EECE 571T Project Proposal: Stock Price Forecasting

Zhaosheng Li, Menghong Huang

Introduction

Prediction of future stock prices has always been an active area of research. A reasonable stock price prediction can identify the patterns in the stock price movement and lead to profitable investment decisions, In recent years, a lot of researchers have utilized different machine learning (ML) techniques to forecast the stock prices. [1] uses support vector regression (SVR) with windowing function to predict stock market prices. [2] compares the performance of using three different kernels of SVR for predicting stock prices. [3], [4], and [5] use the recurrent neural network (RNN) to predict the financial time-series data and finally get promising results. Various adaptations of back propagation neural network (BPNN) are also studied to better model the stock prices. [6] utilizes an improved back propagation (BP) algorithm and compares the single-input and multi-input BPNN for stock prices prediction. In this project, we propose to implement SVR, RNN, and BPNN to forecast the stock price of *GameStop*, an American gaming merchandise retailer. After that, we will evaluate and make comparisons among the performance of those three ML models.

Data

The dataset in this project is publicly available on the GameStop Historical Stock Prices from *Kaggle*. The data series covers the *Gamestop*'s stock prices over the period from 2002-02-13 to 2021-01-28 and contains the key features of daily stock trading, which are opening price, highest price, lowest price, trading volume, closing price, and the adjusted closing price. The dataset has 4773 instances in total. 3773 of those instances will be used as a training dataset for this project, and the remaining 1000 of those will be used for testing purpose.

Approach & Implementation

We will feed the training dataset into the three models mentioned above during the training process. This section illustrates why and how are we implementing those three models in this project:

- **SVR:** We acknowledge that the SVR is a useful ML model to recognize the pattern of time series data. Capturing the pattern of data, SVR can predict the future data reasonably. We will use SVR with three different kernels to predict the stock price. The three kernels include linear, polynomial, and radial basic functions.
- **RNN:** To be specific, we will be implementing the long short-term memory (LSTM) architecture in RNN. This is because the RNN cannot store long time memory. But incorporating LSTM into the RNN can overcome that drawback, because the LSTM has the capability to interpret sequential data and forecast data with long time series.
- **BPNN:** BPNN is effective in identifying nonlinear relationship. It also shows good ability to deal with time series data and predict the short-term trend. In order to avoid the problems in training process, such as easy to reach the local minimum and slow convergence, we will use an improved BP structure that involves all key features of stock trading, along with using stochastic gradient descent method to speed up the learning process.

Evaluation

After the training process is done, we will be feeding the testing dataset into the three models and measure the performance of those three models respectively. The measurement of performance is based on the following two metrics.

- **Root Mean Square Error (RMSE):** The residuals refer to the difference between the actual price and predicted price. Measuring the RMSE can tell us how spread out the residuals are and how concentrated the actual price is around predicted stock price movement.
- **Coefficient of Determination ($\mathbf{R^2}$):** R^2 is a statistic used in the prediction of future outcomes. Measuring the R^2 can tell us goodness of fit between the actual stock price and the predicted stock price movement.

Besides the two metrics above, we will also plot the actual and predicted prices on the same figure for each model to improve the data visualization.

References

[1] P. Meesad and R. I. Rasel, "Predicting stock market price using support vector regression," *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*, pp. 1–6, 2013.

[2] B. Madhu, A. K. Paul, and R. Roy, "Performance comparison of various kernels of support vector regression for predicting option price," *International Journal of Discrete Mathematics*, vol. 4, no. 1, p. 21, 2019.

[3] J. Qiu, B. Wang, and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism," *PLoS ONE*, vol. 15, 2020.

[4] A. Moghar and M. Hamiche, "Stock market prediction using lstm recurrent neural network," *Procedia Computer Science*, vol. 170, p. 1168–1173, 2020.

[5] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of lstm," *International Journal of Machine Learning and Cybernetics*, vol. 11, pp. 1307–1317, 2020.

[6] W. Ma, Y. Wang, and N. Dong, "Study on stock price prediction based on bp neural network," *2010 IEEE International Conference on Emergency Management and Management Sciences*, 2010.

[7] S. Mehtab, J. Sen, and S. Dasgupta, "Robust analysis of stock price time series using cnn and lstm-based deep learning models," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1481–1486, 2020.

[8] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. Menon, and K. Soman, "Stock price prediction using lstm, rnn and cnn-sliding window model," *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1643–1647, 2017.

[9] P. Gao, R. Zhang, and X. Yang, "The application of stock index price prediction with neural network," *Mathematical Computational Applications*, vol. 25, p. 53, 2020.

[10] Y. Guo, S. Han, C. Shen, Y. Li, X. Yin, and Y. Bai, "An adaptive svr for high-frequency stock price forecasting," *IEEE Access*, vol. 6, pp. 11 397–11 404, 2018.

[11] K. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, pp. 307–319, 2003.

[12] F. E. Tay and L. Cao, "Application of support vector machines in financial time series forecasting," *Omega*, vol. 29, no. 4, p. 309–317, 2001.

[13] Y. Song, Y. Zhou, and R. Han, "Neural networks for stock price prediction," *arXiv e-prints*, p. arXiv:1805.11317, May 2018.