

Genetic Neural Network Prediction of Car Ownership Based on Principal Component Analysis

Qing Wang and Yewang Zhou

School of Business, Huanggang Normal University Huanggang, 438000, china

Abstract—The prediction of the car ownership is the basic work for city traffic sustainable development. The paper carry on principal component analysis to influencing factors in the process of prediction of Wuhan City car ownership, determine the main components. Combining genetic algorithm with neural network, using genetic algorithm to optimize the weights of neural network, determine the initial weight values of neural network. Not only to improve the neural network training speed and generalization ability, but also overcome that the network is easy to fall into local minimum to a certain extent, then train the neural network, and carry on the prediction of car ownership. At last use a specific example to verify the prediction effect.

Keywords—principal component analysis; car ownership; genetic algorithm; neural network

I. INTRODUCTION

Study on Car ownership is a basic work of city road traffic planning, how to accurately predict the city car ownership is of great significance to the sustainable development of city traffic planning and management. The prediction of the car ownership is the basic work for city traffic sustainable development. Lam et al set the city road network capacity and parking facilities as two constraints of car ownership development, using the theory of bilevel programming to construct the car ownership prediction model[1][2].Tomas et al established dynamic car ownership prediction model based on the 2D model, Rich established micro economics model on prediction of car ownership based on random utility theory and micro economics in the comprehensive consideration family and work site [4].

The current domestic car ownership prediction method can be divided into the following three categories: one is the extrapolation method, namely using past data to predict the future state, such as time series analysis. The second method is the law of cause and effect, according to the historical data, to find the relation between the variables to be forecasted and related variables, so as to predict the future state, As the regression analysis method. The third method is judgment analysis method, which mainly depends on the expert experience and comprehensive analysis ability to predict future states of [5] [6].

Because of the diversity and complexity of traffic system influencing factors, although the current prediction methods have their advantages, but the data fit is not high, the prediction results have deviation from the actual. Such as time sequence prediction method, using historical data of car ownership to predict the time trend, which does not take relative influencing factors into account. Regression analysis method establish regression model based on some influencing factors that

influence the amount of car ownership, but it only consider several influencing factors, Simulation ability to practical problems is not accurate. With the exception of the above methods, domestic scholars are still using some new artificial intelligence method to predict the car ownership [7].

Genetic algorithm is a nonlinear optimization method. Combining genetic algorithm with neural network, we can have both extensive mapping ability of neural network and rapid convergence properties of genetic algorithm. The paper combine genetic algorithm with neural network, using genetic algorithm to optimize the weights of neural network, determine the initial weight values of neural network. Not only to improve the neural network training speed and generalization ability, but also overcome that the network is easy to fall into local minimum to a certain extent.

II. PRINCIPAL COMPONENT ANALYSIS OF CAR OWNERSHIP

According to influencing factor analysis of car ownership and the feasibility of data, this paper selected some index from 1995 to 2012 at the end of Wuhan city population (person), gross domestic product (100 million yuan), the total industrial output value (100 million yuan), the total import and export volume (million), the total retail sales of consumer goods (100 million yuan), fuel dynamic price index (with a base year of 1995 100), urban residents per capita disposable income (yuan), The total social investment in fixed assets (100 million yuan), total social freight volume (million tons), the whole society of passenger volume (million), pavement, road length (km) and eleven indexes to predict Wuhan city car ownership quantity. These eleven indexes can be divided into population index, economic index, road traffic index.

In this paper X_1, X_2, \dots, X_{11} represent 11 index, using the actual data and principal component analysis can determine the three principal component expressions are as follows:

$$Y_1 = 0.3005X_1 + (-0.4626)X_2 + (-0.1068)X_3 + (-0.0132)X_4 + 0.4888X_5 + (-0.1938)X_6 + (-0.3654)X_7 + (-0.2702)X_8 + 0.1636X_9 + 0.3898X_{10} + 0.0733X_{11}; \quad (1)$$

$$Y_2 = -0.4626X_1 + (-0.1068)X_2 + (-0.0132)X_3 + 0.4888X_4 + (-0.1938)X_5 + (-0.3654)X_6 + (-0.2702)X_7 + 0.1636X_8 + 0.3898X_9 + 0.0733X_{10} + 0.3309X_{11}; \quad (2)$$

$$Y_3 = 0.1535X_1 + 0.0251X_2 + (-0.082)X_3 + (-0.5382)X_4 + 0.138X_5 + (-0.2959)X_6 + (-0.05)X_7 + (-0.2688)X_8 + 0.6625X_9 + 0.2341X_{10} + 0.0284X_{11}; \quad (3)$$

The cumulative variance contribution ratio of the three principal components reached 98.844% can fully reflect the information of the original data.

III. CONSTRUCT PREDICTION MODEL

In this paper, using the global search ability of genetic algorithm to determine the initial weight values of neural network, in instead of random determination method of the initial weights of neural network, the method can solve the slow convergence speed of neural network, easily fall into local minima[9][10].

In this paper, combine the genetic algorithm with neural network to predict the Wuhan city car ownership, the basic idea is as follows:

(1) Principal component analysis of influencing factors, to determine the principal component;

(2) Determine the initial weights of neural network with the genetic algorithm;

(3) To train the neural network, and carry on the prediction of the car ownership.

A. Determine Neural Network Structure

1) Determine the number of neurons

According to the results of PCA, determine three principal components as input data, neuron number of input layer is 3. The output layer is the data of car ownership, neuron number of output layer is 1. This paper identify neurons in hidden layer is 15, thus the structure of neural network is 3 * 15 * 1.

2) Determine the training and learning function

Aiming at the shortcoming of the traditional BP algorithm, this paper select the adaptive learning rate traingda() as the training function of neural network, learngdm as the learning function of neural network.

B. Determine Genetic Algorithm Structure

In this paper, using GAOT toolbox (Genetic Algorithm Optimization Toolbox) to achieve the genetic algorithm.

1) Determine coding mode

This paper select a floating-point number as a method of coding, the individual code string length: $S=M * N + N * L + N + L$. The basic form of coding is shown as follows:

$$W = \{w_1, w_2, b_1, b_2\} \quad (4)$$

w_1 is the connection weights from the input layer to the hidden layer, the length is $M * N$;

w_2 is the connection weights from output layer to the hidden layer, the length is $N * L$;

b_1 is threshold value of the hidden layer, the length is N ;

b_2 is threshold value of the output layer, the length is L ;

M, N, L are neuron numbers of the input layer and hidden layer.

2) Determine fitness function

This paper selects the performance function of neural network as the fitness function of genetic algorithm.

Suppose neural network has n learning samples, then:

$$E = \sum_{i=1}^n (T_i' - T_i)^2 \quad (5)$$

$$f = \frac{1}{E} \quad (6)$$

E is neural network performance function, T_i' is the actual output of the network, T_i is the expected output of the network, F is the fitness function of genetic algorithm.

3) Determine the selection operator

The paper chooses selection probability of operator as q , take the default value is 0.08. The paper chooses arithmetic crossover as the operator of crossover algorithm, and the crossover probability is 0.95.

The paper choose the non-uniform mutation as the mutation operator, the mutation probability is determined as 0.08. The paper chooses the maximum genetic algebra as the termination condition of algorithm, set the maximum genetic algebra as 200 generation.

IV. PREDICTION ANALYSIS

Based on the genetic algorithm and the neural network structure, using the matlab7.0 to program and select data from 1995 to 2009 as training data, select data from 2010 to 2012 as test data of the network prediction performance.

1) The genetic algorithm optimization

By using genetic algorithm for global optimization, get an initial weights, threshold value matrix of neural network, the next is evolutionary graph of error sum of squares.

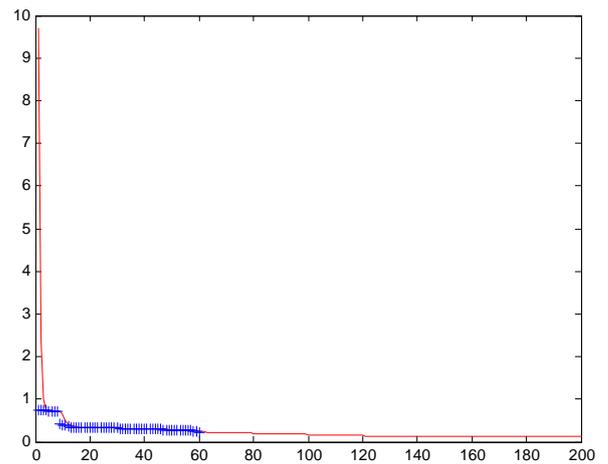


FIGURE 1. EVOLUTIONARY GRAPH OF ERROR SUM OF SQUARES

Error sum of squares show the performance of neural network, solve reciprocal to error sum of squares and get the fitness in the genetic evolution process, as shown below:

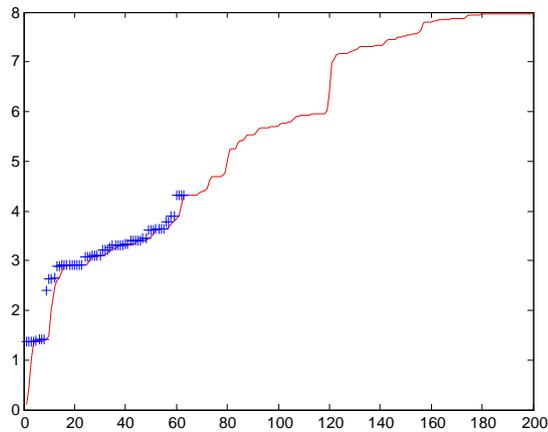


FIGURE II. THE GENETIC EVOLUTIONARY FITNESS MAP

When in the evolution to the 160 generation, fitness gradually become stable, when in the evolution to 200 generation it was obtained the initial weights, threshold value matrix of neural network.

2) Train neural network and test prediction

Put the initial weights and threshold matrix be optimized by the genetic algorithm in neural network, and set the network training accuracy as 10^{-3} , neural network training process is shown below:

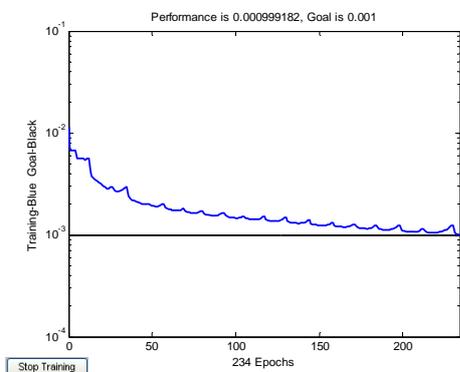


FIGURE III. BP NETWORK TRAINING

Through 234 times training, the neural network begin to convergence, the process spend 19.9 seconds. Use the trained network to predict 2010, 2011, 2012 in Wuhan city car ownership, can get prediction results:

TABLE I. PREDICTION PERFORMANCE TABLE

type	2010	2011	2012
Practical value	370609	418667	484111
Prediction value	380653	424654	485805
Absolute error	2.71	1.43	0.35
The mean absolute error	1.49		

As in 2010 car ownership growth rate dropped from 23.2% in 2009 to 10.8%, change is larger, the prediction result is too large. Prediction results mean absolute error is 1.49%, meet the requirements of precision.

V. CONCLUSION

Genetic algorithm is a nonlinear optimization method. Combining genetic algorithm with neural network, we can have both extensive mapping ability of neural network and rapid convergence properties of genetic algorithm. The paper combine genetic algorithm with neural network, using genetic algorithm to optimize the weights of neural network, determine the initial weight values of neural network. Not only is to improve the neural network training speed and generalization ability, but also overcome that the network easily fall into local minimum to a certain extent. The paper take Wuhan City car ownership as the research object, and an example show that the method can satisfy the prediction demand.

ACKNOWLEDGMENT

The paper is supported by key project of Humanities and social sciences of the Education Department of Hubei Province (11D115).

REFERENCE

- [1] Mei Lam Tam, William H.K. Lam. Balance of car ownership under user demand and road network supply conditions—case study in Hong Kong. Urban Planning & development, March 24-36,2004.
- [2] M.L. Tam, William H.K. Lam. Maximum car ownership under constraints of road capacity and parking space. Transportation Research Part A34,2000.
- [3] Tomas, Soren. Dynamical models of car ownership at the Household level. Working Paper, 2004.
- [4] J.H. Rich, O.A. Nielsen. A microeconomic model for car ownership, residence and work location. Paper for European transport conference PTRC, Cambridge,2001.
- [5] Ma yali. Study on prediction of our country car ownership[J]. traffic science & technology, 2007.1
- [6] Gong huawei, Prediction of car ownership based on econometric models [J]. Journal of transportation systems engineering and information, 2005.5
- [7] Hua zhigang. Prediction of Hunan province car ownership based on BP artificial neural network[J], Shanxi science and technology2005.2
- [8] He ming. Application of principal component analysis and BP neural network method in the prediction of car ownership [J]. computer and communications, 2007.4
- [9] Cheng cisheng,Su lingli. prediction of car ownership in China based on Genetic algorithm BP [J]. Journal of Changsha Institute traffic, 2005.6
- [10] Zhu bin,He changsheng. Prediction of Sichuan civilian vehicle ownership based on G M D H method [J]. modern management science, 2006.6
- [11] Zhang xiaochun,Zhang zhizhe. discussion on prediction method of City vehicle ownership [J]. technology and economy, 2006.4