Extreme Learning Machine Application in Bus Passenger Flow Volume Prediction

Junqin Qi

School of Economics & Management, Xidian University, Xi'an 710126, China

qijunqin@163.com

Abstract

Extreme learning machine（ELM） is a new learning algorithm of single-hidden layer feed-forward neural network（SLFNs）, and overcomes the disadvantages of the classical learning algorithms in huge search space and a large number of calculation. What's more, ELM algorithm has favorable general capability with faster learning speed. It is a fact that bus passenger flow presents strong non-linearity and non-stationary. In this paper, ELM is introduced to predict the bus passenger flow volume. Finally, based on the actual traffic data of a bus line in Xi'an and by comparing with BP and RBF, the effectiveness and advantage of ELM prediction model is tested.

Keywords

Extreme learning machine（ELM）, neural network, bus passenger flow volume, prediction.

1. Introduction

With the rapid development and deep application of Internet of Things and other new information technology, smart city has become an inevitable trend of future development of cities [1]. Public transportation is an important component of smart city. With its large capacity, low cost and high environmental benefits, it is becoming the focus of the smart city construction. Bus passenger flow volume is the basic data of city public transport system scheduling, planning and development. It is of great significance for the construction of smart city to analyze and grasp its characteristics and rules, and predict it scientifically.

Bus traffic forecast is one of the focus areas of public transportation. In the long-term forecast, Wenyuan Xu[2] established prais-winsten AR (1) autoregressive time series model through the analysis of the bus traffic influencing factors, and predicted the total passenger bus in the next two years in Harbin. Yugang Wang et al [3] used the gray correlation method to obtain the correlation between the degree of influence factors and passenger flow volume, and utilized the gray prediction model to predict regular bus volume in Chongqing. As for the short-term forecast, Liu Kai et al. [4] proved that short bus passenger flow sequence has chaotic characteristics, and they proposed combination forecasting method based on the characteristic analysis. Chunhui Zhang et al [5] applied Kalman filter prediction method to forecast the passenger flow of bus station, and verified the advantages of the model by comparison with BP neural network model. Zhao et al. [6] proposed combination forecasting model based on wavelet analysis and neural network aiming at the non-linearity of bus passenger flow sequence. The target of long-term forecast is usually the passenger volume in next year or several years, while the short-term transit traffic prediction method is more focused on short-term traffic forecasts, typically between 5 ~ 15min. In fact, the bus daily passenger flow forecast is assist with scientific, intelligent and accurate bus scheduling, planning and development for the bus company. However, changes in the bus daily passenger flow present strong volatility, which makes it difficult to predict daily passenger flow due to the impact of weather, holidays and other factors. Therefore, how to predict the bus daily passenger flow has important practical and theoretical significance.

As an effective intelligent information processing technology, neural network has a strong non-linear approach capability and self-learning capability, and has been widely used in the short-term prediction of the bus passenger flow volume [5,6]. Huang [7] put forward an effective single hidden layer network learning algorithm, named extreme learning machine (ELM) , in 2006. Compared with the general BP and RBF neural network, this method has faster learning speed and favorable general ability. Therefore, this paper attempts to apply ELM to the bus passenger flow volume forecast, verify its superiority and feasibility by the comparison with BP and RBF neural network, in order to provide a new method for forecast.

2. ELM neural network

ELM neural network is proposed by Huang at. el, which is a new type of SLFNs(single-hidden layer feed forward neural networks) learning algorithm and performs well. The algorithm needn't adjust all the input parameters, selects the input layer weight and hidden layer bias (threshold) randomly, and obtains the hidden layer output weights by the least squares method [7]. In ELM, only the hidden layer neuron nodes are needed to learn to adjust, meanwhile, the entire process is without iterative operation, leading to extremely fast speed. The special learning mechanisms of ELM can effectively overcome the shortcomings of the difficulty in parameter setting and long training time in the traditional machine learning algorithm, and have the advantage of good global search capability and simplicity, which has been proved in the regression analysis, classification, prediction and other fields with a fast learning speed.

Figure 1 shows the typical structure of ELM. Assuming that there are *p* input layer neurons, *L* hidden layer neurons, the activation function between the input layer and the hidden layer is *f*, the training set is $\{\boldsymbol{X_i},\ y_i\}_{i=1}^{N}$, where $\boldsymbol{X_i}=\left[x_i^1,\ x_i^2,...,x_i^p\right]^T \in \boldsymbol{R}^p$, and *N* is the total number of samples. The mathematical model between the input and output layer can be formulated as

$$\sum_{j=1}^{L} f\left(W_j, b_j, \boldsymbol{X_i}\right)\beta_j = y_i \quad i=1,2,...,N \tag{1}$$

Where W_j is the weight vector between the *j*th hidden layer neuron node and the input vector X_i, b_j is the threshold of the *j*th hidden layer neuron node, and β_j is the weight between the hidden layer neurons and output nodes.

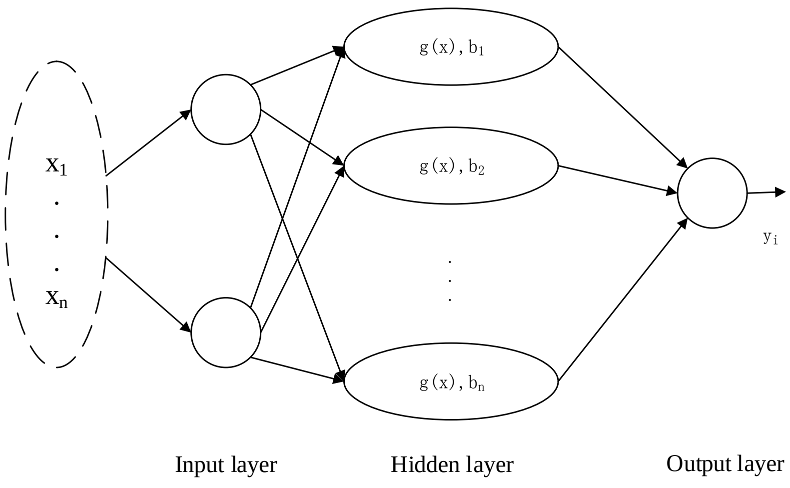

Fig.1 Typical structure of ELM

The equation (1) can be written in the matrix form as follows.

$$\boldsymbol{H\beta}=\boldsymbol{Y} \tag{2}$$

Where

$\boldsymbol{\beta} = \left[\beta_1 \beta_2 \cdots \beta_L\right]^T$

$\boldsymbol{Y} = \left[y_1\ y_2 \cdots y_N\right]^T$

$$\boldsymbol{H} = \begin{bmatrix} f\left(W_1, b_1, \boldsymbol{X_1}\right) \cdots f\left(W_L, b_L, \boldsymbol{X_1}\right) \\ \vdots \quad \cdots \quad \vdots \\ f\left(W_1, b_1, \boldsymbol{X_N}\right) \cdots f\left(W_L, b_L, \boldsymbol{X_N}\right) \end{bmatrix}_{N\times L}$$

So, the parameter $\boldsymbol{\beta}$ can be solved by smallest norm least squares solution through equation (3).

$$\boldsymbol{\beta} = \boldsymbol{H}^{+}\boldsymbol{Y} \tag{3}$$

Where, $\boldsymbol{H}^{+}$ is the Moore-Penrose generalized inverse of the matrix $\boldsymbol{H}$.
In this paper, the activation function is sigmoid function. Assuming that variable is u, the sigmoid function is formulated as

$$f(u) = \frac{1}{1+e^{-u}} \tag{4}$$

3. ELM Prediction of Bus Passenger Flow

The process of ELM prediction of Bus passenger flow is shown in figure 2.

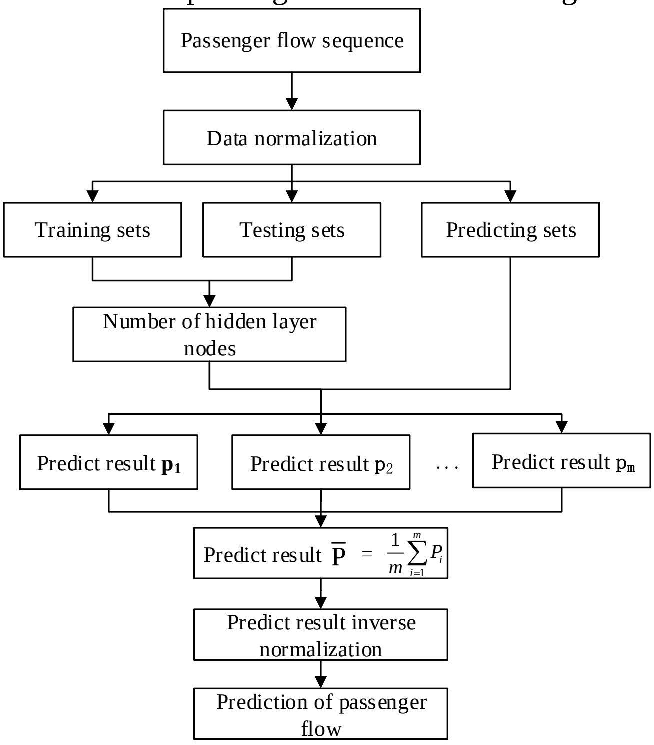

Fig.2 Flow chart of ELM prediction

It should be noted that because of the random parameter setting, ELM often has unstable performance. Repeat ELM forecast, calculate the average of the predicting results, and set the average as the final predicting result, so we can improve its forecasting performance.

4. Instance Analysis

4.1 Data Sources

This article tracks the daily passenger flow volume of a bus line in Xi'an from April 1, 2013 to March 31, 2015, and total number of data is 730, as shown in Figure 3.The first 723 data is regarded as the forecast basis to predict the daily passenger flow volume from March 25, 2015 to March 31, 2015.

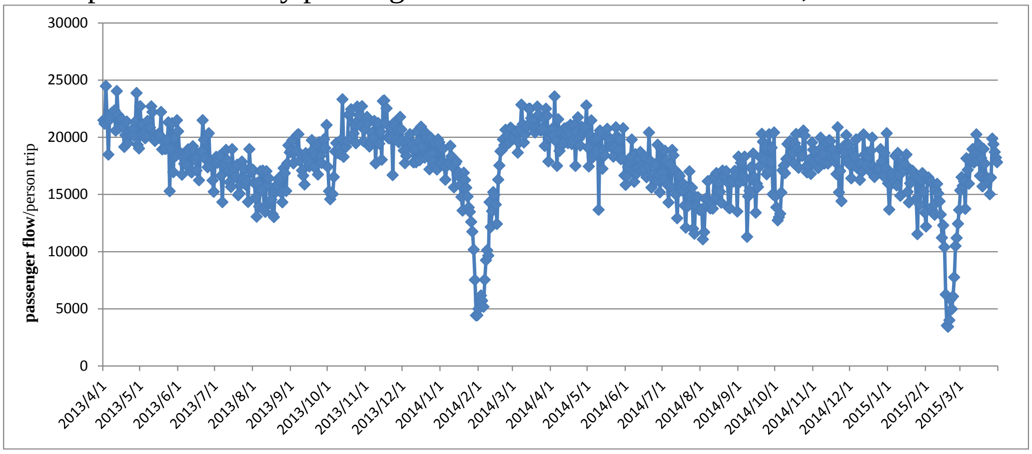

Fig.3 Passenger volume of a bus line in Xi'an from April 1, 2013 to March 31, 2015

4.2 ELM Prediction

Passenger flow volume presents an irregular cycle fluctuations which the period is a week (7 days). Considering the passenger flow volume depends on passenger flow volume a week before, the input nodes number of neural network is set to 7, the input of neural network is the passenger flow volume 7 days ago and the output is passenger volume of forecasting date. Before predicting, we should normalized passenger volume, then use data of the continuous 7 days as the input and the data of eighth day as the output of the network (i.e. target data). In this way, training samples and testing samples of neural network are formed, and a total of 600 training data sets, and 116 test data sets.

ELM neural network is SLFNs, and activation function is sigmoid function. Different hidden layer nodes number has different impact on training and testing sets is shown in figure 4. As shown in figure 4, the hidden layer nodes number of ELM can be set to 30 and the test error would be minimum. Setting the hidden layer nodes number to 30, we can get the corresponding parameter values, and then forecast the passenger flow volume from March 25 to March 31 by using single-step rolling forecasts and setting ELM integrated number to 200.The forecast results are 16976, 16772, 17472, 17214, 16978, 16896, 16759.

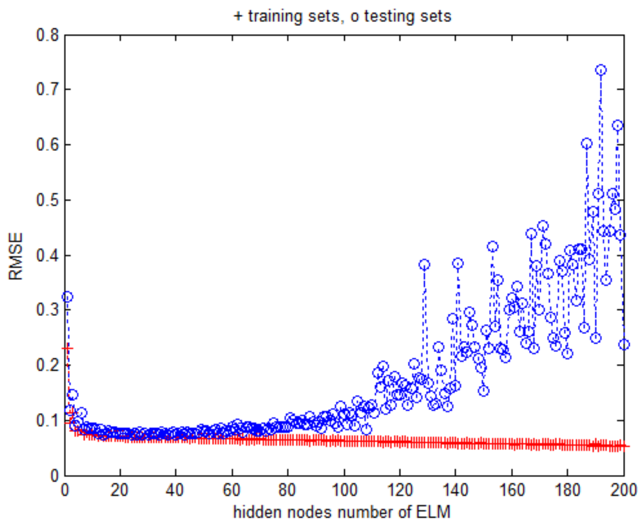

Fig.4 Chart of ELM effect vs. hidden layer neuron number

4.3 Comparison and Analysis

In order to verify the validity of the ELM in the forecast of passenger flow, we compare the results obtained by BP and RBF neural network.

BP neural network forecasting model only contains a single hidden layer. A large number of experiments are conducted to determine the number of hidden layer nodes, and the number of nodes is set to 14 in this study. In addition, the transfer function of the hidden layer neurons is set to logsig, the output layer neuron transfer function is set to purelin, the training function is set to trainlm, and the training times is 2000.The model is used to forecast the passenger flow volume from the March 25 to March 31 in 2015.The forecast results are 17292,17130,17630,17172,17087,16848,16982.

RBF is a three-layer feed forward network with a single hidden layer. In this study, the transfer function of radial basis neuron is set to Gauss function. By a large number of experiments, the expansion coefficient of the radial basis function is set to 600. The model is used to forecast the passenger flow volume from the March 25, to March 31 in 2015.The forecast results are 17052, 16757, 17252, 17113, 16885, 16784, 16709.

Since holidays have great impact on bus daily passenger flow volume, we select the data from April 1 to April 7 in 2015 as example to predict, which the April 5th is the Tomb-sweeping Day, in order to verify the prediction results of different forecasting methods during the holidays.

In order to compare the validity of ELM with BP and RBF neural network, we use mean absolute percentage error（MAPE）, root mean square error（RMSE）and accuracy of trend prediction (ATP)[8] as the evaluation criteria. The MAPE, RMSE and ATP are defined as follows.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i^{'}}{y}\right| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^{'})^2} \tag{6}$$

$$APT = \frac{\sum_{i=2}^{n} T_i}{n-1} \tag{7}$$

Where, the real value is y_i and the predicting value is $y_i^{'}$ （ $i = 1, 2, ..., n$ ）. When $(y_i - y_{i-1})*(y_i^{'} - y_{i-1}^{'}) \geq 0$, then $T_i = 1$.Otherwise $T_i = 0$.

Comparison of the results of the three forecasting methods is shown in Table 1.

Table1 Error analysis of three prediction method prediction

Predicting Method	March 25- March 31			April 1- April 7		
	RMSE	MAPE（%）	ATP	RMSE	MAPE（%）	ATP
BP neural network	1714.9	8.93	0.667	2888.6	16.68	0.167
RBF neural network	1799.2	9.12	0.833	2782.9	15.80	0.500
ELM neural network	1693.6	8.63	0.833	2919.1	16.54	0.500

As can be seen from table 1, whether it is a normal day or holiday, the ELM forecast model is more accurate on the prediction accuracy of the bus daily passenger traffic and the trend of the change of the passenger volume than BP and RBF neural network. Furthermore, in order to verify the efficiency of the ELM neural network in the prediction, we compare and analyze the results with the BP and RBF neural network. The results are shown in Table 2.

Table 2 Comparison of computing time of ELM, BP and RBF (/s)

	ELM neural network	BP neural network	RBF neural network
March 25- March 31	0.034	51.063	1.825
April 1- April 7	0.032	52.366	1.370

As can be seen from table 2, the ELM neural network has the advantages of high accuracy and high speed. ELM neural network is chosen to establish the predicting model of passenger flow volume, which can not only ensure the accuracy of the forecast, but also save the operation time and improve the efficiency of predicting.

5. Conclusions

In this paper, ELM neural network is applied to forecast the bus passenger flow for the first time. The ELM learning method is different from other methods. Under the assumption that the input layer weights and the hidden layer deviation are the random selection, the method can be resolved to obtain the hidden layer output weights .The algorithm is simple and the speed is very fast. Compared with the widely used BP and RBF neural network forecasting method, the method has higher prediction accuracy and faster computing speed. In practical application, the ELM neural network can be used to forecast the bus daily passenger flow volume, and it can provide the basis for the public transportation company to realize scientific scheduling, reduce the cost of transportation and improve the service level.

Acknowledgments

Xi'an science and technology project (SF1404).

References

[1] L. Zhang, Y.W. Yang, M. Zhang: Smart Transportation and Smart City, Science, Vol. 6 (2014) No. 1, P.33-36.

[2] W.Y. Xu, C.Y. Deng, B.Y. Liu, T.C. Qu:The bus passenger traffic time series prediction model, Journal of Liaoning Technical University (Natural Science), Vol. 33 (2014) No. 12, P.1715-1720.

[3] Y.G. Wang, H.Y. Yao, Y.S. Li, X.F. Chen: The influence factors and prediction model of bus passenger volume in Chongqing based on grey theory, Journal of Transport Information and Safety, Vol. 32 (2014) No. 6, P.89-94.

[4] K. Liu, W.Q. Li, J.H. Zhao: Study on wavelet forecast method for short-term passenger flow, Journal of Transportation Engineering and Information, Vol.8 (2010) No.2, P.111-117.

[5] C.H. Zhang, R. Song, Y. Sun: Kalman filter-based short-term passenger flow forecasting on bus stop, Journal of Transportation Systems Engineering and Information Technology, Vol.11 (2011) No. 4, P. 154-159.

[6] S.Z. Zhao, T.H. Ni, Y. Wang, X.T. Gao: A new approach to the prediction of passenger flow in a transit system, Computers and Mathematics with Applications, 2011, No.61, P.1968-1974.

[7] G.B. Huang, Q.Y. Zhu, Chee-Kheong Siew: Extreme learning machine, Theory and applications , Neurocomputing, Vol. 70 (2006)No.1-3, P. 489-501.

[8] J. Bai, J.B. Xia, X.H. Zhao: Prediction model of network traffic based on EMD and RVM, Computer Science, Vol. 42 (2015) No.1, P.122-125.