Advanced Concepts of Data Mining in Crime Investigation

Adithyan S[1] Gautham Krishna M[2] Vignesh G[3] Arun Padmanabhan[4]
[1,2,3,4]Department of Computer Applications
[1,2,3,4]Saintgits College of Applied Sciences, Kottayam, India

Abstract— Crime analysis and prevention is a systematic way of identifying and analyzing popular patterns and methods. Our system can predict regions with the highest risk of crime. With the proliferation of computer programs, crime data analysts can help law enforcement speed up the process of solving crime. Using the data concept mining can extract previously unknown, useful information from random data. The data mining process that can help solve crime faster. Instead of focusing on the causes of crime as the crime scene of the offender, political hostility etc. we focus mainly on everyday criminal activities.

Keywords: Clustering, Classification, Visualization, K-means, Neural Network

I. Introduction

In recent years we can see that the number of criminal activities is increasing rapidly. In addition to that, Worldwide riots and terrorist are emerging. By looking at the motives and the causes, it is possible for us to find a pattern, which can be studied to find the root cause and then preventing them from happening or to reduce the chances of it.

According to Crime Records criminal cases such as burglary and other offenses have been reduced while cases such as murder, assault, etc., we can predict the exact location of the crime.

II. Usage Of Data Mining In Crime Investigation

A. Forensic Investigation

The way of analyzing information and processing information studies is very important in the field of forensic investigation. In the way that data mining is an absolute solution for expertizing the accuracy and level of incredibility in the field of forensic investigation, it includes: -

1) Forensic Analysis

It includes examining data sources by forensic investigators to identify, preserve and analyzing information for data analysis and discovering crime patterns. Using Data mining can make forensic analysis strategies to be an inevitable outcome in the world of Forensic crime investigation

2) Forensic Identification

By using the data mining technique, we can come to a conclusion identification of individuals, group or even a particular organization that involved in a crime. Data mining helps in an easy identification process in which including accurate identification criteria for increasing process reports results in inaccuracy.

B. Cyber Crimes

The impact of cybercrimes varies from individual, business to governments. Data mining is a great criteria technique that can be used for the cyber world for identifying patterns and detecting types of fraud detection, cyber threatening etc. easily by using data mining techniques that are being a very popular method for cybercrime investigation of increased cybercrimes in our day to day developing technological life.

C. Narcotics

The increased use and misuse of drugs and related substance are an increased challenge for youngsters. Data mining helps to analyze the current usage and limit it to a valid range by thorough analysis, for categories drug variants that are very harmful or being less harmful by examining data of effect on health and analyzing the previous and current record of drugs based crimes the government can take necessary actions for controlling and detecting crime increase by wide sources through proper study and research.

III. Technologies and Tools

Data mining can be done using a variety of tools and to be more efficient on the result, using a suitable mining algorithm is very important. Some of the most used algorithms/techniques are:

1) Classification
2) Clustering
3) Regression
4) Outer

By using the above techniques, we are able to predict the number of crimes going to happen or the probability of a certain crime going to take place. The data which we are going to use is also very important and having the right type of data is very crucial in predicting the outcome, the types of data most favourable are:

1) Relational databases
2) Data warehouses
3) Advanced DB or information repositories
4) Object-oriented /Object-relational databases
5) Transactional and Spatial databases
6) Heterogeneous and legacy databases
7) Multimedia and streaming database
8) Text databases
9) Text mining and Web mining

After choosing the right technique it's important to choose the right tools to mine, Data mining is a very hard and time-consuming process as it requires a lot of data and computational power. Two of the most popular data mining tools which are widely used are: -

1) R-Language
2) Oracle Data Mining

R-language is an open-source tool for statistical computing and graphics. R has a wide variety of statistical, classical statistical tests, time-series analysis, classification and graphical techniques. It offers effective data handling and storage facilities.

Oracle Data Mining popularly known as ODM is a module of the Oracle Advanced Analytics Database. This Data mining tool allows data analysts to generate detailed insights and makes predictions. It helps predict customer

behaviour, develops customer profiles, identifies cross-selling opportunities

IV. DATA TRANSFORMATION

The data which we collected has to go through several sorting and cleaning procedures before it can be used as the sample. The data first has to go through smoothing Which is used to clear noise from the data i.e. unwanted or redundant information. It is then aggregated to fit a certain mining style. In this step, Low-level data is replaced by higher-level concepts with the help of concept hierarchies.

There are steps in doing Crime Analysis:

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

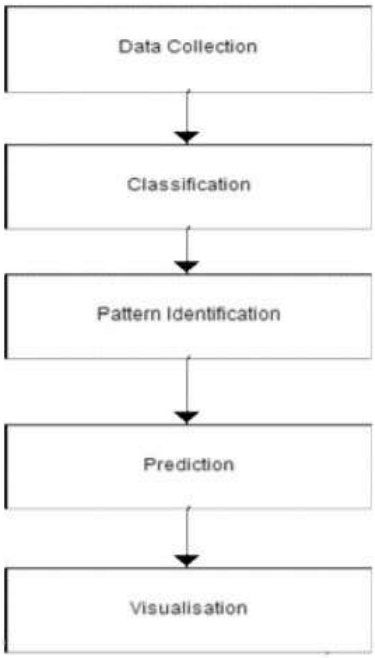

Fig. 1:

V. DATASET COLLECTION

In this step the dataset is collected from the Kaggle website. The dataset named Crime in India by Rajanand llangovan is chosen for this research.

VI. K-MEANS IMPLEMENTATION

Classification, Pattern Identification and Prediction are done with the help of K-means in this case the existing patterns and relationships are searched in the database using k-mean and clustering methods. This method helps to provide an overview of the database and as a result, it helps to search, manage and retrieve required or desired information. From the database 12 regions are selected to form clusters and selections are made on the basis of the IPC value of each province between years 2012-2020.

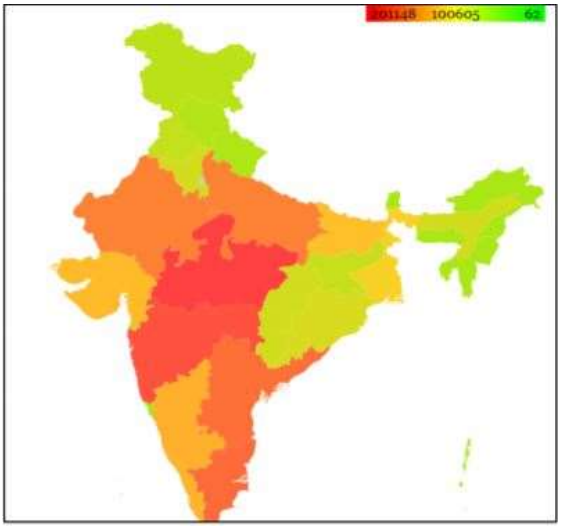

Fig. 2:

As areas are marked on the map it helps to analyse the regions. This information is very useful for investigative agencies and police officers. The integration is achieved using the WEKA tool for cases 1 to 4 - crime in Indian countries, qualifications for "State", "year" and "average IPC crimes" and "total IPC crime" methods.

VII. EXPERIMENTATION AND RESULTS

Since crime is increasing at a rapid rate globally it is important to control it. In order to reduce crime rates, we need to study the crime rates of various regions of a country. In this research all the states and union territories crime rate are studied in detail for different types of crimes. Since unsupervised and supervised learning techniques are used it helps in improving the filtration of large crime databases. Thus, by following the proposed approach the crime rate can be reduced in time and effort.

VIII. CONCLUSION

The Use of Data Mining in crime investigation can help predict criminal patterns, hence can stop/learn different outcome and can take the required precaution. union territories selected are chosen on the basis of their crime rate. This approach is very useful in studying if the crime rate is increasing or decreasing in a particular region. If the crime has increased necessary measures can be taken by the officials to study why the crime has increased and also how to reduce the crime rate in that region. In this research the crime rates during 2001 to 2012 are analysed and this has helped in ranking the states and union territories on the basis of their average IPC crime rate

IX. REFERENCES

[1] Shyam Nath, "Crime Pattern Detection Using Data Mining", IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops, 2006, DOI: 10.1109/WI-IATW.2006.55
[2] Devendra Tayal, Arti Jain, Surbhi Arora, Surbhi Agarwal, Tushar Gupta, Nikhil Tyagi, "Crime detection and Criminal identification in India using data mining

technique", Ai & Society, 2014, https://doi.org/10.1007/s00146-014-0539-6
[3] Shiju Sathyadevan, Devan S, Surya S, "Crime analysis and prediction using data mining", First International Conference on Networks & Soft Computing (ICNSC2014), 2014, DOI: 10.1109/CNSC.2014.6906719
[4] Bhavani Thurasining, Data mining Technologies, Tools and Trends Press, Washington D.C (1994)
[5] Mehmed Kantardzic, Data Mining Concepts, Methods and Algorithm. John Wiley& Sons (2012)

