Forecasting the railway freight volume in China based on combined PSO-LSTM model

Yue-ying Qiu¹, Qiong Zhang² and Ming Lei*¹

¹Tianjin University of Technology, Tianjin 300384, China
²School of International Business of Tianjin Foreign Studies University, Tianjin 300204, China

Leiming0901@tjut.edu.cn  834473682@qq.com

Abstract. Considering the shortcomings of the current railway freight volume prediction model, this paper proposes a railway freight volume forecasting model based on PSO-LSTM. LSTM can learn long-span nonlinear time series effectively, and it is highly efficient in training forecasting of railway freight volume data. In order to improve the accuracy of parameter selection, PSO is used to optimize neural network model. The simulation results show that the PSO-LSTM model proposed has lower prediction error and higher prediction accuracy than the traditional LSTM model GA-LSTM model. This indicates that this model is an effective prediction model for railway freight volume.

1. Introduction

In recent years, the continuous adjustment of China's industrial structure has made the slow growth of domestic railway transport volume. Under the competition of other transportation modes, the market share of railway transportation has decreased year by year. Facing the fiercely competitive transportation market, achieving accurate forecast of railway freight volume can provide reasonable and effective construction planning for railroads, and also play an important role in increasing the proportion of railway transportation in the market.

The prediction methods of railway freight volume can be divided into two categories. One is to analyze the characteristics of railway transportation directly according to experience; the other is to collect relevant historical data and establish corresponding models for prediction. For example, time series model [1], gray model [2-3] and machine learning models, including SVM [4-5], BP [6], RBF [7] and other neural network models [8-9]. The currently proposed railway freight volume forecast models have their own advantages, but in summary, these models have the following shortcomings: (1) Single prediction model has limitations. (2) The input indicator is historical railway freight volume, without considering the influence of relevant factors. (3) Most prediction models are shallow models, which have limited ability to learn the characteristics of time series. For example, the gray model has poor stability in long term training. The support vector machine is too sensitive to the parameters. The BP neural network is easy to fall into a local minimum.

For solving these problems, this paper proposes a combined forecasting model of railway freight volume based on PSO-LSTM model. The input index selects the relevant indicators of railway freight volume and historical data. LSTM is a deep learning model that can effectively process long-span time series and automatically learn data and tap deeper features. PSO algorithm performs adaptive selection
of neural network parameters in order to obtain more accurate prediction results and effective reference basis for railway projects.

2. Algorithms introduction

2.1. Long short-term memory network

Long short-term memory (LSTM) network [11] is a deep learning model for the problem of gradient vanishing of recurrent neural networks (RNN). It has good ability to deal with non-linear data, and is often used for long-term dependencies. Compared with RNN, LSTM introduces complex gate structures to process information, and storage units to determine the forgetting and retention of data. The internal structure of LSTM is as Figure 1. A neuron is composed of input gate $i_t$, forget gate $f_t$, and output gate $o_t$, to conduct selective learning and store information.

The specific formula are as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \cdot \tanh(c_t)$$

where $x_t$ is the input of the cell at time $t$, $h_t$ is the output at time $t$, $\sigma$ is the sigmoid function and $\tanh$ is the hyperbolic tangent function. The terms $W_f, W_i, W_c, W_o$ are the weight matrices. The terms $b_f, b_i, b_c, b_o$ are the corresponding offset vectors.

![Figure 1. The internal structure of LSTM.](image)

2.2. Particle swarm optimization algorithm

Particle swarm optimization (PSO) algorithm derived from the bird flock preying behavior. The information sharing is achieved by simulating the bird's foraging behavior trajectory to obtain the best result of group cooperation in the foraging process of the entire group. PSO uses massless particles to replace the birds in the bird swarm. Only considering the speed and direction of particle movement to find the optimal value of individual particles in the search space, and share information with other particles. The optimal value of all particles is taken as the optimal solution of the global solution space. The remaining particles adjust their moving speed and direction according to this optimal solution to realize the deduction process from disorder to order in the solution space.

In the search space, each particle can be regarded as the potential solution of the global optimal solution, and each particle is given two attribute information: velocity and position. According to the flight experience information of other particles to change their own direction, speed and position. Update the own speed and position according to the current optimal value $g_{best}$ and its own optimal value $p_{best}$ in the search space:

$$v_i = v_i + c_1 r_1 (p_{best} - x_i) + c_2 r_2 (g_{best} - x_i)$$

$$x_i = x_i + v_i$$
where $v_i$ is the speed of particles in the optimization process, where $x_i$ is the position of particles in the optimization process, $c_1, c_2$ are the learning factors, $r_1, r_2$ are the random numbers between 0 and 1.

Increase the inertia weight $\omega$ in the individual velocity of formula (7) to form the current standard particle swarm optimization algorithm:

$$v_i = \omega v_i + c_1 r_1 (p_{best} - x_i) + c_2 r_2 (g_{best} - x_i)$$ \hspace{1cm} (9)

3. Hybrid PSO-LSTM prediction model

The detailed steps of the model are as follows:

Step 1: Collecting the original data and pre-processing the input indicators: the samples are normalized by Min-Max, and the data is mapped to the range of (0, 1).

Step 2: Divide the data set into a training set and a test set in proportion.

Step 3: Initialize the individual particle speed and position of the particle swarm algorithm, the training set is put into the model for training.

Step 4: The fitness function is set as the error between the prediction result and the real value, and the local optimal value and the global optimal value are continuously updated according to the fitness value.

Step 5: Save the model after reaching the maximum number of iterations.

Step 6: Put the test set into the saved model for prediction, and de-normalize the output to get the prediction result.

4. Experimental analysis

4.1. Experimental data

The experiment was conducted in China with sample data from the monthly data released by the National Bureau of Statistics of China. Data information includes total import and export value, resident consumption index, crude steel output, crude oil output, road freight volume, waterway freight volume and civil aviation freight volume. The data covers a total of 120 items from January 2010 to December 2019. The first 80% of the data set were used as training samples for model training. The remaining 20% were used as test samples for prediction comparison.

4.2. Simulation conditions

The parameters of the particle swarm optimization algorithm are set as follows: the inertia weight is 0.8, the learning factors are 1.5, the particle swarm is 10, the maximum number of iterations is 20, and the experiment termination condition is to reach the maximum number of iterations.

4.3. Predictive performance evaluation

$MAPE, RMSE,$ and $MAE$ are used as indicators to evaluate the prediction performance of the model. The calculation equation is as follows:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - y_i'}{y_i} \right| \hspace{1cm} (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_i')^2} \hspace{1cm} (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - y_i'| \hspace{1cm} (12)$$

4.4. Experimental results and analysis

In this paper, PSO is used to optimize the LSTM parameters which mainly include the learning rate, the number of hidden layer neurons and iterations. The value of these parameters will directly affect the final training result. PSO results are shown in Table 1.
Table 1. The result of the parameter optimization.

| iterations | learning rate | The first layer of hidden neurons | The second layer of hidden neurons |
|------------|---------------|----------------------------------|----------------------------------|
| 1903       | 0.00079       | 22                               | 10                               |

In order to verify the effectiveness of the PSO in this experiment, this paper selects the traditional LSTM model and the genetic algorithm and LSTM combination model for comparative experiments. The results of the three prediction models are shown in Table 2 and Figure 2. It can be seen from the figure that the PSO-LSTM model has the highest fitting degree, indicating that its prediction performance is the best. The second is the GA-LSTM model. According to Table 2, compared with the traditional model, the prediction errors of the PSO-LSTM model are reduced by 3.21%, 1403.58 and 1114.73, respectively. The prediction errors of GA-LSTM model prediction were reduced by 2.56%, 2680 and 938.88. This shows that the intelligent optimization algorithm can effectively find higher-precision prediction results than traditional methods.

Table 2. Prediction performance comparison of methods.

| model      | MAPE     | RMSE     | MAE      |
|------------|----------|----------|----------|
| LSTM       | 0.067056 | 2848.822 | 2327.882 |
| GA-LSTM    | 0.041498 | 1678.8189| 1389.0041|
| PSO-LSTM   | **0.034944** | **1445.2468** | **1213.1511** |

Figure 3 shows the fitness of PSO-LSTM model and GA-LSTM model in prediction. It can be seen that the fitness of the PSO-LSTM model at step 7 is minimized and the result is 0.015851. The fitness of the GA-LSTM model at step 14 is minimized and the result is 0.016785. In contrast, PSO algorithm shows better optimization ability and is superior to GA algorithm in optimization degree, convergence speed and prediction accuracy.

Figure 3. The fitness of two optimization algorithms.
5. Conclusion

This paper proposes a prediction model of railway freight volume based on PSO-LSTM model. The MAPE of the PSO-LSTM model is 3.49%, which has high prediction performance. Compared with the traditional LSTM model, the prediction error of the PSO-LSTM model is reduced by 3.21%, 1403.58 and 1114.73. It can be seen from the results that PSO can effectively improve the prediction accuracy of the neural network and is an effective optimization algorithm.

In addition, due to the limitations of data acquisition, it is necessary to find a more comprehensive study on the influencing factors of railway freight volume and the optimization of the model in future research, and additional modeling and discussion are needed to further improve upon the prediction accuracy of the model.

References

[1] Wang Z H and Wang Q Y. (2013) Seasonal adjustment model of china railway monthly passenger traffic volume based on spring festival factors. Journal of the China Railway Society, 35(7): 9–13.

[2] Zuo X Y and Huang X J. (2016) Prediction of Railway Freight Volume Based on Grey Prediction Model in China. Logistics Sci-Tech, 8:82-84.

[3] Zhang C and Zhou X F. (2007) Prediction of Railway Freight Volumes Based on Gray Forecast-Markov Chain-Qualitative Analysis. Journal of the China Railway Society, 29(5): 15-21.

[4] Geng L Y and Chen L H. (2016) Forecast on railway traffic volume using mixed-kernel LSSVM optimized by FOA. Application Research of Computers, 34(2): 409-412.

[5] Zeng M, Lin L and Cheng W M. (2013) Research of regional freight volume forecasting based on LIBSVM and time series. Computer Engineering and Applications, 49(21): 6-10.

[6] Zhu W T. (2014) Forecasting railway freight volume based on improved BP neural network model. Journal of Shijiazhuang Tiedao University, 27 (2): 79–82.

[7] Song S M, Kuang W Z, Xu L and Chang F. (2016) RBF neural network applied to prediction of railway freight volumes. Railway Computer Application, 26(1): 47-51.

[8] Zhang T and Zhang J L. (2014) Study on the Forecast of Freight Volume Based on GRNN. Logistics Sci-Tech, 10: 138-141.

[9] Sun Y, Lang M X, Wang D Z and Liu L Y. (2014) A PSO-GRNN model for railway freight volume prediction: Empirical study from China. Journal of Industrial Engineering and Management, 7: 413-433.

[10] Hochreiter S and Schmidhuber J. (1997) Long short-term memory. Neural Computation 9(8): 1735-1780.