Crime Pattern Detection, Analysis & Prediction using Machine Learning

Rohit Patil[1], Muzamil Kacchi[2], Pranali Gavali[3], Komal Pimparia[4]

[1,2,3,4]U.G. Student, Department of Computer Engineering, LGNSCOE – Sapkal College, Nashik, Maharashtra, India

---***---

Abstract - *Criminal analysis is a methodical approach for identifying and analyzing patterns and trends in crime. With the increasing origin of computerized systems, crime data analysts can help the Law enforcement officers to speed up the process of solving crimes. Using the concept of data mining, system can analyze previously unknown, useful information from an unstructured data. Predictive policing means, using analytical and predictive techniques, to identify criminal and it has been found to be pretty much effective in doing the same. Because of the increased crime rate over the years, system will have to handle a huge amount of crime data stored in warehouses which would be very difficult to be analyzed manually, and also now a day's, criminals are becoming technologically advance, so there is need to use advance technologies in order to keep police ahead of them. In this paper, the main focus is on the review of algorithms and techniques used for identify the criminals.*

Key Words: k-means, Naïve Bayes, crime prediction Regression, Association Rule mining etc.

1. INTRODUCTION

A crime rate has become a topic of major concern certainly to limit the development of good governance and increasing day by day. Crimes are neither systematic nor random otherwise crime cannot be analysis. When crimes like robbery, firebombing etc. have been decreased, crimes like murder, sex abuse, gang rape etc. have been increased. We cannot analyze the victims of crime but can analyze the place where crime occurred or happened. It is difficult to analyze the data to detect crime patterns or predict future crimes by intelligence agencies or local law enforcement agencies. So, there is a need of an effective analyzing tool which can analyze crime data efficiently and quickly to give some useful crime patterns. Predictive policing means, using analytical and predictive techniques to identify criminal and it has been found to be pretty much effective in doing the same. Because of the increased crime rate over the years, we will have to handle a huge amount of crime data stored in warehouses which would be very difficult to be analyzed manually, and also now a day's, criminals are becoming technologically advance, so there is need to use advance technologies in order to keep police ahead of them.

The review is purely based on the crime analysis using data mining. The main aim is to decrease the crime using advance technologies in order to keep police ahead of them. One of the approaches in which crime rates has been calculated from cases registered and persons arrested data during five years. Here, it is quite visible that the cases registered during five years are high. No doubt the persons arrested are low but the aim here is to reduced crimes.

1.1 Literature Review

In the existing crime management system, most of the operations are done manually like send complaints, taking actions against crimes, view status etc. This system need more man power to track the records of crimes. The existing system doesn't have system security. The existing system is time consuming and not very user friendly. So with the existing system if anybody wants to complaint against crimes he must do it through the police. Retrieving old crime records is very time consuming. In the current system all work is done on papers so it is very difficult to secure crime reports data. Error detection in the previous entries made and data cross verification is another important function. These are done manually, and it would take time. The existing system has more workload for the authorized person, but in the case proposed System, the user can register in our site and send the crime report and complaint about a particular city or person.

No	Paper / Publication	Author	Year
1	An overview on Crime Prediction method	Nurul hazwani mohd Shamsuddin	2017
2	Crimecast: A crime pridiction and strategy Direction service	Nafiz Mahud, Khalid ibn Rahman	2016
3	Crime Analysis and prediction Using data mining	Shiju Sathyade van	2014
4	Crime prediction using pattern and context	Nelson Baloian	2017

2. ARCHITECTURE AND WORKING

The four algorithms which were executed on the dataset are
1. Association Mining (Apriori) /Processing
2. Clustering(k-Means)
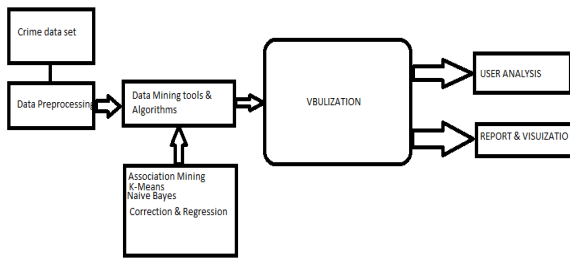3. Classification Techniques (Naive Bayes)
4. Correlation & Regression.

Fig : System Architecture

The dataset embraces number of people arrested and number of crimes committed along with various other attributes. Here we are primarily using four data mining algorithms for analysis of crime and to find hidden patterns of crime in India.

The data mining techniques used are as follows: - 1: Association mining (Apriori Algorithm) along with Clustering (k-mean):- The paper tends to help specialist in discovering patterns, trends, making forecasts, finding relationships and possible explanations, mapping criminal networks and identifying possible suspects. K-means is used to create number of clusters according to values high and low. The data is first imported from the file into local dataset for preprocessing; later k-means is applied on this data set using the same graphic user interface (GUI)

The result of the k-means will acts as an input to the Apriori algorithm for discovering the association among a number of other attributes.

2.1 Apriori Algorithm

Apriori algorithm is a type of association mining, used to find frequent item set. The result obtained after k-means is used as dataset for Apriori as the clusters are now divided into high and low value now we can get the association between various attributes.

2.2 Clustering by K-means Algorithm

K-means is the simplest and most commonly used partitioning algorithm among the clustering algorithms in scientific and industrial software. Acceptance of the K-means is mainly due to its being simple. This algorithm is also suitable for clustering of the large datasets since it has much less computational complexity, though this complexity grows linearly by increasing of the data points. Beside simplicity of this technique, it however suffers from some disadvantages such as determination of the number of clusters by user, affectability from outlier data, high-dimensional data, and sensitivity toward centers for initial clusters and thus possibility of being trapped into local minimum may reduce efficiency of the K-means algorithm

2.3 Classification Techniques

Classification is one of the important features of data mining as a technique for modeling of forecasts. In other words, classification is the process of dividing the data to some groups that can act either dependently or independently. Classification is used to make some examples of hidden and future decisions on the basis of the previous decision makings. Decision tree learning, neural network, nearest neighborhood, Nave Bayes method and support vector machine are different algorithms which are used for the purpose of classification

2.4 Correlation & Regression

Classification is one of the classic data mining techniques, which is used to classify each item in a set of data into one of Predefined set of classes or groups. The idea is to define the Criteria use for the segmentation of the whole database, once this is done, individual dataset can then fall into one or more groups naturally. With the help of classification, existing dataset can easily be understood and it also helps to predict how new individual dataset will behave based on the classification criteria. Data mining creates classification models by observing already classified data and finding a predictive pattern among those data
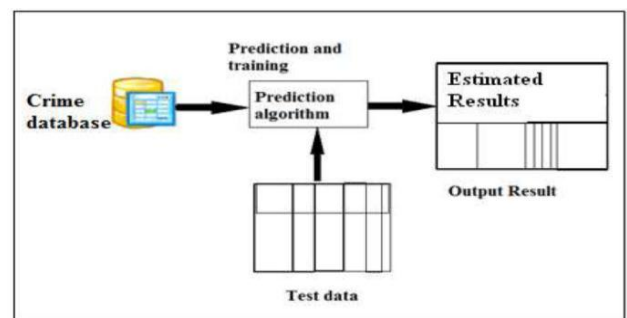
3. FLOWCHART

Chart : Flowchart

4. FUTURE SCOPE

With the increasing popularity towards the technology, machine learning is being adopted by various fields like Finance, Healthcare, Travel and Tourism, Logistic and Real Estate and many more.

We can use detection using machine learning in many fields for maintaining the Tranparency.

In future we can use this pattern by performing certain necessary upgradation in military services on the border campaing technologies.

5. CONCLUSIONS

In this paper, we propose our criminal prediction model in statistical implementation. As we are not aware of such implementation model of criminal prediction system in our country we were unable compare our approach. By implementing real life authentic dataset with precisely increased domain nodes and hidden layers in Neural Network, we would able to figure out more accurate prediction.

The goal of the project reported in this paper was to decide whether or not it would be possible to develop a computer based system to build risk surfaces with sufficient time efficiency as to be appropriate to be actually used by the Maharashtra police in their field operations.

REFERENCES

[1] Shiju Sathyadeven, Deven M.S, Surya Gangadharan. S, Crime Analysisand prediction using data mining, IEEE, 2014

[2] Mugdha Sharma, Z-crime: A data mining tool for the detection of suspicious criminal activity based on decision tree, IEEE, 2014,ISBN:9781-4799-4674-7/14

[3] Arunima S. Kumar, Raju K. Gopal, Data Mining Based Crime Investigation System: Taxonomy and Relevance, IEEE, 2015, ISBN: 978-1-4799-8553-1/15

[4] Ubon Thansatapornwatana, A Survey of Data Mining Techniques for Analyzing Crime Patterns Second Asian Conference on Defense Technology ACDT, IEEE, 2016, ISBN: 978-1-5090-2258-8/16

[5] Dr.Zakaria Suliman Zubi, Ayman Altaher Mahmmud, Using DataMining Techniques to Analyze Crime patterns in the Libyan National Crime Data, Recent advances in image, audio and signal processing.ISBN: 978-960-474350-6

[6] National Crime Records Bureau Website. [online]