International Journal of
Scientific Engineering and Technology Research

ISSN 2319-8885
Vol.03,Issue.16
July-2014,
Pages:3284-3288

www.semargroup.org,
www.ijsetr.com

# A Roadmap on Design Models for Dynamic Traffic Prediction using Multi-Layer Perceptron Networks

PROF.SWARNALATHA.K.S[1], PROF.G.N.SRINIVASAN[2], KIRAN.K[3]

[1]Asst Prof, Dept of CSE, RVCE, Bangalore, Karnataka, India, E-mail: swarnalathaks@rvce.edu.in.
[2]Prof, Dept of CSE, RVCE, Bangalore, Karnataka, India, E-mail: srinivasangn@rvce.edu.in.
[3]Dept of CSE, RVCE, Bangalore, Karnataka, India, E-mail: kirankannarkk.lris@gmail.com.

**Abstract:** Traffic prediction control system must anticipate traffic situations and adjust their strategy on time. Various factors like road demographics, junction designs, potholes, debris, crossings, U-turns influence the short time prediction of the traffic. This paper describes the various design methodologies employed while using multi-layer perceptron in dynamic traffic prediction. The methods described utilize various data collection methods and employ data pre-processing before sending the collected real-time data as input to the neural network. The effectiveness of the various design methodologies is studied.

**Keywords:** Dynamic Traffic Prediction, Artificial Neural Networks, Multi-Layer Perceptron, Data Collection Metrics, Prediction Model, Traffic Flow.

## I. INTRODUCTION

The automobile industry is booming; vehicles are becoming more affordable than ever to everyone. Today, traffic jams have become a daily affair in the metropolitan cities like Bangalore. Traffic is going to increase with time. Hence the need to integrate urban planning with transport arises, i.e. Transit Oriented Development (TOD). Traffic prediction is required both for preventive measures as well as to serve as information for TOD. Accurate traffic reports are essential for congested and overcrowded cities. Traffic prediction is a key to avoid congestion and preventing a breakdown of normal vehicular movement. The main problem is that the traffic is dynamic in nature. There are various factors that affect the traffic flow and speed, for instance, demographics, time, speed, lane width and length. In civil engineering parlance, every road has a unique identifier that is formed by a combination of features that span the above factors.

The amount of parking space, the locality of the road and the type of the road – cross-section/ intersection, tunnels, motorways, highways, subways, multi-lanes or single –lanes are a subset of other features that compound the problem of traffic prediction by emanating different classes of congestions, jams or accidents in different roads [3]. Traffic prediction is taking into account, data from various points to build a holistic view of the traffic in the region specified. It must not be mistaken that traffic prediction is necessary only in urban localities. Traffic prediction is a universal problem. It applies to urban and rural roads alike. Roads are connecting urban to suburbs and rural places, diminishing the factor of distance and time to connect to modernization and development. Hence traffic chaos can easily propagate to rural roads. To predict effectively the traffic in the next intersection, or provide a rerouting mechanism, with congestion controls- both short term and long terms, artificial intelligence (AI) is being employed. AI is seen as the solution to the limitation of processing power of humans. Machine learning has become the term of this decade, changing the meaning of data mining substantially. It has also redefined the limits of what science can do to solve existing problems. Prediction is a part of AI, and data mining technologies like association analysis, decision trees, and artificial neural networks are employed.

## II. MULTI-LAYER PERCEPTRON NETWORKS

An artificial neural network (ANN) is an interconnected collection of artificial neurons that uses mathematic or computational models that can be trained to simulate thinking or learning by experience, which includes knowledge acquisition, recall, synthesis and problem solving as shown in Fig.1. There is an input layer and an output layer, and connecting them are a series of hidden layers which could be just one hidden layer or may be thousands of layers. Each layer can have variable number of neurons, simulating the actual neurons of the brain. Certain weights connect a neuron of one layer to a neuron of another layer. Neurons within a layer do not communicate with each other. A neuron computes an output using an activation function that considers the weighted sum of all its inputs. Inputs given are forwarded to successive layers, each neuron having their activation functions [$A_f$] which introduces the necessary non-linearity to the network. $A_f$ is commonly the sigmoid function, but it need not be restricted to it. The sigmoid function f is defined as:

$$f(x) = \frac{1}{1+e^{-x}} \quad \textbf{(1)}$$

Where f(x) is the output of the neuron and x is the weighted sum of all its input.

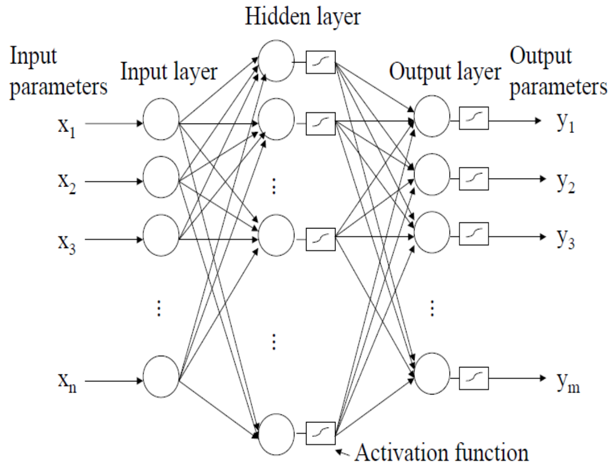


**Fig.1. Representation of an ANN.**

MLPs are simple feed forward neural networks. Except for the input nodes, all other nodes are neurons with activation functions which are non-linear. MLPs employ technique of Back propagation to train the neurons accordingly. This involves feeding input to network and calculating the error rates between expected and actual outputs and going back to adjust weights of each layer of MLP. This process is iterated until convergence is reached, or minimum training gradient is attained (or expected output is gained). The above method of trial/error uses the concept of gradient descent where the iterations are stopped when the errors reach the acceptable minima.

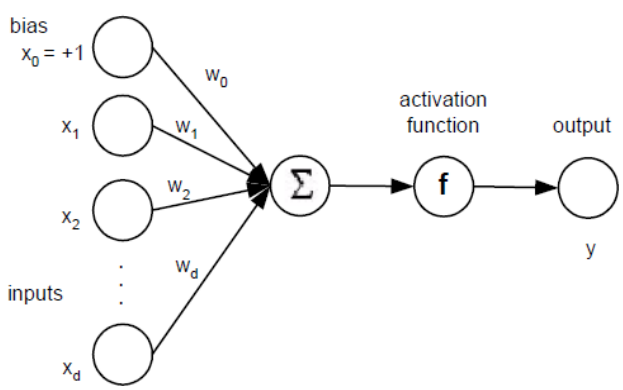


**Fig.2. A single perceptron.**

Fig.2 represents and simply the weighted sum of inputs in most cases:

$$y = \sum_{j=1}^{d} w_j x_j + w_0 \quad \textbf{(2)}$$

Where the weights are shown above the connecting line between neurons of different layers and f is the activation function. This value is fed as x into the sigmoid function or any other activation function.

The architecture of the MLPs depend upon the number of hidden layers in the network, the number of hidden neurons in the network, the connectivity of the input, output and the hidden layers (along with possible bias), the learning rate (this defines the amount of trial/error for correct prediction) and the activation function of each neuron.

## III. CONCEPT OF TRANSPORT FORECASTING

Transport Forecasting is defined as a process of estimating number of vehicles, at a given epoch (time), their speed, and other metrics, that could be used to reroute/predict traffic as an estimate for future costs and calculate capacity of infrastructure [7]. Transport forecasting systems today employ Dynamic Traffic Allocation model that involves four key steps:

- **Trip Generation:** This involves the study of the demographics, land use, residences, social and economic attributes of households surrounding the road network
- **Trip Distribution:** This involves developing "trip table", a matrix that displays the number of trips going from each origin to each destination within the road network.
- **Mode Choice:** This involved predicting different modes of transportation used by different proportion of trips.
- **Route Assignment:** The final step of DTA involves Trip Allocation by a particular mode – a prediction made for the particular trip.

Traffic flow (rate), speed, and density are three basic parameters that describe traffic conditions. Traffic flow can be defined as the total number of vehicles that pass over a given point or section of a lane or roadway during a given time interval. The number can be expressed hourly, daily or in a specific time interval basis. Flow is usually limited by road capacity. Speed is the rate at which a vehicle passes a given point. The speed is limited by number of vehicles occupying the road at the given point. Heterogeneity can also be included as a limiting factor to speed. Density is the number of vehicles observed and measured over a certain road segment over a period of time and not just a given point. Traffic state is influenced by metrics both in time and spatial domain [5]. Static indices can be area type, lane width, number of lanes, parking conditions etc., while the dynamic indices are volume by movement, % heavy vehicles, heterogeneity, operating speed, peak hour factor etc [3]. The roads may be intersection/cross-section, tunnels, straight segments and motorways. There are two problems that are commonly faced while solving traffic congestion problems and traffic prediction.

Wardrop's principle of User Equilibrium explains that each driver takes the shortest path, subject to every other driver who is doing the same. The other problem faced while trying to resolve traffic related problems is the duality existing in nature. The travel time is a function of demand and the demand is a function of travel time, forming a bilevel problem. Traffic flow is stochastic in nature as well as traffic prediction characteristics are non-linear. This works well with neural networks. When working with neural networks, no prior assumptions regarding the problem required and they provide automatic classification. Neural networks are also able to adapt to moderately noisy data. They can scale well with dataset size, which may vary depending on the road network in question.

## IV. DESIGN METHODOLOGY

The design process for any neural networks models may vary depending on the usage of the model and the desired efficiency levels. This section outlines the design paths taken by various MLP models.

### A. Data Collection and Extraction

For any design model, the first step is data collection and extraction. Traffic data that is to be fed into the prediction model can be both historical data and real time data. Real time data is mostly stream video processing data captured by the traffic camera systems deployed at cross-sections or regular identified link points in the grid [1]. The effect of length of links and location of measurement stations and the structure of the entire measurement system are to be taken into account. Data can be collected by setting up 5 minute interval points (or 15 minutes) where cameras are deployed [1]. This could be done at all intersection points, collecting data relating to the indices mentioned above. The objective is to predict the traffic flow status ahead through dynamic indices. It should also take into account the relationships between dynamic and static indices, and also sequential patterns in the data of the whole region [3]. [4] collects data for every minute, and later is aggregated to one hour/ one day depending on the amount of dimensionality reduction required.

This data was used to build a congestion predicting MLP. The metrics were – Day of the Week, Minute of the Day**,** Speed, Volume (Traffic Flow) and target Congestion Level as output. Data preprocessing is necessary because the amount of data collected is huge, and feeding all data into ANN may not be feasible and performance enhancing. Though ANN can moderately adapt to noise, feeding irrelevant data makes no sense in terms of cost and viability. Possible scaling of data to [-1,1] may be necessary to maintain mean of 0 and standard deviation of 1 which is considered to be standard distribution of data. This standardization is useful for the ANN's learning [2]. Normalization/ Dimensionality Reduction is a must considering the number of features that accompany data that spans even 5 minutes. [7] uses sensitivity analysis about the mean to reduce the number of input parameters to be fed into the ANN. were classified in eight categories namely Car/Jeep/Van, Scooter/ Motorcycle, Light Commercial Vehicle, Bus, Truck, 3-wheeler, Tractor and Carts/ other animal drawn vehicles.

Nine most significant inputs parameters were selected, which are speed of Scooter/ Motorcycle, speed of Tractor, number of Scooter/Motorcycle, speed of Carts, traffic density, time, number of Tractor, number of Trucks and speed of Truck. The reduction of 19 parameters to 9 most significant parameters was cost-reducing and also effective in terms of better learning of the ANN. Hence Kranti Kumar et al. in [7] include heterogeneity of the traffic also into the network. Transformation to a different dimensional space is also possible [3].For example, average delay can be a value from 0- 100. If the 10 ranges ([0,9], [10-19]……) can be mapped onto a respective Level of Service (A,B,C,…..), only ten binary values (in terms of range) can be stored instead of 101 binary values.

### B. Data Classification / Design Models

Different approaches and sub-approaches have been taken to understand performance of MLPs under different conditions. Now that we have selected the optimum number of input parameters, and we also know the number of output parameters, the remaining aspect is deciding on the number of hidden layers and the respective number of hidden neurons. For simplification purposes, a single hidden layer is opted, making the MLP a (N)*(1)*(M) neural network, where N is the number of input parameters and M is the number of output parameters. [2] chooses a road of 3 km stretch with 7 cross sections, where the 5th cross-section is where we want to predict the traffic (considering the traffic before and after the target cross-section) Data collected every five minutes in 3 spans. This data is aggregated separately as 15 minute input data as well which consists of mean speed and flow information. Calculating $N_h$ is the first task. To calculate $N_h$, the number of input and output parameters should also be decided. Considering $N_i$, $N_h$, and $N_o$ are number of neurons in input, hidden and output layers respectively, W as no. of parameters (weights and biases) and Tmin as Number of training samples required for the MLP to learn and predict accurately, the maximum number of hidden layer neurons possible is given by the formula for Nhmax:

$$T_{\min} = 10 \cdot W; \quad W = (N_i + 1)N_h + (N_h + 1)N_o$$
$$\Rightarrow \quad N_{h_{\max}} = \frac{0{,}10 \cdot T - N_o}{N_i + N_o + 1} \tag{3}$$

Levenberg –Marquardt Algorithm or also called as damped least squares method is used to solve non-linear least square problem. The RMSEs calculated during back propagation can be used to change the weights using the above algorithm. The idea of the Levenberg-Marquardt algorithm is that the training method shifts from the gradient method towards Newton's method making use of the advantages of both methods. The benefit that is gained by reducing the input parameters is the possibility to have more hidden neurons in the network with the same training set. This is why most approaches prefer to reduce the input space.

The following design studies were made:

**Design Study 1:** Divide into 2 models – flow model and speed model.
**Design Study 2:** Divide into 3 models – each for the 5-min interval data.
**Design Study 3:** Divide into 2 models – one with normalized data, the other is scaled to [-1,1].

**Design Study4:** Study as one model where data is aggregated to a 15 minute interval.

To make the basic model for the prediction, or the models based on both the scaled data and the normalized data, the best combination of functions was a hyperbolic tangent for the hidden layer and a linear function for the output layer [2]. The activation functions need not be sigmoid function and need not be a single non-linear function.

### C. A Compound Back Propagation Model

Bing Wu; et al. in [3] proposed a compound neural network model which uses other data mining technologies along with the neural network for dynamic traffic prediction as shown in Fig.3.

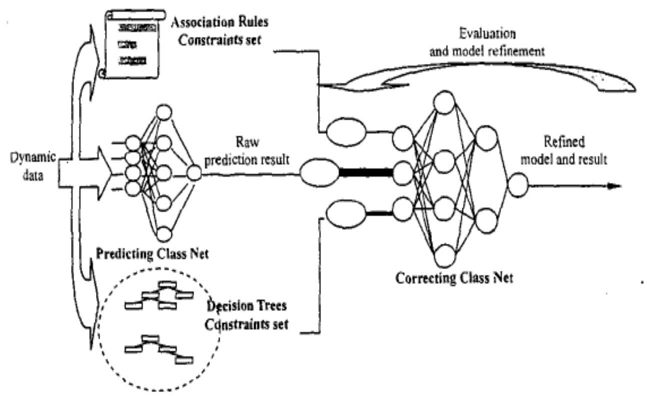


**Fig.3. Compound Back propagation model schema.**

In certain periods of time, changes in parking conditions and bus operation can affect the metrics near the collecting point. To analyze this type of influences, decision trees are used. Neural networks cannot manage such data without manual intervention, while decision trees provide decision paths to process such data. Association analysis results in useful knowledge about relationships between parameters which may be also used for transformation to other dimensional spaces. A correcting class network is used for refinement purposes.

## V. RESULTS ANALYSIS

From the errors calculated for the four design studies, it can be observed that the flow model based on normalized data gave smaller root mean square errors (RMSEs) indicating an even distribution of values. But flow model based on scaled data gave better results when considering the different absolute values of errors which gave more realistic forecasts [2]. When the number of hidden neurons is very small, every new neuron is crucial in improving the results, as quoted by [2]. 99 percent of the predicted speeds were within the range of ten-percent relative error for one month data. The number of hidden layer neurons, as discussed earlier can be increased by reducing the number of input parameters. Pitiphoom Posawang, et al., in [4] observed that the RMSEs remain constant after a threshold value of number of hidden nodes. Hence, an overly increase of the nodes will not increase performance substantially. It was also found out that the optimal learning rate to be 0.3. Learning rate (or 'alpha') indicates the amount of trial/error that may be required to adjust the weights to achieve minima during gradient descent. This accuracy of the MLPs was maintained provided the learning rate is not set to a high value, or the number of neurons is very low. [4]'s congestion prediction model achieved an accuracy of 94.99% in correctly classifying congestion into the three output levels – light, heavy and jam.

The most efficient training algorithm proved to be the Levenberg-Marquardt Back propagation, as it surpassed the others not only in quality of results but also in speed [2]. The number of nodes used in the hidden layer depends on the type and the complexity of the time series, and on the number of inputs. Due to the black-box nature of MLPs, it is not possible to find the mutual interrelation between the variables in MLPs modeling. Hence the compound back propagation model can be employed [3], which includes other data mining technologies with MLPs for traffic prediction. The models that were divided into two sub-models – one for the mean speed forecasts and the other for the traffic flow forecasts – gave better results than one single model predicting both variables simultaneously. In fact aggregating into a fifteen minute model provided better results than handling five minute interval data which fared poorly [2]. For 90 percent of the predicted flows the relative error was 20 percent at most, and for 90 percent of the predicted speeds it was four percent at most.

## VI. CONCLUSION

Neural networks provide a wide variety of freedom to explore for any given problem, considering the number of parameters of the network architecture that need be decided upon. A difference of even 0.1 in the learning rate can significantly alter the results. The dynamic nature of given problem works well with neural networks as they can adapt well to moderately noisy data and need no prior information apart from the training data set to start automatic classification. The MLPs react well to unseen data. It should be mentioned that these prediction models were done with only a small amount of data which span up to a maximum of three months period of observation. These short-term prediction models may enhance may not apply to long-term TOD. The performance also depends upon various factors like accidents, debris, skewed junctions, U-turns and driver behavior. Adverse weather conditions or festival seasons/other events like political events can also compound the traffic woes. This data cannot be fed into the neural network directly. Since data considered in this study consists only off-peak hour's traffic which does not represent total variability in the traffic flow.

A compound neural network partly solves the problem by taking this variability load into other data mining technologies like decision trees and association analysis, but many other factors would be left out in spite of this inclusion. When considering real-time short-term prediction systems, the effectiveness depends mostly on accurate single step ahead predictions of traffic flow. SP (Speed) was observed to be a deciding factor in congestion perception and hence the AI learning technique depends heavily on SP. Based on the results, it was better to increase the number of hidden neurons by reducing the input parameters by decreasing the number of cross-sections rather than by shortening the time-series, i.e. dimensionality reduction is employed. Most designs hover over a single hidden layer, varying over the number of hidden neurons, number of input and output parameters for different desired level of efficiencies. Hence it can be concluded that, many other transportation data prediction studies can be implemented easily and successfully by using Multi-layer perceptrons.

## VII. REFERENCES

[1] Satu Innamaa, "Short Term Prediction of Highway travel time using MLP –Neural Networks", 8th World Congress on Intelligent Transportation Systems, Sydney, Australia, 30 Sept.– 4 Oct. 2001.

[2] Satu Innamaa, "Short Term Prediction of Traffic Situation using MLP- Neural Networks", Finnish National Research and Development Program on Transport Telematics Infrastructure (TETRA) 1998 – 2000.

[3] Bing Wu; Wen-Jun Zhou; Wei-Dong Zhang, "The applications of data mining technologies in dynamic traffic prediction," Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE , vol.1, no., pp.396,401 vol.1, 12-15 Oct. 2003.

[4] Pitiphoom Posawang; et al., "Perception-based Road Traffic Congestion Classification using Neural Networks", Proceedings of the World Congress on Engineering 2009, Vol I WCE 2009, July 1 - 3, 2009, London, U.K.

[5] Bowu Zhang; Kai Xing; Xiuzhen Cheng; Liusheng Huang; Rongfang Bie, "Traffic clustering and online traffic prediction in vehicle networks: A social influence perspective," INFOCOM, 2012 Proceedings IEEE , vol., no., pp.495,503, 25-30 March 2012.

[6] Kranti Kumar, M. Parida, V.K. Katiyar, "Short Term Traffic Flow Prediction for a Non Urban Highway Using Artificial Neural Network", 2nd Conference of Transportation Research Group of India, Procedia - Social and Behavioral Sciences, Volume 104, 2 December 2013, Pages 755-764, ISSN 1877-0428.

[7] Transport Forecasting–http://en.wikipedia.org/wiki/Transportation_forecasting