Stock Price Prediction With Neural Network

1102-CSC0032 Artificial Neural Network Project#1

LIU,TZU-HUNG
Department of Computer Science and Information Engineering
National Taiwan Normal University
Taipei, Taiwan
40747031s@gapps.ntnu.edu.tw

***Abstract*—Stock price prediction is important for value investments in the stock market.On the basis of huge amount of historical data.we can expect to use neural network to find the rule of stock price,at this paper we will focus on 1795(Lotus) of Taiwan Stock.Trying to find a model to predict stock price for the next five days.**

***Index Terms*—Artificial intelligence,Neural networks**

I. Introduction

To predict the prices of 1795(Lotus) for the next five days(3/28 to 4/1).I will use the data set of historical price.And use Long Short-Term Memory(LSTM) as main layer of neural network model. Reaching the goat by adjusting the model structure and parameters.To reduce the loss and MSE.

Long Short-Term Memory (LSTM) [1] is a temporal recurrent neural network (RNN) first published in 1997. Due to the unique design structure, LSTMs are suitable for processing and predicting important events with very long intervals and delays in time series.

At this project I will use Python as programming language,and use packages (Sklearn,Matplotlib,pandas,Numpy and Tensorflow) to reach the goat.Mainly use Tensorflow_keras to construct the hole model.And the full code is on my github(Barry0304/Artificial_Neural_Network_HW1)

II. Methodology

A. Data and Preprocessing

We have a data set about Lotus's historical stock price with its opening ,closing,highest,lowest prices for each day,and also the volume means the trading volume in one day.All these data are in the period from 2012/3/27 to 2022/3/25,with actual 2452 trading days.We can see it as a array with 2452 rows × 6 columns(one for date).At the part of data-preprocessing.First I will drop the date and use MinMaxScaler to rescale the value in a range of [0,1].And separate data with the ratio of 80% for training and 20% for validation.

And essential step in LSTM is to slice the data into multiple input data sequences with associated target values.For this process, I use a sliding windows algorithm moves a window step by step through the time series data, adding a sequence of multiple data points to the input data with each step.So that each input will be prices of past one hundred days and actual value will be the closing price of the next five days.

The result will be a training input array(1857,100,5) with (1857,5) true values,and validation input array(387,100,5) with (387,5) true values.

B. Model structure

With main layer of LSTM and Dense.I set 8 layers in the model:

- first two layer are LSTM layer,with 100*5 units and which takes our mini-batches as input and returns the whole sequence,and the third layer is LSTM layer that takes the sequence from the previous layer, but only return 5 values(wont return sequences)
- Below of them are three Dense layer with two dropout layer set to 0.5 between them,describe in detail,Dense layer's output reduced from 100 to 25 and final dense layer that outputs the five predicted value,all of them are activate with Relu.
- the model's compile use Adam as optimizer and count loss by mean square error.And will fit with 100 epochs.

III. Experiment Result

After fitting for 100 epochs.Cost about 10 minutes.I get average loss about 0.45.The whole loss are reduce belong every epochs show as Fig.1.The loss drops quickly to a lower plateau, which shows that the model has improved throughout the training process.

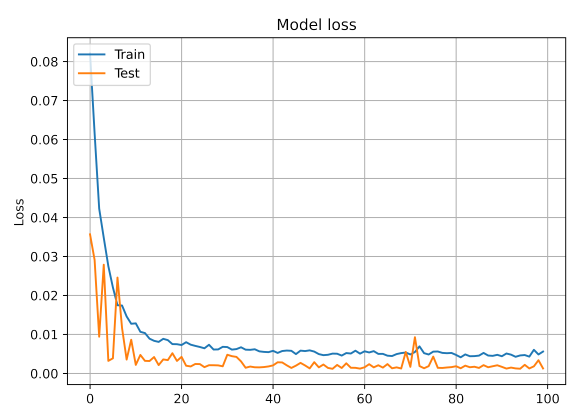

Fig. 1. training and testing loss

The MSE is 26.00,And MAPE is 5.38% that means some means of predictions have deviates from the actual values . The MDAPE is 4.16 % , a bit lower than the mean, thus indicating there are some outliers among the prediction errors. 50% of the predictions deviate by more than 4.16%, and 50% of deviate by less than 4.16% from the actual values.A line plot that shows the forecast and compare it to the actual values is shown as "Fig.2" for reference.

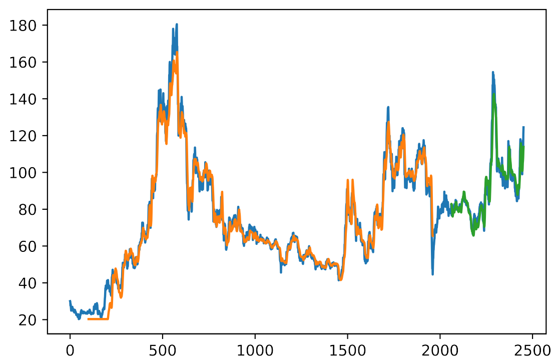

Fig. 2. forecast and actual values

IV. CONCLUSION

We found that the forecast is close to the actual values but also shows some deviations. The deviations are most significant during periods of increased market volatility and least during periods of steady market movement.Singly use one stock's historical prices may not a good choice for stock price.In the future,we can try with more feature like TWSE Capitalization Weighted Stock Index, other stocks in same category or have similar feature.To deal with more faces such as technical analysis or chip analysis,use Financial News [2] as message side also is a kind of choice.

REFERENCES

[1] Sepp Hochreiter, and Jürgen Schmidhuber, "Long Short-Term Memory" NEURAL COMPUTATION 9(8):1735-1780,1997.

[2] Xinyi Li, Yinchuan Li2, Hongyang Yang1, Liuqing Yang,and Xiao-Yang Liu, "DP-LSTM: Differential Privacy-inspired LSTM for Stock Prediction Using Financial News",arXiv:1912.10806v1 [q-fin.ST] 20 Dec 2019