

Prediction of FTSE Bursa Malaysia KLCI Stock Market using LSTM Recurrent Neural Network

Tengku Nurul Aimi Balqis Tengku
Malim Busu
School of Quantitative Sciences,
College of Arts and Sciences,
Universiti Utara Malaysia
06010 UUM Sintok, Kedah, Malaysia
tengku_nurul_aimi@ahsgs.uum.edu.my

Saadi Ahmad Kamarudin
School of Quantitative Sciences,
College of Arts and Sciences,
Universiti Utara Malaysia
06010 UUM Sintok, Kedah, Malaysia
s.ahmad.kamaruddin@uum.edu.my

Nor Aishah Ahad
School of Quantitative Sciences,
College of Arts and Sciences,
Universiti Utara Malaysia
06010 UUM Sintok, Kedah, Malaysia
aishah@uum.edu.my

Nor Azura Md Ghani @ Mamat
Faculty of Computer & Mathematical
Sciences
Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia
azura@tmsk.uitm.edu.my

Abstract— Stock market prediction is vital in the financial world. Investors and people interested in investing would be interested in the future value of the stock market before they invest in it. By using the method of time series, this research gives a contribution to forecast and modelling the FTSE Bursa Malaysia KLCI (FBM KLCI) stock market. In this research, the stock market is forecasted to identify the stock market trend in the future. The FBM KLCI closing prices data was utilized to build Long Short-Term Memory (LSTM) models to predict the stock market. The performance of the model has been evaluated using the root mean squared error (RMSE) and the mean absolute error (MAE) in order to choose the best model. The researcher used the Bursa Malaysia data to forecast the stock market for five years, from October 20, 2016, to October 20, 2021, which has been scrapped from the Yahoo Finance website. The data is analyzed by running Python coding in Google Colab. The result proves that the accuration of the LSTM model by using Recurrent Neural Network (RNN) approach is accurate and the predicted value of the stock market at the date 2021-10-05 is increased by 1.87%. It can be used to predict the future closing stock prices in stock market prediction in FBM KLCI stock market. The results are expected to provide an accurate prediction for a better profit. Thus, prediction in stock market investment can support long-term economic growth, or in other words, it can help economic sustainability.

Keywords— FTSE Bursa Malaysia; Closing stock market; Time series; Long Short-Term Memory (LSTM) networks, forecasting

I. INTRODUCTION

Predicting stock prices is currently one of the biggest issues in the stock market. Due to its dynamic nature and features, a financial time series is represented by the data of stock prices which is getting harder to anticipate [1]. Extreme stock market changes affect the worldwide economy and financial markets. Therefore, we require a more accurate method of forecasting market changes. Some of the numerous forecasting attempts over the past few decades have been successful using quantitative approaches such as Support Vector machines, Artificial Neural Networks (ANN) and Autoregressive Integrated Moving Average (ARIMA) model, but financial economists continue to debate these techniques due to the nonlinear nature of stock market behavior [2-5].

In many research areas, “deep learning” (DL) got a lot of interest recently. Without applying econometric assumptions, DL can extract abstract characteristics and find hidden linkages in financial markets, regardless of the degree of complexity or the presence of nonlinear and linear big data financial market components [6]. The other quantitative methodologies and traditional finance economic methods are unable to accomplish this. Therefore, in this paper, the researcher is interested in predicting the stock market by using Long Short-Term Memory (LSTM). The LSTM is a synthetic recurrent neural network (RNN) architecture with internal neuronal feedback that can be used in deep learning networks. The ability to remember key past events and take into account prior experience is made possible by this internal feedback [7].

The main purpose of this study is to predict the future relative price value FTSE Bursa Malaysia KLCI stock market using the LSTM model. To achieve the main objective, the specific objective, which is to identify the best LSTM model and the stock market trend within five years using the Malaysian stock market, needs to be satisfied.

The following section, the literature review, discusses the review from the other paper regarding the time series forecasting that used the LSTM model. After that, the methodology part will explain the data collection and time series forecasting using the LSTM model. Then, the analysis and results are presented in the third part. Last but not least, the conclusion is discussed in the last part.

II. LITERATURE REVIEW

In this part, the review of the previous research has been discussed. Extrapolation is used in the classical statistical treatment of time series data to make predictions. Nowadays, many fields focus on the topic of time series forecasting. Forecasting entails fitting models to historical data and then utilizing those models to forecast future events.

Big data processing and computational power have increased because of the usage of DL, which also makes it possible for more complex algorithms. Previous research shows that DL can unravel the problems of nonlinear more effectively than conventional techniques [8-10].

Recently, a lot of research has used the LSTM model to predict the stock market as the approach to analyze the stock market. According to [11], the LSTM model can be used to follow the development of opening prices for GOOGL and NKE assets. Moreover, based on the comparison between the ARIMA model and the LSTM model, the accuracy of the LSTM model is more superior compared to the ARIMA model [12].

III. METHODOLOGY

This section provides a detail explanation regarding data collection and the process of analysis using the RNN.

A. Data Collection

This study is quantitative research. This research refers to a set of methods and procedures that describe variables (quantitative approaches are used for prediction analysis). Moreover, the future value can be predicted based on the method of forecasting the time series data. The type of data used in this study is secondary data. In this research, the 1234 daily data from 20th October 2016 until 20th October 2020 were scraped from the Yahoo Finance website and utilized to estimate the future stock market. The data has been analyzed in the Google Colab website by running the Python code using the ANN model. The analysis follows the research framework as illustrated in Fig.1.

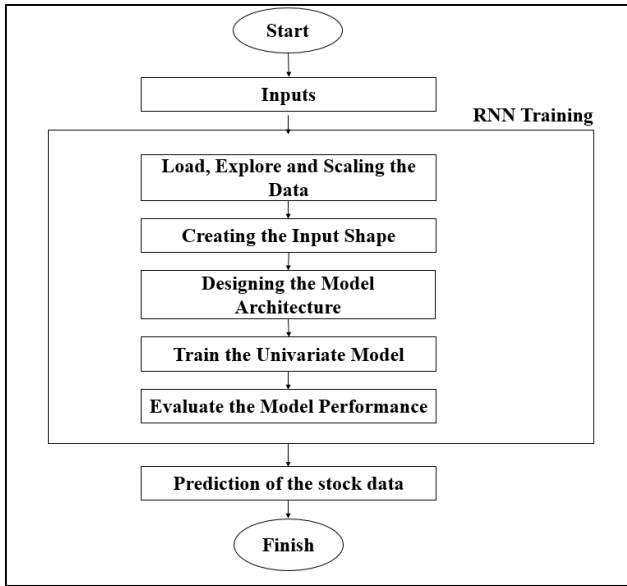


Fig. 1. Research Framework

B. Methods

This study is quantitative research. This research refers to a set of methods and procedures that describe variables (quantitative approaches are used for prediction analysis). Moreover, the future value Recurrent Neural Network (RNN) model is one of the models that can be used in the prediction of future value. RNN consists of at least three layers as shown in Fig. 2.

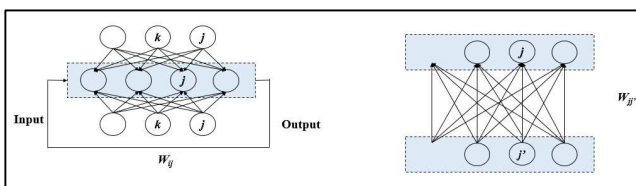


Fig. 2. The structure of the recurrent neural network

The RNN has three types of connection where matrix V , matrix W and matrix U weight connects the hidden-to-output, hidden-to-hidden, and input-to-hidden respectively. The first step in forward propagation is defining the hidden unit j_0 starting state. Then, we use the subsequent update equations for each time step starting at j_0 . The hidden neuron h input value at time t is expressed as

$$h_j^t = \sum_i u_{ji} x_i^t + \sum_{j'} w_{jj'} z_{j'}^{t-1}, \quad (1)$$

where u_{ji} is the weight between hidden neuron j and input neuron i x_i^t is input value at time t . $w_{jj'}$ is the weight between hidden neuron j and j' and $z_{j'}^{t-1}$ is the output value of hidden neuron j' at time $t-1$.

C. Model Performance Evaluation

When making a prediction to find the best model, there are a lot of factors that need to be considered to get the best model [13]. The factor that the researcher wants to check to find the best model is the forecast accuracy measures which are Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

The following is the formula for those two measures [7]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n |(Y_i - \hat{Y}_i) / Y_i| \times 100 \quad (3)$$

where n is the predictions number, \hat{Y}_t is the forecasted value and Y_t is the actual value. A fit model for forecasting is the one that yields the smallest value for each criterion.

IV. DISCUSSIONS

The result in this section has been analyzed using Python in Google Colab and SPSS software.

i. Load, Explore and Scaling the Data

Here the researcher checks the number of available GPUs and the version of TensorFlow. The number of available GPUs is 0 and the version of TensorFlow is 2.8.2.

Before starting the analysis, the data has been load in Google Colab by using Python coding. Table I. presents the data that the researcher has scrapped from Yahoo Finance.

TABLE I. RESEARCH DATA

	Date	High	Low	Open	Close	Volume	Adj Close	Name
0	10/20/2016	1.06	1.05	1.05	1.05	12900	0.916992	0002.KL
1	10/21/2016	1.05	1.05	1.05	1.05	0	0.916992	0002.KL
2	10/24/2016	1.05	1.05	1.05	1.05	10000	0.916992	0002.KL
3	10/25/2016	1.05	1.04	1.05	1.04	52600	0.908259	0002.KL
4	10/26/2016	1.06	1.05	1.05	1.06	87000	0.925725	0002.KL
...
1230	10/13/2021	3	2.95	3	2.95	8300	2.95	0002.KL
1231	10/14/2021	3	2.91	2.92	2.99	7800	2.99	0002.KL
1232	10/15/2021	2.98	2.9	2.9	2.92	12600	2.92	0002.KL
1233	10/18/2021	3	2.92	2.92	3	22000	3	0002.KL
1234	10/20/2021	2.92	2.92	2.92	2.92	400	2.92	0002.KL

In this study, the researcher use the closing prices to predict the stock data. Therefore, before doing the forecast on the stock data, the researcher explore to see the line plot for the FTSE Bursa Malaysia KLCI closing prices. The line plot has been illustrated in Fig 3.



Fig. 3. Line Plot for FTSE Bursa Malaysia KLCI Closing Price

ii. Evaluate the Model Performance

Table II. shows that the value of MAE is 0.08. From the MAE value, the researcher can say that the model generally tends to predict a bit too pessimistic. Moreover, the MAPE value that the researcher get from the analysis using the LSTM RNN model is 2.9%. It indicates that the accuracy of the model is 97.1% which indicates that the accuracy of the model is very good.

TABLE II. ERROR MEASURE

Error Measure	Error (%)
Median Absolute Error (MAE)	8
Mean Absolute Percentage Error (MAPE)	2.9
Median Absolute Percentage Error (MDAPE)	2.68

iii. Prediction of the Stock Data

The prediction of the closing stock price is increasing. Fig 4. also presents the test predictions at the orange zone. The grey area indicates where test predictions and actual data diverge. Many performance measurements have previously shown that the projections are often close to reality.

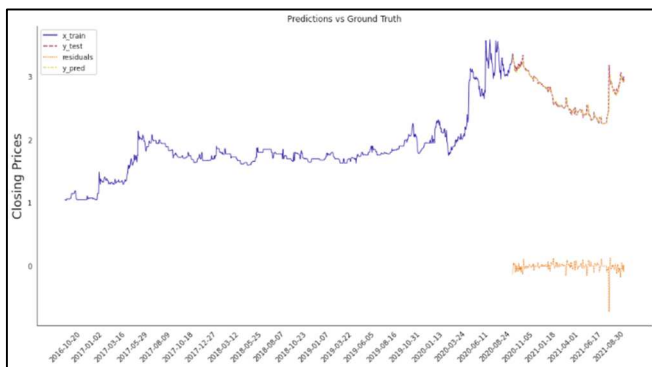


Fig. 4. Stock Market Prediction using LSTM RNN model

The stock market price for FBM KLCI on 2021-10-05 was 1.05, while the predicted FBM KLCI stock market price is 2.92. So, it can be seen that the different between the actual and predicted is 1.87%.

TABLE III. PREDICTION RESULT

Error Measure	Value
The close price for FBM KLCI at the date 2021-10-05	1.05
The predicted close price	2.9200000762939453 (-178.1%)

V. CONCLUSION

All the objectives have been successfully achieved. According to the outcome, the model's performance based on the MAE value indicates that it typically predicts outcomes that are a little too pessimistic. Other than that, the accuracy of the model is very good.

Next, the second objective which is to identify the trend of the stock market within five years using the Malaysian stock market has also been achieved based on the decomposition time series, it shows that there is an increasing trend in the stock market within five years. Then, the forecasts of the FBM KLCI closing prices at the date 2021-10-05 is 2.92. The finding from this study show that the forecast model can be used in the investment of the future stock prices and also can help in economic sustainability.

ACKNOWLEDGMENT

The Fundamental Research Grant Scheme (FRGS/1/2021/STG06/UUM/02/4) of the Malaysian Ministry of Higher Education (MoHE) provided funding for this study. We also like to express our gratitude to the Universiti Utara Malaysia (UUM) for the support to carry out this research work.

REFERENCES

- [1] D. Selvamuthu, V. Kumar, and A. Mishra, "Indian stock market prediction using artificial neural networks on tick data," *Financial Innovation*, vol. 5, no. 1, 2019.
- [2] A. Ang and G. Bekaert, "Stock return predictability: Is it there?" *Rev. Financial Stud.*, vol. 20, no. 3, pp. 651–707, Jul. 2006.
- [3] M. Qi, "Nonlinear predictability of stock returns using financial and economic variables," *J. Bus. Econ. Stat.*, vol. 17, no. 4, pp. 419–429, Oct. 1999.
- [4] M. Aiolfi and C. A. Favero, "Model uncertainty, thick modelling and the predictability of stock returns," *J. Forecasting*, vol. 24, no. 4, pp. 233–254, Jul. 2005.
- [5] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Comput. Oper. Res.*, vol. 32, no. 10, pp. 2513–2522, Oct. 2005.
- [6] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Syst. Appl.*, vol. 83, pp. 187–205, Oct. 2017.
- [7] H. Chung, and K. Shin, "Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction," *Sustainability*, vol. 10, no. 10, pp. 3765, 2018.
- [8] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014.
- [9] X.-W. Chen and X. Lin, "Big data deep learning: Challenges and perspectives," *IEEE Access*, vol. 2, pp. 514–525, 2014.
- [10] X. Jin, J. Shao, X. Zhang, W. An, and R. Malekian, "Modeling of nonlinear system based on deep learning framework," *Nonlinear Dyn.*, vol. 84, no. 3, pp. 1327–1340, May 2016.
- [11] A. Moghar, and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," *Procedia Computer Science*, vol. 170, pp. 1168–1173, 2020.

- [12] S. U. Siame-Namini, and A. S. Namin, "Forecasting economics and financial time series: Arima vs. lstm," arXiv preprint arXiv:1803.06386, 2018.
- [13] D. C. Montgomery, Murat Kulahci, and C. L. Jennings, "Introduction to time series analysis and forecasting," Wiley, 2016.