CRIME SPOT DETECTION

Akshay N Diwate[1], Vidya K Chaudhari[2], Monika R Gaikwad[3], Aishwarya M Sangale[4], Prof. S.R.Bedse[5]

Department of Computer Engineering, Jawahar Education Society's, Institute of Technology, Management and

Research, Survey No 48, Gowardhan, Gangapur Rd, Nashik, Maharashtra, 422222, India[1,2,3,4]

Prof. Department of Computer Engineering, Jawahar Education Society's, Institute of Technology, Management and

Research, Survey No 48, Gowardhan, Gangapur Rd, Nashik, Maharashtra, 422222, India[5]

Abstract: Road crime is most important issue not only for Indian Government but also for common people. Mostly, it is found that road crime happening are more frequent at certain specific locations i.e. black spot. The analysis of these black spot can help in identifying certain road crime factor that make a road crime to occur frequently in
that locations. In this project we apply statistics analysis and Eclat algorithm on the Fatal Crime dataset as an attempt to address this problem. Association rule mining is one of the popular data mining techniques that identify the causes of crime of road crime. In this project, we first applied Eclat algorithm to group the crime locations
into A level, B level, C level crime location. Eclat algorithm takes crime level count as a factor to cluster the locations. Then we will use association rule mining to iden-tify these locations. The rules show different factors associated with road crimes at different locations. For all this we will provide crime data that are issue from Nashik city Commissioner office. Safety driving suggestions will be maked based on crime data, association rules, classification model, and clusters obtained.

I. INTRODUCTION

To identify important factors to road crimes in Nashik we have obtained a large dataset every crime recorded in the Nashik district commissioner office in the Year 2014-2017. The data is currently in an unsorted and scatter format and is stored in a Microsoft excel sheet database table. Unfortunately with the data in its current format, no relevant points or conclusions can be drawn. It is hoped that by applying data mining processes and techniques to the data set, relevant attributes and patterns can be established. And scientific study will also done that will helpful to government authorities and citizen. The main achievements of this project is to create awareness of the conditions affecting road crimes, Establishing which individuals are most likely to be involved in a road crime.

II. PROPOSED SYSTEM

Data mining is basically used to find out unknown patterns from a large amount of data. The prediction of future crime trends involves tracking crime rate changes from one year to the next and use data mining to project those changes into the future. The basic method involves clustering the states having the same crime trend and then using 'next year' cluster information to classify records. To the clustered results, a classification algorithm is applied to predict the future crime pattern. The classification is performed to find in which category a cluster would be in the next year. This allows building a predictive model for predicting next year's records using this year's data.

Step 2: Crime trend analysis/forecasting/ prediction
Crime forecasting helps police to take tactical actions in advance such as targeting patrols to hot spots, conducting surveillance for deployment of special units, scheduling vacation and training of the cops. Crime forecasting requires crime place, ecology of crime and Hot Spot.

Step 3: The Short Term Forecasting Algorithm
The short term forecasting of crime helps in tactical decision making at police station or at district level.
To implement this algorithm, the fitted line equation is used
$$Y_t = + (b * t) \qquad (1.1)$$
Where t = serial number of a month/ year (i.e. 1 for Jan, 2 for Feb, 3 for Mar and so on) Value of a and b can be calculated using $\sum X, \sum Y, \sum XY, \sum X^2$
Where,
$\sum X$ = sum of number of months or years that are to be taken into account to predict the crime values.
$\sum Y$ = sum of number of crimes that took place in previous years/ months.

N = Number of observations

In order to calculate the value for a and b we have line equation

$$\textbf{\textit{Absolute Deviation}}[\textbf{\textit{Z}}] = \textit{Volume }[Y]- \textit{Fitted Value} \quad (1.4)$$

$$\textbf{\textit{Squared Deviation}} = \textit{Square of Absolute Deviation}[Z^2] \quad (1.5)$$

$$\textbf{\textit{MAD}} = \sum Z/N \quad (1.6)$$

$$\textbf{\textit{MSD}} = \sum Z^2/N \quad (1.7)$$

$$\textbf{\textit{MAPE}} = \textit{Mean Absolute Deviation }[MAD]/\textit{ Mean Ratio} \quad (1.8)$$

$$\Sigma Y = N\,a + b\,\Sigma X \quad (1.2)$$
$$\Sigma XY = a\,\Sigma X + b\sum X^2 \quad (1.3)$$

The following flowchart explains the algorithm precisely

.

Step 4: Accuracy of the system

To measure the accuracy of forecasted values we need to check the MAPE, MSD and MADvalues.

III. OBJECTIVES

This Software should be developed for Desktops and Android devices.Client or user has an active internet connection to view the website.Website interface must be friendly and easy to use Must-have requirements are con- straints that should be possessed by the ap-plicant, whereas nice-to-have requirements are preferences that are taken intoconsideration when ranking applicants.The e-recruitment is a system for quickly reaching a large set of potential job-seekers. E-recruiting has attractive growthsince the late 1990s when the rapid economy changes produced a high de-mands for qualified candidates that the labor market could not fully satisfy. The e-recruiting platforms such as corporate homepages and job portal shavedriven this development.

IV.RELATED WORK

YEAR	DESCRIPTION
2017	Athira Mohan, Dr. V.S. Landge,"Identification of Accident Black Spots on National Highway", International Journal of Civil Engineering and Technology (IJCIET), Vol.8, Issue 4,pp. 588-596
2015	Dahlia Sam, Esther Evangelin and V. Cyril Raj, "Improving Road Safety for Pedestrians in Black Spots using a Hybrid Vanet of Vehicular Sensors and Pedestrian Body Unit", ARPN Journal of Engineering and Applied Sciences.

V.METHODOLOGIES

Analysis Models: SDLC Waterfall Model to be applied The Waterfall Model was the first Process Model to be introduced. It is also referred to as a linear-sequential life cycle model. It is very simple to understand and use. In a waterfall model, each phase must be com- pleted before the next phase can begin and there is no overlapping in the phases. The Waterfall model is the earliest SDLC approach that was used for software development. The waterfall Model illustrates the soft- ware development process in a linear sequential flow. This means that any phase in the development process begins only if the previous phase is complete. In this waterfall model, the phases do not overlap. Waterfall Model - Design Waterfall approach was first SDLC Model to be used widely in Soft- ware engineering to ensure success of the project. In "The Waterfall" approach, the whole process of software development is divided into sepa- rate phases. In this Waterfall model, typically, the outcome of one phase acts as the input for the next phase sequentially. The following illustration is a representation of the different phases of the Waterfall Model. The sequential phases in Waterfall model are Requirement Gathering and analysis All possible requirements of the system to be developed are captured in this phase and docu- mented in a requirement specification document. System Design The requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture. Implementation With inputs from the system design, the system is first developed in small programs called units, which

are integrated in the next phase. Each unit is developed and tested for its func- tionality, which is referred to as Unit Testing. Integration and Testing All the units developed in the implemen- tation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures. Deployment of system Once the functional and non-functional test- ing is done; the product is deployed in the customer environment or released into the market

VI.SYSTEM ARCHITECTURE

The basic purposes of activity diagrams is similar to other four diagrams. It captures the dynamic behavior of the system. Other four diagrams are used to show the message flow from one object to another but activity diagram is used to show message flow from one activity to another. • Activity is a particular operation of the system. Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part. • It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single. • The purpose of an activity diagram can be described as • Draw the activity flow of a system. • Describe the sequence from one activity to another. • Describe the parallel, branched and concurrent flow of the system.

VII.ADVANTAGES

• No Central Point of Failure: Since DApps are distributed and don't rely on one single server, there is no central point of failure. DApps allow data stored in them to be decentralized across all their nodes. These nodes are independent of each other. In case of failure of one node, the other nodes won't get affected and will run on the network accordingly.

• Nevertheless, current blockchain-based applications are still limited to utilizing smart contracts for core data and functionality that should be resistant to modifications. Smart contract users need to run their programs locally in order tocomplete the application.

• Voters can vote from home during a pandemic situation instead of going to other places.

• Digital voting based on Blockchain provides private and secure solutions.
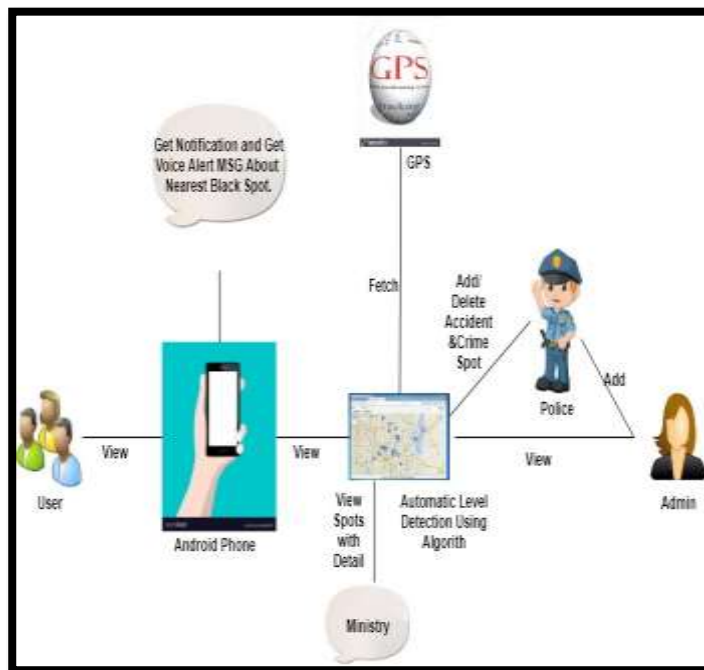
• Voting online saves time and eases vote management.

Fig. 1 Architecture

VIII.CONCLUSION

• In this study, the technique of association rules with a large set of Accident Crimes data to identify the reasons of road Accident Crimes were used. • Analysis showed that producing the association rules, makes identification of factors involved in the Accident Crime that occur to- gether, easier. • It shares a lot in understanding the circumstances and causes of the Accident Crime. So the association rule mining gives the direction to deeper research on the causes of road Accident Crimes.

It helps government to adapt the traffic safety policies with different types of Accident Crime and situations. • The main result of this study is that although the characteristics of humanity and behavior are very impor- tant in occurrence of all road Accident Crimes but we can understand that spatial features and infrastructure play a major role in the Accident Crime. • In this study it is tried to choose the interesting and superior rules to provide a lot of valuable information for policies to provide bet- ter safety policies. • This article can be a step towards providing useful information for highway engineers and transportation designers to design safer roads

IX.REFERENCES

1 T. Abraham,O. de Vel,"Investigating profiling with computer forensic log data and association rules" ,Proc. of the IEEE International Conference on Data Mining (ICDM'06),pp 11-18, 2006.

2 De Bruin, J.S. , Cocx, T.K. , Kosters, W.A. , Laros, J. and Kok,"Data mining approaches to criminal career analysis, in International Conference on Data Mining", Pp. 171-177, 2006.

3 Athira Mohan, Dr. V.S. Landge,"Identification of Accident Black Spots on National Highway", International Journal of Civil Engineering and Technology (IJCIET), Vol.8, Issue 4,pp. 588-596, 2017.

4 Dahlia Sam, Esther Evangelin and V. Cyril Raj, "Improving Road Safety for Pedestrians in Black Spots using a Hybrid Vanet of Vehicular Sensors and Pedestrian Body Unit", ARPN Journal of Engineering and Applied Sciences, Vol.10, 2015.

5 M. Langkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised ¨feature learning and deep learning for time-series modeling", PatternRecognition Letters, vol. 42, pp. 11–24, 2015.

6 R. Maciejewski, R. Hafen, S. Rudolph, S. G. Larew, M. A. Mitchell,W. S. Cleveland, and D. S. Ebert, "Forecasting hotspots—a predictiveanalytics approach", IEEE Transactions on Visualization and ComputerGraphics, vol. 17, pp. 440–453, 2011.

7 A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert,"Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement", IEEE transactionson visualization and computer graphics, vol. 20, pp. 1863–1872, 2014.