Predicting the price of the FTSE/JSE Top40 Index using RNN Based models

Mokgethoa Papola
School of Computer Science and Applied Mathematics
University of the Witwatersrand
Johannesburg, South Africa
2089327@students.wits.ac.za

***Abstract*—Forecasting time series is challenging. Even recurrent neural networks, which can naturally learn sequentiality, find it challenging. This paper presents RNN based models for forecasting time series data (FTSE/JSE Top40 index price), covering data cleaning, EDA and prediction evaluation using RMSE. Comparisons between the Bi-LSTM, GRU, LSTM, and Bi-GRU models are made. These models are compared using different hyper-parameters to see which one performs best.**

***Index Terms*—ML, RNN, EDA, LSTM, BI-LSTM, Attention layer, GRU, BI-GRU, Loss and RMSE**

I. Introduction

These days the world is data driven and the analysis of raw data has helped to better improve life or to do decisive decisions. Just like any other data to the relevant industries, Stock Market data is the most key information to investors, forcing investors to deal with the time series data which is known for being non-linear, non-parametric, and very noisy. Stock price's movement is affected by many factors such as bank rates, corporate policies, political developments, commodity price indices, and overall economic conditions. Not only will it be difficult to predict stock market values, but it will also benefit investors.

A forecasting algorithm is a type of information method that combines prior and current data to estimate future values. Many sorts of decision-making and planning procedures, such as risk management, industrial process control, and finance, rely on forecasting future events. Analyzing time series data is used to forecast economic factors such as commodities, assets, and stock prices. In order to advise clients, stockbrokers attempt to forecast stock values using technical, or time series, or fundamental analysis [1].

The Financial Times Stock Exchange and the Johannesburg Stock Exchange produced the capitalization-weighted FTSE/JSE Top40 Index (JSE). This index includes the top 40 companies by market capitalization from the FTSE/JSE All Shares Index. This Series' goal is to depict the performance of South African companies. This is made possible by providing investors with a detailed and complementary collection of indexes that track the results of the South African market's major capital and industry segments. Subject to minimum free float and liquidity restrictions, the FTSE/JSE all share data reflects 99% of the total market capitalization of all ordinary stocks listed on the JSE's main board.

Artificial intelligence (AI) is a flourishing science with a plethora of practical uses and ongoing research areas. Numerous deep learning, and data scientist researchers are utilizing LSTM to forecast stock values ([2], [3]). This study uses the Gated Recurrent Unit (GRU), Bi-GRU, Bi-LSTM and Long Short Term Memory (LSTM) algorithms to predict the price of FTSE/JSE Top40 Index. It also checks the effect of adding the attention layer to the models.

A. Research Aim

This paper compares the stock market forecasting abilities of the RNN models(LSTM, Bi-LSTM, BI-LSTM with attention layer, GRU, and Bi-GRU).

B. Outline

The remaining of this study is organized as follows: Section II outlines what other researchers did in line with the problem statement. It outlines the results and conclusions from other researchers. Section III discusses the architecture of the model used. It also outlines the techniques that were performed during data cleaning and the statistical tests that were performed to understand the nature of the data set that was used. Section IV Discusses the results that were observed during the study period and gets into details to why we got the kind of results we have. It also highlights the limitations that were encountered during the study. Section V Outlines the conclusion and lesson that was drawn from the study and what can be done to improve our results in future.

II. Related work

There have been numerous studies on stock market predictions using various algorithms such as Support Vector Machine (SVM), Gated Recurrent Unit (GRU), Long Short Term Memory (LSTM), Tree-Based Classifiers, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and traditional/statistical techniques. However, because the objective of this research was to compare LSTM, BI-LSTM, GRU and BI-GRU, additional models and approaches will not be discussed in depth and will only be referred to when there is a need and it is also suitable to do so.

This review was conducted to determine what other researchers discovered about forecasting the stock using algorithmic machine learning and the effectiveness of the LSTM model, as well as what challenges they encountered.

LSTM was employed to forecast stock market prices and it was discovered that due to the nature of the data, the algorithm lost track of the opening price between day 600 and 700 of testing [6]. The model's training epochs is critical because it has a major impact on the model's accuracy [8].

It was discovered that the temporal lag between the expected and actual value is caused by a small period of time [3]. As a result of this discovery, it was concluded that the expected time delay will either have little or no impact on the model's validity. Hyper-parameters are known to be a very important aspect when doing predictions and when doing time series (Stock Market) predictions, a small batch size improves learning accuracy for the LSTM [9].

III. Methodology

A. Data

The data used for this study is primarily available without any cost. The data was collected from https://www.investing.com/. The name of the data is FTSE/JSE Top40 historical Data. The data's time range is from January 1, 2010 up to May 26, 2022. 3097 non-null entries make up the data. Seven features that make up the data are mentioned and described below:

- **Date:** is the date of the trading day.
- **Price:** is the current price that a share of stock is trading for on the market.
- **Open:** is the price from the first transaction of a business day.
- **High:** is the security's highest intraday price during a particular trading day.
- **Low:** is the security's lowest intraday price during a particular trading day.
- **Vol.** is the total number of shares that have been bought or sold during the trading day.
- **Change %:** is the difference between the current price and the last trade of the previous day.

This dataset was chosen because only few to no researchers used this dataset in their studies. Some trends on the dataset were easier to understand because I am a resident of South Africa.

B. Data Pre-processing

These sections explain the process of Data Cleaning, Feature Engineering as well as statistical tests that were used to visually analyze the data, discover trends, patterns, and to check the validity of our assumptions. No Sanity checks were performed in this study.

Before any data cleaning processes, Exploratory Data Analysis (EDA) was done so that we can understand the nature of our data and the features present in our dataset. Firstly, the correlation between features was conducted. It was observed that four features, Low, High, Open and Close have a strong positive correlation. This was the results of the columns having similar entries and it economically evident that they strongly affect each other.

It is further observed that the change in the stock market and the volume are not highly affected by the above mentioned features. This proves that the change in today's stock price is not affected by yesterday's closing and opening price but affected by external factors such as weather. All of this can be seen on figure 1 below.

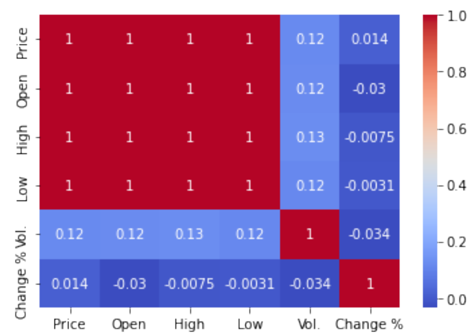


Fig. 1: Correlation Matrix.

Kwiatkowski Phillips Schmidt Shin (KPSS) Test and Dickey Fuller Test were performed to check for stationarity of the dataset and to check if the dataset set has any kind of trend. Based on the findings of the experiment, we may conclude that the dataset is nonstationary according to the ADF test(p-value = 0.6988 which is greater than 0.05) but, trend-stationary according to the kpss test(p-value = 0.01 which is less than 0.05). From the seasonal decompose experiment below (Figure 2), we can see that the dataset has a downward trend.

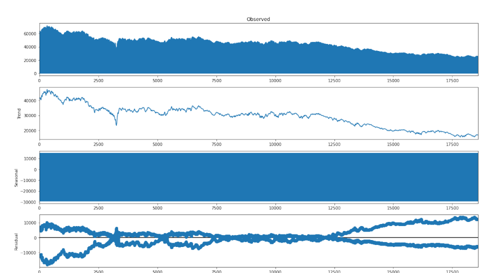


Fig. 2: Seasonal Decomposition.

Interpolation using the time method was used to handle the missing values. This method works similar to the Linear method since our date index is evenly spaced. To remove the trend, rolling mean was used. MinMaxScaler was used to normalise the dataset before it can be used to train the

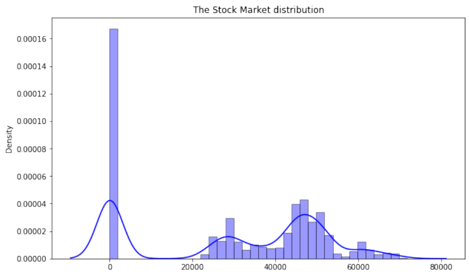


Fig. 3: The distribution.

model. The nature of the dataset and the distribution after data cleaning can be seen on Figure 4 and Figure 5 below.

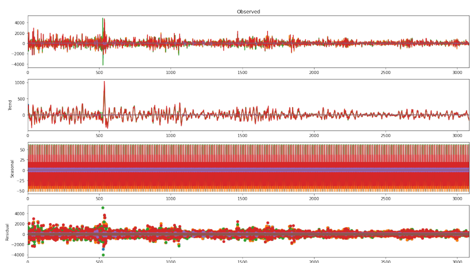


Fig. 4: Seasonal Decomposition

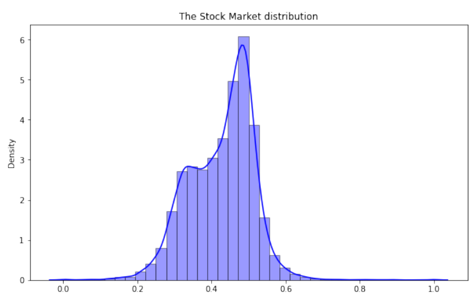


Fig. 5: The distribution.

C. Models

For this study, four (4) models (LSTM, BI-LSTM, GRU and GRU) were used to forecast the FTSE/JSE Top40. The LSTM was used as the benchmark. The benchmark's architecture was the same architecture used by [8]. This model produced the best results in the paper.

1) Long Short Temporary Memory: RNNs can learn long-term correlations, which is very beneficial in sequence prediction challenges. Aside from solitary data points like photographs, LSTM features feedback links that allow it to process the entire sequence of data. This has applications in time-series prediction, machine translation, and speech recognition, to name a few. An LSTM model performs exceptionally well on a wide range of tasks.

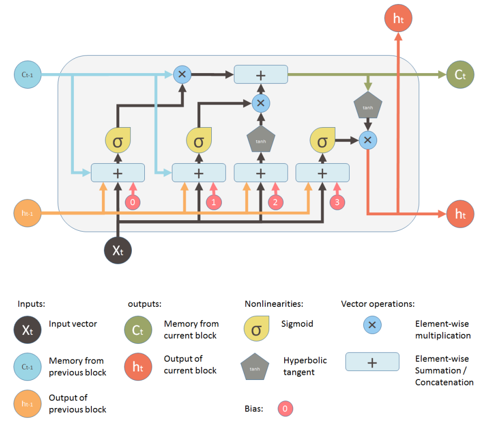


Fig. 6: A Structure of an LSTM

2) Bidirectional-LSTM: Is made up of two LSTMs, one of which receives input forward and the other backward. It is a sequence processing model. With the help of BI-LSTMs, the network has access to more information, which benefits the algorithm's context.

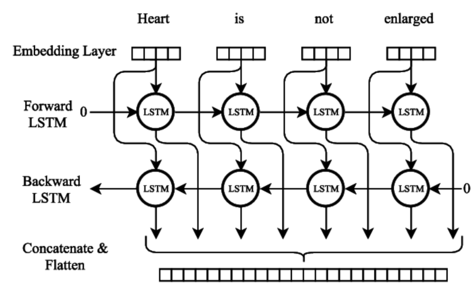


Fig. 7: A Structure of a BI-LSTM

3) Gated Recurrent Unit: GRU attempts to tackle the disappearing gradient problem that recurrent neural networks face. Because of the similarities in their designs and in some cases, GRU might be considered as an adaption of the LSTM. They have the unusual capacity to be trained to store information from the past without it fading away over time, as well as to discard information that is unrelated to the prediction.

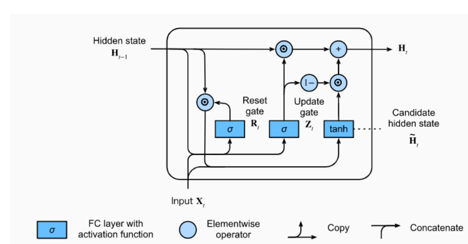


Fig. 8: A Structure of GRU

4) Bidirectional Gated Recurrent Unit: Is determined using the states of two unidirectional GRUs that point in opposite directions . The first GRU goes forward, starting at the beginning of the data series, and the second GRU moves backward, starting at the end of the data sequence. This enables knowledge from the past as well as the future to affect the conditions of the present.

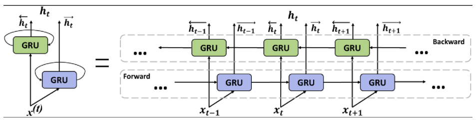


Fig. 9: Structure of BI-GRU

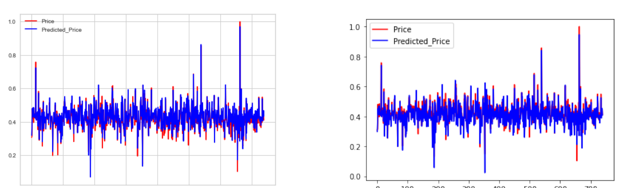
For each model, we used hyper-parameters with the best results.

IV. Results and Discussion

TABLE I: Model's performance

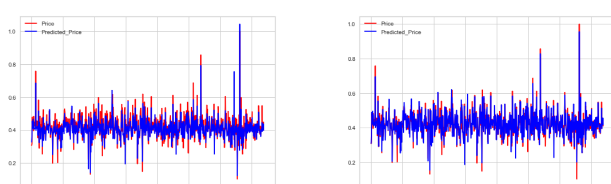
Models	Optimizer	Droput	No. Hidden Layers	RMSE
LSTM	RMSProp	None	3	0.0200
BI-LSTM	Adam	0.5	4	0.0202
GRU	RMSProp	0.5	4	0.0323
BI-GRU	Adam	0.5	4	0.0212
BI-LSTM with Attention	RMSProp	None	5	0.0210

From the Table above (I) we can see that the LSTM performed better than all the model. Just as expected, both the GRU models had the worst performance because we had a large dataset. LSTM model has a memory cell that can store information for lengthy periods of time and allows them to learn long-term dependencies.



(a) LSTM (b) BI-LSTM

Fig. 10: The LSTM and BI-LSTM forecast graphs



(a) GRU (b) BI-GRU

Fig. 11: The GRU and BI-GRU forecast graphs

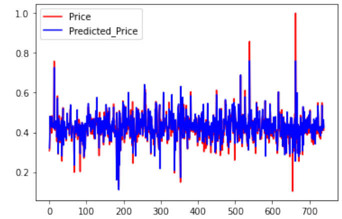


Fig. 12: BI-LSTM with 2 Attention layers

From the figures above we can see that all the models can depict the changes in the FTSE/JSE Top40 index. These shows that our modes can predict the price with a little error. The results we found aligns with what [7] has found during their research. Even though the Benchmark model was a replica of their best model, our RMSE was higher with 0.01141 because various stock sets affects the accuracy of the prediction [11].

A. Limitations

The study was conducted during the academic year, which meant that I had other academic activities that might had affected the depth of this research. The laptop used to run the models is a student HP notebook with Intel(R) Core(TM) i5-4210U CPU processor, 500GB hard drive and 8.00GB install ram which affected how fast we could train, validate and test the models as it took a while to run one model.

V. Conclusion

In order to anticipate the price of the FTSE/JSE Top40 index, this study suggests RNN-based approaches (LSTM, BI-LSTM, GRU, and BI-GRU) models. The outcomes of our models have produced some encouraging findings. The testing results show that our model can track the price evolution of the index. In our future study, we'll look for the optimum hyper-parameters to enhance the accuracy of ore predictions while better suiting our stock. Additionally, we will determine which features have a significant impact on how well our models work and only use those features when making predictions.

References

[1] W. Burdiharto, *"Data science approach to stock prices forecasting in indonesia during Covid-19 using long short-term memory (lstm)"*, Journal of Big Data, 2021. DOI: 10.1186/s40537-021-00430-0.

[2] C. Y. Lai, R.-C. Chen, and R. E. Caraka, *"Prediction stock price based on different index factors using lstm"*, 2020. DOI: 10.1109/ICMLC48188.2019.8949162.

[3] D. Wei, *"Prediction of stock price based on lstm neural network"*, 2020. DOI:10.1109/AIAM48774.2019.00113.

[4] R. Lawrence, *"Using neural networks to forecast stock market prices"*, 1997.

[5] J. Agrawal, V. Chourasia, and D. A. Mittra, *"State-of-the-art in stock prediction techniques"*, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2013.

[6] A. Moghar and M. Hamiche, *"Stock market prediction using lstm recurrent neural network"*, Precedia Computer Science, pp. 1168–1173, 2020.

[7] M. Roondiwala, H. Patel, and S. Varma, *"Predicting stock prices using lstm"*, International Journal of Science and Research(IJSR), 2017. DOI: 10.21275/ART20172755.

[8] M. A. I. Sunny, M. M. S. Maswood, and A. G. Alharbi, *"Deep learning-based stock price prediction using lstm and bi-directional lstm model"*, 2020. DOI:10.1109/NILES50944.2020.9257950.

[9] D. Shah, W. Campbell, and F. H. Zulkernine, *"A comparative study of lstm and dnn for stock market forecasting"*, 2018. DOI:10.1109/BigData . 2018 .8622462.

[10] J. L. Ticknor, *"A bayesian regularized artificial neural network for stock market forecasting"*, Expert Systems with Applications, pp. 5501–5506, 2013.

[11] Kai Chen, Yi Zhou and Fangyan Dai, *"A LSTM-based method for stock returns predictions: A case study of China stock market"*, IEEE International Conference on Big Data, 2015.

List of Figures