



Stock Market Prediction using Time Series Analysis and Neural Networks

Misati Nyambane - 670145

Bethlehem Kebede

Hoziyana Rachael

MTH1060: Analytical Computational Foundation

Dr. Francis Oketch

July 29, 2024

Abstract

Table of Contents

| | |
|--------------------------------------|---|
| Abstract | 2 |
| Methodology | 4 |
| Data Acquisition | 4 |
| Data Preprocessing..... | 5 |
| Machine Learning Models..... | 5 |
| Times Series Analysis | 5 |
| Neural Networks | 5 |
| Results and Discussion..... | 6 |
| Time Series Analysis..... | 6 |
| Neural Networks..... | 6 |
| Training | 6 |
| Testing..... | 6 |
| Validation..... | 6 |
| Conclusion and Recommendations | 7 |
| Time Series Analysis..... | 7 |
| Neural Networks..... | 7 |
| References..... | 8 |
| Appendices..... | 8 |

Methodology

Data Acquisition

Data was obtained from Yahoo Finance. The data spans from 1st July 2014 to 28th June 2024.

The data is in the general form:

Date, Open, High, Low, Close, Adj Close, Volume

2014-07-

01,23.3799991607666,23.517499923706055,23.282499313354492,23.3799991607666,20.68042755126953,152892000

2014-07-

02,23.467500686645508,23.514999389648438,23.272499084472656,23.3700008392334,20.671592712402344,113860000

2014-07-

03,23.417499542236328,23.524999618530273,23.299999237060547,23.50749969482422,20.793209075927734,91567200

The data was obtained using the Yahoo finance package in Python. By use of a dictionary the 20 companies selected were iterated through and their data written to CSV files. Yahoo Finance was preferred since it offers data for each company separately.

Data Preprocessing

Data was imported into MATLAB in bulk and stored in an array. This enables the use of iteration and control structures to access the data. Warnings are suppressed for cleaner output and if there are any missing values they are pointed out.

This allows for missing values to be dropped to ensure that operations are not skewed or curtailed by this data.

Furthermore, data is normalized between the range of 0 and 1 for our Time Series Analysis and Neural Network models to be able to work with our data. Only numerical columns are normalized. The column *date* is not affected in any way.

Machine Learning Models

Times Series Analysis

Neural Networks

We decided to use a Neural Network that:

- Has 40 hidden nodes
- Trains for a maximum of 1000 epochs
- Uses the adam optimizer
- Utilizes 75% of the data to train
- Utilizes 25% of the data to test
- Uses the Mean Squared Error to test the error of the results
- Utilizes validation data from 1st July 2024 to 26th July 2024

Results and Discussion

Time Series Analysis

Neural Networks

Training

Testing

| Company | MSE |
|--------------------------|-------------|
| AMD | 0.005859265 |
| Alphabet | 0.000862567 |
| Amazon | 0.000831416 |
| Apple | 0.007479319 |
| BHP Group Limited | 0.001295159 |
| Coca Cola | 0.006657724 |
| East Africa Metals | 0.003050332 |
| HSBC Holdings | 0.000289854 |
| Intel | 0.001157845 |
| Johnson and Johnson | 0.077928339 |
| L'Oreal | 0.026000365 |
| Meta | 0.123486883 |
| Microsoft | 0.041602214 |
| Nestle | 0.000945337 |
| Novartis AG | 0.007599726 |
| Proctor and Gamble | 0.001207186 |
| Samsung Electronics | 0.000384859 |
| Siemens AG | 0.021622344 |
| Tencent Holdings Limited | 0.000328486 |
| Toyota Motor Corporation | 0.156136595 |

Validation

| Company | MSE |
|--------------------------|----------|
| AMD | 0.011286 |
| Alphabet | 0.036998 |
| Amazon | 0.041988 |
| Apple | 0.044326 |
| BHP Group Limited | 0.015347 |
| Coca Cola | 0.051388 |
| East Africa Metals | 0.483313 |
| HSBC Holdings | 0.102384 |
| Intel | 0.072831 |
| Johnson and Johnson | 0.049414 |
| L'Oreal | 0.174741 |
| Meta | 0.028277 |
| Microsoft | 0.024886 |
| Nestle | 0.049855 |
| Novartis AG | 0.028158 |
| Proctor and Gamble | 0.126288 |
| Samsung Electronics | 0.020738 |
| Siemens AG | 0.154387 |
| Tencent Holdings Limited | 0.01493 |
| Toyota Motor Corporation | 0.017593 |

Conclusion and Recommendations

Time Series Analysis

Neural Networks

References

Appendices

| Company | Testing MSE | Validation MSE |
|--------------------------|-------------|----------------|
| AMD | 0.005859265 | 0.011285854 |
| Alphabet | 0.000862567 | 0.03699829 |
| Amazon | 0.000831416 | 0.041988428 |
| Apple | 0.007479319 | 0.044326173 |
| BHP Group Limited | 0.001295159 | 0.015347291 |
| Coca Cola | 0.006657724 | 0.051387616 |
| East Africa Metals | 0.003050332 | 0.483313323 |
| HSBC Holdings | 0.000289854 | 0.102384032 |
| Intel | 0.001157845 | 0.072831488 |
| Johnson and Johnson | 0.077928339 | 0.049413927 |
| L'Oreal | 0.026000365 | 0.174740765 |
| Meta | 0.123486883 | 0.028276792 |
| Microsoft | 0.041602214 | 0.024885528 |
| Nestle | 0.000945337 | 0.049855178 |
| Novartis AG | 0.007599726 | 0.028157622 |
| Proctor and Gamble | 0.001207186 | 0.126287847 |
| Samsung Electronics | 0.000384859 | 0.020737847 |
| Siemens AG | 0.021622344 | 0.154387357 |
| Tencent Holdings Limited | 0.000328486 | 0.014930324 |
| Toyota Motor Corporation | 0.156136595 | 0.017593062 |

Fig 1.1: Testing and Validation MSE's