# An Optimization Technique Used for Crime Recognition With Fuzzy Clustering

Chhaya Narwariya

M.E. (CSE),

Maharana Pratap College of Technology Gwalior,

Dr. Shivnath Ghosh

Associate Professor, Computer Science & Engineering

Maharana Pratap College Of Technology Gwalior

*Abstract*— Crime analysis can happen at different levels, including strategic, operational, and key. Crime analysts study crime reports, captures reports, and police calls for service to distinguish emerging patterns, series, and patterns as fast as could be expected under the circumstances. They analyze these wonders for every applicable factor, at times predict or forecast future events, and issue announcements, reports, and cautions to their organizations. In this way, a cluster is made out of an arrangement of comparative data which act same as a gathering. One might say that the clustering is equivalent to the arrangement, with just contrast that the classes are not characterized and decided ahead of time, and clustering of the data is managed without supervision. The work also gives a new direction of application of the algorithm with certain limitations for incremental datasets where data are being generated regularly. The present methodology compares with some existing feature selection methods and clusters obtained in each method using same algorithm are evaluated which demonstrates that the proposed technique is more exact than the current methods.

*Keywords*— *Crime, Machine Learning, Clustering, K-Means Algorithm, Crime Rate.*

## I. INTRODUCTION

Crime is a billion-dollar business and it is expanding each year. The PwC global economic crime review of 2009 suggests that close to 30 percent of organizations worldwide have declared being casualties of crime in the earlier year. Crime includes at least one people who purposefully act secretly to deprive another of something of significant worth, for their own advantage. Fraud is as old as humanity itself and can take a unlimited wide range of structures. Be that as it may, lately, the improvement of new technologies has likewise given further manners by which criminals may commit crime [1]. Notwithstanding that, business reengineering, rearrangement or downsizing may weaken or eliminate control, while new data frameworks may show extra chances to commit crime. Crime analysis can happen at different levels, including strategic, operational, and key. Crime analysts study crime reports, captures reports, and police calls for service to distinguish emerging patterns, series, and patterns as fast as could be expected under the circumstances. They analyze these wonders for every applicable factor, at times predict or forecast future events, and issue announcements, reports, and cautions to their organizations [1]. They at that point work with their police agencies to create compelling methodologies and strategies to address crime and disorder. Different duties of crime analysts may incorporate getting ready insights, data inquiries, or maps on demand;  analyzing beat and move designs; getting ready data for gathering or court presentations; answering questions from the overall population and the press; and giving data and data support to a police department's CompStat method. To check whether a crime fits a particular known pattern or

another example is often tedious work of crime analysts, analysts or in police officers cops or representatives themselves. They should physically move however heaps of paperwork and confirmation to predict, expect and ideally prevent crime [2]. The U.S. Division of Justice and the National Institute of Justice as of late propelled activities to support "prescient policing", which is an observational, data driven approach. However this work to distinguish specific patterns of crime carried out by an individual or social occasion, (crime series), remains a manual errand. Over the earlier year, MIT doctoral student Tong Wang, Cambridge (Mass.) Police Department CPD Lieutenant Daniel Wagner, CPD crime analyst Rich Sevieri and Assoc. Prof. of Statistics at MIT Sloan School of Management and the co-maker of "Making sense of how to Detect Patterns of Crime" Cynthia Rudin have illustrated a machine learning methodology called " Series Finder " that can help police in finding crime series in a small amount of the time.  Series Finder grows an pattern of crime, starting from a seed of at least two crimes [2]. The Cambridge Police Department has one of the most seasoned crime analysis units on the world and their historical data was utilized to prepare Series Finder to perceive housebreak outlines. The algorithm tries to construct a modus operandi MO). The M.O. is an arrangement of habits for a criminal and is a sort of conduct used to describe an example. The data of the burglaries incorporate methods for section (front door, window, and so forth), day of the week, qualities of the property (apartment, house), and geographic proximity to other break-ins. Utilizing nine known crime series of robberies Series Finder recuperated the majority of the violations inside these examples and furthermore recognized nine additional crimes. Machine learning is a gigantic instrument for predictive policing. If cases are recognized the police can quickly attempt to stop them. Without such apparatuses it can take weeks and even a very long time of shifting however databases to find an pattern. Series Finder gives a critical data-driven way to deal with an extremely troublesome issue in predictive policing. It's the main scientifically principled way to deal with the automated learning of crime series [2].

## II. CLUSTERING

Clustering Division of an arrangement of data or articles to different clusters is called clustering. In this way, a cluster is made out of an arrangement of comparative data which act same as a gathering. One might say that the clustering is equivalent to the arrangement, with just contrast that the classes are not characterized and decided ahead of time, and clustering of the data is managed without supervision [3].

a) **Clustering by K-means Algorithm** K- means is the simplest and most normally utilized partitioning algorithm among the clustering algorithms in logical

and industrial software [4] [5] [6]. Acknowledgment of the K- means is for the most part because of its being simple. This algorithm is additionally appropriate for clustering of the substantial datasets since it has significantly less computational complexity, however this complexity develops linearly by expanding of the data points [6]. Alongside simplicity of this strategy, it however suffers a few weaknesses, for example, assurance of the quantity of clusters by user, affectability from outlier data, high-dimensional data, and affectability toward communities for initial clusters and therefore probability of being caught into local minimum may lessen effectiveness of the K-means algorithm [7].

### III. LITERATURE SURVEY

Ubon Thongsatapornwatana et al. [2017] this paper proposes the criminal behavior analysis strategy to detect suspect vehicles that are conceivably engaged with criminal movement. It must not depend on the blacklist. The analysis is restrictive on journey path and the contribution of criminal activities. In extra, public officials trust that the suspect vehicle will pick the excursion way without a checkpoint. In this way, we used the journey path analysis methodologies together with the association rule mining to analyze such criminal conduct. From extensive experiments, the results show that the proposed procedure can construct the speculate detection accuracy rate 17.24% past the conventional counterpart [8].

Priyanka Das et al. [2017] the present work proposes an effective methodology for analyzing crime against women in India. The work involves extracting crime reports from online newspaper articles and the documents consisting crime reports of various states and union territories of India are made to undergo several preprocessing techniques. Each document is treated as bag-of-words and finally an exhaustive list of words have been prepared. Similarity has been measured among the words for selecting relevant features for crime trend analysis. Based on the selected crime features, community detection algorithm has been applied for partitioning the states based on crime against women in India. It is a graph based clustering approach and all the states of India have been considered as nodes of the graph. Each community is a group of states which are similar based on crime trends. The work also gives a new direction of application of the algorithm with certain limitations for incremental datasets where data are being generated regularly. The present methodology compares with some existing feature selection methods and clusters obtained in each method using same algorithm are evaluated which demonstrates that the proposed technique is more exact than the current methods [9].

Sunil Yadav et al. [2017] in this paper, crimes are a social irritation and cost our general public deeply in a few ways. Any research that can help in solving crimes rapidly will pay for itself. Around 10% of the criminals perpetrate around 50% of the crimes. The system is trained by feeding previous years record of crimes taken from legitimate online portal of India listing various crimes such as murder, kidnapping and abduction, dacoits, robbery, burglary, rape and other such crimes. As per data of Indian statistics, which gives data of various crime of past 14 years (2001-2014) a regression model is created and the crime rate for the following years in various states can be predicted [8]. We have utilized regulated, semi-supervised and unsupervised learning strategy on the crime records for knowledge discovery and to help in expanding the prescient accuracy of the crime. This work will be helpful to the local police stations in crime suppression [10].

Andrew J. Stop et al. [2017] this paper introduces a structure called Phishing-Detective that identifies phishing sites in view of existing and newly discovered heuristics. For this framework, a web crawler was delivered to scrape the substance of phishing and legitimate sites. These substance were analyzed to rate the heuristics and their dedication scale factor towards the illegitimacy of a site. The data set gathered from Web Scraper was then analyzed utilizing a data mining apparatus to discover examples and report findings. A contextual analysis demonstrates how this structure can be utilized to identify a phishing website. This research is still in progress but shows a new way of finding and using heuristics and the sum of their contributing weights to effectively and accurately detect phishing websites. Further development of this framework is discussed at the end of the paper [11].

S. Sivaranjani et al. [2016] in this work, our utilization different clustering methodologies of data mining to investigate the crime data of Tamilnadu. The crime data is extracted from National Crime Records Bureau (NCRB) of India. It comprises of crime data around six cities namely Chennai, Coimbatore, Salem, Madurai, Thirunelvelli and Thiruchirapalli from the year 2000-2014 with 1760 examples and 9 attributes to represent the instances. K-Means clustering, Agglomerative clustering and Density Based Spatial Clustering with Noise (DBSCAN) algorithms are used to cluster crime activities in perspective of some predefined cases and the consequences of these clustering are contrasted with locate the best appropriate clustering algorithm for crime detection. The consequence of K-Means clustering algorithm is visualized utilizing Google Map for interactive and simple understanding. The K-Nearest Neighbor (KNN) classification is utilized for crime prediction. The execution of each clustering algorithms are assessed utilizing the measurements, for example, precision, recall and Fmeasure, and the outcomes are analyzed. This work helps the law enforcement agencies to predict and detect crimes in Tamilnadu with improved accuracy and thus reduces the crime rate [12].

Nidhi Tomar et al. [2016] in this paper, clustering is a most well known procedure in data mining and is a basic methodology that is performed in perspective of the similarity rule. The segregation of a major database is a stimulating and undertaking of time consuming. It finishes up two unique stages: in the first place, feature extraction maps all documents or record to a point in the space of high-dimensional, at that point algorithms for clustering consequently gathering the focuses into a cluster hierarchy. Clustering has different applications in various fields. Maybe a couple of the fields incorporate criminology, text mining, image res olution, machine learning. Crime detection has become one of the most attractive fields as the crime rate in India and whole world is increasing at a greater pace. We as citizens of a country have to contribute towards its detection and removal. Along these lines, a comprehensive review carried about the

nuts and bolts of clustering has given in this paper. Moreover, proposed work was given that gives the idea of the work going to be done in the upcoming time [13].

## IV. PROPOSED WORK

In the proposed work, different techniques are used for the recognition of crime in various cities. The table below show the cities used for the crime rate detection. Fuzzy C-Means and ACO techniques are used to improve the performance of detection.

Table 1: Name of Cities

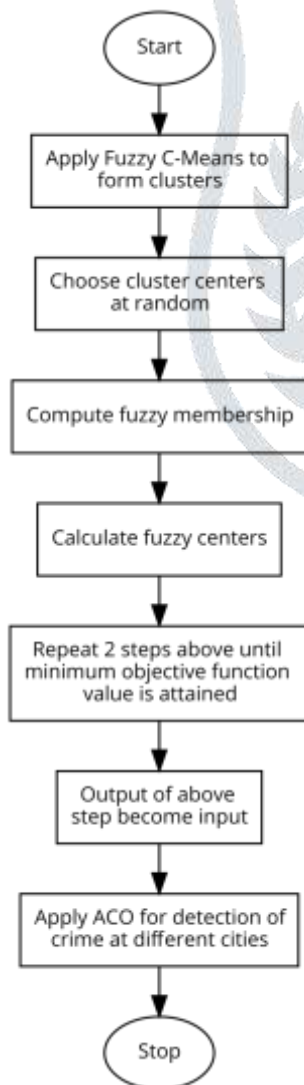| S.No. | Cities Name |
|-------|-------------|
| 1. | Agra |
| 2. | Bangalore |
| 3. | Chennai |
| 4. | Delhi |
| 5. | Hyderabad |
| 6. | Jaipur |
| 7. | Kanpur |
| 8. | Kolkata |
| 9. | Mumbai |
| 10. | Pune |



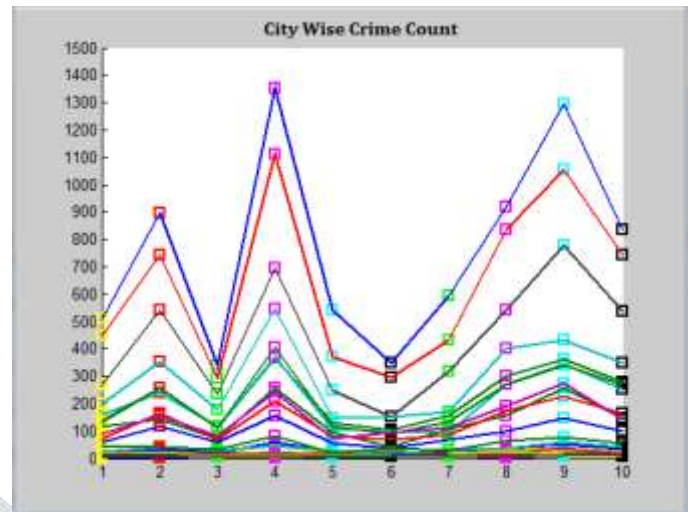Fig.1 Proposed Work Flowchart

## V. RESULT ANALYSIS



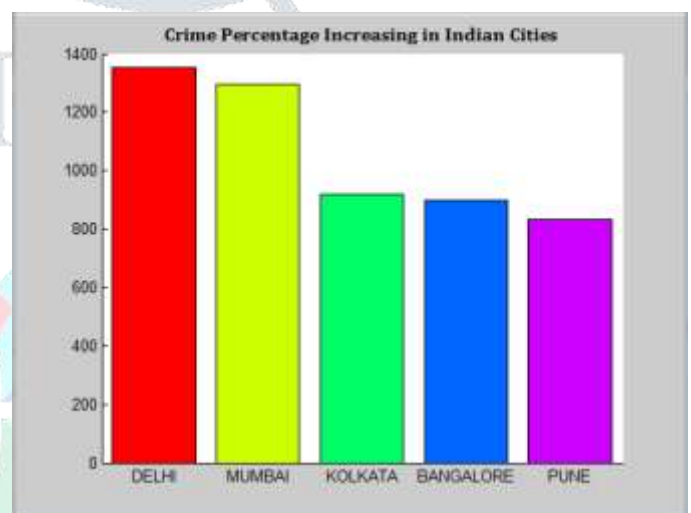Fig.2 Crime Recognition Rate of Different Cities



Fig.3 Bar Graph of Crime in Indian Cities

## Conclusion

Crime includes at least one people who purposefully act secretly to deprive another of something of significant worth, for their own advantage. Fraud is as old as humanity itself and can take a unlimited wide range of structures. Be that as it may, lately, the improvement of new technologies has likewise given further manners by which criminals may commit crime. The clustering algorithm is additionally appropriate for clustering of the substantial datasets since it has significantly less computational complexity; however this complexity develops linearly by expanding of the data points. The analysis is restrictive on journey path and the contribution of criminal activities. In extra, public officials trust that the suspect vehicle will pick the excursion way without a checkpoint. In this way, we used the journey path analysis methodologies together with the association rule mining to analyze such criminal conduct. From extensive experiments, the results show that the proposed procedure can construct the speculate detection accuracy rate more than the past of conventional counterpart

## REFERENCES

[1]    Mrs. B. Umamaheswari, Dr. P. Nithya, Miss.Nair Sarika Chandran "Survey on web crime detection using data mining technique"

International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5 Issue 1, January 2016.

[2] Hosseinkhani, J., Ibrahim, S., Chuprat, S. and Hosseinkhani, N.J. Web Crime Mining by Means of Data Mining Techniques, Research Journal of Applied Sciences, Engineering & Technology (Print ISSN: 2040-7459 Online ISSN: 2040-7467), 2013.

[3] J. Han, and M. Kamber, "Data mining: concepts and techniques" Jim Gray, Series Editor Morgan Kaufmann Publishers, August 2000.

[4] P. Berkhin, Survey of clustering data mining techniques, In: Accrue Software, 2003.

[5] W. Li "Modified k-means clustering algorithm, IEEE Congress on Image and Signal" Processing, pp. 616- 621, 2006.

[6] D.T Pham, S. Otri, A. Afifty, M. Mahmuddin, and H. Al-Jabbouli, "Data clustering using the Bees algorithm" proceedings of 40th CRIP International Manufacturing Systems Seminar, 2006

[7] J. Han, and M. Kamber, "Data mining: concepts and techniques," 2nd Edition, Morgan Kaufmann Publisher, 2001.

[8] Ubon Thongsatapornwatana, Woraphon Lilakiatsakun, Akkarach Kawbunjun and Tossapon Boongoen "Analysis of Criminal Behaviors for Suspect Vehicle Detection" The Twelfth International Conference on

Digital Information Management (ICDIM 2017) September 12- 14, 2017, Kyushu University, Fukuoka, Japan.

[9] Priyanka Das, Asit Kumar Das "Crime Analysis against Women from Online Newspaper Reports and an Approach to apply it in Dynamic Environment" 978-1-5090-6399-4/17/$31.00 c 2017 IEEE.

[10] Sunil Yadav, Meet Timbadia, Ajit Yadav, Rohit Vishwakarma and Nikhilesh Yadav "Crime Pattern Detection, Analysis & Prediction" International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.

[11] Andrew J. Park, Ruhi Naaz Quadari and Herbert H. Tsang "Phishing Website Detection Framework Through Web Scraping and Data Mining" 978-1-5386-3371-7/17/$31.00 ©2017 IEEE.

[12] S.Sivaranjani, Dr.S.Sivakumari and Aasha.M "Crime Prediction and Forecasting in Tamilnadu using Clustering Approaches" 2016 International Conference on Emerging Technological Trends [ICETT].

[13] Nidhi Tomar, Amit Kumar Manjhvar "An Improved Optimized Clustering Technique for Crime Detection" 2016 Symposium on Colossal Data Analysis and Networking (CDAN).