



A Review on Crime Analysis and Prediction Using Dbscan Algorithm

Trupti Thakre, Roshani Sharma, Snehal Pote, Avisha Wankhede, Nikita Meshram

Department of Computer Science and Engineering, Babasaheb Ambedkar College of Engineering, Nagpur University, Maharashtra, India

ABSTRACT

Crime is one of the biggest violations that have been not yet completely solved ever since the evolution of human race. The threat will be reduced if a crime prediction and analysis is concerned on particular areas to find if crime is about to happen or not. The existing system is having trouble about data access speed and it is less efficient so to overcome this problem, we proposed a system in which analysis of crime is done also we use DBSCAN algorithm to find out different cluster of crimes. It shows high accuracy for the given dataset and forms effective cluster. Data mining is an approach that can handle large voluminous datasets and used to predict desired patterns. This aspect will be beneficial for both law enforcement and police organization of our country and to give more correct decision also help in safeguarding an area. This application is useful for police and enforcement of law organizations in order to detect crime and for applying preventive measures.

Keywords : Analysis, Crime, DBSCAN, Existing System, Law Enforcement.

I. INTRODUCTION

Today, time is a concerning factor for sentencing criminals. Many a times criminal is released on bail may yet be a potential threat to the society, even after they have served their sentence. As we all know the rate of crimes is increasing adequately and modern technologies are helping them without knowing such as chatting, videos, news, apps, websites, etc. Criminals cannot be predicted easily so as per the criminals mind not only the persons but also crimes cannot be predicted, so we are analysing the data we have as a dummy dataset and we are going to conclude these data for future awareness for crime records, it includes city(where it has done), number of crimes happens like these crimes, etc. Analysing the data is difficult because there are no sufficient data is available about a particular crime or there may be inconsistent of data availability.

As the crime rates are increasing we be properly analysed and stored. In this paper, we look at the use of frequent pattern mining with association rule mining to analyse the various crimes done by a criminal and predict the chance of each crime that can again be performed by that criminal. This analysis may help the law enforcement of the country to take a more accurate decision or may help in safeguarding an area if a criminal released on bail is very much likely to perform crime. In this paper we look at the use of missing value and clustering algorithm for a data mining approach to help predict the crimes patterns and fast up the process of solving crime. We are going apply these techniques to real crime data. We also use semi supervised learning technique in this paper for knowledge discovery from the crime records and to help increase the predictive accuracy.

II. MAJOR CHALLENGES:

In the present scenario, following major challenges are encountered:

- The crime data is increasing day by day.
- Problem of identifying techniques that can accurately and efficiently analyze this growing volumes of crime data.
- To record crime data different methods and structures are used.
- The data available is redundant and there are too many missing values which makes analysis process more difficult.
- Due to complexity of crime issues investigation of the crime takes longer duration.

III. AIM AND OBJECTIVE

- Aim: Our aim is to develop a user friendly website which can predict regions which have high probability of crime occurrence and can visualize crime prone areas on map.
- Objectives: 1. Performing data mining algorithm on available dataset to find required result. 2. System can predict areas where there is high possibility and probability for crime occurrence. 3. Visualizing crime prone regions in specific areas.

IV. LITERATURE REVIEW

In the study and analysis of criminology data mining can be categorized into two main areas, crime control and crime suppression. In crime control, we use knowledge from the analyzed dataset. And crime suppression used to catch criminal by using his/her history records.

As per the literature survey, crime data is growing very fast and in large amount (running into zeta bytes). So we need advanced and efficient techniques for analysis.

According to [1], the crime alert areas can be represented graphically using maps, which indicates the crime alert in respective area.

The clustering methods are implemented and their performance is tested based on accuracy.

According to the conclusion given in [8], we obtain a result that the DBSCAN clustering algorithm is more accurate than K-means algorithm. DBSCAN algorithm forms effective clusters.

The Experimental setup [6] describes complete setup and steps to manipulate data and change it into useful information. Also result is shown on maps using Google map.

So, according to survey we are having problem statement as, "Existing system do not have provision to predict crime prone regions, less efficiency of algorithm causes trouble in analysis process".

V. CLUSTERING TECHNIQUE

Clustering is the process of partitioning the data into similar classes. It can also formulate a multi-objectives optimization problem. We are considering two clustering techniques, K-means and DBSCAN algorithm for this purpose.

The performance of two clustering techniques was tested based on their accuracy, speed of creating clusters, and identifying crime patterns and crime regions. Henceforth we have found that the DBSCAN algorithm shows improved results than the k-means

algorithm and therefore we are using the DBSCAN algorithm for further study.

VI. PROPOSED WORK

We propose a system which can analyze, classify and predict various crimes, find probability of crime occurrences in a given region. Our system is effective in terms of analysis, speed of crime, classify crime according to their type and show probability of crime occurrences in nearby location.

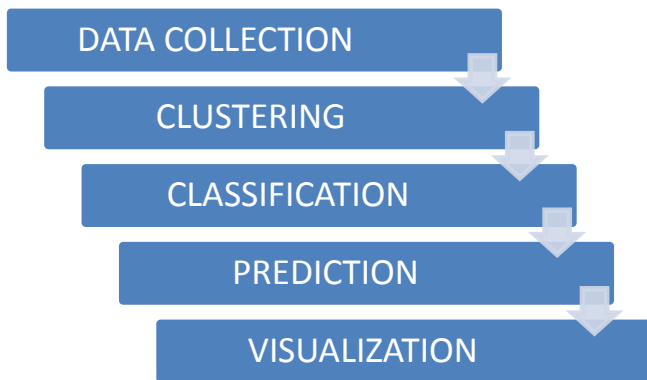


Fig1: Proposed Architecture

Following are steps in doing crime analysis:-

1) DATA COLLECTIONR:

We use dummy dataset because of limitation in getting crime data records from Law Enforcement department. The collected data is stored into database for further process. Since the collected data is unstructured data we use mango DB. Crime data is an unstructured data since the no of field, content, and size of the document can differ from one document to another the better option is to have a scheme less database.

2) CRIME CLUSTURE:

For clustering data we are using DBSCAN clustering algorithm. The DBSCAN algorithm is basically based

on clustering points within the distance of epsilon with some initial minimum number of points. On comparing with K-means the DBSCAN clustering has high accuracy for the given dataset and forms effective clusters.

3) CLASSIFICATION:

For classification we are using an algorithm called Naïve Bayes which is a supervised learning method as well as a statistical method for classification. The algorithm classifies crime based on given training dataset. It is simple, and converges quicker than logistic regression

4) PREDICTION:

For prediction we are using the decision tree concept. A decision tree is similar to a graph in which internal node represents test on an attribute, and each branch represents outcome of a test. Corresponding to each place we build a model. So for getting the crime prone areas we pass current date and current attributes into the prediction software.

5) VISUALIZE:

The crime prone areas can be graphically represented using a map also predicted regions are also shown in another map view. Visualization helps to get more accuracy in performance.

VII. MODULES

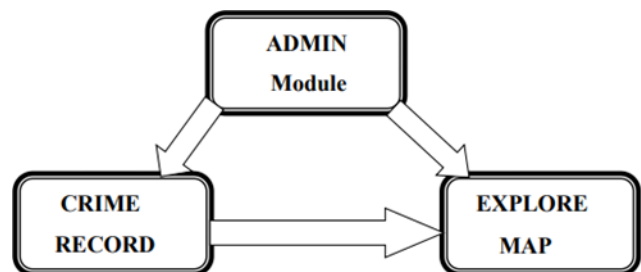


Fig 2:- Project Module

IX. REFERENCES

1. ADMIN MODULE: Admin will register into the system first to generate User ID and Password. As an admin he/she has authorization to register new crime, update information of crime. Admin can also edit his/her profile and reset password only if he/she is logged in to the system.

2. CRIME RECORD MODULE: This module contains detailed information about crime, Such as criminal name, victim gender, crime location (map), crime type, date and time of crime, for e.g. Crime happens at Night. All the necessary information will be stored in this module and the classification and clustering algorithm will apply on this dataset.

3. EXPLORE MAP: It contains two map view, and one list view. Where, in one map cluster view of crime is shown which helps to find out number of crime register in a particular area and another map view shows predicted region or area of crime. List view where description of crime is given along with date and time registered for that crime. We can use the filters like crime date and time and according to type of crime we can fetch crime from dataset.

VIII. CONCLUSION

In this paper, an innovative and fast solution is presented for Crime analysis and prediction. We use DBSCAN data mining algorithm instead of K-means to get more efficient result. Our website predicts crime prone regions, we will use a dummy dataset from **kaggle.com** website, preprocess that data and apply prediction algorithm to obtain result.

- [1]. Devan M.S (TCS), Crime Analysis and Prediction Using Data Mining Conference Paper , August 2014.
- [2]. Malathi. A and Dr. S. Santhosh Baboo. Article: an enhanced algorithm to predict a future crime using data mining. International Journal of Computer Applications, 21(1):1–6, May 2011. Published by Foundation of Computer Science.
- [3]. Anshu Sharma, Raman Kumar, Analysis and Design of an Algorithm Using Data Mining Techniques for Matching and Predicting Crime, IJCST Vol. 4, Issue 2, April - June 2013.
- [4]. Jyoti Agarwal, Renuka Nagpal, Rajni Sehgal, “Crime Analysis using K-Means Clustering”, International Journal of Computer Applications (0975-8887), Vol. 83, No. 04, 2013
- [5]. Anisha Agrwal, Dhanashree Choghule, Arpita Agrwal, Application of Crime Data Mining Using Data Mining Method, International Journal of Advanced Computational Engineering and Networking, ISSN: 2320-2106, Volume-4, Issue-5, May.-2016
- [6]. Veenet Jain, Yogesh Sharma, Vaibhav Arora, Ayush Bhatiya, Crime Prediction using Kmeans Algorithm, Global Research and Development Journal for Engineering | Volume 2 | Issue 5 | April 2017.
- [7]. P. Gera, and R. Vohra, —Predicting Future Trends in City Crime Using Linear Regression, IJCSMS (International Journal of Computer Science & Management Studies) Vol. 14, Issue 07 Publishing Month: July 2014.
- [8]. Raghavendhar T.V, Joslin Joshy, Mahaalakshmi R, Ashutosh Soni M, Crime Prediction and Analysis using Clustering Approaches and Regression Methods | IJCAT - International Journal of Computing and Technology | Volume 5 | Issue 4, April 2018