BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1317-1328
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1317-1328

CRYPTOCURRENCY TIME SERIES FORECASTING MODEL USING GRU ALGORITHM BASED ON MACHINE LEARNING

Melina Melina [1*], Sukono[2], Herlina Napitupulu[3], Norizan Mohamed[4], Yulison Herry Chrisnanto[5], Asep ID Hadiana[6], Valentina Adimurti Kusumaningtyas[7]

[1,5,6]*Department of Informatics, Faculty of Science and Informatics, Universitas Jenderal Achmad Yani*
[7]*Department of Chemistry, Faculty of Science and Informatics, Universitas Jenderal Achmad Yani*
Jln. Terusan Jend. Sudirman, Cimahi, 40531, Indonesia

[2,3]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran*
Jln. Raya Bandung Sumedang KM.21, Jatinangor, 45363, Indonesia

[4]*Department of Mathematics, Faculty of Ocean Engineering Technology and Informatics Universiti Malaysia*
Terengganu, Kuala Nerus Terengganu, 21030, Malaysia

Corresponding author's e-mail: * ***melina@lecture.unjani.ac.id***

Article History:

Received: 22nd October 2024
Revised: 2nd February 2025
Accepted: 5th March 2025
Published: 1st April 2025

Keywords:

Cryptocurrency;
GRU;
Learning Curve;
Machine Learning;
Time Series.

ABSTRACT

The cryptocurrency market is experiencing rapid growth in the world. The high fluctuation and volatility of cryptocurrency prices and the complexity of non-linear relationships in data patterns attract investors and researchers who want to develop accurate cryptocurrency price forecasting models. This research aims to build a cryptocurrency forecasting model with a machine learning-based time series approach using the gated recurrent units (GRU) algorithm. The dataset used is historical Bitcoin closing price data from January 1, 2017, to July 31, 2024. Based on the gap in previous research, the selected model is only based on the accuracy value. In this study, the chosen model must fulfill two criteria: the best-fitting model based on the learning curve diagnosis and the model with the best accuracy value. The selected model is used to forecast the test data. Model selection with these two criteria has resulted in high accuracy in model performance. This research was highly accurate for all tested models with MAPE < 10%. The GRU 30-50 model is best tested with MAE = 867.2598, RMSE = 1330.427, and MAPE = 1.95%. Applying the sliding window technique makes the model accurate and fast in learning the pattern of time series data, resulting in a best-fitting model based on the learning curve diagnosis.

How to cite this article:

M. Melina, Sukono, H. Napitupulu, N. Mohamed, Y. H. Chrisnanto, A. I. Hadiana and V. A. Kusumaningtyas., "CRYPTOCURRENCY TIME SERIES FORECASTING MODEL USING GRU ALGORITHM BASED ON MACHINE LEARNING," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1317-1328, June, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

One of the currency innovations that is currently trending is the presence of cryptocurrency as a digital currency that can be used as a means of payment in addition to conventional currencies [1]. Cryptocurrency is a digital currency obtained from code called blockchain, whose security is guaranteed by cryptography and built on a blockchain that will ensure the security of online transactions. This currency is flexible, transparent, fast, and the transactions are done virtually or over the Internet [2]. In recent years, cryptocurrency has become an increasingly popular investment option among investors due to its properties and high returns [3], [4].

Bitcoin is the most popular type of cryptocurrency, introduced by Satoshi Nakamoto in 2009 [5]. The drastic price fluctuations of Bitcoin have led to many being interested in investing in Bitcoin [6]. Crypto investments present unique challenges and risks to investors due to fluctuations, high volatility, and regulatory risks [7]. Uncertainty factors influence the value of the increase and decrease in the price of Bitcoin, so a method is needed to predict future prices to maximize profits in investing [8]. Methods that can estimate the volatility of cryptocurrencies are essential [9]. Traditional methods have been used to predict the price of Bitcoin, but these methods have problems achieving high accuracy due to heteroscedasticity and nonlinear data characteristics. In recent years, machine learning (ML) has attracted increasing popularity in modern financial applications, and the rapid development of ML can make classifiers more powerful for solving nonlinear problems [10], [11].

Recently, with the increasing use of artificial intelligence in many fields, the application of such models has expanded to research in financial markets, particularly stock and cryptocurrency markets [12]. Traditional time series analysis is often less effective in modeling the complex trends and high volatility associated with cryptocurrencies. Hence, these models fail classical assumptions such as residual normality, and residual variance is constant (homoscedasticity) [13]. Therefore, more sophisticated approaches, such as ML, can provide better solutions for predicting future cryptocurrency price movements [14].

Forecasting time series data is forecasting events or values that will be obtained in the future based on previous data. Time series data is displayed based on time with the characteristics of time series data. One ML-based time series forecasting method is recurrent neural networks (RNNs). The RNNs are an artificial neural network (ANN) that processes sequential data. The RNNs can store memory of previous information in time series data. As the size and complexity of circuits increase, issues such as exploding or missing gradients and computational inefficiency have added complexity to RNN algorithms. Long short-term memory (LSTM) was introduced by Hochreiter and Schmidhuber (1997) as a solution to the missing gradient problem [15]. The LSTM has memory cells that effectively act as long-term memory storage while being able to update information [16], [17]. The GRU is a simpler variation of LSTMs, and both have a similar design developed by Cho et al. (2014) [18].

Cryptocurrency forecasting research conducted by Rizkilloh and Widiyanesti (2022), aims to forecast prices in the cryptocurrency market using the LSTM algorithm. This research data is sourced from the Yahoo Finance site using the Pandas Data reader library through Google Collaboratory. The results of this study show that the best prediction performance is obtained on the type of DOGE coin with the number of Epoch 20 and an RMSE value of 0.0630 [19]. Research by Maliki et al. (2023) predicts the price of Bitcoin cryptocurrency against the rupiah currency using the LSTM algorithm with several processes, namely pre-processing, normalization, LSTM algorithm training, and regression matrix evaluation using mean square error (MSE). The results of this study show that the LSTM algorithm can predict Bitcoin price movements by obtaining an MSE of 0.00374 [20]. Research conducted by Wang et al. (2023), applied ML techniques to forecast cryptocurrency volatility using internal determinants and compared the forecast performance of ML techniques with traditional GARCH volatility models, and explored the determinants of volatility forecasts. Empirical results show that ML techniques outperform traditional methods in forecasting cryptocurrency volatility, with Artificial Bee Colony-LSTM showing the best prediction performance. Random forest and LSTM methods significantly outperform traditional volatility models such as GARCH [21].

Several studies have demonstrated the reliability of the GRU algorithm for forecasting. Research by Bouteska et al. (2024) where large fluctuations in nonstationary cryptocurrency prices motivated the urgent need for accurate forecasting models. The results of this study revealed that the GRU, RNNs, and LightGBM methods outperformed the naive buy-and-hold and random walk strategies [22]. Likewise, Arun Kumar et al. (2022) have conducted research by comparing the performance of ANN models (GRU and LSTM) with statistical techniques (ARIMA and SARIMA) to forecast COVID-19 trends. The result is that the LSTM and

GRU models outperform the ARIMA and SARIMA models [23]. Viéitez et al. (2024) have aimed to test the validity of applying ML techniques to data derived from a highly volatile market. As a result, ANN approaches, in particular LSTM and GRU, predicted the price and trend of the cryptocurrency, Ethereum, with good precision [24]. Likewise, the research of Jaquart et al. (2022) used and analyzed various ML models for daily cryptocurrency market prediction and trading. The results show that the LSTM and GRU models are dominant. LSTM and GRU ensemble model predictions will be used for long-short portfolio strategies [25].

The gap between this research and previous research is that the best model selection is generally only based on the model's accuracy, without diagnosing whether the model has been able to read the learning data pattern accurately. The novelty of this research is that the observation of the learning process is based on the learning curve (LC). The model is evaluated based on the LC performance, which shows how well the model learns. If the diagnosis of the LC is still under-fitting, over-fitting, or not optimal, hyperparameter settings are made until the best-fitting model is achieved. Motivated by this research gap, this study aims to build an ML-based time series forecasting model with the best accuracy value from several model architectures tested. It is a best-fitting model that can read the training data pattern well. This model will forecast cryptocurrency prices for the following seven periods. The framework includes data preprocessing, splitting, setting hyperparameters to build the GRU model architecture, training, LC diagnosis, model testing, accuracy, and forecasting for the following seven periods. This research provides new insights for investors, governments, and researchers and a tool to minimize investment risk [26].

2. RESEARCH METHODS

2.1 Dataset

Daily cryptocurrency time series data is obtained from https://finance.yahoo.com. The most popular cryptocurrency chosen is BTC/USD. Data is obtained from January 1, 2017, to July 31, 2024, with 2769 observations.

2.2 Min-Max Normalization

Min-max normalization of data is used to scale the data to the interval [0,1]. This is generally useful to ease the computation of ML algorithms. Min-max normalization is written as Equation (1) [27]:

$$x''_{min-max} = \frac{x - \min(x)}{max(x) - \min(x)} \tag{1}$$

where, x : the original price, $min(x)$: the minimum value in the data set, $max(x)$: the maximum value in the data set, and $x''_{min-max}$: the scale of the data.

2.3 Gated Recurrent Units

The gated recurrent units (GRU) is another effective alternative to the vanishing or exploding gradient problem in RNNs. The GRU has only one the hidden state (h_t), therefore the GRU architecture is simpler than LSTM. The GRU has a special mechanism to determine h_t must be updated and h_t must be reset. The cell diagram of GRU with reset gate and update gate is shown in Figure 1 [28], [29].

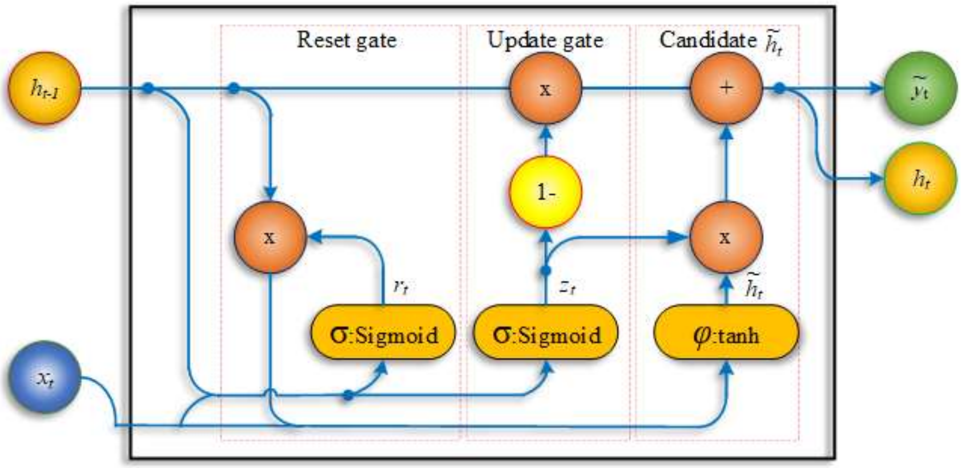


Figure 1. The GRU Architecture

Illustration of the GRU process is shown in **Figure 1**, The following are the steps of the GRU algorithm used in this research:

a) Forward Pass

For each timestep t.

1. Calculate the reset gate (r_t), mathematically the reset gate is written as **Equation (2) [30]**

$$r_t = \sigma(w_r x_t + U_r h_{t-1} + b_r), \tag{2}$$

where r_t : the reset gate, W_r, U_r : parameter weights matrix, x_t : the input vector, h_{t-1} : the previous hidden state, b_r : the bias, σ : the sigmoid activation function.

2. Calculate the update gate (z_t), the update gate is written

$$z_t = \sigma(w_z x_t + U_z h_{t-1} + b_z), \tag{3}$$

where z_t : the update gate, W_z, U_z : parameter weights matrix, x_t : the input vector, h_{t-1} : the previous hidden state, b_z : the bias, σ : the sigmoid activation function.

3. Calculate the hidden state candidate $(\tilde{h}_t)$, the hidden state candidate is written

$$\tilde{h}_t = \varphi(W_h x_t + U_h[r_t \odot h_{t-1}] + b_h), \tag{4}$$

where $\tilde{h}_t$: the candidate hidden state, W_h, U_h : parameter weights matrix, x_t : the input vector, h_{t-1} : the previous hidden state, b_h : the bias, φ : the hyperbolic tangent activation function, operator $\odot$: the Hadamard product.

4. Calculate the hidden state (h_t) , and the hidden state is written

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t), \tag{5}$$

where h_t : the hidden state, z_t : the update gate, h_{t-1} : the previous hidden state, $\tilde{h}_t$: the candidate hidden state, and operator $\odot$: the Hadamard product.

b) Backpropagation and Optimization.

1. Calculate the loss using the MSE function.
2. Perform backpropagation.

3. Update the weights using the Adam or RMSprop optimization algorithm.

c) Prediction.

1. Model input data is test data.
2. Generate predictions for each timestep.
3. Un-normalization of data.
4. Calculate accuracy.

The GRU is designed to overcome the vanishing gradient problem in the RNNs so it can handle long data sets better. The GRU has a more straightforward structure than the LSTM because it only uses a reset gate and update gate, unlike the LSTM, which has an input gate, output gate, and forget gate. This simple structure makes the GRU faster in the training and prediction process, so the GRU requires lower computation time and is more memory efficient.

2.4 Model Accuracy

Model performance evaluation is used to determine how good the model output is, using the mean absolute error (MAE), root mean squared error (RMSE), and the mean absolute percentage error (MAPE) **[31]**. Forecasting error is the difference between actual and forecast, written as **Equation (6)**:

$$e_t = Y_t - \hat{Y}_t, \tag{6}$$

where e_t : the forecasting error at time t, n : the amount of data, Y_t : the actual data of time t, and $\hat{Y}_t$: the result of forecasting time t.

MAE is the average absolute difference between the actual value and the forecasting value. MAE is written as **Equation (7)**:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|e_t|, \tag{7}$$

where MAE: mean absolute error value, n: the amount of data.

RMSE calculates the average of the squared difference between the actual value and the predicted value; then the square root is taken. RMSE is written as **Equation (8) [32]**.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}, \tag{8}$$

where $RMSE$: root mean squared error value, n : the amount of data.

The accuracy measure of forecasting used in addition to MAE and RMSE is MAPE which is written as **Equation (9) [33]**.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t}{Y_t}\right|, \tag{9}$$

where $MAPE$: mean absolute percentage error value, e_t : the forecasting error at time t, n : the amount of data, and Y_t : the actual data of time t.

The assessment of forecasting accuracy is, if MAPE ≤ 0.1 then the assessment is highly accurate, if 0.1 ≤ MAPE ≤ 0.2 then the assessment is a good forecast, if 0.2 < MAPE ≤ 0.5 then the assessment is a reasonable forecast, and if MAPE > 0.5 then the assessment is an inaccurate forecast [34].

2.5 Methodology

This research uses RStudio software for the computational process. Figure 2 shows the methodology of this research.

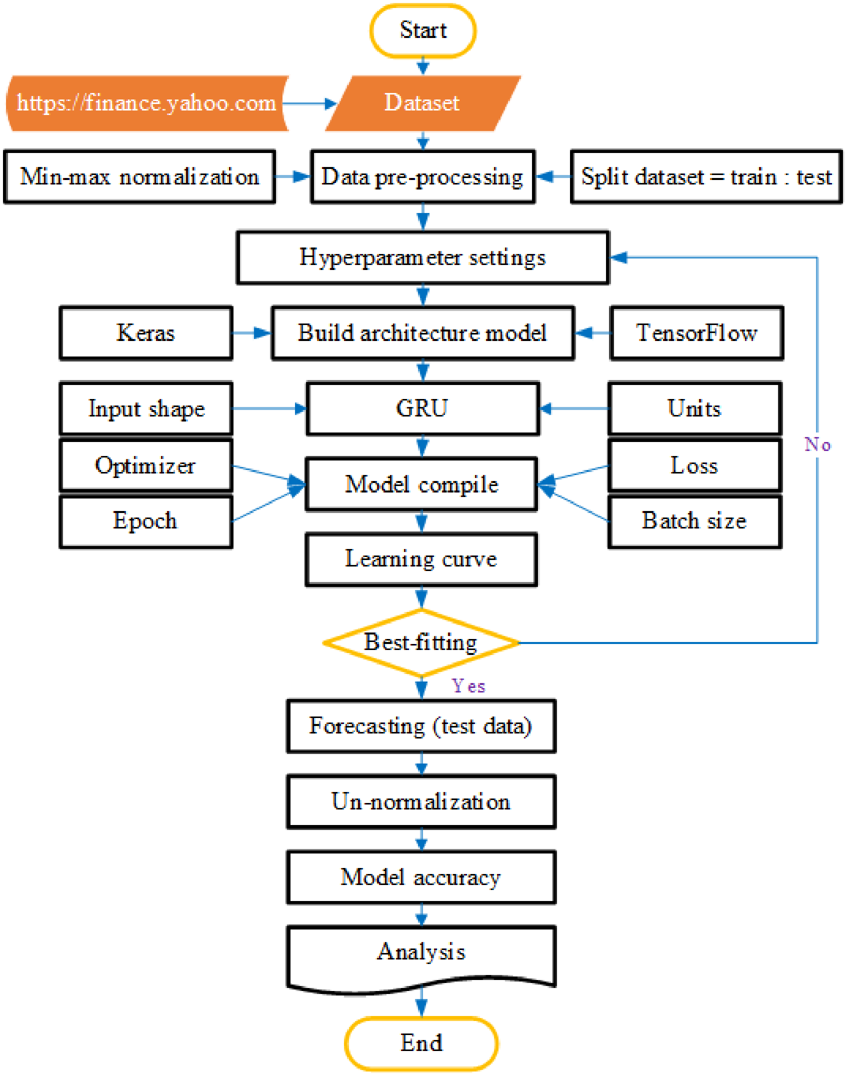


Figure 2. Methodology

Figure 2 shows the stages of the research methodology. In this study, RStudio software is used for computational calculations. The cryptocurrency forecasting model using a machine learning-based GRU algorithm is built using high-level neural network APIs such as 'Keras' and 'TensorFlow.' These Keras and TensorFlow packages are available in RStudio software.

The first step of this research is to download the dataset from the website https://finance.yahoo.com. Next is data pre-processing. Data min-max normalization is performed using Equation (1). at this stage. Next, the data is divided by 80% for training and 20% for test data. Hyperparameter values are set to build the model architecture. Using Keras and TensorFlow functions, the model architecture is built. In the learning stage, the built model learns the pattern of the learning data. The results of this learning are displayed in the form of the LC. A good model is a best-fitting model based on the LC. Hyperparameter settings will be performed until the best-fitting model is obtained. The best-fitting model is selected based on the accuracy value. The best-fitting model with the best accuracy value forecasts the test data. Un-normalization is performed to restore the data into integer numbers to calculate model accuracy. Model performance evaluation is used to determine how sound the model output is. The best-fitting and best model based on model accuracy is used to forecast the Bitcoin price seven periods ahead. Finally, analyses are conducted to conclude.

3. RESULTS AND DISCUSSION

Data pre-processing is the initial stage of preparing data and eliminating some problems that can interfere with data processing, such as missing data and inappropriate data formats. Min-max normalization is done by scaling the data values between the range [0,1] because ML algorithms work well on data with a small and standard scale.

The sliding window (SW) technique is used to derive patterns in time series data, which is commonly used in ML algorithms while reducing the complexity of the algorithms used. Table 1, presents the conversion of training data to the SW matrix with $k = 10$.

Table 1. Conversion of Training Data to SW Matrix With k = 10

Timestep (t)	Observation Variable / Input (X)								Output (y)
	X_t	X_{t+1}	X_{t+2}	X_{t+3}	X_{t+4}	X_{t+5}	…	X_{t+9}	X_{t+k}
1	998.33	1021.75	1043.84	1154.73	1013.38	902.20	…	907.68	777.76
2	1021.75	1043.84	1154.73	1013.38	902.20	908.59	…	777.76	804.83
3	1043.84	1154.73	1013.38	902.20	908.59	911.20	…	804.83	823.98
4	1154.73	1013.38	902.20	908.59	911.20	902.83	…	823.98	818.41
5	1013.38	902.20	908.59	911.20	902.83	907.68	…	818.41	821.80
6	902.20	908.59	911.20	902.83	907.68	777.76	…	821.80	831.53
7	908.59	911.20	902.83	907.68	777.76	804.83	…	831.53	907.94
8	911.20	902.83	907.68	777.76	804.83	823.98	…	907.94	886.62
9	902.83	907.68	777.76	804.83	823.98	818.41	…	886.62	899.07
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2205	20976.30	20976.30	20976.30	20976.30	20976.30	20976.30	…	20976.30	22636.47

Table 1, shows the conversion of training data using the SW technique. The SW technique shifts the window by the window length (k), and the window containing data of length k is used to predict the next value. The time series data is converted into pairs of inputs (X) and outputs (y). Each input (X) is the data in a window of length k, and the output (y) is the next value. The aim is to speed up capturing the time pattern in the data when the model learns the relationship of the variable with the previous time variable. In addition, this technique increases the amount of training data. For $k = 10$, it means that the previous 10-time values will be used to predict the 11th-time value. For timestep = 1 data $X_1 - X_{10}$, with variable values [998.33, 1021.75, 1154.73, ..., 907.68] as input and output is $X_{11} = 777.76$. For timestep = 2 data $X_2 - X_{12}$, with variable values [1021.75, 1043.84, 1154.73, ..., 777.76] are input data, and the output is $X_{12} = 804.83$. This process will stop if $timestep = number\ of\ data - k$. The GRU model is trained using input (X) and output (y) pairs. This SW technique can improve model accuracy, as the model is trained by shifting one data block [35].

The data that has been converted into an SW matrix will be min-max normalized into numbers with interval [0,1]. In building an ML-based sequence forecasting model with a hard library, the input data is converted into an array reshape (sample-batch-size, window-size, target). Sample-batch-size is the number of data samples partitioned (n) based on the window size (k), and the target is the value that will be used as the output reference, in this case, $target = 1$. This dataset is divided with a composition of $80{:}20$. Based on this composition, the number of train data samples is 2215, namely period 1 to period 2215, and the number of test data is 553 periods, starting from period 2216 to period 2769. So, the input frame for training data with k = 10 is array reshape (2205, 10, 1), and array reshapes (2185, 30, 1) for $k = 30$. The input frame for test data for k = 10 is array reshapes (543, 10, 1), and for $k = 30$ is array reshape (523, 30, 1). The parameters are set to obtain the optimal model architecture. The following Table 2 displays the hyperparameter settings.

Table 2. The Hyperparameter Settings

Parameter	Value	Status
Training data	2215	Fixed
Test data	553	Fixed
Window size	10 and 30	Experiment
Target	1	Fixed

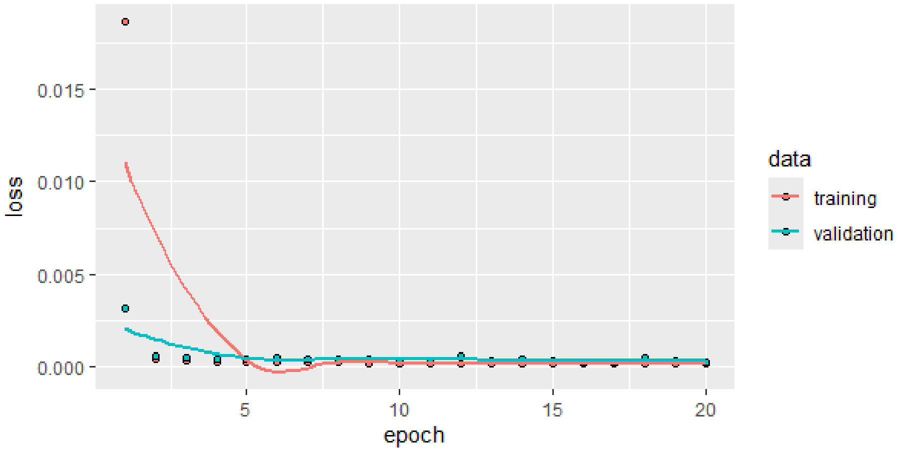
Parameter	Value	Status
ANN type	GRU	Fixed
Hidden layers (HL)	1, 2	Fixed
Units	10 and 50	Experiment
Batch size	32	Fixed
Validation split	0.2	Fixed

Experiments are then conducted to find the significant model architecture that gives the best results. The experiment aims to get the best-fitting model architecture based on the diagnosis of the LC. The best-fitting model that we tested with test data to obtain the model accuracy value based on MAE, RMSE, and MAPE. **Table 3**, shows the model architecture and its parameter settings.

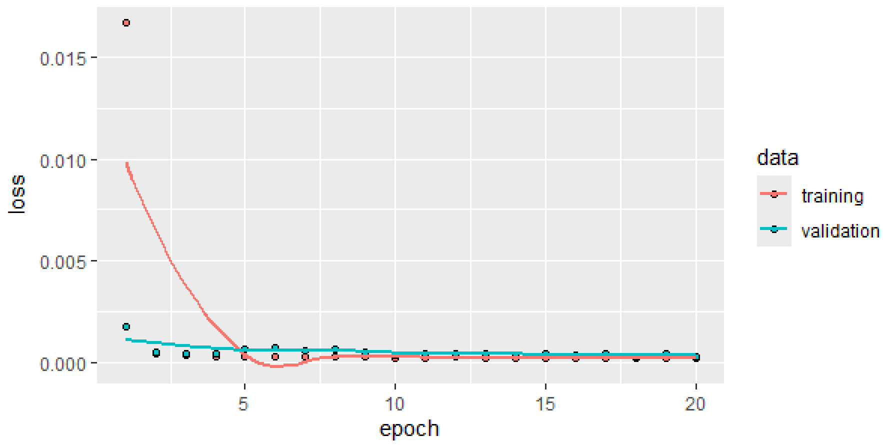
Table 3. Model Architecture and Parameter Settings

Model	Window	HL	
		1	2
GRU 10-10	10	10	
GRU 10-10-10	10	10	10
GRU 10-50	10	50	
GRU 10-50-50	10	50	50
GRU 30-10	30	10	
GRU 30-10-10	30	10	10
GRU 30-50	30	50	
GRU 30-50-50	30	50	50

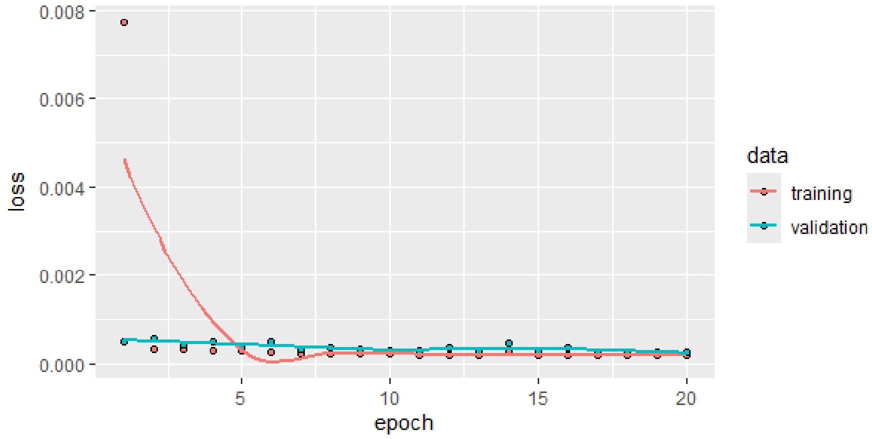
The model was evaluated, and we set the hyperparameter to improve the model's performance when the LC evaluation was not optimal. After repeated experiments, the hyperparameter settings were obtained to get the best-fitting model architecture. **Figure 3** displays the LC of the four models with the best accuracy.



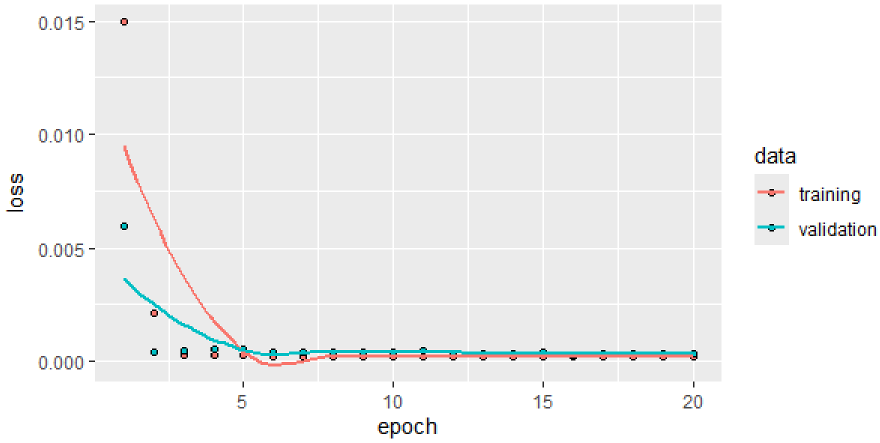
(a)



(b)



(c)



(d)

Figure 3. The LC of (a) The GRU 30-50, (b) The GRU 10-50, (c) The GRU 30-50-50, (d) The GRU 30-10 Models

Figure 3 visually displays the LC of four of the eight models tested, sorted by best accuracy. In general, the LC diagnosis of the tested models all resulted in best-fitting models. The decreasing training and validation loss values identify the LC diagnosis for the best-fitting model. It is stable with a minimal difference between the training and validation loss values until the end of the epoch. Based on the LC diagnosis, it can be concluded that the model successfully reads the time series data pattern in the training stage very quickly with only an $epoch < 20$. We will use this best-fitting model to forecast the test data. The following **Table 4**, displays the LC diagnosis of the model and the MAE, RMSE, and MAPE accuracy values obtained.

Table 4. The Learning Curve Diagnosis and Model Accuracy

Model	LC Diagnosis	Accuracy		
		MAE	RMSE	MAPE
GRU 10-10	Best-fitting	1133.803	1686.217	2.55%
GRU 10-10-10	Best-fitting	1066.422	1490.236	2.58%
GRU 10-50	Best-fitting	867.6518	1338.093	1.98%
GRU 10-50-50	Best-fitting	1247.313	1699.973	2.85%
GRU 30-10	Best-fitting	1036.075	1467.381	2.36%
GRU 30-10-10	Best-fitting	1150.842	1557.443	2.79%
GRU 30-50	Best-fitting	867.2598	1330.427	1.95%
GRU 30-50-50	Best-fitting	949.8794	1417.904	2.10%

Table 4 shows the results of the LC diagnosis and the accuracy of the eight models tested. The method used in this study is GRU. Each model is differentiated based on its parameters: window size (k), number of HLs, and number of units/neurons. All tested models produce best-fitting models based on LC evaluation and are forecasting models with high accuracy based on the acquisition of MAPE < 10%. The GRU 30-50 model outperformed all other candidate models with MAE = 867.2598, RMSE = 1330.427, and MAPE = 1.95%. Therefore, the GRU 30-50 model was selected to forecast the Bitcoin price. **Figure 4**, displays a graph of the

results of predicting training and test data and the forecast results for the following seven periods using the GRU 30-50 model.

Figure 4. Forecasting Graphs of Training and Test Data

Figure 4 visually shows the forecasting results of the GRU 30-50 model with training data. It can be seen that the model successfully reads the training data pattern well, where the forecasting value is almost close to the actual data value. Likewise, when this model is used to forecast the test data, The output of this model produces forecasting values that are close to the test data values. After getting the best performance and accuracy of all tested models, the GRU 30-50 model is used to forecast Bitcoin prices for the following seven periods. **Table 5** shows the forecasting results using the GRU 30-50 model.

Table 5. Bitcoin Price Forecasting Seven Periods Ahead

Period	Date	Forecast
1	8/1/2024	66094.04
2	8/2/2024	65649.83
3	8/3/2024	65227.07
4	8/4/2024	64816.19
5	8/5/2024	64414.34
6	8/6/2024	64021.02
7	8/7/2024	63636.53

Furthermore, **Figure 5** visualizes the output of the model, which is a short-term forecast of the Bitcoin price.

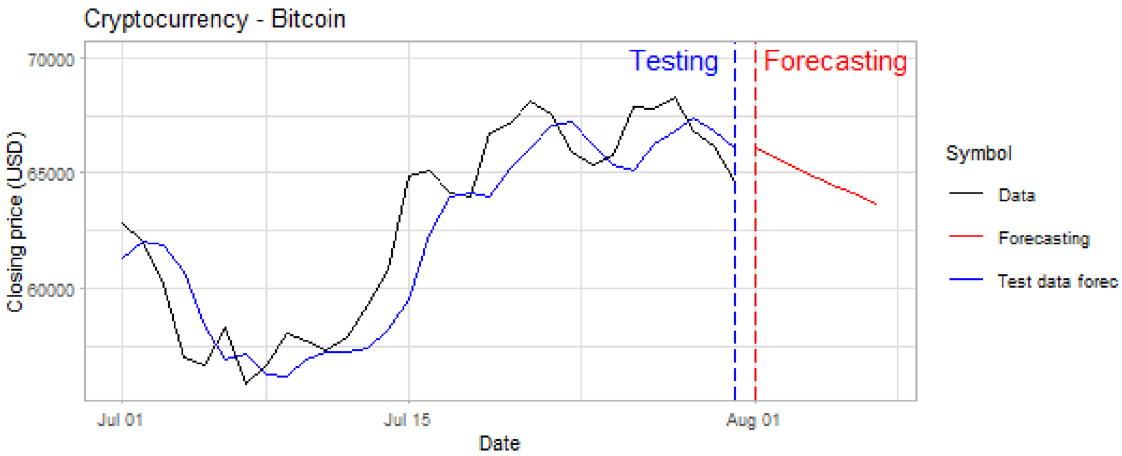


Figure 5. Forecasting Plot of The GRU 30-50 Model

Figure 5 shows the test data and its forecast from July 1-31, 2024, as well as the results of the Bitcoin price forecast for the next seven periods from August 1-7, 2024. It can be seen that the Bitcoin price will experience a slow decline from period 1 to period 7. This forecast result continues the downward trend in

Bitcoin price. The forecast for Bitcoin prices to continue to decline is in line with the trend of test data, where the last 30 periods of Bitcoin prices have decreased.

4. CONCLUSIONS

This research uses the GRU algorithm to build a cryptocurrency forecasting model with an ML-based time series approach. Applying the SW technique allows all models to learn time series patterns very quickly and accurately, as shown in **Figure 3**, the training process is performed with $epoch \leq 20$ and produces best-fitting models based on LC diagnosis while improving model performance. **Table 4** shows that all of these best-fitting models produce high accuracy based on the acquisition of MAPE values. The model selected for forecasting seven days is the best-fitting model based on LC diagnosis and is the best model based on the accuracy value when tested with test data. This research was highly accurate for all models tested with a MAPE < 10%. The GRU 30-50 model is the best-tested model with MAE = 867.2598, RMSE = 1330.427, and MAPE = 1.95%.

ACKNOWLEDGMENT

The authors are grateful to Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Jenderal Achmad Yani (Unjani) for providing an internal research grant for 2024.

REFERENCES

[1] N. N. AlMadany, O. Hujran, G. Al Naymat, and A. Maghyereh, “FORECASTING CRYPTOCURRENCY RETURNS USING CLASSICAL STATISTICAL AND DEEP LEARNING TECHNIQUES,” *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, p. 100251, 2024, doi: https://doi.org/10.1016/j.jjimei.2024.100251.

[2] D. Chuen, L. Guo, and Y. Wang, “CRYPTOCURRENCY: A NEW INVESTMENT OPPORTUNITY?,” *J. Altern. Investments*, vol. 20, pp. 16–40, Dec. 2017, doi: https://doi.org/10.3905/jai.2018.20.3.016.

[3] Y. Wei, Y. Wang, B. M. Lucey, and S. A. Vigne, “CRYPTOCURRENCY UNCERTAINTY AND VOLATILITY FORECASTING OF PRECIOUS METAL FUTURES MARKETS,” *J. Commod. Mark.*, vol. 29, p. 100305, 2023, doi: https://doi.org/10.1016/j.jcomm.2022.100305.

[4] Z. Umar, N. Trabelsi, and F. Alqahtani, “CONNECTEDNESS BETWEEN CRYPTOCURRENCY AND TECHNOLOGY SECTORS: INTERNATIONAL EVIDENCE,” *Int. Rev. Econ. Financ.*, vol. 71, pp. 910–922, 2021, doi: https://doi.org/10.1016/j.iref.2020.10.021.

[5] S. Corbet, B. Lucey, A. Urquhart, and L. Yarovaya, “CRYPTOCURRENCIES AS A FINANCIAL ASSET: A SYSTEMATIC ANALYSIS,” *Int. Rev. Financ. Anal.*, vol. 62, pp. 182–199, 2019, doi: https://doi.org/10.1016/j.irfa.2018.09.003.

[6] F. Iqbal, M. Zahid, and D. Koutmos, “CRYPTOCURRENCY TRADING AND DOWNSIDE RISK,” *Risks*, vol. 11, no. 7, pp. 1–18, 2023, doi: https://doi.org/10.3390/risks11070122.

[7] L. R. Janjua, I. Gigauri, A. Wójcik-Czerniawska, and E. Pohulak-Żołędowska, “RISK MANAGEMENT IN THE AREA OF BITCOIN MARKET DEVELOPMENT: EXAMPLE FROM THE USA,” *Risks*, vol. 12, no. 4, pp. 1–21, 2024, doi: https://doi.org/10.3390/risks12040067.

[8] G. Dudek, P. Fiszeder, P. Kobus, and W. Orzeszko, “FORECASTING CRYPTOCURRENCIES VOLATILITY USING STATISTICAL AND MACHINE LEARNING METHODS: A COMPARATIVE STUDY,” *Appl. Soft Comput.*, vol. 151, p. 111132, 2024, doi: https://doi.org/10.1016/j.asoc.2023.111132.

[9] C. Jin and Y. Li, “CRYPTOCURRENCY PRICE PREDICTION USING FREQUENCY DECOMPOSITION AND DEEP LEARNING,” *Fractal Fract.*, vol. 7, no. 10, pp. 1–29, 2023, doi: https://doi.org/10.3390/fractalfract7100708.

[10] Melina, Sukono, H. Napitupulu, and N. Mohamed, “A CONCEPTUAL MODEL OF INVESTMENT-RISK PREDICTION IN THE STOCK MARKET USING EXTREME VALUE THEORY WITH MACHINE LEARNING: A SEMISYSTEMATIC LITERATURE REVIEW,” *Risks*, vol. 11, no. 3, pp. 1–24, 2023, doi: https://doi.org/10.3390/risks11030060.

[11] M. Melina, Sukono, H. Napitupulu, and N. Mohamed, “MODELING OF MACHINE LEARNING-BASED EXTREME VALUE THEORY IN STOCK INVESTMENT RISK PREDICTION: A SYSTEMATIC LITERATURE REVIEW,” *Big Data*, pp. 1–20, Jan. 2024, doi: https://doi.org/10.1089/big.2023.0004.

[12] B. Amirshahi and S. Lahmiri, “HYBRID DEEP LEARNING AND GARCH-FAMILY MODELS FOR FORECASTING VOLATILITY OF CRYPTOCURRENCIES,” *Mach. Learn. with Appl.*, vol. 12, p. 100465, 2023, doi: https://doi.org/10.1016/j.mlwa.2023.100465.

[13] M. Melina *et al.*, “COMPARATIVE ANALYSIS OF TIME SERIES FORECASTING MODELS USING ARIMA AND NEURAL NETWORK AUTOREGRESSION METHODS,” *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 4, pp. 2563–2576, Oct. 2024, doi: 1 https://doi.org/10.30598/barekengvol18iss4pp2563-2576.

[14] P. L. Seabe, C. R. Moutsinga, and E. Pindza, “FORECASTING CRYPTOCURRENCY PRICES USING LSTM, GRU,

AND BI-DIRECTIONAL LSTM: A DEEP LEARNING APPROACH," *Fractal Fract.*, vol. 7, no. 2, pp. 1–18, 2023, doi: https://doi.org/10.3390/fractalfract7020203.

[15] S. Hochreiter and J. Schmidhuber, "LONG SHORT-TERM MEMORY," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: https://doi.org/10.1162/neco.1997.9.8.1735.

[16] J. A. Adisa, S. Ojo, P. A. Owolawi, A. Pretorius, and S. O. Ojo, "THE EFFECT OF IMBALANCED DATA AND PARAMETER SELECTION VIA GENETIC ALGORITHM LONG SHORT-TERM MEMORY (LSTM) FOR FINANCIAL DISTRESS PREDICTION," *IAENG Int. J. Appl. Math.*, vol. 53, no. 3, pp. 796–809, 2023.

[17] I. Siregar, A. Nugraha, K. Notodiputro, Y. Angraini, and L. Mualifah, "THE COMPARISON OF LONG SHORT-TERM MEMORY AND BIDIRECTIONAL LONG SHORT-TERM MEMORY FOR FORECASTING COAL PRICE," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 19, no. 1 SE-Articles, Jan. 2025, doi: https://doi.org/10.30598/barekengvol19iss1pp245-258.

[18] K. Cho *et al.*, "LEARNING PHRASE REPRESENTATIONS USING RNN ENCODER-DECODER FOR STATISTICAL MACHINE TRANSLATION," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing ({EMNLP})*, A. Moschitti, B. Pang, and W. Daelemans, Eds., Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. doi: https://doi.org/10.3115/v1/D14-1179.

[19] M. F. Rizkilloh and S. Widiyanesti, "PREDIKSI HARGA CRYPTOCURRENCY MENGGUNAKAN ALGORITMA LONG SHORT TERM MEMORY (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 25–31, Feb. 2022, doi: https://doi.org/10.29207/resti.v6i1.3630.

[20] M. A. Maliki, I. Cholissodin, and N. Yudistira, "PREDIKSI PERGERAKAN HARGA CRYPTOCURRENCY BITCOIN TERHADAP MATA UANG RUPIAH MENGGUNAKAN ALGORITME LSTM," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 7, pp. 3259–3268, Aug. 2022.

[21] Y. Wang, G. Andreeva, and B. Martin-Barragan, "MACHINE LEARNING APPROACHES TO FORECASTING CRYPTOCURRENCY VOLATILITY: CONSIDERING INTERNAL AND EXTERNAL DETERMINANTS," *Int. Rev. Financ. Anal.*, vol. 90, p. 102914, 2023, doi: https://doi.org/10.1016/j.irfa.2023.102914.

[22] A. Bouteska, M. Z. Abedin, P. Hajek, and K. Yuan, "CRYPTOCURRENCY PRICE FORECASTING – A COMPARATIVE ANALYSIS OF ENSEMBLE LEARNING AND DEEP LEARNING METHODS," *Int. Rev. Financ. Anal.*, vol. 92, p. 103055, 2024, doi: https://doi.org/10.1016/j.irfa.2023.103055.

[23] K. E. ArunKumar, D. V Kalaga, C. Mohan Sai Kumar, M. Kawaji, and T. M. Brenza, "COMPARATIVE ANALYSIS OF GATED RECURRENT UNITS (GRU), LONG SHORT-TERM MEMORY (LSTM) CELLS, AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA), SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) FOR FORECASTING COVID-19 TRENDS," *Alexandria Eng. J.*, vol. 61, no. 10, pp. 7585–7603, 2022, doi: https://doi.org/10.1016/j.aej.2022.01.011.

[24] A. Viéitez, M. Santos, and R. Naranjo, "MACHINE LEARNING ETHEREUM CRYPTOCURRENCY PREDICTION AND KNOWLEDGE-BASED INVESTMENT STRATEGIES," *Knowledge-Based Syst.*, vol. 299, p. 112088, 2024, doi: https://doi.org/10.1016/j.knosys.2024.112088.

[25] P. Jaquart, S. Köpke, and C. Weinhardt, "MACHINE LEARNING FOR CRYPTOCURRENCY MARKET PREDICTION AND TRADING," *J. Financ. Data Sci.*, vol. 8, pp. 331–352, 2022, doi: https://doi.org/10.1016/j.jfds.2022.12.001.

[26] Sukono *et al.*, "MODELING OF MEAN-VALUE-AT-RISK INVESTMENT PORTFOLIO OPTIMIZATION CONSIDERING LIABILITIES AND RISK-FREE ASSETS," *Computation*, vol. 12, no. 6, pp. 1–18, 2024, doi: https://doi.org/10.3390/computation12060120.

[27] J. H. Lee and J. K. Hong, "COMPARATIVE PERFORMANCE ANALYSIS OF VIBRATION PREDICTION USING RNN TECHNIQUES," *Electron.*, vol. 11, no. 21, p. 3619, 2022, doi: https://doi.org/10.3390/electronics11213619.

[28] Y. Su and C.-C. J. Kuo, "ON EXTENDED LONG SHORT-TERM MEMORY AND DEPENDENT BIDIRECTIONAL RECURRENT NEURAL NETWORK," *Neurocomputing*, vol. 356, pp. 151–161, 2019, doi: https://doi.org/10.1016/j.neucom.2019.04.044.

[29] M. Melina, Sukono, H. Napitupulu, and N. Mohamed, "INVESTMENT RISK FORECASTING MODEL USING EXTREME VALUE THEORY APPROACH COMBINED WITH MACHINE LEARNING," *AIMS Math.*, vol. 9, no. 11, pp. 33314–33352, 2024, doi: https://doi.org/10.3934/math.20241590.

[30] A. Rahmadeyan and Mustakim, "LONG SHORT-TERM MEMORY AND GATED RECURRENT UNIT FOR STOCK PRICE PREDICTION," *Procedia Comput. Sci.*, vol. 234, pp. 204–212, 2024, doi: https://doi.org/10.1016/j.procs.2024.02.167.

[31] M. S. Alam, M. Murshed, P. Manigandan, D. Pachiyappan, and S. Z. Abduvaxitovna, "FORECASTING OIL, COAL, AND NATURAL GAS PRICES IN THE PRE-AND POST-COVID SCENARIOS: CONTEXTUAL EVIDENCE FROM INDIA USING TIME SERIES FORECASTING TOOLS," *Resour. Policy*, vol. 81, p. 103342, 2023, doi: https://doi.org/10.1016/j.resourpol.2023.103342.

[32] H. Ihsan, I. Irwan, and A. Nensi, "IMPLEMENTATION OF BACKPROPAGATION AND HYBRID ARIMA-NN METHODS IN PREDICTING ACCURACY LEVELS OF RAINFALL IN MAKASSAR CITY," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 4, pp. 2435–2448, Oct. 2024, doi: https://doi.org/10.30598/barekengvol18iss4pp2435-2448.

[33] R. Wasono, Y. Fitri, and M. Haris, "FORECASTING THE NUMBER OF AIRPLANE PASSENGERS USING HOLT WINTER'S EXPONENTIAL SMOOTHING METHOD AND EXTREME LEARNING MACHINE METHOD," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 1, pp. 427–436, 2024, doi: https://doi.org/10.30598/barekengvol18iss1pp0427-0436.

[34] C. C. Aggarwal, *NEURAL NETWORKS AND DEEP LEARNING*. Cham: Springer International Publishing, 2018.

[35] Z. Zhan and S.-K. Kim, "VERSATILE TIME-WINDOW SLIDING MACHINE LEARNING TECHNIQUES FOR STOCK MARKET FORECASTING," *Artif. Intell. Rev.*, vol. 57, no. 8, p. 209, 2024, doi: https://doi.org/10.1007/s10462-024-10851-x.