Predicting Road Accidents Using Machine Learning

Keerthi Ellur Arun, B S Sarika, Zuber S, Sahana Udupa S V, Prof. Swathi Darla
[1,2,3,4]Student, R V Institute of Technology and Management, Bangalore, Karnataka
[5]Professor, R V Institute of Technology and Management, Bangalore, Karnataka

ABSTRACT: Traffic accidents are one of the world's most pressing concerns, results in countless casualties, injury each year and notable economic, social losses. These are one of the multiple factors that contribute to traffic accidents. If these are understood, and predicted, it will be possible to take measures to diminish the severity of the damages. elements Automobile engineers and researchers have attempted to design , build safer vehicles, but traffic accidents are inevitable. Dynamics in catastrophic crashes could be predicted if we develop effective predictive approaches to automatically classify the type and degree of harm sustained in various traffic accidents. The goal of this research is to identify these indicators using accident dataset from the district of Setubal, Portugal, from 2016 - 2019. The objective of this study is to build a model that can select a group of major determinants.
that can be used to categorize the enormity of an accident, enabling an evaluation of the accident data to be executed. These cognitive and roadway disaster patterns can help lawmakers develop traffic safety protocols. The performance of four machine learning paradigms - Random Forests , Logistic Regression , and Naive Bayes used to model the severity of injury sustained in traffic accidents are summarised in this research. We considered support vector machines, mixed neural networks, machines, decision trees, and a hybrid model that combines neural networks and decision trees concurrently. Results from experiments show that the hybrid random forest network strategy for better performance the single approaches among the machine learning paradigms taken into consideration. Additionally, this research also proposes a predictive framework for potential road accidents based on past data

KEYWORDS: Machine Learning, Analysis, Prediction, Classification Techniques.

I. INTRODUCTION

Road accidents are responsible for causing significant harm, including physical injuries, fatalities, long-term disabilities, and damage to property. The impact of road accidents goes beyond the economy and extends to the healthcare system by placing a strain on hospitals. To reduce the severity of road accidents, it's important to identify the primary factors that contribute to their severity. Accidents don't occur randomly; rather, they follow identifiable patterns that can be anticipated and prevented. In recent years, data mining has become a hot topic in the information industry because of the abundance of large datasets and the pressing need to transform that data into valuable insights and knowledge. Classification algorithms are employed to categorize vast amounts of data and generate insightful outcomes. The incidence of road traffic accidents is primarily influenced by the geometry of the road, traffic patterns, driver behaviour, and road environment. Numerous studies have been undertaken to forecast accident rates and examine various aspects of traffic accidents, such as identifying hazardous locations or hot spots, analysing the severity of injuries, and assessing accident duration. By extracting data from police reports and state highway-asset-management databases, researchers can analyse traffic safety and predict the likelihood of a traffic crash by examining the relationship between influencing factors and crash outcomes. Accidental severity analysis considers three factors: the count of injuries, the number of casualties, property damage and severity. The authors analyse severity levels independently and consider four options: light injuries, severe injuries, fatalities, and property damage. The level of accident severity is classified into three categories: injury, possible injury, and property damage. Over the past twenty years, accident severity has become a popular area of research. Researchers have been using various statistical approaches to classify road accidents. By using these techniques, the causes of road accidents can be analysed.

The goal is to identify the factors that contribute to an increased likelihood and severity of accidents, create models that can predict both the count and fatality of accidents, and testcase of these models to predict the probability of accidents occurring on specific road area occurring . The format of this essay is as follows: a review of the literature, a methodology, a conclusion, and future work.

II. PROBLEM DEFINITION

Implement a machine learning system that predicts almost prediction of accident using various types of machine learning algorithm.
The goal of the research paper is to get higher prediction analysis. And the major goal is to predict accident and provide precautionary

III. LITERATURE SURVEY

This paper majorly deals with data mining with wide availability of huge amount of data classification, algorithm are used to classify large volume of data, hence data mining are used. Nearly 12 million suffer death rate in accident over 20 and 50 million for non-fatal injuries and mostly 10-19 years children are high among death rates.

Here author [1]uses two popular machine learning algorithm are used Support Vector machine and Artificial Neural Network are used as parameters. Optimized algorithm that is genetic algorithm and particle swarm optimization [PSO] for higher rate accuracy. PSO bases SVM does perform better than other algorithm with high accuracy and robustness.

Analysis for the paper is done in district of Setubal MOPREVIS [modelling and prediction of road accident in Setubal]. This paper contains team study from Concordia University

Team, North South University study for 2015 using Random forest[2], Montreal Vehicle Collision ,XGB boost ,baseline model .Moreover 85% accident occur by Montreal incidents also with negative report of 13% by decision tree[C5.0] model gives more accuracy Algorithms like BRF,RF,XGB -85% RF,KNN,NB,DT,RF,SVMLR,NN,DNN-77%.Data set include severity ,light vehicle, country. Clustering with specific partition with one criterion function model results are given. Clusters involving rules, errors , classes, results tells 21% of accidents and fatalities. Dataset contains 28,102 accidents and random forests contains highest number of accuracy. Hence by final discussion we find the pattern of accidents occurring.

This paper explains software requirement specifications. It also explains IEEE standard specification document that explains the estimation of roads which is poor component in entire world. From road designer to road safety management analysation of fatal injuries and also a complete understanding of accident prediction model. In national road administration dataset contains signin, upload dataset, data visualization, system design, user interface, road accident analyser system. Prediction by KNN, decision tree, naive bayes.

The increase of accidents worldwide needs to be prevented and the severity should be predicted so that the rate of accident being happen should be decreased. The main thing is to analyse the accident areas and get data from those areas so that we can use different machine learning techniques for predicting severity so that the occurrence of accidents should be minimized. The severity should be predicted upon the most correlated attributes such as speed limit, whether-conditions, Light-conditions and Road-surface-conditions, these are the main factors which will affect more on predicting severity level. The other algorithms such as neural networks, SVM also has good ability to predict the severity level but here the random forest is simple and easy to train the dataset and also based on the N number decision tree contribution so that It always ensure that it will predict the right severity level.

The research paper "Road Accident Analysis Using Machine Learning" aims to develop a model that can accurately predict the likelihood of road accidents using machine learning techniques. The study utilizes data from the Indian state of Karnataka to train a decision tree algorithm to identify the factors that contribute to road accidents, including the type of road, time of day, weather conditions, and other environmental factors. The model was trained on a dataset of 12,000 accident records and achieved an accuracy rate of 94.1%[4], indicating that it can be a reliable tool for predicting the likelihood of road accidents. The study also highlights the potential benefits of using machine learning in road safety analysis, including the ability to identify high-risk areas and develop targeted interventions to prevent accidents .The research that machine learning can be an effective tool for improving road safety and reducing the number of accidents on our roads.

One unique aspect of the paper was the use of decision tree algorithms to analyse the factors contributing to road accidents. This approach allowed the researchers to identify the most significant factors that influence the likelihood of an accident, which can help inform targeted interventions to prevent accidents.

However, it only analysed data from the Indian state of Karnataka, which may not be representative of road safety patterns in other regions or countries. Additionally, the study focused on identifying factors that contribute to accidents but did not investigate potential interventions or strategies to mitigate these risks. Further research is needed to explore the effectiveness of different interventions and strategies for improving road safety.

Road safety is a important part of our lives, therefore it is necessity to diminish traffic accidents and their injury severity. In this paper, researchers used unsupervised learning methods with Association rule mining to discover the design of traffic accidents and fatalities from many parts datasets are used for severity prediction. The proposed work is a real-time application that helps the community reduce the frequency of severity occurred and further what actions should be taken to the severity of those incidents' casualties.

The current approach relies on less accurate machine learning methods to identify accidents, accident cars, and injuries, including K-nearest neighbour, random forest , and decision trees. Most unsupervised learning algorithms with association rules just require data sets without training or labelling and may quickly find the most significant links among data elements. The main thing is to analyze the accident areas and get data from those areas so that we can use different machine learning techniques for predicting severity so that the occurrence of accidents should be minimized

In order to identify patterns in traffic accidents and injury severity, researchers in this paper used four different types of algorithms, including Apriori, Apriori Transition ID, set operation for frequent itemset using transaction database-SFIT, and equivalence class clustering and bottom-up lattice traversal-ECLAT.

The simulation results of the proposed methods used in the project for analysing traffic accident severities are the Apriori algorithm find details in 962

milliseconds[7], the Apriori TID algorithm finds in 557 milliseconds, the SFIT used 516 milliseconds and the ECLAT used 124 milliseconds to find patterns. This shows the ECLAT algorithm is superior to find frequent item sets are distinguished by other algorithms because the éclat algorithm takes less frequency to find traffic accident severity patterns from the given accident datasets.

Researchers also suggested that the system proposed in the future can be used by modules like the Public-notification-modules in the project, Query Module and identifying various causes for accident severity.

A hybrid approach for road accident prediction using machine learning and deep learning algorithm

Multilayer perception contains input neurons and one hidden layer and output neuron. The feed signals are added to hidden layers and is fed back to output layer. Tuples of weighs training data is classified. The disadvantage is backward propagation give local minimal values because of random weights they also detect global features. Neural analysis gives 65 hidden neurons giving 63.86% percent for possible injury, for middle level injury the testing performance is 63.56% and for fatal injury testing performance is 78.61%. Other algorithm is classification and regression (CART) for binary decision. Single values are given to each branch. CART gives accurate value by grouping the data. Usually, CART dig for attribute and searches node. The size of the tree increases or decreases. Support Vector Machine (SVM) is another important application used almost in most computer-based application. Usually, SVM is mathematical programming with linear equation, kernel function uses hypothesis space. Hybrid decision tree merges and separate learning model. Performance increases by individual learning. Hence traffic accident predictions are generated.

IV.PROPOSED SYSTEM

Proposed system is system specification or learning the is used in machine learning model.

It also includes what new features should be added so that product gets updated. It includes the following steps dataset collection, pre-processing machine learning model developing, feature selection and final insights. Methods used are random forest, logistic regression, decision tree, nearest neighbour algorithm. The system interface includes system authorization by user id and password.

Interface requirement include sign in for user manipulation, upload dataset base on parameters, data visualization for predicting results with visual graphics, sign out to redirect the page. System design include user interface, database layer, road accident analysis system where classification and prediction model is used.

Classifications are used for grouping, selecting data based on parameters. It includes discrete and classified data mainly used for data management. Prediction is used to find output of the given dataset with trained algorithm for required process. It can be carried out Decision tree, K nearest, adaboost.

Advantages

1. It helps us to predict the reason of the accident and the fatality of the road accident . It also helps us to predict the pattern of accidents occurring in the road.
2. It helps us to examine weather conditions, road status, severity of accidents, factors that are prone to accidents.

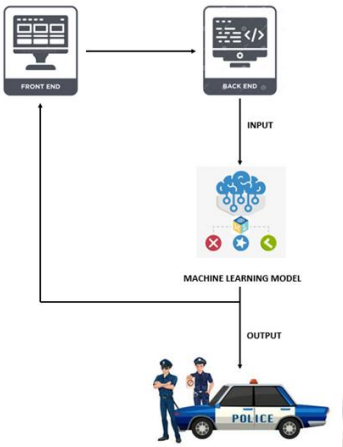


Fig 5.1 System Architecture

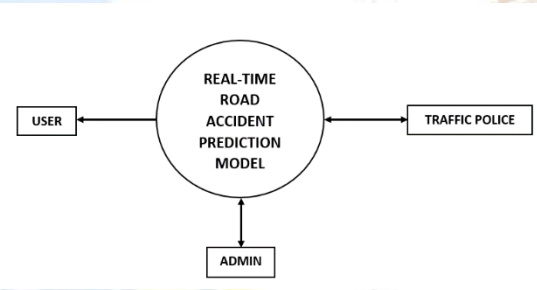


Fig 5.2 DFD Level-0

Fig 5.2 tells us modules used for the analysis , system architecture uses front-end, back-end , back-end datasets are added for implementation using machine learning

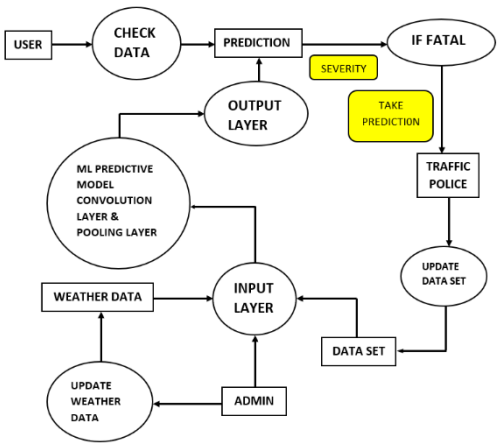


Fig 5.3 DFD Level-1

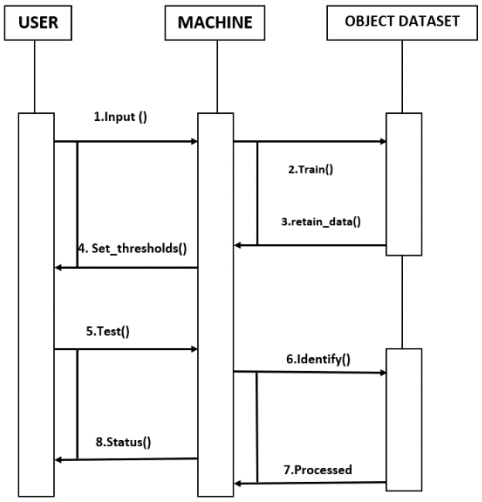
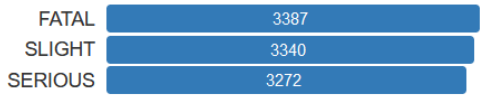


Fig 5.4 USE CASE Diagram



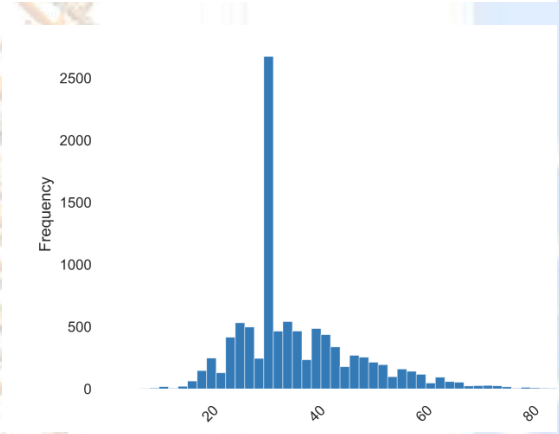


Fig 5.6 Frequency of driver's age

VI. CONCLUSION

Losing your life for unnatural death like road accidents is very painful for the society. By 2030, accidents will be one of the fifth leading death cause rate. So, by early prediction we can take a call for precautionary measure, warning, signatures that may lead us to stop or prevent accident. If detected early by machine learning algorithm the cause can be reduced and damage can be prevented. This paper investigates on how machine learning models are used and how accuracy can be increased by different algorithm. In future we can use GPS, IOT and mobile based applications to contact or predict what is the traffic situation and how it can tackle and how it can be reduced.

VII REFERENCES

[1] Sarker, I. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput.Sci.2021,

[2] Md,Farhan Labib , Ahmed Sady Rifat, Mosabbir Hossain, Amit Kumar Das, Faria Nawrine "Road Accident Analysis and Prediction of Accident Severity by using Machine Learning in Bangladesh", 2019 IEEE 978-1-7281-1557-3/19

[3]. Priyanka A. Nandurge & Nagraj V. Dharwadkar "Analysing Road Accident Data Using Machine Learning Paradigms", 2017 IEEE pp. 604-610.

[4]. Silva, P.; Andrade, M.; Ferreira, S. Machine learning applied to road safety modelling: A systematic literature review. J. Traffic Transp. Eng. (Engl. Ed.) 2020, 7, 775–790

[5]. Theofilatos, A.A.; Chen, C.; Antoniou, C. Comparing Machine Learning and Deep Learning Methods for Real-Time Crash Prediction Transp. Res. Rec. J. Transp. Res. Board2019, 2673, 036119811984157

[6] Hemanth Kanchi,Deepthi Kandagatla, Akshay Baheti Dr. M Shailaja"Predicting Traffic Accidents And Their Injury Severities Using Machine Learning". Vol 10 june 2022

[7] H.L Gururaj, V. Jhanavi ,U Tanujha **"**Prediction of Traffic Accident Severity Based on Random Forest".