2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE) | 979-8-3503-2095-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCoSITE57641.2023.10127847

Energy Sector Stock Price Prediction Using The CNN, GRU & LSTM Hybrid Algorithm

1st Bambang Sulistio
Computer Science Department, Binus Graduate Program – Doctor of Computer Science
Bina Nusantara University
Jakarta, Indonesia
bambang.sulistio@binus.ac.id

2nd Harco Leslie Hendric Spits Warnars
Computer Science Department, School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
shendric@binus.edu

3rd Ford Lumban Gaol
Computer Science Department, School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
fgaol@binus.edu

4th Benfano Soewito
Computer Science Department, School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
bsoewito@binus.edu

***Abstract*— Nowadays, many people are starting to care about early investment. One of the most popular investments lately, especially for millennials, is a stock investment. In investing, there are advantages and risks of loss. One way to reduce the risk of loss is by using price predictions before investing in stocks. This paper proposes the use of deep learning in making stock predictions. We conducted research by calculating the performance of six deep-learning algorithms to predict stock closing prices. The application of the CNN-LSTM-GRU hybrid algorithm combination produces the best performance compared to other methods, based on the value: Root Mean Squared Error (RMSE) decreased by 1.100 by 14%, Mean Absolute Error (MAE) was successfully reduced by 0.798 by 13.4%, and R Square increased by 0.957 by 3.9%. In predicting stock prices on the Indonesian Stock Exchange, especially in the energy sector, CNN-LSTM-GRU is more appropriate for investors than using a single algorithm to make decisions in investing in stocks..**

Keywords*— *Stock prediction, Convolutional Neural Network, Long Short Term Memory, Gated Recurrent Neural Networks, Deep Learning

I. INTRODUCTION

One of the ways companies collect funds is through the stock market, which is a place where company shares can be circulated, traded, and also transferred[1]. Through the issuance of shares, significant capital and public funds will flow into the stock market, increasing the company's capital and encouraging economic commodities' development. This makes the capital market one of a country's financial and economic indicators [2]. In addition, transactions in the capital market are an opportunity for large and small investors or small speculators to increase the amount of capital they have [3]. There are advantages, such as dividends and capital gains, but there are also risks of losses, such as capital loss and liquidation risk. This risk causes depression and stress caused by: Profit Targets, Social Pressure, Workload, and Decision Making Process[4].

To reduce the risk of mistakes in investing in stocks, we need a model that can analyze stocks using price and trading volume so that it can be used to predict stock prices[5]. Many studies have been conducted to find the most accurate method of stock analysis that can minimize the occurrence of investment errors that result in losses. One of them is the use of machine learning algorithms for processing stock data, such as Convolutional Neural Network (CNN) [6], Long Short Term Memory (LSTM) [7] and Gated Recurrent Unit (GRU) [8]. In this research, we aim to combine the advantages of each machine learning algorithm, namely: CNN, LSTM, and GRU, into one hybrid method, using closing prices and stock volume to predict energy sector stock prices on the Indonesian stock exchange. The energy sector is a major factor that influences the performance of the Indonesian stock market [9]. So as our research data, we use stock transaction data from three issuers representing the energy sector: PT. Bukit Asam Tbk (PTBA), PT. Elnusa Tbk (ELSA) and PT. State Gas Company Tbk(PGAS). The data used is daily data consisting of closing prices and trading volume on the Indonesian stock exchange.

II. RELATED WORK

In its development, there are two main methods for predicting stock prices: conducting traditional analysis and analyzing stocks using artificial neural networks.[10]. However, traditional methods are ineffective in predicting stock worth by analyzing stock price movements using technical features with lots of data. On the contrary, analyzing stocks using artificial neural networks has developed rapidly in recent years because it has been proven to be able to take advantage of various data features, with large amounts of data with high transaction frequency.

The use of the neural network model in research to predict IBM stock was carried out in 1988 by making modifications to the learning standard calculation process, which is helpful for further research [11]. The use of the Back-Propagation network achieves a very high level of prediction accuracy within a short prediction time (ten days)[12] . Finding and using the Long Short Term Memory

(LSTM) algorithm is capable of solving long-term problems and remembering a collection of information that has been stored within that time span [7] . The application of Bidirectional LSTM (BLSTM) networks, and gradient modifications to the LSTM learning algorithm with the result that it outperforms unidirectional networks, is faster and more accurate than standard Recurrent Neural Nets (RNN) and Time-windowed Multilayer Perceptrons (MLPs) [13]. Using a variant of the Gated Recurrent Unit (GRU) [8] on RNNs maintains the structure and reduces the parameters on the update and reset gates. The results show that the GRU-RNN variant model works better and reduces the computational process[14].
Research developments in predicting current stock prices combine the best algorithmic methods by applying hybrid methods to Deep Learning algorithms to dig deep into data to improve the accuracy of stock price forecasting based on time series data. For example, use the CNN-LSTM algorithm for stock closing prices and the next day's forecast prices. This study combines the advantages of the Convolutional Neural Network (CNN), which has advantages in extracting compelling features from data; while LSTM can find interdependencies between technical analysis data on time series, this method can effectively increase the accuracy of forecasting stock prices [15]. In addition, developing a hybrid-based forecasting model using CNN and GRU in one framework results in better accuracy and efficiency performance [16].

III. Proposed Method

In this study, four processes are used in predicting stock prices, depicted in Figure 1, which consists of the following processes: Data initialization, Pre-processing, Prediction Model, and Evaluation. Data Initialization: Collecting data on share prices of finance companies sourced from the Indonesian stock exchange on the finance.yahoo.com website, which consists of daily data for the last ten years, with a total of 4973 trading days, transaction values are in IDR. These data are then downloaded using the CSV file format. The stock data features used are the closing price and stock volume. Then the CSV files are read and combined into a dataset with the 2D array dataframe type[17].

Pre-processing is the process of changing stock data from dataname type to NumPy type and forming a stock data dataset prepared as input data in the calculation process [18]. Furthermore, the dataset is normalized using the MinMaxScaler, which functions to calculate the value of all features with the same scale so that it will speed up the computational processing time. Then the dataset is divided into three datasets: training, evaluation, and testing. The training dataset is used in the training process; the evaluation dataset is used for the optimization process to compare prediction errors made by machine learning algorithm models. Finally, the test dataset is used to predict data and evaluate the model performance results in this study. Furthermore, these datasets are formed into a sliding window dataset consisting of input and output, and this dataset is in the form of a 3D array with the same window size [19].

Prediction is a training process using a sliding window dataset, starting with calculating the CNN layer. The data successively passes through the convolution layer and the pooling layer. This CNN process functions to extract the features in the input dataset, then proceed through the LSTM layer in which input gates control new values into cells. Next, data will pass through forget gates that control fixed values in cells, pass through the output gate to control the value in the cell, and generate the output data from the LSTM layer. Then the data will go through the GRU layer, where an input data flow process passes through the reset gate, which functions to combine new input data with past information. Next, the data will go through an update gate to determine how much past information should be stored. Moreover, this gate will generate output data from the GRU layer. Then the data will go through the entire connected layer to get the output of the final stock price estimate.

Evaluation is a process to determine the best method for predicting the closing price of a stock. In this study, we conducted a technical analysis by comparing the error rates of the models: CNN, LSTM, GRU, CNN-GRU, CNN-GRU-LSTM, and CNN-LSTM-GRU. In the process of calculating the evaluation of the error rate of prediction results from the various models above, we use the formula: MAE.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (1)$$

Where $\hat{y}_i$ is the predicted value and y_i is the actual value. The smaller the MAE value, the better the prediction result. The RMSE calculation formula is as follows:

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

Where $\hat{y}_i$ is the predicted value and y_i is the actual value. The smaller the RMSE value, the better the prediction result. The R^2 calculation formula is as follows:

$$\text{R}^2 = 1 - \frac{(\sum_{i=1}^{n}(y_i - \hat{y}_i)^2)/n}{(\sum_{i=1}^{n}(\bar{y}_i - \hat{y}_i)^2)/n} \quad (3)$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value. The range of R^2 values is (0, 1) RMSE values are forecasting accuracy values where the closer to 0, the higher the accuracy. While the value of R^2 is closer to 1, the better the prediction results.

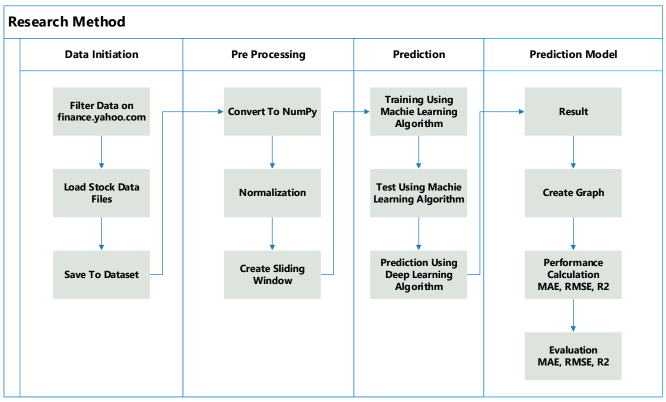


Fig. 1. Flow Process of Research Method

IV. EXPERIMENT RESULT

Based on the research method, we carry out data initiation and pre-processing processes to produce sliding window data consisting of datasets: training, evaluation, and testing. Furthermore, the prediction process for this dataset is carried out based on the architecture and hyperparameters used in this study. Figure 2 illustrates one of the CNN-LSTM-GRU model architectures used in this study, consisting of 2 Convolutional Layers, 2 Pooling layers, 1 LSTM layer, 1 GRU layer, 1 Flatten Layer, 3 Dense Layers, and 2 Dropout layers. The number of neurons in the LSTM layer is 128, and the number of neurons in the GRU layer is 192. The number of layers is part of the hyperparameters carried out for optimization in this study. The activation function used on the layer is the relay function. The dense layer only consists of 1 neuron because the required output is only one value, namely the predictive value of the stock price. The activation function used in the layer is linear. We use this function because it has a high prediction accuracy value [20].

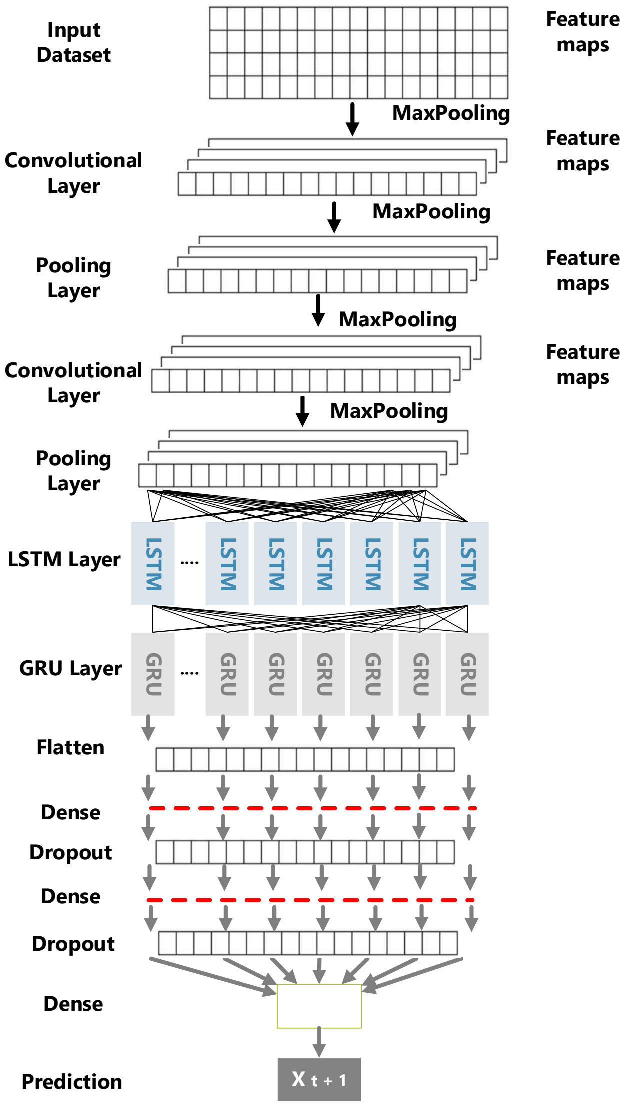


Fig. 2. The proposed framework uses the CNN-LSTM-GRU model for predicting closing stock prices

TABLE I. THE HYPERPARAMETERS TO BE USED IN THE CNN-LSTM-GRU MODEL

Layer (type)	Output Shape
conv1d (Conv1D)	(None, 15, 64)
max_pooling1d (MaxPooling1D)	(None, 14, 64)
conv1d_1 (Conv1D)	(None, 14, 32)
max_pooling1d_1 (MaxPooling 1D)	(None, 13, 32)
lstm (LSTM)	(None, 13, 128)
gru (GRU)	(None, 13, 192)
flatten (Flatten)	(None, 2496)
dense (Dense)	(None, 128)
dropout (Dropout)	(None, 128)
dense_1 (Dense)	(None, 32)
dropout_1 (Dropout)	(None, 32)
dense_2 (Dense)	(None, 1)

After doing the hyperparameters shown in Table 1, each algorithm: CNN, LSTM, GRU, CNN-GRU, CNN-GRU-LSTM, and CNN-LSTM-GRU, is processed using a training dataset. Furthermore, the training model generated by each of these algorithm processes is used to predict the test dataset and produce a model that will issue output in the form of a predicted value. Finally, this value will be compared with the actual value. For example, figure 3 illustrates the actual value and the predicted value of the closing price of shares (PTBA, PGAS, and ELSA) using CNN-LSTM-GRU.

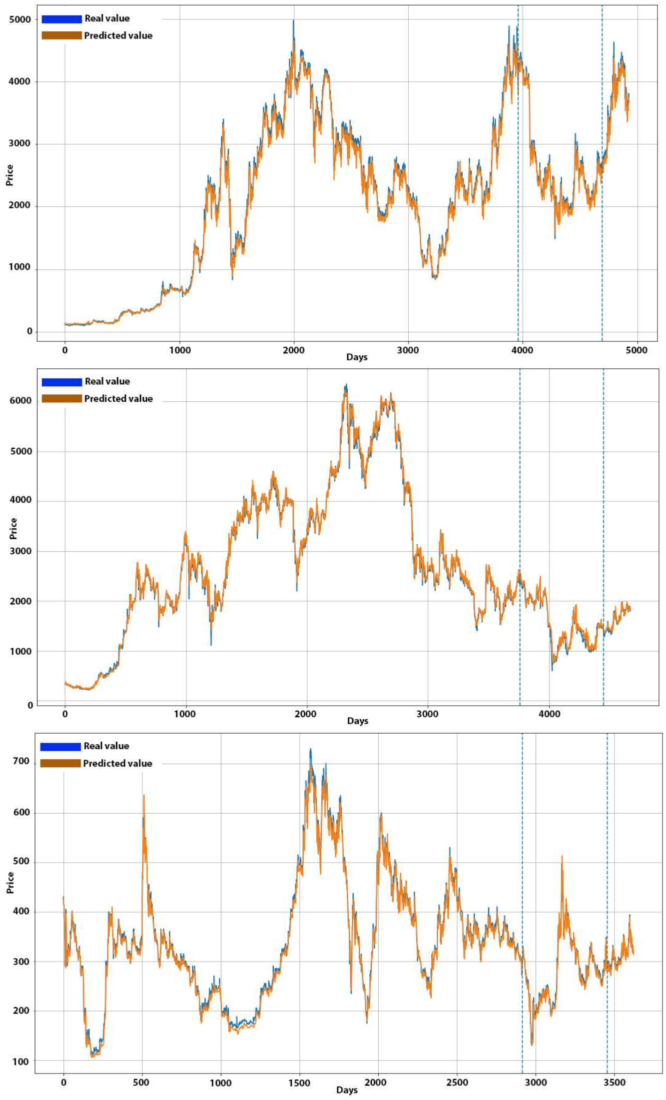


Fig. 3. Predicted value and actual value of PTBA, PGAS and ELSA shares using CNN-LSTM-GRU.

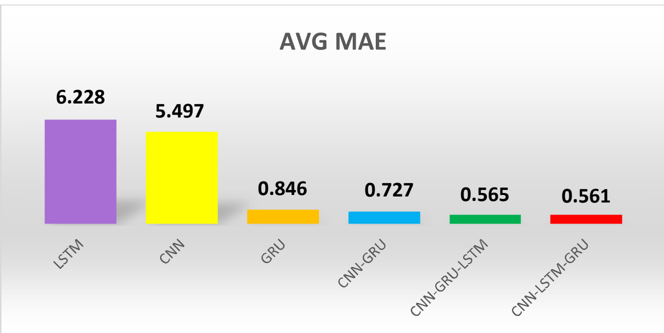
After we compared the prediction results of stock values through the six forecasting methods CNN, LSTM, GRU, CNN-GRU, CNN-GRU-LSTM, and CNN-LSTM-GRU, we concluded that the CNN-LSTM-GRU method has the most accurate prediction value match rate and close to the actual value. The formula to evaluate the suitability of the stock price's predicted value and the actual value is: Mean Absolute Error (MAE). The smaller the MAE value, the better the forecast. Root Mean Square Error (RMSE), the smaller the

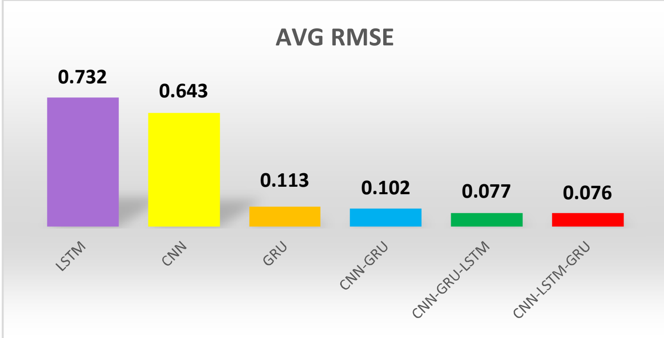
RMSE value, the better the forecast. R-square (R^2), With a range of R^2 values from 0 to 1. The closer the MAE and RMSE values are to 0, the smaller the error between the predicted value and the actual value, and the higher the forecasting accuracy. The closer R^2 is to 1, the better the degree of correspondence between the values. Table 2 contains information on the results of calculations from the error rate prediction of PTBA, PGAS, and ELSA stock prices using MAE, RMSE, and R^2.

TABLE II. MAE, RMSE AND R^2 RESULT

Algo-rithm	Stock	MAE	AVG MAE	RMSE	AVG RMSE	R^2	AVG R^2
CNN	PTBA	5.037	5.497	0.704	0.643	0.103	1.946
	PGAS	10.714		1.14		4.888	
	ELSA	0.74		0.086		0.846	
LSTM	PTBA	10.769	6.228	1.267	0.732	-2.567	-1.645
	PGAS	7.422		0.866		-2.393	
	ELSA	0.492		0.063		0.024	
GRU	PTBA	1.437	0.846	0.188	0.113	0.921	0.933
	PGAS	1.016		0.138		0.913	
	ELSA	0.085		0.012		0.964	
CNN-GRU-LSTM	PTBA	0.66	0.565	0.092	0.077	0.981	0.947
	PGAS	0.904		0.121		0.933	
	ELSA	0.131		0.017		0.926	
CNN-GRU-GRU	PTBA	0.755	0.727	0.104	0.102	0.976	0.924
	PGAS	1.334		0.187		0.842	
	ELSA	0.092		0.014		0.953	
CNN-LSTM-GRU	PTBA	0.834	0.561	0.106	0.076	0.975	0.953
	PGAS	0.731		0.107		0.948	
	ELSA	0.119		0.016		0.937	

The graph in figure 4 illustrates the average values of MAE, RMSE, and R^2 obtained from the predicted closing values of PTBA, PGAS, and ELSA shares.





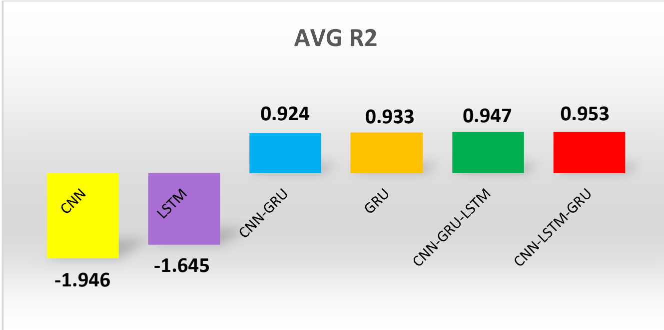


Fig. 4. Comparison of the average value of MAE, RSME and R^2

Based on a comparison of the predicted value and the real value obtained from the experimental results using the six algorithms in predicting the closing price of shares, shown in Table 2 and Figure 4. If you use only one GRU algorithm, you will get an MAE value of 0.846 and an RMSE value of 0.113, which is the smallest and the best compared to the CNN or LSTM algorithms. GRU has an R2 value of 0.93, which is closer to 1. This value is also the best compared to using the CNN and LSTM algorithms. Even though the performance of GRU as a single algorithm is quite good, this performance can be improved by using a combination or hybrid algorithm.

From the tables above, there is an increasing trend in the algorithm performance in the two existing hybrid models, the CNN-GRU algorithm and the CNN-GRU-LSTM, compared to the single GRU algorithm. However, this study found that the combination of three algorithms CNN-LSTM-GRU, can outperform the two previous hybrid methods and produce the most accurate predictions of stock values. The CNN-LSTM-GRU hybrid method has the smallest and the most accurate MAE and RMSE values of 0.561 and 0.076 and the R^2 value of 0.953, which is closest to 1. By comparing the GRU algorithm method and the CNN-LSTM-GRU hybrid algorithm, it was found that the MAE value decreased by 0.285 or 33.69%, for the RMSE value, it decreased by 0.04 or 32.35%, and for the R^2 value, it increased by 0.021 or 2.22%. These values are the best compared to the use of other hybrid algorithms.

V. CONCLUSION

The results of our research in predicting closing prices using the CNN-LSTM-GRU hybrid algorithm produce the most accurate predictive values compared to other prediction methods, either using one algorithm: GRU, CNN, and LSTM, or other hybrid algorithms such as CNN-LSTM and CNN-GRU-LSTM.

Prediction using the CNN-LSTM-GRU hybrid algorithm, when compared with the GRU single algorithm, produces a lower MAE value, which is reduced to 0.285 or 33.69%, a lower RMSE value, which is reduced by 0.04 or 32.35%, and R^2 value is increased by 0.021 or 2.22%. The test results show that the stock value prediction using a hybrid algorithm produce better accuracy than using only single algorithm. The use of the CNN-LSTM-GRU hybrid algorithm is also proven to reduce the error value and improve the prediction results of closing stock prices, even when it is compared to other hybrid algorithms.

[1] Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Comput. Appl.*, vol. 2, 2019, doi: 10.1007/s00521-019-04504-2.

[2] V. Ionescu, "What is the Causal Relationship between STOXX EUROPE 600 Sectors ? But between Large Firms and Small Firms ? West University of Timisoara Director of the Strategy and Financial Stability EUROPE 600 SECTORS ? BUT BETWEEN LARGE FIRMS AND," no. December, 2021, doi: 10.24818/18423264/53.3.19.01.

[3] T. Krulicky, "Innovations in management forecast: time development of stock prices with neural networks," *Mark. Manag. Innov.*, vol. 6718, no. 2, pp. 324–339, 2020.

[4] T. Oberlechner and A. Nimgade, "Work stress and performance among financial traders," *Stress Heal.*, vol. 21, no. 5, pp. 285–293, 2005, doi: 10.1002/smi.1063.

[5] M. Roondiwala, H. Patel, and S. Varma, "Predicting Stock Prices Using LSTM," *Mach. Learn. with Appl.*, no. April 2017, pp. 4–7, 2018, doi: 10.21275/ART20172755.

[6] S. Chen and H. He, "Stock Prediction Using Convolutional Neural Network," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 435, no. 1, 2018, doi: 10.1088/1757-899X/435/1/012026.

[7] 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.

[8] 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

[9] I. Darmawan, H. Siregar, D. B. Hakim, and A. H. Manurung, "The Effect of Crude Oil Price Shocks on Indonesia Stock Market Performance," *Indones. Cap. Mark. Rev.*, vol. 16, no. 1, pp. 11–23, 2020, doi: 10.33830/jom.v16i1.785.2020.

[10] J. Li, S. Pan, L. Huang, and X. Zhu, "A machine learning based method for customer behavior prediction," *Tehnicki Vjesnik*, vol. 26, no. 6. pp. 1670–1676, 2019. doi: 10.17559/TV-20190603165825.

[11] H. White, "Economic prediction using neural networks: The case of IBM daily stock returns," *IEEE 1988 Int. Conf. Neural Networks*, pp. 451–458, 1988, doi: 10.1109/icnn.1988.23959.

[12] E. Schöneburg, "Stock price prediction using neural networks: A project report," *Neurocomputing*, vol. 2, no. 1, pp. 17–27, 1990, doi: 10.1016/0925-2312(90)90013-H.

[13] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, 2005, doi: 10.1016/j.neunet.2005.06.042.

[14] R. Dey and F. M. Salemt, "Gate-variants of Gated Recurrent Unit (GRU) neural networks," *Midwest Symp. Circuits Syst.*, vol. 2017-Augus, no. 2, pp. 1597–1600, 2017, doi: 10.1109/MWSCAS.2017.8053243.

[15] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-Based Model to Forecast Stock Prices," *Hindawi*, vol. 2020, 2020.

[16] M. Sajjad *et al.*, "A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting," *IEEE Access*, vol. 8, pp. 143759–143768, 2020, doi: 10.1109/ACCESS.2020.3009537.

[17] K. A. Althelaya and S. Mohammed, "Evaluation of Bidirectional LSTM for Short- and Long-Term Stock Market Prediction," *IEEE Xplore*, pp. 151–156, 2018.

[18] S. Van der Walt and M. Aivazis, "The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering," *Comput. Sci. Eng.*, vol. 13, no. 2, pp. 22–30, 2011, [Online]. Available: http://aip.scitation.org/doi/abs/10.1109/MCSE.2011.37

[19] T. G. Dietterich, "Machine Learning for Sequential Data : A Review," pp. 15–30, 2002.

[20] M. Rana, "Effects of Activation Functions and Optimizers on Stock Price Prediction using LSTM Recurrent Networks," *Int. Conf. Comput. Sci. Artif. Intell.*, pp. 354–358, 2019.