International Conference on Modern Trends in Engineering Science and Technology
© 2019 IJSRSET | Volume 5 | Issue 5 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

A Review on Crime Rate Analysis Using Data Mining

Akansha A Chikhale[1], Ankita K Dhavale[1], Aparna P Thakre[1], Diksha B Herode[1], Nikita D Nasre[1], Pracheta D Patrikar[1], Prof. Milind Tote[2]

[1]BE, Department of Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur, Maharashtra, India

[2]Assistant Ptofessor, Department of Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur, Maharashtra, India

ABSTRACT

Data Mining is the system which comprises computing and breaking down expansive previous databases so as to deliver new data which might be required to the association. By utilizing the current datasets anticipated extraction of new data. In data digging numerous methodologies for examination and expectation had been performed. Be that as it may, in the field of criminology a numerous couple of endeavors has made. Numerous few have taken endeavors for looking at the data every one of these methodologies produces. The police headquarters and other criminal equity organizations hold numerous extensive databases of data which can be used to foresee or examine the criminal developments and criminal movement support in the general public. For recognizing and breaking down examples and patterns in crime, crime investigation is a deliberate methodology. Our framework can foresee locales which have a high likelihood for crime event and can picture crime inclined zones. With the expanding approach of mechanized frameworks, crime data examiners can help the Law implementation officers to accelerate the way toward understanding crimes. Using the idea of data mining we can remove already obscure, valuable data from an unstructured data. Here we have a methodology between software engineering and criminal equity to build up a data mining strategy that can comprehend crimes exceptionally quick. Rather than concentrating on reasons for crime event like the criminal foundation of the wrongdoer, political hatred and so forth we are concentrating chiefly on crime components of every day.

Keywords : Data Mining Techniques, Crime Analysis, Clustering Algorithms, Bisect K-means, and Feature Selection.

I. INTRODUCTION

The crime rate is increasing considerably day by day. Crime cannot be predicted because it is neither systematic nor random. Likewise, the cutting edge advancements and hi-tech strategies help culprits in accomplishing their wrongdoings. As per Crime Records Bureau wrongdoings like thievery, pyromania and so on have been diminished while violations like homicide, sex misuse, assault and so on have been expanded. Verifiably illuminating violations has been the privilege of the criminal equity and law authorization experts. With the expansion in the utilization of the mechanized frameworks to follow violations and follow lawbreakers, PC information examiners have begun loaning their hands in helping the law implementation officers and investigators to accelerate the way toward illuminating wrongdoings.

Criminology is process that is utilized to identify crime and criminal characteristics. The offenders and the wrongdoing event probability can be evaluated with the assistance of criminology methods. The criminology helps the police division, the analyst offices and wrongdoing branches in distinguishing the genuine attributes of a criminal. The criminology department has been utilized in the proceedings of crime tracking ever since 1800. Crimes are a social nuisance and cost our society dearly in several ways. Even, the Indian Government has taken steps to develop applications and software for the use of State and Central Police in relation with the National Crime Records Bureau (NCRB) [27]. Any research that can help in solving crimes faster will pay for itself. About 10% of the criminals commit about 50% of the crimes [15]. People who study criminology will be able to identify the criminals based on the traces, characteristics and methods of crime which can be collected from the crime scene. In the middle of 1990s, data mining came into existence as a strong tool to extract useful information from large datasets and find the relationship between the attributes of the data [11]. Data mining(DM) originally came from statistics and machine learning(ML) as an interdisciplinary field, but then it was grown a lot that in 2001 it was considered as one of the top 10 leading technologies which will change the world [12]. According to many researchers such as Nath [23], solving crimes is very critical and time consuming task that requires human intelligence and experience and data mining is one technique that can help us with crime detection issues. For solving crimes faster, we have to develop a DM paradigm that performs an interdisciplinary approach between computer science and criminal justice. As said earlier, the Criminology is a procedure that intends to distinguish wrongdoing attributes and it is a standout amongst the most essential fields for applying DM. By utilizing this, DM calculations will probably deliver wrongdoing reports and help in the distinguishing proof of crooks a lot quicker than any human could. As a result of this noteworthy component, there is a developing interest for DM in criminology.

It is only within the last few decades that the technology made spatial DM a practical solution for wide audiences of Law enforcement officials which is affordable and available. Since the availability of criminal data or records is limited we are collecting crime data from various sources like web sites, news sites, blogs, social media, RSS feeds etc.

The inspiration for continuing with this review work is to help some assistance to the youthful analysts who are playing out their examination in criminal investigation and wrongdoing expectation regions. The paper is sorted out in such a way to give experiences about the wrongdoing examination method and afterward produce diverse kinds of wrongdoing investigation activities and those which can be connected together to deliver an end client item which can be connected to the wrongdoing examination in any police headquarters and analyst organizations. This work will be a significant reference to the individuals who go before their examination work in the wrongdoing investigation and Crime expectation utilizing DM methods. Additionally, nature of violations changes after some time, so as to probably identify more up to date and obscure examples in future, bunching procedures work better.

There are steps in doing Crime Analysis:

1) Data Collection
2) Clustering
3) Pattern Identification
4) Prediction
5) Visualization

II. LITERATURE REVIEW

Sharma [1] proposed an idea which portrays zero crime in the general public. For recognizing the suspicious criminal exercises, he has focused on the significance of data mining innovation and planned a proactive application for that reason. In his paper, he proposed an apparatus which applies an upgraded Decision Tree Algorithm to distinguish the suspicious messages about the criminal exercises. An improved ID3 Algorithm with an upgraded highlight determination technique and property significance factor is connected to deliver a superior and quicker Decision Tree dependent on the data entropy which is unequivocally gotten from a progression of preparing data sets from a few classes. He proposed another calculation which is a mix of Advanced ID3 arrangement calculation and upgraded include choice technique for the better productivity of the calculation.

Hamdy et al. [8] portrayed a methodology dependent on the general population's association with interpersonal organizations and versatile use, for example, area markers and call logs. Their work additionally presented a model for identifying suspicious conduct dependent on interpersonal organization feeds and it not just portrays another strategy utilizing the social association of individuals in any case, their work proposes another framework to help crime examination make quicker and exact choices. The suspicious development of the substance can be resolved to utilize the arrangement of induction rules. Their built model can foresee and describe human conduct from reality data sources

Bogahawatte and Adhikari [2] proposed a methodology in which they featured the utilization of data mining procedures, bunching, and order for powerful examination of crimes and criminal ID by building up a framework named Intelligent Crime Investigation System (ICSIS) that could distinguish a criminal put together up with respect to the proof gathered from the crime area. They utilized bunching to recognize the crime designs which are utilized to carry out crimes knowing the way that every crime has certain examples. The database is prepared with an administered learning calculation, Naïve Bayes to foresee conceivable suspects from the criminal records. His methodology incorporates building up a multi-specialist for crime design distinguishing proof. There are specialists for the spot, time, job trademark and substance of crooks which isolates the job of the offenders in segments. The framework is a multi-operator framework and made with oversaw Java Beans. It makes it simple to embody the asked for elements in the work into items and returns it to the bean for uncovering properties. Arranging the hoodlums/suspects depends on the Naïve Bayes classifier for recognizing most conceivable suspects from crime data. Grouping the offenders depends on the model to recognize examples of perpetrating crimes.

Agarwal et al. [3] utilized the fast digger device for examining the crime rates and expectation of crime rate utilizing diverse data mining methods. Their work done is for crime investigation utilizing the K-Means Clustering calculation. The principal goal of their crime examination work is to remove the crime designs, foresee the crime dependent on the spatial circulation of existing data and identification of crime. Their examination incorporates the following manslaughter crime rates starting with one year then onto the next

Kiani et al. [4] played out a crime examination work dependent on the bunching and order systems. Their work incorporates the extraction of crime designs by crime examination dependent on accessible criminal data, the forecast of crimes dependent on the spatial dissemination of existing data and crime acknowledgment. They proposed a model in which the

examination and expectation of crimes are done through the streamlining of anomaly discovery administrator parameters which are performed through the Genetic Algorithm. The highlights are weighted in this model and the low-esteem highlights were erased through choosing an appropriate edge. After which the groups are bunched by the k-implies grouping calculation for the arrangement of crime dataset.

Satyadevan et al. [5] have completed a work which will show high likelihood for crime event and can picture crime inclined regions. Rather than simply concentrating on the crime events, they are concentrating mostly on the crime elements of every day. They utilized the Naïve Bayes, Logistic Regression and SVM classifiers for characterization of crime examples and crime elements of every day. Their strategy comprises of an example distinguishing proof stage which can recognize the patterns and examples in crime utilizing the Apriori Algorithm. The expectation of crime spots is finished with the assistance of Decision Tree calculation which will identify the crime conceivable zones and their examples.

Bruin et al. [7] proposed a procedure which is utilized to decide the bunching of culprits dependent on the criminal professions. The criminal profile per offense every year is separated from the database and a profile remove is determined. From that point onward, the separation network in profile every year is made. The separation framework including the recurrence esteem is made to shape groups by utilizing innocent bunching calculation. They made a criminal profile which is built up in a method for speaking to the crime profile of a wrongdoer for a solitary year. With this data, the vast gathering of lawbreakers is effectively broke down and they anticipated the future conduct of individual suspects. It will be helpful for building up the reasonable picture on various existing kinds of criminal vocations They tried the apparatus on real Dutch National Criminal Record Database for removing the variables for recognizing the criminal professions of an individual.

Huang et al. [6] concentrated on an alternate methodology for criminal action expectation dependent on mining area based Social Network connections. By utilizing these associations, they can gather data utilizing the geological cooperation's and data accumulations from the general population. They formulated a working strategy in which a progression of highlights are sorted from the Foursquare and Gowalla utilized in the San Francisco Bay zone. The crime designs and the crime events are followed the geological highlights which are extricated from the guide and they are investigated to recognize the urban territories with high crime exercises. Their work goes for misusing the area based informal community data to examine the criminal exercises in urban territories. By utilizing the Haversine equation the separation between the two, for example, the crime area and the setting area is determined and appeared in the Google Maps API and OpenStreetMap.

Chen [19] have introduced a general system for crime data mining that draws on experience picked up with the Coplink venture with the specialists at Arizona and their work basically centers around demonstrating the connections between crime types and the connection between the criminal associations. They utilized an idea space approach which will separate criminal from the episode synopses.

Yu [20] have talked about the starter aftereffects of a crime estimating model created as a team with the police branch of a United States city in the Northeast. Their methodology is to planner datasets from unique

crime records. The datasets contain totaled tallies of crime and crime-related occasions arranged by the police division. The area and time of these occasions are inserted in the data. Extra spatial and worldly highlights are collected from the crude data set. Second, a group of data mining arrangement methods is utilized to play out crime estimating. At that point, they dissected an assortment of grouping techniques to figure out which is best to anticipate crime "hotspots". They even explored arrangement on-increment or rise. Last, they have proposed the best estimating methodology which is gone for accomplishing the most steady results.

Rizwan et al. [22] have performed characterization of crime dataset to foresee Crime Category for various conditions of the United States of America. The crime dataset that they utilized in this exploration is genuine in nature. That is, it was gathered from financial data from 1990 US Census, law requirement data from the 1990 US LEMAS review, and crime data from the 1995 FBI UCR. Their work looked at the two changed grouping calculations, in particular, Naïve Bayesian and Decision Tree for foreseeing Crime Category for various states in the USA. The outcomes from their investigation demonstrated that Decision Tree calculation outperformed Naïve Bayesian calculation and accomplished 83.9519% Accuracy in foreseeing Crime Category for various conditions of USA.

Donald [24] have proposed a framework for Crime Analysis which was named by them as The Regional Crime Analysis Program (ReCAP) framework. It was planned by them as a PC application intended to help neighborhood police powers (for Example College of Virginia (UVA), City of Charlottesville, and Albemarle County) in the examination and counteractive action of crime. ReCAP works in participation with the Pistol 2000 records the board framework, which accumulated and housed the majority of the crime data from a locale. Their innovative work was essentially centered on the individual segments of the framework which incorporates a database, geographic data framework (GIS), and data mining devices which comprised of data mining calculations which created spatial mining results over the crime hotspots. Their framework comprises of the consistent incorporation of the considerable number of parts in the framework.

III. CONCLUSION

Crime data is a sensitive and large domain and therefore we need some efficient clustering techniques and algorithms which will help the crime analysts and law enforcers retrieve the data and information and draw patterns and conclude to a result which will help their investigation. The partition clustering algorithm can be developed in such a way that solves the unsolved crimes faster. In this paper, we have worked on some known approaches for crime analysis and prediction concerned with data mining. Although many papers have been studied, only those papers with background in the crime prediction and criminal identification papers are compared with a theoretical study. Each paper has their own advantages and disadvantages. Each paper has its own individual approach for solving the crimes and criminal prediction. Here fast clustering technique is used for identification of crime and criminals which works on text based methods, crime patterns and crime evidence based methods, spatial and geo location based methods, communication based methods and provide the desired output. The data mining techniques studied from this can be applied for identifying the criminals in the society and also for providing a better future to live in.

IV. REFERENCES

[1] Mugdha Sharma, "Z-Crime: A Data Mining Tool for the Detection of Suspicious Criminal Activities based on the Decision Tree", International Conference on Data Mining and Intelligent Computing, pp. 1-6, 2014.

[2] Kaumalee Bogahawatte and Shalinda Adikari, "Intelligent Criminal Identification System", Proceedings of 8th IEEE International Conference on Computer Science and Education, pp. 633-638, 2013.

[3] Jyoti Agarwal, Renuka Nagpal and Rajni Sehgal, "Crime Analysis using K-Means Clustering", International Journal of Computer Applications, Vol. 83, No. 4, pp. 1-4, 2013.

[4] Rasoul Kiani, Siamak Mahdavi and Amin Keshavarzi, "Analysis and Prediction of Crimes by Clustering and Classification", International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No. 8, pp. 11-17, 2015.

[5] Shiju Sathyadevan, M.S. Devan and S. Surya Gangadharan, "Crime Analysis and Prediction using Data Mining", Proceedings of IEEE 1st International Conference on Networks and Soft Computing, pp. 406-412, 2014.

[6] Yu-Yueh Huang, Cheng-Te Li and Shyh-Kang Jeng, "Mining Location-based Social Networks for Criminal Activity Prediction", Proceedings of 24th IEEE International Conference on Wireless and Optical Communication, pp. 185-190, 2015.

[7] Jeroen S. De Bruin, Tim K. Cocx, Walter A. Kosters, Jeroen F. J. Laros and Joost N. Kok, "Data Mining Approaches to Criminal Career Analysis", Proceedings of 6th IEEE International Conference on Data Mining, pp. 1-7, 2006.

[8] Ehab Hamdy, Ammar Adl, Aboul Ella Hassanien, Osman Hegazy and Tai-Hoon Kim, "Criminal Act Detection and Identification Model", Proceedings of 7th International Conference on Advanced Communication and Networking, pp. 79-83, 2015.

[9] Kamal Taha and Paul D. Yoo, "SIIMCO: A Forensic Investigation Tool for Identifying the Influential Members of a Criminal Organization", IEEE Transactions on Information Forensics and Security, Vol. 11, No. 4, pp. 811-822, 2016.

[10] Kevin Sheehy et al., "Evidence-based Analysis of Mentally 111 Individuals in the Criminal Justice System", Proceedings of IEEE Systems and Information Engineering Design Symposium, pp. 250-254, 2016.

[11] David J. Hand, Heikki Mannila and Padhraic Smyth, "Principles of Data Mining", MIT Press, 2001.

[12] Emerging Technologies That Will Change Your World, Available at: http://www.rle.mit.edu/thz/documents/10_emerging_tech.pdf.

[13] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "The KDD Process for Extracting Useful Knowledge from Volumes of Data", Communications of the ACM, Vol. 39, No. 11, pp. 27-34, 1996.

[14] Illhoi Yoo, Patricia Alafaireet, Miroslav Marinov, Keila Pena-Hernandez, Rajitha Gopidi, Jia-Fu Chang and Lei Hua, "Data Mining in Healthcare and Biomedicine: A Survey of the Literature", Journal of Medical Systems, Vol. 36, No. 4, pp. 2431-2448, 2011.

[15] Shyam Varan Nath, "Crime Pattern Detection using Data Mining", Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops, pp. 1-4, 2006.

[16] Hsinchun Chen, Wingyan Chung, Yi Qin, Michael Chau, Jennifer Jie Xu, Gang Wang, Rong Zheng and Homa Atabakhsh, "Crime Data

Mining: An Overview and Case Studies", Proceedings National Conference on Digital Government Research, pp. 1-5, 2003.

[17] Tong Wang et al., "Learning to Detect Patterns of Crime", Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 515-530, 2013.

[18] Karl F. Schuessler and Donald R. Cressey, "Personality Characteristics of Criminals", American Journal of Sociology, Vol. 55, No. 5, pp. 476-484, 1950.

[19] H. Chen, W. Chung, J.J. Xu, G. Wang, Y. Qin and M. Chau, "Crime Data Mining: a General Framework and Some Examples", Computer, Vol. 37, No. 4, pp. 50-56, 2004.

[20] Chung-Hsien Yu, Max W. Ward, Melissa Morabito and Wei Ding, "Crime Forecasting using Data Mining Techniques", Proceedings of 11th IEEE International Conference on Data Mining Workshops, pp. 779-786, 2011.

[21] P. Thongtae and S. Srisuk, "An Analysis of Data Mining Applications in Crime Domain", Proceedings of IEEE 8th International Conference on Computer and Information Technology Workshops, pp. 122-126, 2008.

[22] Rizwan Iqbal, Masrah Azrifah Azmi Murad, Aida Mustapha, Payam Hassany Shariat Panahy and Nasim Khanahmadliravi, "An Experimental Study of Classification Algorithms for Crime Prediction", Indian Journal of Science and Technology, Vol. 6, No. 3, pp. 4219-4225, 2013.

[23] Shyam Varan Nath, "Crime Data Mining", Proceedings of Advances and Innovations in Systems, Computing Sciences and Software Engineering, pp. 405-409, 2007.

[24] Donald E.Brown, "The Regional Crime Analysis Program (RECAP): A Framework for Mining Data to Catch Criminals", Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 2848-2853, 1998.

[25] Colleen McCue, "Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis", Butterworth-Heinemann, 2014.

[26] Arunima S. Kumar and Raju K. Gopal, "Data Mining based Crime Investigation Systems: Taxonomy and Relevance", Proceedings of Global Conference on IEEE Communication Technologies, pp. 850-853, 2015.

[27] Manish Gupta, B. Chandra and M.P. Gupta, "Crime Data Mining for Indian Police Information System", Journal of Crime, Vol. 2, No. 6, pp. 43-54, 2006.