Time Series Analysis of Stock Prices based on Deep Learning

Sukanchalika Boontham
Faculty of Information Technology
King Mongkut's University of Technology North Bangkok
Bangkok, Thailand
sukanchalika.b@email.kmutnb.ac.th

Phayung Meesad
Faculty of Information Technology
King Mongkut's University of Technology North Bangkok
Bangkok, Thailand
phayung.m@it.kmutnb.ac.th

***Abstract*— Stock price time series prediction is a challenging task. In this paper, we report on deep learning usage of forecasting time series analysis of 8 stock price series data from the Stock Exchange of Thailand i.e. PTT, TRUE, GRAMMY, LPN, KBANK, CPF, IVL, MODERN, dating from January 1st, 2013 to December 31st, 2017 with the window series during the past 5 days were used. The deep learning models used in this study were Long Short Time Memory (LSTM) and Gated Recurrent Unit (GRU). RMSE, MAE, and R^2 were used to measure the performance of the resulted models. The results show that LSTM prediction performance is compatible with GRU. However, the running time LSTM takes longer to finish the training process. MODERN stock is the best prediction by LSTM (0.977, 0.046, and 0.036 for R^2, RMSE, and MAE, respectively) with the processing time equals 12.945 sec. For the GRU model performs well MODERN stock prices with the best forecasts (0.978, 0.046, 0.035, for R^2, RMSE, and MAE, respectively) with the processing time equal 10.264 sec.**

Keywords—Time Series, Deep Learning, Prediction Model

I. Introduction

Data stored in a continuous period are called time series data, which may be in the form of daily, annual, quarterly, or monthly data depending on the suitability and use. For example, the SET Index and the Industry Group Index are listed in the Stock Exchange of Thailand (SET). The stock index is used to display high market capitalization[1]. A stock price index is a tool that indicates the price level and outlook of the stock market and the data are stored as a time series. Data in this way will be large amount with behaviors and specific features. The stock price index also has several factors that affect the price of stocks, reflecting the impact on the company, such as economic news, political news and financial news. Stock Patterns of volatile price volatility are constantly difficult to predict future prices.

In this research, we apply long short-term memory and gated recurrent unit deep learning methods for forecasting time series stock prices from SET. The data sets used in the experiment are the Industry Group Index, which are stock price series data from the ranged of January 1st, 2013 to October 31st, 2018 to predicts the next day closing price. A predictive performance of the deep learning predictive modeling is compared.

II. Literature Review

A. Time Series Data

Time series data is a collection of data that is stored periodically over a period. It is usually represented by a line chart. Time series data may be in the form of annual, quarterly, or monthly data.

B. Deep Learning

Deep learning is one of the most popular machine learning algorithms [3]. Deep learning considers a subset of an Artificial Neuron Networks (ANN) model is based on the concepts and techniques of human brain simulations that allow the computer to recognize and memorize information like neurons of humans. ANN learning process will be adjusted parameters for multitasking epoch to the value of the error is reduced. A traditional multilayer neural network can be divided into three layers: input layer, hidden layer, output layer. However, deep learning can be many layers depending on different architectures, such as deep multilayer perceptron, convolutional neural networks, and recurrent neural network. Moreover, deep learning process is different from ANN process which the process increase the number of hidden layer or node in the layer and process more complex but the processing resources will be increased [4].

Deep learning has been applied to several types of tasks, such as real-time crime forecasting using deep learning [5]. Recurrent Neural Network (RNN) applies to sequential data. There are two major parts, input data and the hidden state, which collect the results from the previous node. The two parts are combined and the results are input into the next node. This is often a problem with a sequence data that is too long [6]. For RNN's algorithm include: Long Short-Term Memory (LSTM) is a solution to RNN problems with long sequence time series data [7]. It works by having a cell state that holds the state of each node so that it can reverse the values obtained from the previous state and have the gate as the flow control. The Gated Recurrent Unit (GRU) has the same overall functionality as the LSTM algorithm. Each node has a process by combining the forget gate and the input gate into the update gate. Also, it includes the cell state and the hidden state together. The performance is equivalent to the LSTM algorithm. However, GRU has simpler architecture and fewer parameters to generate sequential data [8].

III. Research Methodology

A. System Requirements Study

This research objective is to apply deep learning models for a short-term stock price prediction. Several deep learning models are compared base on the performance prediction of stock price time series. The data used in this research are the time series of the Industry Group Index of SET. LSTM and GRU deep learning methods are used to build predictive regression models for forecasting stock price time series. To evaluate model performance in time series forecasting Root

Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R Squared (R^2) are used as performance metrices.

B. Data Preparation

The stocks index of 8 industrial groups were selected by a random method from each industry group: PTT, TRUE, GRAMMY, LPN, KBANK, CPF, IVL, and MODERN. The data sets were taken from the ranged of January 1st, 2013 to October 31st, 2018. A sliding window technique, which takes the original time series data and converts the time index into an attribute in a cross-sectional format, was used in the preparing process of time series data required before feeding time series data into the forecasting process.

C. Modeling Machine Learning

Modeling machine learning is the process to build and evaluation the prediction models. Firstly, training and testing data sets using the Industry Groups Index 8 datasets. Then, splitting data into training set and test set. Time series data from January 1st, 2013 to December 31st, 2017 (1,220 days) were used as training set and from January 1st, 2018 to October 31st, 2018 (205 days) were used as test set. Stock closing price was used as the target for forecasting on the next day, while the previous days of Open, High, Low, and Close prices were used as inputs.

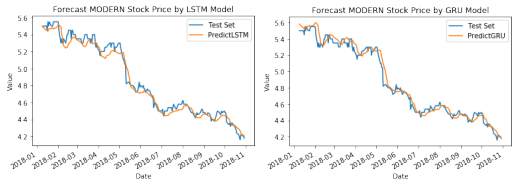
For prediction modeling based on deep learning, LSTM and GRU were used. In the experiment, we set the parameters for the deep learning methods as follows: Input node = '1215', output node = '1', hidden node = '500', epoch = '100', activation = 'linear'.

In the performance evaluation, test data were used to estimate the generalization of the predictive models. Prediction efficiency of models was measured in terms of RMSE, MAE, and R^2 to analyze the relevance of historical data to short-term stock price prediction of time series data.

IV. Experimental Results

SET Industry Group Index stock prices were 8 stocks: PTT, TRUE, GRAMMY, LPN, KBANK, CPF, IVL and MODERN chosen randomly from each industry group. Deep learning predictive models based on LSTM and GRU were built. The performance evaluation for the forecasting models measured by RMSE, MAE and R^2. The results show that the LSTM forecasting models: MODERN, GRAMMY and KBANK, show the top 3 based on R^2, RMSE, and MAE. In addition, the best prediction was on the LSTM on MODERN (0.977, 0.046, 0.036 for R^2, RMSE, and MAE, respectively). Processing time of LSTM was 12.945 sec. For GRU forecasting models, it shows that MODERN, GRAMMY and KBANK were the top 3 performance based on R^2, RMSE, and MAE. The best prediction efficiency was also on MODERN (0.978, 0.046, 0.035, for R^2, RMSE, and MAE, respectively). Processing time of GRU was 10.264 sec. Prediction based the test set is a line graph as shown in Fig.1.

From TABLE I, the results show that deep learning models based on both LSTM and GRU for forecasting stock price time series data are consistent. Considering the direction of trend movement of both the test and predicted data, it is found that the predicted values go along with the real test values. In the prediction performance comparison between LSTM and GRU, the LSTM model can predict with a lower error than the GRU model. However, LSTM takes longer time to process. Based on all 8 stock companies, the GRU predictive model shows a similar performance to the LSTM model with a slightly more tolerance to price volatility.



(a) LSTM (b) GRU

Fig. 1. MODERN stock forecast models with (a) LSTM and (b) GRU

TABLE I. Performance Evaluation Results

Model	Stock	RMSE	MAE	R2	loss	Validate loss	Time (sec.)
LSTM	***PTT***	0.135	0.105	0.697	0.013	0.002	14.808
	TRUE	0.079	0.062	0.867	0.038	0.003	13.546
	GRAMMY	0.059	0.045	0.957	0.107	0.021	13.523
	LPN	0.060	0.044	0.922	0.046	0.011	13.583
	KBANK	0.067	0.049	0.939	0.006	0.003	13.733
	CPF	0.088	0.064	0.842	0.027	0.009	14.032
	IVL	0.161	0.136	0.479	0.035	0.006	13.643
	MODERN	**0.046**	**0.036**	**0.977**	**0.072**	**0.008**	**12.945**
GRU	***PTT***	0.104	0.076	0.819	0.018	0.003	10.300
	TRUE	0.085	0.068	0.847	0.035	0.003	10.396
	GRAMMY	0.060	0.050	0.956	0.084	0.021	10.047
	LPN	0.061	0.043	0.919	0.054	0.010	10.516
	KBANK	0.072	0.055	0.930	0.008	0.004	10.306
	CPF	0.098	0.075	0.802	0.029	0.014	10.412
	IVL	0.132	0.107	0.650	0.043	0.009	10.458
	MODERN	**0.046**	**0.035**	**0.978**	**0.115**	**0.008**	**10.264**

References

[1] "SET - Market Data - Historical Chart." [Online]. Available: https://www.set.or.th//en/market/setindexchart.html. [Accessed: 04-May-2017].

[2] R. Vázquez, A. B. del Río, and V. Iglesias, "Consumer-based Brand Equity: Development and Validation of a Measurement Instrument," *J. Mark. Manag.*, vol. 18, no. 1–2, pp. 27–48, Feb. 2002.

[3] Y. LeCun, Y. Bengio, and G. Hinton, *Deep Learning*, vol. 521. 2015.

[4] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[5] B. Wang, D. Zhang, D. Zhang, P. Jeffery Brantingham, and A. Bertozzi, "Deep Learning for Real Time Crime Forecasting," Jul. 2017.

[6] S. Tangruamsub, "Long Short-Term Memory (LSTM) – Sirinart Tangruamsub," *Medium*, 04-Jul-2017. [Online]. Available: https://medium.com/@sinart.t/long-short-term-memory-lstm-e6cb23b494c6. [Accessed: 14-Nov-2018].

[7] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

[8] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," *ArXiv14123555 Cs*, Dec. 2014.