

# Analysis of Road Fatal Accidents

**Tanisha Sonawane<sup>1</sup>, Siddhi Bhamare<sup>2</sup>, Mayuri Gote<sup>3</sup>, Ujjwala Wale<sup>4</sup>, Prof. M.K.Nivangune<sup>5</sup>**

Student, Comp Dept, Stes Sinhgad Academy of Engineering Kondhwa, Pune, India<sup>1,2,3,4</sup>

Guide, Comp Dept, Stes Sinhgad Academy of Engineering Kondhwa, Pune, India<sup>5</sup>

**Abstract:** Road accidents are a very common issue nowadays. The fatal road accidents causes a life loss of a number of people per day. The present innovation is proposed to make analysis in the same matter and will bring a change accordingly. The road accidents are mainly caused due to the excess light conditions, drunk and driving, the bad road surfaces, and the worse weather conditions. These causes are the prime reason for the road accidents. In the proposed innovation, the accidents causes are made to a fair analysis using the Apriori algorithm, Naive Bayes algorithm, and the K-means algorithm. Some suggestions of the road accidents management are made on the basis of obtained statistic, association and rules, classification models, and the clusters.

**Keywords:** Roadway Fatal Accidents, Association, Classification, Clustering

## I. INTRODUCTION

The rapid urbanization, industrialization and migration along with other social changes have resulted in increasing necessity for travel across all age groups in the entire country. With poor public transportation systems and inability of people to afford cars, the personal modes of transport have increased across Indian cities and in rural areas. This increasing reliance on motor vehicles and motorcycles has also started influencing health of people in a significant way. Accident is an event, occurring suddenly, unexpectedly and inadvertently under unforeseen circumstances. An accident that takes place on the road involving a vehicle is termed as road traffic accident. Each year, road traffic injuries take the lives of 1.2 million people around the world and seriously injure millions more. Road traffic injuries are predicted to rise from ninth place in 2004 to fifth place by 2030 as contributor to the global burden of disease. The classic Apriori algorithm which is used for the frequently addressed item sets. The algorithm uses the bottom up approach and thus the frequently occurring subsets extended one at a time. The vast amount of data is classified into the clusters by making the use of k-means algorithm.

## II. LITERATURE SURVEY

### • **Analysis Of Road Traffic Fatal Accidents Using Data Mining Techniques**

They apply statistics analysis and data mining algorithms on the FARS Fatal Accident dataset as an attempt to address this problem. The relationship between fatal rate and other attributes including collision manner, weather, surface condition, light condition, and drunk driver were investigated. Association rules were discovered by Apriori algorithm, classification model was built by Naive Bayes classifier, and clusters were formed by simple K-means clustering algorithm. Certain safety driving suggestions were made based on statistics, association rules, classification model, and clusters obtained.

### • **Applying Association Rules Mining Algorithms for Traffic Accidents in Dubai**

Association rule mining algorithms are widely used to find all rules in the database satisfying some minimum support and minimum confidence constraints. In order to decrease the number of generated rules, the adaptation of the association rule mining algorithm to mine only a particular subset of association rules where the classification class attribute is assigned to the right-hand-side was investigated in past research. In this research, a dataset about traffic accidents was collected from Dubai Traffic Department, UAE. After data pre-processing, Apriori and Predictive Apriori association rules algorithms were applied to the dataset in order to explore the link between recorded accidents' factors to accident severity in Dubai. Two sets of class association rules were generated using the two algorithms and summarized to get the most interesting rules using technical measures. Empirical results showed that the class association rules generated by Apriori algorithm were more effective than those generated by Predictive Apriori algorithm.

### • **Modelling Severity Level in Multi-Vehicle Collision on Indian Highways**

Road inventory characteristics. Collision partner is the key attribute in determining the severity outcome of crashes, but severity of crashes vary even if the collision partners remain same. The influence of other factors on severity outcome

may be fixed or random. This study makes an attempt to model the crash severity outcome for vulnerable road user involving crashes, crashes involving passenger cars and crashes involving only heavy vehicles using multinomial and ordered models. The model is developed using data from National Highway-6 (NH-6, West Bengal, India), considering crash time, roadside activity characteristics and road inventory details as contributing factors. The crash severity outcomes are categorized into three categories viz. minor injury crashes, major injury crashes and fatal crashes. The study also attempts to identify factors influencing higher severity crashes.

### III. PROPOSE SYSTEM

In Proposed system we are analysing the fatal road accidents which may happens due following Reasons:

- 1) Light conditions - the immense light front he opposite vehicles on the eyes of the drivers make them unable to see for a while.
  - 2) The weather - The weather conditions like very rainy can cause the warning of the roads or there is a great risk on the mountains. This is one of the un-predictable reasons and can occur anytime.
  - 3) Speed limit - As high speed driving not only on the highways but also on the normal roads is becoming a trend. This eventually results into many road accidents.
  - 4) Drunk driving - This is a proposed theory that the driving won't be allowed to the driver, if he is drunken. The sensors would detect and would accordingly deny for the starting of the vehicle.
- From all these parameters we are plotting graph according to analyse data to show the roads having high possibilities of accidents.

### IV. SYSTEM ARCHITECTURE

Following diagram is our system's architecture diagram:

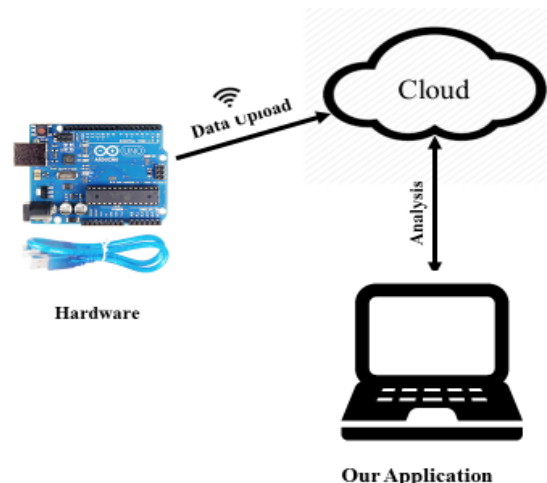


Figure 1: System Architecture

In this system we are taking input from hardwares like sensors, controllers. The data collected by hardwares get uploaded through wifi on cloud. After analysing uploaded data user can see the graph according to area or roads.

### V. METHODOLOGIES

We are going to use sensors, controllers and wifi. Therefore the study consisted of Arduino, controllers, wifi and data mining and storage with different types of algorithms related to data mining for analysis of large collected data from sensors. Arduino is an open-source platform used for building electronics projects. Arduino consists of both a physical programmable circuit board (often referred to as a microcontroller) and a piece of software, or IDE (Integrated Development Environment) that runs on your computer, used to write and upload computer code to the physical board. Arduino provides a standard form factor that breaks out the functions of the micro-controller into a more accessible package. The ESP8266 WiFi Module is a self-contained SOC with integrated TCP/IP protocol stack that can give any microcontroller access to your WiFi network. The ESP8266 is capable of either hosting an application or offloading all Wi-Fi networking functions from another application processor. The ESP8266 module is an extremely cost effective board with a huge, and ever growing, community. This module has a powerful enough on-board processing and storage

capability that allows it to be integrated with the sensors and other application specific devices through its GPIOs with minimal development up-front and minimal loading during runtime. Its high degree of on-chip integration allows for minimal external circuitry, including the front-end module, is designed to occupy minimal PCB area. The ESP8266 supports APSD for VoIP applications. Prediction in data mining is to identify data points purely on the description of another related data value. It is not necessarily related to future events but the used variables are unknown. Predictive modelling is a process that uses data mining and probability to forecast outcomes. Each model is made up of a number of predictors, which are variables that are likely to influence future results. Once data has been collected for relevant predictors, a statistical model is formulated. Cloud storage is a cloud computing model in which data is stored on remote servers accessed from the internet, or "cloud." It is maintained, operated and managed by a cloud storage service provider on a storage servers that are built on virtualization techniques. Apriori is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis.

## VI. ADVANTAGES

- We can easily identify roads or area having high possibilities of fatal accidents.
- To identify easily the place at which accidents happens due to high Light intensity, high speed, weather conditions and drunk drive detection
- We are providing system to analyse the four conditions in one system
- It will help to minimize the road fatal accidents

## CONCLUSION

Road traffic accident continues to be a growing menace, incurring heavy loss of valuable human resources, along with wastage of potential economic growth. This paper has focused some algorithm on all techniques of data mining that can be applied on IOT. In this system we are getting threshold values of parameters like Light intensity, high speed, weather conditions and drunk drive detection. Existing scenario is we need to get data separately for each condition after that we analyse that data and then get separate result for each conditions. So we minimizes our work developing one system to analyse the data of causing accidents.

## REFERENCES

- [1]. Liling Li, Sharad Shrestha, Gongzhu Hu, " Analysis of Road Traffic Fatal Accidents Using Data Mining Techniques", IEEE, 03 July 2017, DOI 10.1109/SERA.2017.7965753.
- [2]. Amira A El Tayeb, Vikas Pareek, and Abdelaziz Araar. Applying association rules mining algorithms for traffic accidents in dubai. International Journal of Soft Computing and Engineering, September 2015.
- [3]. R. V. Ponnaluri, "Road traffic crashes and risk groups in India: analysis, interpretations, and prevention strategies", IATSS Research, 35(2), 2012.
- [4]. MORTH, 2014. "Road Accidents in India 2013", New Delhi: Ministry of Road Transport and Highways Transport Research Wing, Government of India. August 2014.
- [5]. R. Bandyopadhyaya, and S. Mitra, "Modelling Severity Level in Multivehicle Collision on Indian Highways", Procedia – Social and Behavioral Sciences, 104, 1011-1019, 2013.