Research on Premium Income Prediction Based on LSTM Neural Network

Li Diao
School of Insurance,
Central University of Finance and Economics, Beijing, China

Ning Wang
School of Computer and Information Technology,
Beijing Jiaotong University, Beijing, China

ABSTRACT
As one of the four financial pillars, insurance has the functions of risk diversification, loss compensation, financing and social management. It is of great practical significance to predict the level of premium income in the new normal of economy. In this paper, long short-term memory (LSTM) neural network was innovatively applied to the study of premium income prediction. The monthly data of China's premium income from January 1999 to October 2019 was selected for prediction, and the prediction results were compared with BP neural network. The results show that LSTM model can accurately predict premium income, and its performance is better than BP neural network.

Keywords: Premium income prediction; LSTM neural network; Machine learning; BP neural network.

INTRODUCTION

Insurance has the functions of dispersing risks, compensating losses, financing and social management, which is conducive to ensuring people's basic living standards and maintaining social stability. Since the resumption of China's insurance industry, the scale of premium income has maintained a rapid growth, from 460 million yuan in 1980 to 3801.66 billion yuan in 2018. China has also become the third largest premium income country in the world. For the global insurance market, the emerging market dominated by China has become the main driving force of global premium growth. For the domestic economic system, the insurance industry is also one of the four financial pillars and an important factor to promote the stable growth of GDP. At present, China's economy has entered a new normal stage. The economic growth has slowed down, and the driving force of development has shifted from factor driven and investment driven to innovation driven. Under the new situation, the prediction of the scale of premium income has become a topic of concern to all walks of life.

Long short-term memory (LSTM) neural network [1] is one of the classical models of deep learning, which has obvious advantages in mining the long-term relationship of time series data. Financial time series data usually have the characteristics of uncertainty and high noise, and the relationship between variables is also dynamic [2]. The traditional time series model is hard to achieve accurate prediction because of its strict requirements and poor applicability [3]. LSTM neural network model can use its own network structure to "remember" the information that is a long time from now and save the previous state, so as to describe the relationship between the current data and the previous input data, thus affecting the prediction trend of the subsequent data.

LSTM neural network model has been applied to many fields, and it has been proved that it can be used in the prediction of time series. In the field of finance, the prediction of stock market using LSTM neural network model has achieved good results, which shows that LSTM neural network can accurately reflect the time series characteristics of stock market.

MATERIALS AND METHODS

Data Acquiring and Processing

Considering the availability of data, we select the monthly data of China's premium income from January 1999 to October 2019 for training and testing.

We observed the trend of premium income time series data and found that premium income fluctuated in general, showing obvious seasonality, as shown in Figure 1. Therefore, we choose the total premium income and month as the input characteristics of the LSTM model.
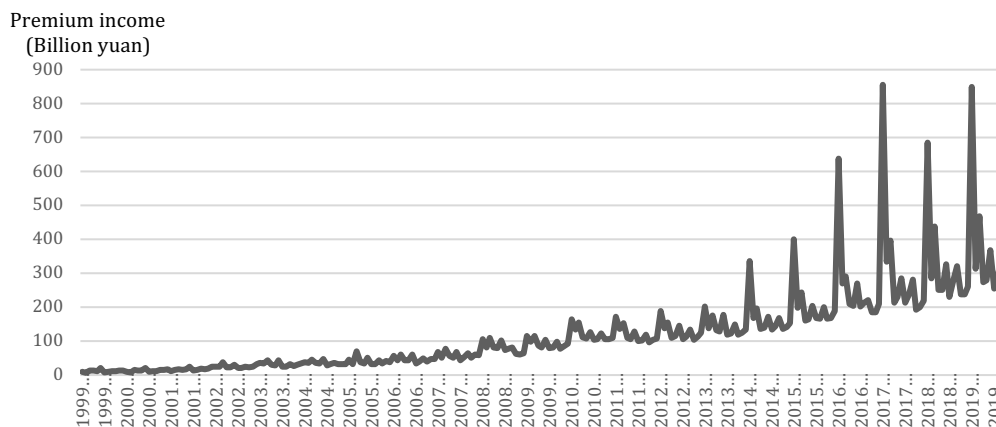

Figure1. Premium income time series data.

LSTM Neural Network

LSTM neural network was proposed by Sepp Hochreiter et al [4] in 1997, which is an improved recurrent neural network (RNN). After that, Graves extended it to many fields, such as unconstrained handwriting recognition [5], speech recognition [6] and handwriting generation [7]. As for the network structure, the LSTM neural network improves the nodes in the hidden layer on the basis of the standard RNN. In the standard RNN model, the hidden layer has only one state, h, which is used to remember the short-term state. In the LSTM neural network, a new state, named c, is added to the original state h to memorize the long-term state. In this way, the long-term dependence of RNN can be well solved.

LSTM neural network selectively controls the input of information by using three "gate" structures, i.e. input gate, forget gate and output gate, so as to realize the function of forgetting or memorizing information and affect the state of each time in the neural network. Its structure is shown in Figure 2. The "gate" structure is to get the value between [0,1] through the calculation of activation function, and control how much information can pass through the structure by multiplying with the original information, so it is similar to the function of a "gate". When the value is 0 (gate closed), the input value is 0; when the value is 1 (gate open), all information is input.
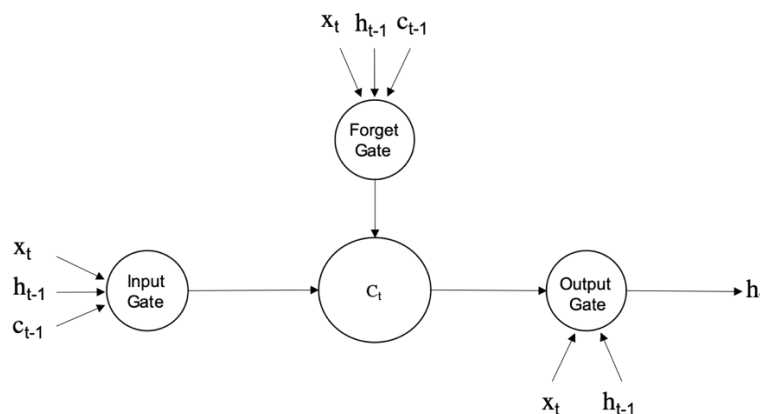
Figure2. LSTM neural network structure.

The function of forgetting gate is to make the neural network forget the useless information before, which is determined jointly according to the current input x_t, the last output h_{t-1} and the last state c_{t-1}. The function of input gate is to supplement the latest information and fill in the forgotten part before which is also determined by x_t, h_{t-1} and c_{t-1}. The function of the output gate is to output h_t which is determined by x_t, h_{t-1} and the current state c_t.

The LSTMCell function provided by TensorFlow framework is used to build the LSTM neural network structure. The three "gate" structures mentioned before have been encapsulated in the structure. Each use of this function is equivalent to defining an LSTM layer. In the process of model training, we use mean square error (MSE) as the loss function of the model, and use adaptive moment estimation (Adam) optimizer to solve the minimum loss. On this basis, we define different layers of LSTM, the number of neurons in each layer of LSTM and the training times of the model, and carry out sufficient experiments on the model from many aspects.

Model Evaluation
In this work, the mean absolute percentage error (MAPE) is selected as the evaluation index to measure the accuracy of prediction, and the standard deviation of absolute percentage error (SDAPE) is selected to measure the stability of prediction.

$$\text{MAPE} = \frac{1}{N} * \sum_{i=1}^{N} \left| \frac{T_i - A_i}{A_i} \right| \qquad (1)$$

$$\text{SDAPE} = \sqrt{\frac{1}{N} * \sum_{i=1}^{N} \left(\left| \frac{T_i - A_i}{A_i} \right| - \text{MAPE} \right)^2} \qquad (2)$$

in which, T_i represents the predicted premium, and A_i represents the true premium. If the MAPE is small, the difference between the predicted value and the real value is small, indicating the better predictions. The smaller the SDAPE, the better the stability of the predictions.

RESULTS
Parameters Adjustment of LSTM
The parameters of LSTM mainly include the number of network layers, the number of neurons in each layer and the number of training.

In order to find the optimal parameters for the sample data, firstly we fix the number of training for 3000, and experiment on different numbers of layers and neurons to minimize the MAPE and SDAPE. Specifically, we set the range of network layers as 1, 2, 3, 4, and the range of

neurons as 10, 30, 50, 70, 100. The experimental results are shown in Table 1. We find that the prediction result is the best when the number of network layers is 3 and the number of neurons in each layer is 50.

Table 1. Experimental results of different numbers of network layers and neurons.

Layers / Neurons		1	2	3	4
10	MAPE	0.099185476	0.12156257	0.10246977	0.09956268
	SDAPE	0.07021626	0.15781445	0.12668743	0.11180736
30	MAPE	0.12215267	0.09383439	0.11396967	0.08259218
	SDAPE	0.11213681	0.09818877	0.12834039	0.11047624
50	MAPE	0.09144791	0.07981141	**0.056452**	0.08808873
	SDAPE	0.09358854	0.10763179	**0.076293**	0.14842728
70	MAPE	0.08916558	0.08867701	0.10281804	0.11315149
	SDAPE	0.11749351	0.11247249	0.15138062	0.15963035
100	MAPE	0.14643573	0.10835095	0.09412221	0.10327301
	SDAPE	0.20046314	0.16421714	0.12607211	0.17061168

Next, we keep the number of network layers as 3 and neurons as 50 to find the best training times. Set the selection range as 100, 300, 500, 700, 1000, 1500, 2000, 3000, 5000, 7500, 10000. The experimental results are shown in Figure 3. Considering both the accuracy and stability, we choose the result of 3000 times training as the final result.
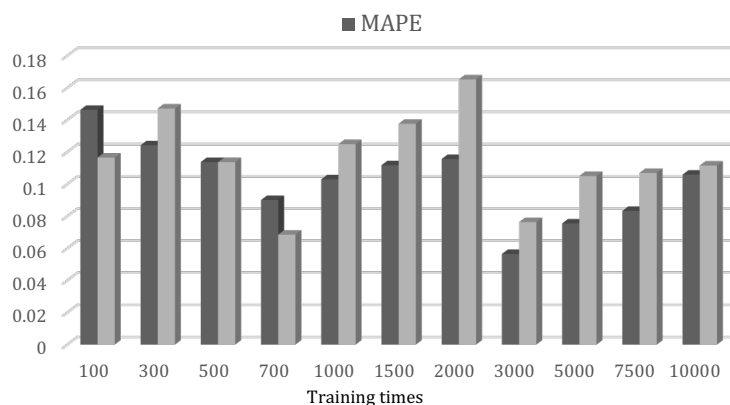

Figure3. MAPE and SDAPE of different training times.

Model Comparation

BP neural network is a machine learning method commonly used in prediction. Here we use both LSTM model and BP neural network to predict the premium income from October 2018 to September 2019 and compare the results with real data. As shown in Figure 4, we find that the predicted result of LSTM is closer to the real data than that of BP neural network, indicating that LSTM model is a more suitable deep learning method for prediction.
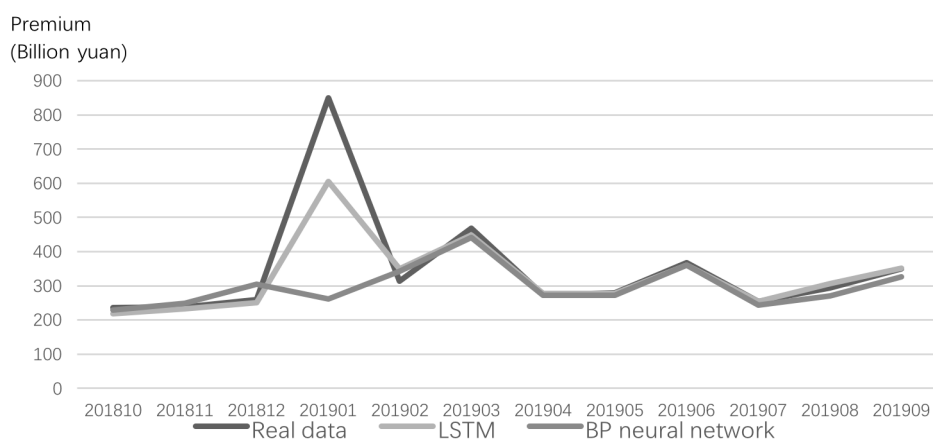
Figure4. Results of LSTM and BP neural network.

CONCLUSION

In this paper, we firstly abstract the prediction of premium income as a time series prediction problem, and describe in detail the deep learning model LSTM neural network which is widely used in the time series prediction problem in recent years, affirming that LSTM model has advantages in the time series prediction.

Secondly, we select the monthly data of the original premium income of the property insurance companies from January 1999 to September 2018 as the training data, and train the LSTM model and BP neural network in turn.

Finally, the above two models are tested on the data from October 2018 to September 2019, and MAPE and SDAPE are selected to evaluate the effect of the model. Compared with the traditional BP neural network model, the LSTM model can better remember the historical information of time series data, which is suitable for premium income prediction.

Therefore, with the continuous development of insurance market in the future, the accumulation of premium income data will become larger and larger. Then the prediction results will be more accurate and reasonable when using LSTM model to predict premium income. In addition, we can consider adding exogenous seasonal factors to control variables as features, so that LSTM model can better identify the seasonality of data and improve the accuracy of prediction.

References

Duan Y J, Lv Y S, Wang F Y. Travel time prediction with LSTM neural network[C]// 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2016.

YANG H M, PAN Z S, BAI W. Review of Time Series Prediction Methods[J]. Computer Science, 2019, 46(01):28-35.

WANG J T. Financial Time Series Prediction Based on LSTM Hybrid Model [D]. Zhengzhou: Zhengzhou University,2019.

HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation,1997, 9(8):1735-1780.

Graves A, Liwicki M, Santiago Fernández, et al. A Novel Connectionist System for Unconstrained Handwriting Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(5):855-868.

GRAVES A, JAITLY N, MOHAMED A R. Hybrid speech recognition with Deep Bidirectional LSTM[C]// Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on. IEEE, 2013.

GRAVES A, SCHMIDHUBER J. Offline Arabic Handwriting Recognition with Multidimensional Recurrent Neural Networks[C]// Advances in Neural Information Processing Systems 21, Proceedings of the Twenty-Second Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 8-11, 2008.