Algorithmic Trading

Mihail Mihaylov

Reg. No: 201888925

1. Introduction and background

The main goal of this project is to come up with an algorithmic trading strategy to be implemented using a machine learning or search-based software engineering technique. The topic chosen for this assignment is to predict the next day low and high prices of a stock based on values from a number of previous days. The prediction will be done using a neural network. Once the model is trained a set of predictions will be obtained, which are going to be supplied to a set of trading rules, which are based on those predictions to determine if a profit can be made. The trading rules would determine whether stocks should be bought, sold or held on that day. Each day has only one decision. This approach is known as short-term trading.

Algorithmic trading involves using a computer program that creates orders and automatically submits them to a market centre or exchange. These orders are based on a set of rules using a trading strategy which is based on technical analysis. A trading strategy is a method of buying and selling in markets that is based on predefined rules used to make trading decisions [3]. On the other hand, technical analysis is a methodology for forecasting the direction of prices through the study of past market data, primarily price and volume [4]. Both of them combined together provide the basis of this project.

There are numerous research projects using neural networks in the stock trading world. Some of them involve building an artificial neural network to build a model for intra-day currency exchange. Even though the currency values are very chaotic, noisy and non-stationary there has been a neural network build on in to predict future values with a 72.5% prediction accuracy [1]. Another project using a neural network involves classifying whether a stocks' price would go up or down in the near future. After the neural network was trained, the model was evaluated in different situations and a profit was made in all of the cases [2]. A third research project involved training several neural networks into a single Boolean decision rule system, to make trading decisions for S&P 500 [5]. These are just a few example of how neural networks have already been incorporated into the stock trading market and how successful they can be in prediction a future.

2. Choice of data

The first part of this section will provide the reader with overview of the data sets chosen as well as some general comments on those data sets and plots of the attributes. The second part will describe the inputs and outputs of the neural network implemented.

2.1. Types of Stock

For this project there were 2 companies considered initially. The first one is Viscount Nelson (VNET) and the second one is Apple Inc. (AAPL). The time period of the data is 2 years, starting on the 1st of January 2016 until the 1st of January 2018. There are 2 main differences between the 2 assets chosen. The first one is that VNET is a low price stock (less then 50 per share) and AAPL is a medium priced stock (more than 100 but less than 200 per share). Another difference is that for the time frame considered the price of AAPL, except for short period trends, has been growing and has

went from 110 in the beginning to ~180 at the end of the period considered. As for VNET, it starts of 20 per share and the stock price has had a big drop at some point and has been fluctuating between 5-10 per share after that until the end of the 2 year period. Below can be seen summary plots for both stocks, containing time series plots of the opening, closing, low and high price as well as the volume, relative strength index (RSI) and the exponential moving average (EMA).

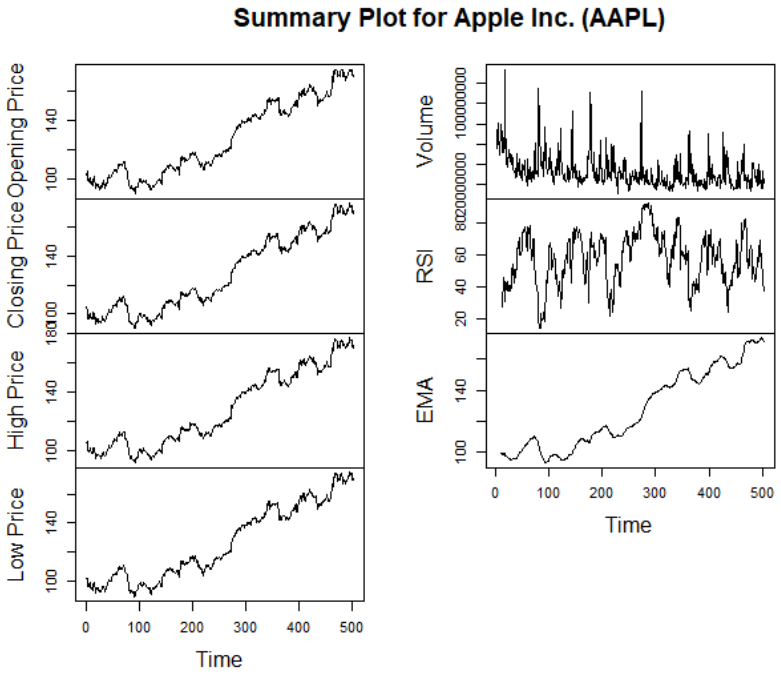

Figure 1. Summary Plot for Apple Inc. (AAPL)

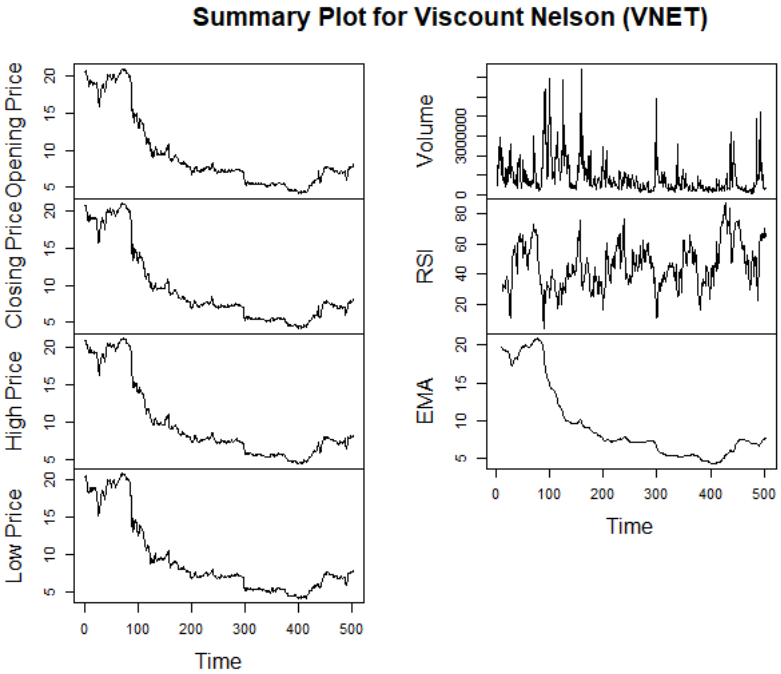

Figure 2. Summary Plot for Viscount Nelson (VNET)

The 4 figures below show the first 6 rows of data and the summary of VNET and AAPL respectively. As it can be seen from the 4 figures below as well as the 2 plot figures above for VNET the initial prices were around 20 per share, but the mean is 9.33 and the minimum value for the low price reached 4.17 at some point. As for AAPL it can be seen from the first 6 rows that the initial prices were around 100 but it keeps going up with every quarter of the data used, reaching a maximum of ~175 per share.

	VNET.Open	VNET.High	VNET.Low	VNET.Close	VNET.Volume	VNET.Adjusted
2016-01-04	20.59	20.87	20.46	20.76	1509800	20.76
2016-01-05	20.74	20.89	20.66	20.70	595900	20.70
2016-01-06	20.48	20.72	20.48	20.62	454600	20.62
2016-01-07	20.17	20.37	19.89	19.97	2434300	19.97
2016-01-08	20.13	20.38	19.75	19.84	1357600	19.84
2016-01-11	19.83	19.97	18.54	18.71	2124200	18.71

Figure 3. First 6 rows of data for VNET

Index	VNET.Open	VNET.High	VNET.Low	VNET.Close	VNET.Volume	VNET.Adjusted
Min. :2016-01-04	Min. : 4.17	Min. : 4.42	Min. : 4.17	Min. : 4.18	Min. : 77400	Min. : 4.18
1st Qu.:2016-07-03	1st Qu.: 5.70	1st Qu.: 5.84	1st Qu.: 5.55	1st Qu.: 5.68	1st Qu.: 356000	1st Qu.: 5.68
Median :2016-12-30	Median : 7.45	Median : 7.61	Median : 7.33	Median : 7.47	Median : 590100	Median : 7.47
Mean :2016-12-31	Mean : 9.52	Mean : 9.70	Mean : 9.33	Mean : 9.51	Mean : 896875	Mean : 9.51
3rd Qu.:2017-07-01	3rd Qu.:10.38	3rd Qu.:10.55	3rd Qu.:10.11	3rd Qu.:10.29	3rd Qu.:1052900	3rd Qu.:10.29
Max. :2017-12-29	Max. :21.05	Max. :21.12	Max. :20.93	Max. :21.03	Max. :6443100	Max. :21.03

Figure 4. Summary of the data set for VNET

	AAPL.Open	AAPL.High	AAPL.Low	AAPL.Close	AAPL.Volume	AAPL.Adjusted
2016-01-04	102.61	105.37	102.00	105.35	67649400	99.50
2016-01-05	105.75	105.85	102.41	102.71	55791000	97.01
2016-01-06	100.56	102.37	99.87	100.70	68457400	95.11
2016-01-07	98.68	100.13	96.43	96.45	81094400	91.09
2016-01-08	98.55	99.11	96.76	96.96	70798000	91.58
2016-01-11	98.97	99.06	97.34	98.53	49739400	93.06

Figure 5. First 6 rows of data for AAPL

Index	AAPL.Open	AAPL.High	AAPL.Low	AAPL.Close	AAPL.Volume	AAPL.Adjusted
Min. :2016-01-04	Min. : 90	Min. : 91.7	Min. : 89.5	Min. : 90.3	Min. : 11475900	Min. : 86.3
1st Qu.:2016-07-03	1st Qu.:105	1st Qu.:106.3	1st Qu.:104.8	1st Qu.:105.8	1st Qu.: 23251750	1st Qu.:100.6
Median :2016-12-30	Median :118	Median :118.2	Median :116.8	Median :117.5	Median : 27940600	Median :113.0
Mean :2016-12-31	Mean :127	Mean :128.4	Mean :126.5	Mean :127.5	Mean : 32785383	Mean :123.3
3rd Qu.:2017-07-01	3rd Qu.:152	3rd Qu.:153.8	3rd Qu.:150.8	3rd Qu.:152.6	3rd Qu.: 37021300	3rd Qu.:148.5
Max. :2017-12-29	Max. :175	Max. :177.2	Max. :174.9	Max. :176.4	Max. :133369700	Max. :173.1

Figure 6. Summary of the data set for AAPL

2.2. Input and Output Data

As mentioned before the use of neural networks was utilized in this project and this section the inputs and outputs of that neural network would be described in detail. In total there are 15 inputs and 2 outputs of the neural network. For best results it was determined that the inputs be the open, close, high and low prices, which were lagged 3 times and the volume, RSI and EMA values, which were lagged once. Other configurations involved lagging 2, 4 or 5 times, as well as lagging the volume, RSI and EMA values the same number of times as the other values, but that proved to overfit the model and did not yield good results for the test set.

The outputs of the neural network are the next day open and close price. It is worth mentioning that all the values were scaled using the scale function from the base package in R and after obtaining the predictions they were unscaled using the unscale function which is from the DMwR package in R.

After scaling the 2 years worth of data and removing the rows with null values, there were 492 instances left to be split into a training/testing set. Since this is a time-series data the split was

performed by keeping the first 75% of the data for the training set and the remaining 25% for testing. That leaves 369 training instances and 123 test instances.

3. Details of the approach taken

This section would focus on describing the neural network implemented and the trading rules used to determine how profitable the predictions are.

3.1. Neural network

The initial neural network was configured for VNET and it consisted of 4 hidden layers with 20 neurons in each layer. The threshold set is 0.01 which tells the neural network to stop training if the fitness function doesn't improve by more than 1% through each iteration. The algorithm used was the "rprop+", which is resilient backpropagation with weight backtracking. The stepmax argument was increased to 1e+06 which allows for more iterations before stopping the neural network. The results from this neural network achieved were really good for the training set but the neural network was underperforming for the test set. After numerous modifications it was determined to use a 3 hidden layer neural network with 10 neurons in each layer. Even though this leads to worse results for the training set it generalizes a lot better for the whole data set. The rest of the necessary arguments to train a neural network were set by default. The error function is the sum of squared errors.

Another argument worth mentioning is the seed argument. By using the build-in set.seed function and giving it an integer parameter, would make the neural network reproducible for future retraining. The neural networks trained in this project used the function with a value of 2.

The figure below shows a plot of a neural network trained for the VNET stock using the parameters described before.

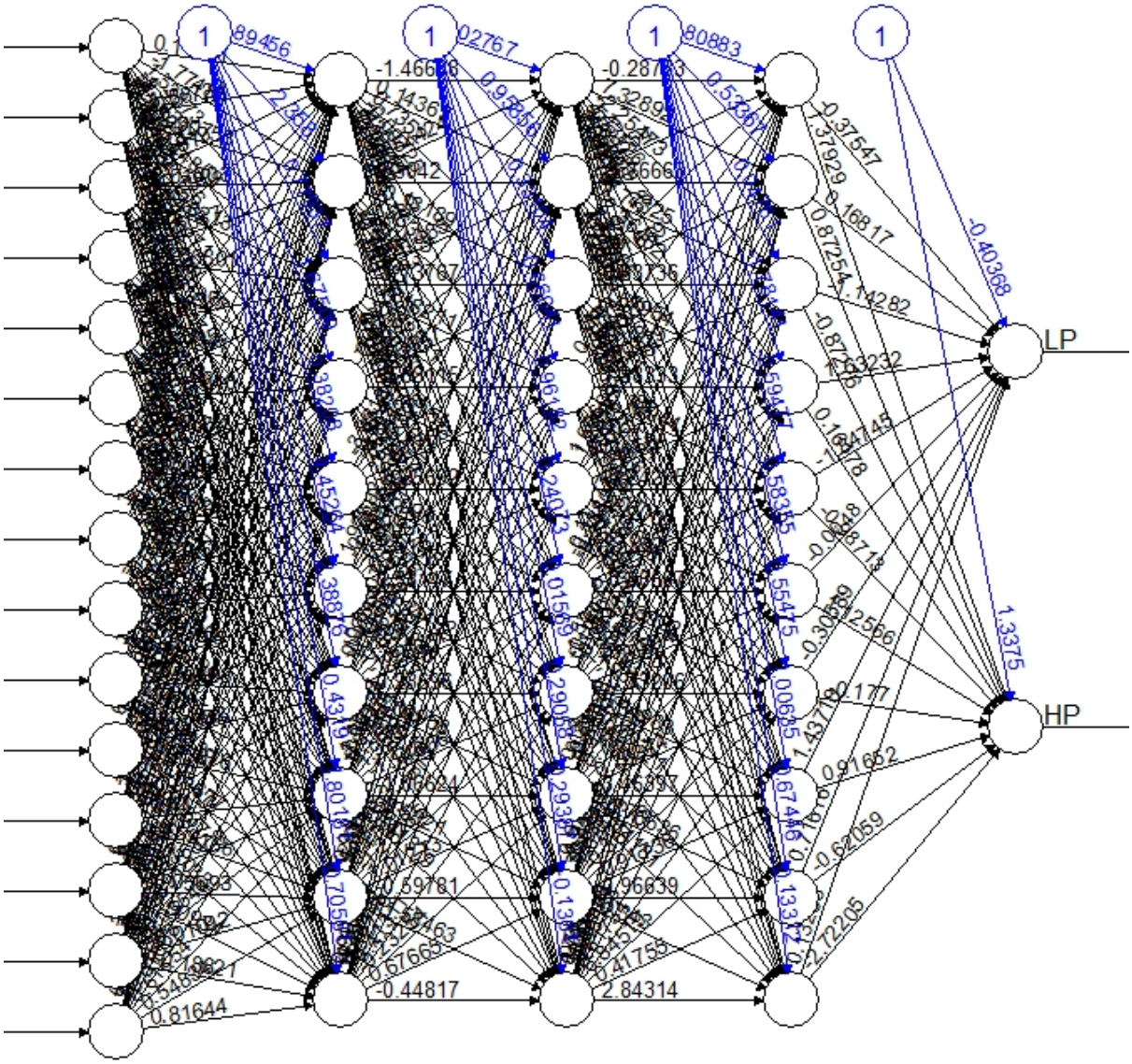

Figure 7. Sample Neural Network trained on the VNET data set

3.2. Trading rules

This subsection describes the trading rules in details. There are 3 actions possible for each day: Buy, Sell or Hold. The trading rules are applied on the predictions based on the test set. That set contains 123 predictions which would yield a 123 decisions, 1 for each day. The trading occurs with a starting capital which has been set to 10000 for testing purposes, a variable to hold the number of stocks set to 0, and a variable to save the last buy price currently set to 0 as well.

The other variables involved in reaching a decision are the predicted high and low values and the actual high and low values for the current day. (The actual high and low values have no influence on the decision or the buy/sell price, they are only used to validate that decision).

For each of the predictions there is a margin set. The chosen margin in the case of VNET is +-1%. That is if the actual price is within the margin of the predicted price the buy or the sale can actually occur. For the buy rules if the actual low price is within the +-1% margin of the predicted low price a buy can occur at the upper limit of the margin. This way it ensures that the said value is going to occur at some point on that day. For the sell of stocks the actual high price would need to be in the margin of the predicted high price. If that occurs the sale is valid and can occur at the lower margin of the predicted high price.

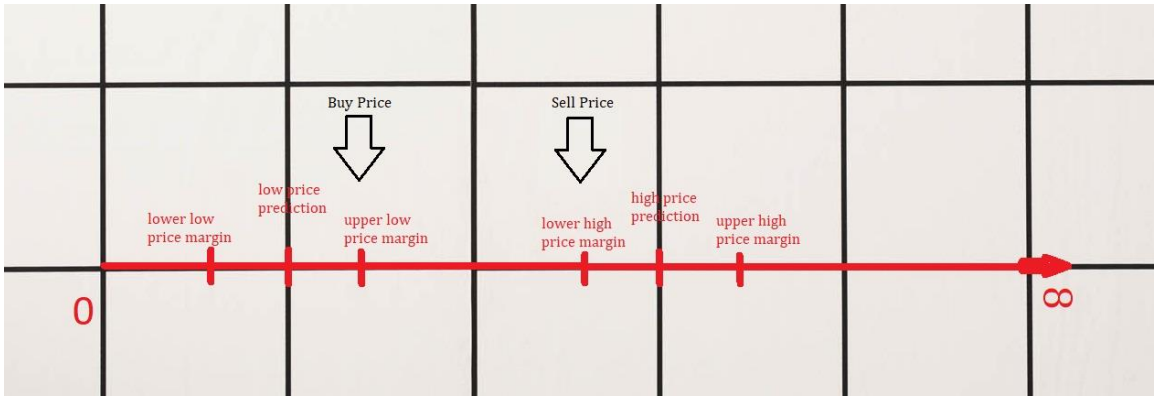

Figure 8. Presentation of predictions and their margins

Another rule implemented is that when a buy occurs, the buy is for the price of the full capital and a buy can not happen again before a sell. And the other way around is also valid, stocks can not be sold in increments, all of them need to be sold when a sell decision is reached. The last selling rule is that a sale can only occur if the predicted high price is at least a set percentage higher than the price the stocks were bought for. In the case of VNET that is 10%, while for AAPL the profit percentage is set to 6%. Another thing taken into consideration is a selling fee. Whenever a sale occurs a 5% fee is taken from the profit made from that sale.

After the last day of trading a summary is prepared, stating the initial, final capital and the profit made as well as the percentage that the capital has been increased by from the 1st Day. In case that there are stocks left after the last day, they are sold after the market is closed at the closing price for that day. The results from applying those trading rules can be seen in the performance of the model section below.

4. Presentation of and comments on the solutions achieved

The neural network described before was trained numerous times using different arguments and after picking the most suitable once the following results were obtained. For the VNET data set the achieved error for the training set was 0.5595, and for the AAPL data set the error was 0.6401. The 4 plots below provide a comparison of the actual data versus the predictions for both data sets. As it can be seen from the comparisons the models are performing really well on the training sets, with very small differences. Both models generalize well over the training data (as well as the testing data, which will be shown in the next section), and do not cause it to overfit or underfit.

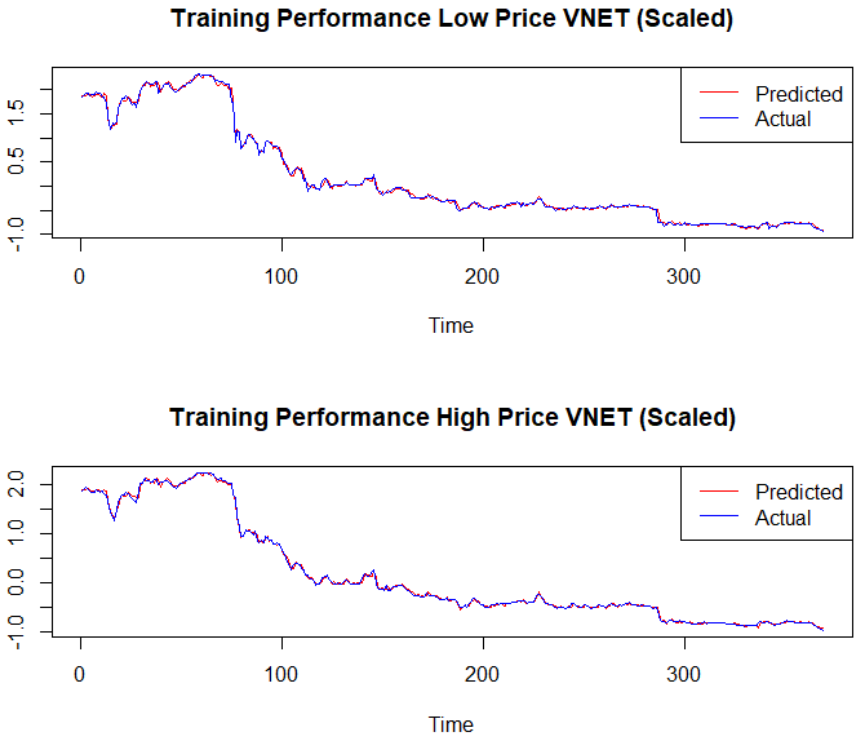

Figure 9. Training Performance for Low and High Prices for VNET

Figure 10. Training Performance for Low and High Prices for AAPL

5. The performance of the model

The performance of the model was evaluated using predictions on the testing data to make decisions of trading actions. Once the model was trained on the data, the testing set (123 instances) was fed into the model to make predictions for the next 123 days. The performance was evaluated using the root mean squared error (RMSE) of the predictions versus the actual values. For the VNET data set with the unscaled low price predictions the RMSE was calculated using a Metrics package to be 0.198 and the unscaled high price predictions matched the actual high prices with a RMSE difference of 0.2328. Considering this was calculated on unscaled data which varies between 4 and 8 those RMSE prove that the model generalizes quite well on future data. The 2 plots below show the comparison between the predicted and the actual values for VNET on the testing set. As it can be seen from the plots the predicted values are quite close to the actual values on future unseen data.

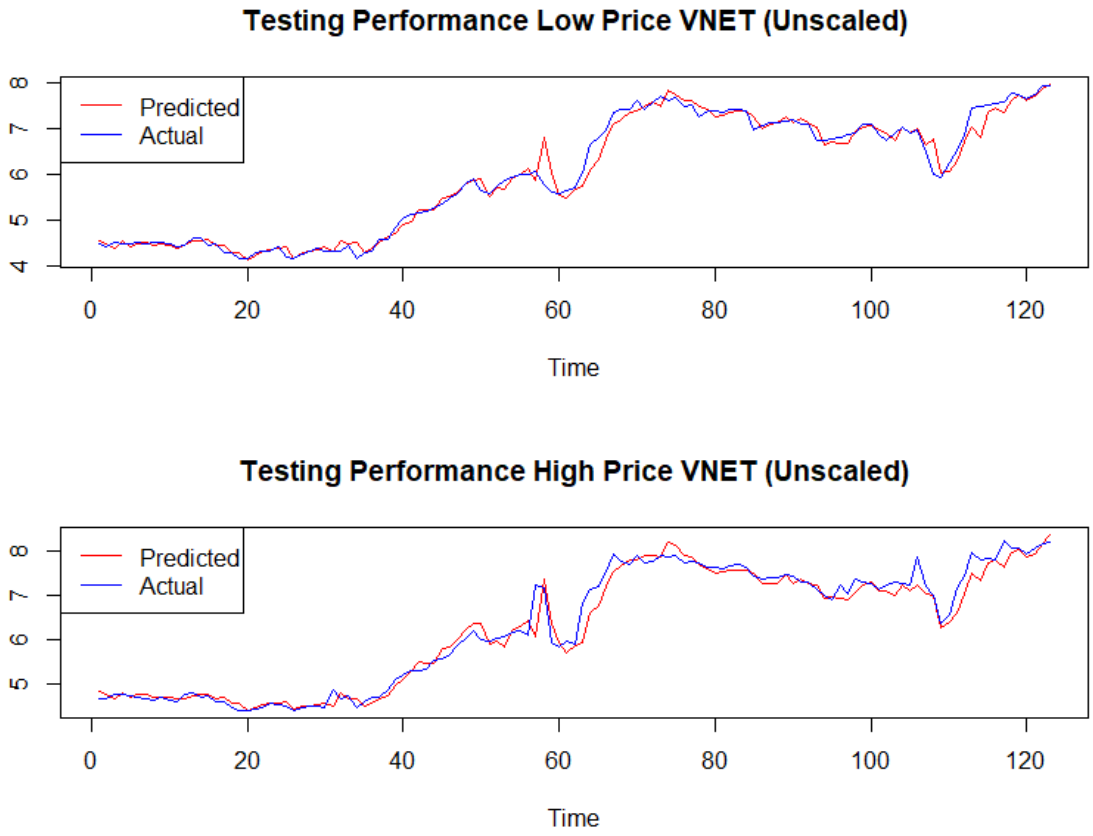

Figure 11. Testing Performance for Low and High Prices for VNET

After obtaining those predictions they were ran against the trading rules described in the previous section to check whether those predictions and the trading rules would yield an actual profit. With an initial capital of 10000 and a minimum sale profit of 10% after 123 days the final capital is 17003 which shows that the rules and the predictions work quite well together to ensure a good profit. A list of the buys and sells can be seen in the figure below. As it can be seen the stocks are sold as soon as they reach at least a 10% profit and new stocks are bought straight after only to be sold for a higher profit after. This is done to minimize the risk of losing the capital waiting for the best price. Based on those predictions it can be seen that a total of 4 buys and 3 sells are performed before the last day of the trading period, where it was sold after the market is closed.

```
Day:  6
After holding for  5 days...
Buy Stocks!
Num stocks: 2196
Capital: 0
Price of buy is: 4.554
-----------------------
Day:  41
After holding for  34 days...
Sell stocks!
Num stocks: 0
Capital: 11436
Price of sell is: 5.243
-----------------------
Day:  43
After holding for  1 days...
Buy Stocks!
Num stocks: 2164
Capital: 0
Price of buy is: 5.284
-----------------------
Day:  54
After holding for  10 days...
Sell stocks!
Num stocks: 0
Capital: 13185
Price of sell is: 6.135
```

```
Day:  55
Buy Stocks!
Num stocks: 2171
Capital: 0
Price of buy is: 6.074
-----------------------
Day:  73
After holding for  17 days...
Sell stocks!
Num stocks: 0
Capital: 16710
Price of sell is: 7.783
-----------------------
Day:  75
After holding for  1 days...
Buy Stocks!
Num stocks: 2131
Capital: 0
Price of buy is: 7.842
-----------------------
Day: 123
Selling on the last day!
The price of sell is: 7.95
The initial capital was: 10000
The final capital is: 17003
The profit made is:  7003
After: 123 days the capital is increased by 70.03 %.
```

Figure 12. List of Buy/Sell Decisions for VNET

The AAPL data set's performance was evaluated the same way as the previous one. Firstly, predictions were obtained for the 123 days testing set and then those predictions were fed into the trading rules to check if there is a profit to be made from them. The predictions for the testing set gave a RMSE value of 7.694 for the low price values and 7.852 for the high price values. These values are higher than the ones for the testing set for VNET, but the price of a share for AAPL varies between 140 and 180 unlike the VNET (4 to 8), which means that those predictions could still be quite good. The figure below shows a plot of the predictions versus the actual values for both the high and the low values of the testing set, where a better understanding of the performance can be obtained.

Figure 13. Testing Performance for Low and High Prices for AAPL

As it can be seen the predictions match the actual values quite well at the beginning but at some point they start to deviate and get quite far from the actual values in the end. The reason for that is a change in the trend of the AAPL stock prices. A sudden change in the actual prices could

cause the model to give out bad predictions. Even though the predictions are not as good for the AAPL stock on unseen future data, they were still ran through the trading rules. The following figure shows the trading decisions obtained using those predictions. As it shows, there are still 2 buys and 1 sell, with a sell on the last day as well, but those are only in the beginning of the trading period (days: 1, 13 and 18). That is due to the fact that the predictions are too far off or the profit margin (set to 6% for AAPL) has not been reached.

```
Day:  1
Buy Stocks!
Num stocks: 69.86
Capital: 0
Price of buy is: 143.1
-----------------------
Day:  13
After holding for  11 days...
Sell stocks!
Num stocks: 0
Capital: 10588
Price of sell is: 152
-----------------------
```

```
Day:  18
After holding for  4 days...
Buy Stocks!
Num stocks: 70.58
Capital: 0
Price of buy is: 150
-----------------------
Day: 123
Selling on the last day!
The price of sell is: 169.2
The initial capital was: 10000
The final capital is: 11944
The profit made is:  1944
After: 123 days the capital is increased by 19.44 %.
```

Figure 14. List of Buy/Sell Decisions for AAPL

6. Comparison

Comparing the neural network approach to other approaches is very important in order to see how well it performs against them. Using a random number generator is not very suitable for this kind of project but the comparison in this case can be done using the mean value. In order to compare properly it is essential to obtain predictions for the same time period as the testing set for the neural network approach. The predictions are based on the mean values of the low and high prices for the previous 7 days. Once those are obtained they are supplied to the same trading rules as the neural network before to see the profit they would yield. Using the mean method gives a 0.3657 RMSE for the low prices for VNET and 0.3734 for the high prices for VNET. Those values, even though they are quite good, are not as good as the values obtained for the test set of the neural network. A plot of the actual against the predicted low and high prices can be seen below.

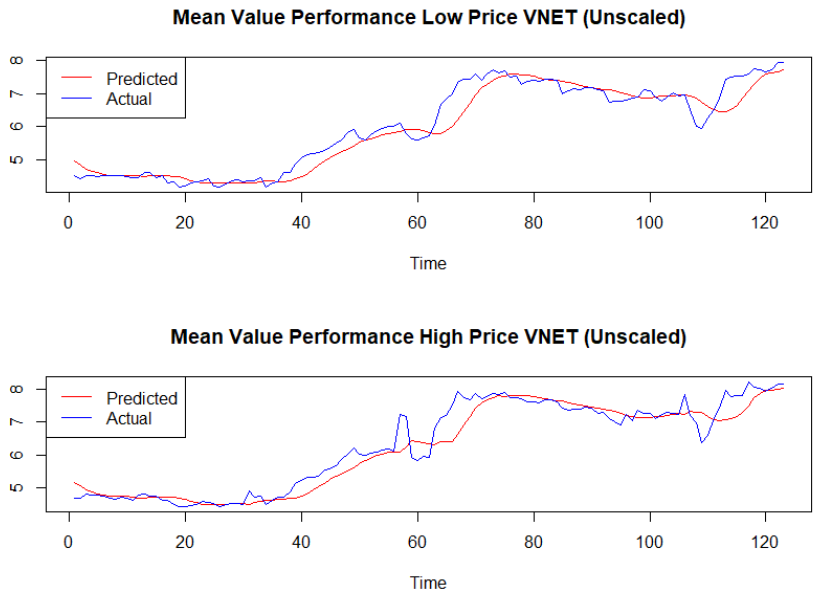

Figure 15. Mean Value Performance for Low and High Prices for VNET

As it can be seen, even though the predictions do not match the actual values there are a few crossovers between them, which means that a few buy/sell decisions will occur. The figure below shows the buy and sell actions that have occurred when the predictions using the mean value have been ran through the trading rules. For the 123 day trading period 2 buys and 1 sell have occurred with a sell after the market has closed on the last day. From those predictions a profit of 34.96% has occurred, which even though it is positive it is not as high as the profit given by the neural network.

```
Day:  6
After holding for  5 days...
Buy Stocks!
Num stocks: 2176
Capital: 0
Price of buy is: 4.595
-----------------------
Day:  56
After holding for  49 days...
Sell stocks!
Num stocks: 0
Capital: 12972
Price of sell is: 6.033
-----------------------
```

```
Day:  77
After holding for  20 days...
Buy Stocks!
Num stocks: 1691
Capital: 0
Price of buy is: 7.670229
-----------------------
Day: 123
Selling on the last day!
The price of sell is: 7.95
The initial capital was: 10000
The final capital is: 13496
The profit made is:  3496
After: 123 days the capital is increased by 34.96 %.
```

Figure 16. List of Buy/Sell Decisions for VNET using Mean Value

The mean value approach was also applied to the second data set (AAPL). The achieved MRSE values from that are 3.257 for the low price predictions and 3.249 for the high price predictions. In comparison to the neural network, the mean value predictions are closer to the actual values than the one from the neural network. This shows that the mean value approach is better when there is a big change in the trend of the stock prices. 2 plots showing a comparison between actual and predicted low and high values for the mean value approach for the AAPL stock can be seen below.

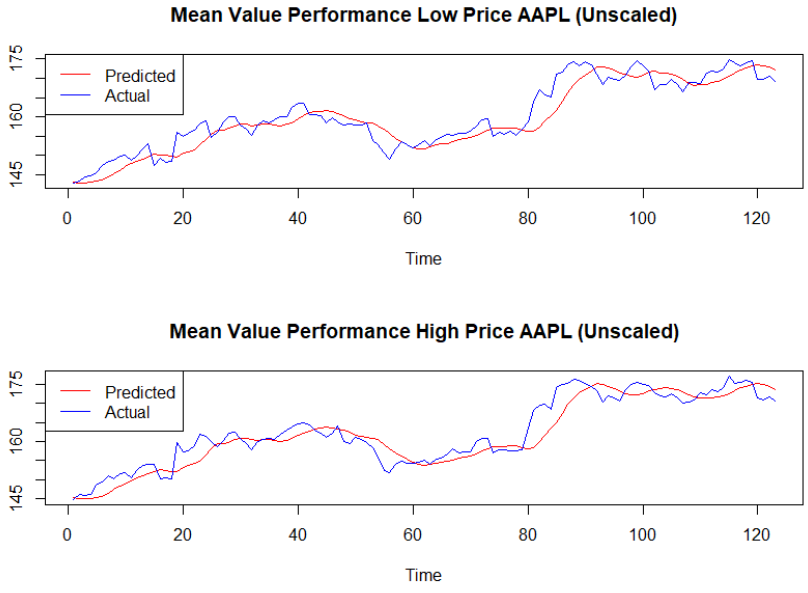

Figure 17. Mean Value Performance for Low and High Prices for AAPL

As it can be seen from the plots the predicted values again do not match the actual values and there are a few crossovers of the predicted and the actual values, which would give a few buy and sell decisions. Another thing to be noted is that even though the mean value approach has lower RMSE values than the neural network approach the neural network has a lot better predictions at first, but then deviates a lot in the end, while the mean value approach even though it has worse predictions their difference with the actual values remains the same throughout the whole dataset and are able to follow the drastic change in value which occurs at some point around day 80. Below are shown the buy and sell decisions based on those prediction.

```
Day:  1
Buy Stocks!
Num stocks: 69.23
Capital: 0
Price of buy is: 144.4
-----------------------
Day:  30
After holding for  28 days...
Sell stocks!
Num stocks: 0
Capital: 10979
Price of sell is: 159.3
-----------------------
```

```
Day:  31
Buy Stocks!
Num stocks: 68.76
Capital: 0
Price of buy is: 159.7
-----------------------
Day: 123
Selling on the last day!
The price of sell is: 169.2
The initial capital was: 10000
The final capital is: 11636
The profit made is:  1636
After: 123 days the capital is increased by 16.36 %.
```

Figure 18. List of Buy/Sell Decisions for AAPL using Mean Value

As it can be seen there are 2 buys and 1 sell occurring for the 123 day trading period, with a final sell on the last day after the market closes. The profit given is 16.36%, which surprisingly is worse than the one given by the neural network approach. The reason for that is even though both approaches buy twice and sell once before the end of the trading period, the neural network buys on day 18 for a lower price than the mean value which buys on day 31.

7. Conclusion

To sum up, the neural network approach described above provides a really good model for predicting daily low and high values, on which trading rules can be based on, but it is prone to errors if the stock is too volatile and does not follow a trend. If the stock's deviation remains within a small margin throughout the whole data set the approach proves to give quite a good profit. It also does a lot better against the mean value approach when using a less volatile stock, but a bit worse if the stock deviates a lot.

References

1. Evans, C., Pappas, K. and Xhafa, F. (2013). *Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation*.
2. Giacomel, F., Galante, R. and Pereira, A. (2015). *An Algorithmic Trading Agent Based on a Neural Network Ensemble: A Case of Study in North American and Brazilian Stock Markets*.
3. Investopedia. (2019). *Trading Strategy*. [online] Available at: https://www.investopedia.com/terms/t/trading-strategy.asp [Accessed 25 Mar. 2019]
4. Kirkpatrick, Charles D.; Dahlquist, Julie R. (2006). *Technical Analysis: The Complete Resource for Financial Market Technicians*. Financial Times Press.
5. Trippi, R. and Desieno, D. (1991). *Trading Equity Index Futures With a Neural Network*.