Development of Urban Road Accident and Forecasting Model

Ms. S. Revathi[#1], Joselin kiruba. J [#2], Gayathri. J [#3], Karthikeyan. S [#4]

[*1]Assistant professor, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

[*2]UG Student, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

[*2]UG Student, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

[*4]UG Student, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract

Road accidents are one of the most relevant causes of injuries and death worldwide, and therefore, they constitute a significant field of research on the use of advanced algorithms and techniques to analyze and predict traffic accidents and determine the most relevant elements that contribute to road accidents. The research of road accident prediction aims to respond to the challenge of offer tools to generate a more secure mobility environment and ultimately save lives. This aims to provide an overview of the state of the art in the prediction of road accidents through machine learning algorithms and advanced techniques for analyzing information, such as Long Short-Term Memory networks, among other deep learning architectures. This project has been done to analyze traffic accidents more deeply to determine the intensity of accident by using machine learning algorithm.

***Keywords*-Long Short -Term Memory, Road accidents.**

I. INTRODUCTION

Transportation, like any other industry and phenomenon, along with its advantages, has its disadvantages and limitations for road users. Traffic and its related predicaments have been on the rise all over the world and have had a detrimental effect on the lives and property of the people of the community. Urban pollution, rising fuel and energy consumption, wasting millions of hours a day in traffic congestion, wasting community service facilities and national assets, and, ultimately, the occurrence of accidents resulting in injury, death, and property damage are the result of poor traffic facilities and conditions. Road traffic accidents now represent the eighth leading cause of death globally. In addition, road traffic fatalities have increased to 1.35 million a year and caused up to 50 million injuries in 2016; that is, nearly 3700 people die on the world's roads every day. In Iran, the reported number of road traffic deaths is 15932 in 2016 statistics also reveals the high number of injury and property damage only (PDO) accidents in urban and suburban roads in the country.

Therefore, the effect of factors affecting the severity of accidents should be investigated meticulously to provide practical solutions for improving safety and reducing the high number of accidents.

An improvement in socio-economic conditions of the people along with industrial and infrastructural development is bound to create an additional burden on roads by means of an increased number of vehicles and associated modes using the roads. Along with the enormous advantages caused by this magnitude of progress, roadway accidents have also become one of the main causes of concerns due to both individual (persons) and economic losses. Therefore, there is a huge demand for assessing these accidents through identifying and analyzing the various causes that are responsible for their occurrence, and also to recommend the various remedial measures to mitigate the accidents. Generally, road accidents are analyzed by means of precisely

defining the event involving damage to the property and/or injury to the road users, which are recorded first-hand by the police and/or emergency services. Accidents are rarely caused by a single factor. Usually, the interaction of the diverse set of factors such as roadway design parameters, road user behavior, environmental conditions, etc., cause accidents; however, one factor can be more responsible than the rest, and can easily be identified

II. PROPOSED SYSTEM

Our proposed model is a prediction of road accidents using accident data records which can help to understand the characteristics of many features like latitude, longitude, roadway conditions, light condition, weather conditions, road side and so on.

Machine learning will help to train the model. Model is trained by using LSTM, after training it helps to predict the severity. It can be illustrated using graphical representation. The model is performed to identify statistically significant factors which can be able to predict the probabilities of crashes and injury that can be used to perform a risk factor and reduce it.

III. METHODLOGY

A. Normalization

The raw data are processed using min-max normalization to eliminate dimensional differences. A linear transformation of the original data causes the result to fall into the [0, 1] interval, and the conversion formula is as follows:

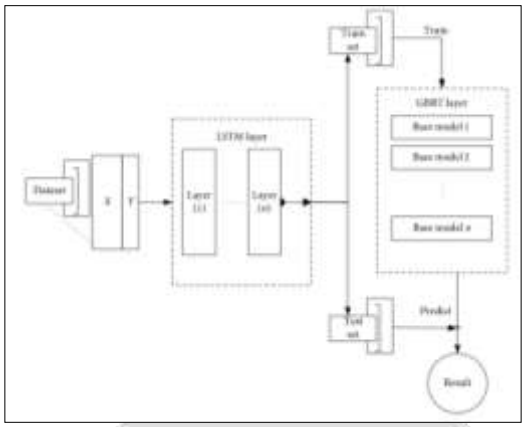
$$X = \frac{x - \min}{\max - \min},$$

where max represents the maximum value of the feature in the sample data and min represents the minimum value of the feature in the sample data; x represents raw data, and X represents normalized data.

B. LSTM

LSTM is one kind of RNN. It is a powerful deep learning method for time series data since its unique design (Hochreiter & Schmid Huber, 1997). RNN suffers from the problem of vanishing gradients, which happens to learning of long data sequences (Aebel, 2018). LSTM can solve it by introducing the memory cell to determine when to forget certain information. A LSTM network is composed of the input layer, the hidden layers, and the output layer. The main characteristics of the LSTM are the memory cells in its hidden layers, which contain memory blocks rather than traditional neuron nodes (Olah, 2015). Each block has several self-connected memory cells and three multiplicative units, input, output and forget gates. These gates provide continuous analogues of write, read and reset operations on the cells.

The LSTM neural network is capable of capturing time-dependent information and has an excellent effect on time series prediction, but it is insufficient in predicting inflection point data. The LSTM neural network is used to extract the features with time-dependent information. The structure of the LSTM model is shown in below figure.



There is no clear theoretical guidance for determining the number of nodes in the hidden layer. In general, use the following formula to select the number of nodes:

$$N = \sqrt{n+m} + a,$$

where N is the number of hidden nodes; n is the number of input nodes; m is the number of output nodes; and a can take a constant of 1 to 10.

In this project, there are 13 input nodes and 1 output node. According to formula, the number of hidden nodes is 5~13. Try a different number of hidden layer nodes using 1 layer of LSTM and judge the degree of deviation according to the error rate and root mean square error, so as to select the number of hidden layer nodes.

The experimental results of the test set show that the LSTM model using 11 hidden nodes has the smallest RMSE value and the best prediction effect. The detailed error rate and root mean square error results of the test set.

Since there are only 19 records in this example, the model depth is too high, which will cause the data to be overfitting. The experiment uses 1~5-layer models for comparison, and 11 hidden nodes are used for each layer. After training, the model performance is judged by calculating the root mean square logarithmic error and the decision coefficient of all records.

The smaller the RMSLE model, the better the fitting effect. The closer the R-square is to 1, the stronger the ability of the variable to interpret *y* and the model fits the data better.

IV CONCLUSION

The proposed model on prediction of road accident in India is developed with Long Short-Term Memory algorithm. The datasets are trained by pandas which gives a good result with perfect accuracy. The data is trained and tested for the model to gain good accuracy. The other researchers have many problems in output, using the dataset, only some were able to get better accuracy. Wrong data are removed successfully from this model since the dataset used is collected from various other sources. The parameters which are used in the dataset is pre-processed well to get better accuracy of the result. After undergoing various training and testing, then the Graph will be displayed with the predictive accuracy, then the accuracy will be chosen and concluded.

V REFERENCE

[1] Dr. L. R. Kadiyali, “Traffic Engineering and Transport Planning”, Khanna Publishers,7th Edition,

2009, pp. 411- 483.

[2] Dr. P. Pramada Valli, "Road accident models for large metropolitan cities of India", International Association of Traffic and Safety Sciences, vol. 29, no. 1, 2010, pp. 57–64.

[3] "Road Accidents in India", A report by Ministry of Road Transport and Highways, 2011.

[4] T.V. Ramachandra and Pradeep Mujumdar, "Urban Floods: Case Study of Bangalore", Disaster & Development, vol. 1, no. 2, November 2006, pp. 1-97.

[5] Smeed. R. J, "Variations in the pattern of accident rates in different countries and their causes.," Traffic Engineering & Control, vol. 10, no. 6, 1968, pp.364- 371.

[6] R. V. Ponazuril, "Modeling Road traffic fatalities in India: Smeed's law, time invariance and regional specificity," International Association of Traffic and Safety Sciences, vol. 36, no. 1, Jul. 2012, pp. 75–82.

[7] Andressean. D, "Development of accident hazard index.," Journal on Transportation Engineering, vol. 114, no. 6, 1988, pp.247-249.

[8] Q. N. Hashmi and T. I. Qayyum, "Accident prediction model for passenger cars," Academic Research International, vol. 2, no. 1, 2012, pp. 164–173.

[9] Fajaruddin Mustakim and Moto Hiro Fujita, "Development of accident predictive model for rural roadway", World Academy of Science, Engineering and Technology, vol 58, 2011, pp. 126- 131.

[10] P. Greibe, "Accident prediction models for urban roads", Accident; analysis and prevention, vol. 35, no. 2, Mar. 2003, pp. 273–85.

[11] M. W. Witczak, K. Kaloush and T. Pellinen, "Simple performance test for super pave mix design", NCHRP, report no.465, 2002.