Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6, 2369-2380
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i6.2482

Research on tourism demand forecast based on BP neural network -- Taking the number of tourists in Mulan Scenic spot in Huangpi District for example

Na Yin[1], Chonlavit Sutunyarak[2*]
[1,2]Chakrabongse Bhuvanarth International Institute for Interdisciplinary Studies (CBIS), Rajamangala University of Technology Tawan-OK, Bangkok, 10400, Thailand; na.yin @ rmutto.ac.th (N.Y.) Chonlavit_su@rmutto.ac.th (C.S.)

Abstract: With the development of the economy and society, the public's demand for tourism is getting higher and higher, followed by the environmental carrying capacity of the destination, the bearing capacity of traffic reception facilities, and the psychological carrying capacity of tourism residents being affected by different degrees. This paper took the analysis of the number of tourists in Mulan Scenic Spot in Huangpi District for the May Day holiday as an example, combined with the BP neural network algorithm, forecast the tourism demand, and put forward the local government's suggestions on holiday tourism management based on the data research.

Keywords: *Tourism demand forecast research tourism demand.*

1. Background

Tourism demand forecasting has essential guiding significance in the policy research of tourism destination planning and economic development, as well as in the operation management practice of tourist destination reception departments. From the 1960s, western scholars began to study the forecasting model and empirical research of tourism demand and developed diversified forecasting methods. Since 2000, domestic scholars have gradually shifted from qualitative research to quantitative research based on the introduction of foreign research methods. However, relevant research is still in the initial stage due to the need for more statistical data related to the number of tourists and the insufficient application of advanced skills upgrading tools. The period from 2000 to 2021 will be 20 years of rapid development since China entered the new century. The tertiary industry has gradually become the largest industry in the economic field. At the same time, tourism plays an increasingly important role in the tertiary sector, and the demand for tourism is increasing. In order to better meet the tourism needs of tourists, improve the tourism experience, and find a balance between economic development and environmental bearing, it is necessary to do a scientific investigation and forecast the demand for tourist destinations.

2. Research Significance

The characteristics of foreign tourism demand literature analysis are the study and analysis of the latest international hot spots or real-time dynamics with quantitative methods and the acquisition of quantitative forecast data for decision-making. The characteristics of domestic literature on tourism demand analysis are usually "small incision" analysis of the tourist flow of a scenic spot, combined with quantitative analysis and qualitative analysis methods, to quantitatively forecast the tourist flow of future scenic spots. However, for some 5A regions with global tourism, it is often not a specific scenic spot that experiences a surge in traffic during holidays. Still, the influx of tourists affects regional traffic, thus making residents' lives inconvenient and their travel experience poor. Therefore, analyzing the

* Correspondence: Chonlavit_su@rmutto.ac.th

regional tourist volume from a more macro perspective is necessary by combining quantitative methods. In the face of major holidays and a surge in the number of tourists, policymakers should do an excellent job in traffic control, public facilities, tourism population diversion, and other measures to achieve the purpose of not only affecting the everyday life of residents but also to allow tourists to play comfortably.

3. Literature Review

3.1. Foreign Research Status

In The Forecasting Research of Beijing Tourism Demand Based on the BP Neural Network (2014) [1], Yang Yu and Shi Min Wang used the MATLAB tool to conduct an empirical study on the number of tourists in Beijing from 1994 to 2012. They believed that the Beijing tourism demand prediction model based on the BP neural network could more accurately predict the number of tourists in Beijing in the future.

In Tourism demand forecasting with online news data mining (2021), Park Eunhye, Park Jinah, and Hu Mingming adopt the combination method of the key themes of the structured theme identification model related to tourism demand. This method involves mining news theme words and combining the selected tourism theme variables to predict tourism demand [2].

Fotiadis Anestis and Polyzos Stathis, in The Good, the Bad and the Ugly on COVID-19 Tourism Recovery (2021) [3], analyzed the impact of COVID-19 on international tourism demand, combined with short-term memory neural networks and generalized additive models, predicted that the decline of international visitors could be between 30.8% and 76.3% by June 2021.

3.2. Domestic Research Status

Luo Chunyan analyzed domestic and foreign tourism demand from a qualitative point of view in the Analysis of Domestic and Foreign Demand for China's Tourism Industry Economy (2017) and believed that domestic demand continued to grow at a high level [4]. That outbound demand would return to rationality. This paper forecasted tourism demand from the macro tourism industry development perspective and obtained the primary direction of tourism development. However, lacking quantitative data and research tools to verify, the analysis is prone to generalities. Chen Tao and Liu Qinglong, in "Research on the Application of Big Data in the Context of Smart Tourism: Taking Tourism Demand Forecasting as an Example" (2015), used "Big data of a tourist attraction" to predict the changes in tourist volume in the future period, providing a reference for managers to make triage and emergency plans [5]. They believe that compared with the vector autoregressive model, the combination of "tourism big data" and artificial neural network model can make a more accurate prediction of tourist volume. In her Research on Shanghai Tourism Demand Forecast Based on BP Neural Network (2020), based on the data of a tourist attraction in Shanghai from 2006 to 2018, Xu Jingna compared the calculation accuracy of four algorithm models, namely linear regression, exponential simulation, grey model and BP neural network, and concluded that BP neural network has a better fitting effect [6]. In the BP Neural Network-based Intelligent Prediction Method for Tourist Flow in Scenic Spots (2021), Wang Chunpeng combined random forest algorithm and BP neural network algorithm to conduct comparative research on part of the historical data of Dujiangyan Tourism Big Data Center and proposed that BP neural network algorithm has higher accuracy and better prediction effect [7].

After selecting the central research direction, before determining the specific research topic, the author consulted a large number of literature materials to understand the relevant research situation and, through comparison and analysis, limited the research content according to the feasibility of the research, the interest and ability of the researcher, determined the research scope of the topic, and better controlled and grasped the topic. After the literature analysis, the research experience at home and abroad provided valuable experience for our research. In particular, many researchers at home and abroad have pointed out from different perspectives that the BP neural network model has higher prediction accuracy than other algorithm models. Based on previous studies, this paper will select a regional perspective, combined with quantitative methods, to study and analyze the demand for tourism.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

4. Huangpi Local Tourism Development Overview and Problems

Huangpi District of Wuhan City has 4400 years of building and more than 1800 years. In the county's long history, it has formed the reputation of being a "thousand-year-old county, the hometown of Mulan, Riverside flower capital and City of Filial piety." It has four cultural business cards: Panlong City culture, Mulan culture, Ercheng culture, and Shouyi culture, with excellent natural landscapes and profound historical and cultural resources. Local governments have been working hard to turn natural resources and historical and cultural resources into tourism resources. Since 2000, the government has taken the development of tourism as a critical task and actively promoted the application for 4A-level scenic spots. In 2012, Huangpi District began to integrate the resources of four 4A-level scenic spots, Mulan Mountain, Mulan Tianchi, Mulan Grassland, and Mulan Yunwu Mountain, and jointly created national 5A-level tourist attractions under the name of "Huangpi Mulan Cultural Ecological Tourism Area."After three years of hard work, Huangpi has nearly 20 scenic spots in Mulan, such as Mulan Mountain, Mulan Sky Pond, and Mulan Grassland. Among them are one national 5A-level scenic spot, 4 4A-level scenic spots, and four 3A-level scenic spots. According to the data disclosed by the Statistics Bureau of Huangpi District, in 2020, the total number of tourist arrivals in the district was 26.6887 million. The comprehensive tourism revenue was 16.278 billion yuan, an increase of 8.02% and 10.01%, respectively, and the visibility, reputation, and influence of the whole region's tourism great Huangpi continued to increase. Tourism has become an essential pillar of Huangpi's industry, vital for revitalizing agriculture and enriching the people. Therefore, Huangpi has realized the transformation from a traditional agricultural area to a critical tourism area.

However, due to the rapid growth of the tourism population, there are also many problems, which are mainly reflected in the following:

1. Inadequate capacity of traffic reception facilities. Taking the May Day holiday as an example, the annual holiday Huangbei Daihuang Expressway, Qianchuan Turntable, Fire Tower Highway, and other main roads will be clogged on May 1, 2020, from 9 am; tourists drive to Mulan scenic area group of Mulan Street, fire tower line and other road vehicles lined up. The traffic "aorta" is obstructed, and the "branch artery" is not smooth. Starting at 10 am on April 30, the bypass roads in and out of the scenic area were congested, the cars in and out could not go, and the tourists were blocked until late at night before going down the mountain. From 3 pm on April 30, the highway on both sides of a well-known scenic spot in Huangpi was completely paralyzed, and there were still many self-driving tourists who could only go down the mountain at 10 pm.
2. The environmental carrying capacity of tourist destinations needs to be increased. The rapid growth of the tourist population leads to the lack of supply of public service facilities such as public toilets and garbage dumps, resulting in a sudden surge of tourists that can only be "solved on the spot," which seriously destroys the natural environment of the scenic spot and has a severe impact on the environmental carrying capacity of the tourist destination.
3. The psychological carrying capacity of residents in tourist destinations needs to be improved. The main road of Huangpi city is connected with the expressway of scenic spots. The congestion of the expressway in scenic spots directly leads to the obstruction of the main road in the town. At the same time, many tourists have brought great inconvenience to the transportation of residents. At the same time, they have also severely impacted the psychology of local residents. They believe that foreigners have crowded out local tourism resources and living resources, and they are not only local people who can not fully enjoy local tourism resources with their families during holidays, but their daily life is affected, thus affecting their psychological status.

5. Overview of Research Methods and Tools

In this paper, the BP neural network and Matlab software tools are used to forecast the demand for tourism. In line with the writing idea of "seeing the big with the small," taking the number of tourists in

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

Mulan Scenic Spot in Huangpi District, for example, to analyze the tourism demand, refine the data, and finally sort out and summarize the valuable information.

(1) A neural network (full name Artificial neural network, English name Artificial Neural Net, referred to as ANN) is a simulation of the operation process of the human brain. "People will use this process of human brain decision-making to solve practical problems; the process is roughly:

1. Input sample (including input and expected output)

2. Conduct sufficient training according to specific training algorithms

3. The training ends when the actual output of the network and the expected output are within a particular range of error.

4. After obtaining the solution, the model can be applied to solve similar problems.

BP neural network. Error Back Propagation Algori (BP neural network) In the practical application of artificial neural networks, neural networks are the most classic and commonly used neural network algorithms. In practical applications, 80% to 90% of artificial neural network models are based on BP neural networks or variations. The idea is to transfer the errors generated by the output layer to the hidden layer through the connection weight, step back, and finally reach the input layer to adjust the weights between the layers according to the errors. BP neural network, as shown in the basic schematic diagram [8]. It has a simple structure and a wide range of investigation objects. It is mainly used in complex system simulation, pattern recognition, classification, nonlinear mapping, and process control.

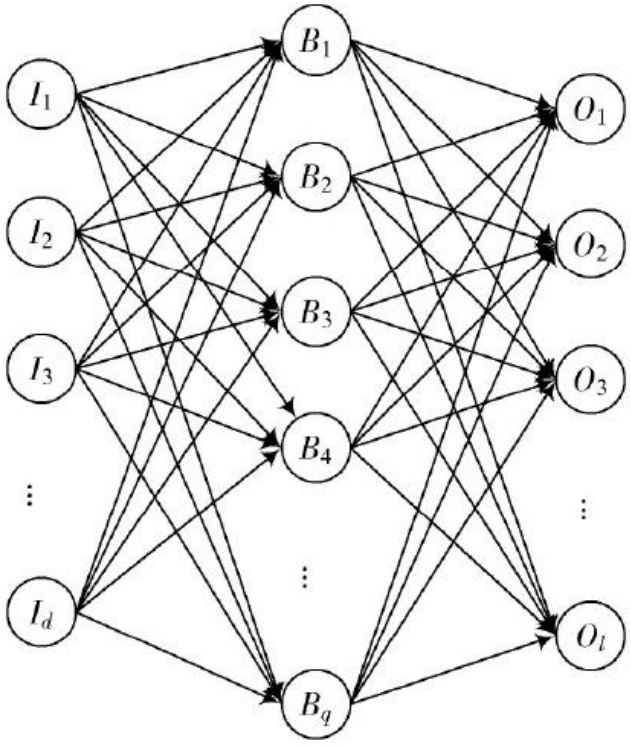

Figure 1.
BP neural network basic schematic diagram.

An artificial neural network consists of three layers (input layer, hidden layer, and output layer). There is no connection between neurons in the same layer, and the mode of complete interconnection between layers is adopted, and the hidden layer can contain multiple layers. Since there is no interconnection between the same layer elements, the primary processing unit of the BP neural network

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

(except the input layer elements) is a nonlinear input-output relationship, and the input and output values of the processing unit can be continuously changed. There are N-dimensional vectors (input nodes) and L-dimensional vectors (output nodes) in the graph, and the sigmoid function is their nonlinear activation function $f_u = \frac{1}{1+e^{-u}}$.

BP neural network is composed of two propagation directions: positive and negative.

5.1. Forward Propagation Process

The input sample (N) passes through the input layer, reaches the hidden layer, and gradually spreads. Each layer's neuron state only affects the neurons' state in the next layer. If the output (L value) matches the expected value, if the output (L value) does not match the expected value, it enters the backpropagation.

5.2. Backpropagation Process

Determine the error between the actual output (L value) and the expectation and modify it layer by layer from back to front, adjusting the threshold of linking each neuron until it reaches the allowable error range.

As shown in the flow chart of the prediction method based on the BP neural network, the BP neural network is an algorithm that combines the forward propagation process and the backpropagation process, and constantly adjusts parameters and modifies the deviation through machine learning and multiple iterations to control the error within a reasonable range.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

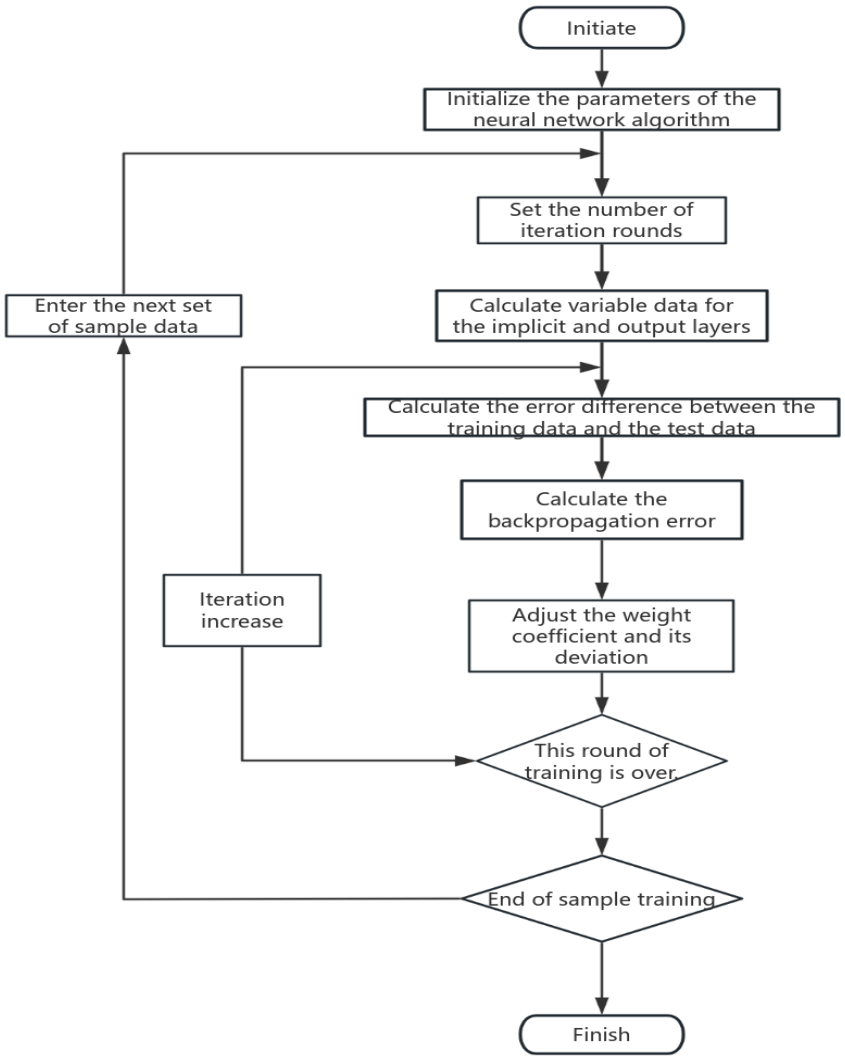

Figure 2.
Flow chart of prediction method based on BP neural network.

5.3. Research Tools

MATLAB is an interface program by Cleve Moler of the University of New Mexico in the 1980s to facilitate students' learning of linear algebra. It is named MATLAB (Matrix Lab), which is matrix Lab. MATLAB is a powerful teaching software that integrates matrix calculation, numerical analysis, visualization, and dynamic system simulation modeling into a development environment, providing great convenience for scientific research. In addition, MATLAB also includes a powerful toolbox, which provides users with a significant role in simulation modeling.

The BP neural network model design using MATLAB follows the following seven steps: 1. Information collection. 2. Data normalization processing. Third, the selection of output nodes and hidden layer nodes is determined. 4. Establish the sample output value. 5. Set training parameters. 6. Create a BP neural network and initialize and call the training function. 7. Error test. 8. Test results. MATLAB software is used for data analysis and processing.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

6. Empirical Research

6.1. Time Series Analysis

6.1.1. Data Collection

Due to the lack of statistical data on some tourists, this project combined with the "Foreign economic trade and tourism" part of the Statistical Yearbook of Huangpi District (2000 to 2010) and the Huangpi Tourism Statistical Bulletin (2010 to 2020) issued by Huangpi District Culture and Tourism Bureau about the tourism population statistics of the May Day Golden Week. The data for each of the two years in the sample data is used to predict the third year, i.e., the data for 2000-2001 predicts 2002, the data for 2001-2002 predicts 2003, and so on. The program randomly divides these data sets into training sets and test sets according to the proportions of 70% and 30%. The test set tests the trained model to reflect the fitting situation.

Table 1.

Huangpi District from 2000 to 2021.

A particular year	Number of people (people)
2000	284823
2001	290519
2002	296330
2003	302256
2004	311324
2005	320664
2006	330284
2007	346798
2008	364138
2009	382344
2010	420579
2011	462637
2012	508900
2013	763351
2014	839686
2015	1007623
2016	1229322
2017	1507145
2018	1955467
2019	2100287
2020	2168125

6.1.2. Normalization Treatment

After obtaining the first-hand data, to facilitate comparison and processing, the data in the table is concentrated between (0,1), and the data is normalized according to the normalization formula:

$$X = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$

Among them, X is the quantized index value, the actual data of an indicator, the maximum value, and the minimum value. MATLAB normalizes the data.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

6.1.3. Select the Number of Hidden Layers

Because of neural networks' vast parallel distribution structure and nonlinear dynamic characteristics, obtaining a simple, accurate, and general analytical expression or a formula for determining hidden layer nodes is complex. However, some qualitative conclusions obtained through extensive and long-term application will help rationally arrange the number of hidden layer nodes. Therefore, combined with the actual situation in the paper, this paper refers to the empirical formula to determine the number of hidden layer nodes:

$$p = \log_2 n$$

Where p is the number of nodes and n is the total number of samples. The number of samples is 44, so the number of hidden layer nodes is 5.

Table 2.
Sample expectation table 2000-2021.

A particular year	Expected number (Person)
2000	287671
2001	293424
2002	299293
2003	306790
2004	315994
2005	325474
2006	338541
2007	355468
2008	373241
2009	401462
2010	441608
2011	485769
2012	636126
2013	801518
2014	923654
2015	1118461
2016	1368200
2017	1731250
2018	2027800
2019	2134100
2020	1952900

6.1.4. Determine Sample Output

Standard methods in MATLAB include the gradient decline method (function trained), conjugate gradient method (function training), LM algorithm (function trail), and gradient descent method with driven quantity (function trained). This project selects a gradient descent method with momentum, good generalization ability, and high precision. Based on the tourism population of Huangpi from 2000 to 2020, the corresponding number of tourists in 2021 was predicted, the maximum number of iterations was set to 1000, the target convergence accuracy was set to 0.001, the tang function was selected for the hidden layer, and the trail function was chosen for the output layer.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

6.1.5. Determine the Sample Parameters

Different training parameters have different influences on the results. After many attempts, the author finally set the parameters as follows:

Table 3.

Implied layer transfer function	**logsig**
The output layer is a transfer function	logsig
Training function	traingdx
Show the interval	net.trainParam.show=100
Network learning rate	net.trainParam.ir=0.05
Momentum coefficient	net.trainParam.mc=0.9
Maximum training times	net.trainParam.epochs=1000
Target error	net.trainParam.goal=1e-5

6.1.6. Create BP Neural Network, Initialize and Call Training Function

```
net = newff(minmax(p), [5, 1], {'logsig', 'logsig'}, 'traingdx')
% Call the training function
traingdx
% Displays the interval
net.trainParam.show=100 ;
% momentum coefficient
net.trainParam.mc=0.9 ;
% Displays the interval
net.trainParam.epochs=1000 ;
```

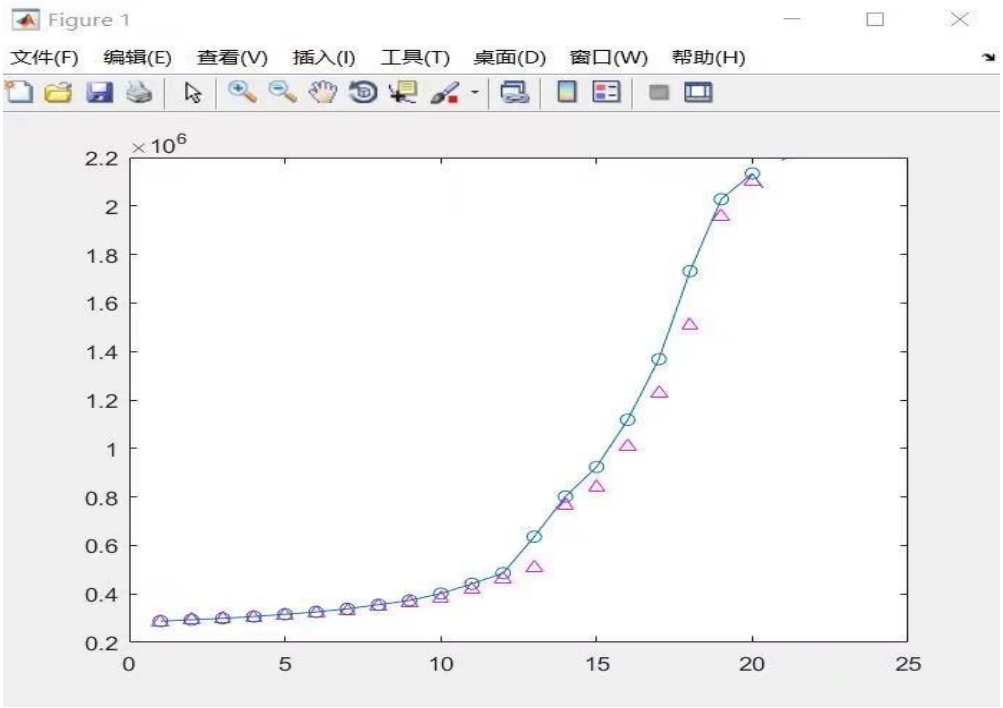

Figure 3.
The BP neural network training graph in Figure.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

As can be seen from the BP neural network training diagram, the circle points represent the actual number of tourists to Huangpi, and the triangle points represent the number of tourists predicted by the BP neural network algorithm after machine learning. From the perspective of change trend, after 22 times of machine learning, the two curves are highly fitted, and the model can be preliminarily determined to be formed and have a certain forecasting function. However, to further verify the model's prediction accuracy, it is necessary to conduct further error tests on the fitting results to determine whether the predicted value is within the error range.

6.1.7. Error Test

% Target error

net.trainParam.goal=1e-5 ;

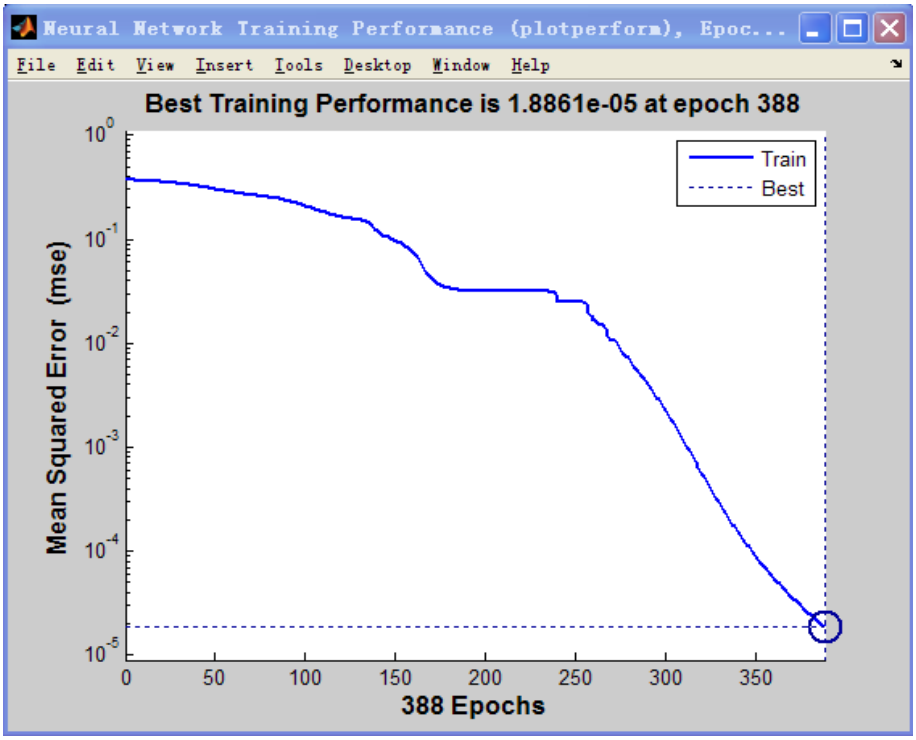

Figure 4.
Error plot.

The error curve shows that after 388 training, the error value is close to the error range, and the training purpose is achieved.

6.1.8. Test Results

Test samples were l and input was 1737857

P_test=[1737857]

The inspection function is:

y=slin(net, P_test) ;

The output result is:

Y=1737871

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482

The error value of 0. 9 * 10 6, less than the target error of 1e-5, is consistent with expectations and indicates that the experiment is valid. Therefore, the practice proves that the tourism demand prediction model based on the BP neural network designed in this paper is scientific, perfect, and reasonable.

7. Conclusion

1. Huangpi tourism population maintains a fluctuating growth. With the increase in per capita disposable income and the improvement of the paid holiday system, tourists are becoming more and more willing to travel and mature. Therefore, for tourists whose primary motivation is "promotion," the Huangpi local government should further improve reception facilities and service management, such as setting up large traffic warning signs at the main urban area's entrance or at the expressway's exit. Arrange traffic police to command on the road, encourage and organize local volunteers to guide vehicle diversion, and prevent large areas of traffic jams so as to improve tourists' sense of experience and satisfaction.
2. Huangpi public facilities should be adequately equipped in advance. In 2012, Huangpi Mulan Cultural Eco-Tourism Area was successfully established as a national 5A-level tourist attraction, marking the entry of Huangpi tourism into the top national tourist scenic area and a milestone in the development history of Huangpi tourism. Since then, Huangpi has had a "golden signboard" to accelerate the process of internationalization and enhance comprehensive competitiveness. This golden signboard makes Huangpi's tourist population jump from "100,000" to "millions" during the May Day Golden Week, and the lack of public facilities will affect the gold content of this gold signboard. During the holidays, the nodes of key scenic spots should be equipped with garbage recycling buckets, toilets, tea rooms, and other equipment to meet the needs of tourists.
3. The Huangpi local government should give residents a sense of gain. The free admission policy will help boost economic development after the epidemic. In the first half of 2020, Wuhan needs particular economic policies to stimulate the economy after the lockdown. Since launching the "Huiyu Hubei · Punch Wuhan" activity that year, Huangpi has become a popular tourist destination in China. Due to the exemption from tickets, the tourist population of Huangpi during the May Day Golden Week in 2020 reached 2.168 million, exceeding the number of 2.100,200 tourists in 2019. The influx of foreign tourists brings economic benefits to residents and a lot of inconvenience to local residents. The government should learn to guide them and balance the inner needs of residents. For example, Huangpi locals can visit scenic spots free of charge when they hold ID cards, and locals and foreigners do not rush into scenic spots simultaneously. It can also ensure the comfort of tourists and allow residents to enjoy the resources of the picturesque area.

References

[1] Y. Yu and S. M. Wang, "The forecasting research of beijing tourism demand based on the BP neural network," *Applied Mechanics and Materials*, vol. 571, pp. 128-131, 2014, https://doi.org/10.4028/www.scientific.net/AMM.571-572.128

[2] E. Park, J. Park, and M. Hu, "Tourism demand forecasting with online news data mining," *Annals of Tourism Research*, vol. 90, p. 103273, 2021/09/01/ 2021, https://doi.org/10.1016/j.annals.2021.103273

[3] A. Fotiadis, S. Polyzos, and T.-C. T. C. Huan, "The good, the bad and the ugly on COVID-19 tourism recovery," *Annals of Tourism Research*, vol. 87, p. 103117, 2021/03/01/ 2021, https://doi.org/10.1016/j.annals.2020.103117

[4] C. Luo, "Analysis of the domestic and foreign demand of China's tourism industry economy," *Reform and Strategy*, vol. 33, no. 11, pp. 150-152+170, 2017, https://doi.org/10.16331/j.cnki.issn1002-736x.2017.11.030

[5] T. Chen and Q. Liu, "Research on big data application under the background of smart tourism: Taking tourism demand forecast as an example," *E-Government*, no. 9, pp. 6-13, 2015, https://doi.org/10.16582/j.cnki.dzzw.2015.09.013

[6] J. Xu, "Shanghai Tourism Demand Forecast Research based on BP neural network," *China Economic and Trade Guide (Middle)*, no. 7, pp. 38-40, 2020, https://doi.org/10.3969/j.issn.1007-9777.2020.20.014

[7] C. Wang, "Intelligent prediction method for tourist traffic in scenic spots based on BP neural network," *Modern Electronic Technology*, vol. 44, no. 16, pp. 175-178, 2021, https://doi.org/10.16652/j.issn.1004-373x.2021.16.035

[8] Y. Hu and Y. Wang, "Application of artificial neural network in e-commerce," *Microcomputer and Application*, no. 4, pp. 4-6, 2003, https://doi.org/10.3969/j.issn.1674-7720.2003.04.001

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 2369-2380, 2024
DOI: 10.55214/25768484.v8i6.2482