

# Accident Blackspot Location And Analysis System

Alka Sinha<sup>1</sup>, Ashwini M<sup>2</sup>, Haritha NC<sup>3</sup>, Swathi S<sup>4</sup>, Rajendra M<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup> Dept of Computer Science and Engineering

<sup>1, 2, 3, 4, 5</sup> Atria Institute of Technology/VTU, ASKB Campus, Anandnagar, Hebbal, Bengaluru, Karnataka 560024

**Abstract-** Accident analysis studies aim at the identification of high rate accident locations and safety deficient areas. In this study, an effort has been made to identify the accident prone zones. For this purpose, the road accident data have been used. Accident particulars like date, location, type of vehicle involved, number of persons injured or died are included in the database. Accident analysis studies aim at the identification of high rate accident locations and safety deficient areas. The concept of cluster has become a popular practice in the field of road safety, mainly for the identification of prone zones. The objective of this study is to contribute to the ongoing research effort on accident identification and analysis. Employing the concept of Naïve Bayes Classifier and Clustering algorithm which finds the posterior probability based on the attribute values that is in the datasets, which predicts whether accident occurs or not and also provides the overall information about these attributes which causes accidents. Hence this proposed model also maps the accident prone location as a black spot. Based on the result, suggestions are provided to reduce the accidents in the future.

**Keywords-** Accident locations, Black spots.

## I. INTRODUCTION

The economic growth of any country depends on its transportation network, including roads, rail, and water and air connectivity. Of these roads is the critical one. Improving and providing great ideas on road safety is essential. Road traffic accidents are current growing issue. As surprised way it increased in its number from thousands to lakhs and so on. Road transportation is one of the busines paths for country's economic purpose. Incrementation in road accidents introducing significant social and economic losses. According to world wide road traffic record the death reached more than 1.28 million people per year. A good road network is important as it provides connectivity between rural and urban areas. Along with this, road safety is a most important aspect. It plays a key role towards a sustainable transportation development strategy. The adverse impact of modern road transportation system is injury and loss of life due to road accidents. While the road accident situation is improving in high industrialized countries, most developing countries are facing a severe situation. The continuous socio-economic

growth over the years is causing an increase in demand for transport. With the number of vehicles growing rapidly, more is the road conflicts develop traffic accidents. It is observed that most of these accidents result from human actions and carelessness on the part of drivers or pedestrians. However, the probability of occurrence of accidents, and its severity, can often reduced by the application of proper traffic control devices and sound roadway design practice. The use and maintenance of such control devices and design specifications however, depends upon the analysis of traffic accident records at specific location. It has been recognized that one of effective means towards accident reduction lies in a systematic and scientific approach based on the use of accurate and reliable accident data. So, the data required for such an analysis is not always available. Most of the accident information available in police records is incomplete. In addition, records are also needed to provide actual facts on guide programs including education, maintenance, vehicle inspection, emergency medical services and engineering improvements that can be provided to the streets and highways. The reasons for accidents can be listed to the lack of signage, raised cover with trees and bushes, making pedestrians not visible to drivers, improper design of pedestrian crossings, improper roads, and lack of enforcement to control wrong side entry and movement of vehicles. There are however, four factors which contribute to road accidents directly or indirectly like road, vehicle, road user and environmental factors. In Indian cities, the traffic is highly in nature, ranging from pedestrians and animal drawn carts to motorized vehicles like cars, buses, and as well as over sized vehicles. Analyzing of traffic data indicates that drivers fault is main reason for majority of the accidents. The economic and social costs of traffic accidents are very huge. Property damage, lost productivity, medical expenses, and inflated motor insurance rates imposed an estimated loss to the Indian economy. Therefore there is need to study and analysis the accident pattern and reasons of such location area along with the identification of black-spots location in the city. Accidents causing reasons can be classified as

- Road aspects
- Vehicle aspects
- Road user aspects
- Environmental aspects.

## II. LITERATURE SURVEY

This section provides a background to the research through a review of some of the literature which are central to the scope of this paper. According to Sachin Kumar and Durga Toshniwal in “Analyzing Road Accident Data Using Association Rule Mining”, data mining techniques can be used to analyze the data provided by EMRI (Emergency Management research Institute) in which the accident data is first clustered using K-modes clustering algorithm and further association rule mining technique is applied to identify circumstances in which an accident may occur for each cluster. An Apriority algorithm has then been applied on every cluster using WEKA3.6 to generate association rules.

The paper “Data Integration and Clustering for Real Time Crash Prediction” presents a methodology to develop a classifier which uses historical as well as real-time data for any given road and outputs the probability of accident on any given time on that road. By using clustering method, a Bayesian network (BN) is constructed by mining the database using Kernel Density Estimation (KDE). This BN can act as a classifier to predict the probability of accidents for any given street. Data has been extracted from several simulation runs by PARAMICS micro-simulator.

As most of the current studies did not pay enough attention to the time factor when studying the relationship between traffic state and crashes on highways, the paper “Evolution of Traffic Flow Analysis under Accidents on Highways Using Temporal Data Mining” proposed a method to construct time-series data using traffic flow data when accidents happened. This paper proposes a model to describe traffic flow evolution when accidents happen on highways. A Discrete Fourier Transform was implemented to turn the time series from time domain to frequency domain. By clustering analysis, the traffic dynamics can then be studied. The newly developed method provides a better insight into the evolution of traffic flow on highways and the impact of highway crashes.

In “Application of Spatial Data Mining in Accident Analysis System” , a system based on data mining of Geographical Information System (GIS) has been proposed. A three-layer architecture displays the whole process of accident data extracting, preprocessing and mining and it applies spatial data mining to GIS. The first layer is data storage layer which includes databases. The second layer is business logic layer, including spatial data mining function module and accident analysis module. A design module can inquire about, add and delete spatial data, attribute data and accident data. The accident analysis module is used to identify and display

accident black-spots on the map according to a certain accident rate and accident level. The third layer is the interactive interface. ArcGIS Engine and C# have been used to develop the system

In his paper, Andreas Gregoriades [5] proposed a method which is split into two phases, (A) the development of the microscopic simulation model and (B) the development of the agent-based monitoring system phase. During Phase A, a preliminary micro-simulation model of a road network in Cyprus is developed using statistical data. Based on these a topology of the BBN model was created. Phase B of the method addresses the development of an agent-based monitoring system to track changes in traffic volumes, densities, behaviour and speed. This technique can test whether the model would generate predictions similar with the known scenario outcomes.

“Large Imbalance Data Classification Based on Map Reduce for Traffic Accident Prediction” [6] suggests using Hadoop framework to process and analyze big traffic data efficiently. Based on this, the predicting system first preprocesses traffic big data and analyzes it to create data for the learning system. The imbalance of created data is corrected by a sampling method. To improve accuracy, corrected data is classified and analysis is applied. Using Map Reduce, processing performance for generating training data set and conducting classification can be improved. The overall system is based on Hadoop, and which implements data preprocessing, learning data creation, over sampling by Hive. The cluster and classification analysis is operated by Mahout. First, input the accident and traffic data to Hive and process them with Hadoop.

## III. METHODOLOGY

**Data Collection:** Data collection is the collection of accident locations records frequently occurred. These data is collected from reports maintained by police department.

**Uploading the Dataset:** The data collected is uploaded.

**Training of Dataset:** The dataset which is uploaded is trained.

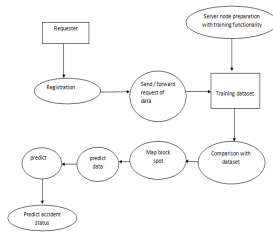


Fig: Dataflow

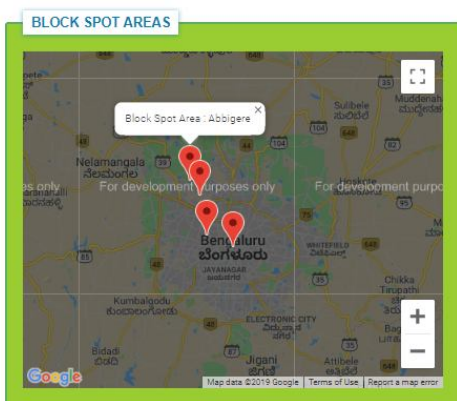
**Analyzing the roles:** The admin can analyze the dataset which is trained.

**Testing the dataset:** The dataset is tested by the user providing the attributes that is causing the accident.

**User Visualizing:** The map with Top accident locations are identified by the data that is provided by the admin.

#### IV. RESULTS

The fig below shows the top accident location that is obtained by the use of algorithm and the dataset that is used in this paper.



#### V. CONCLUSION

Machine learning approach has been proven as a reliable technique in analyzing road accident data. The rules generated for every group exposed the various factors associated with road accidents in these locations. However, the data with more number of attributes can reveal more information using our approach. Employing the concept of Naïve Bayes Classifier and Clustering algorithm which finds the posterior probability based on the attribute values that is in the datasets, which predicts whether accident occurs or not and also provides the overall information about these attributes which causes accidents. Hence this proposed model also maps the accident prone location as a black spot. Based

on the result, suggestions are provided to reduce the accidents in the future.

#### REFERENCES

- [1] <https://www.hindustantimes.com/india-news/road-accidents-claimed-nearly-400-lives-every-day-in-india-in-2016/story-7DlmtDnVMYLLZVGxXKOaJN.html>
- [2] <https://www.thehindu.com/news/cities/bangalore/bbmps-colourcoded-map-shows-denguerampant-areas/article7525135.ece>
- [3] [https://www.researchgate.net/publication/315796881\\_IDE NTIFICATION\\_AND\\_ANALYSIS\\_OF\\_ACCIDENT\\_B LACK\\_SPOTS\\_ON\\_NH147\\_USING\\_GIS](https://www.researchgate.net/publication/315796881_IDE NTIFICATION_AND_ANALYSIS_OF_ACCIDENT_B LACK_SPOTS_ON_NH147_USING_GIS)
- [4] <https://link.springer.com/article/10.1007/s40534-016-0095-5>