International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

CRIME ANALYSIS AND PREDICTION USING OPTIMIZED LOGISTIC REGRESSION ALGORITHM

K Banupriya[1], B Devendiran[2], P Divesh[3], R Sanjay[4], K Logeshwaran[5]

[1]Assistant Professor, Dept of Information Technology, Adhi College Of Engineering AndTechnology, [2345]Students of Dept of Computer Science And Engineering, Adhi College Of Engineering AndTechnology

[1]banupriya.cse@adhi.edu.in, [2]baskardeva7@gmail.com, [3]diveshk2468@gmail.com,[4]Sanjayr21092022@gmail.com, [5]logeshwaran2563@gmail.com

ABSTRACT

The objective of this project is to tackle a vital issue in the society - Crimes. Analyzing and examining of crimes happening in the world will give us a Broadview in understanding the crime regions and can be used to take necessary precautions to mitigate the crime rates. Identifying Crime patterns will allow us to tackle problems with unique approaches in specific crime category regions and improve more security measures in society. Current studies show the reason of increase in crime rates is more in areas that are economically backward. In few decades' property crime will be a target. The following approach involves predicting crimes classifying, pattern detection and visualization with effective tools and technologies. Use of past crime data trends helps us to correlate factors which might help understanding the future scope of crimes.
Keywords: Crime Prediction, K-Means, Clustering, Data Mining, Crime Prone Areas

I. INTRODUCTION

In present scenario criminals are becoming technologically sophisticated in committing crime and one challenge faced by intelligence and law enforcement agencies is difficulty in analysing large volume of data involved in crime and terrorist activities therefore agencies need to know technique to catch criminal and remain ahead in the eternal race between the criminals and the law enforcement. Crooks are irritation for the general public in all edges of world for quite a while now and measures are required to destroy crimes from our reality. Our main goal is to offer crime counteractive action application to keep open safe. Current policing techniques work towards finding the culprits, essentially after the crime has happened. Be that as it may, with the assistance of innovative headway, we can utilize notable crime data to perceive crime examples and utilize these examples to anticipate crimes already.

We are utilizing clustering calculations to anticipate crime-prone areas. There are many clustering calculations to aggregate the important data into wanted groups. The vast volumes of crime datasets just as the multifaceted nature of connections between these kinds of data have made criminology a fitting field for applying data mining systems. Criminology is a region that centers the logical investigation of crime and criminal conduct and law

implementation and is a procedure that intends to recognize crime qualities.

It is a standout amongst the most essential fields where the use of data mining procedures can deliver critical outcomes. Distinguishing crime qualities is the initial step for growing further examination. The knowledge picked up from data mining approaches is a valuable apparatus which can help and bolster police powers. Clustering systems change over dataset to groups which are additionally inspected for determining crime-prone areas. These bunches outwardly speak to the gathering of crimes overlaid on the guise of police ward. Groups store area of crimes alongside different accreditations of crime like sort and time. These bunches are characterized based on their individuals. Thickly populated bunches move toward becoming crime-prone areas while groups with fewer individuals are overlooked. Preventive measures are executed by crime type in crime-prone areas. K-means is the easiest and most ordinarily utilized clustering calculation in logical and modern programming. Because of less computational multifaceted nature, it is reasonable for clustering vast data sets. In that capacity, it has been effectively utilized in different points, including market division, PC vision, and geostatistics, stargazing, and farming. It frequently is utilized as a pre-processing venture for different calculations, for instance, to locate a beginning arrangement.

We picked clustering strategy over some other directed method, for example, arrangement, since crimes shift in nature generally and crime database are frequently loaded up with unsolved crimes. Accordingly, characterization procedure that will depend on the current and known fathomed crimes, won't give great prescient quality for future crimes.

II. RELATED WORK

Data mining in the examination and investigation of criminology can be ordered into primary areas, crime control and crime concealment. De Bruin et. al. [1] presented a framework for crime patterns utilizing another separation measure for looking at all people dependent on their profiles and after that clustering them as needs be.

Manish Gupta et. al. [2]. Features the current frameworks utilized by Indian police as e-administration activities and furthermore proposes an intuitive inquiry based interface as crime investigation apparatus to help police in their exercises. He proposed interface which is utilized to separate helpful data from the immense crime database kept up by National Crime Record Bureau (NCRB) and discover crime problem areas utilizing crime data mining procedures, for example, clustering and so forth. The adequacy of the proposed interface has been delineated on Indian crime records.

Nazlena Mohamad Ali et al.[3] examine an improvement of Visual Interactive Malaysia Crime News Retrieval System (I-JEN) and depict the methodology, client contemplates and arranged, the framework design and future arrangement. Their primary targets were to build crime-based occasion; research the utilization of crime based occasion in enhancing the characterization and clustering; build up an intuitive crime news recovery framework; envision crime news in a viable also, intelligent way; coordinate them into a usable and strong framework and assess the ease of use and framework execution and the examination will add to the better comprehension of the crime data utilization in the Malaysian setting just as the created framework with the perception highlights to address crime data and the inevitable objective of fighting the crimes.

Sutapat Thiprungsri [4] looks at the utilization of bunch examination in the bookkeeping area, especially error discovery in review. The motivation behind his investigation is to inspect the utilization of clustering innovation to mechanize

misrepresentation separating amid a review. He utilized bunch investigation to enable reviewers to center their endeavors while assessing bunch disaster protection claims.

A. Malathi et al.[5] look at the utilization of missing quality and clustering calculation for a data mining way to deal with assistance foresee the crimes examples and quick up the way toward understanding crime.

Malathi. An et. al.[6] utilized a clustering/arrange based model to foresee crime patterns. The data mining strategies are utilized to break down the city crime data from Police Department. The aftereffects of this data mining could possibly be utilized to reduce and even avert crime for the imminent years.

Dr. S. Santhosh Baboo and Malathi. A [7] look into work concentrated on building up a crime examination apparatus for Indian situation utilizing distinctive data mining strategies that can help law authorization division to proficiently deal with crime examination. The proposed apparatus empowers offices to effectively and financially perfect, describe and break down crime data to distinguish noteworthy examples and patterns.

Kadhim B. Swadi Al-Janabi [8] presents a proposed framework for the crime and criminal data investigation and recognition utilizing Decision tree Algorithms for data order and Simple K Means calculation for data clustering. The paper will in general help authorities in finding examples and patterns, making figures, discovering connections and conceivable clarifications, mapping criminal networks and distinguishing conceivable suspects.

Aravindan Mahendiran et al. [9] apply a bunch of instruments on crime datasets to dig for data that is avoided human observation. With the assistance of cutting edge representation procedures, we present the examples found through our calculations in a perfect and natural way that empowers law implementation offices to channelize their assets as needs are.

Sutapat Thiprungsri[10] look at the likelihood of utilizing clustering innovation for examining. Computerizing extortion separating can be of incredible incentive to nonstop reviews. The goal of their examination is to inspect the utilization of bunch investigation as an option and creative inconsistency location strategy in the wire exchange framework.

K. Zakir Hussain et al. [11] attempted endeavour to catch long periods of human experience into PC models through data mining and by structuring a reproduction show.

III. METHODOLOGY

After literature review there is need to used an open source data mining tool which can be implemented easily and analysis can be done easily. So here crime analysis is done on crime dataset by applying k means clustering algorithm.

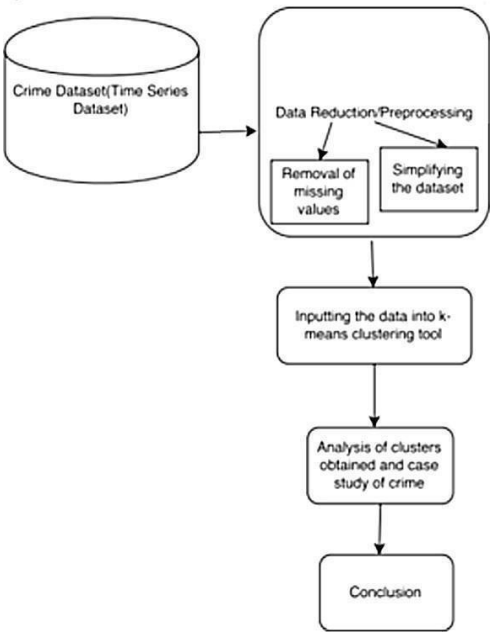

Figure 1. Flowchart for Proposed Methodology

The above flowchart depicts the methodology followed for the analysis of time series crime data set.

1. Collection of Dataset from the Bureau of Crime Statistics and Research of New South Wales Government's Justice Department Website.
2. Preprocessing Of Dataset (Data Cleaning, Data Selection, Data Transformation)
3. Analysis Of K-Means Using Clustering Tool A. Identification Of K Using Silhouette Measure.
4. Inputting the Data into K-Means Clustering.
5. Cluster 0 To Cluster 4 Obtained Using K-Means
6. Analysis Of Clusters Obtained Using K-Means And Case Study Of Crime At Various Locations.

IV. IMPLEMENTATION

4.1 Dataset Used

Crime dataset used for crime analysis is an offences recorded by the police in England and Wales by offence and police force area from 2011 to 2015.

4.2 Algorithm

K-means clustering is one of the method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

Process:

1. Initially, the number of clusters must be known let it be k
2. The initial step is to choose a set of K instances as centers of the clusters.
3. Next, the algorithm considers each instance and assigns it to the cluster which is closest.
4. The cluster centroids are recalculated either after whole cycle of re-assignment or each instance assignment.
5. This process is iterated.

K means algorithm complexity is O(kn), where n is instances, c is clusters, and t is iterations and relatively efficient . It often terminates at a local optimum. Its disadvantage is applicable only when mean is defined and need to specify c, the number of clusters, in advance. It unable to handle noisy data and outliers and not suitable to discover clusters with non-convex shapes.

V. CONCLUSIONS

This project focuses on crime analysis by implementing clustering algorithm on crime dataset and here we do crime analysis. From the clustered results it is easy to identify crime trend over years and can be used to design precaution methods for future. From the encouraging results, we believe that crime data mining has a promising future for increasin the effectiveness and efficiency of criminal and intelligence analysis. Visual and intuitive criminal and intelligence investigation techniques can be developed for crime pattern. As we have applied clustering technique of data mining for crime analysis we can also perform other techniques of data mining such as classification. Also we can perform analysis on various dataset such as enterprise survey dataset, poverty dataset, aid effectiveness dataset, etc.

VI. REFERENCES

[1]. De Bruin, J. S.,Cocx,T.K,Kosters,W.A.,Laros,J. and Kok,J.N(2006) Data mining approaches to criminal carrer analysis ,"in Proceedings of the Sixth International Conference on Data Mining (ICDM"06) ,Pp. 171-177

[2]. Manish Gupta1*, B.Chandra1 and M. P. Gupta1,2007 Crime Data Mining for Indian Police Information System

[3]. Nazlena Mohamad Ali1, Masnizah Mohd2, Hyowon Lee3, Alan F. Smeaton3, Fabio Crestani4 and Shahrul Azman Mohd Noah2 ,2010 Visual Interactive Malaysia Crime News Retrieval System

[4]. Sutapat Thirprungsri Rutgers University .USA ,2011 Cluster Analysis of Anomaly Detection in Accounting Data : An Audit Approach

[5]. A.Malathi, Dr. S. Santhosh Baboo. D. G. Vaishnav College, Chennai ,2011 Algorithmic Crime Prediction Model Based on the Analysis of Crime Clusters.

[6]. Malathi., Dr.S.Santhosh Baboo 2 and Anbarasi . A 31 Assistant professor, Department of Computer Science, Govt Arts College, Coimbatore, India. 2 Readers, Department of Computer science , D.G. Vaishnav Collge ,Chennai , India , 2011 An intelligent Analysis of a city Crime Data Using Data Mining

[7]. Malathi , A; Santhosh Baboo , S, 2011 An Enhanced Algorithm to Predict a Future Crime using Data Mining

[8]. Kadhim B.Swadi al-Janabi . Department of Computer Science . Faculty of Mathematics and Computer Science .University of Kufa/Iraq , 2011 A Proposed Framework for Analyzing Crime DataSet using Decision Tree and Simple K-means Mining Algorithms.

[9]. Aravindan Mahendiran, Michael Shuffett, Sathappan Muthiah, Rimy Malla, Gaoqiang Zhang,2011 Forecasting Crime Incidents using Cluster Analysis and Bayesian Belief Networks

[10]. Sutapat Thiprungsri,2012 Cluster Analysis for Anomaly Detection in Accounting Data : An Audit Approach, K. Zakir Hussain, M. Durairaj and G. Rabia Jahani

[11]. Farzana ,2012 Application of Data Mining Techniques for Analyzing Violent Criminal Behavior by Simulation Model

Cite this article as :

Akansha A Chikhale, Ankita K Dhavale, Aparna P Thakre, Diksha B Herode, Nikita D Nasre, Pracheta D Patrikar, Prof. Milind Tote , "Crime Rate Prediction using K-means Algorithm", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 6 Issue 1, pp. 313-317, January-February 2019.

Journal URL : http://ijsrst.com/IJSRST196153