Prediction of Road Accidents in Delhi using Back Propagation Neural Network Model

Sushant Sikka
Department of Applied Mathematics
Delhi Technological University
New Delhi, India
sushantsikka.506@gmail.com

Abstract : Road accidents cause more than 1,35,000 deaths in India every year which is higher than any other country in the world. Increasing motorization without adequate institutional mechanism is a major cause of this problem. Although, the number of fatalities caused by road accidents in Delhi has been declining over the years, but the situation is still alarming. In this paper we use the back propagation neural network model for predicting the road accidents and fatalities caused by them in Delhi based on the number incidents of drunken driving, over speeding, driving without wearing seat belts and helmets and the number of vehicles registered in Delhi. The data of these traffic violations and year wise vehicle registration was collected for the years 2008 to 2012. Out of 5 data sets, 3 were used for training and 2 data sets were used for testing. The data was fed to a neural network of 1 hidden layer and another of 2 hidden layers consisting 11 nodes each. The lowest average error obtained was 6.3728 % for the neural network with a single hidden layer.

Index Terms : Artificial Neural Networks; Road Accidents; Back Propagation Algorithm; Prediction; Linear Regression; Nonlinear Regression

I. INTRODUCTION

Artificial neural networks use mathematical formulations to mimic the nervous system found in animals and thereby perform tasks like image recognition, voice recognition, predictions, etc that would otherwise be very difficult and time consuming to perform. While a conventional digital computer is a serial machine and is fast in numeric computations, the brain is highly parallel and much slower. Unlike computer systems, brain can learn and also teach itself. An ANN being a computational model that is based on biological neural networks has nonlinearity that is distributed throughout. This introduces the learning capability in the ANN and makes it adaptive, that is, the ANN can adapt the free parameters to change based on the information that flows through it. An ANN also has the advantage of fault tolerance that makes it to degrade performance in case some part of the network is malfunctioning or not working rather than the whole network crashing. Also an ANN is efficient in knowledge acquisition even under uncertainty.

The number of traffic collision related fatalities in New Delhi are nearly 40 times the deaths by the same cause in London. The major causes for these accidents identified by World Health Organization in “Global Status Report on Road Safety” were over speeding, driving under influence and not wearing helmets and seat belts. The large number of motor vehicles plying on roads has also been a major contributor to the frequency of road accidents. The Planning Commission of India in its research conducted in 2002 found that traffic collisions caused a monetary loss of nearly $ 10 billion during 1999 – 2000.

In this paper, we examine the ability of Back Propagation Neural Network in predicting the road accidents in Delhi. The rest of the paper is organized as follows. Section II deals with Artificial Neural Networks, the Artificial Neurons and Fully Connected ANN. Section III contains an elaboration on Back Propagation Algorithm. Section IV talks about the normalization of data and the percentage of data sets used for training and testing. In the following sections we discuss the errors in training and testing for ANN with 1 hidden layer and for the one with 2 hidden layers. We also compare these errors to the ones observed using Linear and Non-Linear Regression. In the final section of this paper we have the conclusion.

II. ARTIFICIAL NEURAL NETWORK

Under this section we talk about the structure of a basic artificial neuron and how these artificial neurons are inter-connected to form an ANN.

A. Artificial Neurons

An artificial neuron is to an ANN as a biological neuron is to the brain. An artificial neuron is the basic building block of the ANN. The basic model of an artificial neuron is shown below.

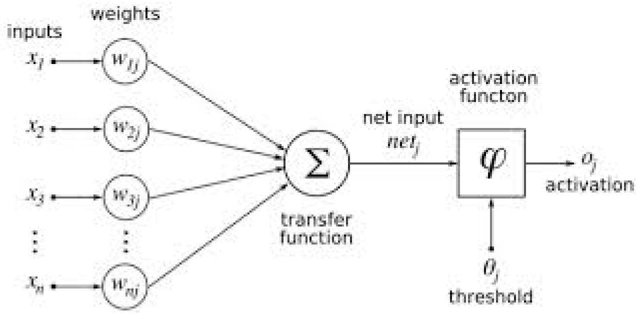

Fig 1 Artificial Neuron Structure

In the diagram shown above $x_1, x_2, x_3, ..., x_n$ are the inputs to the neural network and the weights to which these inputs are multiplied are $w_1, w_2, w_3, ..., w_n$. The activation function then generates output for the net input $x_1w_1 + x_2w_2 + x_3w_3 + ... + x_nw_n$. The result obtain is the output of the artificial neuron.

B. Fully Connected Artificial Neural Network

A fully connected ANN has all its artificial neurons or nodes in the input and each of the hidden layers connected to every node of the following layer. That is, the output of each node in the input and hidden layer acts as the input to each node in the following layer of nodes in the network. In this paper, we have used a fully connected ANN.

III. BACK PROPAGATION ALGORITHM

Back Propagation Algorithm is a member of the family of gradient descent algorithms and is an implementation of Delta rule. By iteratively adding the negative of the slope of the function to the value of the function at a given point, the algorithm tends to reach the maxima or the minima. This is how the gradient descent algorithm minimizes the error.

Back Propagation is a supervised learning method since, the output for a given set of data inputs should be known before proceeding with the algorithm. The output is the key with which the algorithm ascertains the loss function gradient.

Consider a fully connected artificial neural network with 1 hidden layer as shown below.

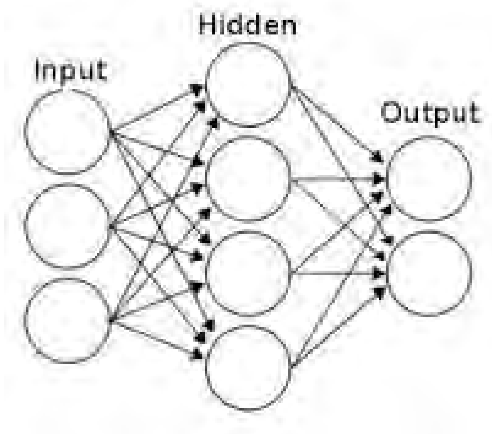

Fig 2 A fully connected ANN

Let the number of nodes in input layer be I, the number of nodes in hidden layer be J and number of them in output layer be denoted by K.

Then Error in output layer is $\frac{1}{2}\sum_{k=1}^{K}(O_k - t_k)^2$. The back propagation algorithm aims to minimize $\frac{\delta E}{\delta W_{jk}^{l}}$.

For unipolar sigmoid function $\sigma(x) = \frac{1}{1-e^{-\lambda x}}$,

$$\frac{\delta E}{\delta W_{jk}} = (O_k - t_k)\left[\sigma(x_k)\left(1-\sigma\left(x_k\right)\right)\right]\frac{\delta x_k}{\delta W_{jk}} \tag{1}$$

That is for the output layer $\frac{\delta E}{\delta W_{jk}^{l}}$ is ,

$$\frac{\delta E}{\delta W_{jk}} = (O_k - t_k)\left[\sigma(x_k)\left(1-\sigma\left(x_k\right)\right)\right]O_j \tag{2}$$

And for the intermediate hidden layer,

$$\frac{\delta E}{\delta W_{ij}} = \frac{\delta}{\delta W_{ij}}\left(O_j\right)\sum_{k\in K}\left[(O_k - t_k)O_k\left(1-O_k\right)W_{jk}\right]. \tag{3}$$

For output layer $\frac{\delta E}{\delta W_{jk}} = \left(O_j\right)\delta_k$ where $\delta_k = O_k\left(1-O_k\right)\left(O_k - t_k\right)$.

For hidden layer $\frac{\delta E}{\delta W_{ij}} = O_i\delta_j$ where $\delta_j = O_j\left(1-O_j\right)\sum_{k\in K}\delta_k W_{jk}$. (4)

The back propagation algorithm is as follows:

I. Assign arbitrary initial values of weights. Feed into the network the input data.

II. For each output node compute $\delta_k = O_k\left(1-O_k\right)\left(O_k - t_k\right)$

III. For each hidden node compute $\delta_j = O_j\left(1-O_j\right)\sum_{k\in K}\delta_k W_{jk}$

IV. Update $W = W + \left(-\eta\right)\delta_l O_{l-1}$.

When steps I – IV are followed iteratively for a given set of data we train the weights such that they predict a fairly accurate value for different data sets of same attributes.

Some characteristics associated with back propagation algorithm are:

- It is imperative that the activation function of the artificial neurons be differentiable.
- The learning process by back propagation algorithm is very slow.
- It requires the data input to be normalized before being fed to the network.

IV. DATA PRE – PROCESSING

The data of the number of road accidents in Delhi, number of over speeding challans, number of drunken driving challans, number of challans for not wearing seat belts and helmets and the number of vehicles registered in Delhi and year wise fatalities and traffic collision incidents for the years 2008 to 2012 was collected. All the data was normalized between 0 and 1.0 by taking the maximum and minimum values from their respective attributes. 60% of the data was used for Training and 40% was used for Testing.

V. DATA USED

Table 1 Data used for Training and Testing

Year	Over-speed	Seat-belts	drunken driving	helmet	vehicles registered in Delhi	Fatalities	Total accidents
2008	189622	158932	8296	191612	359196	2093	8435
2009	222659	134494	12784	255686	413967	2325	7516
2010	143242	69860	11388	207474	485659	2153	7260
2011	44618	224368	18073	331463	512420	2110	7280
2012	261377	319640	25678	341686	529268	1866	6937

VI. NEURAL NETWORK ARCHITECTURE

Prediction values for the given values of data sets is obtained for neural network architectures with single hidden layer and another with two hidden layers. The number of inputs for the network is 5 so the input layer has 5 nodes. Each hidden layer has (2n+1) nodes that is, 11 nodes and the output layer has 1 node.

VII. RESULTS OBTAINED

Table 2 RMS Errors for Neural Networks with 1 Hidden Layer and 2 Hidden Layers

Number of hidden layers	Training RMS Error	Testing RMS Error	Combined RMS Error
1	7.1848%	8.389%	7.6894%.
2	7.2138%	15.528%	11.299%

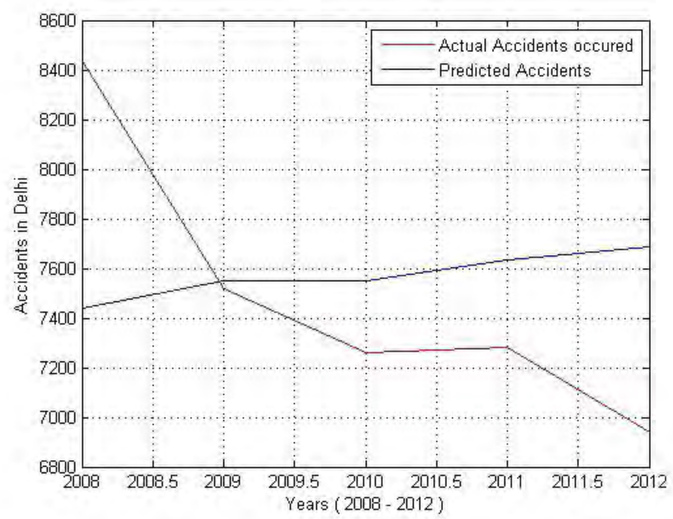

Fig. 3 Actual Accidents in Delhi vs. Predicted Values

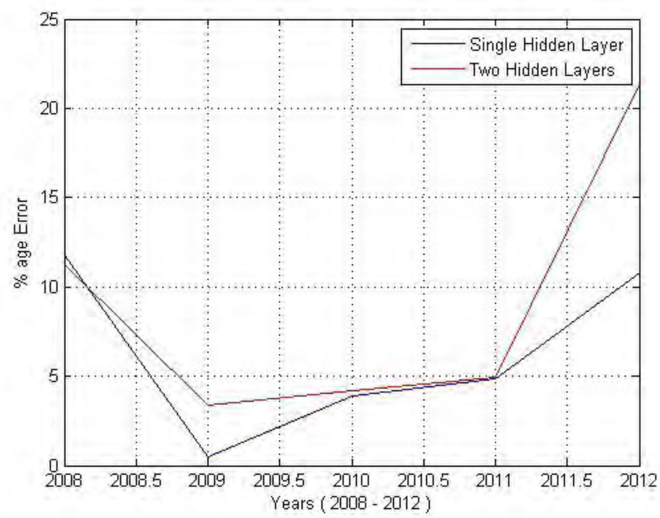

Fig 4 Comparing the errors for neural network with 1 hidden layer and 2 hidden layers

VIII. COMPARISON WITH LINEAR REGRESSION

Using the data of year wise accidents from 2008 – 2012, a linear regression model was used to predict the number of accidents in Delhi for the year 2013.

The predicted accidents for the year 2013 were 6633 while the actual accidents are 7566. The % Error obtained is 12.33%.

Fig 5 LabVIEW Front Panel showing Best Fit, Intercept, Slope, Residue and Error for Linear Regression

As can be seen from LabVIEW Front Panel, the year wise accidents in Delhi for the period of 2008 – 2012 has been taken in Table Y and numbers from 1 to 5 have been assigned for years 2008 to 2012. The slope and intercept obtained are -303.64 and 8455.2.

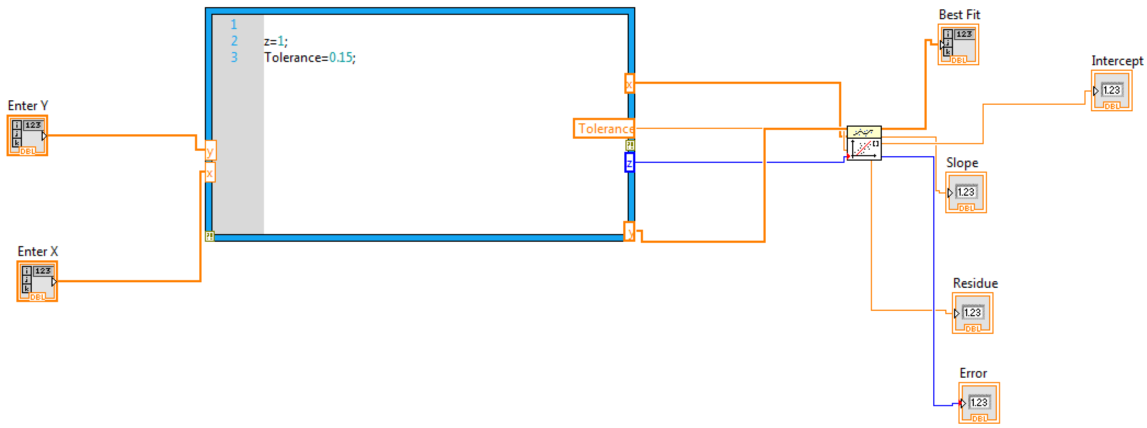

Fig 6 LabVIEW Block Diagram for Linear Regression

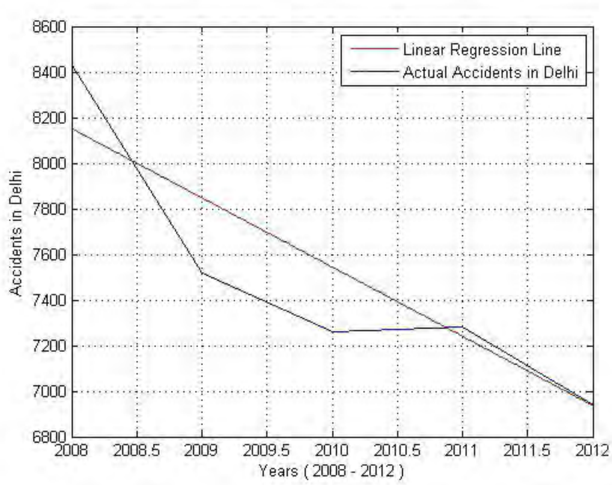

Fig 7 Linear Regression line for the given data

IX. COMPARISON WITH NONLINEAR REGRESSION

Using the data of year wise accidents from 2008 – 2012, a linear regression model was used to predict the number of accidents in Delhi for the year 2013.

The predicted accidents for the year 2013 were 5340. The % Error obtained is 29.42%.

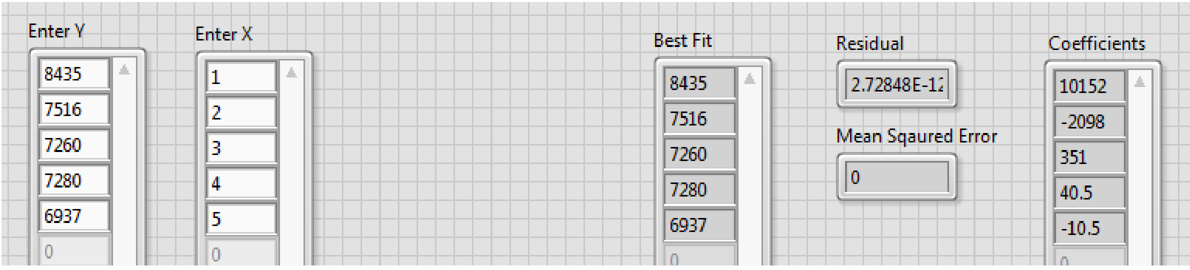

Fig 8 LabVIEW Front Panel showing the Best Fit, Coefficients and Residual for Nonlinear Regression

The equation of the non linear regression curve obtained is Y=(-10.5*X.^4)+(351*X.^2)+(40.5*X.^3)+(-2098*X)+10152.

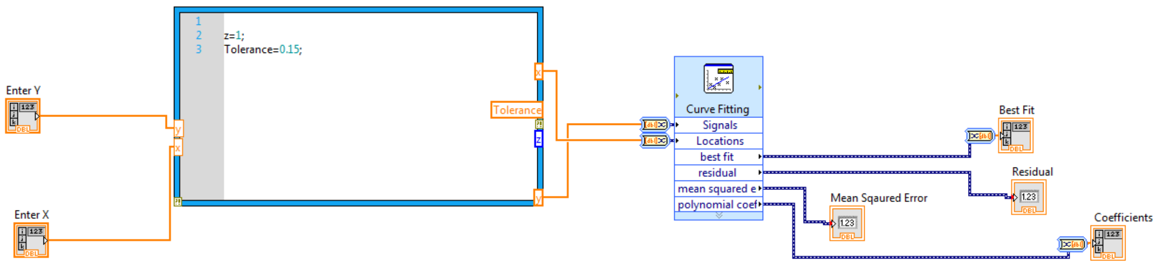

Fig 9 LabVIEW Block Diagram for Nonlinear Regression

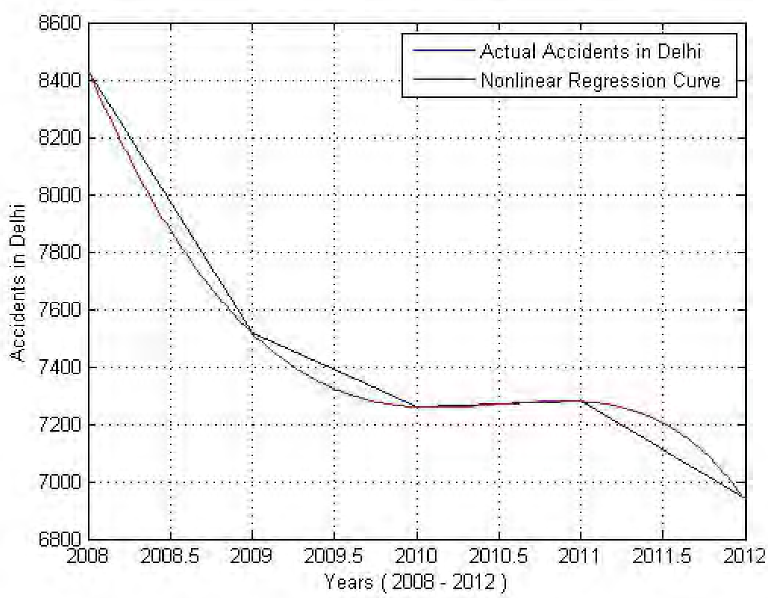

Fig 10 Nonlinear Regression line for the given data

X. CONCLUSION

In this paper, an Artificial Neural Network was applied for predicting the road accidents in Delhi. We also laid out the theory of ANN and Back Propagation Algorithm. According to results, the neural network was found to be fairly accurate as can be seen from Fig 3, Fig 4 and Table 1. We hence, conclude that back propagation neural networks model can be used for prediction of road accidents. The least error observed was as low as 0.479%. The RMS error of neural network is much lower than the results obtained from Linear and Non-Linear regression models as inferred from Fig 7 and Fig 10. Hence, artificial neural networks are better for prediction than linear and non-linear regression.

XI. REFERENCES

[1] List of countries by traffic related deaths, Wikipedia http://en.wikipedia.org/wiki/List_of_countries_by_traffic-related_death_rate

[2] 2013 saw a decline in road accidents in Delhi, Zee News http://zeenews.india.com/news/delhi/2013-saw-decline-in-road-accidents-in-delhi_901334.html

[3] Shankar Kumar Jha ,Road Safety in Delhi: A critical analysis of the role of public institutions "The Planning Commission of India in its research conducted in 2002 found that traffic collisions caused a monetary loss of nearly $ 10 billion during 1999 – 2000."

[4] Traffic Collisions in India, Wikipedia "The frequency of traffic collisions in India is amongst the highest in the world. A National Crime Records Bureau (NCRB) report revealed that every year, more than 135,000 traffic collision related deaths occur in india". http://en.wikipedia.org/wiki/Traffic_collisions_in_India

[5] Abhishek Kar, Stock Prediction using Artificial Neural Networks

[6] Sudarshan K. Valluru, T. Nageswara Rao Introduction to Neural Networks, Fuzzy Logic and Genetic Algorithms – Theory and Applications ,2011,pp.63-64, 2-4

[7] Enireddy Vamsidhar, K.V.S.R.P Varma, P.Sankara Rao, Ravikanth Satapati Prediction of Rainfall using Back Propagation Neural Network Model