UNP JOURNAL MANAGEMENT CENTER

math.ppj.unp.ac.id
p-ISSN 2716-0726
e–ISSN 2716-0734

Indexed

Article History
Received: Sep 23, 2025 | **Revised**: Oct 30, 2026 | **Accepted**: Oct 31, 2026 | **Published Online**: Oct 31, 2026

Subject Areas:
Statistics

Keywords:
Dividend, Forecatsing, Time Series, Macroeconomic Factors, Artificial Intelligence

Author for correspondence:
Zamahsary Martha
e-mail: zamahsarymartha@fmipa.unp.ac.id

DOI:
https://doi.org/10.24036/rmj.v5i2.100

How to Cite:
Wudd, R., A., Martha, Z., Amalita, N., Mukhti, T., O. (2026). Time Series-Based Transformer for Forecasting SOE Dividends as Danantara's Investment Fund Source. *Rangkiang Mathematics Journal*, 5(2), 161-171. https://doi.org/10.24036/rmj.v5i2.100

Time Series-Based Transformer for Forecasting SOE Dividends as Danantara's Investment Fund Source

Raihan Athaya Wudd[1], Zamahsary Martha[1], Nonong Amalita[1], and Tessy Octavia Mukhti[1]

[1] Department Of Statistics, Faculty of Mathematics and Science, Padang State University, West Sumatra

Abstract- In recent years, dividend allocation from State-Owned Enterprises (SOEs) has become increasingly important following the establishment of Daya Anagata Nusantara (Danantara) on February 4, 2025, a national investment management institution tasked with optimizing state assets outside the state budget, where SOE dividends contribute approximately 40% to 45% of total managed assets. Accurately estimating future dividends is crucial for Danantara to plan and allocate productive investments efficiently and establish a basis for future evaluation. SOE dividend distribution is influenced by internal factors such as net profit and external macroeconomic conditions, including inflation, Bank Indonesia's benchmark interest rate (BI Rate), the Rupiah exchange rate against the US dollar, Gross Domestic Product (GDP), and the money supply. This applied research uses a time series-based Transformer model with a self-attention mechanism to capture short-term and long-term dependencies across multivariate financial and macroeconomic data. The secondary dataset consists of annual SOE dividends, net profit, and Indonesian macroeconomic indicators obtained from the Ministry of SOEs and Bank Indonesia covering 35 observations from 1990 to 2024. Model accuracy is evaluated using the Mean Absolute Percentage Error (MAPE). The results show that the best model is achieved with parameters of 450 epochs, a d_model dimension of 128, four multi-head self-attention heads, and four layers, yielding a testing MAPE of 31.81%, which is categorized as reasonably suitable for forecasting. Forecasting results for the five-year horizon show an initial downward adjustment in SOE dividend values before increasing significantly in the fourth projection year and moderating in the final year, providing a valuable quantitative reference for strategic investment planning by Danantara.

Copyright ©2020 The Authors. Published by Rankiang Mathematics Journal which permits unrestricted use, provided the original author and source are credited

This is an open access article under the CC–BY-SA license

1. Introduction

As corporate entities whose capital is predominantly owned by the state, State-Owned Enterprises (SOEs) play a vital role in the Indonesian national economy (Kementerian BUMN, 2024). Beyond carrying out public service obligations, SOEs actively drive national economic growth, maintain stability across strategic sectors, and spearhead development in infrastructure, energy, and technology (Kementerian BUMN, 2024; Bank Indonesia, 2024). Consequently, strengthening the strategic role of SOEs represents an essential step in supporting Indonesia's continuously expanding economy (Bank Indonesia,2024).

SOEs demonstrate their financial performance through two primary indicators: net profit and dividends (Kementerian BUMN, 2024). Net profit reflects the managerial capability of SOEs to operate businesses and generate earnings, whereas dividends represent the portion of profit contributed to the state treasury as Non-Tax State Revenue (Kementerian BUMN, 2024). Historically, dividends contributed by SOEs to the state showed a consistent upward trend from 1990 to 2019, rising from approximately IDR 1.9 trillion to more than IDR 50 trillion (Kementerian BUMN, 2024). Between 2020 and 2021, dividend payments experienced a sharp contraction due to the economic disruptions caused by the COVID-19 pandemic, before rebounding drastically between 2022 and 2024 to surpass IDR 80 trillion (Kementerian BUMN, 2024). In 2024, all SOEs recorded a combined net profit of IDR 304 trillion, representing an approximate 7% decrease compared to IDR 327 trillion in 2023 (Kementerian BUMN, 2024). Despite this slight decline in net profit, total SOE dividends paid to the state increased to a record IDR 85.5 trillion in 2024, up from IDR 81.2 trillion in 2023, demonstrating the substantial capacity of SOEs in financing public infrastructure and national development (Kementerian BUMN, 2024).

According to the Ministry of SOEs (Kementerian BUMN, 2024), the primary dividend contributions originate from seven major SOEs that have been integrated into Danantara as core sources of investment funds: Bank Rakyat Indonesia or BRI (IDR 25.715 trillion), Bank Mandiri (IDR 17.179 trillion), Mining Industry Indonesia or MIND ID (IDR 11.214 trillion), Pertamina (IDR 9.357 trillion), Telkom Indonesia (IDR 9.211 trillion), Bank Negara Indonesia or BNI (IDR 6.277 trillion), and PLN (IDR 3.09 trillion). Among these seven enterprises, BRI recorded the highest dividend contribution due to its consistent financial performance, lower risk profile, and stable revenue streams (Bank Rakyat Indonesia, 2023). Conversely, Pertamina and Telkom Indonesia operate in capital-intensive industries (Kementerian BUMN, 2023). Pertamina must contend with global crude oil price fluctuations and substantial operational costs, while Telkom Indonesia must continuously reinvest in network infrastructure such as fiber optics, data centers, and digital services, resulting in a smaller proportion of net profits distributed as dividends (Kementerian BUMN, 2023).

Although the number of active SOEs exceeds 65 companies, only a subset regularly distributes dividends to the state (Kementerian BUMN, 2023). This disparity is driven by variations in financial performance across SOEs, business scale, and corporate status (Kementerian BUMN, 2023). Several SOEs are currently executing business expansion or financial recovery programs, requiring their profits to be allocated as retained earnings for internal investment needs (Kementerian BUMN, 2023). Furthermore, within holding company structures such as Mining Industry Indonesia (MIND ID), profits generated by subsidiaries are first consolidated at the parent company level before being distributed to the state (Kementerian BUMN, 2023).

The dividends generated by SOEs are influenced by various interrelated internal and external factors (Bank Indonesia, 2024). Internally, net profit is the most influential determinant, as dividends are distributed from residual earnings after deducting reinvestment requirements and other obligations (Bank Indonesia, 2024). Externally, SOE dividend policies are shaped by macroeconomic factors, including inflation, Bank Indonesia's benchmark interest rate (BI Rate), the Rupiah exchange rate against the US dollar, Gross Domestic Product (GDP), and the money supply (Bank Indonesia, 2024; Kementerian BUMN, 2023). High inflation tends to suppress purchasing power and corporate profit margins, thereby reducing a company's ability to pay dividends (Bank Indonesia, 2024). Similarly, an increase in the BI Rate raises borrowing costs, which reduces net profit and lowers dividend payouts (Bank Indonesia, 2024). Fluctuations in the Rupiah exchange rate against the US dollar affect the cost of imported raw materials and foreign debt servicing, directly impacting corporate profitability (Badan Pusat Statistik, 2023). Conversely, positive GDP growth reflects healthy macroeconomic conditions that

generally correlate with higher SOE profits and dividends, while the money supply reflects economic liquidity that stimulates consumption and investment activities, ultimately supporting overall SOE financial performance (Kementerian BUMN, 2023).

In recent years, however, investment decisions undertaken by several SOEs have exhibited inaccuracies in fund allocation triggered by weak corporate governance, limited transparency, decision-making without in-depth risk analysis, and inadequate supervision of investment portfolios (Badan Pemeriksa Keuangan, 2021). The clearest examples occurred in the cases of PT Asuransi Jiwasraya and PT Asabri (Badan Pemeriksa Keuangan, 2021). The Audit Board of Indonesia (BPK) recorded that PT Asuransi Jiwasraya incurred losses of IDR 16.81 trillion due to poorly managed investments in stocks and mutual funds, while PT Asabri suffered losses of IDR 22.78 trillion resulting from the placement of investment funds in high-risk and detrimental financial instruments (Badan Pemeriksa Keuangan, 2021). These substantial state losses indicate that internal SOE governance requires comprehensive reform in financial reporting transparency, internal auditing, and managerial accountability (Badan Pemeriksa Keuangan, 2021). If these problems are not addressed promptly, the sustainability of state investments through SOEs will be threatened, and their contribution to national development could decline significantly (Badan Pemeriksa Keuangan, 2021).

To overcome these challenges in national investment management, the Indonesian government established a national investment management institution named Daya Anagata Nusantara (Danantara) on February 4, 2025, under the Third Amendment to Law Number 19 of 2023 concerning SOEs and Presidential Regulation (Perpres) Number 10 of 2025 (Peraturan Presiden, 2025). Initiated by President Prabowo Subianto, Danantara holds the mandate to manage and optimize government investments separated from the State Budget (APBN), consolidate dispersed government assets across ministries, and ensure that investments in strategic sectors generate maximum and sustainable economic returns (Peraturan Presiden, 2025; Danareksa, 2024). Danantara manages not only SOE dividends but also strategic state shares, state property assets, and State Capital Participation (Kementerian BUMN, 2024; Danareksa, 2024). Across its entire portfolio, SOE dividends contribute approximately 40% to 45% of the total managed asset value, while the remainder originates from non-dividend state assets (Kementerian BUMN, 2024; Danareksa, 2024).

From a global perspective, the Financial Services Authority (OJK) noted that the establishment of Danantara aligns with the international Sovereign Wealth Fund (SWF) model used by many countries to manage strategic state capital (CNN Indonesia, 2025). Globally, Danantara ranks eighth among the world's largest SWFs with total managed assets of USD 600 billion, positioned alongside major global funds such as Norway GPFG (USD 1,740 billion), China CIC (USD 1,330 billion), China SAFE (USD 1,090 billion), UAE ADIA (USD 1,060 billion), Kuwait KIA (USD 1,030 billion), Saudi Arabia PIF (USD 925 billion), and Singapore GIC (USD 800.80 billion) (CNN Indonesia, 2025). Given its substantial scale and reliance on SOE dividends for 40% to 45% of its portfolio, it is crucial for Danantara to ensure that available funds are directed toward productive and strategic investments (Kementerian BUMN, 2024; Danareksa, 2024). One essential effort to achieve this objective is accurately forecasting the amount of SOE dividends to be received, enabling Danantara to plan and allocate investment capital more efficiently while establishing an objective baseline for future evaluation (Kementerian BUMN, 2024; Danareksa, 2024).

The application of artificial intelligence and machine learning has become a critical strategy for improving prediction accuracy in the financial sector (Greff, 2017; Box, 2015; Goodfellow, 2016). Although methods such as Long Short-Term Memory (LSTM) and Autoregressive Integrated Moving Average (ARIMA) have been widely applied, both exhibit structural limitations (Greff, 2017; Box, 2015). LSTM models frequently encounter difficulties when handling long sequences due to gradient degradation, whereas ARIMA is effective only for linear and stationary data, rendering it less capable of capturing complex non-linear patterns in volatile financial time series (Greff, 2017; Box, 2015; Goodfellow, 2016).

As a solution to these limitations, the Transformer architecture was introduced by Vaswani et al (2017) and has emerged as a highly promising approach for time series analysis. The Transformer is a neural network architecture that relies entirely on a self-attention mechanism without recurrent layers, enabling it to capture long-term dependencies efficiently and process sequential data in parallel to

math.ppj.unp.ac.id Rangkiang Mathematics Journal Vol 5, No. 2, 2026

accelerate training time (Vaswani et al, 2017). Consequently, Transformers handle complex patterns more flexibly than LSTM or ARIMA and do not require stationarity assumptions, making them well-suited for dynamic and irregular financial time series (Wu et al, 2021; Zeng et al, 2023). Furthermore, Wang et al (2023) demonstrated that self-attention-based approaches are highly effective in multivariate time series forecasting because they capture complex interactions across interrelated macroeconomic variables.

Empirical studies further substantiate the effectiveness of Transformer-based architectures in time series forecasting (Madani, 2024; Hartanto & Gunawan, 2024; Shi et al, 2022). Madani (2024) and Hartanto and Gunawan (2024) showed that Transformer-based deep learning models such as the Temporal Fusion Transformer (TFT) achieved high effectiveness in multivariate stock price forecasting for major Indonesian banking stocks (BBCA, BBRI, BMRI, BBNI, and BRIS), reaching a Mean Absolute Percentage Error (MAPE) as low as 0.23%. Similarly, Shi et al (2022) concluded that the Transformer model achieved the lowest Mean Absolute Error (MAE) of 14.60 and Root Mean Squared Error (RMSE) of 23.57 in multivariate time series forecasting, outperforming both LSTM and GRU models on the same dataset. In the context of SOE dividend forecasting, predictive accuracy directly impacts the effectiveness of strategic investment planning by Danantara (Kementerian BUMN, 2024; Danareksa, 2024). Therefore, this study applies a time series-based Transformer model to forecast Indonesian SOE dividends as a source of investment funds for Danantara.

2. Methods

This study is an applied quantitative research project that utilizes secondary annual time series data obtained from official publications of the Ministry of State-Owned Enterprises (via Katadata Databoks) and the official website of Bank Indonesia (Kementerian BUMN, 2024; Bank Indonesia, 2024). The dataset comprises 35 annual observations spanning from 1990 to 2024, consisting of SOE dividends as the target variable and six explanatory variables: net profit, inflation rate, BI Rate, Rupiah exchange rate against the US dollar, Gross Domestic Product (GDP), and money supply (M1). The structure and measurement units of the research variables are presented in Table 1.

Table 1. Research Variables

Variable	Description	Unit
Target (Y)	SOE Dividends	Trillion Rupiah
X_1	Net Profit	Trillion Rupiah
X_2	Inflation	Percent (%)
X_3	BI Rate	Percent (%)
X_4	Rupiah to USD Exchange Rate	Rupiah
X_5	Gross Domestic Product	Billion/Trillion Rupiah
X_6	Money Supply (M1)	Billion/Trillion Rupiah

1. Data Normalization and Denormalization

Data normalization transforms original numerical values into a uniform scale without altering the underlying information patterns (Budianita & Prijodiprodjo, 2013). In machine learning and Transformer-based time series forecasting, normalizing multivariate inputs is essential because variables measured in vastly different units can cause numerical instability during training (Han & Pei, 2022). This study applies the Min-Max normalization method to scale all variables into the [0,1] interval, which enhances convergence speed and model accuracy (Chamidah & Salamah, 2012).

$$z = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

Conversely, after the forecasting process is completed, the normalized predictions are returned to their original scale using the denormalization formula.

$$x_i = z(x_{\max} - x_{\min}) + x_{\min}$$

where z is the normalized data value, x_i is the i-th original observation, $x_{\min}$ is the minimum value of the variable, and $x_{\max}$ is the maximum value of the variable

2. Sliding Window and Data Splitting

A standard approach in neural network time series forecasting is the sliding window (windowing) method, which restructures sequential observations into supervised input-target segments to minimize prediction error (Wahyuni, 2021). In this study, historical observations from previous periods ($t-3$ to $t-1$) across the six input features are segmented into overlapping windows to predict future SOE dividend values. Following normalization and windowing, the dataset is partitioned into 70% training data and 30% testing data.

3. Time Series Transformer Architecture

Following Vaswani et al (2017), the Time Series Transformer processes sequential multivariate inputs through six main architectural stages:

a. Input Embedding: Multivariate time series data is first converted into a continuous vector representation of dimension d_{model} through a linear embedding layer so that it can be processed by the network (Vaswani et al, 2017).

$$E = X \cdot W + b$$

Where:

X : *Input* multivariate with dimensions $[n, t, k]$
n : Number of samples
t : Length of the time sequence
k : Number of features *Input*
W : Weighting *Embedding*
b : Bias, additional vectors for each dimension *Output*
E : *Output embedding*
d_{model} : Model dimensions (feature vector length in *Transformer)*

b. Positional Encoding: Because the Transformer architecture does not contain sequential recurrence like RNN or LSTM models, temporal position information must be added to the input embeddings using positional encoding (Vaswani et al, 2017).

$$PE_{(t,2i)} = \sin\left(\frac{t}{1000^{\frac{2i}{d_{model}}}}\right); \; PE_{(t,2i+1)} = \cos\left(\frac{t}{1000^{\frac{2i}{d_{model}}}}\right);$$

Where:

$PE_{(t,i)}$: Value *positional encoding* at the time t and the dimension of i
t : Time to *tsequence*
i : Dimension of features in vector
d_{model} : Model dimensions (feature vector length in *Transformer)*

where $PE_{(t,i)}$ is the positional encoding value at time step t and feature dimension i, t is the time step in the sequence, i is the feature dimension index, and d_{model} is the model feature vector dimension. Sine and cosine functions generate unique positional patterns across different frequencies, which are then added to the input embedding.

$$E' = E + PE$$

c. Self-Attention and Multi-Head Self-Attention: The core of the Transformer is the self-attention mechanism, which calculates relevance weights between elements across the entire sequence so the model can capture both short-term and long-term dependencies (Vaswani et al, 2017). The scaled dot-product attention formula is defined as

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

Where:

Q : *Query*, Representation of *Input*
K : Key, Available information (historical data)
V : Value, Value to be combined *Query*, *Key*and *Value matrix* (result of *Input* multiplied by the weight matrix)

softmax converts relevance scores QK^T into probability weights, and d_k is the key dimension. To

math.ppj.unp.ac.id Rangkiang Mathematics Journal Vol 5, No. 2, 2026

allow the model to observe relationships across features from multiple perspectives simultaneously, Multi-Head Attention combines h parallel attention heads (Vaswani et al, 2017).

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

Where:

$head_i$	: $Attention\ (QW_i^Q, KW_i^K, VW_i^V)$
h	: Number of attention heads
QW_i^Q, KW_i^K, VW_i^V	: Projection Weight Matrices for each head
W^o	: Output Weight Matrix
$Concat$	: Denotes the concatenation of all attention heads

d. Feedforward Neural Network (FFN): **Feedforward Neural Network** (FFN): After the self-attention sublayer, each sequence element is processed independently at each time step (t) by a position-wise feedforward network to enrich data representation and introduce non-linearity via the ReLU activation function (Vaswani et al, 2017).

$$\text{FFN}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2$$

Where:

x	: *Output* of the process *Self-attention*
$ReLU$	: Activation function
W_i	: Matrix of the weight of the layer$i$
b_i	: Bias of the layer

e. Layer Normalization: Each attention and FFN sublayer is connected via a residual connection followed by layer normalization to standardize values (mean 0, variance 1), ensure gradient stability, accelerate convergence, and preserve critical information from the original input (Vaswani et al, 2017).

$$h_t = \text{LayerNorm}\Big(x + \text{Sublayer}(x)\Big)$$

Where:

x	: Sublayer Input
$Sublayer$	: Output of the FFN process

f. Output Layer: The final representation from the stacked blocks is mapped to the target dividend dimension through a linear output layer.

$$\hat{y}_{t+1} = h_t \cdot W + b$$

where h_t represents the final encoded representation at time t.

4. Training Parameters and ADAM Optimization

Training the Time Series Transformer involves tuning key hyperparameters d_{model} (embedding vector dimension), number of layers (network depth), number of attention heads (h), and number of epochs (complete iterations over the training dataset). During the forward and backward propagation phases, network weights are updated using the Adaptive Moment Estimation (ADAM) optimization algorithm to ensure stable and adaptive convergence.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla_t; \quad v_t = \beta_2 v_{t-1} + (1-\beta_2)\nabla_t^2$$
$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}; \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}; \quad \theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} - \epsilon}$$

where $\hat{m}_t$ and $\hat{v}_t$ are the bias-corrected first and second moment estimates of the loss gradient ∇_t at iteration t, β_1 and β_2 are momentum constants, θ_t represents the updated model weights, and α is the learning rate.

5. Model Accuracy Evaluation Mean Absolute Percentage Error (MAPE)

To evaluate forecasting accuracy during training and testing, this study uses the MAPE, which measures the average absolute percentage error relative to actual values (Montgomery et al, 2015).

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} |\,PE_i|$$

where PE_i is the percentage error at observation i and n is the number of observations. The forecasting capability of the model is interpreted based on the MAPE intervals established by Chang et al (2007), as shown in Table 2.

Table 2. MAPE Evaluation Criteria

MAPE Range	Forecasting Capability Significance
MAPE < 10%	Excellent forecasting capability
10% < MAPE ≤ 20%	Good forecasting capability
20% < MAPE ≤ 50%	Reasonable / suitable for forecasting
MAPE ≥ 50%	Poor forecasting capability

3. Results and Discussion

(a) Descriptive Statistical Analysis and Data Exploration

Before conducting further modeling, a descriptive statistical analysis was carried out across the 35 annual observations spanning from 1990 to 2024 to obtain an overview of the data characteristics. The summary statistics for each research variable are presented in Table 3.

Table 3. Descriptive Statistical Value

Variable	***Mean***	***Minimum***	***Maximum***
Dividend	24.33 Trillion IDR	1.5 Trillion IDR	85.5 Trillion IDR
Net Profit	86.72 Trillion IDR	2.1 Trillion IDR	327 Trillion IDR
Inflation	9.02%	1.57%	77.63%
BI Rate	11.66%	2.75%	70%
Exchange Rate	9.930 IDR	1.901 IDR	16.242 IDR
Gross Domestic Product	6.72 Trillion IDR	14.8 Billion IDR	22.21 Trillion IDR
Money Supply (M1)	659.78 Billion IDR	22.1 Billion IDR	2.56 Trillion IDR

Table 3 shows wide historical variations across both corporate financial indicators and macroeconomic variables over the observation period. SOE dividends and net profits reached their lowest levels during the post-Asian Financial Crisis period and achieved their highest realizations in the final years of the dataset, reflecting substantial long-term corporate expansion. Meanwhile, inflation and the Bank Indonesia benchmark interest rate (BI Rate) recorded extreme spikes during the monetary crisis before stabilizing in recent years, whereas the Rupiah selling exchange rate, Gross Domestic Product (GDP), and M1 money supply demonstrated structural upward growth alongside national economic expansion. Preliminary visual exploration of the historical series (Figure 1) illustrates the trajectory of SOE dividend payments over time, where the solid line represents historical dividend realizations and the vertical red dashed line marks the boundary between the training subset and the testing subset. While dividends and net income exhibit a consistent long-term upward trend, macroeconomic fluctuations and external shocks such as the pandemic period temporarily disrupted dividend stability.

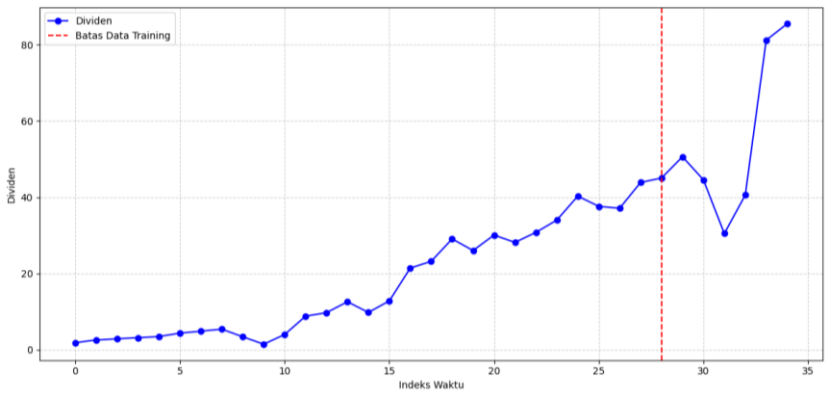


Figure 1. Trajectory of SOE Dividend Payments Over Time

(b) Data Normalization and Sliding Window Modification

Prior to model training, Min-Max normalization was applied to scale all variables into a uniform interval from 0 to 1, preventing numerical instability caused by differences in measurement units while preserving the original data distribution (Budianita & Prijodiprodjo, 2013; Han & Pei, 2022). The normalized values for the initial and final observation periods are presented in Table 4.

Table 4. Normalized Data Results

Period	X_1	X_2	X_3	X_4	X_5	X_6	$Target\ (Y)$
1	0.012	0.105	0.222	0.007	0	0	0.004
2	0.015	0.104	0.250	0.008	0.006	0.001	0.013
3	0.016	0.044	0.219	0.009	0.011	0.003	0.016
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
35	0.929	0	0.049	1	1	1	1

Following normalization, the dataset was restructured using the sliding window method to form sequential multivariate input-output segments suitable for the Transformer model (Wahyuni, 2021). Using the six explanatory features across consecutive historical periods to predict the target dividend in the subsequent period, the time series was segmented into sequential windows as shown in Table 5.

Table 5. Sliding Window Modification Results

Window	X1	X2	X3	⋯	X18	$Target$ (Y)
1	0.012004	0.104654	0.222305	⋯	0.381	0.020238
2	0.015389	0.104523	0.250558	⋯	0.003	0.02381
⋮	⋮	⋮	⋮	⋮	⋮	⋮
13	0.020	0.023	0.034	⋯	1	1

Each window in Table 5 represents a sequential input-output set that enables the Transformer architecture to learn temporal patterns, inter-variable relationships, and long-term dependencies across historical periods. The resulting windowed dataset was then divided chronologically into 70% training data for weight and bias optimization and 30% testing data for out-of-sample model evaluation.

(c) Hyperparameter Tuning and Model Evaluation

In the Time Series-based Transformer training process, model quality is determined by combining four core hyperparameters: the number of epochs (determining how many times the entire dataset passes through the network), d_{model} (the embedding vector dimension), the number of attention heads in multi-head self-attention, and the number of encoder-decoder layers (Vaswani et al, 2017). The optimal parameter configuration was identified through a trial-and-error grid search based on the MAPE, as summarized in Table 6.

Table 6. Transformer Model Evaluation on Training Data

No	***epoch***	$\boldsymbol{d_{model}}$	***head***	***layer***	**MAPE**
1	300	64	2	2	34,18%
2	300	64	4	4	39,11%
3	300	128	8	8	31,84%
4	300	128	8	2	30,86%
5	350	64	2	4	31,56%
6	350	64	2	8	33,75%
7	350	128	4	2	28,89%
8	350	128	8	4	25,32%
9	400	64	2	8	29,49%
10	400	64	2	2	32,07%
11	400	128	4	4	26,77%

12	400	128	4	8	29,39%
13	450	64	2	2	30,70%
14	450	64	4	8	25,21%
15*	**450**	**128**	**4**	**4**	**24,19%**
16	450	128	8	2	30,31%
17	500	64	2	4	29,99%
18	500	64	4	8	31,42%
19	500	128	8	2	32,13%
20	500	128	8	4	24,56%

Table 6 demonstrates that training the Transformer model with different hyperparameter combinations produces varying levels of accuracy. Increasing the number of layers, attention heads, or training epochs does not automatically reduce the error rate, indicating that greater architectural complexity does not guarantee higher predictive accuracy. Among all tested configurations, Model 15 produced the lowest training MAPE and was selected as the optimal model.

Next, the optimal Transformer model obtained from the training phase was evaluated on the testing dataset to assess its ability to predict unseen observations. Figure 2 illustrates the comparison between actual historical dividends (green line) and model predictions (yellow line). The prediction trajectory closely follows the general pattern of the actual data, capturing the downturn beginning in 2019 as well as the subsequent recovery in the final period. On the testing dataset, the optimal model achieved a MAPE of 31.81%. Based on the evaluation criteria of Chang et al (2007), this accuracy level falls within the reasonable and suitable category for forecasting SOE dividends.

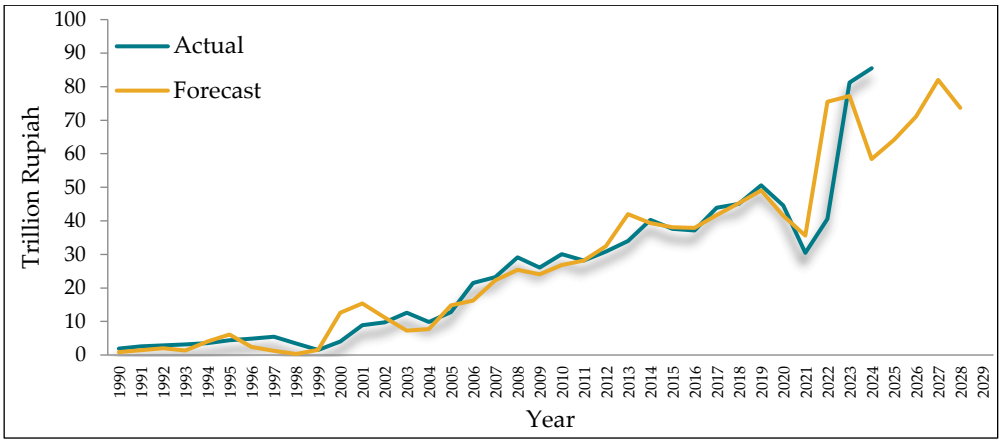


Figure 2. Comparison between actual historical dividends and model predictions

(d) SOE Dividend Forecasting and Discussion

The optimal trained Transformer model was subsequently applied to forecast SOE dividends for the five-year horizon from 2025 to 2029 by incorporating macroeconomic indicators and SOE financial performance as model inputs. After denormalizing the predicted outputs back to their original Rupiah scale, the final dividend forecasts were obtained as presented in Table 7.

Table 7. SOE Dividend Forecast Results

Period	Forecast Results
2025	58.38 Trillion Rupiah
2026	64.11 Trillion Rupiah
2027	71.08 Trillion Rupiah
2028	82.02 Trillion Rupiah

math.ppj.unp.ac.id Rangkiang Mathematics Journal Vol 5, No. 2, 2026

2029	73.64 Trillion Rupiah

Based on the forecast results in Table 7, projected SOE dividends experience an initial decline during the early forecast horizon compared to the 2024 peak realization, followed by a steady upward recovery from 2026 through 2028 before experiencing a moderate adjustment in 2029. Although SOE dividend distributions fluctuate over the five-year horizon, the strong upward trajectory culminating in 2028 indicates substantial medium-term growth potential as a strategic investment funding source for Danantara. These projected variations reflect the dynamic interaction between internal SOE net profit performance, corporate dividend payout policies, and external macroeconomic conditions—including inflation, the BI Rate, exchange rates, GDP, and money supply—in alignment with corporate finance theory (Kementerian BUMN, 2024; Bank Indonesia, 2024).

As Indonesia's national investment management institution, Danantara plays a strategic role in ensuring that capital derived from SOE dividends is allocated toward productive, high-impact investments (Peraturan Presiden, 2025), (Danareksa, 2024). Because Danantara's operational effectiveness depends heavily on anticipating annual dividend inflows, the consistency of the Time Series Transformer model in capturing complex multi-variable relationships and long-term dependencies provides a reliable scientific foundation for medium- and long-term capital planning (Vaswani et al, 2017; Wu et al, 2021). Furthermore, adopting artificial intelligence through Transformer-based forecasting supports the modernization and data-driven transformation of national state asset management.

4. Conclusion

Based on the research results and discussion conducted using the Time Series-based Transformer method to forecast dividends of State-Owned Enterprises (SOEs), twenty models with different training hyperparameter combinations were evaluated through a trial-and-error process. Among all tested configurations, the optimal model was determined based on the lowest Mean Absolute Percentage Error (MAPE) value. The best Time Series-based Transformer model consists of the parameter combination of 450 epochs, a *d_model* dimension of 128, four attention heads in multi-head self-attention, and four layers, which produced a training MAPE of 24.19% and a testing MAPE of 31.81%. Based on this accuracy level, the model is categorized as reasonably good and suitable for forecasting SOE dividends. The optimal model was subsequently applied to forecast SOE dividends for the five-year period from 2025 to 2029. The forecasting results indicate a downward adjustment in dividend values at the beginning of the projection period compared to the previous peak realization, followed by a steady recovery that reaches IDR 82.02 trillion in 2028, before experiencing a moderate decline to IDR 73.64 trillion in 2029. These findings provide an empirical overview of the future dynamics of potential state revenue from SOE dividends and serve as a strategic reference for Daya Anagata Nusantara (Danantara) in planning and allocating national investment funds.

This study still has limitations, particularly the relatively limited observation length of 35 annual periods and the computational complexity of tuning Transformer architectures. Therefore, future researchers are advised to use data with a longer observation horizon and further develop the model by exploring broader parameter combinations such as the number of layers, attention heads, and learning rates. Additionally, automated hyperparameter optimization methods such as Bayesian Optimization or Grid Search can be implemented to obtain model configurations with higher predictive accuracy and generalization performance.

References

1. Kementerian BUMN. (2024). Laporan Tahunan BUMN Tahun 2024.
2. Bank Indonesia. (2024). Laporan Perekonomian Investasi Tahun 2024.
3. Bank Rakyat Indonesia. (2023). Laporan Tahunan Bank Rakyat Indonesia. Jakarta: PT Bank Rakyat Indonesia (Persero) Tbk.
4. Kementerian BUMN. (2023). Paparan Program Strategis Pengelolaan Aset Negara dan Investasi Jangka Panjang. Jakarta: Kementerian BUMN.
5. Badan Pusat Statistik. (2023). Nilai Tukar Valuta Asing. Jakarta: BPS RI.
6. Badan Pemeriksa Keuangan. (2021). Ikhtisar Hasil Pemeriksaan Semester II. Jakarta: BPK RI.

7. Peraturan Presiden (2025). Perpres Nomor 10 Tahun 2025 tentang Perubahan Ketiga atas Undang-Undang Nomor 19 Tahun 2023 tentang Badan Usaha Milik Negara. Jakarta.
8. Danareksa. (2024). Company Profile & Press Release: Pembentukan Danantara sebagai Special Mission Vehicle. Jakarta: PT Danareksa.
9. CNN Indonesia. (2025). Kepala Eksekutif OJK: Kehadiran DANANTARA bukan hal baru dalam dunia investasi global [Wawancara]. Diakses dari https://www.cnnindonesia.com/ekonomi/danantara-investasi-ojk.
10. Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A search space odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924.
11. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). New Jersey: Wiley.
12. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. Cambridge, MA: MIT Press.
13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
14. Wu, H., et al. (2021). Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. Advances in Neural Information Processing Systems, 34.
15. Zeng, A., et al. (2023). Are Transformers Effective for Time Series Forecasting? arXiv preprint arXiv:2305.09646.
16. Wang, H., Mo, Y., Yin, N., Dai, H., Li, B., & Fan, S. (2023). An Efficient Attention-Based Approach for Multivariate Series Forecasting. arXiv preprint arXiv:2312.06220.
17. Madani, A. R. (2024). Penerapan Temporal Fusion Transformer (TFT) untuk prediksi data multivariate deret waktu: Studi kasus harga saham BBCA, BBRI, BMRI, BBNI, dan BRIS (Undergraduate Thesis, UIN Sunan Kalijaga Yogyakarta).
18. Hartanto, S., & Gunawan, A. A. S. (2024). Temporal Fusion Transformers for Enhanced Multivariate Time Series Forecasting of Indonesian Stock Prices. International Journal of Advanced Computer Science and Applications, 15(7).
19. Shi, X., et al. (2022). Transformer-based multivariate air quality forecasting.
20. Budianita, E., & Prijodiprodjo, W. (2013). Penerapan Learning Vector Quantization (LVQ) untuk Klasifikasi Status Gizi Anak. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 7(2), 155–166.
21. Han, J., Kamber, M., & Pei, J. (2022). Data Mining: Concepts and Techniques (4th ed.). Cambridge: Morgan Kaufmann.
22. Chamidah, N., & Salamah, M. (2012). Implementasi algoritma backpropagation dalam jaringan syaraf tiruan untuk pengenalan pola huruf braille. Jurnal EECCIS (Electronic, Electrical, Communication, Control, and Information System), 6(1), 65–70.
23. Wahyuni, R. E. (2021). Optimasi Prediksi Inflasi Dengan Neural Network Pada Tahap Windowing Adakah Pengaruh Perbedaan Window Size. Technologia: Jurnal Ilmiah, 12(3), 176–181.
24. Montgomery, C. D., Jennings, C. L., & Kulahci, M. (2015). Introduction to Time Series Analysis and Forecasting (2nd ed.). New Jersey: John Wiley and Sons, Inc.
25. Chang, P. C., Wang, Y. W., & Liu, C. H. (2007). The Development of Weighted Evolving Fuzzy Neural Network for PCB Sales Forecasting. Expert Systems with Applications, 32(1), 86–96. https://doi.org/10.1016/j.eswa.2005.11.021.

math.ppj.unp.ac.id Rangkiang Mathematics Journal Vol 5, No. 2, 2026