



Road Accident Prediction Using Data Mining Model

¹Mr.P.M.Suresh, ²Dr.Subba Rao Kolavennu ,

³T.Kushuma, ⁴P.Supraja, ⁵A.Vandana

¹Computer Science & Engineering(AIML) Engineering (Assistant Professor) Sphoorthy Engineering College. (JNTUH)

^{2,3,4,5}Computer Science & Engineering(AIML) Engineering Sphoorthy Engineering College. (JNTUH)

Abstract

Road accidents generate a significant number of serious injuries reported every year and are of great concern worldwide, mainly in underdeveloped countries. Many people have lost loved ones because of accidents on this road. Therefore, a system capable of saving lives is needed. The system detects the important factors that contribute to the accident or associates the accident with various accident factors. This study offers an accident prediction system that can help analyze potential safety problems and predict whether accidents will occur.

1.Introduction

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. Data-mining techniques have evolved significantly in recent decades and are being widely applied to several real-world problems. Current data-mining methods can be used on a database to rapidly extract knowledge that can help to guide policing methods and thus improve accident-prevention techniques and awareness campaigns produced by the security forces.

2. Literature Survey

We have surveyed the existing projects and finally thought of making necessary modifications for getting the latest edition. The steady increase in the rate of accidents in India have prompted many researchers to look into the factors affecting road accidents and study about it. Since data mining techniques do not require certain assumptions between dependent and independent variables which are required in traditional statistical techniques, various categories of data mining techniques have been made use of in creating prediction models for road accidents in the past. Researchers have focused on different sets of attributes in developing such models. Srivastava et al. [2] and Ghazizadeh et al. [3] have mainly concentrated on studying the accidents occurring at intersection points.

Existing System:-

Over the past years, multiple approaches have been proposed to solve the problem of road accident prediction. The traditional

analysis method mainly depends on database system and the education of customers. The database system are limited to size, inaccurate and takes more time for huge data set. Database systems can process only structured data Therefore using a traditional database will not be efficient. Road accidents are a very difficult challenge in almost every part of world. In most part of the world, local authorities will put sign boards and put speed limits. Several attempts to use ML also has been attempted

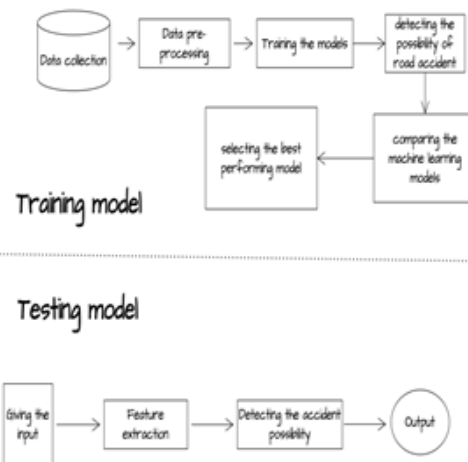
Proposed System:-

In this paper, we have 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). Clustering is performed on this dataset. The clusters are then subjected to other algorithms like Support Vector Machines (SVM).

Multiple Machine Learning Models will be trained and compared to find the one with best performance. The system will be able to predict the possibility of Road Accident by using the given inputs like Location, Time etc.

The datasets utilized for making the system is a Government available datasets for a particular area. We will be using the datasets, pre process the data and train multiple models to find the best performing one from them and gives best result

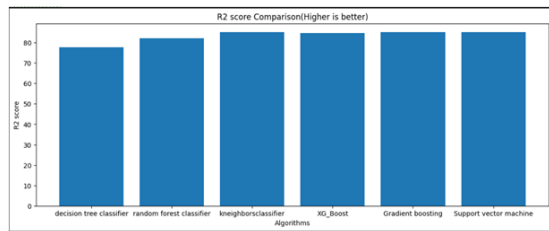
4.IMPLEMENTATION



To implement a well-designed road framework management system for looking into road security aspects, it is often desired to have an optimized accident prediction model which can analyze potential issues arising due to infrastructure fallbacks and to estimate the effect of existing models in reducing the occurrence of accidents. The main challenges involved in the creation of such a model include the evaluation of the weight that can be attributed to the impact of each variable in contributing to the accident and assessing how the model can be best designed to incorporate the effects of all such variables. Data mining techniques and models have in the past been found useful for the purpose of data interpretation in a variety of domains including but not limited to credit risk management, fraud detection, healthcare

informatics, recommendation systems and so on.

5.Results-



6.Conclusion-

In our proposed system, we are training different machine-learning algorithms to predict road accidents. & finally, k-nearest neighbour & XG-boost give the best performance in predicting road accidents.

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.

7.References -

<https://towardsdatascience.com/live-prediction-of-traffic-accident-risks-using-machine-learning-and-google-maps-d2eeff9389e>

[1]

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

[2] Srivastava AN, Zane-Ulman B.

(2005). Discovering recurring anomalies in text reports regarding complex space systems. In

Aerospace Conference, IEEE. IEEE 3853-3862.

[3] 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>.

[4] 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>

8.Appendix-

SVM =Support vector machine

KNN=K-Nearest Neighbor algorithm

XG Boost = Extreme Gradient boosted decision tree