Parallelization of LSTM-GRU Architectures for Multivariate Prediction of Stock Prices

Gabriel Asael Tarigan
Computer Science Department
BINUS Graduate Program, Master of Computer Science
Bina Nusantara University
Jakarta, Indonesia, 11480
gabriel.tarigan@binus.ac.id

Eric Savero Hermawan
Computer Science Department
BINUS Graduate Program, Master of Computer Science
Bina Nusantara University
Jakarta, Indonesia, 11480
eric.savero@binus.ac.id

Abba Suganda Girsang
Computer Science Department
BINUS Graduate Program, Master of Computer Science
Bina Nusantara University
Jakarta, Indonesia, 11480
agirsang@binus.edu

***Abstract*—Anticipating stock market movements is challenging, as hasty prediction risks large financial losses. Stock price prediction is complex due to the high volatility and many influencing external factors. Recurrent Neural Network (RNN) models are made to learn time series data making it ideal for stock prediction based on records. Well-known RNN architectures for forecasting problems are Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU). However, past studies only explored standalone models and have not generated convincing performance results. By integrating the efficiency of GRU to effectively handle shorter data sequences with LSTM's ability to capture long-term dependencies, our model aims to improve stock price forecasting accuracy by collaborating the strength between both architectures. The proposed model was used to forecast multivariate stock prices and evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). We found that parallelization of LSTM-GRU offers better prediction, up to 9.1% improvement compared to sequential LSTM-GRU and Bidirectional LSTM (Bi-LSTM).**

***Index Terms*—LSTM, GRU, stock prices, short-term prediction**

I. Introduction

Stock price prediction advances have earned significant relevance among experienced analysts and investors, as correct predictions lead to attractive profits. However, making proper stock price predictions poses a formidable challenge, given the intricate interplay of high volatility and myriad influencing variables, spanning politics, economics, and social dynamics [1]–[3]. Since such data are spread out and semantically diverse, past price fluctuation patterns can be seen as the summarized outcome of all the variables for prediction. Predicting the closing price offers information related to the stock price fluctuation to assist the decision-making process of buying and selling stocks [4].

Deep learning, specifically recurrent neural networks (RNN), is increasingly used in stock market predictions. RNN has the adaptability to sequential data, Long Short-Term Memory (LSTM) [5] and Gated Recurrent Unit (GRU) [6] designs are among the most extensively utilized forms of RNNs [7]. Multiple studies have proposed approaches using standalone (consisting only of one model) LSTM and GRU with their respective advantages and disadvantages, however, the result of using such standalone models in a time series with volatile patterns of fluctuation has achieved at best results deemed only satisfactory [8]. Thus, this research focuses on the Parallelized LSTM-GRU (PLG) model for stock market prediction. To measure the PLG model stock prediction performance, the model is trained and tested on varied stock price data as a generalized performance measure when learning multiple stock data. The PLG model is also compared with the standalone and sequential variants of the commonly used RNN models: LSTM and GRU, and the Bidirectional variant of RNN [9] using Bidirectional LSTM (Bi-LSTM) which is popular for time series tasks such as stock price prediction [10]. Evaluation of the PLG model improvement against the comparator models uses Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These metrics provide a comprehensive assessment of the magnitude of the difference between each model's prediction and the actual data [11], [12].

II. Background and Literature Review

Prior studies have shown how machine learning (ML) techniques have numerous advantages over statistical forecasting models, for the ability to give more precise predictions and the adoption of RNN has led to practical use to assist in financial studies for analyzing the complex changes in the stock market. The growing robustness of deep learning and its accessibility has led to an increase in usage by investors to study stock price patterns and make calculated decisions [13], [14].

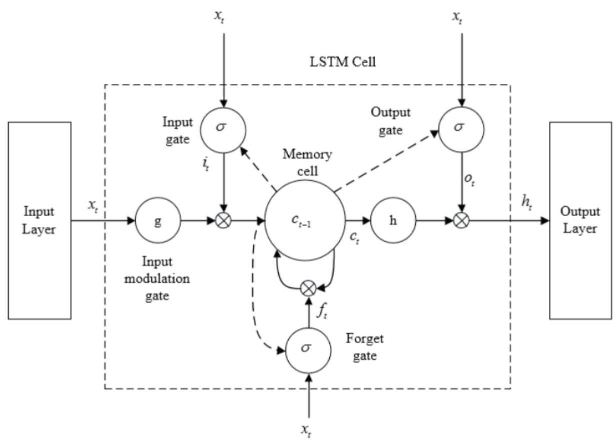

Fig. 1. LSTM architecture.

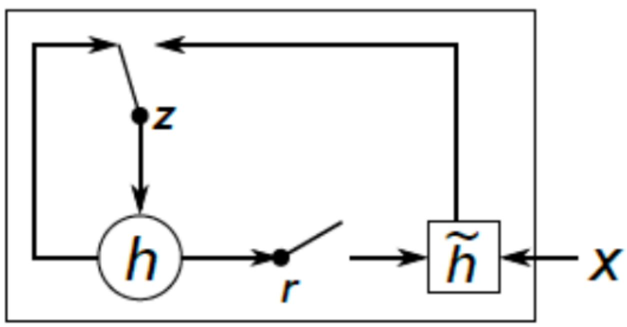

Fig. 2. GRU architecture.

A. Long Short-Term Memory (LSTM)

LSTM is an RNN architecture notable for employing the cell state as a continuous data sequence to counter vanishing gradient in standard RNN and three gates for filtering information from time series data [5]. The forget gate selects which information is forgotten or carried through the cell state from the past state, while the input and output gates select the information to be added and passed respectively. The architecture is summarized in Figure 1.

B. Gated Recurrent Units (GRU)

GRU is also an RNN refined to address the vanishing gradient problem. It is lighter than LSTM by having fewer tensors and two gates for learning [6]. GRU only transfers information through the hidden state, the update gate filters information from the previous hidden state and new input and the reset gate calculates previous state information to be omitted. The architecture is shown in Figure 2.

C. Model Evaluation Metrics

The models in this paper are evaluated using root mean squared error (RMSE) and mean absolute error (MAE) denoted by the equations below. RMSE and MAE calculate the average difference magnitude and the absolute difference between prediction ($\hat{y}_i$) and ground truth (y_i), the lower the score (closer to 0) means the prediction output is closer to the actual value.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

D. Related Studies

More volatile fluctuations seen in the cryptocurrency market were proven to be learnable by the LSTM-GRU parallelization as shown in multiple studies [14]–[16]. Stock price prediction studies have used the LSTM and GRU models for time series data in market regions such as the Moroccan market [7]. The most observed input variable was the closing price to forecast the future price as the stock market value. Their proposed model showed satisfactory performance evaluated to improve annual returns by 27%.

Regularized LSTM-GRU was proposed in [17] and it was found to perform better for a short-term estimate of the stock closing price. It was able to converge quickly and offered high forecast accuracy. For forecasting cryptocurrency fluctuations [15] LSTM was found to be the state-of-the-art model for time series data. The model is desired for its recall capacity and ability to extract temporal properties from data. The study also explored the hybrid GRU-LSTM approach, which proved superior to the standalone LSTM based on the prediction error rate. However, [16] found that the ensemble approach using the hybrid approach from [15] did not outperform the LSTM network, which was present in all the top-performing ensembles. The LSTM-GRU ensemble still outperformed the hybrid non-sequential LSTM-GRU model. A bidirectional variation of LSTM, Bi-LSTM [18] was shown to perform better than regular LSTM. Its additional training in left-right and right-left steps improved the learning which also resulted in longer training time to reach equilibrium.

The overall comparison of LSTM and GRU in [19] found the GRU model has a higher return ratio, with a more robust approach for forecasting trends, but it predicts with less confidence than LSTM. LSTM, however, produces excellent outcomes, with the downside of the inferior simulation of dramatic market falls or rapid price spikes.

III. Methodology

The architectures analyzed in this project consist of LSTM and GRU with different arrangements; LSTM-GRU sequential, Bi-LSTM, and LSTM-GRU Parallel. Each model is trained and tested using three popular stock price data collected from Yahoo Finance Dataset, and the testing performance is analyzed based on prediction accuracy.

TABLE I
REPRESENTATION OF STOCKS DATASET

close	date	scaled_close	future
75.247…	2020-08-06	0.0602…	0.0448…
74.918…	2020-08-07	0.0560…	0.0616…
74.841…	2020-08-08	0.0550…	0.0675…
74.027	2020-08-09	0.0448…	0.0599…

A. Dataset

In total, three stock price data are used for the model training and testing to see if the PLG model's performance can be applied and is not only limited to one stock price. The datasets used are Google, Amazon, and Tesla, as their stock prices are popular and have many fluctuations. All the stock prices were collected from the Yahoo Finance Dataset, the training data used starts from the 6th of August 2020 until the most recent data for the collection process is on the 27th of November 2023. The training set starts from the 6th of August 2020 until the month of April 2023, and the test data uses data from the month of May 2023 until the 27th of November 2023. The input data is the daily close price, and the model outputs a prediction of the close price for three consecutive days. Min-max scaler is used as a normalization technique to maintain the large variation of prediction and forecasting, the normalization technique is required to reduce the range of data into a common scale [20] to reduce the impact of outliers. The input and the prediction are three days apart. As shown in Table I, the scaled_close column is the min-max scaled price, and the future column contains the price after three days from the scaled_close value offset by 3.

B. Model Architecture

LSTM and GRU are well-known RNN architectures capable of passing information from past stages and using it for future predictions [21], as such, the usage of both models (mainly LSTM) is essential in stock prediction due to the nature of the task using time-series data. LSTM and GRU are both widely used in a hybrid manner to predict stock prices [7], [14], [17], [19], [22].

1) Parallelized LSTM-GRU (PLG) Model: The parallel architecture of PLG combines both LSTM and GRU with the expectation of the model having a more nuanced result to the prediction by utilizing the complexity of LSTM and the more streamlined GRU model. Both of their outputs are concatenated to produce better prediction results, the model can be seen in Figure 3.

2) LSTM-GRU Sequential Model: The LSTM-GRU Sequential Model is used to gauge its performance when carrying time series information between the different GRU and LSTM architectures. The LSTM comes in front of the more complex model to retain more information, followed by GRU to streamline the information output, the expected result is the more information-packed output of LSTM can be generalized more when being used as the input of GRU, as shown in Figure 4.

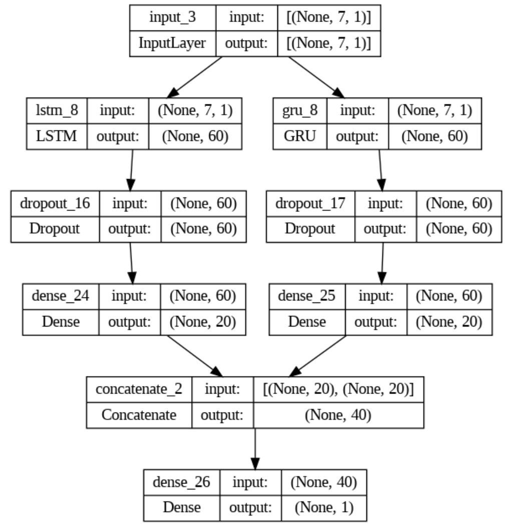

Fig. 3. PLG Model.

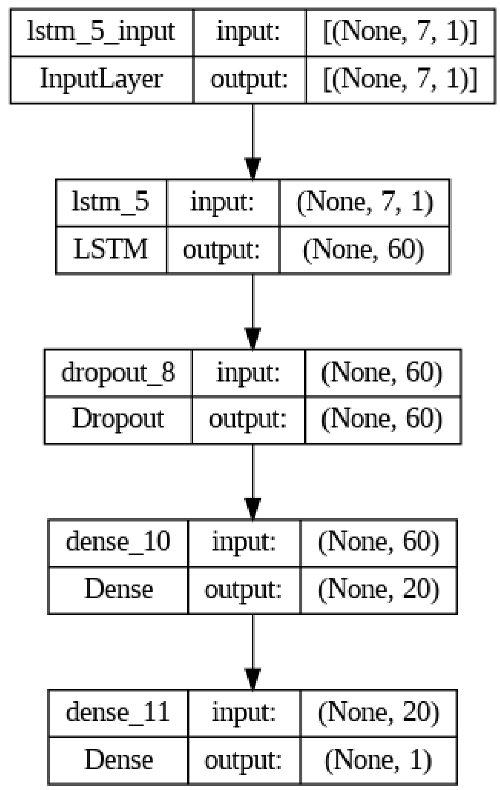

Fig. 4. Sequential Model.

3) Bi-LSTM Model: The Bi-LSTM model is a widely used time series forecasting model [14]. The Bi-LSTM is employed for time series and stock prediction due to its capacity to capture long-term dependencies and intricate patterns within sequential data. Its bidirectional nature allows it to consider information from both past and future time steps, enabling a more comprehensive understanding of temporal relationships and enhancing the model's ability to discern subtle trends

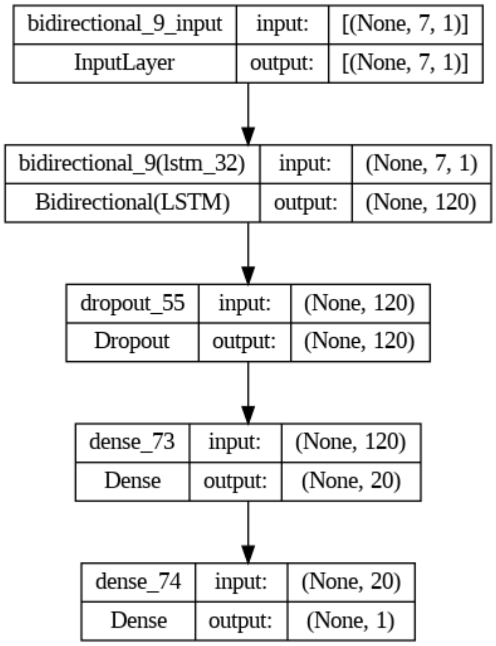

Fig. 5. Bi-LSTM Model.

TABLE II
HYPERPARAMETER SETTING

Parameter	Value
Epoch	100
Batch size	32
Loss Function	RMSE
Optimizer	Adam

in the time series data. This makes Bi-LSTM particularly effective for tasks where the prediction relies on a nuanced analysis of historical patterns and dependencies over extended periods. The model architecture used for this research is shown in Figure 5.

C. Training Hyperparameters

To minimize the variables and chances of inconsistency in the results, all the models used in this paper are trained using the same hyperparameters as elaborated in Table II.

IV. RESULTS

A. Models Evaluation

The models are evaluated using RMSE and MAE to calculate the average difference magnitude and the absolute difference between prediction and ground truth stock prices. Training and test set loss is also shown, as incorporating those metrics also provides a quantitative measure of the model's performance and generalization ability, offering readers valuable insights into its effectiveness on unseen data and overall optimization dynamics. The results are shown in Table III.

The loss results of the models are provided in Table III, where the RMSE and MAE results are compiled. The PLG model consistently achieved superior RMSE and MAE, while

TABLE III
MODEL RESULTS

Stocks	Evaluation Metrics	PLG	Sequential LSTM-GRU	Bi-LSTM
Google	Training Loss	0.0395	0.0488	0.0411
	Test Loss	0.0619	0.0686	0.0657
	RMSE	0.0625	0.0691	0.0666
	MAE	0.0484	0.0536	0.0513
Amazon	Training Loss	0.0351	0.0416	0.0391
	Test Loss	0.0557	0.0597	0.0583
	RMSE	0.0564	0.0601	0.0590
	MAE	0.0428	0.0469	0.0465
Tesla	Training Loss	0.0364	0.0354	0.0415
	Test Loss	0.0560	0.0615	0.0581
	RMSE	0.0567	0.0626	0.0587
	MAE	0.0429	0.0466	0.0448

TABLE IV
MULTIPLE MODEL PREDICTION GRAPH LEGEND

Actual stock price
Predicted price of PLG
Predicted price of Sequential LSTM-GRU
Predicted price of Bi-LSTM

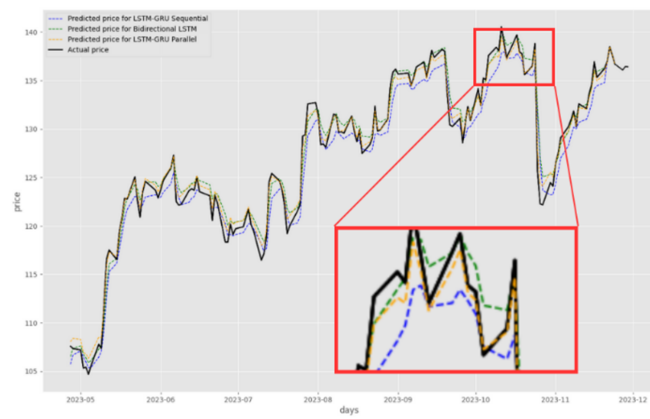

Fig. 6. Google (GOOGL) stock price prediction from all the models.

both Bi-LSTM and sequential models performed likewise with slightly less accuracy. From the results gathered in Table III, the PLG model shows the best RMSE of 0.0564 on the Tesla (TSLA) data, and it is closer to 0 when compared to other comparators, which can be interpreted as the model having the best accuracy.

B. Prediction Graph

The result graph shows the prediction result starting from May 2023 and ending in November 2023. The legend for the multiple-model prediction graphs is given in Table IV, and the prediction plots are shown in Figures 6, 7, and 8.

The stock prediction was run on the 27th of November 2023. Therefore, the models predict the stock price of 28th – 30th of November 2023. With the visualizations in Figure 9. The predictions are then compiled in Table V to make it easier to interpret, which shows what model predicts the results of

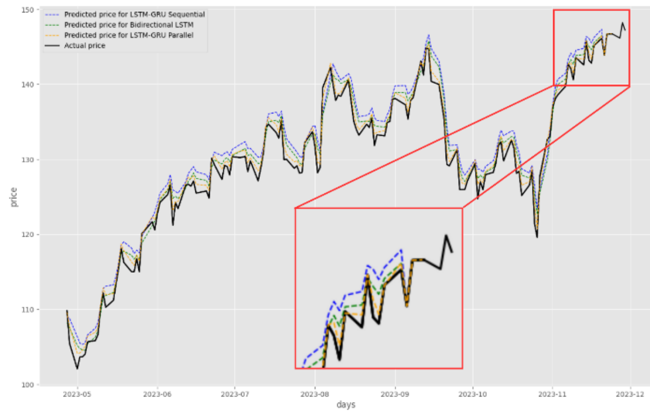

Fig. 7. Amazon (AMZN) stock price prediction from all the models.

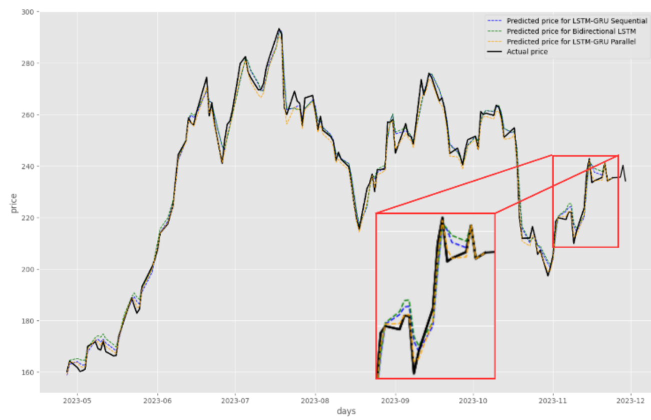

Fig. 8. Tesla (TSLA) stock price prediction from all the models.

TABLE V
PRICE PREDICTIONS ON NOVEMBER 28–30

Stocks	Model	Daily price (In dollars)		
		Day 1	Day 2	Day 3
Google	PLG	136.08	136.44	136.42
	Sequential LSTM-GRU	137.13	137.25	136.08
	Bi-LSTM	136.83	136.51	137.62
Amazon	PLG	146.17	148.23	147.23
	Sequential LSTM-GRU	146.2	146.76	148.48
	Bi-LSTM	146.22	147.36	147.64
Tesla	PLG	235.71	240.26	234.26
	Sequential LSTM-GRU	234.48	240.41	236.41
	Bi-LSTM	232.82	237.27	236.7

the 28th–30th of November as Day 1–3. Figure 9 shows the PLG performance predicting each stock data. Notice that only the PLG model predictions are observed here as this paper is focused on the PLG model performance.

V. CONCLUSION

Stock price prediction has been a very challenging task for researchers due to the external social and psychological factors that affect its price outcome. This research found that the Parallel LSTM-GRU model was able to learn the fluctuation patterns from multiple stock prices, the performance was validated by the low error rate. As a result, it can be concluded that the PLG model has the highest confidence in predicting stock prices closer to real life compared to the comparator models. Our findings were in line with [15] where their proposed hybrid LSTM-GRU was shown to be superior for predicting the more volatile cryptocurrency market. The LSTM-GRU Sequential model performed with the least accuracy even compared to Bi-LSTM. The reason for this can be due to the output from LSTM used as the GRU input, in which the LSTM's more complex output features got lost when processed by GRU which reduced the complexity.

REFERENCES

[1] D. P. Gandhmal and K. Kumar, "Systematic analysis and review of stock market prediction techniques," *Comput Sci Rev*, vol. 34, p. 100190, 2019, doi: 10.1016/j.cosrev.2019.08.001.

[2] N. Petrusheva and I. Jordanoski, "Comparative analysis between the fundamental and technical analysis of stocks," *Journal of Process Management. New Technologies*, vol. 4, no. 2, pp. 26–31, 2016, doi: 10.5937/jpmnt1602026p.

[3] A. Sharaff and M. Choudhary, "Comparative Analysis of Various Stock Prediction Techniques," *Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018*, no. Icoei, pp. 735–738, 2018, doi: 10.1109/ICOEI.2018.8553825.

[4] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques," *Procedia Comput Sci*, vol. 167, no. 2019, pp. 599–606, 2020, doi: 10.1016/j.procs.2020.03.326.

[5] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[6] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," pp. 1–9, 2014, [Online]. Available: http://arxiv.org/abs/1412.3555

[7] Y. Touzani and K. Douzi, "An LSTM and GRU based trading strategy adapted to the Moroccan market," *J Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00512-z.

[8] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," *Proceedings - 2016 31st Youth Academic Annual Conference of Chinese Association of Automation, YAC 2016*, no. November 2016, pp. 324–328, 2017, doi: 10.1109/YAC.2016.7804912.

[9] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, Apr. 1997, doi: 10.1109/78.650093.

[10] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," in *2019 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2019, pp. 3285–3292. doi: 10.1109/BigData47090.2019.9005997.

[11] A. Jadon, A. Patil, and S. Jadon, "A Comprehensive Survey of Regression Based Loss Functions for Time Series Forecasting," 2022, [Online]. Available: http://arxiv.org/abs/2211.02989

[12] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 4, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[13] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Applied System Innovation*, vol. 4, no. 1, pp. 1–30, 2021, doi: 10.3390/ASI4010009.

[14] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," *Fractal and Fractional*, vol. 7, no. 2, p. 203, Feb. 2023, doi: 10.3390/fractalfract7020203.

[15] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, "A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions," *Journal of Information Security and Applications*, vol. 55, no. June, p. 102583, 2020, doi: 10.1016/j.jisa.2020.102583.

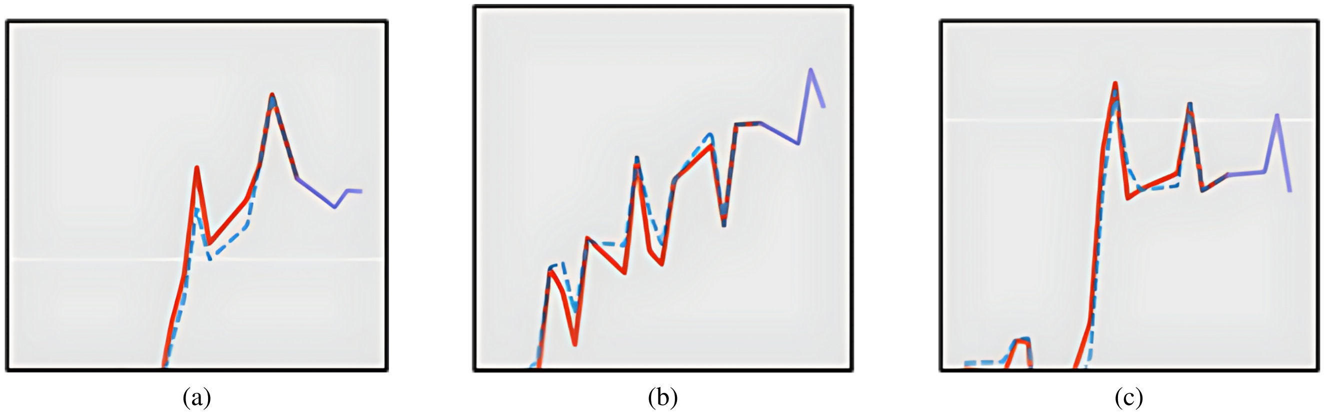

Fig. 9. PLG prediction of (a) Google, (b) Amazon, and (c) Tesla stock prices on November 28–30.

[16] K. Murray, A. Rossi, D. Carraro, and A. Visentin, "On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles," *Forecasting*, vol. 5, no. 1, pp. 196–209, 2023, doi: 10.3390/forecast5010010.

[17] Y. Liu, Z. Wang, and B. Zheng, "Application of Regularized GRU-LSTM Model in Stock Price Prediction," *2019 IEEE 5th International Conference on Computer and Communications, ICCC 2019*, pp. 1886–1890, 2019, doi: 10.1109/ICCC47050.2019.9064035.

[18] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," in *2019 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2019, pp. 3285–3292. doi: 10.1109/BigData47090.2019.9005997.

[19] A. Sethia and P. Raut, *Application of LSTM, GRU and ICA for stock price prediction*, vol. 107. Springer Singapore, 2019. doi: 10.1007/978-981-13-1747-7_46.

[20] S. G. K. Patro and K. K. sahu, "Normalization: A Preprocessing Stage," *Iarjset*, pp. 20–22, 2015, doi: 10.17148/iarjset.2015.2305.

[21] A. Moghar and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," *Procedia Comput Sci*, vol. 170, pp. 1168–1173, 2020, doi: 10.1016/j.procs.2020.03.049.

[22] H. Song and H. Choi, "Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models," *Applied Sciences (Switzerland)*, vol. 13, no. 7, 2023, doi: 10.3390/app13074644.