Analysis in Bangladesh,” International Journal of Computer Science and Mobile Computing, Vol.7 Issue.7, July- 2018, pg. 145-148

36. Awal, Md. Abdul, Jakaria Rabbi and Imran Ahmad Rana. “Using Data Mining Technique to Analyze Crime of Bangladesh.” (IJCSN - International Journal of Computer Science and Network, Volume 6, Issue 4, August 2017

37. S. Kim, P. Joshi, P. S. Kalsi and P. Taheri, "Crime Analysis Through Machine Learning," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2018, pp. 415-420.

38. Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 11155(1):119–139, August 1997.

39. Freund, Y., Robert, S., Abe, N.: A short introduction to boosting. J. Jpn. Soc. Artif. Intell. 14, 771–780 (1999)

40. Friedman, Jerome H. “Greedy Function Approximation: A Gradient Boosting Machine.” The Annals of Statistics, vol. 29, no. 5, 2001, pp. 1189–1232.

41. Chen, Tianqi & Guestrin, Carlos. (2016). XGBoost: A Scalable Tree Boosting System. 785-794. 10.1145/2939672.2939785.

42. Mitchell, T. M. (1997). Machine Learning. New York City: McGraw Hill.

43. Rokach, Lior & Maimon, Oded. (2005). Decision Trees. 10.1007/0-387-25465-X\_9.

44. Ho, Tin Kam (1995). Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.

45. Ho TK (1998). "The Random Subspace Method for Constructing Decision Forests" (PDF). IEEE Transactions on Pattern Analysis and Machine Intelligence. 20 (8): 832–844.

46. Breiman, L. Random forests. Mach. Learn. 2001

47. Biau, Gérard & Scornet, Erwan. (2015). A Random Forest Guided Tour. TEST. 25. 10.1007/s11749-016-0481-7.

48. G. Biau. Analysis of a random forests model. Journal of Machine Learning Research, 13:1063–1095, 2012.

49. <https://machinelearningmastery.com/a-gentle-introduction-to-the-bootstrap-method/>

50. <https://medium.com/m/global-identity?redirectUrl=https%3A%2F%2Fblog.exploratory.io%2Fintroduction-to-extreme-gradient-boosting-in-exploratory-7bbec554ac7#:~:text=XGBoost%20is%20one%20of%20the,it%20helps%20to%20reduce%20overfitting.>