Probablity Prediction for Road Accidents & Traffic Congestion using Support Vector Machine

Adheena Babu
UG Scholar, CSE, Mangalam College of Engineering, Kerala.

Aishwarya Lakshmi
UG Scholar, CSE, Mangalam College of Engineering, Kerala.

Ganga P Chandran
UG Scholar, CSE, Mangalam College of Engineering, Kerala.

Athulya Surendran
UG Scholar, CSE, Mangalam College of Engineering, Kerala.

Nimmymol Manuel
Assistant Professor, Dept of CSE
UG Scholar, CSE, Mangalam College of Engineering, Kerala.

***Abstract*— In this era of fast-moving automobiles, the rate of traffic congestion and accidents occurring on a daily basis is also increasing at an alarming rate. With the rising number of congestions, accidents, and death these days, the transportation sector's ability to predict the number of traffic accidents throughout a specific time is critical for making accurate decisions. In this case, it would be vital to carry out a framework that could distinguish the reasons for the ongoing congestion, and accidents and alert the driver about them. This paper consists of two models, i.e. traffic congestion and road accident prediction, incorporated into a single system. On the app side, information is fed into the model by the user noting questions regarding the road conditions, driver age, and so forth. On the server-side, a machine learning model created using the Support Vector Machine library, is hosted and predicts the data in view of that model. A python-based web framework, Django is used here to deploy the model.**

Keywords*— *Support Vector Machine (SVM), Traffic Congestion Prediction, classification algorithms, accident analysis

I. INTRODUCTION

In recent years, the quantity of traffic is speedily increasing, and also the alarming rate of accidents increasing in India is currently a cause for serious concern. In keeping with some recent statistics, India accounts for roughly 6% of worldwide road accidents whereas owning only 1% of the world vehicle population. There are a lot of accident cases rumoured because of the negligence of two-wheelers, whereas over-speeding is additionally another tributary factor. whereas the key explanation for road accidents is attributed to the increasing range of *vehicles*, the role compete by the condition of the roads and different environmental factors cannot be overlooked. The number of people who died in India as a result of traffic accidents is alarming. The situation is bleak, with more than 137,000 people killed or injured in traffic accidents. As a result, having a proper estimate of accidents as well as knowledge of accident hotspots and contributing factors will aid in taking steps to reduce them. This necessitates a thorough investigation of accidents and the development of accident prediction models. It is frequently desired to have an optimised accident prediction model that can analyse potential issues arising due to infrastructure fall backs and estimate the effect of existing models in reducing the occurrence of accidents in order to implement a well-designed road framework management system for looking into road security aspects. The evaluation of the weight that can be attributed to the impact of each variable in contributing to the accident, as well as determining how the model can be best designed to incorporate the effects of all such variables, are the main challenges involved in the creation of such a model.

For this paper, we have investigated the inter-relationship between the occurrences of road accidents and the roles played by the underlying road conditions and environmental factors in contributing to the same. Since such a study requires us to cover several aspects affecting accidents, we can make use of data mining techniques to analyse this data to extract relevant details from them, as these huge volumes of data would otherwise be meaningless without the right interpretation applied to them. Congestion and the problems that it causes are the banes of our time, particularly in developing countries with poor infrastructure. Traffic congestion impact in terms of lost productivity, energy costs, health risk cost through pollution, safety costs, and other factors are grossly underestimated, and it presents a critical opportunity to raise the standard of living and improve the quality of life in general. The accurate estimation of traffic flow in a given region at a specific interval of time in the future is the problem of traffic flow prediction.

II. RELATED WORKS

Several algorithms and strategies have been employed by many researchers to anticipate traffic accidents and congestion on roads and in rural regions. The passengers were able to prevent accidents in traffic regions thanks to the unintentional data. Some of the literary works are described in this section. Data mining and clustering techniques were employed by the Department of Enterprise Engineering at the

University of Rome Tor Vergata [2] to investigate accident statistics from the 15 districts of Rome Municipality from 2016 to 2019. They gathered information on 97,297 road accidents from a nearby police station. The kind (e.g., rear-end collision), the number of individuals involved, the time and date, the location, the severity (unharmed, confidential prognosis, deceased, other), and the type and number of involved cars are all available for each accident. Other data is based on the features of the road, the weather, the illumination, and the traffic conditions. For descriptive analysis, the author utilized the K-means method and the Kohonen network algorithm, which are excellent for identifying clusters of data objects that are more similar to each other, and for predictive analysis, the author used the C5.0 and CHAID algorithms. By mapping a series of input values to an output value, they were utilized to predict future events. [3] To predict the likelihood of accidents and accident-prone locations, data mining tools and techniques are used. By evaluating the severity of accidents based on the type of accident, type of spot, and utilizing the R tool, the study sheds light on projecting the probability of accidents on roads, with a special emphasis on STATE HIGHWAYS and ORDINARY DISTRICT ROADS. RStudio is an integrated development environment for using R. The limitation of this approach is that, while exploratory visualization techniques depict many important aspects such as frequency distribution of large data categories and summarize datasets in pictorial form, additional explanation is required to expose key trends, hidden patterns, classify new instances, define similar and homogeneous, and interesting relationship areas, which can be accomplished using various algorithms such as K nearest neighbor and K-means. Using correlation analysis and exploratory visualization approaches, the incidence of traffic collisions on roads was examined. The road conditions that aid in determining the relationship between two numerical variables are investigated in correlation analysis. The research's specific goal is to use exploratory visualization techniques and regression analysis to assess and predict black patches, accident prone zones, and road conditions. The authors proposed a "Convolutional Neural Network model" for deep learning-based new road traffic accident prediction in [4]. The paper uses the India Accident 2016-2018 data set from the Java high-level neural network API framework to simulate the proposed traffic accident prediction system. When the suggested Convolutional neural network is compared to a classic neural network-based back propagation prediction approach, the proposed CNN model outperforms the traditional neural network-based back propagation prediction method in terms of prediction accuracy and loss. The authors found that the suggested CNNs prediction technique has less loss and higher prediction accuracy than the conventional machine learning prediction algorithm based on the experimental findings. The authors of [5] focused on using auto-machine learning for hyper parameter adjustment to learn traffic statistics in highway system key regions. Deep learning models, in particular, have lately been proposed as emergent traffic forecast tools. Convolutional Neural Networks , Recurrent Neural Networks , Long Short-Term Memory , Restricted Boltzmann Machines , and Stacked Auto Encoders are examples of deep learning architectures .The goal of the article was to develop a DL approach for predicting congestion in highway networks utilizing LSTM-RNN models. They first proposed the Hyper Net framework, which they used to create a deep learning model with a long short term memory network based on the Hyper Net framework for learning the temporal variation of traffic datasets at major highway traffic systems. They collected highway traffic data from the Korea transportation system's main routes. The research [6] developed a model for traffic congestion prediction that is built using a machine learning classification technique called random forest to forecast traffic congestion state. The random forest algorithm is known for its great robustness, high performance, and practicality. To create a road traffic forecasting model, weather conditions, time period, special road conditions, road quality, and holidays are used as model input variables. The prediction accuracy of the traffic prediction model created using the random forest classification method is 87.5 percent. The paper [6] established a successful method for traffic congestion forecasting, as well as a scientific foundation for traffic management to avoid traffic congestion the objective of [7] is to discover factors using accident data from the district of Setbal, Portugal, from 2016 to 2019. The authors suggest a predictive model for future road accidents based on historical data, as well as constructing models that can pick a set of relevant elements that may be used to determine the severity of an accident, supporting a study on the accident data. These models are created using a variety of machine learning techniques. Unsupervised machine learning techniques such as DBSCAN and hierarchical clustering are employed alongside supervised machine learning methods such as decision trees, random forests, logistic regression, and naive Bayes. As a consequence, a rule-based model employing the C5.0 algorithm was able to reliably determine the most important elements indicating the severity of a road collision. Furthermore, the predictive model's findings show that the RF model could be effective for predicting accident hotspots.

III. PROBLEM DEFINITION

To manage the massive wide variety of avenue injuries in a locality a unique analysis is required. This analysis could be accomplished greater deeply to determine the intensity of the street accidents by using data mining and machine learning algorithms. In this paper, this will classify the severity of the accidents as fatal, bone breaking, deadly, simple injury and motor collision. Many of companies specifically authorities agencies are become aware of the elements that make a contribution to the accident roads or highways. The measurements to prevent twist of fate speed reduction, widen divider, or different else. The model that has been developed to research the accident can examine all of the variables. It is hard for the model that has been developed because of infrequently complicated mathematical mode. In this project, we are attempting to create an app including two models, i.e. traffic congestion and road accident prediction. This will help to analyze traffic congestion and road accidents in a single system.

IV. METHODOLOGY

The system introduces a model that could predict the probability of occurrence of both traffic congestions and road accidents. The system analysis the circumstances of a road accident and other factors that leads to an accident using the Support Vector Machine library. The system also predicts traffic congestion and alerts the driver to avoid traffic jams. Most data mining algorithms have been used in previous research on road accidents, and based on a literature review, clustering, classification, and linear regression algorithms have been applied to road accident data for analysis. But they haven't achieved to introduce a system that predicts accidents and congession. The system consist of an user interface in the form of a Flutter App by using Dart programming language. The app requires the driver to go through a one-time registration, where the driver has to provide various details like his age, gender, experience, type of vehicle etc. The data is sent to the server using Django.The app then studies the location of the driver using location API and weather API. It then understands the traffic congestion situation and accident prediction of that area using Support Vector Machine library. The predicted data is given to the user in the form of alert messages in the app.

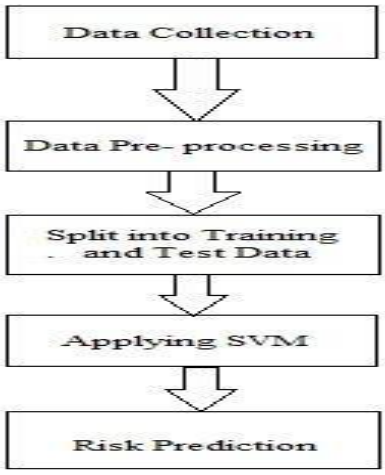

Fig. 1. Model processing stage

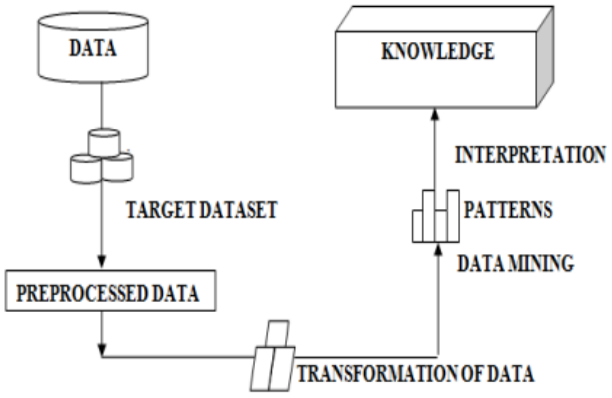

Fig. 2. System architecture

CONCLUSION

The project demonstrates the working of the APP with a year-based dataset.The data is fed into a Support Vector Machine that classifies the data to deduce the needed result.The result is then reached to the user using an APP created by the framework Django and hosted using the same.The proposed system is a demonstration of the actual app. Here, we deal with the data at a particular location, i.e., our current location.In the actual app, we will collaborate with location and weather APIs to predict the traffic congestion and accident prediction level at the live location.The proposed system introduces a model that could predict the probability of occurrence of both traffic congestions and road accidents.The system will analysis the circumstances of a road accident and other factors that leads to an accident using the Support Vector Machine library. The system also predicts traffic congestion and alerts the driver to avoid traffic jams.

REFERENCES

[1] Dhanya Viswanath, Preethi K,Nandini R,Bhuvaneshwari R, "A Road Accident Prediction Model Using Data Mining Techniques", Preceedings of the fifth International Conference on Computing Methodologies and Communication (ICCMC 2021) IEEE Xplore Part Number:CFP21K25-ART.

[2] Ms. Gagandeep Kaur, Er. Harpreet Kaur ," Prediction of the cause of Accident and Accident Prone Location on roads Using Data mining Techniques", IEEE – 40222.

[3] Prajakta S. kasbe, Apeksha V. Sakhare, "A Review On Road Accident Data Analysis Using Data Mining Techniques", 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).

[4] Amani Thaduri, Vijayakumar Polepally, SwathyVodithala," Traffic Accident Prediction Based on CNN Model", Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021) IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2.

[5] Daniel Santos * , José Saias, Paulo Quaresma and VítorBeires Nogueira," Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction", Informatics Departament, University of Évora, 7002-554 Évora, Portugal; jsaias@uevora.pt (J.S.);