Coal Price Prediction based on LSTM

Sitong Pan*

The University California, San Diego Revelle College, 9500 Gilman Dr, La Jolla, CA, 92093

*Corresponding author: sitong_pan6357@foxmail.com

Abstract. The price of coal has always been the focus of the government and is affected by various factors. This paper uses days, weeks and months as time units to establish a Long-Short Term Memory Recurrent Neural Network Model (LSTM RNN) to predict the change trend of coal prices. The accuracy and sensitivity of the model are tested, and the results show that the model is stable and has high accuracy.

1. Introduction

With the increasing consumption of energy, the price of coal fluctuates. The price of coal is not only subject to the supervision of relevant state departments, but also to the influence of the domestic coal market. In addition, other factors such as climate change, travel patterns, energy consumption patterns, and the international coal market will also affect coal prices. The analysis and prediction of the change trend of coal prices play an important role in the analysis of my country’s economic trends, which is of great research significance [1].

2. Long-Short Term Memory Recurrent Neural Network Model

For the prediction of thermal coal price, this paper establishes Long-Short Term Memory Recurrent Neural Network Model for prediction. Different from the traditional neural network [2, 3], LSTM RNN is a neural network based on time series. The network structure mainly includes input gates, output gates, and forget gates. Its structure is shown in the figure below:

![Figure 1. LSTM RNN Structure diagram.](image-url)

The rectangular box in the middle of Figure 1 is called a memory block, which mainly contains three gates (forget gate, input gate, and output gate) and a memory cell. The horizontal line at the top of the
box is called the cell state. It is like a conveyor belt that can control the transfer of information to the next moment [4].

3. Algorithm flow

The process of LSTM RNN is as follows:

Step 1: Decide what information can pass cell state. This decision is controlled by the "forget gate" layer through sigmoid, which will pass or partially pass based on the output at the previous moment. As follows:

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

Figure 2. The first step of the LSTM neural network process.

Step 2: Generate new information that we need to update. This step consists of two parts. The first is an "input gate" layer that uses sigmoid to determine which values to update, and the second is a tanh layer that is used to generate new candidate values. The candidate values are obtained as shown in Figure 3. The two steps are the process of discarding unnecessary information and adding new information;

$$C_t = f_t \ast C_{t-1} + i_t \ast \tilde{C}_t$$

Figure 3. The second step of the LSTM neural network process.

Step 3: To determine the output of the model, we firstly obtain an initial output through the sigmoid layer, then use tanh to scale the value between -1 and 1. What’s more, we need to multiply the output from the sigmoid pair by pair to obtain the output of the model, as shown in Figure 4.

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

$$h_t = o_t \ast \tanh(C_t)$$

Figure 4. The final step of the LSTM neural network approach.
4. Model solving and result verification

In this paper, the data are processed as days, weeks, and months as the unit to establish a neural network, 80% of the data is taken as the training sample, and 20% of the data as the test set. In LSTM RNN, for the forecast model with the unit of day, the data of the 11th day in the time series is defined by the previous 10 days; for the forecast model with the unit of the week, the data of the 8th day in the time series is defined by the 7 Weekly data forecast; for a forecast model with a month as a unit, the data of the 6th month in the defined time series is predicted by the data of the previous 5 months. The hidden layer of the neural network is set to 30.

For the test of the prediction model, this paper selects the mean square error MSE and the mean absolute error MAE as the two test indicators of model accuracy [5], the calculation formula is as follows:

\[
MSE = \frac{1}{m} \sum_{i=1}^{m} (y - \hat{y})^2
\]

\[
MAE = \frac{1}{m} \sum_{i=1}^{m} |y - \hat{y}|
\]

Where \(y\) is the actual value, \(\hat{y}\) is the predicted value, and \(m\) is the number of samples.

Establish LSTM RNN models with day, week, and month as the time series in MATLAB software. The final training model results are shown in Table 1:

|                      | MSE  | MAE  |
|----------------------|------|------|
| Forecast model 1 (day)| 0.2402 | 0.1243 |
| Forecast Model 2 (week) | 13.1241 | 1.2114 |
| Forecast Model 3 (month) | 34.4221 | 2.1453 |

4.1. Forecast model 1 (day)

![LSTM training process diagram](a)
From the above results, it can be seen that the forecasting model 1 with days as the unit has achieved very high accuracy, so it can be used to predict the Qinhuangdao thermal coal price in the next 31 days.

4.2. Forecast Model 2 (week)
From the above results, it can be seen that the accuracy of the prediction model 2 with the unit of week is lower than that of the prediction model 1 with the unit of day. The reason is that the number of training samples is greatly reduced, but the model can also meet the accuracy requirements of prediction. Therefore, it can be used to predict the Qinhuangdao thermal coal price in the next 31 weeks.

4.3. Forecast Model 3 (month)
Figure 7. (a) And (b) below are the training process diagram and error diagram of the prediction model 3 (month).

5. Conclusion
In this paper, the coal price prediction model based on LSTM is established in the unit of day, week and month, and has achieved good accuracy. Therefore, the model established in this paper has certain practical significance and reference value.

References
[1] Willmott C, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance [J]. Climate Research, 2005, 30(1):79-82.
[2] Gers F A, Schmidhuber E. LSTM recurrent networks learn simple context-free and context-sensitive languages [J]. IEEE Transactions on Neural Networks, 2001, 12(6):1333-1340.
[3] Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM [J]. Energy, 2018, 148(APR.1):461-468.
[4] Yildirim, ?zal. A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification [J]. Computers in Biology & Medicine, 2018:189.