Online Car-hailing Service Quality Evaluation Based on BP Neural Network

Mingzhi Wang¹, Xianyu Wu¹,*

¹School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China

*Corresponding author’s e-mail: 18120900@bjtu.edu.cn

Abstract: In order to reduce the subjectivity and make sure the online car-hailing evaluation results more authentic and credible, an evaluation index system is constructed from five dimensions of the safety, cost, time, reliability and empathetic, and then a BP neural network model is proposed to solve it. Through the supervised model training of SP survey data, the maximum error between the network output and the actual results no more than 3.68%, which can meet the requirements of online car-hailing service quality evaluation. Finally, a mathematical statistical analysis method is proposed to comprehensive evaluate the online car-hailing service quality. The results show that the evaluation index system is scientific and reasonable, and the model can effectively improve the objectively of the evaluation results and reflect the quality level of online car-hailing service.

1. Introduction

At present, many problems in online car-sharing operation make passengers question the service quality, and make the supervision mode more complex. How to establish a good evaluation system of service quality is an urgent task in the construction of the online car-sharing industry.

Yao Zhigang¹ Based on the investigation of taxi service quality in Hangzhou, it is found by factor analysis that the passenger perceived service quality can be evaluated from four aspects: tangible, guaranteed, reliable and caring, and the passenger perceived service quality, and this indicators can be evaluated by calculating the passenger perception and expected value by SERVQUAL method. Jin Meng² Based on the analysis of the whole process of passenger travel, Delphi method and questionnaire method are used to determine the evaluation system of ride-hailing service quality including comfort, reliability, value, safety and responsiveness, and the weight of each influencing factor is determined by AMOS method. Cui Qingan et al³ Through the SERVQUAL scale, the factors affecting the service quality of ride-hailing network are set up, and the correlation and weight of each index are studied by DANP method. Jiang Jiating et al⁴ Taking the two dimensions of platform service and driver service as the first level index, the five elements of service quality as the second level index, using analytic hierarchy process to analyze the index, finally put forward constructive suggestions on the development of online car-sharing from the government level, platform level and user level.

It can be seen that online car-sharing is a new industry in China, and the evaluation of service quality is relatively few. Hierarchical analysis and other traditional methods are mainly used in the research methods, but these methods are very difficult to determine the weight, have greater subjectivity, and when there are too many indicators, it is difficult to judge the importance of the indicators. The BP neural network model can overcome this defect.
2. Construction of Evaluation Index System of Service Quality of Net Car-sharing

The research on service quality evaluation can be regarded as whether the service quality of online car-sharing has reached the passenger consumption expectation. Therefore, the evaluation index system needs to select the appropriate index from the perspective of passenger satisfaction. In general, a complete service process in the process of using online ride-hailing mainly includes the following points: passenger demand for taxi, platform response demand, driver receiving order, driver arriving at the passenger location, driver transporting passengers, arriving at the destination and complete payment, passenger service quality evaluation. Therefore, this paper will analyze the passenger demand and expectation according to these seven service steps, and then select the appropriate evaluation index. The evaluation index system of service quality is shown in Table 1:

| Level I indicators | Secondary indicators | Level I indicators | Secondary indicators |
|--------------------|----------------------|-------------------|----------------------|
| Safety validity B1 | C1 Compliance driving rate | Time Validity B3 | C7 Efficiency of drivers |
| Cost validity B2 | C4 Prices | Reliability validity B4 | C8 Trustworthiness |
| C6 Response time | C5 Penalty | C9 On time | C10 Punctuality |
| Time Validity B3 | C11 Service attitude | C12 Environment of vehicles |

3. Model Establishment and Parameter Calibration

3.1. Model network node determination

3.1.1 Input layer nodes.

set to 12 according to the number of secondary indicators

3.1.2 Implicit layer nodes.

In this paper, the "improved gold method" is used to determine the hidden layer neurons. The steps are as follow[8]:

**Step1:** range of hidden layers

\[ a = \log_2 1 \leq L \leq \sqrt{1 + J} + c = b \quad (1) \]

where \( I, J, L \) are the number of input layer nodes, output layer nodes and hidden layer nodes difference; \( c \in [1,10] \). In this paper, \( I = 12, J = 1 \) so \( L \in [4,14] \) can be calculated by formula (1).

**Step2:** trial and error comparison

\[
\begin{align*}
g_1 &= 0.618 \cdot (b - a) + a \\
g_2 &= 0.382 \cdot (b - a) + a
\end{align*}
\]

\[
[a', b'] = \begin{cases} 
[g_1, g_1], E(g_1) > E(g_2) \\
[g_1, b], E(g_1) < E(g_2) \\
g_1, g_2, E(g_1) = E(g_2)
\end{cases}
\]

(3)

Where \( E(g_1), E(g_2) \) are the mean square of \( g_1, g_2 \)

**Step3:** repeat the Step2 until can't get any smaller ones \( g_1, g_2 \)

**Step4:** number of hidden layer nodes

\[
E(L) = \min \{E(a), E(g_1), E(g_2), E(a'), E(b'), \cdots \}
\]

(4)
Step5: determine the number of hidden layer nodes
A BP neural network is set to train 2000 times and the error accuracy is $10^{-7}$. Through cyclic training, the number of hidden layer nodes with the smallest number of training times, the fastest convergence speed and the smallest error are selected. The results are shown in Table 2:

| Number of neurons | MSE      | MSE       |
|-------------------|----------|-----------|
| 4                 | 0.19     | 1.51×10^{-6} |
| 5                 | 0.05     | 6.92×10^{-7}  |
| 6                 | 0.03     | 1.69×10^{-5}  |
| 7                 | 0.18×10^{-2} | 3.39×10^{-9} |
| 8                 | 3.39×10^{-9} | 0.33×10^{-2}  |
| 9                 | 3.39×10^{-9} | 8.05×10^{-5}  |
| 10                | 3.39×10^{-9} | 3.06×10^{-8}  |

From Table 2, when the hidden layer number is 11, the minimum training mean square difference is 3.3923×10^{-9}. Therefore, the number of neurons in the hidden layer is set to 11.

3.1.3 Output layer nodes.
This paper focuses on the overall evaluation of the sample, so set to 1.

3.2. Selection of transfer functions
The output results range from 0 to 10, so the unipolar S function logsig is selected as the transfer function between the input layer and the hidden layer, and the purelin as the transfer function between the hidden layer and the output layer.

3.3. Selection of training functions
The selection of training function is related to the size of training sample size. In order to select the appropriate training function, this paper compares the training efficiency of different training functions by inputting training samples and setting the same parameters, so as to select the most suitable training function. Let the number of input neurons be 12, hidden neurons 9, output neurons 1, the number of training is set to 2000 times, the iteration accuracy is 0. Compare the number of iterations of each training function, and the iteration accuracy as shown in Table 3:

| Function   | Number of iterations | Iterative Precision |
|------------|----------------------|---------------------|
| TRAINGLM   | 87                   | 2.41×10^{-17}       |
| TRAINGDA   | 1981                 | 0.026472            |
| TRAINGDX   | 1994                 | 0.017538            |
| TRAINGD    | 2000                 | 0.064658            |
| TRAINGDM   | 2000                 | 0.053349            |

As can be seen, TRAINGLM functions have advantages over other functions in both training efficiency (iteration times) and iteration accuracy. Therefore, the TRAINGLM function is chosen as the training function of BP neural network.

4. Case studies
4.1. Sample processing
A total of 320 questionnaires were distributed and 302 valid questionnaires were recovered. In order to make the evaluation process of service quality more scientific and reasonable. First of all, the evaluation index should be quantified. The set of impact assessment indicators $D$ can be written as follows

$$D = \{d_1, d_2, \cdots, d_i\}, i = 1, 2, \cdots, n$$

(5)

Where $d_i$ is the first $i$ index of the second level evaluation set; $n$ is the number of the second level evaluation index.
Then the evaluation result grade is quantified by the standard. The evaluation result grade is the set of the final evaluation results of the evaluator's service quality of the online car-sharing operation. It can be written as follows

\[ V = \{v_j \} , j = 1, 2, \ldots, m \]  

(6)

Where \( v_j \) is the first \( j \) evaluation results; \( m \) is the number of evaluation grades. In setting the evaluation level of the domain, usually take odd digits. Therefore, the evaluation grade of service quality of net-car-sharing is five grades:

\[ V = \{ \text{Fine}, \text{ OK}, \text{ General}, \text{ Difference}, \text{ Poor} \} \]  

(7)

In order to facilitate calculation, the grade is quantified. The corresponding evaluation grade and its membership evaluation score threshold range are shown in Table 4.

| Evaluation rating | Fine | OK   | General | Difference | Poor |
|-------------------|------|------|---------|------------|------|
| Membership score  | [9,10]| [8,9)| [6,8)| [4,6)| [0,4)|

**4.2. Training sample establishment**

The training sample is the expert sample. The quality of the training sample determines the scientific nature of the model. So, the training sample should reflect the evaluation object from different dimensions as much as possible. Therefore, the evaluation indexes involved in the investigation process are included in the training set. The training set is as follows:

\[ G = (g, h) \]  

(8)

Among them:

\[ g = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix} \]  

(9)

\[ h = (h_1, h_2, \ldots, h_m)^T \]  

(10)

In the formula: \( g_{mn} \) indicates the first \( m \) user to the first \( n \) item index score; \( h_m \) indicates the first \( m \) user to the net car service quality total score, namely expects the output result.

**4.3. Results analysis**

The results obtained by entering the test sample into the trained model are shown in Table 5:

| Serial number | Actual score | Network output | Relative error | Serial number | Actual score | Network output | Relative error |
|---------------|--------------|----------------|----------------|---------------|--------------|----------------|----------------|
| 1             | 5            | 5.0000         | 0.00%          | 11            | 6            | 6.0000         | 0.00%          |
| 2             | 9            | 9.3998         | 2.17%          | 12            | 8            | 8.1185         | 1.46%          |
| 3             | 7            | 7.1448         | 2.03%          | 13            | 8            | 8.2269         | 2.76%          |
| 4             | 6            | 6.1836         | 3.06%          | 14            | 7            | 7.0000         | 0.00%          |
| 5             | 7            | 7.0199         | 0.28%          | 15            | 6            | 5.8571         | 2.44%          |
| 6             | 10           | 10.3823        | 3.68%          | 16            | 7            | 6.9857         | 0.20%          |
The maximum error between the network output and the actual score is 3.68 and the average error is only 1.71. It can be seen that the evaluation results of the trained BP neural network model on the service quality of the network ride-hailing are basically consistent with the actual evaluation. The evaluation results of the BP neural network model are reliable.

By statistical analysis of the satisfaction degree of each index in the sample, the average score radar distribution map of 12 indexes can be obtained as shown in figure 1.

By analyzing the chart, we can see that the penalty index has the lowest score. Secondly, the complete license rate, vehicle emergency facilities, ride price and other indicators score low. Therefore, it is suggested to take the following measures to improve the service quality satisfaction: 1) standardize the operation system of the ride-sharing platform and ensure the rights and interests of consumers; 2) standardize the pricing mode of the ride-sharing network; 3) strictly supervise the legal operation of the ride-sharing network. Create a green ride environment.

5. Conclusion
A service quality evaluation model based on BP neural network model is proposed in this paper. The model can effectively shorten the modeling time, reduce the workload and improve the objectivity in the evaluation process. The parameters of BP neural network model are calibrated by SP survey data, and the factors affecting the evaluation results of service quality are analyzed. The results show that: (1) the maximum error between the network output and the actual results is 3.68 and the average error is 1.71, which can meet the requirements of passenger service quality evaluation, indicating that the model can evaluate the service quality of ride-hailing; (2) at present, default compensation, ride price and the corresponding time are the bottleneck factors restricting the further improvement of ride-hailing service quality; (3) some constructive suggestions are put forward according to the research results, which have certain practical significance to promote the healthy and stable development of ride-hailing.

Acknowledgments
This research is funded by the Fundamental Research Funds for the Central Universities Education of China(2019JBM028). We hereby express our gratitude for this.
References

[1] Yao, Z.G. (2010) Measuring Passengers Perceptions of Taxi Service Quality with SERRQUAL Weighted by Factor Loadings. Journal of Wuhan University of Technology, 01: 82-86.

[2] Jin, M. (2019) Research on the service quality evaluation of car-hailing platforms from the perspective of customer experience. Wuhan: Wuhan University of Technology.

[3] Ren, Q.L., Wang, L. (2017) Research on the evaluation of ride-hailing service quality. Highway &Automotive Applications. 02: 44-49.

[4] Jiang, J.T. (2018) Research on evaluation index system of online Car-Hailing service quality in beijing. Beijing: Beijing University of Chemical Technology.

[5] Gu, T.Q., Zhou, Y.M., Li, J., et al. (2019) An Empirical Study of Online Car-hailing Users' Characteristics and Service Quality Evaluation. Shanghai Urban Management., 05: 61-68.

[6] Wang, Y.J., Tu, J. (1998) Neural Network Control. China Machine Press, Beijing.

[7] Yang, G.Y., Shi, T.Y., Zhang, Q.L. (2016) Railway Passenger Service Quality Evaluation Based on BP Neural Network. Journal of Transportation Systems Engineering and Information Technology, 05: 33-39.

[8] Nie, C.C. (2016) Research on evaluation of taxi informatization service quality based on multivariate data. Xian: Chang'an University.