Sains Malaysiana 52(3)(2023): 993-1009
http://doi.org/10.17576/jsm-2023-5203-22

A Comparative Study of Deep Learning Algorithms in Univariate and Multivariate Forecasting of the Malaysian Stock Market

(Kajian Perbandingan Algoritma Pembelajaran Mendalam dalam Peramalan Univariat dan Multivariat Pasaran Saham Malaysia)

Mohd.Ridzuan Ab. Khalil[1] & Azuraliza Abu Bakar[2,*]

[1]Malaysian Administrative Modernisation and Management Planning Unit (MAMPU), Federal Government Administrative Centre, 62502 Putrajaya, Federal Territory, Malaysia
[2]Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

Received: 8 August 2022/Accepted: 28 December 2022

ABSTRACT

As part of a financial institution, the stock market has been an essential factor in the growth and stability of the national economy. Investment in the stock market is risky because of its price complexity and unpredictable nature. Deep learning is an emerging approach in stock market prediction modeling that can learn the non-linearity and complexity of stock market data. To date, not much study on stock market prediction in Malaysia employs the deep learning prediction model, especially in handling univariate and multivariate data. This study aims to develop a univariate and multivariate stock market forecasting model using three deep learning algorithms and compare the performance of those models. The algorithm intends to predict the close price of the Malaysian stock market using the Axiata Group Berhad and Petronas Gas Berhad from Bursa Malaysia, listed in Kuala Lumpur Composite Index (KLCI) datasets. Three deep learning algorithms, Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM), are used to develop the prediction model. The deep learning models achieved the highest accuracy and outperformed the baseline models in short and long-term forecasts. It also shows that LSTM possessed the best deep learning model for the Malaysian stock market, achieving the lowest prediction error among the other models. Deep learning demonstrates the ability to handle univariate and multivariate data in preserving important information, thus forecasting the stock market. This finding is relatively significant as deep learning works well with high-dimensional datasets.

Keywords: CNN; deep learning; LSTM; MLP; multivariate; stock forecasting; time series; univariate

ABSTRAK

Pasaran saham merupakan sebahagian daripada institusi kewangan yang menjadi faktor penting dalam pertumbuhan dan kestabilan sesebuah ekonomi negara. Pelaburan dalam pasaran saham adalah sangat berisiko disebabkan oleh perubahan harganya yang rumit dan sifatnya yang sukar untuk diramal. Pembelajaran mendalam adalah satu pendekatan baharu yang semakin menonjol dalam ramalan pasaran saham kerana ia mampu mempelajari data pasaran saham yang tidak linear dan rumit. Sehingga kini, tidak banyak kajian yang dilakukan mengenai ramalan pasaran saham di Malaysia menggunakan pendekatan pembelajaran mendalam khususnya yang melibatkan pendekatan data univariat dan multivariat. Penyelidikan ini dijalankan untuk membangunkan model ramalan pasaran saham univariat dan multivariat menggunakan tiga algoritma pembelajaran mendalam dan seterusnya membuat perbandingan prestasi antara model tersebut. Ia akan meramal harga tutup di pasaran saham Malaysia menggunakan data saham Axiata Group Berhad dan Petronas Gas Berhad dari Bursa Malaysia dan turut tersenarai di dalam Indeks Komposit Kuala Lumpur (KLCI). Tiga algoritma pembelajaran mendalam iaitu *Multilayer Perceptron* (MLP), *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM) digunakan untuk membangunkan model ramalan. Hasil uji kaji menunjukkan model pembelajaran mendalam mencapai ketepatan yang tinggi dan mengatasi kesemua model dasar bagi ramalan untuk tempoh jangka pendek dan panjang. Ia juga menunjukkan LSTM merupakan model pembelajaran mendalam yang terbaik untuk pasaran saham Malaysia dengan ralat ramalan yang paling rendah berbanding kesemua model lain. Pembelajaran mendalam menunjukkan keupayaan yang ketara dalam membuat ramalan pasaran saham menggunakan data univariat dan multivariat. Penemuan ini adalah signifikan dengan keupayaan pembelajaran mendalam terutamanya dalam mempelajari set data yang bersifat multidimensi dan mempunyai fitur yang banyak.

Kata kunci: CNN; LSTM; MLP; pembelajaran dalam; multivariat; ramalan saham; siri masa; univariat

Introduction

The stock market is a financial institution that directly influences national economic growth. It also acts as an indicator for each country to represent the economy's stability. The long and short-term influence of the stock market could drastically affect and transform the country's economic growth and investment rate (Azam et al. 2016; Masoud 2013). The ability of a government to maintain a sustainable and viable stock market is critical since it could ensure the stability and growth of its economy. A strong stock market will attract the investor involved in the market transaction, increasing the volume of trade in the market, raising capital and opportunities for the related company to expand their growth, and indirectly strengthening the country's economic development. Pradhan (2018) has described the relationship and bilateral effect between the stock market and economic growth, where both directly impact each other. Hence, the stability of the stock market and the country's economy is crucial to ensure the growth and development of the country.

In Malaysia, Bursa Malaysia is the organization responsible for coordinating and managing the stock market. It provides the platform and facility for the investor to transact efficiently in the stock market. The existence of this organization will ensure a transparent and open market which will give the investor confidence and interest to involve in the Malaysian stock market (Hafizah & Saiful 2013). Formerly known as Bursa Saham Kuala Lumpur (BSKL), it was established in 1930, while the public trading of shares commenced in 1960. As a benchmark and indicator to represent the stock market's strength in a country, the stock market index is used as a reference by the investor (Hafizah & Saiful 2013). Kuala Lumpur Composite Index (KLCI) was launched in 1986 to represent the Malaysian stock market performance. In 2009, Bursa Malaysia took the initiative to maintain its relevance in the global market by improving the stock market index using the standard methodology of the Financial Times Stock Exchange (FTSE) index. This approach is an internationally accepted index calculation methodology that will provide a much better representation of the overall stock market and transparently index. KLCI was rebranded with a new name and was known as FTSE Bursa Malaysia KLCI. It comprises the largest 30 companies listed on the Main Board and meets the terms and requirements of the FTSE Bursa Malaysia Ground Rules.

The ability to analyze and understand the behavior of the stock market is critical to the investor and should be the primary approach before they can invest in the market. Investment in the stock market is high-risk because price changes are unpredictable and could drastically lead to huge losses. If the analysis and prediction of the stock market movement are made accurately, investors could earn a high-profit return from their investment significantly when the purchased stock's price increases. Investors usually use two traditional approaches to forecast the stock market price: fundamental analysis and technical analysis (Abu-Mostafa & Atiya 1996; Bustos & Pomares-Quimbaya 2020; Chung & Shin 2020). This approach helps the investor predict the future stock price movement, guiding them to make a precise decision for their investment and minimizing the risk of loss. However, both approaches consume much time to observe and analyze detailed information about the stock market. The stock market and its price movement are difficult to predict because of its dynamic and random behaviour. It is also filled with noise, non-stationary, and nonlinear, making stock market price forecasting difficult (Abu-Mostafa & Atiya 1996). Hence, forecasting the stock market with high accuracy is challenging in the investment industry. However, this does not hinder various parties from exploring and researching different methods and techniques to forecast the stock market, either from the academic or investment industry.

Deep learning is a modern approach widely used in various fields, including the financial industry. It is an evolution from a machine learning technique developed based on the Artificial Neural Network (ANN) algorithm.

It uses a modern and complex architecture to learn complex data for accurate prediction. Deep learning tries to imitate and mimic the human brain by combining large data, weightage, and bias configured in the algorithm. These elements will try to identify, classify, and learn the information and pattern in the data precisely. Deep learning is widely used in various fields, such as image recognition (Pak & Kim 2018), health (Zeroual et al. 2020), and the financial industry (Livieris, Pintelas & Pintelas 2020). It has proven excellent ability and performance compared to the traditional machine learning approach (Rama Krishna et al. 2020; Sezer, Gudelek & Ozbayoglu 2020). Much recent research on time series forecasting uses deep learning for accurate stock market forecasting (Ismail Fawaz et al. 2019). However, the deep learning approach in Malaysia's stock market forecasting research is still open.

This study aims to develop univariate and multivariate stock market forecasting using deep learning algorithms. It shows the ability of the methods to handle the univariate and multivariate time series data, which seek to preserve important information, thus, able to forecast the stock market accurately. We showed that although the multivariate forecasting model offers slightly lower accuracy, it can preserve more information than the univariate model. This finding is relatively significant as deep learning works well with high-dimensional datasets.

This paper is organized into six sections; next section explores the work on deep learning in stock market forecasting. We present the deep learning methods and the experiment design in the following sections. Subsequently, we discussed the deep learning forecasting results and comparative analysis. Finally, the concluding remarks are presented in the last section.

RELATED WORK

The deep learning approach in stock market forecasting and prediction has been gaining interest among researchers because of its ability to efficiently learn nonlinear and complex stock market data. Its superiority has been proven through much research related to stock market forecasting with various types of stock data and deep learning algorithms. Mehtab, Sen and Dutta (2021), in their study, compared the different traditional machine learning algorithms with a deep learning algorithm for stock market forecasting. Tata Steel and Hero Moto Corporation data from India Stock Exchange are used to train and test all prediction models. The data consist of stock price every 5 minutes from 2013 to 2014. Multivariate Regression, DTR, Bagging, Boosting, RFR, ANN, Support Vector Machine (SVM), and LSTM algorithms are used to develop a prediction model. Open price, close price, low price, high price, volume, and 11 new variables are input features for the prediction models to predict the future price. The results show that the LSTM model outperforms the other machine learning models, including the ANN model, with the lowest Root Mean Square Error (RMSE) value compared to the others. While this study has proved that deep learning performs better than machine learning models, there is no comparison between different deep learning algorithms to evaluate their performance. In another study, Lee et al. (2021) used the LSTM algorithm with four hidden layers to forecast the Taiwan Stock Exchange Index. They also compared four additional technical indicators as input features for the algorithm to predict the stock index. The scope of the study covers the prediction of the future index movement and the stock price using daily data from 2017 until 2019. The previous 20 days' stock prices are used as the window size of the input feature for the prediction model. The result shows that the LSTM model with the Moving Average Convergence/Divergence (MACD) indicator achieved an average accuracy of 75% and the highest compared with other indicators. In comparison, combining all four additional indicators with other features achieved a prediction accuracy of up to 83%. While the study shows an approach to boost the accuracy of the LSTM model with additional technical indicators as input features, there is no comparison made with the baseline model or another deep learning algorithm to analyze the performance of the proposed LSTM model.

In another study, Mehtab, Sen and Dasgupta (2020) compares the performance of two deep learning algorithms, CNN and LSTM, to forecast the future stock price of Bharat Forge on the Indian Stock Exchange. Daily stock data from December 2012 until January 2015 are used to forecast the price for the next five days. The previous 5 and 10 days of data are used as the window size of the input features for both deep learning algorithms. The input features used are open, close, high, low, and volume for the multivariate model, while close for the univariate model. Various configurations and input features are used to see the performance of the models to forecast the future closed price. Experimental results show that all deep learning models with various configurations and input features achieve excellent performance with low RMSE values. The result also indicates that the CNN models outperform the LSTM model with the lowest RMSE values.

Nabipour et al. (2020) used three deep learning algorithms to forecast stock market prices on Tehran Stock Exchange, Iran. ANN, Recurrent Neural Network (RNN), and LSTM are used to develop the prediction model using stock data from 2009 until 2019. Several machine learning models are used as a baseline model to compare the performance of those deep learning models. Ten technical indicators are used together with the stock market price data as the input features for all the prediction models. Each model will forecast the stock price for the next 1, 2, 5, 10, 15, 20, and 30 days. Experimental results show that the LSTM model achieved the best performance for every prediction length with the lowest RMSE, MAPE, and MAE values.

Liu and Long (2020) proposed a framework to forecast the next day's close price of the stock market using the LSTM algorithm. They employed the Empirical Wavelet Transform (EWT) and Outlier Robust Extreme Learning Machine (ORELM) methods for data pre-processing. Three different univariate stock data S&P 500, Dow Jones Index, and China Minsheng Bank, and previous ten days' close price are used as the input features for the prediction models. Results from several previous studies are used as a baseline to compare the performance of the proposed framework and model. Experimental results and analysis show that the proposed LSTM model outperforms the results obtained by the previous study with the lowest RMSE, MAE, and MAPE values.

Maiti and Shetty (2020) used LSTM with four hidden layers to predict the price movement and the close price of five different stocks on the Turki Stock Exchange using the datasets with five-minute intervals from 2014 until 2019. Open price, close price, high price, low price, volume, and technical indicator are the input features with a window size of 50. The model will predict the price movement and forecast the close price for the next interval. The accuracy of the LSTM model in predicting the price movement for all five stocks is between 97.53% and 98.91%. The model shows excellent performance for price forecasting with a low RMSE value between 0.024 and 0.0048. Both results have proved the LSTM algorithm's ability to learn the stock data's complexity.

Sismanoglu et al. (2019) used the LSTM to forecast stock prices from New York Stock Exchange (NYSE), NASDAQ, and NYSE MKT. The datasets used were from 1968 to 2018, which consist of daily stock price data, with input features such as open price, close price, high price, low price, and volume. The prediction result shows LSTM obtained good performance in forecasting the stock price with lower error. The RMSE value is 0.04, which indicates that LSTM could predict the stock price with minimum error. However, no comparison is made with any baseline model or previous research to prove the performance of the proposed LSTM model.

Fister et al. (2019) proposed an automated stock trading model using a deep learning approach, specifically LSTM. The LSTM model predicts the daily price movement of BMW stock in Germany. Comparison is made with several traditional techniques, such as passive and rule-based methods, technical indicators, and machine learning methods, to evaluate the performance. The LSTM model will predict the price movement and simulate the stock transaction to buy, sell or hold based on the prediction. Stock price and several technical indicators are used as input features to predict the price movement from the following first day until the fifth day. Experimental result shows that the LSTM model outperforms the other baselines, including the traditional machine learning model, in predicting price movements. The simulation also proved that this model could gain and achieve higher profit than the other methods.

Cao et al. (2019) proposed the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) algorithm to enhance the prediction performance of the LSTM model. Four stock composite indexes from different countries are used; S&P 500 from the United States, Hang Seng Index (HSI) from Hong Kong, Deutscher Aktien Index (DAX) from Germany, and Shanghai Stock Exchange Composite Index (SSE) from China. Dataset was collected from 2007 until 2017, where only close prices were used as the input feature to forecast the next day's price. The performance of the CEEMDAN-LSTM model is compared with LSTM, SVM, CEEMDAN-SVM, and CEEMDAN-MLP to evaluate the ability of the model to forecast all the composite indexes. The final analysis shows that CEEMDAN-LSTM outperforms the other models with the lowest RMSE, MAE, and MAPE values.

In another study, Chung and Shin (2020) focused on the CNN algorithm to predict the movement of the Korea Composite Stock Price Index (KOSPI). The index's history from 2000 until 2016 was used as the dataset, where seven additional technical indicators are created to be used as the input features. The Genetic Algorithm (GA) systematically enhances all the parameters in the CNN model. ANN and CNN without GA are used as baseline models to evaluate the performance of the proposed model to predict the KOSPI index. Experimental result shows that CNN without GA outperforms the ANN model with accuracy achieved at 70.16% while ANN only at 58.62%. When the GA is applied to the CNN model, the accuracy increases to 73.74%, which is much better and outperforms the original CNN model.

Zulqarnain et al. (2020) proposed a CNN algorithm with Gated Recurrent Unit (GRU) to predict the next day's index movement of the Hang Seng Index (HSI), S&P 500, and DAX. Historical data from 2008 until 2016 is used. The previous 250 days to the index are utilized as the size of the input feature windows. GRU, Deep Neural Network (DNN), and CNN as baseline models to evaluate the proposed model's performance. Experimental result shows that the GRU-CNN model can predict all three indexes with more than 50% accuracy. The model scores 56.2% for HIS, 56.1% for DAX, and 56.3% for S&P 500. These results are much better than what has been achieved by the baseline model for all three indexes.

Mehtab and Sen (2020) in their study, used CNN to forecast the NIFTY 50 index in India using nine multivariate input features from 2015 until 2019. The previous five days' data are used as the window size for the input feature to forecast the index for the next fifth and tenth days. The Multivariate Regression, DTR, Bagging, Boosting, RFR, ANN, and SVM are baseline models to evaluate the CNN model. Analysis and result show that the CNN model using the multivariate input feature outperform all the baseline model with the lowest RMSE value.

A comparison of several deep learning model performances with the ARIMA model to forecast the stock market price is made by Hiransha et al. (2018). This study uses different datasets to train and test the model. Stock data from the National Stock Exchange of India is used to train the model, while three other stock data from NSE and two from NYSE are used to test all the models. The previous 200 days' close price is used as the window size for the input feature, and the model predicts the close price for the next ten days. Four deep learning algorithms, MLP, RNN, LSTM, and CNN, were used to develop the prediction model and the ARIMA model as a baseline model. The result shows that all deep learning models outperform the ARIMA statistic model with the lowest MAPE value. Overall, the CNN model achieved the best performance predicting all five stock prices with the lowest MAPE value compared to the other deep learning models.

Yong, Abdul Rahim and Abdullah (2017) chose the MLP algorithm in their study to predict the movement of the Straits Time Index in Singapore Exchange for decision support in a stock transaction. Historical data from 2010 until 2017 are used with a window size of 10 days as the input feature configuration. The input features consisting of open price, close price, high price, and low price are used to predict the index movement for the next day until the fifth day. The MLP model was developed using three hidden layers with ten neurons assigned for each layer. The result shows that the model could predict the index movement with a high accuracy rate and profit from the stock transaction simulation.

Nikou, Mansourfar and Bagherzadeh (2019) compared the ability of the deep learning model with several traditional machine learning models to predict the MSCI United Kingdom stock index. Historical data from 2015 until 2018 are used with a window size of 10 days chosen as the feature input configuration. LSTM is developed as the prediction model with the MLP, SVR, and RFR as the baseline model. Analysis and results show that LSTM is the best model to predict the stock index compared to MLP, SVR, and RFR models. This study suggests that deep learning is the best method to predict the future stock index compared to other traditional methods. It also indicates that the deep learning model could be used as decision support for investors and stock market analysts for investment purposes.

Meanwhile, research on stock market prediction in Malaysia has been done by Al-Mashhadani et al. (2021), where Maybank stock data from 2016 until 2017 are used as the dataset. Several sets of input features are used in this study: close price, differences between high and low price, and volume. The previous five days' data is used as the window size for the input features. ANN is used to predict the price movement for the next day. The experimental result shows that the proposed model could predict price movement with a high accuracy rate. A univariate approach using only close price as the input feature achieved the highest accuracy rate compared to the other input feature set. Soon et al. (2018) also used the ANN algorithm to predict the stock price of CIMB. Historical data from 2000 until 2015 are used as the dataset with a window size of 5 days chosen as the input feature configuration to predict the next day's price. The input features are close price, KLCI index, interest rate, and foreign exchange rate. Nonlinear Autoregressive Network with Exogenous Input (NARX) is used as a baseline model to evaluate the performance of the ANN model. Analysis and results show that ANN outperforms the NARX model with better prediction performance and lower MSE value. Yiing and Thim (2015) compared the performance of the ANN model with the ARIMA model to predict the KLCI index. Data from the years 2012 until 2015 are used in this study to develop the prediction model. Experimental result shows that ANN has the lowest RMSE value compared to the ARIMA model.

There is also some evolved version of deep learning algorithm, such as AlexNet and ResNet, based on the CNN architecture. AlexNet was proposed by Krizhevsky, Sutskever and Hinton (2012), while ResNet was introduced by He et al. (2016). Even though both are preferred approaches for image recognition tasks, some researchers have used them in stock market prediction. Zhao and Khushi (2020) proposed a wavelet denoised ResNet model to predict the Forex price rate of change. The price dataset is denoised using wavelet denoising, which is then used to calculate the technical indicators. Different configurations of the ResNet and CNN models are also used to compare the performance of the proposed model. The result shows that the proposed model outperforms the other baseline model with better prediction ability.

Pan, Li and Li (2020) transformed raw time series data of the China stock market into an image dataset by plotting it into candlestick charts. AlexNet, ResNet, DeseNet, ShuffleNet, and VGG are used to extract the features from the dataset. It is then used to find a low-risk and high-return portfolio. Experimental results show that ResNet achieved the best results in most indicators and the best overall performance.

Liu and Song (2018) used ResNet to predict stock price trends from China Market Stock Index. The raw dataset was transformed into a candlestick graph as an image dataset for the ResNet model. Six additional models are also used to evaluate the proposed model's performance comprehensively. The results show that ResNet could provide more practical information and find the hidden trend of the stock price concealment compared to the other models.

This discussion of the related study showed the significant performance of various deep learning algorithms in stock market prediction around the globe. MLP, CNN, and LSTM are much preferred for stock market time series data because they manage hidden information efficiently. The evolved AlexNet and ResNet seem to be used more significantly on image recognition than the stock market price dataset. As deep learning is yet to be employed in the Malaysian stock market, this study offers three deep learning algorithms, MLP, CNN, and LSTM, for modeling the univariate and multivariate stock market data. The significant difference of this work is it evaluates the prediction performance in several periods to determine the capability of deep learning in predicting at the earliest time, which positively impacts potential investors.

DEEP LEARNING ALGORITHMS

This study employed three deep learning algorithms for predictive model development. The algorithms are Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM). The algorithms have specific features which are based on neural network algorithms.

Multilayer Perceptron (MLP)

MLP is a subset of the ANN algorithm, where both have the same fundamental architecture and can be classified under the same category (Jiang 2021). The concept of ANN was founded and introduced by Warren McCulloh dan Walter Pitts until it was widely used to model various types of nonlinear processes (Nikou, Mansourfar & Bagherzadeh 2019). It was developed based on the concept of neuron structure in the human brain. MLP is a neural network consisting of one input layer, one output layer, and at least one hidden layer. Each hidden layer has neurons that will process the input with the weightage and bias configured in the algorithm. Each neuron is connected directly to all neurons from the previous layer and creates a fully connected neural network. MLP and ANN are among the most popular deep learning approach used in stock market forecasting (Bustos & Pomares-Quimbaya 2020). Both are quite simpler compared to other deep learning algorithms and can efficiently learn nonlinear data's complexity. However, high dimensional datasets and the increment of neurons in the architecture will affect the model's performance in learning the data (Sezer, Gudelek & Ozbayoglu 2020). The structure of MLP will treat time series data as independent and has no relationship with the previous information. This will affect the learning process, where past information will be ignored to predict future values.

Convolutional Neural Network (CNN)

CNN was introduced by a researcher named Yann LeCun in the era of 1980s (Hoseinzade & Haratizadeh 2019). It has a different architecture that focuses on the input features and does not has any records or memory about past information (Selvin et al. 2017). CNN is usually used with grid topology data such as images where the information will be represented in two dimensional grid for rows and columns. CNN architecture has three main layers: a convolutional layer, a pooling layer, and a fully connected layer. The convolutional layer is the primary layer where the process of extracting meaningful information from the data will be done. In this layer, the convolutional operation is performed between the input feature and filter, known as the kernel. The kernel will slide over rows and columns of the input feature to select the most relevant and important information. The final output, known as the feature map, will then be fed to the pooling layer. In this stage, the size of the convolved feature map will be decreased to reduce the computational costs. A feature sample with important information representing the whole data will be selected (Chung & Shin 2020). This process will also minimize the noise in the data and create much better and clean data before being fed to the next layer. The convolutional layer and pooling layer can be repeated in the architecture to develop multiple and deep layers. The fully connected layer is the final layer in CNN which has the same design as MLP and CNN. It has neurons connected to the pooling layer and convolutional layer. The final output or result will be processed and generated from this layer based on the features received from the previous layer.

Even though CNN is widely used for analyzing visual images, its application in stock market forecasting has become popular among researchers because of its ability to extract the important features from the dataset (Hoseinzade & Haratizadeh 2019). It allows the algorithm to learn the past important information and predict future values. It also has the advantage of noisy data where only relevant information from the data will be selected.

Long Short-Term Memory (LSTM)

LSTM is an enhanced version of Recurrent Neural Network (RNN) and was introduced in 1997 by Sepp Hochreiter and Jurgen Schmidhuber. They invented LSTM to solve the missing gradient problem that occurred in the RNN algorithm (Cao et al. 2019). RNN and LSTM can store past information from the input features and use it to learn and predict future values. However, LSTM has huge advantage where it can store past information for a long term in the memory. This advantage has made the LSTM algorithm the most chosen approach for time series data. LSTM has a memory block or cell that represents the neural network. Each cell has three gates: the input gate, the output gate, and forget gate. It allows LSTM to manage the information efficiently, whether storing or forgetting past information. This architecture and design could overcome the long-term dependency problems faced by most other algorithms. In stock market data, LSTM has the advantage of learning the dependency and relationship between the data (Livieris, Pintelas & Pintelas 2020).

Materials and Methods

The implementation and development of the proposed prediction model are done in Python programming language version 3.7. We utilized the Google Colab platform, which offers Python development tools in the cloud to do the data pre-processing and model development. Google Colab offers an easy-to-configure environment and free access to GPUs to run deep learning algorithms.

The methods for this study is based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. CRISP-DM is an open standard process model that serves as a base and describes approaches used for a data science process. The lifecycle of this model consists of six primary phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. Figure 1 shows the CRISP-DM model used in this study.

Business Understanding

Stock market forecasting, prediction, and the needs of the stakeholders have been the significant factor and goals for the implementation of this study. Many methods and techniques predict the movement and the stock market price. However, deep learning has been widely explored because of its robustness and high prediction accuracy. This approach is much more flexible and reduces the cost and time to develop such a model compared to other traditional methods. The advantages of the deep learning model meet the investor's need to have a fast and precise reference for them to gain and increase the chances to maximize their profit. Hence, based on the findings and understanding of the business goals, this study aims to fill the gap and requirement by developing a stock market prediction model based on a deep learning algorithm. Investors can use it as guidance and reference to decide the best possible option and action in their stock market investment.

Data Understanding

Dataset used in this study is obtained and downloaded from the Yahoo Finance portal. Axiata Group Berhad from the telecommunication sector and Petronas Gas Berhad from the gas and utility sector are chosen. Five

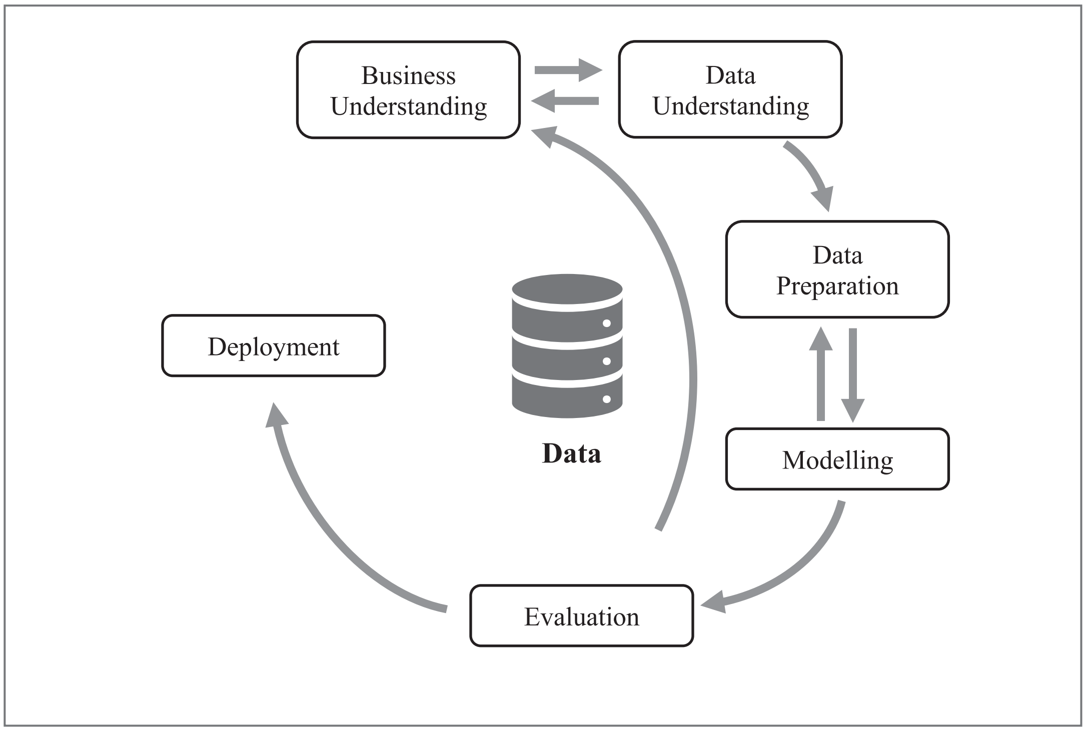

FIGURE 1. Lifecycle of CRISP-DM model (Source: http://www.ibm.com)

years of historical data for both profiles from January 2017 until December 2021 is used. The five years is selected because it can give the model enough precise information and details to learn and predict the stock price. The extended period will risk the model learning irrelevant information and price pattern, affecting the future prediction results (Shen & Shafiq 2020). Both datasets consist of six numerical features and a date which acts as the time stamp. The values of Open, Close, High, Low, and Adj Close features are the stock price-related information in Ringgit Malaysia (RM), while volume is the feature that records the daily trading volume for that stock. The date feature is time stamp information that records the stock transaction's daily date.

The Open feature stores the opened stock price, while the Close feature is the stock closing price on that day before the market closes. High and Low refer to the highest and lowest price during that transaction day. Adj Close feature is the closing price after deducted by certain charges. Usually, the investor will use the close price as an indicator for their decision to sell, buy or hold the stock. The close price will determine whether the investment on that transactional day will gain profit or loss (Table 1).

TABLE 1. List of features in the dataset

Feature name	Data type	Description
Date	Interval	Date of the transaction
Open	Numeric	Open price (RM)
High	Numeric	Highest price (RM)
Low	Numeric	Lowest price (RM)
Close	Numeric	Closing price (RM)
Adj Close	Numeric	Adjusted closing price
Volume	Numeric	Volume of trading

DATA PREPARATION

Data pre-processing is crucial because it will prepare the data in the most meaningful way and make it easier to interpret and use. The quality of the data and the information derived from it will affect the performance and ability of the model to learn. It also prepares the data in a format and structure that allows the model to learn and interpret it efficiently. In this phase, we conduct three main processes; data cleaning, data reduction, and data transformation. Data cleaning is the main activity in the early data pre-processing phase, improving the dataset's quality and representation. Data imputation and correction were performed to fill and replace the missing or incorrect values in the dataset. The missing value should be treated with data imputation or removal for the affected records as long as it will not affect the dataset's quality. Since the closing price is the target feature for the prediction, this feature will remain and be used for both univariate and multivariate predictions. Open, High, Low, and volume features are used in multivariate prediction, while the Adj Close feature is removed from the dataset. The final structure of the features after implementing data reduction is shown in Figure 2.

	Open	High	Low	Close	Volume
Date					
2017-01-02	4.318914	4.318914	4.318914	4.318914	0
2017-01-03	4.181661	4.199961	4.117609	4.135910	3061100
2017-01-04	4.145060	4.145060	4.081008	4.126760	4707200
2017-01-05	4.117609	4.264013	4.108459	4.254863	2626100
2017-01-06	4.254863	4.511069	4.227412	4.511069	11006200

FIGURE 2. Structure of the dataset after data reduction

Data transformation refers to the process and technique to convert the structure and representation of the data into a suitable and ideal format for the model to learn. It also will enhance the ability of the model to learn and interpret the information and pattern of the data efficiently. The feature values in the dataset are numeric and in different ranges; therefore, normalization will be done to scale the values into a smaller and similar range between 0 and 1. This will allow the model to learn the data efficiently and faster compared to the original values. It will also avoid bias among the features since the original data have a different range of values.

The raw data structure is not prepared in a supervised format for the model to learn. Thus, the sliding window technique that will restructure the time series dataset as a supervised format is applied. This technique will use specific last-time steps or historical data known as lag as input variables and specific next time steps as the target or output variables. This transformation will allow the data to be used to train the model before it can predict the future price. This study uses a window size or lag of 10 days to transform the dataset into the supervised format. The size is chosen based on the previous related study done by Liu and Long (2020), Nikou, Mansourfar and Bagherzadeh (2019), and Yong, Abdul Rahim and Abdullah (2017). Figure 3 shows the concept of a sliding window in a time series problem that will transform data structure into a supervised format.

For predictive modeling, the sliding window technique will be applied to prepare the dataset in univariate and multivariate forms. The univariate dataset consists of only one parameter, which is the close price target features, for the prediction. Meanwhile, in the multivariate dataset, several features are included in the dataset. We use the window size of 10 days lag for both datasets.

Model Development

We employed three deep learning algorithms for the predictive model development; MLP, LSTM, and CNN. The performance of the models is compared with three baseline algorithms. Each model will be trained to learn and interpret that data's hidden information, relationship, and pattern to map the output values.

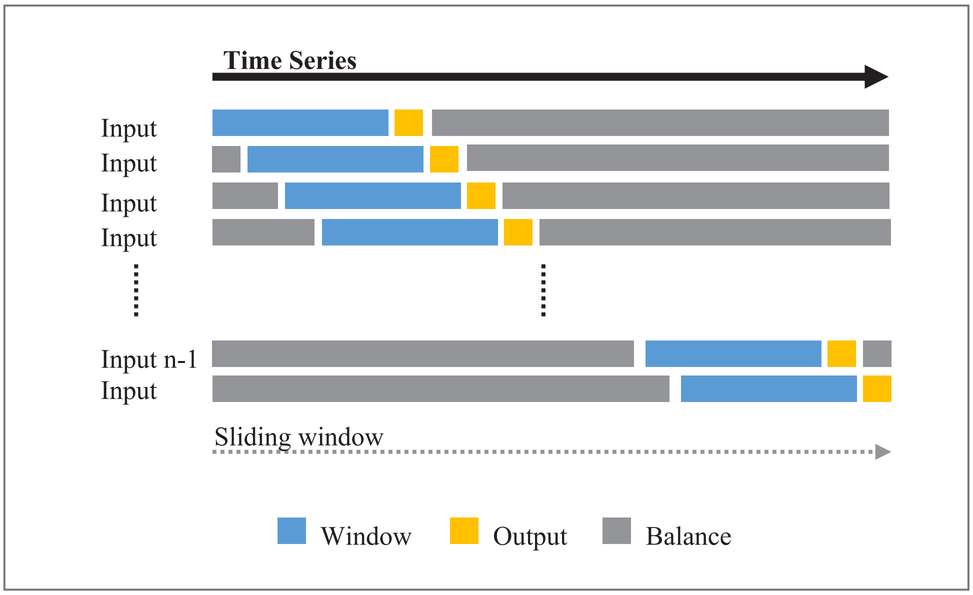

FIGURE 3. Sliding window technique

The models are tested on unseen data to evaluate the prediction accuracy. We used the training and testing split as an 80:20 ratio giving the 987:246 data split.

Table 2 shows the hyperparameter setting used for this study's MLP, CNN, and LSTM models. Two hidden layers are applied to the MLP model, and additional input and output layers. This configuration will create a deep learning layer to interpret and learn the information from the data and predict future values. For each hidden layer, 200 neurons are assigned to process the input and pass that information to the output layer.

We also used two convolutional layers in the CNN model to create the deep neural network. It allows the algorithm to extract the data's most important and relevant information. These layers will be combined with a pooling and fully connected layer to create the design of the proposed CNN model. The matrix size known as the kernel is set to 3 for filtering and extracting the information from the data. For the LSTM model, two hidden layers are applied together with an input and output layer. The proposed design will create a deep neural network allowing the algorithm to learn and predict efficiently. One hundred nodes are assigned to each hidden layer to process and learn the data and pass the information to the output layer. Figure 4 shows the topological design of the proposed deep learning models.

TABLE 2. Hyperparameter configurations for MLP, CNN, and LSTM model

Hyperparameter configurations for MLP model		Hyperparameter configurations for CNN model		Hyperparameter configurations for LSTM model	
Parameter	Configuration	Parameter	Configuration	Parameter	Configuration
nodes	200	filters	256	nodes	100
epochs	150	kernel	3	epochs	200
batch_size	24	epochs	100	batch_size	32
activation	relu	batch_size	32	activation	relu
optimizer	adam	pool_size	2	optimizer	adam
loss	mse	activation	relu	loss	mse
learning_rate	0.001	optimizer	adam	learning_rate	0.001
		loss	mse		
		learning_rate	0.001		

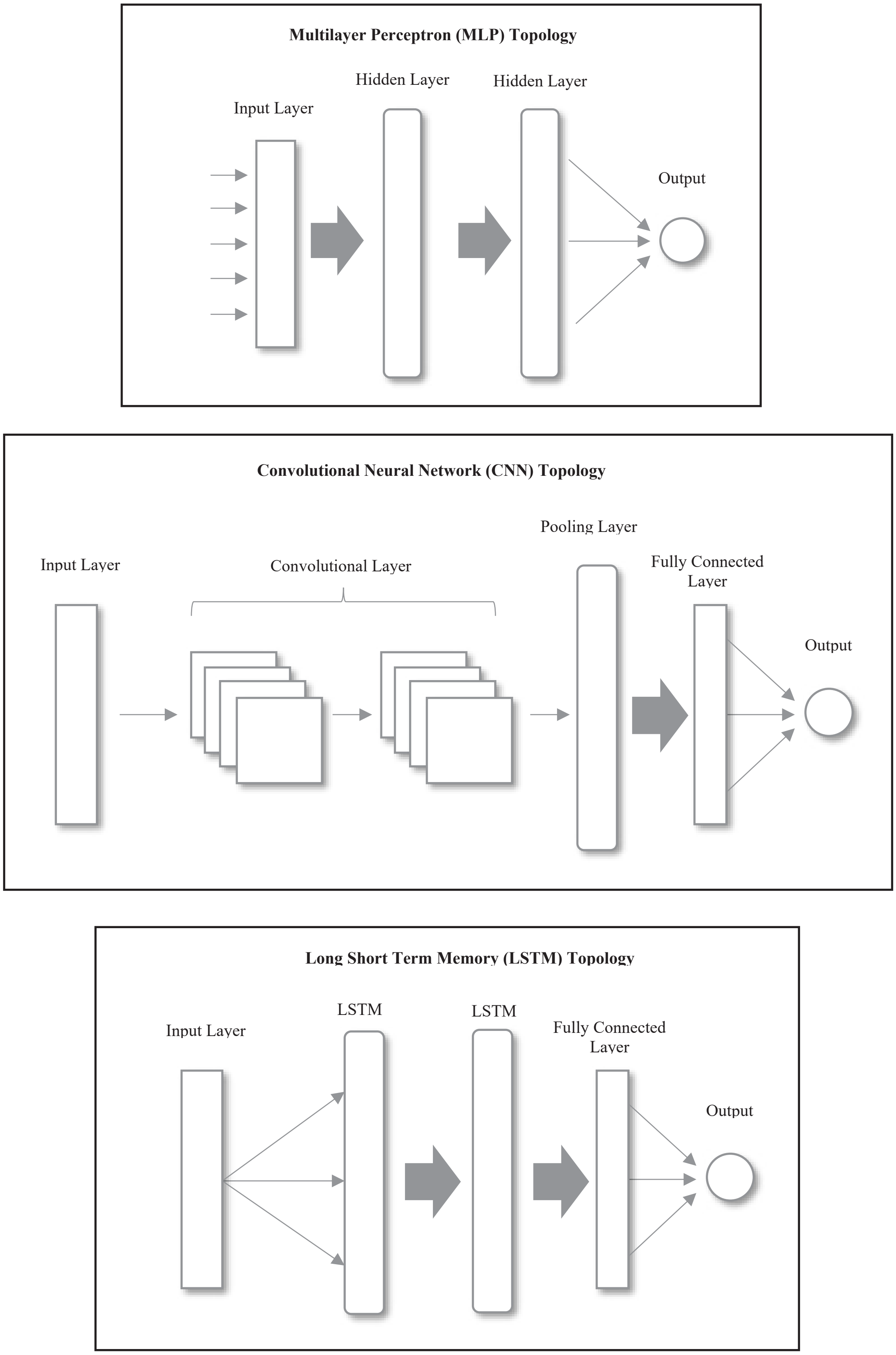

FIGURE 4. MLP, CNN and LSTM topological design

The univariate and multivariate models will predict the closing price of both Axiata Group Berhad and Petronas Gas Berhad for the 1st, 3rd, 5th, and 7th days. Using the univariate and multivariate approaches, six models, and different prediction periods, a total of 96 predictions are made.

Evaluation

The trained models are tested with the same test dataset to analyze the closing price prediction. Based on the predicted value, performance evaluation is made using three standard evaluation metrics for regression tasks; Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These evaluation metrics will represent an indicator to measure each model's prediction error and period. RMSE, MAE, and MAPE are widely used in regression tasks, including stock market prediction (Bustos & Pomares-Quimbaya 2020; Gandhmal & Kumar 2019; Jiang 2021). The RMSE, MAE, and MAPE use the difference between predicted and observed values as a base for each metric. Thus, a model with the lowest RMSE, MAE, and MAPE values is the best model with high ability and prediction performance compared to a model with higher values.

Results and Analysis

The closing price predicts both companies' stock data for the 1st, 3rd, 5th, and 7th days. Both univariate and multivariate approaches are applied to each model to see the ability and performance obtained using different input features. Based on the prediction made, RMSE, MAE, and MAPE are calculated and used to evaluate the performance achieved by each model. The univariate and multivariate predictions will also be compared to determine the differences in performance between both approaches. The overall evaluation results of each model for all univariate and multivariate prediction periods are shown in Figure 5. The experiment results show a consistent pattern of errors where the metric evaluation values increase with the prediction period. The prediction accuracy decreases once the period increases from day 1 to day 7. Overall, deep learning models obtained consistent results for all prediction periods and outperformed other models with lower RMSE, MAE, and MAPE values. The performance also seems consistent for both stock market data for all prediction periods using the univariate and multivariate approaches.

The univariate and multivariate models give a slight difference in prediction errors. The univariate

Prediction on 1st day for Axiata Group Berhad

	Axiata Group Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.1084	0.0829	2.2096	0.1176	0.088	2.3452
RFR	0.0803	0.0595	1.5886	0.0761	0.0574	1.5397
SVR	0.0978	0.0769	2.0995	0.1121	0.0925	2.5148
MLP	0.0739	0.0555	1.4873	0.08	0.0607	1.6296
CNN	0.0734	0.0557	1.4999	0.0784	0.0608	1.6317
LSTM	**0.0692**	**0.0517**	**1.3861**	**0.0717**	**0.0538**	**1.4496**

Prediction on 1st day for Petronas Gas Berhad

	Petronas Gas Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.2394	0.1794	1.1198	0.2349	0.1744	1.0894
RFR	0.1806	0.1268	0.7873	0.1786	0.1245	0.7726
SVR	0.1849	0.1318	0.8209	0.2015	0.1478	0.916
MLP	0.1677	0.1211	0.7543	0.1659	0.1192	0.7424
CNN	0.1646	0.1207	0.7521	0.1673	0.125	0.7825
LSTM	**0.1623**	**0.1171**	**0.7281**	**0.1646**	**0.1241**	**0.7805**

Prediction on 3rd day for Axiata Group Berhad

	Axiata Group Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.1785	0.1423	3.805	0.432	0.3744	9.9356
RFR	0.1276	0.102	2.7344	0.192	0.1644	4.4026
SVR	0.1902	0.1501	4.1241	0.2001	0.1733	4.5186
MLP	0.1122	0.0896	2.418	0.1236	0.097	2.6265
CNN	0.1124	0.0903	2.4364	0.1265	0.0985	2.6616
LSTM	**0.11**	**0.0884**	**2.3734**	**0.1182**	**0.0912**	**2.4731**

Prediction on 3rd day Petronas Gas Berhad

	Petronas Gas Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.3509	0.2543	1.5818	0.65	0.5167	3.2797
RFR	0.2479	0.1823	1.132	0.3085	0.2372	1.4968
SVR	0.2871	0.2208	1.3828	0.6496	0.4711	2.8812
MLP	0.2348	0.1736	1.0798	0.2591	0.1957	1.2202
CNN	0.2401	0.1769	1.1013	0.2546	0.1916	1.197
LSTM	**0.2314**	**0.1685**	**1.0502**	**0.247**	**0.1796**	**1.1212**

Prediction on 5th day for Axiata Group Berhad

	Axiata Group Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.2225	0.1727	4.6478	0.4049	0.3521	9.3403
RFR	0.1537	0.1248	3.3466	0.1616	0.1318	3.5087
SVR	0.2814	0.2279	6.2718	0.1831	0.1549	4.0603
MLP	0.1381	0.1095	2.9761	0.1421	0.1144	3.0914
CNN	0.1509	0.1261	3.3546	0.1615	0.1283	3.4823
LSTM	**0.1394**	**0.1105**	**3.0073**	**0.1368**	**0.1086**	**2.9348**

Prediction on 5th day for Petronas Gas Berhad

	Petronas Gas Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.4421	0.3396	2.1128	1.1292	0.8743	5.5997
RFR	0.3005	0.2238	1.3909	0.4591	0.377	2.4002
SVR	0.3524	0.283	1.7781	0.6122	0.4526	2.7802
MLP	0.2514	0.1912	1.1901	0.2858	0.2116	1.3202
CNN	0.3112	0.235	1.4457	0.2873	0.2125	1.3271
LSTM	**0.2677**	**0.1979**	**1.2353**	**0.281**	**0.2098**	**1.3159**

Prediction on 7th day for Axiata Group Berhad

	Axiata Group Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.268	0.1967	5.2591	0.4408	0.3715	9.882
RFR	0.166	0.13	3.4638	0.1891	0.1487	3.9303
SVR	0.3686	0.3065	8.427	0.2748	0.2419	6.2955
MLP	0.1645	0.1312	3.4834	0.1503	0.1208	3.2546
CNN	0.1405	0.1113	3.0039	0.1586	0.1276	3.4301
LSTM	**0.1396**	**0.111**	**2.9799**	**0.1464**	**0.1161**	**3.1093**

Prediction on 7th day for Petronas Gas Berhad

	Petronas Gas Berhad					
	Univariate			Multivariate		
Method	RMSE	MAE	MAPE	RMSE	MAE	MAPE
DTR	0.4977	0.3828	2.3796	0.9341	0.7657	4.9175
RFR	0.3342	0.2519	1.5655	0.5073	0.4065	2.5915
SVR	0.4007	0.3295	2.0746	0.6261	0.4558	2.7928
MLP	0.3312	0.2551	1.5816	0.3798	0.2924	1.814
CNN	0.3168	0.2424	1.5053	0.3509	0.2649	1.6461
LSTM	**0.3122**	**0.2384**	**1.4826**	**0.3112**	**0.2386**	**1.4941**

FIGURE 5. Overall performance of prediction algorithms for Axiata and Petronas dataset in four different days

models show lower RMSE, MAE, and MAPE values than the multivariate. This difference is well known as the univariate only uses one feature to predict the price, which has a minimal feature. The performance of the multivariate model prediction is also significant as the error difference is small and insignificant in choosing the best model. The results also show that multivariate prediction is still competitive enough to give the best prediction results for both stock market data. Even though univariate has a slight advantage over the multivariate with lower errors, both approaches have their advantages and disadvantages, which influence the prediction results.

Figures 6, 7, and 8 compare RMSE, MAE, and MAPE for both univariate and multivariate models' performance achieved by the best model for each prediction period. The RMSE, MAE, and MAPE values achieved by univariate and multivariate prediction are competitive. The error changes patterns and the increment of the prediction period are consistent for both models giving a slight difference in all metrics. Multivariate models offer competitive results with low and minimum prediction error. It also has the advantage and ability to preserve and carry other information and variables along with the model. For further analysis of univariate and multivariate prediction, a statistical test is done to compare the prediction errors of both approaches. The T-test is chosen to calculate the statistically significant values of both models. The null hypothesis stated an insignificant difference between both prediction accuracy, while the alternative hypothesis stated that the difference significantly exists between both approaches. The test is done using the predicted values for both univariate and multivariate prediction with confidence intervals $\alpha = 0.05$. The test will accept or reject the null hypothesis with 95% confidence. The results from the test are shown in Table 3.

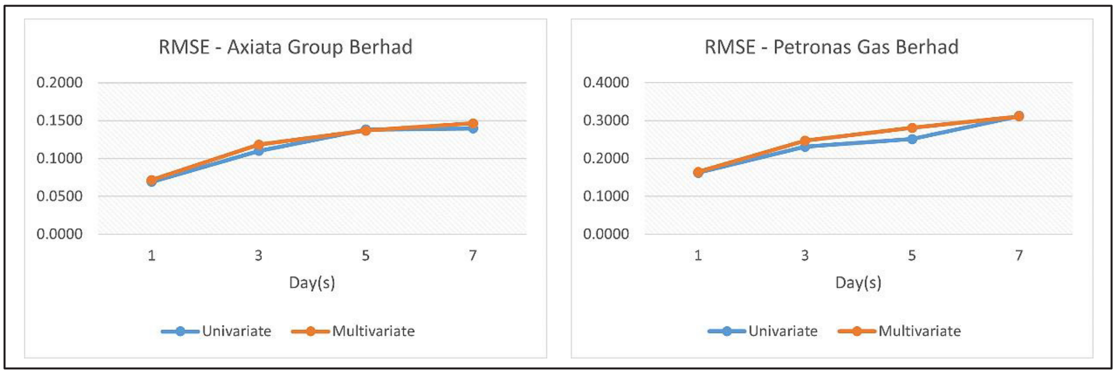

FIGURE 6. Comparison of best RMSE value between univariate and multivariate

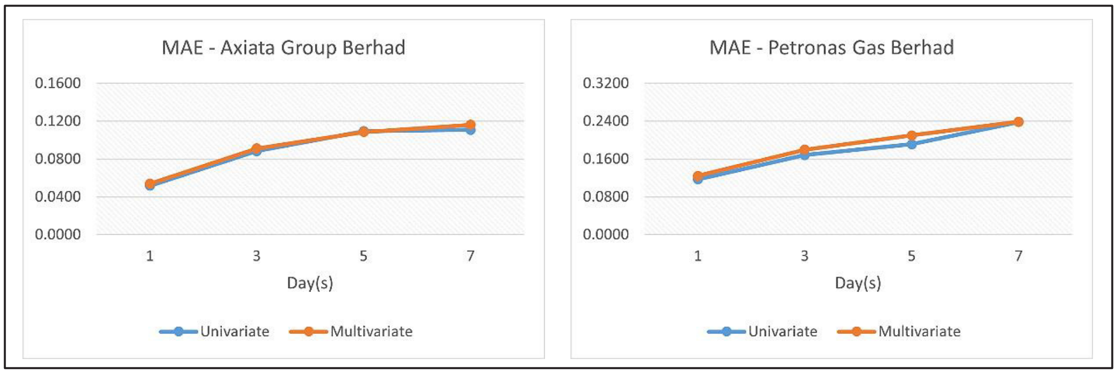

FIGURE 7. Comparison of best MAE value between univariate and multivariate

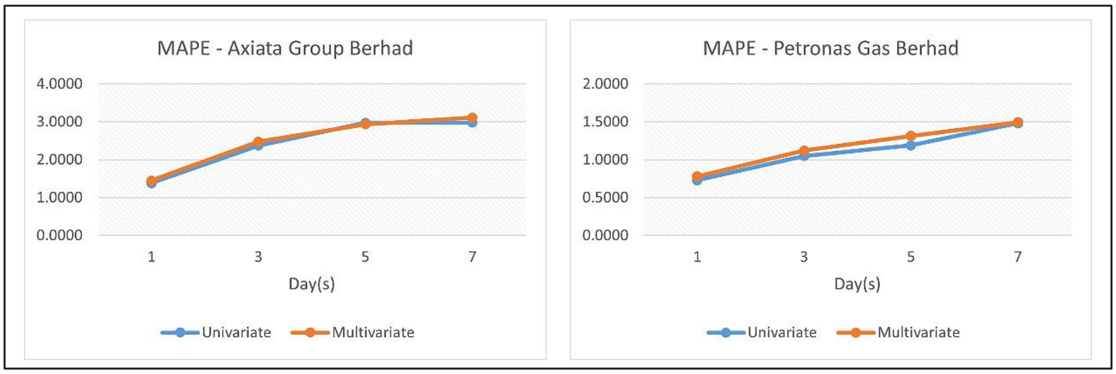

FIGURE 8. Comparison of best MAPE value between univariate and multivariate

TABLE 3. Statistical T-test on univariate and multivariate prediction

Prediction period	p Value	Hypothesis result
	Axiata Group Berhad	
1 Day	0.4147	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
3 Day	0.7114	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
5 Day	0.5937	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
7 Day	0.4871	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
	Petronas Gas Berhad	
1 Day	0.2559	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
3 Day	0.2928	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches
5 Day	0.0015	Reject $H_{\emptyset}$: There is significant difference between both approaches
7 Day	0.6114	Fail to reject $H_{\emptyset}$: There is no significant difference between both approaches

P-value from the test against the confidence intervals α = 0.05 will decide whether a significant difference exists between univariate and multivariate predictions. The p-value that is less than the α value shows a significant difference between univariate and multivariate prediction, while the opposite is when the p-value is more than the α value. Based on the test results, it can be concluded that there are no significant differences between the univariate and multivariate predictions for all prediction periods except for the 5th day of Petronas Gas Berhad stock. Hence, this test has proved and supported the finding that multivariate prediction is vital in stock prediction models, specifically for the Malaysian stock market. In conclusion, univariate has a slight advantage over multivariate prediction with better prediction accuracy. Multivariate models benefit from considering many features via information from the dataset with minor prediction errors.

The three best models from each prediction period are ranked for stock data to determine the model's consistency and best prediction performance. Table 4 shows the rank of the three best models for each prediction period and stock data. Therefore, the best model with the highest ranking can be determined.

TABLE 4. Ranking of three best model for each prediction period and stock data

	Axiata Group Berhad			Petronas Gas Berhad		
Number of Day(s)	1	2	3	1	2	3
			Univariate			
1	LSTM	CNN	MLP	LSTM	CNN	MLP
3	LSTM	MLP	CNN	LSTM	MLP	CNN
5	MLP	LSTM	CNN	MLP	LSTM	CNN
7	LSTM	CNN	RFR	LSTM	CNN	MLP
			Multivariate			
1	LSTM	CNN	MLP	LSTM	MLP	CNN
3	LSTM	MLP	CNN	LSTM	CNN	MLP
5	LSTM	MLP	CNN	LSTM	MLP	CNN
7	LSTM	MLP	CNN	LSTM	CNN	MLP

Based on ranking for both stock data using univariate and multivariate prediction, the deep learning model outperforms all the machine learning models for each prediction period. There is only one univariate prediction on Axiata Group Berhad, where RFR ranks as the third-best model on the 5th day. Overall, the best model for each period is still obtained by the deep learning model for both stock data. The analysis also shows that the LSTM model performs best for both stock data using multivariate prediction. LSTM was also ranked as the best model for almost all prediction periods using univariate prediction. Even though MLP obtained the best performance for the 5th-day prediction, LSTM remained competitive and ranked as the second-best model for that prediction period.

The overall prediction performance clearly shows that LSTM is the best model either for short or long-term prediction for the Malaysian stock market. It outperforms almost all the other models with the lowest error. The architecture and design of the LSTM algorithm that can store the historical information sequence help it predicts with a high accuracy rate. It also has the mechanism to remove irrelevant information from memory, which gives LSTM a significant advantage compared to the other models. These help the LSTM algorithm to learn the complexity and non-linearity of stock data efficiently and predict the future price with a high accuracy rate.

The architecture of LSTM, which contains the memory block or cell that can store past information for the long term, is the huge advantage that boosts the prediction performance of this model. Its capability to manage the information and to keep or remove irrelevant past information also has a leap the performance compared to the others deep learning models. In the context of stock market data, the relevant past information, such as stock price, will be stored in the memory and available for the algorithm to learn with the new input features. It allows the dependency and relationship of the past, and present information remains in the memory for the algorithm to learn and predict accurately. Meanwhile, MLP and CNN only learn the pattern of the stock price movement based on the input features fed to the algorithm. Without the memory function, the past information is ignored, which affects the dependency and relationship of the present information. Thus, the performance of the prediction made by LSTM is much better compared to MLP and CNN.

CONCLUSIONS

This paper presents the performance of the deep learning method to predict the Malaysian stock market. Univariate and multivariate approaches with different input features are used to make the prediction and evaluate both methods' effect on the prediction model's performance. Two stock market profiles are obtained from Bursa Malaysia, Axiata Group Berhad, and Petronas Gas Berhad as datasets for this study. Both data were obtained and downloaded from Yahoo Finance, where five years of historical data for both profiles from January 2017 to December 2021 is used. We employed three deep learning and traditional machine learning methods for comparative evaluation. Experimental results show univariate prediction has a slight advantage over multivariate with a higher accuracy rate. However, multivariate seems still competitive, with prediction results at par with univariate. It also carries other information and variables from the dataset along with the model compared to the univariate approach. Analysis of the prediction performance shows that deep learning models have advantages in terms of accuracy and low prediction error compared to machine learning models. Deep learning, specifically the LSTM, outperforms other machine learning models for all prediction periods by using univariate or multivariate predictions. The lowest error obtained indicates that LSTM is the best model with high accuracy for Malaysian stock market prediction.

This study can be further enhanced using the hybrid or ensemble approach, combining different algorithms to predict the Malaysian stock market. A combination of CNN and LSTM algorithms could further improve the prediction performance of the Malaysian stock market. AlexNet, ResNet, and other advanced deep learning algorithms could also be considered to predict the stock market price. Another improvement is having additional input features such as technical indicators, commodity data, and other financial or market data influencing the stock price, which could also be used to boost the prediction performance.

REFERENCES

Abu-Mostafa, Y.S. & Atiya, A.F. 1996. Introduction to financial forecasting. *Applied Intelligence* 6(3): 205-213.

Al-Mashhadani, A.F.S., Hishan, S.S., Awang, H. & Alezabi, K.A.A. 2021. Forecasting Malaysian Stock Price using Artificial Neural Networks (ANN). *Journal of Contemporary Issues in Business and Government* 27(1): 4466-4482.

Azam, M., Haseeb, M., Samsi, A.B. & Raji, J.O. 2016. Stock market development and economic growth: Evidences from Asia-4 countries. *International Journal of Economics and Financial Issues* 6(3): 1200-1208.

Bustos, O. & Pomares-Quimbaya, A. 2020. Stock market movement forecast: A systematic review. *Expert Systems with Applications* 156: 113464.

Cao, J., Li, Z. & Li, J. 2019. Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and Its Applications* 519: 127-139.

Chung, H. & Shin, K.S. 2020. Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. *Neural Computing and Applications* 32(12): 7897-7914.

Fister, D., Mun, J.C., Jagric, V. & Jagric, T. 2019. Deep learning for stock market trading: A superior trading strategy? *Neural Network World* 29(3): 151-171.

Gandhmal, D.P. & Kumar, K. 2019. Systematic analysis and review of stock market prediction techniques. *Computer Science Review* 34: 100190.

Hafizah Bahaludin & Saiful Hafizah Jaaman. 2013. Peta pasaran saham Malaysia. *Journal of Quality Measurement and Analysis* 9(2): 27-36.

He, K., Zhang, X., Ren, S. & Sun, J. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December.* pp. 770-778.

Hiransha, M., Gopalakrishnan, E.A., Menon, V.K. & Soman, K.P. 2018. NSE stock market prediction using deep-learning models. *Procedia Computer Science* 132(Iccids): 1351-1362.

Hoseinzade, E. & Haratizadeh, S. 2019. CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications* 129: 273-285.

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L. & Muller, P.A. 2019. Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery* 33(4): 917-963.

Jiang, W. 2021. Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications* 184(March 2020): 115537.

Krizhevsky, B.A., Sutskever, I. & Hinton, G.E. 2012. ImageNet classification with deep convolutional neural networks. *Communications of the ACM* 60(6): 84-90.

Lee, M.C., Chang, J.W., Hung, J.C. & Chen, B.L. 2021. Exploring the effectiveness of deep neural networks with technical analysis applied to stock market prediction. *Computer Science and Information Systems* 18(2): 401-418.

Liu, H. & Song, B. 2018. Stock price trend prediction model based on deep residual network and stock price graph. *Proceedings - 2018 11th International Symposium on Computational Intelligence and Design* 2: 328-331.

Liu, H. & Long, Z. 2020. An improved deep learning model for predicting stock market price time series. *Digital Signal Processing: A Review Journal* 102: 102741.

Livieris, I.E., Pintelas, E. & Pintelas, P. 2020. A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications* 32(23): 17351-17360.

Maiti, A. & Shetty, D.P. 2020. Indian stock market prediction using deep learning. *IEEE Region 10 Annual International Conference, Proceedings/TENCON 2020-Novem*: pp. 1215-1220.

Masoud, N.M.H. 2013. The impact of stock market performance upon economic growth. *International Journal of Economics and Financial Issues* 3(4): 788-798.

Mehtab, S. & Sen, J. 2020. *Stock Price Prediction Using Convolutional Neural Networks on a Multivariate Timeseries*. https://doi.org/10.36227/techrxiv.15088734.v1

Mehtab, S., Sen, J. & Dasgupta, S. 2020. Robust analysis of stock price time series using CNN and LSTM-based deep learning models. *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology.* pp. 1481-1486.

Mehtab, S., Sen, J. & Dutta, A. 2021. Stock price prediction using machine learning and LSTM-based deep learning models. *Communications in Computer and Information Science* 1366: 88-106.

Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., Salwana, E. & Shahab, S. 2020. Deep learning for stock market prediction. *Entropy* 22(8): 840.

Nikou, M., Mansourfar, G. & Bagherzadeh, J. 2019. Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management* 26(4): 164-174.

Pak, M. & Kim, S. 2018. A review of deep learning in image recognition. *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology, CAIPT 2017 2018-Janua*. pp. 1-3.

Pan, W., Li, J. & Li, X. 2020. Portfolio learning based on deep learning. *Future Internet* 12(11): 1-13.

Pradhan, R.P. 2018. Development of stock market and economic growth: The G-20 evidence. *Eurasian Economic Review* 8(2): 161-181.

Rama Krishna, V., Subhamastan Rao, T., Narayana, G.V.S. & Rachapudi, V. 2020. A model for stock price predictions using deep learning techniques. *International Journal of Advanced Trends in Computer Science and Engineering* 9(5): 8266-8271.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E.A., Menon, V.K. & Soman, K.P. 2017. Stock price prediction using LSTM, RNN, and CNN-sliding window model. *2017 International Conference on Advances in Computing, Communications, and Informatics, ICACCI 2017 2017-Janua*. pp. 1643-1647.

Sezer, O.B., Gudelek, M.U. & Ozbayoglu, A.M. 2020. Financial time series forecasting with deep learning: A systematic literature review: 2005-2019. *Applied Soft Computing Journal* 90: 106181.

Shen, J. & Shafiq, M.O. 2020. Short-term stock market price trend prediction using a comprehensive deep learning system. *Journal of Big Data* 7: Article No. 66.

Sismanoglu, G., Onde, M.A., Kocer, F. & Sahingoz, O.K. 2019. Deep learning based forecasting in stock market with big data analytics. *2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science.* pp. 1-4.

Soon, G.K., On, C.K., Rayner, A., Patricia, A. & Teo, J. 2018. A CIMB stock price prediction case study with feedforward neural network and recurrent neural network. *Journal of Telecommunication, Electronic and Computer Engineering* 10(3-2): 89-94.

Yiing, A.T.S. & Thim, C.K. 2015. Prediction of Bursa Malaysia Stock Index using autoregressive integrated moving average and artificial neural network. *Malaysia Statistics Conference* (MyStats 2015) 1997: 95.

Yong, B.X., Abdul Rahim, M.R. & Abdullah, A.S. 2017. A stock market trading system using deep neural network. *Communications in Computer and Information Science* 751: 356-364.

Zeroual, A., Harrou, F., Dairi, A. & Sun, Y. 2020. Deep learning methods for forecasting COVID-19 time-Series data: A comparative study. *Chaos, Solitons and Fractals* 140: 110121.

Zhao, Y. & Khushi, M. 2020. Wavelet Denoised-ResNet CNN and LightGBM method to predict forex rate of change. *IEEE International Conference on Data Mining Workshops, ICDMW 2020-November*. pp. 385-391.

Zulqarnain, M., Ghazali, R., Ghouse, M.G., Hassim, Y.M.M. & Javid, I. 2020. Predicting financial prices of stock market using recurrent convolutional neural networks. *International Journal of Intelligent Systems and Applications* 12(6): 21-32.

*Corresponding author; email: azuraliza@ukm.edu.my