ISSN No:-2456-2165

Development of a Predictive Model for Road Accident Prediction in India

[1] Indraja E, [2] Smitha M, [3] Tabitha Paul R, [4] Chandini S B
Department of Information Science and Engineering
Vidyavardhaka College of Engineering, Mysuru, India

Abstract:- India accounts for about 10% of road accident fatalities of the world every year. This makes India rank 1st in road accident related deaths. There could be many reasons which lead to the occurrence of accidents in India, some of these reasons are - unfavourable weather conditions, vehicular defects, over speeding, alcohol intoxication of the driver etc. In some cases, the accident could be minor but unfortunately, sometimes it could also be fatal and can lead to major destruction of property. The main aim of the present study is to suggest the development of a predictive model (dashboard) whose goal is to predict the number of road accidents that can occur due to each of the above mentioned elements. The methodology most suitable for this is to use data mining technologies. This vital information if given to the concerning authorities, will help them take necessary precautions to prevent the incidence of accidents thereby reducing the fatality rate and destruction caused by road accidents in India.

Keywords:- *Road Accident Fatalities; Predictive Model; Data Mining.*

I. INTRODUCTION

The Transport Department of India has information of its commuters. The Transport Research Wing (TRW) of Ministry of Road Transport and Highways has aggregated road accidents related data. The data is in terms of 19 items format devised under Asia Pacific Road Accident Data(APRAD)/Indian Road Accident Data(IRAD) project of the United Nations Economic and Social Commission for the Asia and Pacific (UN-ESCAP). The information extracted from this data refers to State/Union Territory wise count of persons injured in road accidents. From this information we can confer that 150,785 people were killed another 494,600 injured, 480,600 road crashes, in the year 2016. These accidents roughly translate into 1317 crashes and 413 deaths every day and 55 crashes, 17 deaths every hour. The government would want to reduce these numbers to keep the countries residents safe and secure. Precautionary methods can be taken by the concerned authorities to reduce the accident rates. The main question for them is "When should these methods be implemented?" this is when the idea of a prediction model comes into picture. The use of data mining techniques can be useful to find some information thus predicting a result for the future.

II. LITERATURE SURVEY

The model proposed by the author uses four psychological instruments to assess the behaviour of the driver and personal traits in the hopes of identifying a relationship between the considered constraints and the occurrence of road accidents. The participants were made to fill a demographic survey and in addition to the described behaviours, correlation coefficients were calculated and four hierarchical regressions were designed to throw light on the predictive power that the personality traits can have on the occurrence of road accidents. In terms of impulsivity and aggressiveness it has been found that the personality traits have the highest influence.

The advantage of this model is that it can be mainly used in fields like the design of procedures for selecting a professional driver, building of a system for the prevention of traffic accidents and violations of law, rehabilitation of drivers who have been deprived of the driving license.

This research mainly focuses on the behaviour of the driver to draw a conclusion as to whether an accident is going to occur or not. Hence, this model cannot be deployed in case there are other factors responsible for an accident to occur. [1]

This research determines associations that road accidents have with each other and predicts the type of accidents that might occur for current roads as well as upcoming roads by finding ways in mining for frequent patterns from relevant data sets which are responsible for road accidents. For this purpose, association and classification rules have found to be the most suitable. Apriori algorithm is used to regulate associations that can be found among road accidents. Also with the help of Naïve Bayes algorithm prediction of common accidents that can occur with the cause of it being new roads was made. This system overcomes the disadvantage of the current system used by the government that maintains a ledger and the data is analysed manually. [2]

This paper mainly focuses on the road crashes that occurred in the Anambra state of Nigeria. Their main intention was to develop a predictive model with maximum accuracy for forecasting crash frequency using ARIMAX and ARIMA modelling techniques. They have considered eleven contributing factors, some of these are over speeding, weather conditions, loss of control, etc. All of these have significant effects on the number of traffic crashes. The advantage found was an increase in any of

these considered parameters will lead to the increase the number of accidents in Anambra State. Result: Finally, the time series predictive model is developed, which performs well for crash analysis. [3]

In this article, they have collected weather datasets of Visakhapatnam city for a period of 97 days, from the Indian Meteorological Department. This dataset consists of several attributes, such as wind pressure, minimum and maximum temperature, humidity, forecast and type.

The main objective of their project was evaluating weather conditions for a short period of days that is for fifteen days. This was intended towards predicting the weather forecast. The Forecasting experiment achieved this by using the ARIMA model prediction algorithm. Interventional analysis was done based upon autocorrelation models and partial autocorrelation models. These two models were used to predict wind pressure and humidity for the next 15days. The result was a comparison between predicted results and real data is displayed only for a single city. The prediction of unfavourable weather conditions can be a valuable information to take precautions to prevent an accident. [4]

Regularizing of traffic in both developed as well as under developed countries is done by traffic signals. The author has proposed a model that combines cluster analysis and Intelligent Agent Based System that regulates the traffic signal based on the traffic volume. Traffic signal changes colour in a regular interval. But an intelligent traffic light system senses the presence or absence of vehicles and reacts accordingly. it provides an intelligent transportation system to control and avoid accidents. This article gives a proposal which avoids traffic accidents from the perspective of Data Mining.

The results of clustering were presented by making use of WEKA and other dimension reduction techniques. The use of Data Mining tools such as Waikato Environment for Knowledge Analysis makes it convenient for the Transportation Department to devise a plan to avoid future accidents. [5]

This article suggests that data mining has proven to be a very useful technique for analysing vehicle crashes and to render effective outcomes. Many models use these techniques by identifying the elements that influence the criticality of the crash. A factor considered in this article is the location of the occurrence of accidents. In this paper, they have first made use of the K-means technique to cluster crash areas into three divisions, the largest number of accident-prone areas, relatively fewer number of accident-prone areas and the least number of accidentprone areas. It takes the count as a factor to group the areas. Association mining rule which is a popular technique was used to characterize these areas. While the data limits itself to only a particular type of parameter, this analysis was useful in finding hidden information. [6]

The main objective of this article is develop a simple and practical accident prediction model which can predict an expected number of accidents. This is specific to urban junctions and road links. The great deal of importance is also given to make accurate predictions. This system is also meant to help the local authorities find the causes for the accident occurrence. The model uses data from 1036 road junctions and 146 road links in the urban regions. In this model they have made use of generalized linear modelling techniques. It was found that modelling for road links involved relatively lesser efforts as compared to road junctions. The reason being that there were not enough explanatory variables for junctions. This also proved to be one of the most powerful models. [7]

Under this article, the dataset they have examined is taken from the NHAI which consists of the past vehicle crash data from 2014 to 2017. This data set consists of 19,167 accident data and ten attributes after pre-processing. The first task carried out is data pre-processing followed by the application of data mining techniques namely classifying algorithms like Random Tree and Naïve Bayes for predictions and associations between the data stored in a large database. Then the outcomes obtained from operations are visualized.

The resulting form shows the accident severity prediction. Further, the Apriori algorithm is applied to invent the relationship among various parameters that frequently affect the severity of an accident. In the end, the result is obtained from Association Rule mining and classification techniques. [8]

In this article, the dataset they have examined is taken from California Polytechnic State University and whole data originally came from FARS (Fatal Accident Reporting System) is analysed by applying statistical analysis and data mining algorithms as an effort to address the number of fatalities. Those attributes of the datasets include Speed limit, light condition, weather condition, condition of the roads. Here before implementing the algorithms, the tuples with missing value in collected attributes was removed, the numerical values were converted to nominal values. Classification of states is done using Naïve Bayes Classifier and clustering of the states using K-means Algorithm. Result of the analysis: Association rules amongst the variables, clustering of states in the USA based upon their population, number of fatal accidents based on every single attribute, and classification of the regions as existing high or low risk of disasters is displayed using bar graphs. Conclusion: The environmental factors like roadway covering, climate and light condition do not entirely affect the fatal rate, while the human factors like being intoxicated or not, and the collision type, have a greater effect on the fatal rate. [9]

This article concentrates on the PRACT (Predicting Road Accidents – a Transferable methodology across Europe) project. The aim of this project is to develop an accident prediction model structure for the continent of Europe. It is said to be applicable for the different types of

road networks all across Europe. The main ideas around which the PRACT project structure revolve around are – The development of an ideal Accident Prediction Model (APM) model with a unique set of Crash Modification Factors (CMFs) which can be valid for all of Europe and for all the different type of networks is quite unrealistic. The development of such type of a specific APM model which relies on local data is expensive even in terms of time consumption. It will also need adequate data and experience which is usually not found in most road administrations. The development of a model, "local" CMFs which are solely based on historical data sets prevents the possibility of checking for the effectiveness of new technologies. The PRACT project gives a solution to these problems by designing and developing a user friendly tool which will allow easier implementation in different countries across the continent. [10]

III. CONCLUSION

Devising a predictive model for road accident prediction could be challenging. The reason being that, an accident could happen not only based on the expected parameters, but also under unanticipated circumstances. Predicting an occurrence leads to taking precautionary measures and hence reducing the risk of a casualty. For this model data mining has proven to be a reliable technique which can be used to analyse the data and to identify the major reasons which cause road accidents. The resulting model can help enhance the road safety in India.

REFERENCES

[1]. "A Model for Traffic Accidents Prediction Based on Driver Personality Traits Assessment" MARJANA ČUBRANIĆ-DOBRODOLAC, KRSTO LIPOVAC, SVETLANA ČIČEVIĆ, BORIS ANTIĆ, University of Belgrade, Faculty of Transport and Traffic Engineering Vojvode Stepe 305, 11000 Belgrade, Serbia.

[2]. "Analysis of road accidents using data mining techniques" Poojitha Shetty, Sachin P C, Supreeth V Kashyap, Venkatesh Madi Dept. of Computer science and engineering, National Institute of Engineering, Karnataka, India.

[3]. "Road traffic accidents prediction modelling: An analysis of Anambra State,Nigeria" Chukwutoo C. Ihueze, Uchendu O. Onwurah, Department of Industrial and Production Engineering, Nnamdi Azikiwe University, Awka, Nigeria.

[4]. "An Integrated Approach for Weather Forecasting based on Data Mining and Forecasting Analysis", G.Vamsi Krishna Research scholar, Department of CSE, GITAM University.

[5]. "Application of Data Mining Techniques for Analyzing Road Traffic Incidents A Case Study of Libya's Road Traffic", Sayed Mujahed Hussain, Department of Computer Science, Faculty of Arts and Science Ajdabiya, University of Benghazi,Ajdabiya,Libya, Faraj A. El-Mouadib ,Department of Computer Science, Faculty of Information Technology, University of Benghazi, Benghazi, Libya.

[6]. "A data mining approach to characterize road accident locations", Sachin Kumar, Durga Toshniwal.

[7]. "Accident prediction models for urban roads", Poul Greibe, Danish Transport Research Institute, Knuth Winterfeldts Allé, DK-2800 Kgs. Lyngby, Denmark.

[8]. "Survey on Analysis and Prediction of Road Traffic Accident Severity Levels using Data Mining Techniques in Maharashtra, India", Baye Atnafu and Gagandeep Kaur Dept of CS/IT, Symbiosis Institute of Technology, Pune, India.

[9]. "Analysis of Road Traffic Fatal Accidents Using Data Mining Techniques", Liling Li, Sharad Shrestha, Gongzhu Hu, Department of Computer Science, Central Michigan University, USA.

[10]. "Development of a transnational accident prediction model", Francesca La Torre , Lorenzo Domenichini, Monic Meocci , Daniel Graham , Niovi Karathodorou , Thomas Richter , Stephan Ruhl , George Yannis , Anastasios Dragomanovits , Alexandra Laiou ,University of Florence, Via S. Marta, 3, 50139, Firenze, Italy, Imperial College London, SW7 2AZ, London, United Kingdom, Technische Universität Berlin, Gustav-MeyerAllee 25, 13355, Berlin, Germany, National Technical University of Athens, 5, Heroon Polytechniou st., 15773, Athens, Greece.