Machine learning algorithm for the financial time series data forecasting

Lin Feng, Ziren Chen, Harold A. Lay, Jr.

Department of Mathematics, Middle Tennessee State University

ABSTRACT

We present a Machine Learning algorithm, the deep learning Long Short-Term Memory (D-LSTM), for the prediction of financial time series data. This algorithm aims to predict the future time series based on the existing ones. The outcomes show it is superior to some traditional time series analysis models, like Single and Double Exponential Smoothing. The crucial process of D-LSTM is using the Backpropagation algorithm to update the weight and bias matrices as new information comes, which helps to understand the information gradually. It is easy to find that D-LSTM uses the same recursive formula as the Single Exponential Smoothing when updating the information. The recursive formula shows that the estimation of next moment is the weighted average of previous information. However, there is a big difference between them. In the traditional Exponential Smoothing method, generally, we choose weight parameters by hands, picking a sequence between 0 and 1, then plug them into the model to select the relatively best one. The weights are fixed also, that is, nothing will change to the weight as new information comes. It increases the systematic error and lead to an inflexible characteristic, which is a fatal shortcoming for a universal model. But D-LSTM does good in these two parts. In this algorithm, the weight parameters are selected by the neural network and keep renewing as new information entered, which helps to make a better combination of long-term and short-term information and helps to show a more accurate and more flexible characteristic compared to the traditional Exponential Smoothing. In another side, because of preciously weights selection and renewal, the CPU time of D-LSTM is longer than exponential smoothing. However, the outcomes are more reliable and more acceptable. And we got all the outcomes by Python and the TensorFlow in it.

MATERIALS AND METHODS

Exponential Smoothing

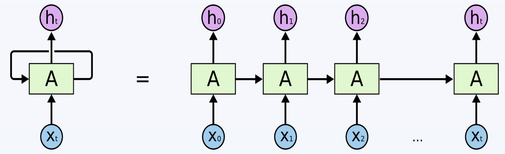
Exponential Smoothing is a very popular scheme to produce a smoothed Time Series. Whereas in Single Moving Averages the past observations are weighted equally, Exponential Smoothing assigns exponentially decreasing weights as the observation get older.

In other words, recent observations are given relatively more weight in forecasting than the older observations.

MATERIALS AND METHODS (cont.)

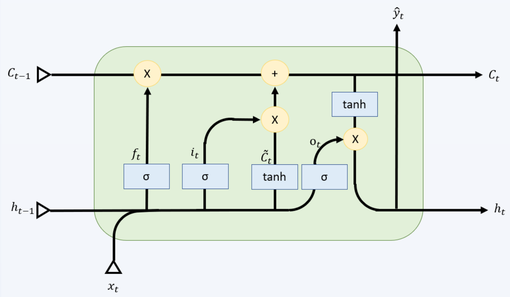
Recurrent Neural Networks(RNNs)

Recurrent neural Networks(RNNs) is a type of Artificial Neural Network(ANNs) where the input of next step depends on the output from previous step. For tasks that involve time series, such as language and finance data, it is often better to use RNN.



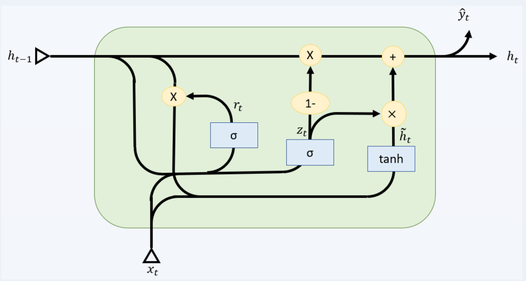
Long-short-erm Memory (LSTM)

LSTMs are an extension of RNNs. Due to the existence of cell and hidden state, compared with the traditional RNNs, it is able to learn how to combine long-term and short-term information perfectly, which is very helpful for us to analyze the various time series cases.



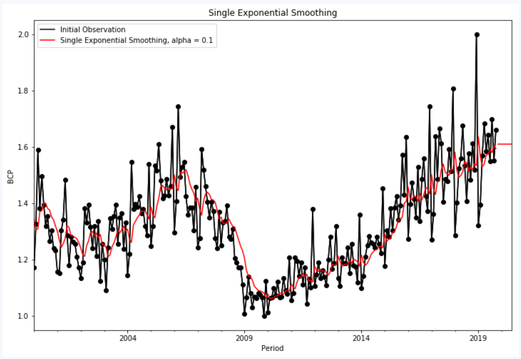
Gated Recurrent Unit (GRU)

GRU is a variant on LSTMs. It combines the forget and input gates into a single “update gate.” It also merges the cell and hidden state. So GRU helped more efficient compared with the traditional LSTMs. GRU tend to absorb short-term information due to the mixture of cell and hidden state, which is consistent with the fluctuation law of some time series cases. Therefore, the performance of GRU is even better than LSTMs in some special cases.

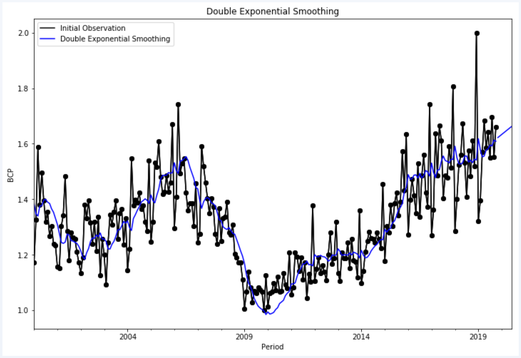


RESULTS

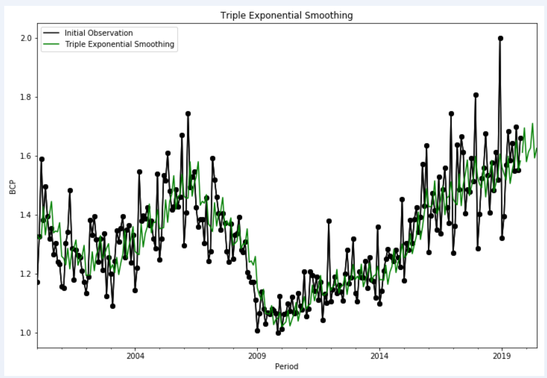
Results of Single Exponential Smoothing



Results of Double Exponential Smoothing

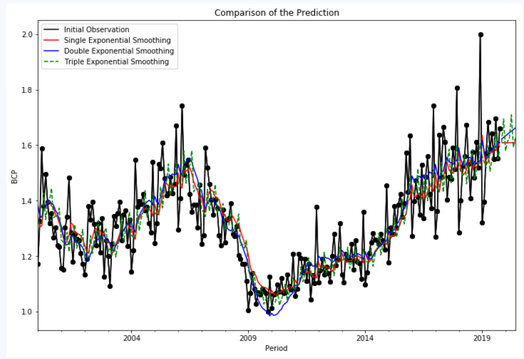


Results of Triple Exponential Smoothing

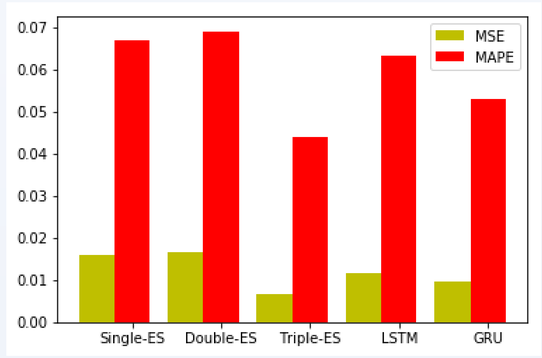


Results (cont.)

Comparison for all the method



Assessment



Error Comparison