Evaluating Machine Learning Algorithms for Stock Prediction

Daiki Akiyoshi (dakiyo14@earlham.edu)
Department of Computer Science, Earlham College, Richmond, IN

Motivation and Background

Stock market prediction and developing profitable trading strategies have always attracted businesses and academia. Previous studies demonstrated high accuracy in predicting directional movements of stock prices[1]. However, achieving high directional accuracy does not directly translate into great profits, as this ignores the magnitude of stock price fluctuations.

To address this limitation, this study evaluated predictive models based on three metrics: directional accuracy, closeness, and profit generated by trading simulation. These three metrics allow for effective comparison of machine learning algorithms and help determining the potential applicability of machine learning in predicting stock prices in a real setting.

Datasets

- S&P 500 Index data (an American stock market index) from Jan 2000 to Jan 2019.
- The data was pre-processed to generate ten financial indicators (e.g. simple moving average)

Alpha Vantage API [2]

Project Framework

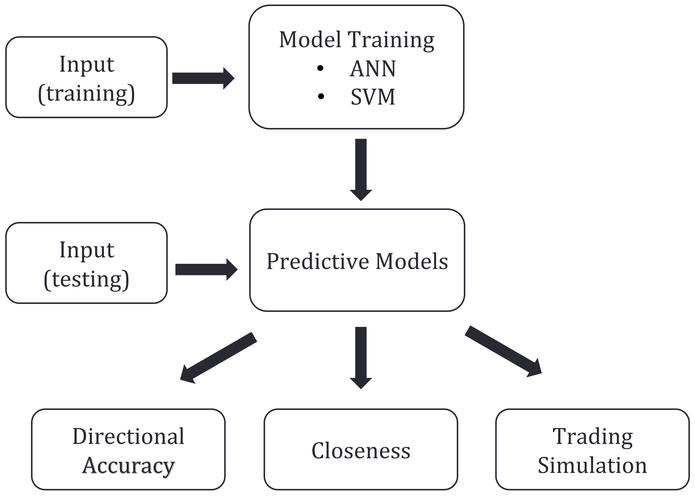

Figure 1: Project Framework

Methods

This project focused on comparing prediction performances of ANN and SVM based on three evaluation metrics: **directional accuracy**, **closeness**, and **trading simulation (**see Figure 1). Table 1 describes the formulas used for the three metrics. Moreover, this study examined how the values yielded from these three metrics change as the models predict distant futures 1, 3, 5, 10, and 20 days in advance. To set the model hyperparameters efficiently, different levels of hyperparameters were tested. Table 2 shows the hyperparameters that were determined based on the experiments.

Metrics	Formula
Directional Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$
Mean Absolute Percentage Error (MAPE)	$\frac{1}{n}\sum\left\|\frac{Actual-Forecast}{Actual}\right\|$
Trading Simulation	$Ln(\frac{Total\ exit\ price\ for\ trades}{Total\ cost\ for\ trades})$

Table 1: Evaluation Metrics

Hyperparameters	ANN	SVM
Neurons	80	-
Epochs	1000	-
Momentum constant	0.2	-
Learning rate	0.1	-
Kernel	-	Polynomial
Degree	-	1
Regularization	-	1

Table 2: Hyperparameters

References

1. Jigar Patel, Sahil Shah, Priyank Thakkar, and K Kotecha. 2015. Predicting Stock and Stock Price Index Movement Using Trend Deterministic Data Preparation and Machine Learning Techniques. Expert Syst. Appl. 42, 1 (Jan. 2015), 259–268. https://doi.org/10.1016/j.eswa.2014.07.040
2. Alpha Vantage. 2018. Alpha Vantage. https://www.alphavantage.co/ Accessed: 2018-10-30.

Results

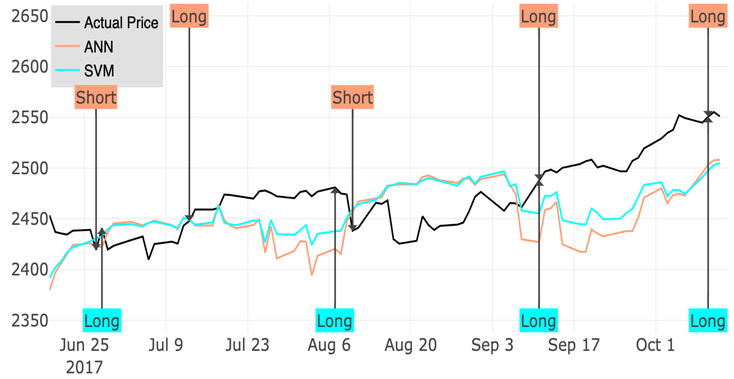

Figure 2: ANN vs SVM for 20-day ahead predictions

	1-day	3-day	5-day	10-day	20-day
Accuracy	51.88%	59.66%	51.17%	66.52%	57.52%
MAPE	0.83%	1.26%	1.36%	1.81%	2.54%
Predicted return (P)	0.59%	0.81%	0.52%	0.83%	0.50%
Actual return (A)	-0.14%	0.22%	-1.32%	-0.14%	-0.74%
P-A	0.73%	0.47%	1.84%	0.97%	1.24%

Table 2: Experiment Results for ANN

	1-day	3-day	5-day	10-day	20-day
Accuracy	50.00%	56.72%	59.07%	65.03%	71.46%
MAPE	0.66%	1.03%	1.31%	1.81%	2.45%
Predicted return (P)	0.29%	0.53%	0.4%	0.33%	0.34%
Actual Return (A)	-0.78%	-0.48%	-1.22%	-2.01%	0.20%
P-A	1.07%	1.01%	1.62%	2.34%	0.14%

Table 3: Experiment Results for SVM

In terms of directional accuracy and profit prediction, ANN performed better for 1, 3, 10-day predictions, while SVM performed better for 5 and 20-day predictions. In terms of MAPE, SVM outperformed ANN in all of the five different predictions. Next steps for this study is to understand the performance differences between ANN and SVM.

Acknowledgements

I would like to thank Ajit Chavan and David Barballa for providing me detailed feedback to improve this research. I also would like to thank Xunfei Jiang for providing me feedback on topic proposal last semester.