

Industrial Engineering Journal ISSN: 0970-2555

Volume : 53, Issue 11, November : 2024

A NEW APPROACH FOR PREDICTION OF ROAD ACCIDENTS BY USING

DATA MINING TECHNIQUES

PRIYADARSINI DASARI, Student, Department of CSE, Vikas College of Engineering & Technology, AP, India

Mr. LOKENDRA GAUR, Senior Assistant Professor, Department of CSE, Vikas College of Engineering & Technology, AP, India

E-Mail id: lokendra.gaur@gmail.com

Abstract-Due to the exponentially increasing number of vehicles on the road, the number of accidents occurring on a daily basis is also increasing at an alarming rate. With the high number of traffic incidents and deaths these days, the ability to forecast the number of traffic accidents over a given time is important for the transportation department to make scientific decisions. In this scenario, it will be good to analyze the occurrence of accidents so that this can be further used to help us in coming up with techniques to reduce them. Even though uncertainty is a characteristic trait of majority of the accidents, over a period of time, there is a level of regularity that is perceived on observing the accidents occurring in a particular area. This regularity can be made use of in making well informed predictions on accident occurrences in an area and developing accident prediction models. In this paper, we have studied the inter relationships between road accidents, condition of a road and the role of environmental factors in the occurrence of an accident. We have made use of data mining techniques in developing an accident prediction model using Apriori algorithm and Support Vector Machines. Bangalore road accident datasets for the years 2014 to 2017 available in the internet have been made use for this study. The results from this study can be advantageously used by several stakeholders including and not limited to the government public work departments, contractors and other automobile industries in better designing roads and vehicles based on the estimates obtained

1. Introduction:

The alarming rate of increase of accidents in India is now a cause for serious concern. According to some recent statistics [1], India accounts for roughly six percent of global road accidents while owning only one percent of the global vehicle population. There are a lot of accident cases reported due to the negligence of two-wheelers, whereas over-speeding is also another contributing factor. Accidents caused while under the influence of alcohol or during general traffic violations are also common. In spite of having set regulations and the highway codes, the negligence of people towards the speed of the vehicle, the vehicle condition and their own negligence of not wearing helmets has caused a lot of accidents. While the major cause of road accidents is attributed to the increasing number of vehicles, the role played by the condition of the roads and other environmental factors cannot be overlooked.

The number of deaths due to road accidents in India is indeed a cause for worry. The scenario is very dismal with more than 137,000 people succumbing to injuries from road accidents. This figure is more than four times the annual death toll from terrorism. Accidents involving heavy goods vehicles like trucks and even those involving commercial vehicles used for public transportation like buses are some of the most fatal kind of accidents that occur, claiming the lives of innocent people. Weather conditions like rain, fog, etc., also play a role in catalysing the risk of accidents. Thus, having a proper estimation of accidents and knowledge of accident hotspots and ca sing factors will help in taking steps to reduce them. This requires a keen study on accidents and development of accident prediction models.

2. Literature Survey:

https://www.statista.com/topics/5982/road-accidents-in-india/

India is one of the busiest countries in the world in terms of road traffic. The automotive industry across the south Asian country became the fourth largest in the world in 2017. In 2022, there were more than <u>3.8</u> million new car registrations in the country. The



Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Issue 11, November : 2024

Indian road network, spanning over five million kilometers, carried almost 90 percent of the country"s passenger traffic and about 65 percent of the goods. With the rapid increase in the number of cars and the mercilessly congested Indian roads, road safety has turned into a factor of utmost importance for the country"s citizens.

Srivastava AN, Zane-Ulman B. (2005). Discovering recurring anomalies in text reports regarding complex space systems. In Aerospace Conference, IEEE. IEEE 3853- 3862.

Many existing complex space systems have a significant amount of historical maintenance and problem data bases that are stored in unstructured text forms. The problem that we address in this paper is the discovery of recurring anomalies and relationships between problem reports that may indicate larger systemic problems. We will illustrate our techniques on data from discrepancy reports regarding software anomalies in the Space Shuttle. These free text reports are written by a number of different people, thus the emphasis and wording vary considerably.

Ghazizadeh M, McDonald AD, Lee JD. (2014). Text mining to decipher free- response consumer complaints: Insights from the nhtsa vehicle owner's complaint database. Human Factors 56(6): 1189-1203.

http://dx.doi.org/10.1504/IJFCM.2017.089439.

This study applies text mining to extract clusters of vehicle problems and associated trends from free-response data in the National Highway Traffic Safety Administration''s vehicle owner''s complaint database. Background: As the automotive industry adopts new technologies, it is important to systematically assess the effect of these changes on traffic safety.

3. Existing System:

Williams et al. [5] have found through their studies that the age and experience of a driver also play a major role in the occurrence of accidents. Suganya, E. and S. Vijayarani [6] in their paper have analysed the road accidents in India and compared the performance of different classification algorithms such as linear regression, logistic regression, decision tree, SVM, Naïve Bayes, KNN, Random Forest and gradient boosting algorithm using accuracy, error rate and execution time as a measure of performance. They have found the performance of KNN to be better than that of the others. **Disadvantages:**

The system doesn"t have facility to train and test on large number of numbers.

The system doesn"t measure an accurate road accident due to poor classification models.

4. Proposed System

In the proposed system, the system has built an application that is capable of predicting the possibility of occurrence of accidents based on available road accident data. Data pre-processing is done on this road accident data to obtain a dataset. The data preprocessing step includes cleaning to remove the null and garbage values, and normalization of the data, followed by feature selection, where only relevant features from the original dataset are selected to be included in the final dataset. The dataset is then subjected to different data mining techniques. Clustering is performed on this dataset. The clusters are then subjected to other algorithms like Support Vector Machines (SVM) and Apriori. Since the data being used for the study has an unknown distribution and we need to sort out the frequent and infrequent items in the dataset, the former (SVM) is used to predict the probable risk of accidents while the latter (Apriori) is applied to perform rule mining, that is, to generate a frequent item set based on given support and confidence values.

Advantages

These optimized models can be efficiently utilized by the government to reduce road accidents and to implement policies for road safety.





Fig.1: Proposed model flow



Industrial Engineering Journal ISSN: 0970-2555

Volume : 53, Issue 11, November : 2024

5. Dataset Description: Service Provider:

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Browse Data Sets and Train & Test, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View All Road Accident Prediction, Find Road Accident Prediction Type Ratio, View Road Accident Ratio Results, Download Predicted Data Sets, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user"s details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like PREDICT ROAD ACCIDENT STATUS, VIEW YOUR PROFILE.

6. Result & Analysis:



Some Trace Presedual Than	
-	deschard later
Base Service	Stree Res John . X
Lanna, Continue	See (p6) (See
Status Justices	Sectors (1995) - 1
Terrary Spec	And Party of
	And the st
1.000	(iem
Same Trans, of Street, or	- berthe Alexandria
And that A start from	

7. CONCLUSION:

An accident can change the lives of many people. It is up to each of us to bring down this increasing number. This can be made possible by adopting safe driving measures to an extent. Since all instances of accidents cannot be attributed to the same cause, proper precautionary measures will also need to be exercised by the road development authorities in designing the structure of roads as well as by the automobile industries in creating better fatality reducing vehicle models. One thing within our capability is to predict the possibility of an accident based on previous data and observations that can aid such authorities and industries. This project was successful in creating such an application that can help in efficient prediction of road accidents based on factors such as types of vehicles, age of the driver, age of the vehicle, weather condition and road structure, This model was implemented by making use of several data mining and machine learning algorithms applied over a dataset for Bangalore and has been successfully used to predict the ris probability of accidents over different areas with high accuracy.

8. **REFERENCES**:

https://www.statista.com/topics/5982/road-accidentsin-india/

Srivastava AN, Zane-Ulman B. (2005). Discovering recurring anomalies in text reports regarding complex space systems. In Aerospace Conference, IEEE. IEEE 3853-3862.

Ghazizadeh M, McDonald AD, Lee JD. (2014). Text mining to decipher free-response consumer complaints: Insights from the nhtsa vehicle owner's complaint database. Human Factors 56(6): 1189-1203.

http://dx.doi.org/10.1504/IJFCM.2017.089439.

Chen ZY, Chen CC. (2015). Identifying the stances of topic persons using a model-based expectationmaximization method. J. Inf. Sci. Eng 31(2): 573-595.

http://dx.doi.org/10.1504/IJASM.2015.068609

Williams T, Betak J, Findley B. (2016). Text mining





ISSN: 0970-2555

Volume : 53, Issue 11, November : 2024

analysis of railroad accident investigation reports. In 2016 Joint Rail Conference. American Society of Mechanical Engineers V001T06A009- V001T06A009. http://dx.doi.org/10.14299/ijser.2013.01.

Suganya, E. and S. Vijayarani. "Analysis of road accidents in India using data mining classification algorithms." 2017 International Conference on Inventive Computing and Informatics (ICICI) (2017): 1122-1126.

Sarkar S, Pateshwari V, Maiti J. (2017). Predictive model for incident occurrences in steel plant in India. In ICCCNT 2017, IEEE, pp. 1-5. http://dx.doi.org/10.14299/ijser.2013.01.

Stewart M, Liu W, Cardell-Oliver R, Griffin M. (2017). An interactive web-based toolset for knowledge discovery from short text log data. In International Conference on Advanced Data Mining and Applications. Springer, pp. 853-858. http://dx.doi.org/10.1007/978-3- 319- 69179-4_61.



Industrial Engineering Journal ISSN: 0970-2555 Volume : 53, Issue 11, November : 2024