BAREKENG: Journal of Mathematics and Its Applications
March 2024 Volume 18 Issue 1 Page 0095–0104
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss1pp0095-0104

# IMPLEMENTATION OF THE BIDIRECTIONAL GATED RECURRENT UNIT ALGORITHM ON CONSUMER PRICE INDEX DATA IN INDONESIA

**Andjani Ayu Cahaya Tanjung[1], Dewi Retno Sari Saputro[2], Nughthoh Arfawi Kurdhi[3*]**

[1,2,3]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret*
*St. Ir Sutami No 36A, Jebres, Surakarta 57126, Indonesia*

*Corresponding author's e-mail: *arfa@mipa.uns.ac.id*

***Article History:***

*Received: 17th July 2023*
*Revised: 16th November 2023*
*Accepted: 17th December 2023*

***Keywords:***

*Bidirectional Gated Recurrent Unit (BiGRU);*
*Consumer Price Index (CPI);*
*Mean Absolute Percentage Error (MAPE).*

*ABSTRACT*

*The Consumer Price Index (CPI) is the main index in measuring the inflation rate. Changes in the CPI from time to time reflect inflation and deflation; namely, the higher the CPI value, the higher the inflation rate. This study aims to apply the Birectional Gated Recurrent Unit (BiGRU) model to the CPI data in Indonesia. BiGRU comprises two GRU layers so it captures sequences that are ignored by the GRU. The research data is in the form of CPI data in Indonesia from January 2006 to December 2022 sourced from the website of the Central Bureau of Statistics, totaling 204 data. The data is divided into training data and testing data. Training data was taken from January 2006 to July 2019, with as much as 163 data. Data testing was taken from August 2019 to December 2022, with as many as 41 data. Before the data is processed, a sliding window process is carried out by dividing the data into segments to reduce the error value. The window size value used is 10. In the sliding window process, the number of segments is 194 data segments. Based on the experiment results, it was concluded that the application of BiGRU to the CPI data was carried out in an experiment with 20 BiGRU architectures. BiGRU architecture was obtained, which produced the lowest MAPE value, namely an architecture with two BiGRU layers having 256 neurons, 400 units, and one dense layer. In addition, the epochs used are 200 epochs, the ReLU activation function, and Adam optimization. The experimental results of the BiGRU architecture obtained a MAPE value of 0.24%, which indicates that the architectural performance is very good.*



***How to cite this article:***

A. A. C. Tanjung, D. R. S. Saputro and N. A. Kurdhi., "IMPLEMENTATION OF THE BIDIRECTIONAL GATED RECURRENT UNIT ALGORITHM ON CONSUMER PRICE INDEX DATA IN INDONESIA," *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0095-0104, March, 2024.


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

The price index is a barometer of general economic conditions. With the price index, the government can manage the data obtained so that it can know the development of the business, for example, to determine the policy of raising or lowering prices. A system of continuous price increases is referred to as inflation. One of the price indices used to measure the inflation rate is the Consumer Price Index (CPI) [1].

The CPI is the main indicator of the level of continuous price increases in Indonesia. The CPI comes from the value of consumption of each commodity by the community, which results in a cost of living index. The CPI is calculated and informed to the public every month by the Badan Pusat Statistik (BPS). Information about the CPI is used to review price changes over time, which illustrates the level of increase (inflation) or the level of decrease (deflation) in the price of goods and services. The higher the CPI value, the higher the inflation rate. The increase and decrease in prices of goods and services can affect the value of the currency. Historically, the inflation rate in Indonesia is relatively higher compared to other developing countries. Where in the period 2005 - 2014, Indonesia's inflation rate averaged 8.5%, while in other developing countries, the average was around 3%-5%.

CPI data is included in time series data, where the data depends on the time period. The need for goods and services in certain seasons, such as during Eid, causes a surge in prices. This situation affects changes in CPI values and forms seasonal patterns in certain time periods. Time-dependent changes in the CPI cause the CPI level to fluctuate so that the CPI value is uncertain. Therefore, a method is needed to obtain a good model of CPI data.

Research related to CPI data has been conducted by [2] with the ARMA model. The results of this study state that the ARMA model has a relatively small error value in long-term forecasting results. However, the ARMA model is only suitable for stable time series data. Still, according to [2], they suggested using other models on CPI data to stabilize the series.

According to [3] conducted research on forecasting using BiGRU on tropical cyclone wave height. The results showed that the BiGRU accuracy is 0.78, indicating that it not only gets accurate prediction results in short-term predictions but can also maintain stable accuracy in long-term predictions. Furthermore, [4] also conducted research using BiGRU for the prediction of remaining useful life on mechanical devices and concluded that BiGRU could improve prediction accuracy. In 2022, [5] predicted the host load using Discrete Wavelet Transformation and BiGRU. The results of his research state that BiGRU significantly reduces MAPE, MAE, and RMSE prediction errors because BiGRU combines forward and backward units with the final results of MAPE 3.96, RMSE 0.027, and MAE 0.022. Therefore, this research applied the BiGRU algorithm to CPI data in Indonesia.

## 2. RESEARCH METHODS

In this research, secondary data from the Badan Pusat Statistik website is used, namely Consumer Price Index (CPI) data in Indonesia from January 2006 - December 2022 [6]. The following research stages were carried out including data pre-processing, data splitting, modeling, and model evaluation. An illustration of the research stages is in **Figure 1**.

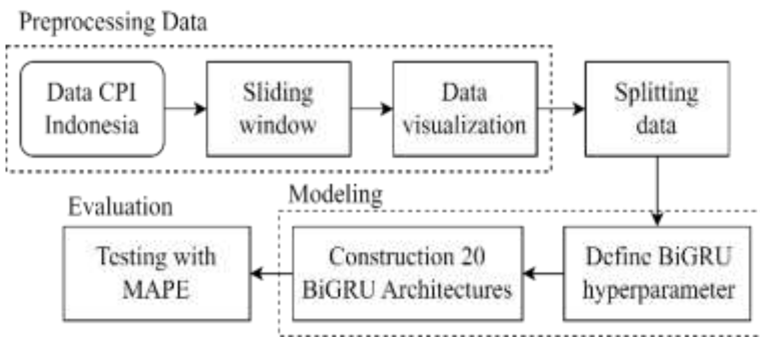


**Figure 1. The Research Method**

### 2.1 Pre-processing Data

In the pre-processing process, CPI data in Indonesia is used which is sourced from the BPS website. The CPI data taken is monthly data from January 2006 - December 2022. The amount of data used is 204 data. The scatter plot of CPI data is shown in **Figure 2**.

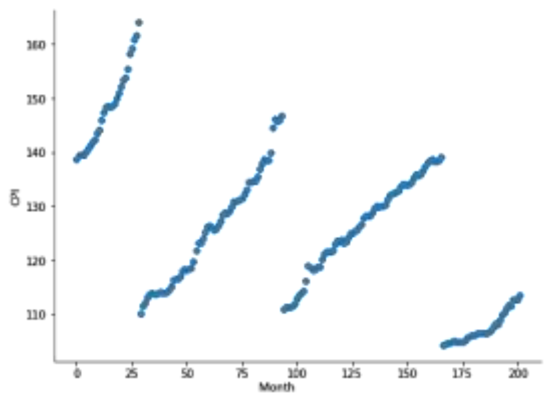


**Figure 2. CPI data plot**

Based on **Figure 2**, the CPI data for January 2006 - December 2022 has fluctuated. In June 2008, January 2014, and January 2020, the CPI value decreased drastically. Furthermore, a sliding window is performed by determining the window size value. A sliding window is part of the data structure at a certain point in time, and window size is data that is generated with a certain size by producing a dataset in the form of a set. Sliding windows are used to calculate past statistics **[7]**.

In this research, 10 window sizes are used so that a dataset is generated. The sliding window process starts from the first data by taking ten data in one segment. Then, the segment is shifted to the right by one data so that the next segment contains the second data until the eleventh data. This process continues until the last segment, and the resulting number of segments is 194 in the CPI data used.

### 2.2 Splitting Data

In the data splitting process, the data is divided into training data and testing data. Training data is the data used to train the algorithm in finding the appropriate model, while test data is the data used to test and determine the performance of the model obtained at the testing stage. Splitting data helps find the most efficient set of model parameters that have the right balance between the generalization ability of the model and its complexity **[8]**. In this research, the CPI data resulting from pre-processing was divided into 80% training data, totaling 163 data, and 20% test data, totaling 31 data.

### 2.3 Gated Recurrent Unit

The Gated Recurrent Unit (GRU) was proposed by [9] and [10] in 2014 on statistical translation engines. GRU is a development architecture of Recurrent Neural Network (RNN). As one of the developments of RNN, GRU has the ability to overcome vanishing gradient and gradient exploding problems [11]. A vanishing gradient occurs when the calculated gradient becomes very small because the GRU structure is so complex that the gradient must go through a long sequence. Meanwhile, gradient exploding occurs when the gradient becomes increasingly large which causes the training process to be longer. GRU has fewer weights and parameters, so the time required in the training process will be faster and produce a small error value. GRU has good performance on sequential problems, including natural language processing, image classification, and time series prediction [12]. GRU has two gates, namely the update gate and the reset gate. The update gate is used to determine how much past information should remain stored and remember new information. While the reset gate is used to combine new input and past information and determine whether the new information should be forgotten or not [13]. The gate structure in GRU is shown in Figure 3.

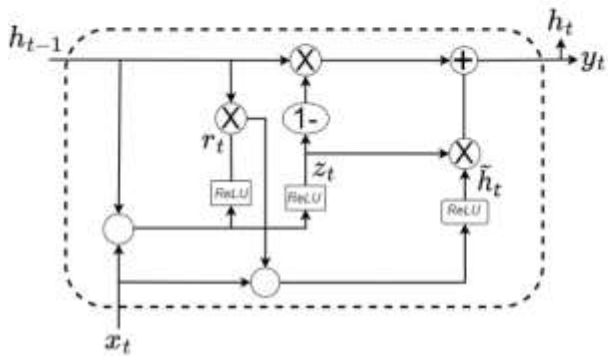


**Figure 3. The Gate Structure of GRU**

Based on Figure 3, the process in the gate GRU structure starts from combining the hidden state at the previous time ( $h_{t-1}$) and the input value at time t ($x_t$) to produce the output value ($y_t$). The equations of the gate GRU structure are written as

$$
\begin{aligned}
z_t &= ReLU(\mathrm{W_z} \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}]) \\
r_t &= ReLU(\mathrm{W}_r \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}]) \\
\tilde{h}_t &= \mathrm{ReLU}\left(\mathrm{W}_{\tilde{h}} \cdot (r_t * h_{t-1}, x_t)\right) \\
\tilde{h}_t &= \mathrm{ReLU}\left(\mathrm{W}_{\tilde{h}} \cdot (r_t * h_{t-1}, x_t)\right) \\
h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \\
y_t &= ReLU(W_o \cdot h_t)
\end{aligned}
$$

where, $\mathrm{z_t}$ : update gate, $\mathrm{r_t}$ : reset gate, $\tilde{\mathrm{h}}_\mathrm{t}$ : candidate hidden state, $\mathrm{h_{t-1}}$: hidden state at the previous time, $\mathrm{h_t}$ : the current updated hidden state, $\mathrm{W_z}$ : weight matrix on update gate, $\mathrm{W_r}$ : weight matrix on reset gate, $\mathrm{W}_{\tilde{\mathrm{h}}}$ : weight matrix on candidate hidden state, $\mathrm{W_o}$ : weight matrix of input values at time t, $\mathrm{x_t}$ : input at time t, and $\mathrm{y_t}$ : output at time t.

Based on the equations of the GRU gate structure, there are equations of gate update $z_t$, gate reset $r_t$, candidate hidden state $\tilde{h}_t$, hidden state $h_t$, and output $y_t$. In equations $z_t$ and $r_t$, $[h_{t-1}, x_t]$ is the concatenation of vector $h_{t-1}$ and vector $x_t$. The vector $x_t$ has a dimension of n with the form of vector $x_t$ is $[x_1, x_2, \dots, x_n]$. The vector $h_{t-1}$ as dimension m with the form of vector $h_{t-1}$ is $[h_1, h_2, h_3, \dots, h_m]$. The composite vector $[h_{t-1}, x_t]$ has dimension $m + n$ with the form $[h_1, h_2, h_3, \dots, h_m, x_1, x_2, \dots, x_n ]$. $W_z, W_r, W_{\tilde{h}}$, and $W_o$ are the weight matrices at the update gate, reset gate, candidate hidden state and output with dimension $(u, m + n)$. The value of u depends on the number of neuron units of the GRU architecture, while m and n are the dimensions of the vectors $h_{t-1}$ and $x_t$. The symbol $\cdot$ represents the matrix multiplication operation. The * symbol represents the element-wise multiplication operation, which is a mathematical operation between two vectors or matrices where each element at the corresponding position is multiplied by each other.

### 2.4 Bidirectional Gated Recurrent Unit

Bidirectional Gated Recurrent Unit (BiGRU) is a development of GRU. BiGRU consists of two layers of GRU [14]. The performance possessed by BiGRU is higher than GRU because BiGRU has units with forward propagation and backward propagation so that it can access previous information and subsequent information. BiGRU can capture sequences of information that may be ignored by GRU and can overcome gradient problems, one of which is vanishing gradients. The gate structure of BiGRU is shown in Figure 4.

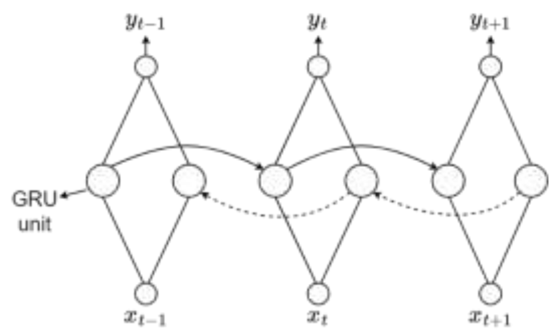


**Figure 4. The Gate Structure of BiGRU**

Based on Figure 4, the process in the BiGRU gate structure starts by entering information at the previous time and then proceeds with the forward and backward pass processes. In the forward pass, information is passed from the initial layer to the final layer, while in the backward pass, information is passed from the final layer to the initial layer. After that, the outputs generated at each layer are combined to produce the final output. In Figure 4, $x_{t-1}$ is the input at the previous time, $x_t$ is the input at time t, $x_{t+1}$ is the input at the next time, $y_{t+1}$ is the out-put at the previous time, $y_t$ is the output at time t, and $y_{t+1}$ is the output at the next time. The output at the current state is obtained from the two-position hidden state at the previous time and the next time. The gate structure equations in BiGRU are written as

$$\overrightarrow{h_t} = GRU(x_t, \overrightarrow{h_{t-1}})$$
$$\overleftarrow{h_t} = GRU(x_t, \overleftarrow{h_{t-1}})$$
$$h_t = \alpha \overrightarrow{h_t} + \beta \overleftarrow{h_t} + b$$
$$y_t = ReLU\ (W_b\ h_t)$$

where $\overrightarrow{h_t}$ : hidden layer for the positive direction at time t, $\overleftarrow{h_t}$ : hidden layer for the negative direction at time t, $h_t$ : hidden state at time t, $\alpha$ : weight for positive direction, $\beta$ : weight for negative direction, $b$ : bias, and $W_b$ : weight matrix.

### 2.5 Evaluation

In the evaluation process, Mean Absolute Percentage Error (MAPE) is used. According to [15], MAPE is a measurement tool that shows significant prediction errors in percentage form. The smaller the MAPE value, the smaller the estimated error value. The MAPE formula is as follows.

$$MAPE = \frac{\sum_{t=1}^{n} \frac{|x_t - y_t|}{x_t}}{n} \times 100\%$$

where $x_t$: actual data, $y_t$: prediction data, and $n$: the number of prediction data.

The range of MAPE values that can be used as a measurement material regarding the ability of a prediction model is shown in **Table 1 [16]**.

**Table 1. Range of MAPE Value**

| Range MAPE | Interpretation |
|---|---|
| < 10% | Highly accurate forecasting |
| 10 – 20% | Good forecasting |
| 20 – 50% | Reasonable forecasting |
| >50% | Inaccurate forecasting |

## 3. RESULTS AND DISCUSSION

In this study, 20 BiGRU architectures were built with different number of layers and number of BiGRU units. Before building the architecture, BiGRU hyperparameters that are influential in obtaining good performance are determined. The hyperparameters used include the number of layers and BiGRU units, ReLU activation function, and 200 epochs. In addition, each architecture uses one dense layer and Adam optimization. Dense serves to add a fully connected layer. Adam's optimization serves to optimize the value of each layer so that the results are more accurate. The loss function used is Mean Absolute Percentage Error (MAPE) to measure how well the architecture performs. The MAPE results of the 20 BiGRU architectures are shown in **Table 1**. The MAPE results of the 20 BiGRU architectures are shown in **Table 2**.

**Table 2. The Experimental Result of 20 Architectures BiGRU**

| No | BiGRU Architecture | MAPE |
|---|---|---|
| 1 | BiGRU 32 unit neuron | 0.432% |
| 2 | BiGRU 64 unit neuron | 0.307% |
| 3 | BiGRU 80 unit neuron | 0.321% |
| 4 | BiGRU 100 unit neuron | 0.292% |
| 5 | BiGRU 120 unit neuron | 0.294% |
| 6 | BiGRU 128 unit neuron | 0.350% |
| 7 | BiGRU 150 unit neuron | 0.315% |
| 8 | BiGRU 200 unit neuron | 0.284% |
| 9 | BiGRU 300 unit neuron | 0.281% |
| 10 | BiGRU 400 unit neuron | 0.260% |
| 11 | BiGRU 64, 100 unit neuron | 0.275% |
| 12 | BiGRU 100, 128 unit neuron | 0.298% |
| 13 | BiGRU 100, 200 unit neuron | 0.280% |
| 14 | BiGRU 128, 100 unit neuron | 0.276% |
| 15 | BiGRU 128, 120 unit neuron | 0.302% |
| 16 | BiGRU 200, 300 unit neuron | 0.260% |
| 17 | BiGRU 300, 200 unit neuron | 0.288% |
| 18 | BiGRU 100, 64 unit neuron | 0.400% |
| **19** | **BiGRU 256, 400 unit neuron** | **0.24%** |
| 20 | BiGRU 400, 200 unit neuron | 0.259% |

Based on **Table 2**, 20 BiGRU architectures are shown with the MAPE value generated from each architecture. The best BiGRU architecture is selected with the architecture that produces the smallest MAPE value. In the **Table 2** number 19 shows that the architecture with two BiGRU layers totaling 256 and 400 neuron units has the smallest MAPE value of 0.24%. The CPI data is processed in BiGRU with 256 and 400 neuron units. The first five weight values of the dense layer in the best architecture are [-0.00209816, -0.02615988, 0.00880788, -0.01637967, 0.04520454]. The architecture is obtained at the 194th epoch with the results of the loss value on the training set is 1.3583 and the validation loss on the testing set is 0.24897. The results of loss and validation loss can be seen in **Figure 5**.

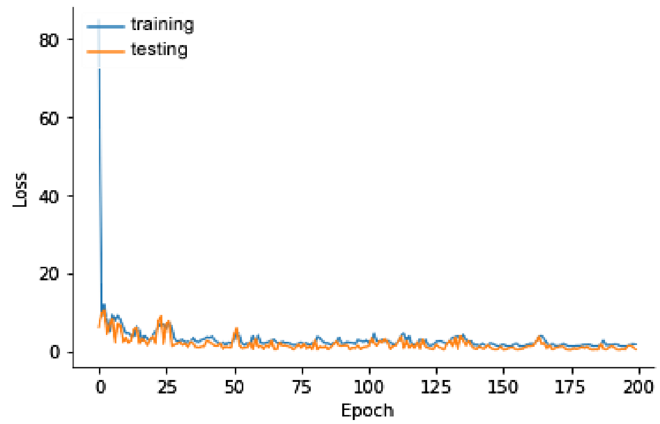


**Figure 5. The Loss and Validation Loss of BiGRU 250, BiGRU 400**

Furthermore, data prediction is carried out from June 2020 - December 2022. The comparison graph between actual and predicted data is shown in **Figure 6**.

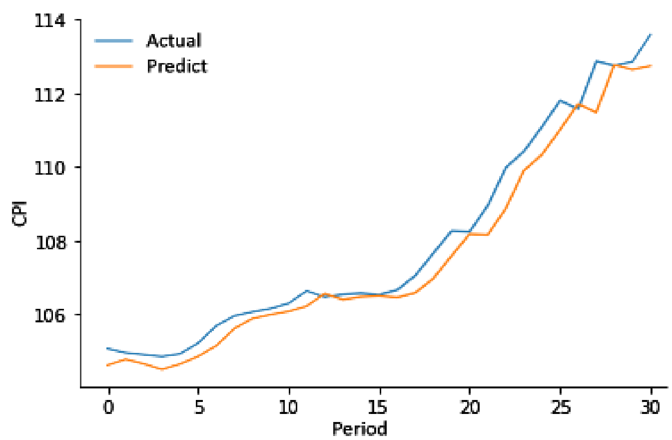


**Figure 6. Actual data and predict data**

Based on **Figure 6**, the prediction results using BiGRU are close to the actual value of the CPI data. By using the MAPE formula, the MAPE value is 0.24%. Based on the MAPE range in **Table 1**, it shows that the prediction ability with BiGRU on CPI data in Indonesia is classified as highly accurate forecasting.

The best BiGRU architecture that has been obtained is used to forecast CPI data for the next seven months, namely June 2023 - December 2023. The forecasting results obtained are also compared with the latest actual CPI data, namely January - May 2023, which has been published by the Badan Pusat Statistik. The results of data forecasting are shown in **Table 3**.

**Table 3. CPI Forecasting Results in Indonesia for the Next Seven Months**

| Month, Year | Actual CPI | Forecasting CPI | Error |
|---|---|---|---|
| January 2023 | 113.98 | 113.81 | 0.17 |
| February 2023 | 114.16 | 114.02 | 0.14 |
| March 2023 | 114.36 | 114.24 | 0.12 |
| April 2023 | 114.47 | 114.46 | 0.01 |
| May 2023 | 114.84 | 114.68 | 0.16 |
| June 2023 | - | 114.91 | - |
| August 2023 | - | 115.36 | - |
| September 2023 | - | 115.59 | - |
| October 2023 | - | 115.82 | - |
| November 2023 | - | 116.05 | - |
| December 2023 | - | 116.29 | - |

Based on Table 3, the forecasting results for January - May 2023 have a small error value. It can be concluded that the forecast has performed well in approximating the actual values. In addition, it is known that during the next seven months, namely June - December 2023, the range of CPI values in Indonesia ranges from 114.19 to 116.29, and the CPI has increased. This suggests that the economy may be experiencing growth or inflation during that period. In other words, prices of goods and services are likely to rise within that range.

## 4. CONCLUSIONS

Based on the results and discussion, it is concluded that the application of BiGRU on CPI data in Indonesia produces the best architecture. The best BiGRU architecture is obtained in a two-layer BiGRU architecture, with each layer having 256 neuron units and 400 neuron units, and one dense layer. The hyperparameters used are 200 epochs, ReLU activation function, and Adam optimization. This best architecture can be used for the forecasting process. The first five weight values of the dense layer in the best architecture are [-0.00209816, -0.02615988, 0.00880788, -0.01637967, 0.04520454]. The results of the architecture experiment obtained the smallest MAPE value of 0.24%. Based on MAPE, it is stated that the performance of the architecture is very good.

## ACKNOWLEDGMENT

Thank you to the Softcomputing Mathematics Research Group for their moral support and constructive comments in improving this paper. We also express our gratitude to the Institute for Research and Community Service (LPPM) Universitas Sebelas Maret (UNS) for the research group grant through the letter of approval for research implementation of non-APBN funds No. 228/UN27.22 /PT.01.03/2023.

## REFERENCES

[1] E. Riyanto, "Peramalan Harga Saham Menggunakan Jaringan Syaraf Tiruan Secara Supervised Learning Dengan Algoritma Backpropagation," *J. Inform. Upgris*, vol. 3, no. 2, pp. 137–142, 2017, doi: 10.26877/jiu.v3i2.1899.

[2] F. W. Zhang, W. G. Che, B. B. Xu, and J. Z. Xu, "The research of ARMA model in CPI time series," *Appl. Mech. Mater.*, vol. 347–350, pp. 3099–3103, 2013, doi: 10.4028/www.scientific.net/AMM.347-350.3099.

[3] F. Meng, T. Song, D. Xu, P. Xie, and Y. Li, "Forecasting tropical cyclones wave height using bidirectional gated recurrent unit," *Ocean Eng.*, vol. 234, no. January, p. 108795, 2021, doi: 10.1016/j.oceaneng.2021.108795.

[4] D. She and M. Jia, "A BiGRU method for remaining useful life prediction of machinery," *Meas. J. Int. Meas. Confed.*, vol. 167, no. July 2020, p. 108277, 2021, doi: 10.1016/j.measurement.2020.108277.

[5] S. Dogani, J., Farshad, K., Mehdi, "Host Load Prediction in Cloud Computing with Discrete Wavelete Transformation (DWT) and Bidirectional Gated Recurrent Unit (BiGRU) Network," vol. 198, pp. 157–174, 2022.

[6] Badan Pusat Statistik, “Indeks Harga Konsumen dan Inflasi Bulanan Indonesia, 2006-2022.” https://www.bps.go.id/statictable/2009/06/15/907/indeks-harga-konsumen-dan-inflasi-bulanan-indonesia-2006-2022.html (accessed May 29, 2022).
[7] J. Gama, R. Sebastião, and P. P. Rodrigues, “On evaluating stream learning algorithms,” *Mach. Learn.*, vol. 90, no. 3, pp. 317–346, 2013, doi: 10.1007/s10994-012-5320-9.
[8] D. E. Birba, “Study of Data Splitting Algorithms for Machine Learning,” 2020.
[9] K. Cho *et al.*, “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1724–1734, 2014, doi: 10.3115/v1/d14-1179.
[10] D. Bahdanau, K. H. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.
[11] J. Rosindell and Y. Wong, “Biodiversity, the tree of life, and science communication,” *Phylogenetic Divers. Appl. Challenges Biodivers. Sci.*, no. 2, pp. 41–71, 2018, doi: 10.1007/978-3-319-93145-6_3.
[12] J. Yuan and Y. Tian, “An intelligent fault diagnosis method using GRU neural network towards sequential data in dynamic processes,” *Processes*, vol. 7, no. 3, 2019, doi: 10.3390/pr7030152.
[13] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling,” no. December, 2014, [Online]. Available: http://arxiv.org/abs/1412.3555
[14] H. He, H. Wang, H. Ma, X. Liu, Y. Jia, and G. Gong, “Research on Short-term Power Load Forecasting Based on Bi-GRU,” *J. Phys. Conf. Ser.*, vol. 1639, no. 1, 2020, doi: 10.1088/1742-6596/1639/1/012017.
[15] A. De Myttenaere *et al.*, “Mean Absolute Percentage Error for regression models To cite this version : HAL Id : hal-01312590,” *Adv. Artif. neural networks, Mach. Learn. Comput. Intell. - Sel. Pap. from 23rd Eur. Symp. Artif. Neural Networks (ESANN 2015)*, pp. 38–48, 2017.
[16] P. C. Chang, C. H. Liu, C. H. Yeh, and S. H. Chen, “The development of a weighted evolving fuzzy neural network,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4114 LNAI, pp. 212–221, 2006, doi: 10.1007/978-3-540-37275-2_28.