# Imputation Missing Stock Prices using Generative Adversarial Networks and Attention Mechanism


Agus Tri Haryono
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
7025231020@student.its.ac.id

Riyanarto Sarno*
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
riyanarto@if.its.ac.id

Ratih Nur Esti Anggraini
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
ratih_nea@if.its.ac.id

Kelly Rossa Sungkono
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
kelly@its.ac.id



2024 2nd International Conference on Technology Innovation and Its Applications (ICTIIA) | 979-8-3503-5161-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICTIIA61827.2024.10761233



***Abstract*—Missing value in time series data can disrupt the analysis of historical data, especially stock market data, which requires the sequence of data points. This research proposes a method for imputation of missing stock prices in multivariate time series. The Generative Adversarial Networks (GAN) method, such as Generative Adversarial Imputation Networks (GAIN), and attention mechanism fill the missing value. The simulated missing pattern as dataset training is produced using the missing completely at random (MCAR) method. The GAN method learns the missing value pattern using generator and discriminator parts and attention mechanism, which are self-attention and multi-head attention. The proposed method was tested using various lookbacks, missing rates, hint rates, and evaluation methods on 100 stock issuers. The GAIN multi-head attention significantly outperforms baseline GAIN and GAIN self-attention methods.**



***Keywords*—attention mechanism, generative adversarial networks, imputation methods, missing value, and stock price.**


## I. Introduction

Time sequence in stock transactions is crucial to discovering patterns, associations, trends, sequences, and dependencies in stock forecasting using multivariate time series [1]. Time sequences in time series can be harmed by missing values produced by the data provider [2]. In the Healthcare Physionet Challenge 2012 dataset, the missing value rate achieved 80%, which decreases model performance for predicting missing values [3]. The Indonesian Stock Exchange (IDX) historical data has a 5.9% missing value, collected from Yahoo Finance from 2000 to 2024. The imputation of the missing value process is essential to reduce errors in stock forecasting [4], [5].

Missing values are separated into three types: missing not at random (MNAR), missing at random (MAR), and missing completely at random (MCAR) [6]. In MNAR, the missing value depends on the values of the missing data, regardless of the observed data. In MAR, the probability of missing values depends on the observed data. In MCAR, missing values are entirely random and do not rely on observed and unobserved data [7]. This research focuses on the MCAR type, similar to missing value patterns on stock transaction data, and uses various missing value rates to evaluate imputation methods.

Several studies have resolved missing values on time series data using imputation methods such as statistical methods [8], machine learning [9], deep learning [10], and generative models [11], [12]. The statistical method uses the mean value from the nearest neighbour or interpolates using known values but is limited to univariate data [8]. Machine learning, such as support vector machines, has advantages in computational time but has underperformed with less data training [13]. Gated Recurrent Unit (GRU) is used to impute missing values on river water quality dataset with short lookback size, low missing rate, and single variable, not yet discussed long lookback size that requires for long sequence time series forecasting [14]. Generative models such as Generative Adversarial Nets (GAN) employ generator and discriminator parts to learn missing value patterns and then generate value with similar patterns of actual data [15]. The self-attention and multi-head attention mechanism from Transformer architecture has tackled various limitations on vision, natural language processing, and time series domain. The attention mechanism can tackle limitation memory constraints and better time training than the GRU method [16].

This research proposes a new approach to imputation missing stock prices using GAN with an attention mechanism. The MCAR type used has five missing rates with two kinds of lookback. Five hint rates are used in GAN to reveal a correlation between missing and hint rates. Two attention mechanism types are added to GAN's generator and discriminator parts to improve the imputation of multivariate missing value results.

The rest of the articles are organized as follows: Section II presents related works, and Section III presents the proposed method. Section IV presents the results and discussion, followed by a conclusion in Section V.


979-8-3503-5161-3/24/$31.00 ©2024 IEEE

## II. Related Works

The GAN method already tackles the problem of imputation missing values on various datasets. The previous research analyzed missing values with missing rates of 30%, 50%, and 70% on the Stock, Energy, Beijing Air-Quality, Letter, and Spam datasets using GAN inversion with Graph Convolution Network (GCN) and decay connection. The decay connection is employed to construct the graph of missing value time series [12]. The Conditional GAN (cGAN) is developed to compute missing values on the MRI dataset. The generator part uses three missing channels of four-channel 2D MRI, and the discriminator part calculates loss on each channel with 16 x 16 patches [11].

The attention mechanism is already used to imputation missing values in time series data. The attention mechanism tackles imputation missing values on time series datasets. Transformer architecture with an attention mechanism called Self-Attention-based Imputation for Time Series (SAITS) learns missing values using two diagonally masked self-attention blocks. The first block takes the missing value matrix and masking matrix and produces the imputed matrix. The imputed matrix and the same masking matrix pass into the second block to produce final imputed data with weighted combination [16].

## III. Proposed Method

### A. Dataset

The dataset comes from our previous research, which contains the transaction date, open, low, high, close price, and transaction volume (OLHCV) dataset [17]. The dataset uses 100 stock issuers from the IDX, selected randomly with requirements that the minimum amount of data is 200, the variance is zero and does not have missing values. Table I shows the statistics of the dataset used in this research. The dataset contains 100 stock issuers and 679,225 data points across 11 sectors. Each stock issuer has a different amount of data because each stock issuer has a diverse Initial Public Offering (IPO) date.

TABLE I
Statistic of stock issuer grouped by sector.

| Sector | Issuer | Data Point | Mean | |
|---|---|---|---|---|
| | | | Close | Volume |
| Energy | 9 | 61,305 | 1,755.65 | 9,640,534.74 |
| Finance | 9 | 62,180 | 732.08 | 18,830,483.55 |
| Health | 4 | 26,175 | 1,608.89 | 2,618,599.14 |
| Infrastructure | 14 | 100,515 | 441.89 | 6,847,811.00 |
| Industry | 5 | 30,810 | 371.28 | 8,522,300.39 |
| Non-primary consumer goods | 19 | 122,340 | 801.06 | 20,864,172.31 |
| Primary consumer goods | 11 | 71,090 | 478.08 | 10,338,622.91 |
| Property | 9 | 53,955 | 880.39 | 4,001,738.82 |
| Raw goods | 9 | 83,465 | 545.03 | 41,684,933.54 |
| Technology | 5 | 31,350 | 2,627.87 | 7,259,711.71 |
| Transportation | 6 | 36,040 | 357.94 | 7,088,315.52 |
| Total | 100 | 679,225 | | |

### B. Pre-processing

This research compares the performance of GAN models using two lookback periods: 50 and 100 days. Each OLHCV dataset derived from stock issuer slicing is transformed into input format using a lookback value. The input format contains three types of data: rows of data, lookbacks, and indications. Because the GAN input only accepts two-dimensional data, three-dimensional data is converted to two-dimensional data in rows and lookback multiplication indicators. The transformation of data dimension is shown in Fig. 1.

Before the GAN training process, the GAN input uses a min-max scaler for normalization data and replaces the actual value with the missing value. The missing value is created using the MCAR method with missing rates 10%, 30%, 50%, and 70%. The process of missing value imputation requires a hint from the actual missing value. Hint rates of 90%, 80%, 70%, and 60% are from masked actual value and used to help generate new value to replace missing value. Various missing and hint rates are used to determine the correlation between the high and low rates of missing and hints.

### C. Generative Adversarial Network

Generative Adversarial Imputation Nets (GAIN) is a method for imputation of missing values using baseline GAN [18]. Shown in Fig. 1 GAIN has two parts: the Generator, denoted by $G$, and the Discriminator, denoted by $D$. The $G$ generates a new matrix $\bar{\mathbf{X}}$ using an input matrix with missing value $\tilde{\mathbf{X}}$, masking matrix $\mathbf{M}$ and random matrix $\mathbf{Z}$. The imputed matrix value $\hat{\mathbf{X}}$ is produced by element-wise multiplication $\odot$ using an $\bar{\mathbf{X}}$, $\mathbf{M}$, and $\tilde{\mathbf{X}}$. The missing value in $\tilde{\mathbf{X}}$ will be replaced by a value from $\bar{\mathbf{X}}$. The process of $G$ is shown in Equations (1) and (2).

$$\bar{\mathbf{X}} = G(\tilde{\mathbf{X}}, \mathbf{M}, (1 - \mathbf{M}) \odot \mathbf{Z}) \tag{1}$$

$$\hat{\mathbf{X}} = \mathbf{M} \odot \tilde{\mathbf{X}} + \bar{\mathbf{X}} \odot (1 - \mathbf{M}) \tag{2}$$

The Discriminator $D$ takes $\hat{\mathbf{X}}$ and hint matrix $\mathbf{H}$ to produced matrix probability of each component $\hat{\mathbf{D}}$. Training $D$ is used contrary to training $G$ similar baseline GAN. The difference $D$ in GAIN and baseline GAN is that the $D$ in GAN used to judge $\hat{\mathbf{X}}$ is completely fake or completely actual. The $D$ in GAIN attempts to distinguish which components are observed or imputed based on probability. The calculation of $D$ is shown in Equations (3).

$$\hat{\mathbf{D}} = D(\hat{\mathbf{X}}, \mathbf{H}) \tag{3}$$

The hint mechanism produced $\mathbf{H}$ based on $\mathbf{M}$ to control the $\bar{\mathbf{X}}$ in the $G$ with a distribution similar to the actual data. Hint mechanism using hint rates $\mathbf{h}$ to control information about $\mathbf{M}$ reveal to $\mathbf{H}$. Different $\mathbf{h}$ are used to discover minimum $\mathbf{h}$ without significantly lowering GAIN performance. The process of hint mechanism using Equation (4).

$$\mathbf{H} = \mathbf{M} \odot \mathbf{h} \tag{4}$$

The $G$ and $D$ has loss function to calculate probability of predicting $\mathbf{M}$. Shown in Fig. 1, the $G$ produces Mean Squared Error (MSE) loss using $\tilde{\mathbf{X}}$ and $\hat{\mathbf{X}}$, and the $D$ produces Cross Entropy (CE) loss using $\hat{\mathbf{D}}$ and $\mathbf{M}$. Training $G$ employs MSE loss and CE loss to minimize the probability, and training D

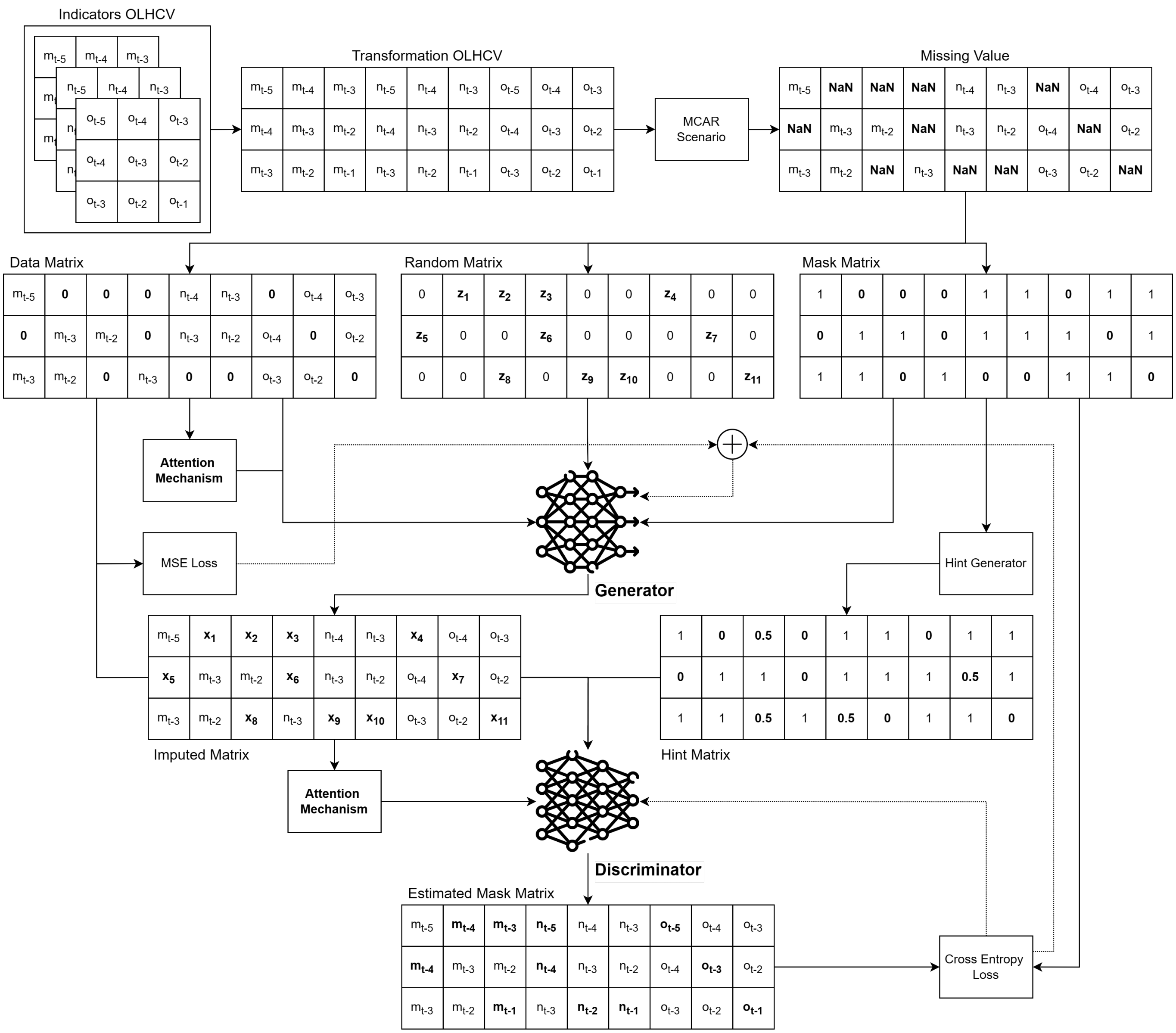


Fig. 1. The architecture of generative adversarial imputation network and attention mechanism.

employs CE loss to maximize the probability. The calculation loss function is shown in Equation (5) and (6).

$$\min_{G} \max_{D} V(D,G) \tag{5}$$

$$V(D,G) = \mathbf{M} \odot \log \hat{\mathbf{D}} + (1-\mathbf{M}) \odot \log(1-\hat{\mathbf{D}}) \tag{6}$$

## D. Self-attention

Self-attention mechanism learns the input sequence of vector $X = [x_{t-n}, ..., x_{t-2}, x_{t-1}]$ with missing value. In the Generator the $X$ is $\tilde{\mathbf{X}}$ and in the Discriminator the $X$ is $\hat{\mathbf{X}}$. The input sequence X learned linear transformation by $Q$ (Query), $K$ (Key), and $V$ (Value) with weight matrix $W_Q$, $W_K$, and $W_V$. The attention weight produced by Equation (7), with $QK^T$ is matrix multiplication between query and key, and $\sqrt{d_k}$ is a key dimension. The attention weight is summed up with the input sequence, and then the process is similar to baseline GAIN.

$$Att(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

## E. Multi-head attention

Multi-head attention employs multiple heads (h) for linear projection query, key, and value instead of using single attention. The process multi-head attention using Equation (8) and (9). Each projection query, key, and value is split into some heads, then attention weight is calculated. This research employs ten heads. The multi-head attention apply to $\tilde{\mathbf{X}}$ in

Generator and $\hat{\mathbf{X}}$ in Discriminator.

$$MultiHead(Q,K,V) = \text{Concat}(\text{h}_1,\ldots,\text{h}_n)W_O \quad (8)$$

$$\text{h}_i = Att(QW_Q^i, KW_K^i, VW_V^i) \quad (9)$$

### F. Evaluation

This research employs three evaluation metrics to ascertain the optimal method for the compared imputation missing value methods on the stock transaction. The performance evaluation relies on a comprehensive set of three evaluation metrics, which is the Root Mean Square Error ($RMSE$), to evaluate larger errors with higher weight, suitable for financial models [19]. The Mean Absolute Error ($MAE$) is employed to give a straightforward indication of average error magnitude without direction consideration and outliers sensitivity, respectively [20]. The Mean Relative Error $MRE$ is used to compare the error proportionally across different scales [16]. The $RMSE$, $MAE$, and $MRE$ calculation shown sequentially in (10) to (12).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (10)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (11)$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{|y_i|} \quad (12)$$

The actual values are denoted as $y_i$. The predicted values are represented as $\hat{y}_i$. These values are used to produce performance value on each metric. These metrics collectively provide a robust and comprehensive means of assessing each stock issuer's imputation method performance and suitability.

### G. Experiment Setup

This research uses the architecture shown in Fig. 1 for comparison performance imputation missing value methods. Each stock issuer passes into the preprocessing section to dimension transformation and produces a missing value dataset with a lookback of 50 and 100, missing rates of 10%, 30%, 50%, and 70%, and hint rates of 90%, 80%, 70%, and 60%. Baseline GAIN using three dense layers with a gelu activation function. GAIN self-attention uses one attention layer, one normalization layer, and five dense layers with a gelu activation function. GAIN multi-head attention uses one multi-head attention layer, one normalization layer, and five dense layers with a gelu activation function. The training process uses 1000 epochs, 128 batch sizes, and validation every 20 epochs. Each imputation method is examined using three evaluation methods, and the evaluation process only uses predicted imputation value with actual value.

## IV. Result and Discussion

### A. Pre-processing

The preprocessing produces a missing value dataset with various missing rates. Statistics of preprocessing results are shown in Table II. The maximum consecutive missing value increases as the missing rate increases. A lookback of 50 and a missing rate of 70% produces a minimum consecutive missing value of 1, a maximum consecutive missing value of 36, and a missing value of 91,811,820 on 130,945,000 data points. A lookback of 100 and a missing rate of 70% produces a minimum consecutive missing value of 1, a maximum consecutive missing value of 37, and a missing value of 176,454,200 on 251,890,000 data points.

TABLE II
STATISTICS OF MISSING VALUE DATASET AFTER PREPROCESSING.

| Lookback | Missing Rate | Consecutive Min | Consecutive Max | Missing Value | Data Point |
|---|---|---|---|---|---|
| 50 | 10% | 1 | 5 | 13,014,956 | 130,945,000 |
| | 30% | 1 | 12 | 39,263,220 | 130,945,000 |
| | 50% | 1 | 22 | 65,503,164 | 130,945,000 |
| | 70% | 1 | 36 | 91,811,820 | 130,945,000 |
| 100 | 10% | 1 | 5 | 25,207,168 | 251,890,000 |
| | 30% | 1 | 14 | 75,591,820 | 251,890,000 |
| | 50% | 1 | 23 | 125,925,432 | 251,890,000 |
| | 70% | 1 | 37 | 176,454,200 | 251,890,000 |

TABLE III
COMPARISON OF IMPUTATION MISSING VALUE METHODS PERFORMANCES.

| Method | Lookback | RMSE | MAE | MRE |
|---|---|---|---|---|
| GAIN | 50 | **0.1013** | 0.0658 | 0.2866 |
| | 100 | 0.3467 | 0.1670 | 0.7536 |
| GAIN self-attention | 50 | 0.1464 | 0.0943 | 0.4004 |
| | 100 | 0.1494 | 0.0954 | 0.3984 |
| GAIN multi-head attention | 50 | 0.1047 | **0.0654** | **0.2779** |
| | 100 | **0.1001** | **0.0637** | **0.2679** |

### B. Comparison of Imputation Methods Performance

This research compares the baseline method GAIN with a proposed method, which is GAIN with an attention mechanism. The attention mechanism uses self-attention and multi-head attention. Three evaluation methods are used for comprehensive comparison performance. Each method's performance is the average from 100 stock issuers, five missing rates, and five hint rates.

As shown in Table III, GAIN multi-head attention outperformed the method of the others for imputation missing value. The GAIN multi-head attention method has evaluation performance RMSE, MAE, and MRE sequentially at 0.1047, 0.0654, and 0.2779 using a lookback of 50. The GAIN method with a lookback of 50 was better than the other methods when using RMSE evaluation, with a performance of 0.1013. The GAIN multi-head attention method performed best on three evaluations using a lookback of 100; the performance value is RMSE of 0.100, MAE of 0.0637, and MRE of 0.2679.

The attention mechanism influences better performance on imputation missing value, especially using a lookback of 100. The GAIN self-attention method performance increased by 42.87% than The GAIN method using MAE evaluation. The GAIN multi-head attention method performance increased by 61.86% than The GAIN method using MAE evaluation. The

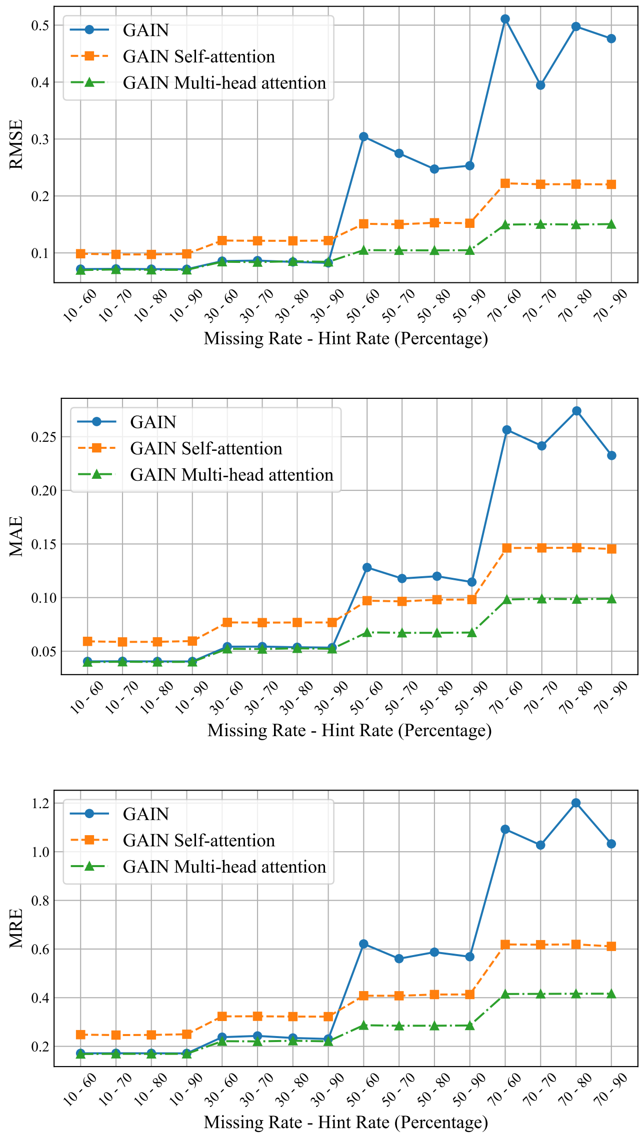


Fig. 2. Comparison of performance based on missing and hint rate using RMSE, MAE, and MRE.

attention mechanism does not significantly improve performance imputation missing value when using a lookback of 50, especially on the GAIN self-attention method, which is underperformed than the GAIN method.

### C. Correlation Missing Rates and Hint Rates

Further investigation was conducted to understand why the GAIN attention mechanism outperformed baseline GAIN. Fig. 2 shows a performance comparison of imputation missing value methods based on missing and hint rates using RMSE, MAE, and MRE evaluations. The baseline GAIN using the hint mechanism competes with GAIN multi-head attention until the missing rate of 30%. The baseline GAIN has high errors when using higher missing rates of 50% to 70%, and the hint mechanism is insufficient to increase performance. The attention mechanism performs excellently on 50% and 70% missing rates, with multi-head attention having the best performance. The collaboration of the hint and attention mechanism produces the best performance, with the higher hint rates increasing the performance.

Table IV shows stock issuer's performance's top five and bottom five. The performance was sorted ascending by RMSE, MAE, and MRE evaluations. The BELL, PZZA, TCPI, WOOD, and MARK stocks were selected as top and bottom

TABLE IV
TOP 5 AND BOTTOM 5 GAIN MULTI-HEAD ATTENTION SORTED BY RMSE, MAE, AND MRE.

| Issuer | Lookback | Rate (%) | | Evaluation | | |
|---|---|---|---|---|---|---|
| | | Missing | Hint | RMSE | MAE | MRE |
| BELL | 50 | 10 | 90 | 0.0379 | 0.0203 | 0.1752 |
| BELL | 50 | 10 | 60 | 0.0383 | 0.0207 | 0.1788 |
| BELL | 50 | 10 | 80 | 0.0387 | 0.0208 | 0.1800 |
| BELL | 50 | 10 | 70 | 0.0393 | 0.0209 | 0.1803 |
| BELL | 100 | 10 | 60 | 0.0440 | 0.0212 | 0.1815 |
| PZZA | 50 | 70 | 70 | 0.2227 | 0.1589 | 0.4593 |
| TCPI | 100 | 70 | 90 | 0.2272 | 0.1768 | 0.3768 |
| TCPI | 50 | 70 | 80 | 0.2437 | 0.1900 | 0.3949 |
| WOOD | 50 | 70 | 90 | 0.2651 | 0.1968 | 0.5202 |
| MARK | 50 | 70 | 70 | 0.2753 | 0.2056 | 0.5795 |

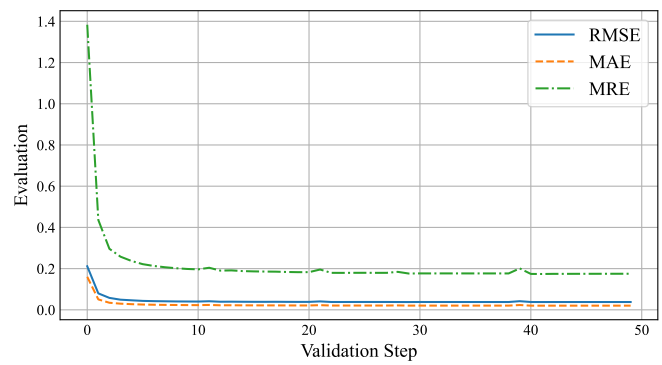


Fig. 3. BELL issuer validation result on training GAIN multi-head attention.

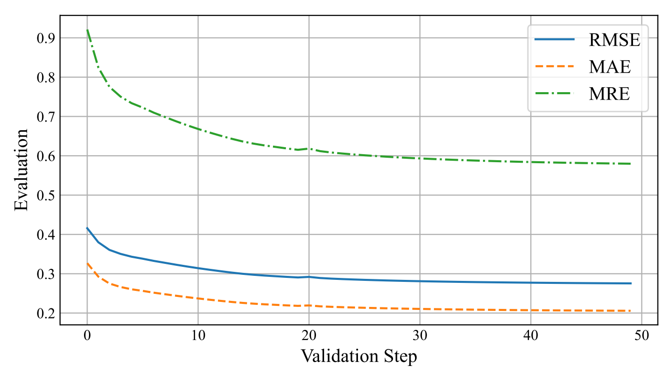


Fig. 4. MARK issuer validation result on training GAIN multi-head attention.

performance to reveal influence parameters on imputation performance. The top five were achieved by BELL issuers, with the top four using a lookback of 50 and a missing rate of 10%. Increasing the hint rate does not have a linear correlation with improving performance, but the best performance is when using a hint rate of 90% and a missing rate of 10%. The bottom five are achieved by various stock issuers: PZZA, TCPI, WOOD, and MARK. The most lookback use is 50, and the missing rate is 70%. The hint mechanism is not enough to increase the performance of the imputation missing value. The missing rate 70% on the dataset is not recommended for stock price forecasting because it produced the worst performance on imputation missing value. As confirmed using Fig. 2, the highest missing rate of a dataset that can be used on stock forecasting is 50% because RMSE and MAE evaluation oscillates on 0.1.

The convergence of performance on the imputation missing value training model is shown in Fig. 3 and Fig. 4. The BELL issuer in Fig. 3 uses a lookback of 50, a missing rate of 10%, a hint rate of 90%, and GAIN multi-head attention. The BELL training has performance convergence on the validation step of 20, which is 200 epochs. The MARK issuer in Fig. 4 uses a lookback of 50, a missing rate of 70%, a hint rate of 70%, and GAIN multi-head attention. The MARK training has performance convergence on the validation step of 30, which is 600 epochs. The GAIN multi-head attention method training needs epochs from 200 to 600.

## V. Conclusion

This research proposes a novel method for imputing missing stock prices in multivariate time series using GAN and multi-head attention. The GAN learns missing values produced by the MCAR method in the training and learns the pattern of non-missing values by hint mechanism and multi-head attention. The generator part from GAN produced a predicted value to replace the missing value. The GAIN multi-head attention method outperformed the baseline GAIN and GAIN self-attention. The proposed method for imputation of missing stock price can be expanded further in the following research because each indicator's positional information disappears when the input is transformed from 3-dimensional to 2-dimensional. The following research can also employ multi-head attention to generate a random matrix.

## Acknowledgment

This work was funded by Institut Teknologi Sepuluh Nopember (ITS) under Penelitian Dana Departemen Program, Penelitian Keilmuan, Penelitian Flagship and under the Project Scheme of the Publication Writing and Intellectual Property Rights (IPR) Incentive (Penulisan Publikasi dan Hak Kekayaan Intelektual/PPHKI) Program 2024; and the Directorate General of Higher Education, Research and Technology of the Republic of Indonesia under Penelitian Terapan Program.

## References

[1] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators," *IEEE Access*, vol. 11, pp. 77 132–77 146, 2023, doi: 10.1109/ACCESS.2023.3298445.

[2] M. Alabadla, F. Sidi, I. Ishak, H. Ibrahim, L. S. Affendey, Z. C. Ani, M. Jabar, U. Bukar, N. Devaraj, A. Muda, A. Tharek, N. Omar, and M. I. Jaya, "Systematic review of using machine learning in imputing missing values," *IEEE Access*, vol. PP, pp. 1–1, 2022, doi: 10.1109/access.2022.3160841.

[3] A. Y. Yildiz, E. Koç, and A. Koç, "Multivariate time series imputation with transformers," *IEEE Signal Processing Letters*, vol. 29, pp. 2517–2521, 2022, doi: 10.1108/DTA-12-2020-0298.

[4] E. Cahan, J. Bai, and S. Ng, "Factor-based imputation of missing values and covariances in panel data of large dimensions," *Journal of Econometrics*, 2021, doi: 10.1016/j.jeconom.2022.01.006.

[5] S. Jäger, A. Allhorn, and F. Biessmann, "A benchmark for data imputation methods," *Frontiers in Big Data*, vol. 4, 2021, doi: 10.3389/fdata.2021.693674.

[6] A. Raj, A. Dubey, A. Rasool, and R. Wadwani, "Machine learning based missing data imputation techniques," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–6, 2023, doi: 10.1109/ICCCNT56998.2023.10307783.

[7] S. Zhuchkova and A. Rotmistrov, "How to choose an approach to handling missing categorical data: (un)expected findings from a simulated statistical experiment," *Quality & Quantity*, vol. 56, pp. 1 – 22, 2021, doi: 10.1007/s11135-021-01114-w.

[8] P. Saeipourdizaj, P. Sarbakhsh, and A. Gholampour, "Application of imputation methods for missing values of pm10 and o3 data: Interpolation, moving average and k-nearest neighbor methods," *Environmental Health Engineering and Management*, 2021, doi: 10.34172/ehem.2021.25.

[9] Y. Deng and T. Lumley, "Multiple imputation through xgboost," *Journal of Computational and Graphical Statistics*, 2021, doi: 10.1080/10618600.2023.2252501.

[10] S. Li, H. Huang, and W. Lu, "A neural networks based method for multivariate time-series forecasting," *IEEE Access*, vol. 9, pp. 63 915–63 924, 2021, doi: 10.1109/ACCESS.2021.3075063.

[11] S. Festag, J. Denzler, and C. Spreckelsen, "Generative adversarial networks for biomedical time series forecasting and imputation," *Journal of Biomedical Informatics*, vol. 129, p. 104058, 2022, doi: 10.1016/j.jbi.2022.104058.

[12] L. Xu, L. Xu, and J. Yu, "Time series imputation with gan inversion and decay connection," *Information Sciences*, vol. 643, p. 119234, 2023, doi: 10.1016/j.ins.2023.119234.

[13] A. Palanivinayagam and R. Damaševičius, "Effective handling of missing values in datasets for classification using machine learning methods," *Inf.*, vol. 14, p. 92, 2023, doi: 10.3390/info14020092.

[14] W. Shu, J. Li, Z. Lan, G. Xu, J. Nie, and S. Liu, "Missing value imputation and prediction of river water quality based on gru–autoencoder with input-decay," *2021 40th Chinese Control Conference (CCC)*, pp. 8169–8174, 2021, doi: 10.23919/CCC52363.2021.9550206.

[15] S. Festag and C. Spreckelsen, "Medical multivariate time series imputation and forecasting based on a recurrent conditional wasserstein gan and attention," *Journal of Biomedical Informatics*, vol. 139, p. 104320, 2023, doi: 10.1016/j.jbi.2023.104320.

[16] W. Du, D. Côté, and Y. Liu, "Saits: Self-attention-based imputation for time series," *Expert Systems with Applications*, vol. 219, p. 119619, 2023, doi: 10.1016/j.eswa.2023.119619.

[17] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Stock price forecasting in indonesia stock exchange using deep learning: a comparative study," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 14, p. 861, 2024, doi: 10.11591/ijece.v14i1.pp861-869.

[18] J. Yoon, J. Jordon, and M. V. D. Schaar, "GAIN: Missing data imputation using generative adversarial nets," *Proceedings of the 35th International Conference on Machine Learning*, pp. 5689–5698, 2018, doi: 10.48550/arXiv.1806.02920.

[19] D. W. Firmansyah and R. Sarno, "Data augmentation technique using two step smote for electronic-nose signal in breath ketone level detection," *International Journal of Intelligent Engineering and Systems*, pp. 523–536, 2023, doi: 10.22266/ijies2023.0831.42.

[20] J. Luo, G. Zhu, and H. Xiang, "Artificial intelligent based day-ahead stock market profit forecasting," *Computers and Electrical Engineering*, vol. 99, p. 107837, 2022, doi: 10.1016/j.compeleceng.2022.107837.