Survey on Crime Interpretation and Forecasting Using Machine Learning

Ankita A Khartmol[1], Nikki[2], Swastika Arya[3], Trushika Arya[4]
Mahalakshmi Manasa[5]

[1] Student, Dept. of Computer Science engineering, Dayananda Sagar College of Engineering, Karnataka, India
[5] Professor, Dept. of Computer Science engineering, Dayananda Sagar College of Engineering, Karnataka, India

---***---

Abstract - *Understanding the crime pattern and taking safety measures to solve the crime problems has become the important factor in today's world. The main aim here is to locate the crime location and find the pattern based on the time and place in Bangalore city. There is a lot of rise in the software systems which helps the officers to solve the crime issues. We will be looking at the various machine learning algorithms which helps to detect the crime pattern and helps in solving the crime in very less time. There are different algorithms in machine learning like k-means clustering algorithm which can be used in the process of identification of crime. The Data mining approach is used to predict the features of crime dataset that affects the high crime rate. There are two types of algorithms i.e., supervised and unsupervised, supervised algorithms that is used to evaluate the training dataset for its accuracy and unsupervised algorithms helps in solving the unlabelled data into classes or clusters. There are other few different learning methods like random forest in data mining based on unknown and the previous year collected datasets which can be used for predicting the crime pattern based on the time and place.*

Key Words: **Crime data, Decision tree, KNN, Machine learning, Naïve bayes and Random Forest**

1. INTRODUCTION

Criminal activities are one of the most important and common problems in the existing society and the prevention of it is a key hassle. A huge number of crimes take place each day. With the present scenario, no solution can grant the best safety to an individual to prevent any harm that can be caused by such an activity. Therefore, this necessitates prediction of crimes in advance by storing all such criminal data and maintaining a good database for the future and for further analysis. Crime has become the huge problem in our society and its safety is an important task. Machine Learning algorithms have been increasing that made crime prediction possible based on the past data. To improve more security measures in society and in explicit crime category regions recognizing crime patterns will allow us to solve problems with different approaches. To understand the future scope of crimes past crime data trending factors may help us to detect the future crimes. Gradually crime rate is increasing very much. These days crimes are difficult to predict as they are not systematic. Even the modern technologies and other high-level techniques helps criminals in accomplishing the

crime. As per the crime records bureau crimes like robbery, stealing, theft and other few have reduced but the other crimes except these types of crime have been increased. It is difficult to predict who may be the victims but it is possible to predict the probability of occurrence. The outcome tells that our application helps in reducing the crime to a possible extent by providing security in sensitive areas as the results predicted cannot give clarity with full perfection. For collecting the crime records and evaluating it, we need to build a strong crime analytics tool. We will be using the various ML models to predict crimes. To understand the pattern of crimes we need to keep area wise track for the geographical analysis. The other visualization techniques and different plots can help law enforcement agencies and to the police department fields to predict the crimes and detect them with higher accuracy. By using data mining algorithms, it will be helpful to extract the previous unknown data and useful information from the dataset and find out the crime pattern. The challenges faced will be to analyse the dataset and for maintaining a proper data of the criminal activities and using the different statistics method to solve and predict the crimes in future. There is an approach to solve between the criminal justice and computer science for developing the data mining procedure which helps to solve crime faster. We will be focusing on the crime factors on each day instead of finding out the criminal background of the offender and other politics etc. The aim here is to analyse the data that integrate a vast number of crimes and to identify the number of crimes which may occur in future. Therefore, the crime evaluation a prediction are well known approaches for finding out and analysing the crime styles and the trending factors in crime.
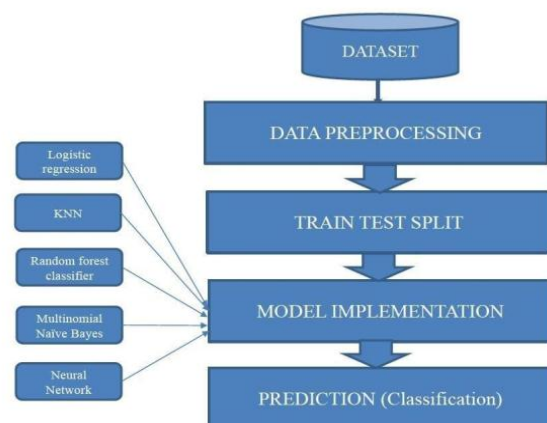

Figure – 1: Model Implementation

2. MOTIVATION

With high growth in popularity of crime analysis and forensics, engineers and researchers have been able to find a large number of patterns taking place in the crime. The meaning of a crime is any act which is punishable by law under the government, since it is a gross violation of the law. It refers to several variants of illegal activities that are forbidden by law. Most crimes are known to be committed today by an individual for personal pleasure or satisfaction, lack of monetary benefits, lack of basic necessities such as food and shelter, or for seeking revenge. Machine learning now provides us with the mechanism to use artificial intelligence to create automated analysis techniques that were considered impossible for computers. The danger of crimes occurring is high for a lot, especially in developing countries where millions face poverty, dwell in slums, unemployment and are hence exposed to increased risky environments. Criminal actions are defined by the laws of specific jurisdictions and many times, there are vast differences between countries and even within the same country regarding what kind of activity is prohibited and illegal and can be classified as criminal. Recently, the National Criminal Records Bureau of India accounted that 2,82,171 criminal cases were reported in a single year in Uttar Pradesh alone, the highest for any state in India. Therefore, it is important to classify what activities constitute criminal offences and hence, the prevention of the same and to safeguard the public is thus need of the hour.

3. RELATED WORK

[A] Lalitha and Suresh Babu, "Cluster based Zoning of Crime Info", (2017)

A CICD (Criminal Identification and Crime Detection) method is used to identify illegal activities in India. In this method, criminals are identified based on their facial recognition, type of crime and its motive, weapon used etc. It consists of six primary steps: 1) Data extraction 2) Clustering, 3) Pre-Processing, 4) Mapping, 5) Classification and the use of WEKA tool (Collection of ML Algorithms). In this method, K-Means and KNN together are used for filtration in big data collections. In the United Kingdom, Cambridge Police Department have performed a project with the name Series Finder to find out the crime patterns in theft. Patterns in crime have been extracted to attain the criminals by using modulus operandi. The information included are the way the criminal illegally enters the area of the theft, timestamp of theft and the kind of property (independent, lease, rental), along with the geographical location with the other break-ins. Using different crime series, it was able to identify most kind of crimes happening. The prediction's accuracy was high (above 80%) in London, United Kingdom, mobile activities along with the location information is speed to predict important crime spots. In USA, Machine learning is applied on live stream of video

coming from CCTV footage to check any unusual activity occurring, for example, K means algorithm is used in a room which allows limited access to check if any unusual access has occurred at a particular time. KDD – a collective method which includes statistical modelling using machine learning algorithms, one of the most important techniques in any machine learning application, is used in crime analysis. The main function of KDD is prediction of human behaviour and finding patterns. [1]

[B] Chris Delaney, "Crime Pattern Definitions for Tactical Analysis", (August 2011)

Criminal patterns are a foundational concept in crime analysis. By elaborating and classifying the various crimes occurring in patterns, the IACA looks to increase communication and enhance knowledge amongst police practitioners. [2]

[C] Prajakta Yerpude and Vaishnavi Gudur, "Predictive Modelling of Crime Dataset Using Data Mining", (July, 2017)

Due to increase in crimes across the world, there is a need to decrease the occurrence of such events that threaten the right to one's life. It helps society and cops to take adequate steps to solve crimes faster. Data mining approach has been taken to predict the crime dataset that result in high crime rates. Regression, Naive Bayes and Decision trees are machine learning methods performed on past data and used for forecasting that cause crime in a locality. The Police Department and Crimes Record Bureau can take measures to reduce crimes rates using the crime patterns and results generated through predictions. [3]

[D] Shyam Nath, "Crime Pattern Detection Using Data Mining", (2006)

Data mining is used as a pattern for crime inspection problems. Any study that can assist in resolving crimes faster will remedy the cruel acts in society. 10% of the criminals constitute over 50% of the crimes that occur. Here, clustering algorithm approach to identify the criminal trends is used to fasten up the process of resolving offences. Then further K-means clustering has been taken and enhanced it to predict crime. [4]

[E] Dawei Wang, Wei Ding, Henry Lo, Tomasz, "Crime hotspot mapping using the crime related factors", (2012)

The concept of the paper helps in analysing spatial factors of activities revolving crime. Spatial distribution of crime involves a distribution of crime opportunity and socio-economic factors. Current approaches live more on the density of crime to store the information without analysing the factors. A new crime hotspot mapping tool - Hotspot Optimization Tool (HOT) is introduced. It is an application of spatial data mining. HOT is Geospatial Discriminative

Patterns (GD Patterns) that captures dissimilarity between the datasets. [5]

[F] Shiju Sathyadevan, Surya S, "Crime analysis and prediction using data mining", (2014)

The main objective is to focus on crime factors instead of crime occurrences. With the help of data mining, previous unknown data is extracted from raw data. One can synthesize between criminal justice and computer science to model a data mining technique to resolve crimes faster. This paper explains various types of crime prediction and analysis using several data mining. [6]

[G] Sathyadevan, S., & Gangadharan, S., "Crime Analysis and Prediction Using Data Mining", (2014, August)

Paper let us know about the cluster analysis which is done by k-means algorithm on crime dataset. The custom map generates hotspots of crime and displays crime details of different states which helps the crime departments to take precautionary measures to fight against the crime plan advanced investigation strategies. Zoning awareness could aid in cautioning police to take increased levels of action. [7]

[H] Tayebi, M. A., Gla, U., & Brantingham, P.L., "Learning where to inspect: Location Learning for Crime Prediction", (2015, May)

The motive here is to look into the machine learning methods for predicting crime which helps to yield better results with good accuracy using better techniques with importance of the crime datasets that helps in analyzing the crime. Data set is analyzed with the help of the supervised machine learning techniques to collect the information with labelled data and do further validation process of it and visualization will be on the complete dataset. [8]

[I] P. Thongtae and S. Srisuk, "An analysis of data mining applications in crime domain", (2008)

Examining and evaluating pre-existing databases and evaluating them to generate meaningful information underlies data mining. New information involving crime trends is extracted and analysed from criminal dataset. Increased number of techniques to predict and to do analysis in data mining is taken place. However, only a rare few efforts were invested in the field of criminology. [9]

[J] Yadav, S., Timbadia, A., Vishwakarma, R., & Yadav, N., "Crime Patteren Detection, Analysis & Prediction". (2017)

Series finder, a technique to find patterns in criminal activity was developed by the police department, Cambridge UK. They used the modus (M.O.) of offender as well as extracted illegal patterns that were seen all the way through the person responsible for. The police built M.O. of the guilty. It is a fixed behaviour of a criminal and is variant of action

which represents a type of pattern. The information included is the way the criminal illegally enters the area of theft, the timestamp of theft and kind of property (independent, lease, rental), along with the location etc. It identifies several of the majority of crime patterns and styles and also identifies several illegal activities. The correctness of it was higher than 80%. Therefore, a similar kind of approach is to be implemented by finding patterns in crime. [10]

4. COMPARISON TABLE

Author and Year	Title	Concept	Limitations
Shiju Sathyadevan, Surya S, 2014	Crime Analysis and Prediction using Data Mining:	Crime analysis has been a strong and a valuable technique for analysing crime and other factors. Here we can extract the data from the unknow sources.	To get better results for prediction there is a need to make a search for many crime attributes that is of different places not by just fixing it. Need to include more attributes for better accuracy.
Prajakta Yerpude and Vaishnavi Gudur, 2017	Predictive Modelling of Crime dataset using data mining	As there is continuous increase in crime across the globe, crime data should be analysed to decrease the crime rate. Therefore, it helps the police force in solving the crime issues faster.	Higher the value of accuracy the data become insignificant for the model compared to significant data which shows that regression needs continuous data including the sparse values.
P. Thongtae and S. Srisuk, 2008	An analysis of Data Mining Applications in Crime Domain	This paper gives the efficient and effective way to solve the crime problems i.e., the methods or and different types of approaches on data mining for crime data analysis. Here the idea is to find the	The data mining techniques and the meta- data has to be chosen based on background awareness gained of the analyst. Hence, there is a probability to choose the integrated data which can be performed repeatedly.

		criminal activities of professional identity of the fraud criminals based on the cognizance found of their own background.	
Shyam Varan Nath, 2006	Crime Pattern Detection Using Data Mining	In today's world crimes have been a social bother which cost's our society in many ways. Researches which can help in solving crime faster will be effective. Machine learning works with geo-spatial plot which helps in the improvement in the productivity of the detectives and the other crime departments.	Crime pattern analysis can only help the detective. The quality of input data is very sensitive for data mining which may not be the exact value and that may have not had the important information.
Dawei Wang Wei Ding, Henry Lo, Tomasz 2012	Crime hotspot mapping using the crime related factors—Spatial Data Mining Approach	The technique of hotspot mapping has been used in vast for analysing the spatial behaviour of crimes. The crime is considered to be the related with the socio-economic and other community of crime factors.	The hotspot mapping technique is based on the grid the matic mapping which doesn't permit for the demonstration of the hotspots. The conversion of points which indicates the crime incident cells with crime counts and the other details across the cells can be lost.

Table – 1: Comparison Table

5. CONCLUSION

We can see just how easy it is to get started with a revolution to safeguard our society from the evil eye, through crime prediction which can change the whole crime scenario in the world. It is remarkable to see the success of machine learning in such varied real-world problems. The aim of society must not be just to catch criminals but to prevent crimes from occurring in the first place. In the future, the system could be extended to predict who will commit a crime based on the behavioural patterns observed in the person. It can also be used to correlate the relationship between crimes occurring in different places and thereby truly acting as a prevention to the public. In this project, here it has been demonstrated as a classification of the different criminal acts wherein the user inputs a location and given time for which he wants to predict if he is safe or not. The model he, which separates it from other methods that rely heavily on transfer learning approach and is user-friendly.

REFERENCES

[1] Lalitha and Suresh Babu, Cluster based Zoning of Crime Info, IEEE, 2017, DOI: ieeexplore.ieee.org/document/7905269

[2] Chris Delaney, "Crime Pattern Definitions for Tactical Analysis", (August 2011)

[3] Prajakta Yerpude and Vaishnavi Gudur, -- International Journal of Data Mining & Knowledge Management Process, Predictive Modelling of Crime Dataset Using Data Mining, (IJDKP), Vol. 7, No.4, July 2017, DOI: 10.5121/ijdkp.2017.7404

[4] Shyam Nath, —Crime Pattern Detection Using Data Mining, IEEE/WIC/ACM International Conference on Web Intelligence, 2006, DOI: ieeexplore.ieee.org/document/4053200

[5] Dawei Wang, Wei Ding, Henry Lo, Tomasz, a spatial data mining approach, Applied Intelligence, 2012, Josue Salazar, Melissa Morabito, —Crime hotspot mapping using the crime related factors, DOI: dl.acm.org/doi/10.1007/s10489-012-0400-x

[6] Shiju Sathyadevan, Surya S, —Crime analysis and prediction using data mining, First International Conference on Networks & Soft Computing (ICNSC2014), 2014, DOI: 10.1109/cnsc.2014.754321

[7] Sathyadevan, S., & Gangadharan, S. (2014, August). Crime analysis and prediction using data mining. In Networks & Soft Computing (ICNSC), 2014 First International Conference on (pp. 406- 412). IEEE, DOI: 10.1109/CNSC.2014.6906719

[8] Tayebi, M. A., Gla, U., & Brantingham, P. L. (2015, May). Learning where to inspect: location learning for crime prediction. In Intelligence and Security Informatics (ISI), 2015 IEEE International Conference on (pp. 25-30). IEEE, DOI: academia.edu/37571766/IRJET

[9] P. Thongtae and S. Srisuk, "An analysis of data mining applications in crime domain", IEEE 8th International Conference on Computer and IT Workshops, 2008, DOI: 10.1109/CIT.2008.Workshops.80

[10] Yadav, S., Timbadia, A., Vishwakarma, R., & Yadav, N. Crime pattern detection, analysis & prediction. In Electronics, Communication and Aerospace Technology (ICECA),2017 International conference of (Vol. 1, pp. 225-230). IEEE, DOI: 10.1109/iceca.2017.8203676

[11] 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', CoRR, vol. 14092983, 2014, DOI: 10.1109/iceca.2017.8203575