



A Comparative Study of Stock Forecasts by LSTM and RNN Neural Networks

Siqi Wu

The Affiliated High School of Peking University, Beijing 100086, China
Email: wusiqi@i.pkuschool.edu.cn

Abstract: In the financial market, stocks have always played a very important role. The economic situation is closely related to the stock market. If people could make an effective prediction of the future trend of the stock market, it is of great significance to prevent the financial crisis and guide the investment direction. From this point of view, this paper uses artificial intelligence to obtain a feature representation through the analysis of massive stock price data, to predict the future stock price. Specifically, it uses recurrent neural network (RNN) and long short-term memory networks (LSTM) to predict the stock trend. Under the same experimental conditions, the experimental results predicted by the two methods are compared and analyzed. The experimental results show that RNN has an effective prediction for the trend of stock price, but LSTM has a better prediction accuracy, especially in the short-term prediction.

Keywords: LSTM, RNN, stock-price, prediction

1. Introduction

At present, the development of Chinese market economy is getting better and better, and the stock market has gradually become an integral, important part. As a large number of people have joined the stock market investment, making an effective prediction of the coming situation of the stock market has become a topic of continuous concern. Especially for some private equity firms or government financial institutions, an effective forecast can help them develop and formulate a better winning strategy.

In recent years, with the breakthrough of artificial intelligence technology, the use of artificial intelligence deep learning technology to make a reasonable and effective prediction of stock trends has become an engineering project. Because the stock price trend has a time series, combined with artificial intelligence technology in the time series representation, this paper employs the RNN and LSTM to predict the niche of stocks, and compares the results of these two scenarios under the same experimental conditions.

In this paper, the TensorFlow framework is used to construct the above two model structures, to predict the closing price of the stock on the next day, after many parameters debugging, to find the most suitable parameters, and to train an optimal model.

2. Related theoretical basis

2.1 RNN structure model

Circulatory neural network is a kind of neural network that can model time correlation, the original purpose of which is to solve the problem of long-term dependence of words in the field of natural language processing. In natural language processing, words are often converted to the form of word vectors, and correlations exist between different word vectors called sequence correlations. Figure 1 is a structural model diagram of RNN, and it is clear from the figure that the input of the entire neural network is x , and U is a weight from the input layer to the hidden layer of the data, and V and W are also a weight matrix, but represents a weight of the data from the hidden layer to the output layer and a weight of various memory units S to the hidden layer respectively, and the entire neuron output is O . On the right part of Figure 1 is a specific expansion of a circular neural network, with time $t-1$ series, t and $t+1$, different inputs in different sequences, corresponding to x_{t+1} , x_{t-1} , x_t , through the weight matrix, u , v , w can generate the corresponding output, o_{t-1} , o_t , o_{t+1} , characterizing the correlation of time.

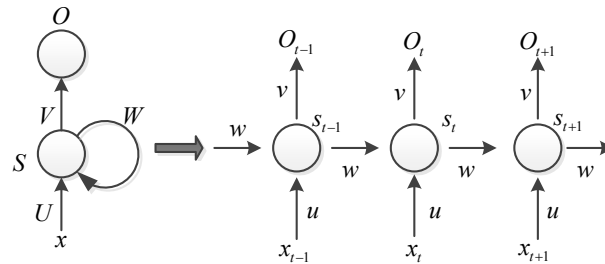


Figure 1. RNN model structure diagram

2.2 LSTM neural network model

LSTM is also a kind of circular neural network in essence, and the most significant difference from RNN is that the hidden unit in the RNN structure is replaced by a more special memory unit, which can greatly alleviate the problem of gradient disappearance caused by the RNN structure in the training process of the model. Gradient disappearance refers to the phenomenon that the model's accuracy decreases as the hidden layer increases during the training process of the model in the training process. In contrast, Convolutional Neural Networks (CNN) can solve this problem well. And inspired by this, LSTM incorporates the principles of convolutional neural networks, increasing the model's ability to express features. In predicting the trend of a particular stock, the robustness of the model in this way is greatly enhanced and is important for extremely sensitive data such as stocks. Shown in Figure 2, the model diagram of LSTM, with three gating units, f_t , i_t and o_t representing the forgetting door control unit, the output gate control unit and the hidden layer control unit, C_t representing updated memory cells, x_t representing inputs, \tanh and δ representing the "tanh" activation function and the "sigmoid" activation function, respectively.

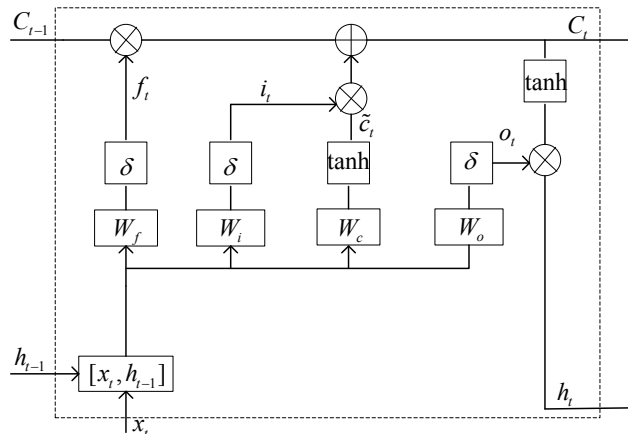


Figure 2. LSTM Model Structure Diagram

Formula (1) is the parameter update representation of the LSTM network, where "" represents convolution operators; "*" and "o" represents dot product, " δ " represents the sigmoid activation function; \tanh is also an activation function. Activation functions are used to further improve the expressive power of the overall model.

$$\begin{cases} f_t = \delta(W_f \circ [h_{t-1}, x_t] + b_f) \\ i_t = \delta(W_i \circ [h_{t-1}, x_t] + b_i) \\ c_t = f_t * C_{t-1} + i_t * \tanh(W_c \circ [h_{t-1}, x_t] + b_c) \\ o_t = \delta(W_o \circ [h_{t-1}, x_t] + b_o) \\ h_t = o_t * \tanh(c_t) \end{cases}$$

(1)

3. Details and results of the experiment

3.1 Experimental data

The experimental data in this article are primarily obtained by a crawling program, obtaining some shares and stock market data per share trading in Shanghai and Shenzhen Stock Markets from January 25, 1999 to July 15, 2020 as a dataset, characterized by time, closing price, opening price, day trading high, low price, current transaction maximum price, minimum price, price change, price change amount, volume, turnover, and amplitude, and labeling the closing price of stocks the following day. At the same time, 90% of the overall data is taken as a training set and 10% as a validation set. The test set is also used with stock data from Nvidia (NVDA) and CATL from July 16, 2020 to July 16, 2021.

3.2 Experimental environment

All of the experiments in this article are based on the TensorFlow deep learning framework, and the software environment includes: the operating system is Windows 10, the TensorFlow version is 1.9.0, the CUDA version in the GPU Acceleration Library is 9.0, and the CUDNN version is 7.1. Hardware configurations include, CPU for AMD Ryzen 7 2700, and training and reasoning use the RTX 2070 8G model GPU, with 16G,512G SSDs.

3.3 Normalization of data

Before training the data samples, they also need to be normalized. Normalization means that each eigenvector is scaled to the same range. All data should be uniformly zoomed into the range of [0,1] or [-1,1], due to the great differences between the above stock characteristic attribute values. If the data is directly trained, it would be, as a result, the network convergence speed slows down and the training process oscillates violently, and these property values unified into a small interval can be greatly alleviated. The standardized method used in this model is z-score standardization, as shown in formula (2). x represents the input sample property value, μ represents the average value of the property, σ represents the standard deviation of the property value, and the standardized result is used x' to represent the value of the property.

$$x' = \frac{x - \mu}{\sigma} \tag{2}$$

3.4 Feature learning and evaluation indicators

In the past, when machine learning was used to accomplish a task, the sample feature information required was primarily manually designed, so this process was also called "feature engineering". Because of the decisive relationship between the generalization performance of the model and the feature design, there is a great deal of uncertainty in the method of designing the features manually, so the results of the model often fall short of the expected requirements. In contrast, "feature learning" is to dig the characteristic information from its own point of view. According to the general law of statistical learning, the more model parameters, the greater the amount of data will be involved in training. In this paper, when the data set is used to train RNN network and LSTM network, the number of iterations is set to 50, and a loss function is used as a cross-entropy loss function, which, as shown in the formula (3) X_i represents a predicted value, X_i^* represents an actual value, which is \log with e bottom, $J(\theta)$ is to indicate the similarity between the actual output and the desired output, θ is the network layer to learn parameters .

$$J(\theta) = X_i^* \log(X_i) + (1 - X_i^*) \log(1 - X_i) \tag{3}$$

In order to evaluate the performance of the prediction model, the evaluation index used in this paper is custom error. The error of relative real value is more fully expressed for the overall model performance, using the following formula (4) as the error evaluation index of y the model, which represents the true value, and "predict" represents the forecast value, and it is stipulated that the smaller the error is, the higher the prediction accuracy of the model is.

$$error = \frac{|predict - y|}{y} \tag{4}$$

3.5 Results analysis

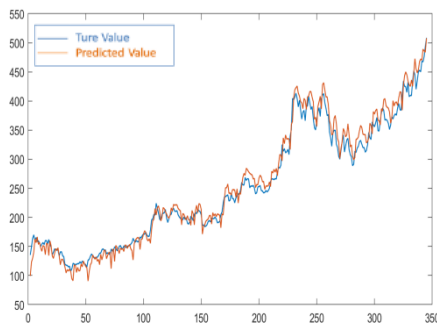
Table 1 shows the error between the stock price forecast and the actual situation of different companies under the two

model structures of LSTM and RNN. And it is clear from the figure that the LSTM approach is generally smaller than the RNN's method error, and the RNN's error is 4% to 6% higher than that of LSTM. At the same time, relatively speaking, the stock price forecast error in the CATL is smaller than Nviday's, and the error under the best result is only 10.4%, which may be a deviation of different stocks, resulting in a certain difference in forecasting.

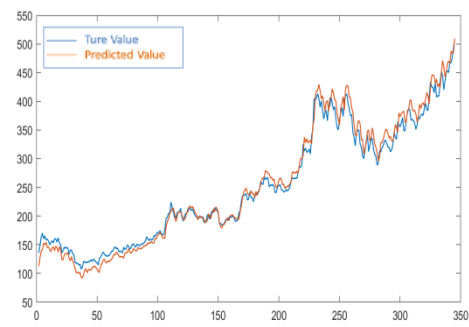
Table 1. Share price forecast errors under different models

Experimental methods	The subject of the experiment	Error (%)
LSTM	CATL	10.4
RNN	CATL	15.3
LSTM	NVDA	12.6
RNN	NVDA	19.3

As shown in Figure 3 below, a comparison of stock price forecasts for the CATL shows that the LSTM and RNN models are broadly consistent with the trend of stock price forecasts in the CATL, as well as the trend of real share prices, illustrating the relative effectiveness of the two approaches in time series forecasting of share prices. On closer inspection, however, the predicted movement of the RNN is more volatile than that of LSTM, largely due to the RNN model structure itself. In addition, in a relatively short period of time, the prediction results between the two different models are close to the real results, but in a longer period, the prediction results and the real results are much different, which also reflects the characteristics of the time series.



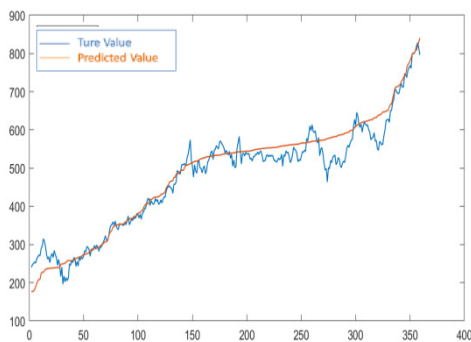
(a) LSTM predicts the CATL



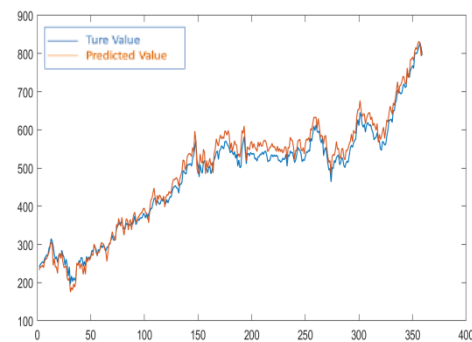
(b) RNN predicts the CATL

Figure 3. Comparison of stock price forecasts for the CATL

As shown in Figure 4 below, NVDA's share price forecast comparison chart shows that the RNN model is not very good for the stock price forecast, but the overall trend is basically correct. It may be that the RNN model in the training process was affected by gradient-vanishing, which results in anomalies that weights cannot be reasonably predicted according to the established route, and LSTM with the gated thinking, to some extent, alleviates the loss of gradient phenomenon.



(a) LSTM predicts NVDA



(b) RNN predicts NVDA

Figure 4. Comparison of INVISTA's share price forecasts

4. Conclusion

This article analyzes two stocks in Shanghai stock market and U.S. stock market and uses artificial intelligence time series models to predict their stock prices and integrates information by extracting data features to achieve the prediction of the closing prices of the selected stocks on the next trading day. The experimental results show that both LSTM and RNN models can basically describe the future stock trend, but the LSTM model is better in terms of accuracy, especially for short-term stock price prediction.

References

- [1] Thaler H. Mental accounting and consumer choice. *Marketing Science*, 2008 (1): 12-25.
- [2] Richard Thaler. Mental Accounting and Consumer Choice[J]. *Marketing Science*, 1985: 4(3).
- [3] Daniel Kahneman. Consumers' "Mental Accounting" in Responses to Unexpected Price Savings at the Point of Sale[J]. *Marketing Intelligence & Planning*, 1981: 4(24).
- [4] Tversky Kahneman D. Social Norms and Mental Accounting Processes: A Virtual Community Context[J]. *American Marketing Association*, 1984: (10).
- [5] Zhang Changhua A deep learning model for financial market forecasting: the FEPA model University of Electronic Science and Technology, Chengdu, 006.
- [6] Su Zhi, Lu man, Li Dexuan Financial Empirical Applications of Deep Learning: Dynamics, Contributions, and Prospects. *Financial Research*, 2017 (05): 111-126.
- [7] Li Chunwei Stock Intelligence Predictions Based on Neural Networks and Genetic Algorithms. Northwestern University of Technology, Xi'an, 2006.
- [8] Liang Xia. Application of artificial neural network with self-correction function in stock rolling prediction. *Computer Application Research*, 1999 (1): 76-80.