Performance Analysis of Traffic Congestion Using Designated Neural Network Training Algorithms

Ituabhor, Odesanya¹,a* and Joseph Femi, Odesanya²,b

¹Department of Physics, Federal University Lokoja, Kogi State, Nigeria.
²Department of Transport, Federal University of Technology Akure, Ondo State, Nigeria.

a*e-mail: ituabhor.odesanya@fulokoja.edu.ng. b- mail: jfodesanya@futa.edu.ng

Keywords: Traffic flow congestion, Neural Network, Levenberg-Marquardt, Bayesian Regularization, Scaled Conjugate Gradient.

Abstract. A lot of neural network training algorithms on prediction exist and these algorithms are being used by researchers to solve evaluation, forecasting, clustering, function approximation etc. problems in traffic volume congestion. This study is aimed at analysing the performance of traffic congestion using some designated neural network training algorithms on traffic flow in some selected corridors within Akure, Ondo state, Nigeria. The selected corridors were Oba Adesida road, Oyemekun road and Oke Ijebu road all in Akure. The traffic flow data were collected manually with the help of field observers who monitored and record traffic movement along the corridors. To accomplish this, three common training algorithms were selected to train the traffic flow data. The data were trained using Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG) and Levenberg-Marquardt (LM) training algorithms. The outputs/performances of these training functions were evaluated by using the Mean Square Error (MSE) and Coefficient of Regression (R) to find the best training algorithms. The results show that, the Bayesian regularization algorithm, performs better with MSE of 2.37e-13 and R of 0.9999 than SCG and LM algorithms.

Introduction

Traffic flow and congestion problem in road transport system has turn out to be a significant real time problem in both developed and under developed countries [1]. Neural networks algorithm offers new architectural ideals for computing machines and new approaches for learning. However, the advancement of the understanding of the nervous system is the most important objective of neural network investigation [2]. The simple processing elements called neurons operates base on their local connection with other elements are known as Artificial Neural Networks (ANNs). The stable states of artificial neural networks are consistent with the current input and output values. ANNs design was motivated by the configuration and the function of the human brain, but elements and the architectures that are used in artificial neural networks processing, do not have anything in common with their biological analogue. Connections between the elements and the weights are simply adjustable parameters that enables the network to realize and perform different functions. When the weights of the network has a decision function that updates it, the network is said to have a training algorithm.

There are basically two training concepts in a network; the batch training and the classic training [3]. A study was carried out by [4] on back-propagation neural network (BPNN) in predicting traffic flow, occupancy and speed. The results were quite promising. Due to this, Neural Network methods have been commonly used for traffic flow prediction. There are several types of training algorithms in neural networks, such as Gradient descent, Bayesian regularization, Resilient, Levenberg-Marquardt, Scaled conjugate gradient etc, of these, this research focus on Bayesian regularization, Scaled conjugate gradient and Levenberg-Marquardt training algorithms being the most frequently used algorithms in prediction analysis. It is in fact, a herculean task estimating which of the training algorithms performs best [5]. The main objective of this research is to analyse the performance of traffic congestion using some designated neural network training algorithms on traffic flow in some selected corridors within Akure, Ondo state, Nigeria. This paper is organized as follows: Literature
review; this section takes a look at the relevant works carried by other researchers. The Materials and methods section, consists of data source and the neural network algorithms employed. The Result and discussion section, presents the result obtained from the various algorithm and discusses the implications of each of the algorithms outputs. Then, conclude the research in the Conclusion section.

Brief Information on Transport System and Traffic Congestion

Urban Transportation systems management and operation is key to improving the management of traffic movement within an urban centre, recently development have put Intelligent Transport System (ITS) as the fulcrum for a coordinated traffic network. Traffic flow projecting is very important deployment in traffic management and control strategies which is dominated by the emerging use of intelligent transportation systems [6]. Proper deployment of intelligent management of traffic flow will deliver travellers with more precise information on traffic and road condition which can further reduce the negative impact of congestion [7]. [8] posited that predicting of traffic flow is important and key to determining several aspects of road traffic capacity such as design of road way, planning, management and operations of urban roadways noting that the main reason is to increase traffic operational efficiency, capacity and level of service of the road transport system through wide and versatile use of advanced technological and telecommunication systems. Neural networks have played an extensive role in the forecasting of traffic volume data, in the Netherlands [9] used back propagation neural network in The Hague (Utrecht/Rotterdam) district to design a model that predicted traffic movement, speed and traffic occupancy of a selected road corridor. They observed that elasticity test serves as a good possibility to understand the established neural network model. Likewise, [10] did a study in which they compare between statistical models and neural network traffic flow prediction on motorway traffic data in France. Neural network approach for short-term traffic forecasting in Queensland, Australia for a highway stretching between Brisbane and the Gold Coast, which stood as an object focused approach was developed by [11]. [12] used input parameters of traffic flow difference and applied the concept of dynamic rolling prediction to project a new short term traffic flow prediction method by deploying Back Propagation neural network model, they then use the actual observed field data of traffic flow to develop a model structure, by repeating the calculation steps by step again. The consequences of the result they obtain show that the technique is viable, consistent, and practicable. While [13] present an approach of selecting a neural network structure by providing a wide-ranging basis for modelling complex interactions such as collaborating supervised and unsupervised learning rules in a network layer or deriving a new learning equation by integrating a recurring processing element into hidden layer of a feed forward topology without generating new equations. They thereby developed an innovative, genetic algorithm founded, multi layered structural optimization strategy which can assist in demonstrating traffic flow data with temporal and spatial characteristics. The performance of the network developed was validated by using it for both univariate and multivariate traffic flow data from an urban road traffic road. The result shows the capability of static neural network when fully optimized with the appropriate approach will provide satisfactory result. [14] developed a neural network model based on dynamic wavelength for traffic flow prediction in order to view the dynamics of traffic flow and pattern recognition which have enhanced ability to detect special features. Study by [15] have been able to applying different artificial neural networks (ANN) architectures with various input factors inputted to live data obtained from the field, in which they reveal that ANN modelling is an operational potent approach for short term traffic flow modelling. [16] employed ANN and ANFIS to solve a transhipment problem having nondeterministic input data and [17] for cost optimization in a multi-dimensional two-way system. In addition, [18] discovered that accurateness is very significant in predicting short term traffic flow using Kalman filtering they went ahead to proposed two (2)-dimensional forecasting method for traffic data. Their proposed method delivered an improved accuracy than the normal Kalman filtering methodology. Traffic stream are stochastic and highly non-linear in behaviour, machine learning techniques [19] have received a great attention and hence are taken as an alternative for traffic flow prediction Recently, [20] investigate the significance of deep learning which is a function of artificial neural network for automobile traffic
density prediction through traffic flow by carrying out a comparative analysis of orthodox (periodic) models, and multiple alternatives of repeated neural models, based on a forty (40) day-long traffic volume data from fifty-eight (58) cameras located at different locations around London, they found out that result obtained from the reconstructed traffic density had a minimal error based on the predicted data output using a spatial forecast model similar to the multivariate model by using supervise neural network (tree structured Parzen estimator (TPE)) algorithm. [21] used fuzzy info granulation method together with Artificial Neural Network (ANN), support vector machines (SVM) and KNN methods, they opine that interim prediction is more significant and thought-provoking than point prediction for traffic analyst when considering the future of ITS their data were collected from field study in carried out in America. However, their result shows that an increase in time sequence would transmit to steady predictor systems. [22] also studied Artificial Intelligent (AI) in traffic control system of mixed traffic for non-autonomous vehicles at some signalized road intersection, their investigation reviewed that AI techniques, and artificial neural network (ANN) model produced the best result for traffic congestion in a mixed traffic condition by eliminating or plummeting traffic volume in the case of non-autonomous vehicle in mixed traffic at the study site in South Africa. In Nigeria, few studies have considered two lanes divided urban roadway, studies carried out by [23] only considered traffic intersections and carried out manure traffic flow study. Therefore, it is imperative that a study like this should consider mixed traffic flow in Nigeria urban roads. The objective of this work is to analyse traffic data using some designated neural network algorithms to predict traffic flow and to know the algorithm that perform optimally by employing the Mean Square Error (MSE) and coefficient of regression.

Neural Network Training Algorithm

The simple processing elements called neurons which operates and connects with other elements are known as Artificial Neural Networks (ANNs). ANNs design was encouraged by the structure and the function of the human brain, but the elements and the architectures that are used in artificial neural networks processing, do not actually have anything in common with their biological models. Neural networks algorithm offers new architectural ideals for computing machines and new approaches for learning. There are basically three training algorithms in neural network [22]. Mathematical model of the neuron in Figure 1 has N number of inputs, denoted by \(X_1, X_2… X_n\). A weight \(W\) is assigned to each line connecting these inputs to the neuron. The assigned weights are denoted as \(W_1, W_2… W_n\) respectively. The action that decides if the neuron will be fired or not, is given by the expression in equation (1).

\[
a = \sum_{j=1}^{n} W_j X_j
\]  

(1)

The output is dependent on the activity or action of the activation function \(\varphi\) of the neuron:

\[
Y = \varphi(a)
\]  

(2)

Where the output function of the neuron \(\varphi(a)\) is a threshold function.
In other words, a neuron model simply receives a vector input $\vec{X}$ that has the components $x_i$. These inputs are multiplied by the proper weights $w_j$ and accumulated sum also known as the weighted sum, $\sum_j w_j x_j$. The scalar output $Y$ is defined by the non-linear mapping,

$$Y = \varphi\left(\sum_j w_j x_j\right)$$

(3)

Notation for NN Model

$W = \text{Weight of connection}$

$W_{k1} = \text{Weight of connection from neuron 1 to } k$

$W_{k2} = \text{Weight of connection from neuron 2 to } k$

$W_{kn} = \text{weight Connection from neuron n to } k$

$X_1 = \text{Input from 1 to neuron n}$

$X_2 = \text{Input from 2 to neuron n}$

$X_n = \text{Input from n to neuron n}$

$j = \text{inputs number identifying index}$

$n = \text{Index identifying destinations}$

$k = \text{processing element identifier}$

$Y_k = \text{output}$

$\varphi = \text{activation function}$

**Materials and Methods**

Data were collected manually and with the aid of video recording along Oba Adesida road, Oyemekun road and Oke Ijebu road all in Akure, (Figure 2) Ondo State, Nigeria. The duration of data collection spans from November 2018 to April 2019. Peak period data were gotten from 7.00 am - 9.00am (morning peak), and 3.00 pm - 5.00pm (evening peak) while data were collected around 11 am – 1 pm (Afternoon session) this is not regarded as a peak period. Data collection procedures such as field observation data logging and video recording extraction method were deployed.

**Video Recording:** The selected point along Oba-Adesida corridor (Cathedral junction) was part of the points in which traffic flow data were collected for a period of 6 months to be able to get a fair measure of traffic movement along the point. Video camera were placed on a tripod stand close to point of studied this was then proceeded by manual mining of the flow based on category of the automobiles. Extraction was achieved off site with the camera played severally to achieve this.

**Field Observation:** Field observers were trained and paid to collect data for the selected corridor (Cathedral junction) for a period of six (6) months, Two (2) observers per lane, one person collected entry traffic flow while another was accountable for booking and inputting the data. Manually, the trained observers prudently count and input data of the traffic flow for motorcycles, cars, buses and articulated vehicles (Trucks) in a well-articulated data sheet designed for that purpose, the recorded data were collected real time.

![Figure 1: The neuron Mathematical model.](image)
The categories of means of movement were motorcycles (M), cars (C), buses (B) and trucks (T) as can be found in Table 1. The table shows the statistical characteristics such as the mean, maximum (Max), minimum (Min) and standard deviation (Sd) calculated from the traffic data. The traffic data were collected along Cathedral, Ilesha garage, oyemekun road, First bank junction, Oke-Ijebu, A-Division, Fiwasaye, Ijapo junction and Ijumu junction respectively.

Table 1: Statistical characteristic of the traffic data.

<table>
<thead>
<tr>
<th>Month/Location</th>
<th>Cathedral</th>
<th>Ilesha Ga</th>
<th>Oyemekun</th>
</tr>
</thead>
<tbody>
<tr>
<td>November</td>
<td>Min 1016</td>
<td>1006</td>
<td>1107</td>
</tr>
<tr>
<td></td>
<td>Max 1357</td>
<td>1464</td>
<td>1532</td>
</tr>
<tr>
<td></td>
<td>Mean 1172</td>
<td>1225</td>
<td>1307</td>
</tr>
<tr>
<td></td>
<td>Sd 143</td>
<td>191</td>
<td>177</td>
</tr>
<tr>
<td>December</td>
<td>Min 1067</td>
<td>1126</td>
<td>1235</td>
</tr>
<tr>
<td></td>
<td>Max 1425</td>
<td>1532</td>
<td>1602</td>
</tr>
<tr>
<td></td>
<td>Mean 1230</td>
<td>1308</td>
<td>1393</td>
</tr>
<tr>
<td></td>
<td>Sd 150</td>
<td>173</td>
<td>158</td>
</tr>
<tr>
<td>January</td>
<td>Min 1203</td>
<td>1109</td>
<td>1071</td>
</tr>
<tr>
<td></td>
<td>Max 1451</td>
<td>1469</td>
<td>1419</td>
</tr>
<tr>
<td></td>
<td>Mean 1329</td>
<td>1291</td>
<td>1245</td>
</tr>
<tr>
<td></td>
<td>Sd 131</td>
<td>173</td>
<td>164</td>
</tr>
<tr>
<td>February</td>
<td>Min 1176</td>
<td>1094</td>
<td>1064</td>
</tr>
<tr>
<td></td>
<td>Max 1489</td>
<td>1509</td>
<td>1458</td>
</tr>
<tr>
<td></td>
<td>Mean 1337</td>
<td>1304</td>
<td>1256</td>
</tr>
<tr>
<td></td>
<td>Sd 131</td>
<td>173</td>
<td>164</td>
</tr>
<tr>
<td>March</td>
<td>Min 1319</td>
<td>1173</td>
<td>1119</td>
</tr>
<tr>
<td></td>
<td>Max 1478</td>
<td>1511</td>
<td>1452</td>
</tr>
<tr>
<td></td>
<td>Mean 1383</td>
<td>1328</td>
<td>1271</td>
</tr>
<tr>
<td></td>
<td>Sd 69</td>
<td>142</td>
<td>139</td>
</tr>
<tr>
<td>April</td>
<td>Min 1307</td>
<td>1166</td>
<td>1114</td>
</tr>
<tr>
<td></td>
<td>Max 1632</td>
<td>1585</td>
<td>1503</td>
</tr>
<tr>
<td></td>
<td>Mean 1448</td>
<td>1357</td>
<td>1289</td>
</tr>
<tr>
<td></td>
<td>Sd 140</td>
<td>177</td>
<td>164</td>
</tr>
</tbody>
</table>
The collected traffic data were preprocessed and fed into the network for prediction. The network was allowed to randomize and divide the data into training, validation and testing. This was to enable the algorithms to be evaluated have the same network conditions.

**The Scaled conjugate gradient (SCG):** This algorithm make use of the step size rather than the line search method in error estimation [24]. The algorithm which minimize error function $E(w)$. Can be as represented below;

The initial weight vector $w_I$ is chosen and set $k = 1$.

The search direction $p_k$ is determined and a step size $\alpha_k$ such that,

$$ E(w_k + \alpha_k p_k) < E(w_k). $$

(4)

The update vector:

$$ w_{k+1} = w_k + \alpha_k p_k. $$

(5)

Now, if $E(w_k \neq 0)$ then set $k = k + 1$ and go to 2 else return $w_{k+1}$ as the required minimum. If $p_k$ in the pseudo algorithm above is negative, that is, $-E'(w)$ and $\alpha_k$ is set to a constant $\epsilon$, this process becomes the gradient descent algorithm [25]. The scaled conjugate gradient algorithm is represented in equations (6) and (7).

$$ s_k = E''(w_k)p_k $$

(6)

The scaled conjugate gradient algorithm is basically another approach in estimating the step size in a function. The aim here, is to estimate $s_k$ in conjugate gradient with an approximation that is non symmetric in nature. Hence, equation (6) approximates to equation (7)

$$ s_k \approx \frac{E'(w_k + \partial_k p_k) - E'(w_k)}{\partial_k}, \quad 0 < \partial_k \ll 1 $$

(7)

Where $\partial_k$ is the step size in conjugate gradient.

The network was trained using the preprocessed traffic data which serves as the input data. Since scaled conjugate gradient is a supervised neural network, targets were set based on the traffic data collected. Both the pre-processed and target data were fed into the network. The network consists of a two-layer feed forward network with ten hidden weight neurons using trainscg as the training function. The training stopped when the validation check was completed in 175 number of iterations at a gradient of 0.072. Figure 2 shows the error versus output of the scaled conjugate gradient algorithm.

![Figure 2: Network output of SCG algorithm.](image-url)
**The Bayesian regularization (BR):** this algorithm is used to modify the mean sum of square network error ($E_d$) in order to improve the network generalization ability. This algorithm is suitable for overcoming the problem of over-fitting.

The function objective is:

$$F = E_d = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2$$

(8)

Expanding equation (8), by adding the term $E_w$ being the sum of squares of weights in the network, equation (8) becomes,

$$F = \beta E_d + \alpha E_w$$

(9)

Where $\alpha$ and $\beta$ are the parameters to optimized by Bayesian regularization [26].

The network consists of a two-layer feed forward network with ten hidden weight neurons using trainbr as the training function. The training stopped when the minimum gradient was reached at 310 number of iterations. Figure 3 shows the error versus output of the Bayesian regularization algorithm.

---

**Levenberg-Marquardt (LM):** this algorithm is an iterative technique used primarily in the least squares curve fitting problem. It expresses the minimum of multi-function as the sum of squares of real value non-linear functions [27].

In a set of $n$ empirical datum pairs of dependent and independent variables, $(x_i, y_i)$, in optimizing the parameter $\beta$ of the model curve $(x, \beta)$ so that the sum of the squares of the deviations becomes minimal.

$$S(\beta) = \sum_{i=1}^{n} [x_i - f(y_i, \beta)]^2$$

(10)

The network consists of a two-layer feedforward network with ten hidden weight neurons using trainlm as the training function. The training stopped when the validation check was completed in 15 number of iterations at a gradient of 0.00091. Figure 4 shows the error versus output of the Levenberg-Marquardt algorithm.
Results and Discussion

Three neural network algorithms were used to trained traffic flow data in a network. These algorithms were trained and the network stopped training at different conditions. The scaled conjugate gradient network stopped training when the validation checks were complete. As it is with SCG algorithm, the network require less memory and the system stops when the generalization does not improve due to increase in the mean square error of the samples being validated. The Bayesian regularization algorithm stopped training when the minimum gradient was reached. This algorithm took more time in training the network and when the adaptive weights were minimized, training stopped. Levenberg-Marquardt took more memory but trained in a short time. The training stopped when the validation checks were reached due to increase in the mean square error of the validation samples. From the analysis, it could be seen in table 2 that the Bayesian regularization algorithm (BR), better predict the traffic flow data. It has the less MSE of 2.37e-13 meaning that the average square difference between outputs and targets does not have much significance. It also shows that it has the highest coefficient of regression (R) approximately 1 implying that there is a close relationship between the output and the target data.

Table 2: Network outputs of the different algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MSE</th>
<th>R</th>
<th>Best performance</th>
<th>Epoch</th>
<th>gradient</th>
<th>Training time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCG</td>
<td>3.76e-2</td>
<td>0.99660</td>
<td>0.107</td>
<td>175</td>
<td>0.072</td>
<td>3</td>
</tr>
<tr>
<td>BR</td>
<td>2.37e-13</td>
<td>0.999999</td>
<td>2.006</td>
<td>310</td>
<td>8.19e-9</td>
<td>313</td>
</tr>
<tr>
<td>LM</td>
<td>3.67e-4</td>
<td>0.99966</td>
<td>0.165</td>
<td>15</td>
<td>0.1.3e-4</td>
<td>10</td>
</tr>
</tbody>
</table>

Likewise, figure 5 shows the trend over time of the predicted data for the mixed traffic flow of the corridors studied. The predicted output for LM, BR and SCG for the month of November 2018 to April 2019 is also shown. From the predicted traffic flow data, we see that Bayesian regularization algorithm predicted the traffic flow better all through the six (6) months that the traffic data were collected and the predicted projection added. Figure 5, shows clearly that Bayesian regularization algorithm when used for the prediction of traffic flow will give a more accurate result.
Conclusion

This research aimed at analysing the performance of traffic congestion using some training algorithms in neural network on traffic flow data. Three neural network training algorithms were examined using the RMS and R parameters to evaluate which of the algorithms perform best given the same network conditions. Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG) algorithms were considered using heterogeneous traffic data collected in Akure, Ondo State, Nigeria. From the analysis, BR algorithm gave the best result with a better RMS and R values. However, the limitation of this work is the unavailability of automated data using automated traffic count equipment which could have improved data gathering. The future of traffic flow congestion analysis and prediction lies with the ability to simulate traffic congestion and neural network for better intelligent transport system.

Conflict of Interest

The authors declares that, there is no conflict of interest.

Acknowledgement

The authors acknowledged the field observers in persons of;
Michael,
Toba Babatope-Ojo,
Tolu Odekunle,
Kolade Awoyelu,
Tolu Ajibade,
Abdulahi Abdulfatai,
Abdulsalam Azeez.O,
Ajayi Taiye Samuel,
Lawrence Faith Morenikeji
Oguntimehin Oluwaseun Funmi
Who are undergoing their undergraduate study at the Federal University of Technology, Akure, Ondo State, Nigeria.
References


