International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

# A Study on the Use of Artificial Neural Networks in the Study and Modeling of Road Accidents

Kanamala Priyanka, B Khadar Mohammed

[1,] *M.Tech (Transportation Engineering) Student, MVR College of Engineering and Technology, Paritala, Andhra Pradesh, India.*

[2]*Assistant Professor of Civil Engineering, MVR College of Engineering and Technology, Paritala, Andhra Pradesh, India.*

Submitted: 15-09-2021 Revised: 25-09-2021 Accepted: 28-09-2021

**ABSTRACT:** The present research work is aimed to identify the significant factors in relation to crash and develop an Artificial Neural Network (ANN) model to estimate crash severity on urban highways. The ANN has profound applications in various information technologies. In traffic and transportation engineering, ANN is used in modeling travel behavior, flow estimation and management. It can be employed to relate driver injury with crash leading factors including driver, vehicle characteristics, roadway characteristics and environmental characteristics. Road Traffic Accidents (RTA) have been identified as one of the major causes of deaths across Telangana. The rate of RTAs is being reduced in developed countries. With the increasing vehicles, driver inability to understand the traffic behavior and roadway conditions, the RTAs are not decreased in developing countries like India. The number of vehicles, number of accidents and the population are considered as explaining parameters. The feed-forward-back propagation is used in the sigmoid and linear functions activation. Results showed that the ANN model showed better performance in comparison to the statistical methods in use.

**Key words**: Vehicles, Accident, roadway, Artificial Neural Networks, Linear Function, Back Propagation.

## I. INTRODUCTION

Artificial Neural Networks (ANN) replicates the behavior of the biological nerves system which acquires the information, processes and derives the results after the pattern recognition is properly trained. The Neural network consists of layers of parallel processing elements called neurons. Multiple layers may exist between the input and output. The neurons of a hidden layer are interconnected with adjacent neurons by weighted factors which care adjusted in the training process of model. The network organization varies with training methods for a specific application. An example of ANN model can be considered as a three layer network with four neurons in the input layer, hidden layer having four neurons and output layer with two neurons and the neurons are interconnected with weighting factor (wii) in between the layers of neurons.

The 'training' of ANN is a process by which ANN processes a set of test data by changing the weights according to a predetermined algorithm to improve its performance and back-propagation is well popular in ANN model training. The back-propagation supervises with back feeding the output error through the network, varying the weights to reduces the error between input and target output. The weights are corrected as follows.

- The structural member, surface layer thickness, defect indicator construction, traffic and cracking retardation occurring due to maintenance with time are the significant independent variables of cracking model.
- Age of pavement, traffic loading, raveling retarding factors of road maintenance and construction defects are significant variables defining raveling model.
- Traffic loading, bituminous surface thickness, construction defects and precipitation are significant variables in potholes model.
- Structural number, traffic loading and bituminous surface thickness, construction defects are significant variables explaining rutting model.
- Age of pavement, structural number, traffic

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

loading, surface layer thickness, and environmental conditions- temperature and moisture variations are significant variables explaining roughness model.

- Traffic loading and initial texture depth are significant independent variables of the texture depth model.
- Traffic loading and average annual texture depth are significant independent variables of the skid resistance model.

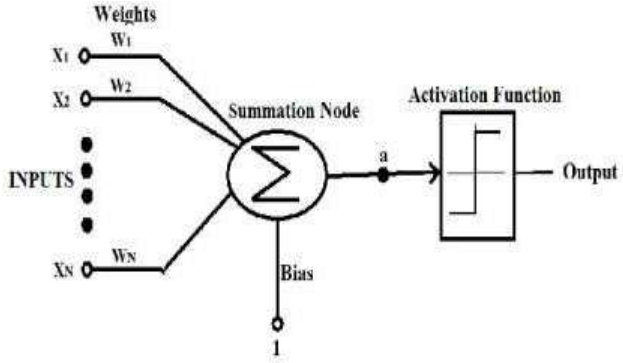


**Figure 1: Simple perception Neural Network**

## II. LITERATUREREVIEW

**(1) Ozgan, E. and Demirci. R. (2014)** the transportation infrastructure system plays a major role in developing the countries 'economic, social, cultural and employment opportunities. A countries development depends on the road network growth. The road networks projects a significant proportion of nation's wealth. With over 4.4 million kms of road network, India stands as a second largest country in the world in terms of road network. Indian road network categorized in to five types consisting of expressways, national highways, state highways, other district roads, and village roads. The village roads are observed to have less traffic, with less than 150 CVPD in comparison with other category roads in India. The village roads in India are interconnected with other villages and other type of roads and has about 2,750,000 km of network. These roads connect the agriculture and production sectors. Indian rural areas have an incurable disease of poverty which can be curbed by the village roads.

**(2) Abdelwahab,H. T.andAbdel-Aty,M.A. (2001)** ANN is potentially used in modeling the travel behavior, traffic flow and its management.ANN is employed to model the relationship between the driver injury severity and factors causing the crash which includes the driver characteristics, the vehicle characteristics, roadway characteristics and the environment characteristics. The use of ANN reveals the relationship between vehicle characteristics, roadway characteristics and environment characteristics and driver injury severity. ANN can solve the complex interrelationship among traffic system parameters in forecasting traffic. ANN provides realistic and faster means in developing models with sufficient data. The study explains the application of ANN in the modeling of the number of persons. The ANN models also facilitate the comparison of the states' road safety performance with the fatalities from motor vehicles.

**(3) Manicom R., J. F., Page, Y., (2006)** Most of the models for pavement performance are meant for the high-volume roads in India. The maintenance of low volume roads became a challenging task with the unavailability of models for scientific performance of low volume roads. In the absence of scientific performance models, estimation of quantitative condition of pavement in terms of roughness value in m/km can be used as maintenance factor. The roughness value indicates the cumulative surface unevenness in m/km of the road. So, it could be extremely useful in specifying the roughness value of pavements so as to make strategy decisions on well-founded maintenance. Also in India, the agencies lack skilled manpower in maintaining low volume village roads, and it is very difficult to them intake appropriate precautions while defining roughness value using bump integrator.

**(4) Wei, C.H.andLee,Y.(2007)**ANNs are being widely used in solving complex problems, resource intensive problems. They are an alternative to traditional techniques. Ann is being used as a tool in pavement structural analysis in making rapid and precise prediction of deflections and responses of a flexible pavement when subjected to the typical highway loadings. The processing speed is observed to be increased by 42 times by using back propagated ANN in comparison to ELP analysis for calculating the module of pavement layers. ANN applications are found in pavement deterioration simulating, in

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

predicting pavement performance, predicting flexible pavement cracking, concrete pavement condition rating. ANN is also used to `current pavement condition and predict future deterioration of pavement which assist engineers in making maintenance decisions and rehabilitation activities. This enables in making choice of best available strategies in resource allocation.

**(5) Pande, A., and Abdel-Aty, M. (2006)** For any country, a productive road transportation framework is of fundamental significance. The Road transportation in India, has predominant position in the nation's general transportation arrangement as it has favorable circumstances in terms of accessibility, operation adaptability and reliability. The importance of the road network is majorly recognized in preserving an adequate condition. However, creating and keeping up a valuable road network is not a simple assignment. This requires a careful arrangement, vast assets, advanced construction technique and other related perspectives. Pavement Depreciation models simulate the pavement depreciation procedure conditions and forecast pavement conditions to forecast over time, plays a vital part in Pavement Management System (PMS). The study aims to develop pavement performance models using neural network algorithm to urban road net work in India.

**(6)Sharda, R., and Bessonov, M. (2006**) The PSI and PCI models for pavement deterioration was developed for low volume roads in India. An ANN model was developed with explaining variables slope variance, patches, rut depth and longitudinal cracking. It was observed that the SVM model developed using AASTHO panel data gave better estimates of PSR and PSI. Psi model developed for the asphalt pavement section was located in the Noida, India. Pavement age was considered as a function to develop the PSI. An attempt was made to calibrate the AASHTO equation for PSI and its suitability of thee quotation for Indian pavement conditions for the selected urban roads

**(7) Onakoya A. O. (2007)** An ANN model was developed to examine and predict the accident rate in a case study. The model parameters were number of vehicles, population and accidents. A feed-forward back propagation approach was adopted using functions of sigmoid and linear. The results showed that the ANN model performed better in comparison to statistical methods.

## III. METHODOLOGY

ANN has solved wide variety problems in communications, traffic and transportation engineering. ANN is used to predict travel behavior, traffic management and traffic flow. It is also employed in relating the driver injury severity and the crash with influencing parameters vehicle characteristics, driver characteristics, roadway characteristics and environment characteristics. The use of ANN used to relate the factors vehicle characteristics, roadway characteristics and environment characteristics and driver injury severity. ANN can be used to solve problems in Traffic forecasting which involve complex inter relationships between traffic system variables. ANN offers realistic and continue always in building models, and provides enough data. The present studies the use of ANN in the relating the number of fatalities injured in vehicle accidents in data sets of the states of the US. The ANN models enables and assists in comparing the safety of the state roads with the number of fatalities of motor vehicle. The ANNs use has been effective in a wide variety of areas including transportation engineering. Researchers identified the factors concerning driver injury severity in relation to driver characteristics, vehicle characteristics, roadway characteristics, and environment characteristics by adopting two profound ANN algorithms, the multilayer perception algorithm and the fuzzy adaptive resonance theory ANN. Recently ANN is used in chronological forecasting of event duration from the point of event notification to the event road clearance. The lane-change occurrence in relation to freeway crashes has been estimated by using traffic inspection data which is composed using a pair of dual loop detector sand this study establishment will provide clear understanding on drivers and passenger's likelihood of death and injury in RTAs.

The intent of this chapter is to determine the procedure which is used to adopt in this present study. A flow chart involving proposed methodology is shown below and each step is briefly explained.

**Collection of accident data:**

Accident data of selected stretch will be collected from secondary sources. Secondary datainclude sinformationlikenumberofaccidentsoccurredoverthepasteightyearsforNH7andfive years accident data for NH9 stretch. The data such as accused vehicle type, victim vehicle type, nature of accident, accident type classification etc,

**Collection of road geometric, intersection, traffic and socio economic data**

On the selected stretch of NH7 and NH9 geometric data such as number of culverts, number

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

of vertical curves and degree of curve is collected from secondary sources. Detailed project report of the selected stretch will be collected from the National Highway Authority of India (NHAI).Geometric parameters extracted from DPRs

**Interpretation of significant causative factors**

After preliminary analysis of accident data is performed, Student test (one sample) was performed to find the significant variables for developing the model.

**Design of the ANN**

Network elements are utilized as the inputs to the ANN. The temperatures of the cores, switches and links are the output of Base ANN. The threshold producing a single bit output is compared to the Base ANN output. The output signifies the element which crossed the threshold or not. The ANN trained is a concatenation of three subdivided ANN streams. Core streams, Link stream and Switch stream are the subdivided ANN streams. The inputs to all the streams remain unchanged, i.e. the utilization of all the network elements and time to be predicted

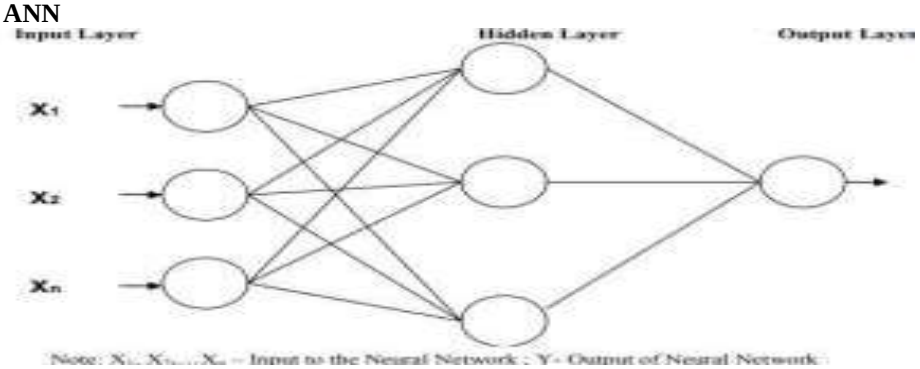


**Figure 2: Three layered neuron network**

## IV. RESULTS AND DISCUSSIONS:

**Table 4.1.** Data coding of input and output Variables

| FACTORS | CODING |
|---|---|
| Time Of Accident | DAY: 1; NIGHT:0 |
| Age Of Driver | Age of thedriver:0-35y:1; >=35:0 |
| Nature Of Accident | overturn:1;headon collision:0.5;rearend collision:0.75 |
| Person Driving Vehicle | Owner of private vehicle: 1; paidvehicle: 0 |
| Type Of Vehicle | 2w:1;Auto,3w:0.5;car/lorry/dcm/zeep:0 |
| Classification Of Accident | fatel:0;serious:1;injure:-0.5 |

**TRAININGAND TESTING ACCURACY**

For the present study a total of 462 accident sample data from 13 police stations is used. Out of 462 data points 370 data points were used for training phase and 92 data points were used for testing phase. The accuracy of ANN is done with the help of program written in MATLAB R2007a. The accuracy of the model is tested by drawing graph between performance of and number of epochs. One cannot achieve zero error so the performance goal was kept as 0.1 in the program i.e. if the performance level is reached nearer to 0.1 then one can say that the model is accurate. The training phase was performed with different combinations of hidden functions and training algorithms. The performance level or the mean squared error of training functions is checked.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

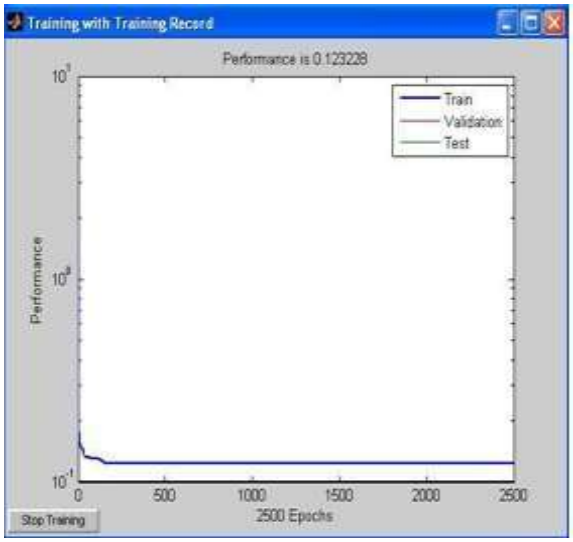


**Figure3: Variation of epochs with means quared error for best trained neural network.**

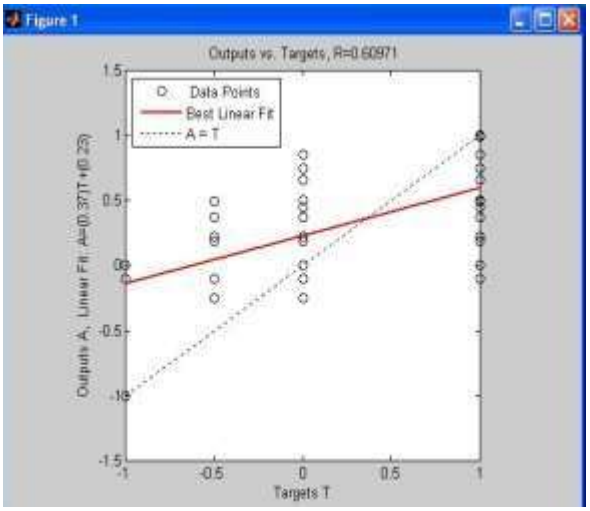


**Figure 4:** Linear fit between input and output variables

After number of trails it was found that the Levenbeg-Marquardt (trainlm) is giving least means quare error and reasonable coefficient of correlation

Levenberg Marquardt algorithm was selected for both training and testing with number of hidden nodes being 6,4and the means quared error wegotas0.123228 and correlation coefficient as 0.60.

**MODELDEVELOPMENT TO PREDICT ACCIDENT RATE BASED ON ROAD GEOMETRICS**

**Variables for NH9from40-160 km**

An attempt has been made to develop relationship between number accidents per year per km and different road geometric conditions using Microsoft Excel trend line approach for different stretches in NH9and NH7.

**Accident/yr/km vs Degree of curve**

A scatter plot was drawn which relates to number of accidents per year per km with radius of curvature for NH9 stretch from 40 to 160 km. It is clearly shown that as the degree of curve varying parabolic with Accident/yr/km

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

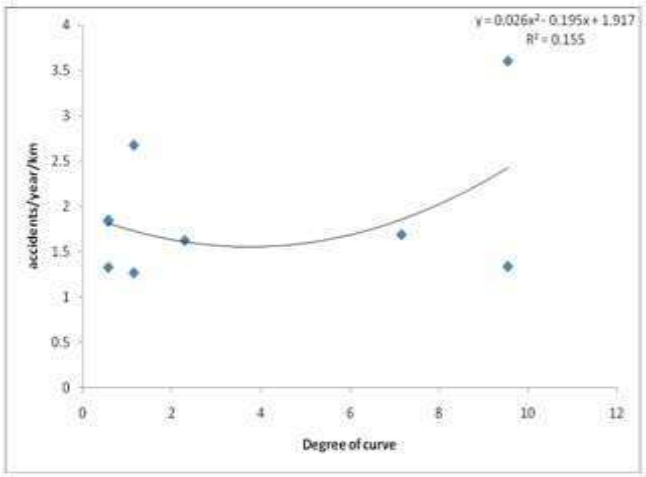


**Figure 5: Accidents/year/km as a function of degree of curve**

**Accident/yr/km vs number of vertical curves**

A scatter plot was drawn which relates to number of accidents per year per km with number of vertical curves is varying parabolic form with respect to accident/yr/km

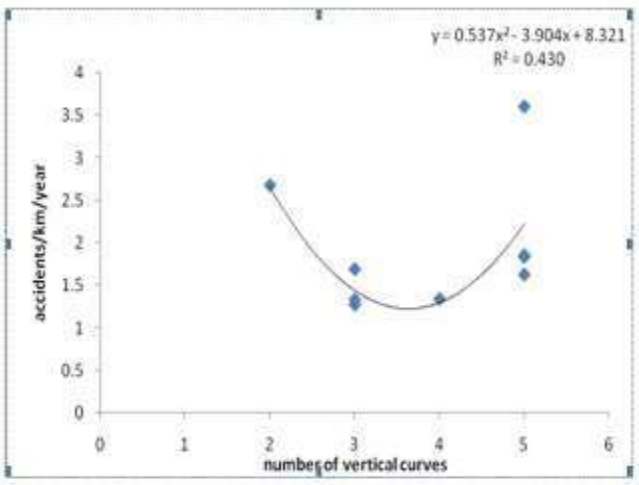


**Figure6: Accidents/year/km as a function of number of vertical curve**

**Accident/yr/km vs number of the culverts (m)**

A scatter plot was drawn which relates to number of accidents per year per km with number of the culverts is varying parabolic form with respect to accident/yr/km

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

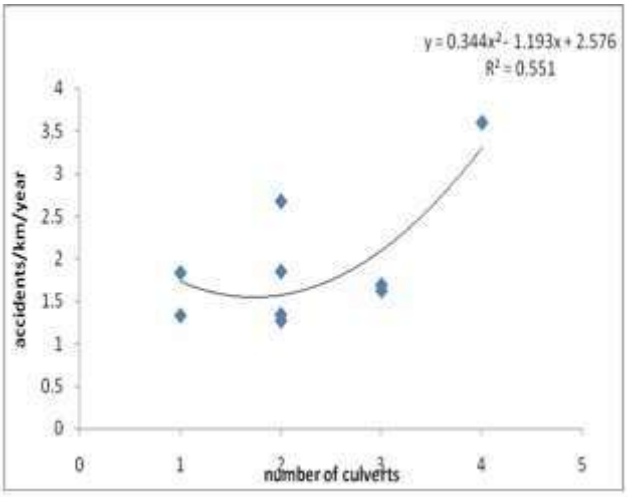


**Figure 7: Accidents/year/km as a function of number of culverts**

**Variables for NH7 from 308-454km**

**Accident/yr/km vs degree of curve**

A scatter plot was drawn which relates to number of accidents per year pee km with degree of curve for NH7stretch .It is clearly shown that as the degree of curvature varying in parabolic form w.r.t Accident/yr/km.

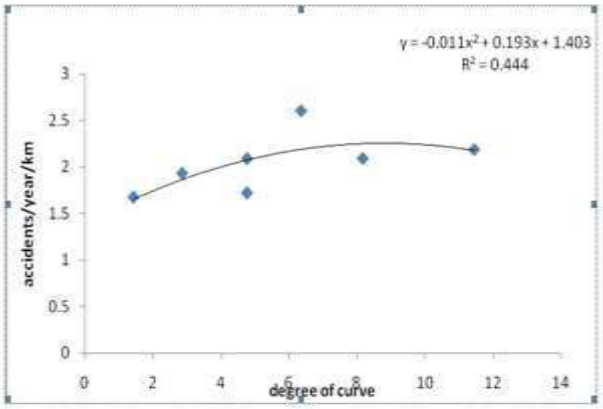


**Figure 8: Accident/yr/km as a function of degree of curve**

**Accidents/year/km vs number of vertical curves**

The relation between numbers of vertical curves w.r.t. accidents/year/km is shown in the figure 9, which is varying linearly.

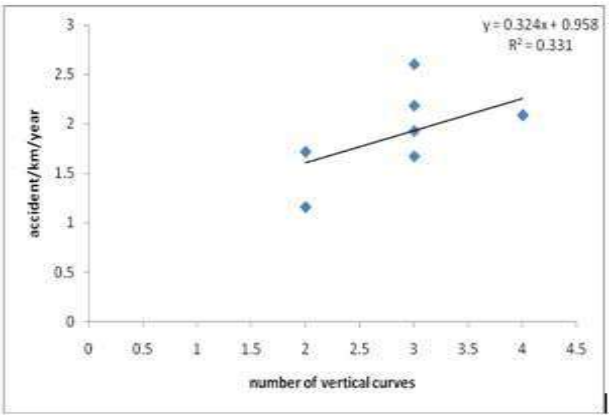


**Figure 9: accidents/km/year as a function of number of vertical curves**

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 1556-1564 www.ijaem.net ISSN: 2395-5252

**Accidents/year/km vs. number of culverts**

The number of culverts w.r.t. accidents/yr/km gives linear relationship as shown in figure10

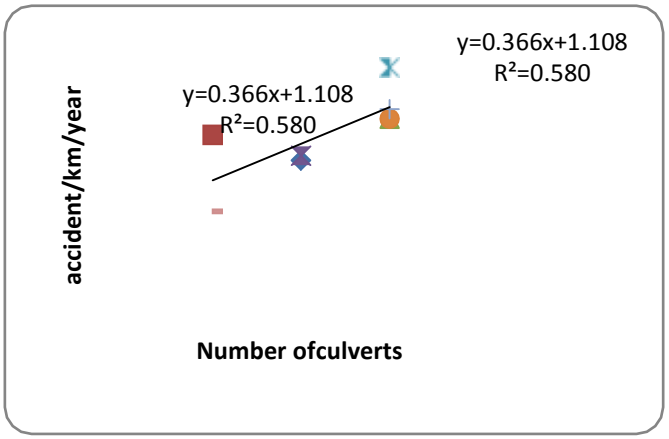


**Figure 10: Accidents/km/year vs number of culvert**

In this chapter test for one sample was used to interpret the independent variables for the development of the model. After observing the significant variables accident modeling was done by using artificial neural networks tool to find how much the input variables are influencing the output variable by selecting Levenberg-Marquardt algorithm. The mean squared error and correlation coefficient also estimated from the model. The mean squared error was 0.1232 and the coefficient of correlation was 0.60. it means that the input variables are having a strength of 60 percent to explain the output variable with least error of0.1232.

The influence of geometric parameters accident rate is also done by using regression analysis for NH9 and NH7.

## CONCLUSIONS:

The following are the conclusions from this study:

- Maximum numbers of accidents are reported in the month of May.
- Most accused type of vehicles causing accidents are the heavy vehicles like trucks and the victims are pedestrians and two wheeler riders.
- Head on and Rear end collisions are observed in more number of accidents.
- The accidents occurring during day and night are almost equal in proportions.
- Highest numbers of accidents have occurred near the villages where the National Highway is passing.
- From the model it is found that time of accident, type of vehicle, person driving vehicle, nature of accident, age of the driver are showing 60 percent influence on the accident severity and with a mean squared errorof0.123228.

The present study is mainly intended to determine the influence of variables like age of driver, person driving vehicle, type of vehicle, nature of accident on driver accident severity. The dependence of these input variables on the output variable (accident severity) is checked through soft tool called artificial neural networks (ANN). Regression graphs were drawn for various geometric parameters vs accidents/year/km.

## REFERENCES

[1] Ozgan, E. and Demirci. R. (2008) Neural Networks-based Modeling of Traffic Accidents in Inter urban Rural Highways, Duzce Sampling. J.Applied Sci., 8,pp. 146-151

[2] Abdelwahab, H. T. and Abdel-Aty, M. A. (2001) Development of ANN Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections, Transport. Res. Rec.,1746,pp.6-13

[3] Manicom R., J. F., Page, Y., (2006) A Bayesian Neural Network Approach to Estimating the Energy Equivalent Speed. Accident Analysis and Prevention,38,pp. 248-259

[4] Wei, C. H. and Lee, Y. (2007) Sequential Forecast of Incident Duration Using ANN Models. Accident Analysis and Prevention. 39(5),pp.944-954

[5] Pande, A., and Abdel-Aty, M. (2006) Assessment of Freeway Traffic Parameters Leading To Lane-Change Related

Collisions. Accident Analysis and Prevention.38,pp. 936-948

[6] Sharda, R., and Bessonov, M. (2006) Identifying Significant Predictors of Injury Severity in Traffic Accidents Using A Series of Artificial Neural Networks Accident Analysis &Prevention,38 (3),pp.434–444

[7] OnakoyaA.O.(2007)Visual Functions of Commercial Drivers in Relation to Road Accidents in Nigeria. IndianJOccupEnvironMed. 11,pp.71-75

[8] Bergan,A.,andZhang,G.(1996)RoadSafetyModel:SomeFundamentalIdeas,inTransportationResearch Record. 1563, pp. 48-57

[9] Ozgan, E. and R. Demirci. (2008) Neural Networks-based Modeling of Traffic Accidents in Inter urban Rural Highways, duzce sampling. Applied Sci., 8:146-151