Traffic demand Forecast of online car-hailing based on BP Neural Network

Junji Jiang[1]

[1]College of Management and Economy, Tianjin University Tianjin, China

***Abstract*—**To accurately predict passengers' demand for ride-hailing, increase transport capacity in some areas directionally, make it easier for passengers to book ride-hailing, and thus enhance passengers' travel experience, based on BP neural network model, combined with 2015-Mak-2019 ride-hailing demand change data, and based on MATLAB platform, the demand trend of 20202024 is forecasted. The results show that the demand for shared cars and rides will increase rapidly. Demand for shared bikes and taxis is rising slowly.

1 Introduction

Traffic is not only the basis of urban survival and development but also an essential guarantee of urban economic activities and people's normal life. With the development of society and cities, people's demand for transportation is also increasing. To meet the growing demand for travel, major cities advocate vigorously developing public transport and improving the public transport system to solve the contradiction between supply and demand for travel. The urban public transport system has gradually developed from a single bus to a comprehensive system composed of urban buses, BRT, taxis, subways, light rail, and other travel modes. As an indispensable part of urban public transport, ride-hailing enriches the urban public transport system and meets the diversified travel needs because of its convenience and flexibility. [1]

In recent years, the research on online car-hailing demand forecasting includes online car-hailing supply and demand matching research and short-term demand research. In the matching of taxi supply and demand, this paper studies the influencing factors of online car-hailing demand, including macro factors such as urban GDP, per capita income and so on, and then uses the model to determine the number of taxi demand in each region of the city. The specific research includes the following: through the research and analysis of the characteristics and main influencing factors of taxi demand in 2017, Lu Yi and others set up a taxi demand forecasting simulation model based on BP neural network by taking five main factors that affect online car-hailing demand, GDP, urban population, residents' consumption level, average waiting time and a total length of bus lines, as inputs. In 2019, Wang Shujia and others selected six factors affecting ride-hailing demand, including urban per capita disposable income, urban population, urban area, number of buses, taxi prices, and whether the city has subway variables as explanatory variables, and establish a linear regression model to forecast taxi demand, and finally adjust the model to screen variables. Finally, the urban per capita disposable income and the number of buses are selected as the input variables of the model. In 2018, Liu Lina and others used Beijing ride-hailing records to extract and analyze the data, statistically analyze the travel rules and characteristics of Beijing ride-hailing from the IC card data, and find out the factors that affect the demand for ride-hailing. In 2019, Filipe et al. proposed two deep learning data fusion architectures based on heterogeneous time series and text data, thereby improving the accuracy of the time series forecast of online ride demand. [2,3,4]

BP neural network has arbitrarily complex pattern classification ability and excellent multi-dimensional function mapping ability, which solves the XOR and some other problems that cannot be solved by simple perceptron. The three-layer neural network can approximate any non-linear continuous function with arbitrary precision. It makes BP neural network particularly suitable for solving complex internal mechanisms. During training, the BP neural network can automatically extract the reasonable rules between the output and the output data through learning, and adaptively memorize the learning content in the weights of the network.

---

e-mail: 13371576712@163.com

TABLE I. 2015 TO 2019 VARIOUS TYPES OF TRANSPORT TRANSACTIONS (100BILLIONYUAN)

Index	*2015*	*2016*	*2017*	*2018*	*2019*
Private car transaction	370.64	902.74	1746.69	2224.96	2116.76
Taxi	119.8	222.12	353.31	386.23	639.9
Hitchhiker	123.82	163.43	191.93	232.14	288.3
Car rental	1.32	3.62	17.29	36.48	66.39
Bike sharing	0.52	0.6	90.64	132.96	153.71

The BP neural network has a high degree of self-learning and self-adaptive capabilities. Structurally, the BP network has an input layer, a hidden layer, and an output layer.

Primarily, the BP algorithm uses the network error square as the objective function and uses the gradient descent method to calculate the minimum value of the objective function. Therefore, this article selects the BP neural network, combined with 2015-2019 online ride-hailing data, on the MATLAB platform to predict the demand trend of 2020-2024.

2   Data collection

Based on the data of China Statistical Yearbook, the transaction data of chauffeured car transactions, taxis, free rides, car timeshare rentals and shared bicycles from 2015 to 2019 are shown in Table 1. According to figure 1, it can be seen that chauffeured car transactions grew rapidly in 2015 and 2019, while shared bikes grew rapidly in 2016-2017 but reached saturation by 2019. The change in the growth of taxis is not very obvious. [5]

3   Model building

3.1 BP neural network

Error backpropagation neural network, referred to as BP (BackPropagation) network, is a kind of multi-layer feedforward neural network trained according to the error backpropagation algorithm, and it is the most widely used neural network. It is a kind of multi-layer neural network with three or more layers, and each layer is composed of several neurons. As shown in figure 3, it is the structure diagram of a BP neural network, in which each neuron in the left and right layers is fully connected, that is, each neuron in the left layer is connected to each neuron in the right layer, but there is no connection between the upper and lower neurons. The neurons are arranged in layers. The leftmost layer is called the input layer, which is responsible for receiving input data; the rightmost layer is called the output layer, and the neural network output data is obtained through the output layer. The layer between the input layer and the output layer is called the hidden layer because they are not visible to the outside. [7]
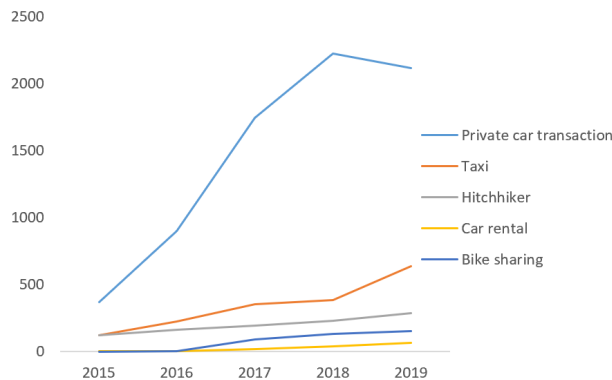

Figure 1. 2015 to 2019 various types of transport transactions

The BP neural network is trained according to the supervised learning model. When a pair of learning modes are provided to the network, first calculate the state and activation value of each layer until the last layer, the activation value of its neurons will propagate from the input layer to the output layer through each hidden layer, and each neuron in the output layer will output the network response corresponding to the input mode. Then, calculate the error of each layer, and the calculation process of the error is advanced from the last layer. According to the principle of reducing the error between the desired output and the actual output, each connection weight is modified layer by layer from the output layer through each hidden layer, and finally back to the input layer. Because this correction process is carried out layer by layer from output to input, it is called the "error backpropagation algorithm". With the continuous training of this kind of error backpropagation, the correct rate of the network response to the input mode will continue to improve. [7]

3.2 Genetic algorithm

Genetic algorithm is a computational model that simulates the biological selection process of Darwin's biological evolution theory and genetic mechanism, and is a method to search for the optimal solution by simulating the natural evolution process. This algorithm mathematically uses computer simulation operations to transform the problem-solving process into a process similar to the crossover and mutation of chromosome genes in biological evolution. When solving more complex combinatorial optimization problems, relative to some conventional optimization algorithms, it is

usually possible to obtain better optimization results faster.

The connection weights and thresholds of each layer are randomly assigned during the BP neural network training process, and the selection of connection weights and thresholds often affects the final prediction accuracy, therefore, this paper introduces a genetic algorithm. Genetic algorithm is a search heuristic algorithm for solving optimization problems [9]. This paper uses GA to optimize the initial weights and thresholds of the BP neural network, so that the BP neural network has better prediction accuracy.

The process of using genetic algorithm to optimize the BP neural network is to first determine the BP network topology, input the training set and test set data, and perform normalization processing; set the parameter values required for BP neural network training and perform the initial training. According to the first training of BP neural network and the result error value, the initial population of GA is determined, and the fitness function is set, and the fitness function is determined using the prediction error of BP neural network. Determine the number of iterations, iterate, and perform operations such as selection, crossover, and mutation. Finally, the optimal individuals obtained by GA are assigned to the BP neural network as weight and threshold input, and then BP is run to obtain the best prediction result.

Figure 2. Genetic algorithm optimizes BP neural network process

The alternating number of this learning is 25000 times. Through the repeated iterative operation, a correction value for the signal passing through the connection, called weight, is shown. The weight value of each layer can be adjusted through learning. The learning rate can determine the weight change in each cycle, and a fixed learning rate of 0.05 is used in this study. The pure linear function (purlin) is applied to the output layer, and the training function (trained), is used to set the maximum number of alternating learning times as 25000. Through the repeated iterative operation, the weight coefficient and threshold are determined, the learning and training process is over, and the model is established.
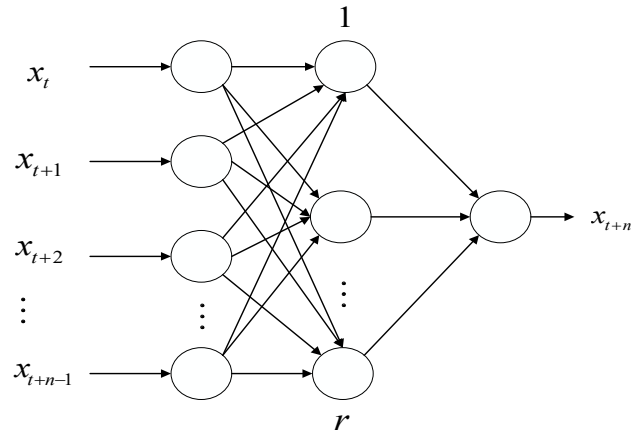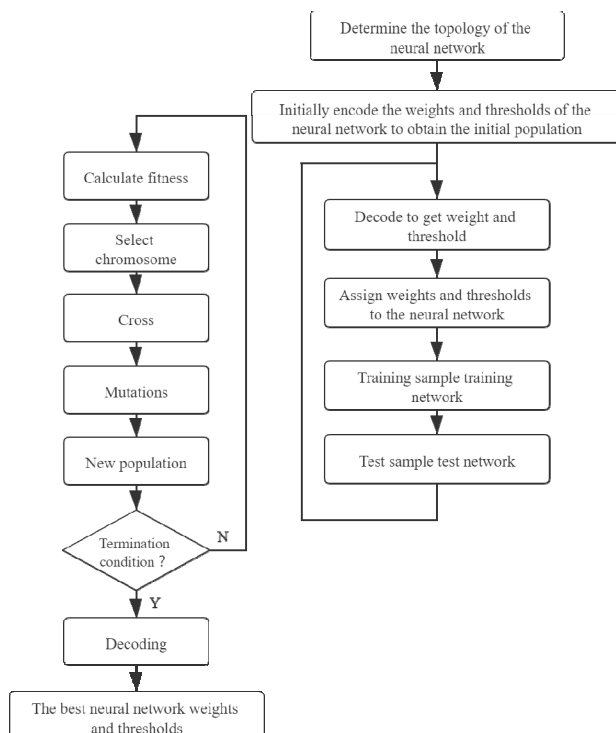

Figure 3. BP neural network conduction mode

4   Results and result analysis

4.1 Result calculation

Based on the MATLAB platform, BP neural network is constructed. All kinds of transaction data of 2015 Mel 2019 are substituted, and the algorithm parameters are adjusted. The flow chart of the partial BP neural network is shown in figure 4. [8]

4.2 Result analysis

Based on the MATLAB platform, all kinds of transaction data in 2015 and 2019 are substituted by the BP neural network. The projected demand for 2020-2024 is shown in the
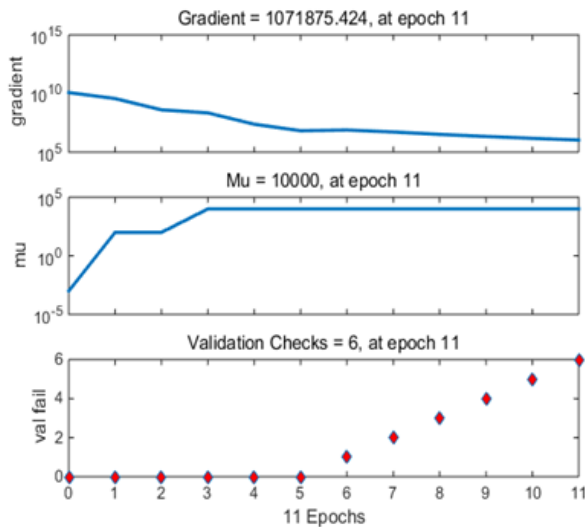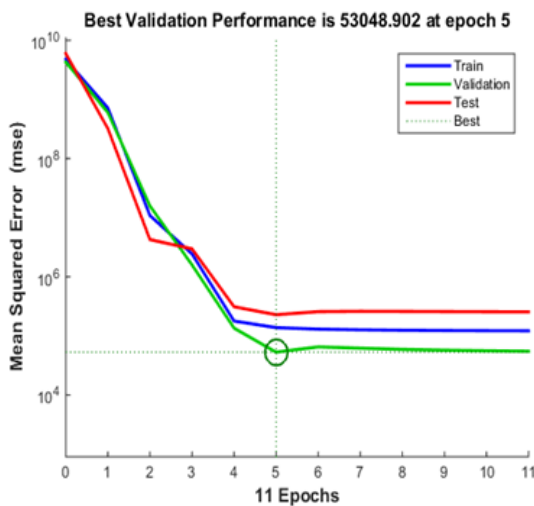
Figure 4. Selection of transmission times

Figure 5. Selection of prediction nodes

Figure 6. As can be seen from the picture, the transaction volume of chauffeured cars and taxis is rising rapidly. This shows that the taxi market still has potential, and competition in the taxi industry will gradually intensify in the future. The growth of hitchhiking, car timeshare and shared bikes are slow. It may be that the saturation of shared bicycle resources leads to the saturation of the selected quantity.
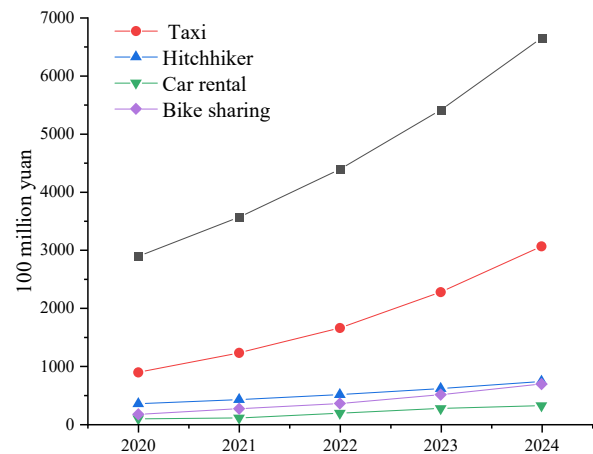

Figure 6. BP neural network forecast trend

5 Conclusion

Based on the BP neural network model and online car-hailing data of 2015 and 2019, this paper predicts the demand trend of the 2020 car and 2014 on the MATLAB platform. The results show that the demand for shared cars and rides will increase rapidly, while the demand for shared bicycles and taxis will rise slowly. To sum up, the development of ride-hailing service and ride-sharing services will accelerate the progress and rapid development of the city.

References

[1] Xinze Hu, Qihong Zhang, Hanpu Sun, Jianing Wang. Urban traffic congestion prediction based on BP neural network [J]. Electronic Technology and Software Engineering, 2020 (01): 186-187.

[2] Haitao Li. Network traffic prediction of the optimized BP neural network based on Glowworm Swarm Algorithm [J]. Systems Science & Control Engineering, 2019,7 (2).

[3] Qi Gao, Jianyu Chu, Yinfeng Li. Prediction model of terminal area congestion level based on PSO-BP neural network [J]. Aeronautical Computing Technology, 2019, 49 (06): 57-61.

[4] Jie Yang,Jing Dong,Zhenhong Lin,Liang Hu. Predicting market potential and environmental benefits of deploying electric taxis in Nanjing, China[J]. Transportation Research Part D,2016,49.

[5] Peng Guo. Intelligent traffic light control system based on LTE-V technology and neural network [C]. China Intelligent Transportation Association. Proceedings of the 14th China Intelligent Transportation Annual Conference (2). China Intelligent Transportation Association: China Intelligent Transportation Association, 2019: 197-206.

[6] Liwei Hu, Jinqing Yang, Yueren He, Ling Meng, Zhenwu Luo. Research on vehicle operation risk identification in urban traffic congestion environment based on improved BP neural network [J]. Highway Transportation Science and Technology, 2019, 36 (10): 105-113 .

[7] Song Li, Lijun Liu, Yongle Xie. Genetic algorithm optimized BP neural network for short-term traffic flow chaotic prediction [J]. Control and Decision, 2011, 26 (10): 1581-1585.

[8] Zifeng Jiang, Xiaokun Liu. Traffic incident detection algorithm based on neural network [J]. Journal of Xi'an Highway and Transportation University, 2000 (03): 67-69 + 73.

[9] Guochun Liu, Qiang Fei, Wuyun Zhao, Fei Dai. Application of optimization toolbox based on Matlab genetic algorithm [J] .Mechanical Research and Application, 2014,27 (02): 71-73.