ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 7, Issue 2, pp: (412-420), Month: April - June 2019, Available at: www.researchpublish.com

ACCIDENT MODELLING ON NH-7 USING ARTIFICIAL NEURAL NETWORK

Nirpinder Jain[1], Dr. Sanjiv Kumar Aggarwal[2]

[1]Research Scholar, IK Gujral Punjab Technical University, Kapurthala, Punjab, India

[2]Professor, Department of Civil Engineering, GZSCCET, MRSPTU, Bathinda, Punjab, India

Abstract: **There has been a tremendous growth of vehicles in developing countries in the last decade. The growth of vehicles has led to increase in traffic volume and increased number of road trips. Further, this scenario has led to increase in road traffic accidents. The deaths and injuries due to road accidents are posing a major concern and challenge, globally, due to its socio-economic costs. The causative factors of accidents are the complex interaction between the various road-user and vehicle-environment related factors. However, prediction of future accidents and variables responsible for occurrence of accidents is of utmost importance to appreciate the quantum of problem and speed up the decision making. Moreover, the Indian two lane highways carry mixed traffic which has a lot of variability. This further affects the explanatory variables responsible for accidents. Accident prediction modelling enables traffic engineers to analyze why an accident happens and correlate mathematically the causal factors to accident occurrence. In this study, an attempt is made to explore the potential of Neural Networks for highly complex and nonlinear relationship of variables influencing accidents on the study stretch of NH-7. The results suggested that ANN approach is a very useful computational tool for analyzing and predicting traffic accidents.**

Keywords: **Traffic, Artificial Neural Network, Accident Modelling, Minitab, Two Lane Roads.**

1. INTRODUCTION

In last decade, remarkable effort and money have been invested to improve the highway safety. The current challenge to transportation engineers is to plan and reduce the loss of life and property in the transportation system. Accidents lead to a huge financial burden on the society especially in low income countries. In Indian scenario the deaths and injuries due to road accidents have been increasing at an alarming rate during last 10 years. Fatal accidents, on Indian multi lane highways have increased consistently since 2006 (93,917) to 2016(1, 36,071) [1]. Consequently, number of persons killed per 100 accidents, has gone up from 22.9 in 2006 to 31.4 in 2016. The current circumstances demand that highway safety engineers develop techniques and methods for roadway and roadside improvements to reduce the probability and severity of crashes. Even though the identification of basic cause of road accident is difficult, an understanding of accident causing situations will help in formulation of accident prevention strategies. In this study an attempt is made to develop model using Artificial Neural Network that explains the relationship between the number of accidents and geometric and traffic characteristics of roadways.

2. ACCIDENT CAUSING FACTORS

Influencing factors for cause of accidents can be classified as those related to behaviour of driver, road geometrics, vehicles, traffic and environment. Further road, road user and traffic factors are the three prominent factors that can be controlled by highway/traffic engineers and can be incorporated in the safety oriented design to minimize the road accident causing factors.

2.1 ROADWAY AND TRAFFIC VARIABLES

The roadway and traffic variables can be concluded as (i) Lane/Pavement Width (ii) Width of Shoulder (iii) Condition of Pavement and Shoulder (iv) Horizontal Curves (v) Traffic Volume (vi) Speed. The literature study pointed that there is negative correlation between accidents and width of pavement and width of shoulders [6], [3]. Researchers also correlated accidents with paved and unpaved shoulders and reported that paved shoulders showed reduction in accidents [4], [10]. A few studies showed an increase of accidents at curves due to head on collision or loss of control [5]. Traffic volume is another significant variable in predicting accidents. The studies on traffic variables pointed that 27% higher AADT is predicted to increase accident rate by 16.4% [6], traffic flow strongly affects the accident rate [7], [8]. The number of single vehicle accidents decreases as density and volume to capacity ratio increases and number of multi vehicle accidents increases with vehicle density and capacity ratio [9]. Further, the literature revealed that passenger cars, night-time, and rural areas are more dangerous in terms of driver injury severity [2].

2.2 MODELLING FOR PREDICTION AND ANALYSIS OF ACCIDENTS

For monitoring the effectiveness of various road safety measures that have been initiated to minimize accident occurrences, Accident Prediction Models can be utilized. Also, the prediction models give an idea to transportation planners in determining new policies and strategies about road safety. Over the past years, many models have been developed to estimate traffic accidents all over the world [11]. In recent years, research related to analyses and prediction of accident frequencies have adopted statistical models like linear regression, Poisson or negative binomial regression models assuming accidents on highway as random events [12]. However, recent research has pointed that linear regression has undesirable statistical properties when applied to accident analysis [12]. Also, another drawback is that linear regression is not restrained from negative accident frequency. This would be a significant factor where a main highway section has a low or no accident frequency for some period of time. Negative accident prediction will bias the estimated coefficients, invalidating the model unless corrective measures are taken [13]. Research also showed that Poisson Regression approaches have one important constraint that the mean must be equal to the variance which if violated, the standard errors estimated by the maximum likelihood method, will be biased, and the statistical analysis from the model will be incorrect. Many studies have shown that accident data are usually over-dispersed, when the variance exceeds the mean, therefore, incorrect estimation of the likelihood of accident occurrence could result in applications of the Poisson regression model [14]. In efforts to overcome the problem of over-dispersion, researchers began to employ the Negative Binomial (NB) distribution instead of the Poisson distribution, which relaxes the mean, equals to variance constraint, and hence can accommodate over-dispersion in crash data counts [14]. However, NB models have some limitations such as the inability to handle under-dispersion of accident counts when the mean of the accident counts is higher than the variance. Rarely this phenomenon can arise when the sample size is very small, leading to erroneous parameter estimates [15].

The relationship between the dependent variables and the associated independent variables in accident modeling is not sufficiently explained by linear functions. Hence, non-linear approximators such as neural networks have also been explored. Artificial Neural Networks (ANNs) are a class of computational intelligence tools that can be used for prediction and classification problems. ANNs can model very complex non-linear functions to high accuracy levels using a process of learning that is similar to the learning procedure of the cognitive system in the human brain. The neural network body is composed of input layers, hidden layers, and output layers. These models can be trained to approximate any nonlinear function to a required degree of accuracy using a learning algorithm (such as back propagation) that would give the desired output, in a supervised learning process. ANNs have some advantages over the statistical models. For instance, regression models need a pre-defined relationship or functional form between the dependent variable (crash frequency) and the independent explanatory variables that can be estimated by some statistical approaches, whereas the ANNs do not require the establishment of these functional forms, and can be easily applied in the analysis [16].

3. LITERATURE REVIEW

H.T. Abdelwahab and M. A. Abdel-Aty (2001) deployed ANN learning to predict driver injury severity in traffic accidents at signalized intersections. The study concluded that rural intersections are more dangerous in terms of driver injury severity than urban intersections. The speed ratio increases the likelihood of injury severity and vehicle type plays an important role in driver injury severity. Further, ANN showed a better generalization performance of 65.6 and 60.4 percent for the training and testing phases respectively as compared with ordered logit model performance of 58.9 and 57.1 percent for the training and testing phases respectively [17].

Semeida et al. (2002) have used ANN and statistical models in investigating the factors contributing to accidents on Rural Roads in Egypt. The results of this study showed that in general Multi-layer Perceptron models perform better than the other models used for data prediction. The study concluded that the decrease in shoulder width lead to more and severe accidents, the increase of heavy vehicles percentage in traffic composition lead to more accidents, the increase in trailers and motor cycles percentage lead to more fatal accidents, and also, the increase in the number of openings in road sections leads to more fatal injury accidents[18].

Chang (2005) in his comparative study used Negative Binomial Regression versus Artificial Neural Networks (ANN) to analyze the accident frequencies of the freeways in Taiwan. He contributed ANN as a "consistent alternative method" in analyzing freeway accident frequencies. He underlined the erroneous estimation risk of Negative Binomial Regression model, when its assumptions were violated. However ANN is more powerful since this method do not need any predefined underlying relationship between dependent and independent variables [13].

Akgungor and Dogan (2009) introduced accident prediction models for Turkey using ANN and nonlinear regression approaches to estimate the number of accidents, injuries and deaths. They compared ANN and nonlinear models in terms of various error expressions. The study showed that ANN model had better results against the nonlinear regression models [11].

Dr Jain S.S et al (2011) in their study RSA for four lane National Highways in India used Multiple Linear Regression Analysis and ANN Model for prediction of Accident Rate. On the basis of developed models, it was found that ANN models revealed very good correlation between observed and predicted accident rates [19].

Sharaf Alkheder et al. (2016) explored the effectiveness of ANN for severity prediction of Traffic Accident in Abu Dhabi along with ordered probit model. The researchers used the modelling techniques for four injury severity variables namely death, severe, moderate and minor. They concluded that ANN has performed better with 74.6% accuracy value as compared to 59.5% for ordered probit model in accident prediction [20].

4. DATA COLLECTION

The accident prediction model developed in this study is based on the segments of non urban two lane highway NH-7 from Bathinda to Sangrur in the State of Punjab, India. Each segment is having a length of 5 kilometres. The various data related to road accident, road geometrics and traffic were collected for each segment.

4.1 Accident Particulars

Data was collected from the database of concerned police stations from First Information Reports for last 5 years. Other details have been studied, blackspot have been identified on study stretch by the authors and can be referred by readers [21].

4.2 Roadway Data

The variables related to roadway considered in this study are Land-use (L), Width of Pavement (WP), Width of Shoulder (WS), Shoulder Type (ST), Horizontal Curves (HC), Number of Junctions (NJ), Bus Stops (BS) and Road Side Friction (RSF). Pavement Width and Shoulder Width were measured physically in the field for all segments with tape. Rating for pavement and shoulder condition were rated for all segments by visual inspections. Road inventory details like curves, intersections, bus stops, access roads were counted manually.

4.3 Traffic Data

The traffic volume data was collected at mid block of all the segments for 12 hours on three days (Monday, Thursday and Saturday). Manual counting was done to understand the movement and flow of traffic at the site. Spot speed data was obtained from the field by recording the observations over a 50m stretch with two observers at each end with stopwatch.

TABLE 1: INPUT VARIABLES

Sl.No.	Input Parameter	Minimum Value	Maximum Value	Mean
1	AADT (PCU/Day)	14.565	18.547	17.167
2	Heavy Vehicle (%)	34.5	41.5	37.70833
3	Speed (Kmph)	69	76.2	71.7833
4	Width of Pavement (m)	6.8	7.2	6.9833
5	Width of Shoulder (m)	0.5	1.75	1.0833
6	Number of Junctions	2	4	3.0833
7	Number of Curves	1	3	2.16667

5. ARTIFICIAL NEURAL NETWORK MODELLING

The data set for accident modelling using ANN with data set of 761 records collected from the concerned police stations on study stretch of NH-7 was divided into two parts in random order. The data was collected from year 2006 to 2013 [21]. In the process of model development the data was used in proportion of 70%, 15%, and 15% for model training, testing and validation respectively. The data set from year 2014 to 2016[21] was used for cross validation of the developed model as shown in Figure 6.

5.1 Development of Artificial Neural Network Model

The Neural Network (Fig 1) [22] used in the study is feed forward error-back propagation neural network with one hidden layer. Artificial neural network tool box available in MATLAB R2015b was used in this study. Modelling was done using each variable and the most significant variables were found out. The significant variables found out from the study are as tabulated in Table 1.

- AADT (X1)
- Width of Pavement (X4)
- Number of Curves (X7)
- Heavy Vehicle (X2)
- Width of Shoulders (X5)
- Speed (X3)
- Number of Junctions (X6)

The training process involved, setting initial values for weights, measuring the error, adjusting the weights and further the weights continue to be modified as each error is computed and then the validation performance of the network model where the predicted values are compared with the actual as given by the validation data. Further, the values predicted are compared with input using cross validation data that was not used either in training or validation. Graph obtained trough data simulation is shown in Figure 2.

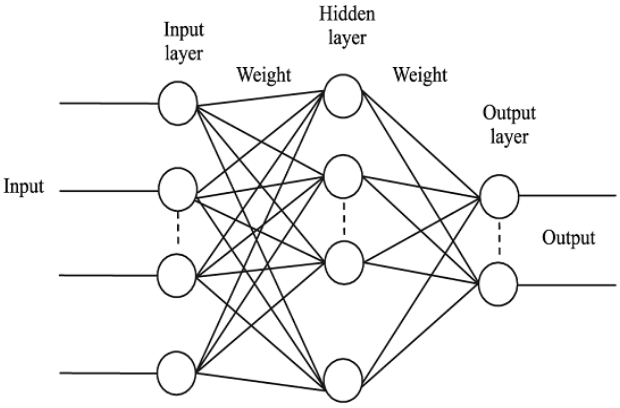


Fig 1: Neural Network Architecture

ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 7, Issue 2, pp: (412-420), Month: April - June 2019, Available at: www.researchpublish.com

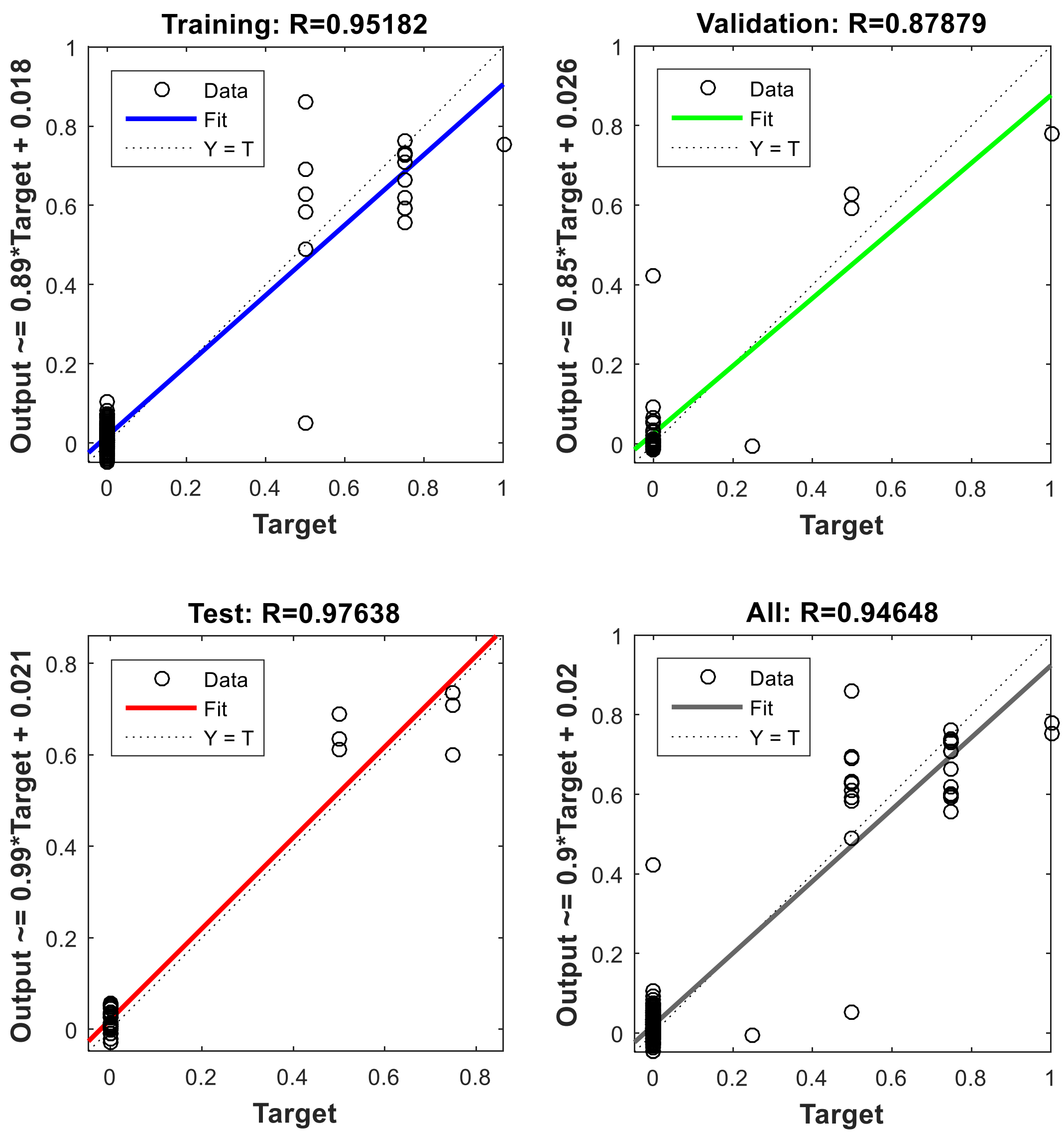


Fig 2: ANN Output

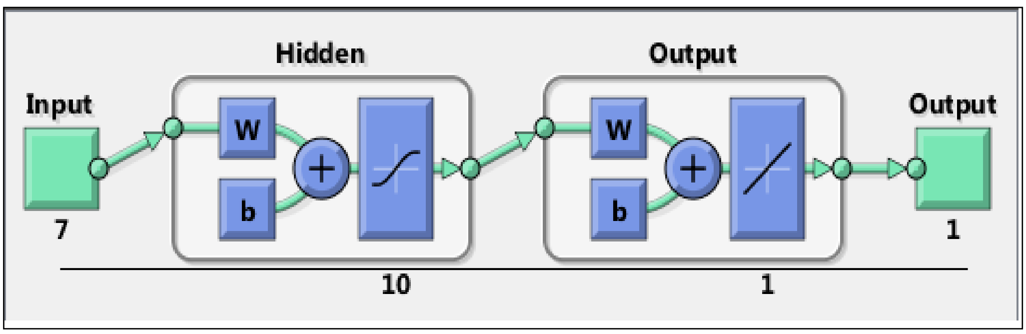


Fig 3: Neural Network Diagram

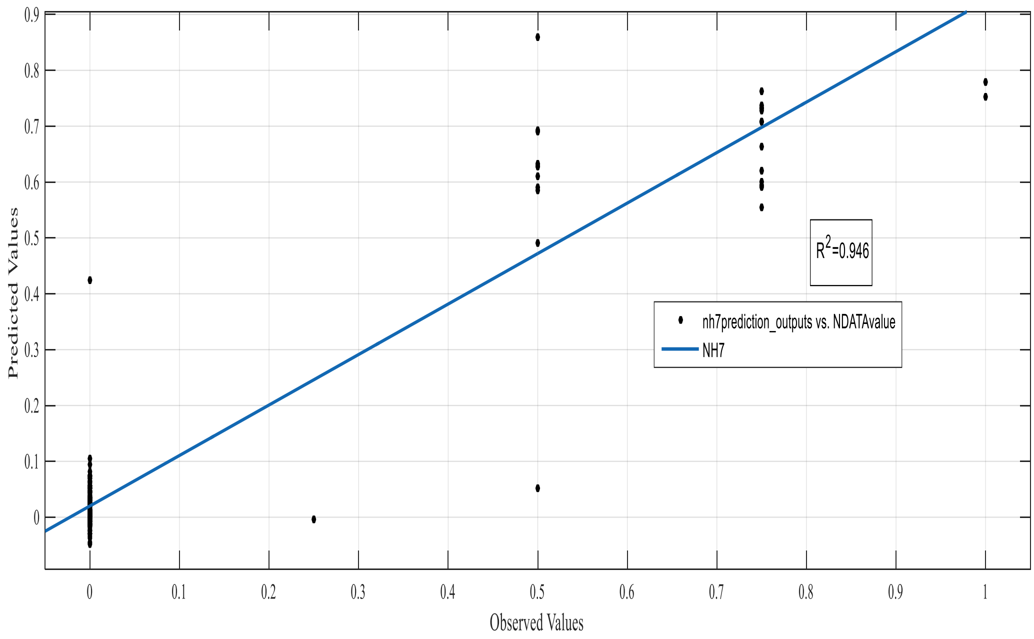


Fig 4: ANN Model between Observed and Predicted Number of Accidents

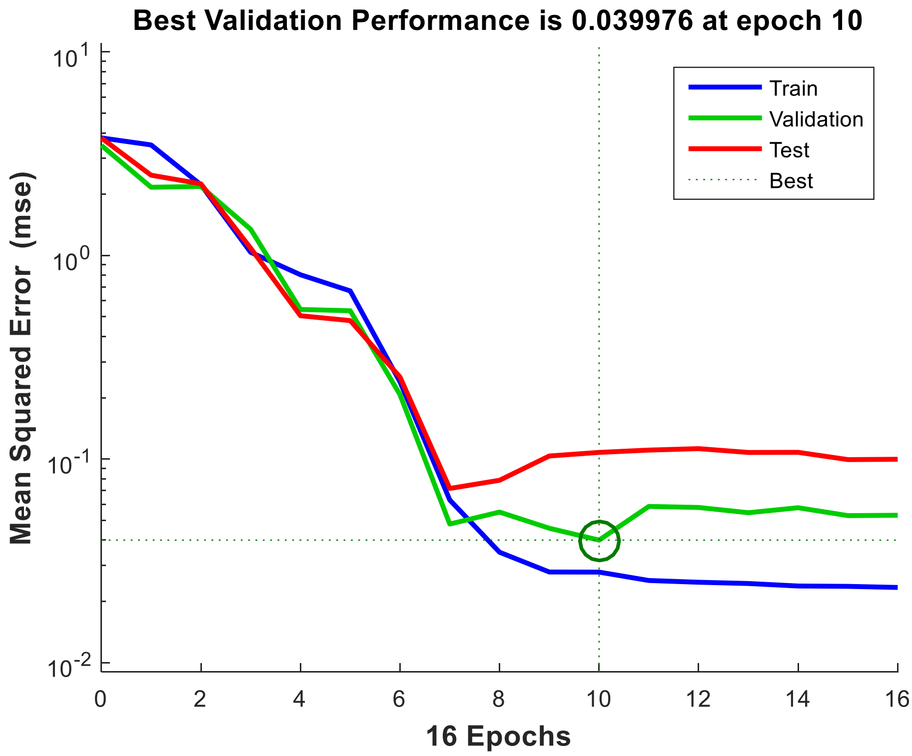


Fig 5: Validation Performance Analysis

6. RESULT ANALYSIS

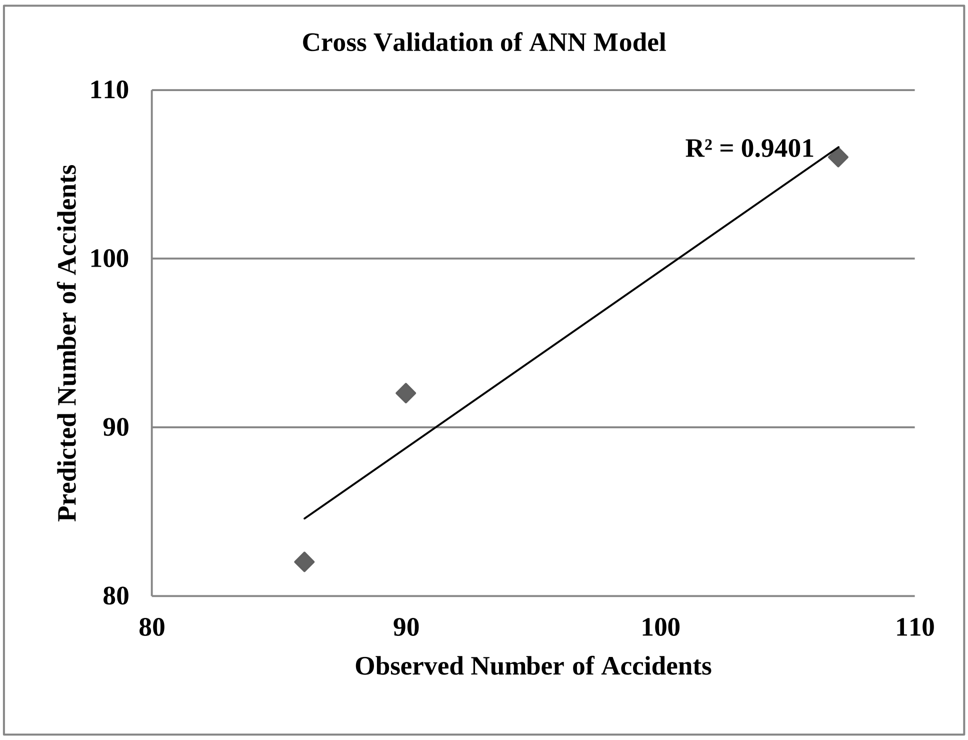
Statistical summary of the variables using MINITAB package revealed that parameters like shoulder type, bus stop and road side friction had less correlation to independent variable number of accidents leading to omission of these variables. The Neural Network Toolbox from the MATLAB library was used to train, validate and test the dataset. The training algorithm used Levenberg Marquardt optimization. The network used is as shown in Figure 3. Validation vectors are used to stop training early if the network performance on the validation vectors fails to improve or remains the same

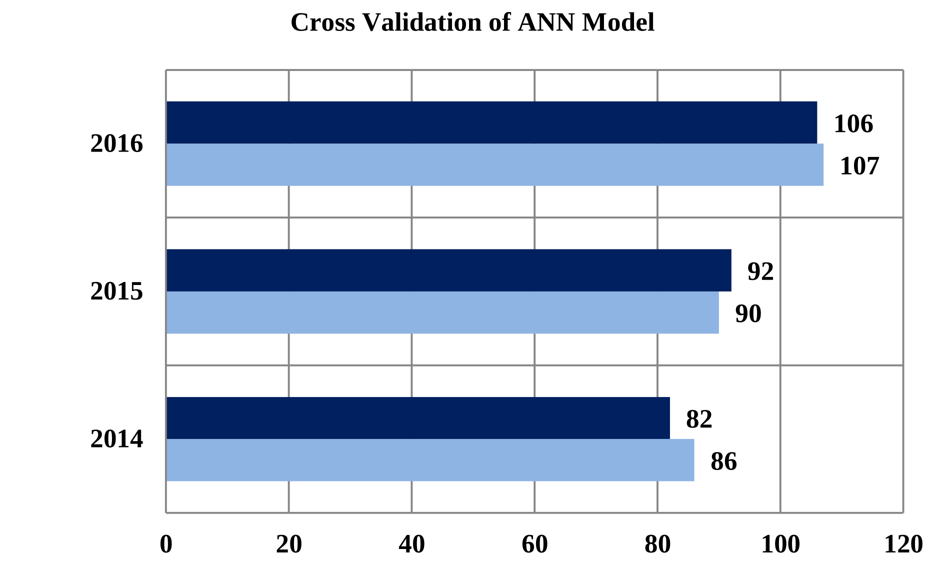
ISSN 2348-1218 (print)
International Journal of Interdisciplinary Research and Innovations ISSN 2348-1226 (online)
Vol. 7, Issue 2, pp: (412-420), Month: April - June 2019, Available at: www.researchpublish.com

for *max_fail* repetitions in a row. Test vectors are used as a further check that the network is generalizing well, but do not have any effect on training. The training parameters for *trainlm* are stipulated in *Table 2*. The input layer neurons that represent seven different explanatory variables and the output layer have one neuron representing dependent variable – accident number. The transfer function at the hidden layer is *sigmoid* and the output layer is *purelin.* To select the number of layers and hidden nodes, experiments with different values was carried out. The number of hidden nodes that gives the best network performance is selected using coefficient of correlation as shown in Figure 2, MSE and performance as shown in Figure 5. The validation performance of the network model where the predicted values are compared with the actual as given by the validation data is shown in Figure 4 indicating validity of developed model. Further, the developed model was cross validated with three years data which was not used in model calibration process. The result in Figure 6 shows R^2 value of 0.94 for cross validation of model, indicating the robustness of the developed model.

TABLE 2: TRAINING PARAMETERS

net.trainParam.epochs	1000	Maximum number of epochs to train
net.trainParam.goal	0	Performance goal
net.trainParam.max_fail	6	Maximum validation failures
net.trainParam.min_grad	1.00E-07	Minimum performance gradient
net.trainParam.mu	0.001	Initial mu
net.trainParam.mu_dec	0.1	mu decrease factor
net.trainParam.mu_inc	10	mu increase factor
net.trainParam.mu_max	1.00E+10	Maximum mu
net.trainParam.show	25	Epochs between displays
net.trainParam.showCommandLine	FALSE	Generate command-line output
net.trainParam.showWindow	TRUE	Show training GUI
net.trainParam.time	inf	Maximum time to train in seconds





	2014	2015	2016
Predicted	82	92	106
Observed	86	90	107

Fig 6: Cross Validation of ANN Model between Observed and Predicted number of Accidents

7. CONCLUSIONS

The study presented a computational tool for analysis of variables leading to cause of accidents on National Highway having heterogeneous movement of traffic. From data simulation, it was found that Traffic Volume, Speed, Width of Pavement and Width of Shoulder were the variables having major role in accident occurrence. Further, the results indicated that the developed model fitted well in predicting the number of accidents on study stretch. Also, the data used for cross validation of the model from 2014 to 2016 gave a R^2 value of 0.94 validating the Neural Network Model. Hence, the developed model could be used on roads of same traffic and geometric conditions. Moreover, Researchers would also employ the same model for other NH and SH of the State of Punjab in India.

REFERENCES

[1] Ministry of Road Transport and Highways, Transport Research Wing "Road Accidents in India," 2016.

[2] Mohamed A. Abdel-Aty and Hassan T. Abdelwahab .2004, "Predicting Injury Severity Levels in Traffic Crashes: A modeling approach", Journal of Transportation Engineering, Vol. 130, No. 2, March 1, 2004

[3] Milton J and Mannering F., 1996, "The Relationship between Highway Geometrics, Traffic Related Elements and Motor Vehicle Crashes", Report WA-RD 403.1, Washington State Department of Transportation.

[4] Ogden K.W, 1997, "The effects of Paved Shoulders on Crashes on Rural Highways", Accident Analysis and Prevention, 29, 352-362.

[5] Elvik, R., Muskaug. R., 1994. Konsekvensanalyser og trafikksikkerhet. Metode for beregning av konsekvenser for trafikksikkerheten av tiltak pa˚ vegnettet. TØI-rapport 0281. Transportøkonomisk institutt, Oslo.

[6] Jianming M., Kockelman, M.K., Damien, P., 2008, "A Multivariate Poisson-Lognormal Regression Model for Prediction of Crash Counts by Severity-using Bayesian Methods. Accident Analysis and Prevention 40, 964-975.

[7] Greibe, P. (2003), "Accident Prediction Models for Urban Roads". Accident Analysis and Prevention, 35, 273-285.

[8] Haynes, R., Lake, L.R., Kingham, S., Sabel, C.E., Pearce, J., Barnett, R., 2008. " The Influence of Road Curvature on Fatal Crashes in New Zealand". Accident Analysis and Prevention, 40, 843-850.

[9] Lord, D., Manar, A.,Vizioli, A., 2005. Modeling Accident Flow Density and Accident Flow V/C ratio Relationships for Rural and Urban Freeway Segments". Accident Analysis and Prevention, 37, 185-199.

[10] Zongzhi Li, et al (2013), "Safety Effects of Shoulder Paving for Rural and Urban Interstate, Multilane, and Two-Lane Highways". Journal of Transportation Engineering, Vol. 139, No. 10, October 1, 2013.

[11] Akgungor, A. P. & Dogan, E. (2009), "An application of modified Smeeds, adapted Andreassen and artificial neural network accident models to three metropolitan cities of Turkey", Proc. of Scientific Research and Essay, vol.4 (9). Kirikkale, Turkey: Academic Journals. pp. 906-913.

[12] Chen W., Jovanis P.P., "Method for Identifying factors contributing to driver-injury severity in traffic crashes", Transportation Research Record, Issue 1717, 2000, PP.1-9.

[13] Chang L. (2005), " Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network", Safety Science 43:541-557.

[14] Lord, D., and Mannering, F. (2010) The Statistical Analysis of Crash Frequency Data: A Review and assessment of methodological alternatives. Accident Analysis and Prevention, 44, 291-305.

[15] Lord, D. (2006) Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter. Accident Analysis and Prevention, 46, 751-766.

[16] Abdulhafedh, A. (2017) Road Crash Prediction Models: Different Statistical Modelling Approaches. Journal of Transportation Technologies, 7, 190-205.

[17] Abdel wahab, H.T. and Abdel-Aty, A.A., 2001. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record*; 1746, pp. 6–13.

[18] A. Semeida, M. El-shabrawy, M. Sabry, M. Sadek, and I. Hashim, "Investigation of Factors contributing to accidents on Rural Roads in Egypt,"Journal of Advanced Research,Vol. 2, 2002.

[19] Dr Parida M, P.K Singh, Dr S.S Jain (2011), " Road Safety Audit for Four Lane National Highways", 3rd International Conference on Road Safety and Simulation, Indianapolis, USA,2011.

[20] Sharaf Alkheder, Madhar Taamneh Salah Taamneh (2016), "Severity Prediction of Traffic Accident Using an Artificial Neural Network", Journal of Forecasting, 2016.

[21] Jain Nirpinder and Dr. Sanjiv Kumar Aggarwal (2019), " Identification and Prioritization of Blackspots on NH-7 in state of Punjab, India", International Journal of Technical Innovation in Modern Engineering and Science ,Vol 5, Issue 3, 2019.

[22] Bayata F.H, Fatih Hattatoglu, Neslihan Karsli (2011), "Modelling of monthly traffic accidents with the artificial neural network method", International Journal of the Physical Sciences Vol. 6(2), pp. 244-254, 18 January, 2011.