# Crash Prediction Modeling of Two Lane Undivided Highways Using Artificial Neural Network

Nivea John, Archana S

**Abstract**— Road accidents are increasing day by day, causing losses to human life and economy. This scenario focuses the necessity of understanding why these accidents occur and how to prevent future ones. Crash prediction models have been developed which are able to correlate accident frequency with infrastructure characteristics and to support the planning and design of countermeasures to enhance road safety. In this paper crash prediction model is developed using Artificial Neural network (ANN). Neural Networks have the ability to describe high complex systems and helps engineers to learn about, all the factors which promotes road crashes. ANN were developed in Python with Keras library. The factors considered in the model are crash data, speed, volume, landuse type, pavement width and condition, shoulder width, number of horizontal curves, vertical curves, intersections and bus stops. Results shows that estimated traffic accidents, based on the input data are close enough to actual road accidents hence it is reliable to predict future accidents in two lane undivided state highways. The performance of ANN is found to be better than other statistical methods.

**Index Terms**— Accident prediction model, Artificial neural network (ANN), Python, Keras

## 1 INTRODUCTION

IN Indian cities, the traffic is highly heterogeneous in nature, promotes a large accident rates. Analysis of traffic data Indicated that drivers fault is responsible for majority of the accidents. The economic and social costs of traffic accidents are tremendous. Property damage, lost productivity, medical expenses, and inflated motor insurance rates imposed an estimated loss to the Indian economy.

Crash injury severity has always been a major concern in highway safety research. During road safety analysis of a road, a major target is to locate hazardous segments which are dangerous, and then to identify the factors influencing its safety level. Evaluation of road safety measures appears to be the weakest component of road safety management systems. To improve road infrastructure in the concern of safety management, road authorities, road designers and road safety practitioners need prediction tools, commonly known as Accident Prediction Models (APMs). Accident prediction model will helps to analyse the potential safety issues and helps to improve safety measures.

Two lane undivided state highways have a great contribution to the number of accidents due to several factors. This paper focuses on the identification of factors which promotes accidents and its relationship with the number of accidents. Development of accident prediction model will helps to know how these accidents occur and how to prevent future ones.

- *Nivea John, PG Scholar, Department of Civil Engineering, Jyothi Engineering College, Cheruthuruthy Email: niveajohn6@gmail.com*
- *Archana S., Assistant Professor, Department of Civil Engineering, Jyothi Engineering College, Cheruthuruthy Email: archanas@jecc.ac.in*

## 2. LITERATURE REVIEW

Francisca Nonyelum Ogwueleka et.al. [1], focused on the design of an Artificial Neural Network (ANN) model for the analysis and prediction of accident rates in a developing country. The development of an Artificial Neural Network (ANN) model is done for the prediction of accidents rate using Nigeria as a case study. Number of vehicles, accidents, and population are used as model parameters. By using self- organizing map, road accident was categorized. The Multi-Layer Perceptron Neural Network (MLPNN) for the development of ANN. Results shows the ability of ANN over conventional programming in the study and this makes neural networks to relate input with output, by allowing large number of variables and error are tolerant.

Muhammed Yasin Codur et.al. [2], studies an accident prediction model of Erzurum's Highways in Turkey using artificial neural network (ANN) approaches. The geometric characteristics of the highway such as AADT, the degree of horizontal and vertical curvatures in each section, lane, median, and shoulder widths were considered. There are 8 input variables containing 31 neurons which are the input variables representing the potential risk factors for accidents. ANN were developed in MATLAB and using coefficients of determination (R2) mean square error (MSE), and the root mean square error (RMSE) performance of ANN model is evaluated.

Ebrahim. S et.al [4], studies the application of Artificial Neural Network for road accident prediction in Khulna metropolitan city. In the design, vehicle type, accidents type, junction type and collision type were selected as model parameters. In the ANN model development, the sigmoid activation function was used with Levenberg-Marquardt algorithm. The results shows that the estimated traffic accidents, based on sufficient data, are close enough to actual

traffic accidents and thus are reliable to predict future traffic accidents in Khulna Metropolitan city.

Khair S. Jadaan et.al [5] studies the Prediction of Road Traffic Accidents in Jordan using Artificial Neural Network (ANN). , data collection phase includes the parameters number of registered vehicles, population, total length of paved roads and the gross domestic product .Training, validation and testing of the network was performed using MATLAB. The statistical analysis was done using the SPSS statistical software. Artificial Neural Networks (ANN) proved to be successful in solving engineering problem. Rezaie Moghaddam et.al [3], focused on the Prediction of accident severity using artificial neural networks in Israel. Input variables for model creation includes traffic volume, speed, geometrical characteristics and crash data. Training, validation and testing of the network was performed using MATLAB. The statistical analysis was performed using the SPSS statistical software.

## 3 METHODOLOGY

### 3.1 Study Area

Thrissur is one of the districts in Kerala which is vulnerable to road accidents. The selected road includes state highways, SH 69, SH 50 and SH 74. These highways have a major contribution to the number of accidents that happening in the Thrissur district. SH 69 is the Thrissur –Kuttipuram state highway, SH 50 is Chavakkad- Wadakkanchery state highway and SH 74 is Vazhakode- Alathur state highway. Three road stretches of two lane undivided state highways are identified based on their hazardous nature. The road stretch selected from the state highway 69, of about 10 km starting from Kunnamkulam bus stand to Thrissur boundary with Malappuram. From state highway 50, road stretch of about 15 km length starting from Ottupara to vellarakad is taken as the second one. Third road stretch, Vazhakode to Pazhyannur of about 20 km is taken from the SH 74. All these three road stretches consists of vulnerable horizontal and vertical curves. Number of intersections in these highways are high and it may be an important factor for the increase in the number of accidents. These stretches have similar characteristics and main thing is that they are two lane undivided highways.

### 3.2 Data Collection

Data from each road stretch is collected by dividing the roads into 500 metre segments. All three road stretches were divided into 500 metre segments, thus forming a data set of 84 samples. The parameters used for the model development include the type of landuse, pavement width and condition, width of right and left shoulder, number of vertical curves, horizontal curves, intersections and bus stops were taken along with speed and volume. The data were collected from each 500 metre segment of the road along with that number of accidents which occurred in each segments. Past five year accident data (2013-2018) were collected from the District Crime Records Bureau (DCRB), Thrissur district. Road inventory data were collected using odometer and pavement conditions and type of landuse rated manually. Speed data were collected as free flow speed in off peak hours and classified volume data were collected from peak hours.

### 3.3 Preliminary Analysis

Preliminary analysis of the data collected reveals that these study stretches shows the vulnerability of accidents. Most of the intersections and some midblock identified as hazardous zones. The presence of sharp vertical and horizontal curves promotes accidents in midblock. The results shows that volume and speed have a significance importance in the road accidents. Increase in volume and speed have the chances in the increase in number of accidents. The variation in number of accidents along with volume and speed is given below in Fig 1. and Fig 2.

fig1. Relationship between road accidents and volume

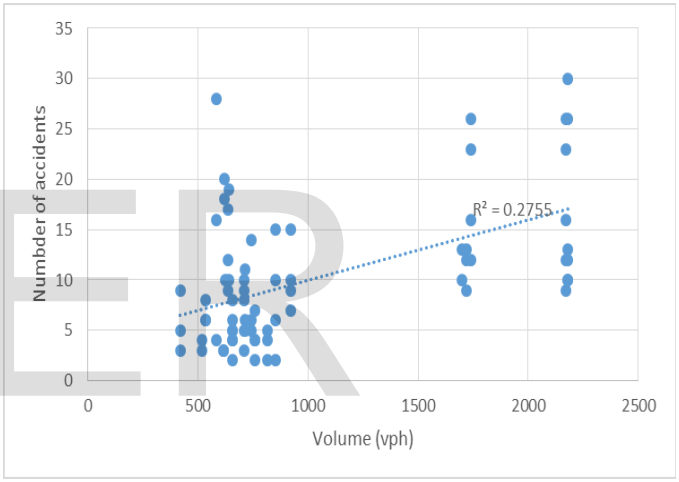
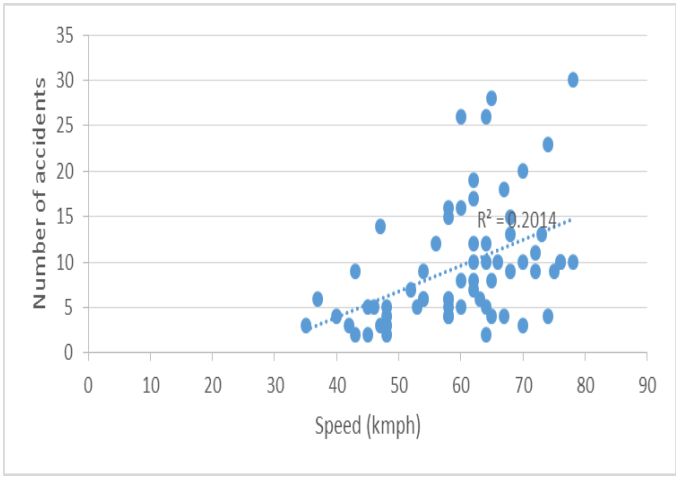



fig2. Relationship between road accidents and speed




The major observation found out in the analysis that all these state highways doesn't consists of road markings, proper signboards and essential precautions that has to be provided in an intersection. The speed limit of a state highway is 60 kmph, but most of the vehicles are running above the speed

limit causes increase of accident number especially at intersections. The lack of road humps in minor road that joining the major roads promotes a large number of road accidents. High rate of speed over the vertical and horizontal curves becomes also a factor for road accidents.

Some remedial measures for the hazardous to decrease the accident severity as per the observations are given below,

1. Providing proper road markings, signboards.
2. Midblock sections with horizontal and vertical curves can be treated with transverse bar markings as a traffic calming measure.
3. At large intersections, speed tables can be provided on the major roads and road humps should be provided in the minor roads.
4. At junctions speed limit should be reduced to 30-40kmphr and provision of interlocked or colored pavements can be provided for catching attractions to the drivers and thus reducing the speed.
5. At small intersections, speed breakers should be provided for major roads and speed bumps to minor ones with proper markings and signboards.
6. Narrowing of roads or provision of chicanes is another traffic calming measure which can be provided for high accident zones in mid blocks.
7. Position of bus stops is another cause of accidents, thus provision of bus bay may be remedy for this.
8. Closure of minor roads which provide access to the accident zones is a good option to reduce accident severity.

## 4 CRASH PREDICTION MODEL

Accident prediction model is developed by using python software with the help of Keras library. Keras is an open-source neural-network library written in Python and designed to enable fast experimentation with deep neural networks. The python codes are written in Colaboratory which is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud. It is a type of regression problem where accident causing factors taken as independent variables and number of accidents taken as dependent variable. The total sample set of 84 was divided into two parts as training data and testing data in random order. Out of these 84 samples one by third data is used for testing purpose i.e., 45 samples for training and 28 samples for testing purpose. Within these 45 training samples one by fifth data is used for the validation process. The model is created by the simultaneous training and validation of the random data samples.

Multi-layer perceptron neural network (MLPNN) is used to develop model by using Backpropagation algorithm. The designed model consists of total 20 input nodes. In these 20 nodes, 9 nodes consists of input parameters as pavement width, right shoulder with, left shoulder width, number of vertical curves, horizontal curves, intersections, bus stops, speed and volume. The input parameter 'land use' is categorical value with eight different types and 'pavement condition' parameter consists of three categorical values. In python categorical values should be converted into numerical value using label encoder. Thus the total count of input parameter becomes 20. The model is designed of one hidden layer with 5 nodes and one output node. The activation function used in input hidden layer is Relu and in output layer, linear transfer function is used. ANN model created is depicted in Fig3 and Fig4.

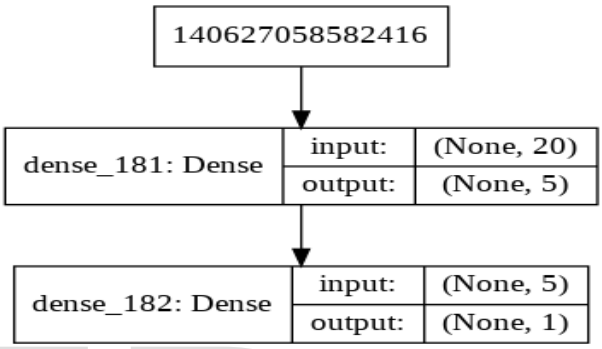
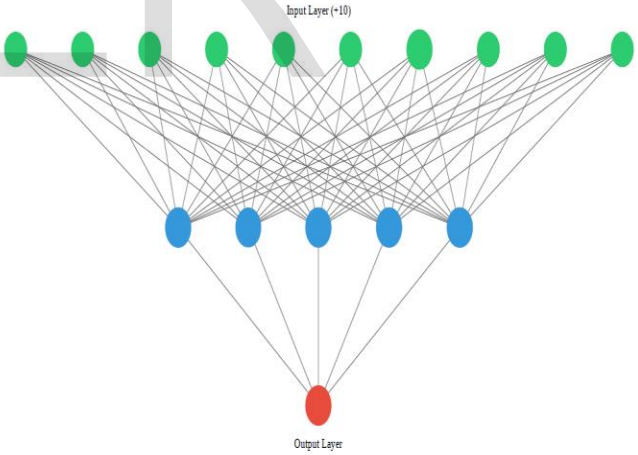



fig3. Model plot

fig4. Structure of ANN model




The model development includes the following steps,

1. Setting of initial weights to input synapses and evaluating the output based on the initial weights.
2. Calculate the error by using mean absolute error
3. In Backpropagation algorithm, errors are minimised by adjusting weights using stochastic gradient descent 'Adam 'as an optimizer.
4. Weights are continued to be modified until the error becomes very less (eventually to zero).
5. Training were done in 1600 epochs.
6. No adjustments of weights are occurred in validation phase and cross validation is done by k-fold validation method. Validation is integrated with training process to improve the performance

of the model.

7. In testing process, predicted values are compared with actual values in test data set that is not used in training and validation process.

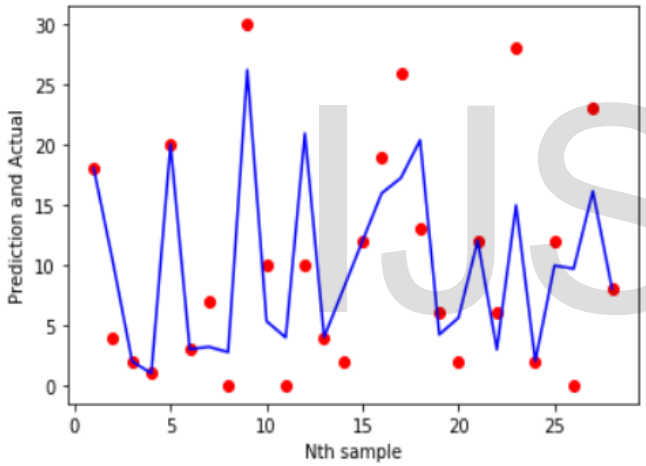
The variables in the test data set is provided to the model to get the predicted number of accidents. The actual number of accidents and predicted values are evaluated from the model and the corresponding error and coefficient of determination $r^2$ is calculated. This will shows the performance of the model.

## 5 RESULTS AND DISCUSSION

The designed model consists of 20 input nodes, one hidden layer consist of 5 nodes and one output layer. The output layer provides the prediction of road accidents with respect to the factors given from the test data set. The graph showing predicted and actual values are depicted below.

Fig5. Predicted and actual number of accidents

The red dot shows the actual values and blue line represents the predicted number of accidents. 28 samples were




taken for the test purpose. Out of these, 18 samples have closeness between predicted and actual values, which shows a good model. The mean absolute error of train and test data are 3.662 and 3.626, which is approximately equal, which indicates the model is in good fit. The coefficient of determination $r^2$ have the value 0.66.

The results which got from the ANN model is also compared with the result obtained from the Statistical Package for Social Science (SPSS), in order to check the reliability and performance of ANN. The mean absolute error which got from SPSS is 6.425 and $r^2$ value is 0.447.Thus ANN have high reliability than other statistical methods.

## 6 CONCLUSION

The study investigated the factors which causes the accidents in two lane undivided highways and remedial measures for most hazardous zones were recommended. The low value of error and $r^2$ value shows the superiority of the model. Comparison with the SPSS software also reveals the high predictive performance of the ANN model.

Thus artificial neural networks can be used as a forecasting techniques to predict the number of accidents in the two lane undivided highways with similar characteristics. This, also act as a platform for linking ANN to other planning models to get the best outcomes in transportation planning field also.

## ACKNOWLEDGMENT

I owe a great debt of gratitude to teaching staffs and Department of Civil Engineering, Jyothi Engineering College, for great support throughout this project. I am deeply indebted to Githin V George, Nighisha John and Anirudh Rajan V for their valuable suggestions and encouragement.

## REFERENCES

[1] Ogwueleka. F.N, S. Misra, Ogwueleka, T. C and Fernandez Sanz, L" An Artificial Neural Network Model for Road Accident Prediction: A Case Study of a Developing Country", *Acta Polytechnica Hungarica*, Vol.11, pp.177-197, 2014.

[2] M.Yasin Çodur, and A. Tortum, "An artificial neural network model for highway accident prediction: A case study of Erzurum, Turkey" *Promet- Traffic&Transportation*, Vol.3, pp. 217-225,2015

[3] Moghaddam, F. R, S.Afandizadeh and M. Ziyadi, "Prediction of accident severity using artificial neural networks", *International Journal of Civil Engineering*, Vol. 9, pp.41-45,2011

[4] S.Ebrahim and Q.S Hossain, "An Artificial Neural Network model for road accident prediction: A case study of Khula Metropolitian City", *Proc. Fourth Ann. Civil Engineering for Susutainable Development* , pp. 9-11, 2018. (Conference proceedings)

[5] Jadaan, K. S, M. Al-Fayyad and Gammoh, H. F," Prediction of road traffic accidents in jordan using artificial neural network (ANN)",*Journal of Traffic and Logistics Engineering*, Vol.2.pp.23-26,2014

[6] A.,Benedetto ,C.Benedetto and De Blasiis, M. R ,"A new effective approach to accidents prediction to improve roads'design and rehabilitation",*Advances in Transportation Studies*, Vol.2,pp.4-6,2004

[7] Kunt, M. M, Aghayan, I and N.Noii,"Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods", *Transport*, Vol.4, pp.353-366, 2011.

[8] G.Wang and J .Kim,"A Large Scale Neural Network Model for Crash Prediction in Urban Road Networks", *Australasian Transport Research Forum Proceedings*, Vol.2, pp. 27 – 29, 2017.

[9] D.Akin and B. Akba," A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics", *Science Research. Essays*, Vol.5, pp. 2837-2847, 2010.

[10] Chong, M.M, A.Abraham and M. Paprzycki,"Traffic accident analysis using decision trees and neural networks", *Applied Sciences*, Vol.8, pp.664, 2004.