Analysis Of Crime Patterns And Prediction

Mayuri M. Menkudle, Mrs . R.S.Potpelwar

Dept. of Information Technology

Shri Guru Gobind Singhji Institute of Engineering and Technology

431606, Nanded (MH)

--***---

Abstract - Analysis of crime data aims to the extraction of patterns of crimes and trends of crimes from previously stored data of crimes. Analysis of crime data plays an important role to prevent crimes that occur in the country and simultaneously that increase the security of people. The paper aims to seek out a frequent occurrence of crime patterns with knowledge discovery and its prediction. We have used machine learning techniques on criminal records for knowledge discovery and to assist in increasing the predictive accuracy of the crime. This work is going to be helpful to the local police stations in crime suppression.

*Key Words***:** Crime Pattern, Big data analytics, Machine Learning, Data Mining, Prediction.

1.INTRODUCTION

In recent years, due to continuous growth of population and urbanization cities play an important main role in our society. So that there are more violent accidents and incident of crimes are occurring. To solve such incidents, sociologists, analysts, and security institutions have put a lot of effort into mining potential patterns and factors [2]. To deal with such a large amount of heterogeneous data in several metros may be a challenging task, so big data analytics are used. Big data analytics (BDA) has become an developing methodology for extracting knowledge of data and analyzing their relationship over a wide range of application areas [3]. Consequently, new methods and technologies are required to analyze these heterogeneous and multisource data. Big data analysis provides us to effectively monitor events, identify similarities from events, resource deployments and to take decisions accordingly [4].

This also help us to understand both current conditions and historical issues, eventually increasing better safety/security and quality of life, simultaneously increasing growth of economic and culture. Crime rates accelerate continuously and patterns of crime change continuously so that the detection of behaviors of crime patterns is difficult. The mostly dataset of crime incidents includes following some features attributes: IncidentNum, Dates, Address, Descript, PdDistrict, DayOfWeek, arrest or booked, Category, X, Y, Coordinate, Resolution, Dome, Arrest.

The most basic techniques of BDA, machine learning is knowledge base, an innovative and developing exploration zone, which may build paradigms and techniques across varied fields for extracting hidden patterns from knowledge and deducing helpful information from dataset [5].. With the help of such systems, big data analytics can help us effectively distinguish particular crime patterns that happen in a specific zone and how they are linked with the time. Machine learning is an application of computing (AI) that provides systems the power to mechanically learn and improve from expertise while not being fully programmed. Machine learning is further categorized into three methods as supervised, unsupervised, semi-supervised machine learning. The supervised machine learning is technique in which we train the machine by using labeled data. In unsupervised machine learning there is no need of any well labeled dataset.

The implications of statistical techniques and machine learning on big data applications or crime like time-series data or traffic accidents will enable the analysis, extraction, and trends and understanding of associated patterns, ultimately helping us in prevention and management of crimes. For extracting pattern of crime, data should collect from government crime data sites and after collecting data we sorted and integrate actual data from raw data with preprocessing technique after that we use various types of data mining or machine learning technique for analyzing the data and by analyzing data we can easily predict the pattern of crimes and rate of crime which occurs mostly. The systematical flow of the project is shown in figure 1.

Mining of data is utilized with the end goal of the revelation of new information and furthermore for improving our comprehension. The rates of crimes and patterns increase day by day constantly so it is difficult to clarify behaviors of a particular crime. The aim is to produce a comprehensive review of theory and research concerning the prevention of the crime within the society and to implement different data analysis algorithms that address the connections between crime and its pattern. In this paper, state-of-the-art machine learning and big data analytics algorithms are utilized for the mining of crime data. First of all, we collect the crime data from three US cities, i.e. San-Francisco, Chicago and Philadelphia. After preprocessing, including data altering and normalization, we extracted various patterns of crimes. There are various types of crimes such as Theft, Robbery, Assault, Fraud, Drugs. In various cities different kinds of crime occur frequently. As shown in fig.2.
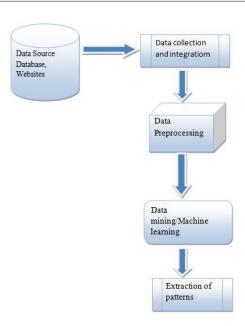
Fig 1: **System Flow Diagram**

Various techniques of time series modeling and machine learning are utilized for analyzing future trends. The most important contribution of this paper is often summarized as follows: 1)A series of investigatory exploration is used to explore crime data. 2) We used various methods of machine learning and data mining to analyze and find crime patterns of incidents. The remaining part of the paper is organized as follows. A previous work explained related work done based on this paper in Section 2, the various techniques used for analyzing and extraction of crime patterns explained section 3 followed by some remarkable conclusion summarized in section 4.
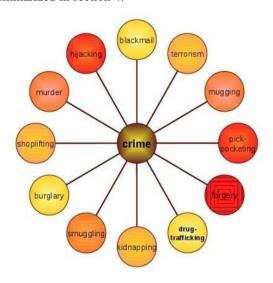

Fig. 2. **Types of crimes**

II. PREVIOUS WORK

Crime prevention and detection are important in order to detect crimes. Various studies have discovered several techniques in order to detect crimes, and that used by many applications. Such studies can help speed up the process of solving crime and help the computerized systems detect the criminals automatically. In addition, the rapidly advancing technologies can help address such issues. Data mining is the analysis process used to analyze the historical data to find trends, patterns and knowledge. Thongsatapornwatana[5] reviewed, data mining techniques such as association Rule Mining is an unsupervised learning method that used to find the hidden bits of knowledge in unlabeled data. It is used to solve the issues if the learners get the unlabeled example data. Clustering method is used to analyze data. In this technique we divide the same data into the same group and the different data into the other group. For extraction patterns of crimes various classification algorithm such as Decision Tree, Nearest Neighbor, Neural Network. As tested result, system can classify and predict the crimes more than 90% accuracy.

Yadav, Timbadia, Vishwakarma[7] uses various tools like Weka tool and R tool for prediction of crimes. Weka is a tool which consists of collection of various kinds machine learning algorithms for mining the data. R is a language which is used for graphics and gives environment for computing statistics. The dataset used in this paper depends on number of crimes occur in the particular area with number of criminal arrested along with many other features. Four algorithms implemented in the paper with the help of this Tool: Association Mining (Apriori) algorithm is used to find patterns, forecasts, possible explanations, trends, relationships, mapping criminal networks and identifying suspects. K-means is used to make several groups corresponding to higher and lower values. One of the classic method of data mining technique is the classification which is used to classify each object into a set of predefined classes or groups with the help of a set and regression, the number of people we can acquit. Offense on the number of trials completed during the year. This gives 0.98 (strong correlation) between trials completed during the year and the person convicted.

The paper [8] described the development of a computer-based system for building hazardous surfaces with sufficient efficiency and time efficiency to be used by Chilean police in their field operations. We take the crime data of the big cities of Chile. This approach consists of three software tool which is independent to each other that make predictions based on various algorithms. The ultimate assessment is the collaborative integration of the individual. The developed system was tested on previous (historical) data and checks performance of system and this performance is considered for police field. The accuracy of this paper is over 50%. The paper used Dempster Shaffer Theory Method, a multi-kernel method for predicting high-risk offenses in the form of grids and estimating the natural interval of the risk attendant risk surface.

In this paper, we exploit spatiotemporal models of urban data in a borough in a city and then influence transfer learning practices to reinforce the crime

assessment of other boroughs. In particular, we first confirm the existence of spatiotemporal patterns in urban crime. We then extract crime-related features from cross-domain datasets. Finally, we propose a novel transfer learning framework to integrate these features and model spatiotemporal models for crime assessment Wu, Dragut described the ability to include geographic information systems with crime log data. The use of visuals, which include hot charts and charts, is less difficult to recognize than undeniable textual content or numerical facts whilst representing crime patterns in this region and analyzing early crime styles. They automatically collect and analyze crime data from available resources and manage and search for mining. Google Maps is used to provide data geographically Vineet and Ayush Pandey [11] have often used the FP Max algorithm to create crime sets by building an FP tree. The Crime Intensity Point (CIP) states the rate of crime related with a particular state. The training dataset of CIP values with the labeled format is fed to the random forest for classification rules and gives accuracy 97.62%. Rodríguez, Gomez, Rey[12] explained a way to forecast spatiotemporal patterns of criminal activity, through a completely unique time series method from fuzzy clustering, A memetic algorithm is introduced a problem-oriented fitness function as well as execution of the series forecast. Joshi, Sabitha, Choudhury[15] presented a Rapid miner software tool that gives an suitable environment for machine learning and used for mining and includes results as representations, validations, and optimizations. With the help of k-means we calculate crime rates of each type of crimes and cities with high crime rates. A decision support system(DSS) with association rules(AR) to determine the factor and effect for specific patterns. In this, we use confidence and support value to find what kind of crime occurs most frequently. If value of confidence is more we can predict, that type of crime is occurring more frequently.

III. TECHNIQUES FOR ANALYZING CRIME PATTERNS

In this section, we present a overview of the various data mining and machine learning techniques used for analyzing and extraction of crime. Spatiotemporal pattern technique is used for finding crimes in intra and inter regions. In the intra region if two adjacent time slots are the same then they share the same crime number and if the distance between timeslot is increased then they share different crime numbers. For a different region, we considered it as inter-region. In the inter region, two close spatially regions gives a similar crime number otherwise it takes different crime number. Prediction model is used for solving the problems of crime which are related to time series. Apripri algorithm is used for finding frequent criminal activity. FP max is used to find frequent crimes that occur in a city. K-mean is a clustering algorithm used for creation of clusters in order to tackle high or low numbers of persons involved in criminal activity. Naive

bayes is a classification algorithm used for predicting the crimes. Overview is shown in fig.3.

Techniques	Tasks
Spatio-temporal pattern technique	Used for crime prediction.
Predictive Model	To tackle the problems of forecasting time series model of crime.
Apriori Algorithm	Apriori algorithm is a type of association mining, used to find frequent crime pattern.
FP Max	Gives maximum frequent crime pattern set which has better performance.
Clustering	K-mean is unsupervised machine learning which is used to cluster same groups of crimes in various cluster to analyze crime pattern.
Classification	Naive Bayes is supervised machine learning used for prediction of the number of crime committed by a particular age group.

Fig. 3: **Summary of various techniques for analyzing crime pattern**

IV.CONCLUSION

In this paper, we studied various kinds of crimes and which kinds of crimes occur frequently and big data analytics (BDA) techniques were used to analyze crime big data and various kinds of techniques of data mining and machine learning used to extract patterns of crimes. We compared variously supervised and unsupervised techniques to

decide which technique is better to predict patterns of crime. We can say the supervised technique gives more accuracy than the unsupervised technique.

REFERENCES

1) M. Feng, J. Zheng, J. Ren, A. Hussain, X. Li, Y. Xi, and Q. Liu, "Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data," *IEEE Access*,2019

2) Hassani, X. Huang, M. Ghodsi, and E. S. Silva, "A review of data mining applications in crime," Stat. Anal. Data Mining, ASA Data Sci. J., vol. 9, no. 3, pp. 139154, Apr. 2016.

3) A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," Int. J. Inf. Manage., vol. 35, no. 2, pp. 137144, Apr. 2015.

4) A. Agresti, An Introduction to Categorical Data Analysis, 3rd ed. Hoboken, NJ, USA: Wiley, 2018.

5) U. Thongsatapornwatana, "A survey of data mining techniques for analyzing crime patterns," in Proc. 2nd Asian Conf. Defence Technol., Chiang Mai, Thailand, 2016, pp. 123128.

6) S.Wang, X.Wang, P. Ye, Y. Yuan, S. Liu, and F.-Y.Wang, "Parallel crime scene analysis based on ACP approach," IEEE Trans. Computat. Social Syst., vol. 5, no. 1, pp. 244255, Mar. 2018.

7) S. Yadav, A. Yadav, R. Vishwakarma, N. Yadav, and Timbadia, "Crime pattern detection, analysis & prediction," in Proc. IEEE Int. Conf. Electron., Commun. Aerosp. Technol., Coimbatore, India, Apr. 2017, pp. 225230.

8) N. Baloian, C. E. Bassaletti, M. Fernández, O. Figueroa, Fuentes, R. Manasevich, M. Orchard, S. Peñael, J. A. Pino, and M. Vergara, "Crime prediction using patterns and context," in Proc. 21st IEEE Int. Conf. Comput. Supported Cooperat. Work Design, Wellington, New Zealand, Apr. 2017, pp. 29.

9) X. Zhao and J. Tang, "Exploring transfer learning for crime prediction," in Proc. IEEE Int. Conf. Data Mining Workshops, New Orleans, LA, USA, Nov. 2017, pp. 11581159.

10) S. Wu, J. Male, and E. Dragut, "Spatial-temporal campus crime pattern mining from historical alert messages,"in Proc. Int. Conf. Comput.,Netw. Commun., Santa Clara, CA, USA, 2017, pp. 778782.

11) K. R. S. Vineeth, T. Pradhan, and A. Pandey, "A novel approach for intelligent crime pattern discovery and prediction," in Proc. Int. Conf. Adv. Commun. Control Comput. Technol., manathapuram, India, 2016, pp. 531538.

12) C. R. Rodríguez, D. M. Gomez, and M.A. M. Rey, "Forecasting time series from clustering by a memetic differential fuzzy approach: An application to crime prediction," in Proc. IEEE Symp. Ser. Comput. Intell., Honolulu, HI, USA, Nov./Dec. 2017, pp. 18.

13) Y.gao, Y. Xia, J. Qiao, and S. Wu, "Solution to gang crime based on graph theory and analytical hierarchy process," Neurocomputing, vol. 140, pp. 121127, Sep. 2014.

14) A. Joshi, A. S. Sabitha, and T. Choudhury, "Crime analysis using K-means clustering," in Proc. 3rd Int. Conf. Comput. Intell. Netw., Odisha, India, 2017

15) M. Lofstrom and S. Raphael, "Crime, the Criminal Justice System, and Socioeconomic Inequality," J. Econ. Perspect., vol. 30, no. 2, pp. 26103, Mar. 2016.

BIOGRAPHIES

Menkudle Mayuri M. is currently pursuing masters degree program in SGGS College, Nanded(MH), India, PH-8329366988.

Email:mmmenkudle@gmail.com/2018mit002@sggs.ac.in

*R.S.Potpelwar is Associate professor in SGGS College, Nanded(MH), India, PH-9970284506.
E-mail: rspotpelwar@sggs.ac.in*