# FORECASTING THE CURRENCY PRICE

**TRINH THI MY CHUNG[1], TRAN PHUONG ANH[2], AND CHE DUY KHANG 3[3]**

[1]Faculty of Information Systems, University of Information Technology, (e-mail: 21520653@gm.uit.edu.vn)
[2]Faculty of Information Systems, University of Information Technology, (e-mail: 21520595@gm.uit.edu.vn)
[3]Faculty of Information Systems, University of Information Technology, (e-mail: 21522187@gm.uit.edu.vn)

## ABSTRACT

The rapid fluctuations in currency exchange rates pose significant challenges for investors and policymakers. This study aims to forecast currency prices using a range of statistical, machine learning, and deep learning algorithms. The models employed in this research include Linear Regression, Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing State Space Model (ETS), Stacking, Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Gated Recurrent Units (GRU), and Long Short Term Memory (LSTM). Evaluation metrics such as Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are utilized to assess the performance of each forecasting model on various currency datasets. The findings indicate that deep learning models, particularly GRU and LSTM, outperform other methods in predicting currency prices.

## INDEX TERMS

forecasting the currency price,

## I. INTRODUCTION

The currency price, often referred to as the exchange rate between different national currencies plays a crucial role in global financial markets and economies. Its fluctuations significantly impact various aspects of the economy, including trade, investment, and banking. While currency price movements can present opportunities for investors and businesses, they also introduce uncertainties and risks.

Traditionally, forecasting currency prices has been challenging due to the complex and unpredictable nature of the forex market. However, the advent of machine learning has opened new possibilities in this field. Machine learning algorithms excel at processing large volumes of data and identifying complex patterns that may not be discernible to humans.

In recent years, several algorithms have gained prominence in currency price forecasting. Exponential Smoothing State Space Model (ETS), Stacking Model, and PatchTST are among the notable ones. ETS provides a framework for modeling and forecasting time series data, while Stacking Model combines multiple models to improve predictive accuracy. PatchTST, on the other hand, utilizes a patch-based approach for time series forecasting, leveraging both local and global patterns in the data.

By leveraging statistical models and machine learning techniques for currency price forecasting, significant benefits can be derived for investors, businesses, and even entire nations. Understanding market trends and fluctuations enables informed decision-making regarding investment, risk management, and business planning based on more accurate currency price forecasts.

This article explores the application of statistical models and machine learning algorithms in forecasting currency prices, highlighting their potential to enhance decision-making processes and drive economic growth.

**GBP (Great British Pound)**

The British Pound Sterling (GBP), dating back to its introduction in the late 17th century, boasts a rich and enduring history. Its evolution from the establishment of paper money by the Bank of England to becoming one of the oldest and most widely traded currencies globally signifies its importance. Despite facing numerous economic and geopolitical challenges over the centuries, including wars and financial crises, the GBP has maintained its position as a symbol of stability and strength in the international financial system. Today, the GBP remains a cornerstone of global finance, with its exchange rate closely monitored by investors, traders, and policymakers worldwide. Its value reflects not only the economic health of the United Kingdom but also broader trends in international trade and finance. Thus, the GBP's legacy and resilience underscore its significance in shaping the landscape of global economics for generations.

**EUR (Euro)**

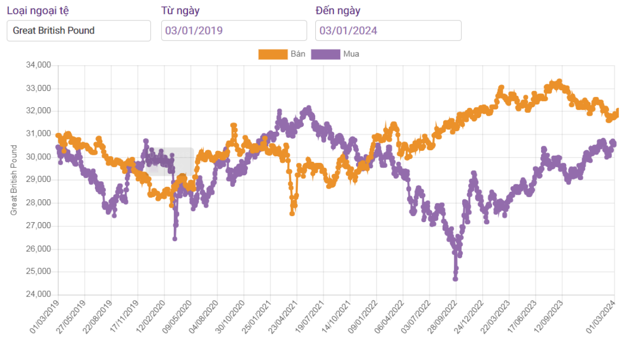


FIGURE 1. Exchange rate of GBP from 2019 to 2024

The Euro is the common currency of the member countries in the Eurozone, a monetary union consisting of 19 of the 27 European Union (EU) member states. It is widely used in trade and finance, serving as a primary medium of exchange for transactions within the Eurozone and beyond. The Euro's exchange rate is closely monitored by economists, investors, and policymakers, as it serves as a key indicator of the region's economic health and stability. Its value against other major currencies, such as the US dollar and the British pound, is often used as a benchmark for global trade and investment. Despite occasional fluctuations, the Euro has maintained a relatively stable exchange rate, contributing to its reputation for stability and credibility in the international financial system. This stability has attracted trust from investors and businesses worldwide, making the Euro one of the most widely accepted and respected currencies in the world.

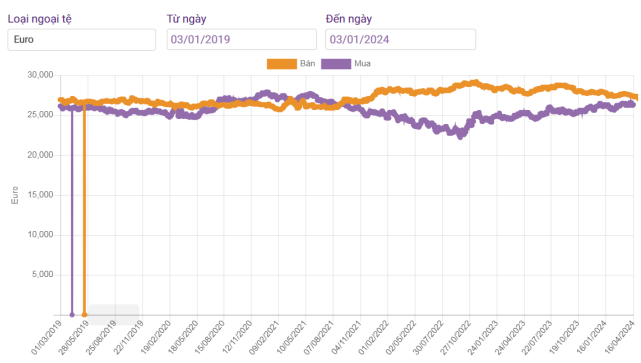


FIGURE 2. Exchange rate of EUR from 2019 to 2024

**EUR (Japanese Yen)**

The yen is the official currency of Japan and is widely recognized as one of the major currencies in the world. Its exchange rate is closely monitored in global financial markets due to Japan's significant role in international trade and finance. The yen's value can be influenced by various factors, including Japan's economic performance, monetary policies, and geopolitical developments.

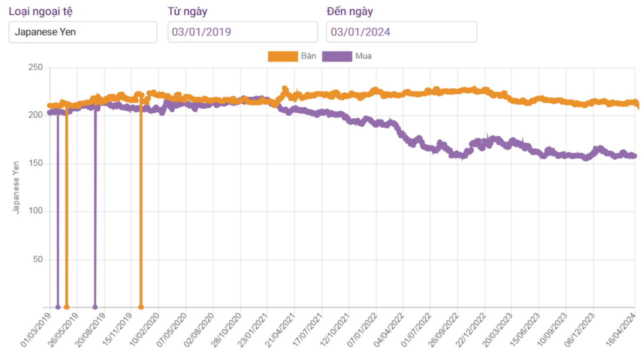


FIGURE 3. Exchange rate of EUR from 2019 to 2024

## II. RELATED WORKS

Accurately forecasting currency exchange rates has been an enormous challenge because of the complexity and dynamism of financial markets. Researchers have worked in this area so as to explore different methodologies to achieve reliable predictions. There have been many research articles on predicting currency exchange rate, such as:

Pengfei Liu et al. [1] study on currency exchange rate prediction by implementing multiple forecasting model to forecast and analyze the daily currency exchange rate of USD/RMB. This study uses CNN, STLSTM, AM model to estimate the accuracy of models. The experiments show that all three models above have higher forecasting accuracy and fitting degree that other models and they are appropriate for forecasting the closing price of the USD/RMB exchange rate.

M.S. Islam, E. Hossain [2] focus on forecasting the currency exchange rate by presenting a new model that combines two powerful neural networks used for time series prediction: Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM). It is used for predicting future closing prices of FOREX currencies. The performance of the model is validated using MSE, RMSE, MAE and R2 score. Researchers applied the model to predict the closing price of four currency pairs in 10 and 30 minutes before the actual time. The model is considered a promising area and has a good predictive capability

Siyuan Liu et al. [3] research on predicting the USD/CNY exchange rate using the novel LASSO-BiLSTM-based ensemble learning method by integrating least absolute shrinkage and selection operator (LASSO) and bidirectional long short-term memory (LSTM). The model performed well and the LASSO-BiLSTM-based ensemble learning method demonstrated high potential in forecasting exchange rates

Qimian Zhu [4] has paper aims to forecast the change of exchange rate of USD/EUR in 2022 using ARIMA model. Researchers discussed the performance of the univariate ARIMA model and the multivariate regression model with ARIMA errors, i.e. four macroeconomic variables, influence the exchange rate incorporated in the AR part of the ARIMA model. The performance of model depends on the quality of the predicted predictors, which are the four macroeconomic variables in the study

Kamruzzaman et al. [5] apply ANNs to predict currency exchange rate of Australian Dollar against other currencies, such as US Dollar (USD), Great British Pound (GBP), Japanese Yen (JPY), Singapore Dollar (SGD), New Zealand Dollar (NZD) and Swiss Franc (CHF). The research focuses on three different ANNs based model, which is Standard Backpropagation (SBP), Scaled Conjugate Gradient (SCG) and Bayesian regularization (BPR). Five different indicators, which are Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS), Correct Up trend (CU) and Correct Down trend (CD), were used to compare the result of three models, for 35 and 65 weeks. The results show that SCG and BPR forecast more accurately than SBP and follow closely to the actual exchange rate, in term of the metrics calculated.

## III. MATERIALS

### A. DATASET

This study will utilize historical exchange rate data of Euro (EUR) to Vietnamese Dong (VND), British Pound (GBP) to Vietnamese Dong (VND), and Japanese Yen (JPY) to Vietnamese Dong (VND) from 1/3/2019 to 1/3/2024. The dataset includes columns such as Date, Purchase Price, Sale Price, and Transfer Price. Since the objective is to forecast foreign currencies' sale prices, only data related to the Sale (VND) columns will be processed.

### B. DESCRIPTIVE STATISTICS

TABLE 1. EUR-VND, GBP-VND, JPY-VND dataset's Descriptive Statistics

| | EUR-VND | GBP-VND | JPY-VND |
|---|---|---|---|
| Count | 1825 | 1825 | 1825 |
| Mean | 26715.9 | 30411.2 | 200.8 |
| Std | 1062.08 | 1268.95 | 20.34 |
| Min | 23533 | 25979 | 166.27 |
| 25% | 26151 | 29537 | 179.55 |
| 50% | 26647 | 30448 | 210.46 |
| 75% | 27469 | 31311 | 217.99 |
| Max | 29180 | 33305 | 228.6 |

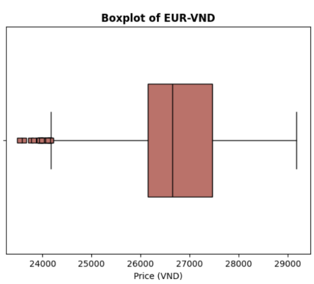


FIGURE 4. EUR-VND price's boxplot

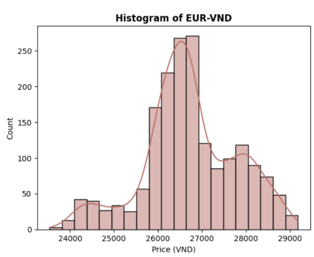


FIGURE 5. EUR-VND price's histogram

From the descriptive statistics of the EUR-VND dataset, we can observe that the selling price of the Euro (EUR) to Vietnamese Dong (VND) currency pair from March 1, 2019, to March 1, 2024, exhibits a skewed distribution with the primary concentration around the mean and median values, but with uneven fluctuations.

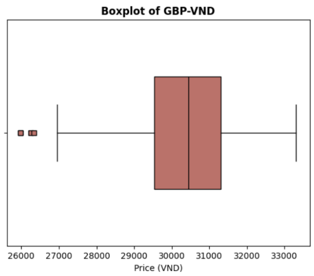


FIGURE 6. GBP-VND price's boxplot

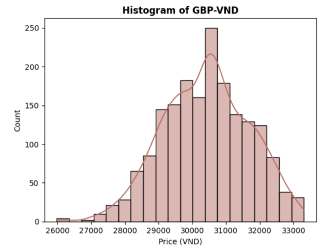


FIGURE 7. GBP-VND price's histogram

From the descriptive statistics of the GBP-VND dataset, we can observe a certain level of volatility in the selling prices of the British Pound (GBP) to Vietnamese Dong (VND) currency pair from March 1, 2019, to March 1, 2024. Most selling prices are concentrated at lower levels, with some higher values contributing to an increase in standard deviation and causing a skewed right distribution on the histogram.

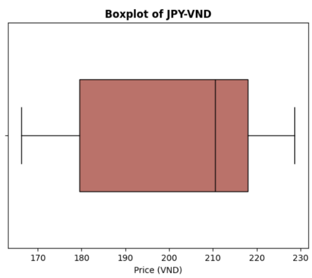


FIGURE 8. JPY-VND price's boxplot

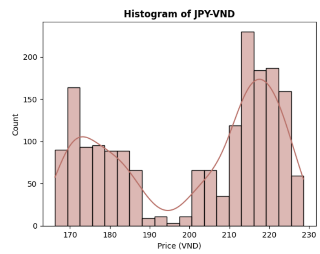


FIGURE 9. JPY-VND price's histogram

From the descriptive statistics of the JPY-VND dataset, we can observe that the market selling price of the Japanese Yen (JPY) to Vietnamese Dong (VND) from March 1, 2019, to March 1, 2024, exhibits variability and fluctuations. The histogram of the data indicates instability and variability in the values. The boxplot of the dataset reveals that the majority of values concentrate at higher price levels.

## IV. METHODOLOGY

### A. LINEAR REGRESSION

Regression analysis is a tool for building mathematical and statistical models that characterize relationships between a dependent variable and one or more independent, or explanatory, variables, all of which are numerical. This statistical technique is used to find an equation that best predicts the y variable as a linear function of the x variables. A multiple linear regression model has the form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Where:

- $Y$ is the dependent variable (Target Variable).
- $X_1, X_2, \ldots, X_k$ are the independent (explanatory) variables.
- $\beta_0$ is the intercept term.
- $\beta_1, ..., \beta_k$ are the regression coefficients for the independent variables.
- $\varepsilon$ is the error term.

### B. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

ARIMA model is a form of regression analysis that gauges the strength of one dependent variable relative to other changing variables. The ARIMA model is used to make predictions about future values of the time series based on its past values. The ARIMA model consists of three parts including Autoregressive (AR), Integrated (I), and Moving Average (MA).

The AR component with order p utilizes the preceding p values of the time series for current value prediction. The AR(p) model has the form:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \epsilon_t$$

Where:

- $Y_t$ is the current observed value.
- $Y_{t-1}, \ldots, Y_{t-p}$ are past observed values.
- $\alpha_0, \alpha_1, \ldots, \alpha_p$ are regression analysis parameters.
- $\epsilon_t$ is the random forecasting error of the current period. The expected mean value is 0.

Integrated (I) represents the differencing of raw observations, allowing the time series to become stationary.

- First Difference I(1): $dY_t = Y_t - Y_{t-1}$
- Second Difference I(2): $dY_t = Y_t - 2Y_{t-2} + Y_{t-3}$

The MA model with order q analyzes the past q forecast errors to anticipate the current value. The MA(q) model has the form:

$$Y_t = \beta_0 + \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}$$

Where:

- $Y_t$ is the current observed value.
- $\epsilon_t$ is a random forecasting error of the current period. The expected mean value is 0.
- $\epsilon_{t-1}, \ldots, \epsilon_{t-q}$ are forecast errors.
- $\beta_0, \beta_1, \ldots, \beta_q$ mean values of Y(t) and moving average coefficients.

### C. EXPONENTIAL SMOOTHING (ETS)

Exponential smoothing is one of the most popular models used for demand forecasting in practice. It includes Error, Trend and Seasonal components, thus being called 'ETS' [6]. The trend component (T) represents the tendency of increasing or decreasing of data over the time. The seasonal (S) shows the periodic fluctuations at fixed intervals within the data. The fluctuations are affected by specific times of the year like holidays, seasonal changes or events. The error (E) is also known as residual. It represents the unpredictable or fluctuations in the data that cannot be explained by the trend or seasonal components. There are three main methods to estimate exponential smoothing, which are:
• Simple exponential smoothing: used when the data has no trend and no seasonal pattern.
• Double exponential smoothing: used for forecasting the time series when the data has a linear trend and no seasonal pattern.
• Triple exponential smoothing: used for forecasting the time series when the data has both linear trend and seasonal pattern. This method is also called Holt-Winters exponential smoothing [7]

$$\begin{aligned} y_t &= l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t \\ l_t &= l_{t-1} + \alpha\varepsilon_t \\ b_t &= b_{t-1} + \beta\varepsilon_t \\ s_t &= s_{t-m} + \gamma\varepsilon_t \end{aligned}$$

| | |
|---|---|
| $y_t$ | Actual value at time $t$. |
| $l_{t-1}$ | Level estimate at time $t-1$. |
| $b_{t-1}$ | Trend estimate at time $t-1$. |
| $s_{t-m}$ | Seasonal component at time $t-m$ (where $m$ is the number of seasons). |
| $\varepsilon_t$ | Random error at time $t$ [8] |

### D. STACKING MODEL

The Stacking Model, also known as Ensemble Learning, is a machine learning technique that combines multiple machine learning models to create a more accurate predictive model.

- Improved accuracy: The Stacking Model can help improve the accuracy of predictions by combining the strengths of multiple machine learning models.
- Reduced overfitting: The Stacking Model can help reduce overfitting by using multiple different machine learning models to learn from the data.
- Increased flexibility: The Stacking Model can be used with many different types of machine learning models, making it a versatile tool for different prediction tasks.

How the Stacking Model Works:

- Train base models: First, you need to train several base machine learning models on the data. These models can be of any type, such as linear regression, decision trees, random forests, etc.
- Create metadata: After training the base models, you need to create metadata. Metadata is a new dataset that includes the predictions of the base models as input data.
- Train meta model: Finally, you need to train a meta-model on the metadata. The meta-model will learn how to combine the predictions of the base models to.

The Stacking Model is a powerful machine-learning technique that can significantly improve the accuracy of predictive models. By combining the strengths of multiple different machine learning models, the Stacking Model offers a flexible and effective approach to solving complex prediction tasks.

### E. GATED RECURRENT UNIT (GRU)

A gated Recurrent Unit (GRU) is a special kind of RNN (Recurrent Neural Network). For every element of a sequence, GRU performs a similar task. That is the reason why it is called recurrent. [8]
A GRU cell consists of two gates: the update gate and the reset gate. These gates control how information flows through the cell, deciding what to keep and what to discard. The update gate determines how much of the past information needs to be passed along to the future. The reset gate determines how much of the past information to forget. These operations are performed by the following equations:

$$z_t = \sigma\left(W^{(z)}x_t + U^{(z)}h_{t-1}\right)$$
$$r_t = \sigma\left(W^{(r)}x_t + U^{(r)}h_{t-1}\right)$$

Where, $h_t$ and $h_{t-1}$ represent the output of the current and previous states, respectively, while $r_t$ and $z_t$ indicate the reset and update gates, respectively. $\sigma$ is the logistic sigmoid function while $W_r$ and $U_r$ are the weight matrices.

### F. LSTM

LSTM (Long Short-Term Memory) networks are a specialized type of recurrent neural network (RNN) architecture designed to address the challenge of learning and remembering over long sequences of data. They excel at capturing long-term dependencies in sequential data, making them well-suited for tasks such as time series forecasting.
**Memory Cells:** LSTM networks utilize memory cells to store and update information over time, allowing them to retain relevant information over long sequences.
**Gating Mechanisms:** They incorporate gates to regulate the flow of information, including input gates, forget gates, and output gates, enabling precise control over what information is stored or discarded.
**Update Cell State:** LSTM networks update their cell state by combining the previous cell state with new information, controlled by the input and forget gates. The equation for updating the cell state is:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

Where:

- $C_t$ is the updated cell state.
- $f_t$ is the forget gate activation.
- $C_{t-1}$ is the previous cell state.
- $\tilde{C}_t$ is the candidate cell state.

**Memory Retention:** LSTM networks are adept at retaining and selectively updating information over long sequences, making them effective for tasks requiring memory of past events.
**Effective Sequence Modeling:** Their ability to capture long-term dependencies makes them highly effective for modeling sequential data, enabling accurate predictions in time series forecasting.

### G. RECURRENT NEURAL NETWORK (RNN)

RNN is a type of neural network model designed to handle sequential data. The distinctive feature of RNNs is their ability to maintain information across time steps, allowing them to utilize information from previous steps to influence the processing of the current step. This makes RNNs particularly useful in applications where the order of data is important, such as natural language processing (NLP), machine translation, and time series forecasting.
RNNs operate by processing sequential data and maintaining a hidden state to capture information from previous time steps. The formula for updating the hidden state is:

$$\alpha_t = \psi_0(W_{\alpha x}x_t + W_{\alpha\alpha}\alpha_{t-1} + b_\alpha)$$

Where:

- $\alpha_t$ is a hidden layer state at each time step t
- $\psi_0$ is the activation function
- $W_{\alpha x}$ and $W_{\alpha\alpha}$ are weight matrices
- $x_t$ is an input data
- $b_\alpha$ is a bias vector

The formula predicts the output at each time t:

$$y_t = \Psi_1(W_{y\alpha}\alpha_t + b_y)$$

Where:

- $y_t$ is an output data at each time t
- $\Psi_1$ is the activation function
- $W_{y\alpha}$ is weight matrices
- $\alpha_t$ is a hidden layer state
- $b_y$ is a bias vector

### H. MULTILAYER PERCEPTRON (MLP)

MLP is a type of artificial neural network that belongs to the feed-forward neural network group. It consists of multiple layers: an input layer, one or more hidden layers, and an output layer. Each neuron in an MLP uses a nonlinear activation function, such as the sigmoid, ReLU, or tanh function. These neurons are fully connected, meaning each neuron in one layer is connected to every neuron in the adjacent layer. MLP can learn and model nonlinear relationships between inputs and outputs, making it effective for many complex problems.
The formula to calculate the input of a layer (except the input layer):

$$z^{(l)} = W^{(l)T}\alpha^{(l-1)} + b^{(l)}$$

Where:

- $z^{(l)}$ is a matrix containing the input values for each neuron in layer (l)
- $W^{(l)}$ is a matrix containing the connection weights between neurons in layer (l-1) and neurons in layer (l)

- $\alpha^{(l-1)}$ is a matrix containing the output values of layer (l-1) and serves as the input for layer (l). For the layer immediately following the input layer, it will be replaced by matrix X
- $b^{(l)}$ is a matrix containing the threshold values for each neuron in layer (l)

When the input value exceeds the threshold, meaning the neuron's z value is greater than 0, the neuron will produce an output value. The formula to calculate the output of a layer (except the input layer):

$$\alpha^{(l)} = f(z^{(l)})$$

Where:

- $\alpha^{(l)}$ is a matrix containing the output values of each neuron belongs to layer l
- $f()$ is an activation function

## V. RESULT

### A. EVALUATION METHODS

**Mean Percentage Absolute Error** (MAPE): is the average percentage error in a set of predicted values.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| = 1$$

**Root Mean Squared Error** (RMSE): is the square root of the average value of squared error in a set of predicted values.

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

**Mean Absolute Error** (MAE): is a measure of the average difference between predicted values and actual values in a dataset.

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

Where:
• $n$ is the number of observations in the dataset.
• $y_i$ is the true value.
• $\hat{y}_i$ is the predicted value.

### B. EUR-VND DATASET

| EUR-VND Dataset's Evaluation | | | | |
|---|---|---|---|---|
| Model | Train:Test:Validate | RMSE | MAPE (%) | MAE |
| LN | **7:3** | **937.211** | **2.687** | **696.015** |
| | 8:2 | 1057.561 | 3.334 | 902.972 |
| | 9:1 | 1152.592 | 3.981 | 1084.636 |
| ARIMA | 7:3 | 2289.514 | 7.992 | 2133.657 |
| | 8:2 | 854.828 | 2.691 | 728.866 |
| | **9:1** | **372.325** | **1.197** | **325.546** |
| ETS | 7:3 | 1019.858 | 3.532 | 937.412 |
| | **8:2** | **367.357** | **1.103** | **297.673** |
| | 9:1 | 486.213 | 1.603 | 432.97 |
| RNN | 7:3 | 1.5398 | 0.603 | 1.209 |
| | 8:2 | 1.465 | 0.541 | 1.105 |
| | **9:1** | **1.45** | **0.524** | **1.095** |
| GRU | **7:3** | **131.149** | **0.369** | **98.491** |
| | 8:2 | 135.706 | 0.388 | 104.839 |
| | 9:1 | 137.436 | 0.383 | 104.969 |
| Stacking Model | 7:3 | 1928.668 | 1750.873 | 6.751 |
| | **8:2** | **493.815** | **372.973** | **1.407** |
| | 9:1 | 521.951 | 440.77 | 1.636 |
| LSTM | 7:3 | 138.613 | .405 | 108.009 |
| | 8:2 | 155.868 | 0.447 | 120.601 |
| | **9:1** | **136.237** | **0.387** | **106.114** |
| MLP | **7:3** | **?** | **?** | **?** |
| | 8:2 | ? | ? | ? |
| | 9:1 | ? | ? | ? |

TABLE 2. EUR-VND Dataset's Evaluation

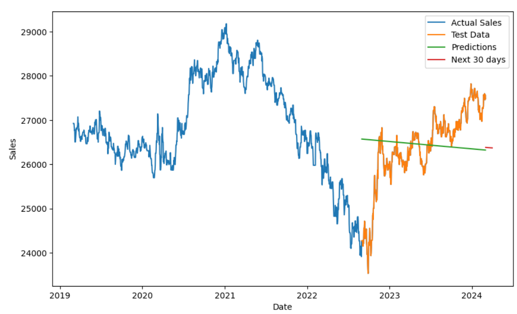


FIGURE 10. Linear model's result with 7:3 splitting proportion

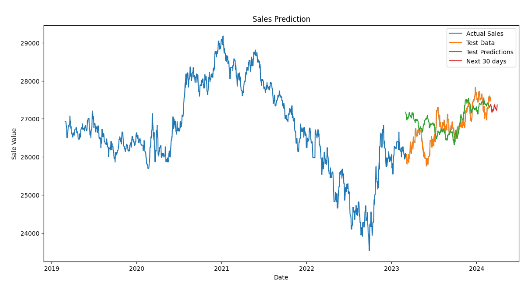


FIGURE 11. Stacking model's result with 8:2 splitting proportion

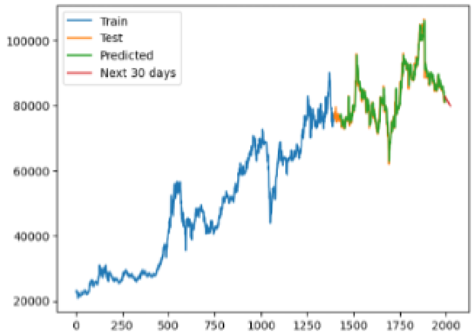


FIGURE 12. GRU model's result with 7:3 splitting proportion

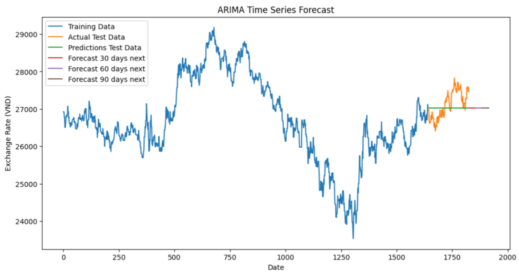


FIGURE 13. ARIMA model's result with 9:1 splitting proportion

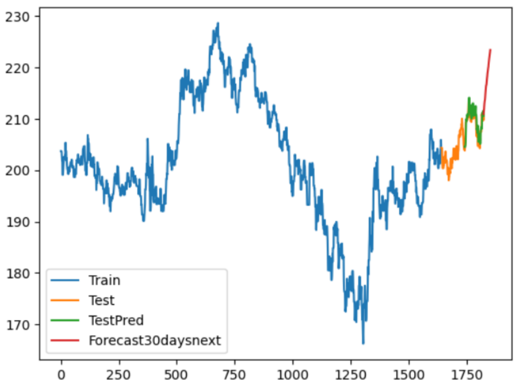


FIGURE 14. RNN model's result with 9:1 splitting proportion

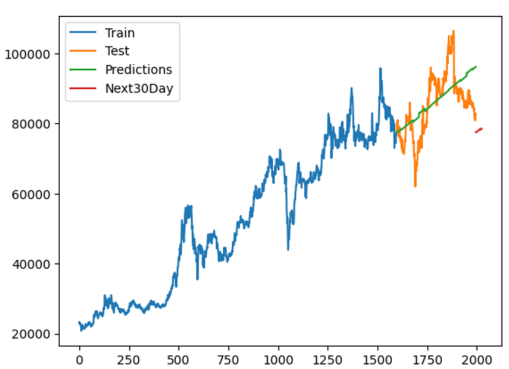


FIGURE 15. DLM model's result with 8:2 splitting proportion

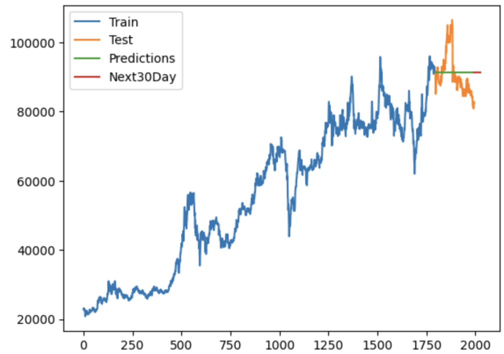


FIGURE 16. SES model's result with 9:1 splitting proportion

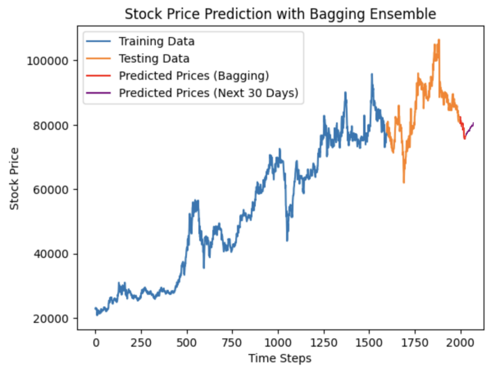


FIGURE 17. Bagging-GRU model's result with 8:2 splitting proportion

### C. GBP-VND DATASET

| Model | Train:Test:Validate | RMSE | MAPE (%) | MAE |
|---|---|---|---|---|
| LN | 7:3 | 1625.412 | 4.390 | 1277.908 |
| | **8:2** | **950.107** | **2.598** | **804.7** |
| | 9:1 | 1165.593 | 3.26 | 1026.264 |
| ARIMA | 7:3 | 2391.128 | 7.061 | 2160.296 |
| | 8:2 | 1865.457 | 5.445 | 1689.997 |
| | **9:1** | **546.792** | **1.575** | **490.115** |
| ETS | 7:3 | 1403.073 | 3.987 | 1213.524 |
| | **8:2** | **741.095** | **1.927** | **597.069** |
| | 9:1 | 1173.279 | 3.614 | 1126.355 |
| RNN | 7:3 | 1.465 | 0.544 | 1.107 |
| | 8:2 | 1.462 | 0.528 | 1.106 |
| | **9:1** | **1.354** | **0.476** | **1.023** |
| GRU | 7:3 | 174.233 | 0.437 | 132.637 |
| | 8:2 | 177.784 | 0.435 | 135.318 |
| | **9:1** | **158.711** | **0.363** | **115.036** |
| Stacking Model | 7:3 | 1620.452 | 1429.037 | 4.713 |
| | 8:2 | 817.089 | 680.45 | 2.234 |
| | **9:1** | **733.845** | **594.598** | **1.893** |
| LSTM | 7:3 | 187.121 | 0.477 | 145.069 |
| | 8:2 | 183.637 | 0.452 | 140.525 |
| | **9:1** | **179.470** | **0.44** | **139.233** |
| MLP | **7:3** | **?** | **?** | **?** |
| | 8:2 | ? | ? | ? |
| | 9:1 | ? | ? | ? |

TABLE 3. GBP-VND Dataset's Evaluation

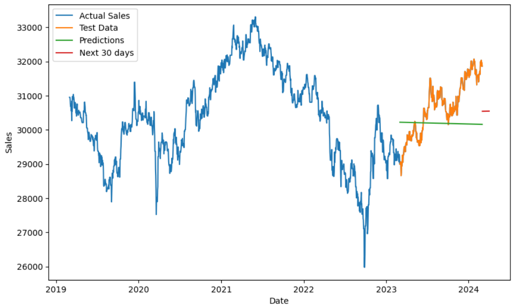


**FIGURE 18.** Linear model's result with 8:2 splitting proportion

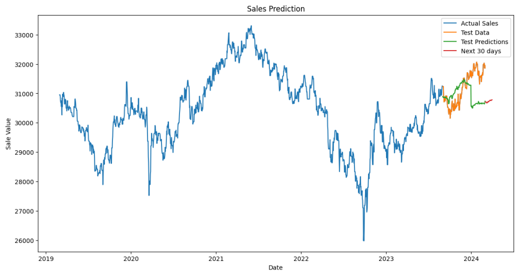


**FIGURE 19.** Stacking model's result with 9:1 splitting proportion


**FIGURE 20.** GRU model's result with 9:1 splitting proportion

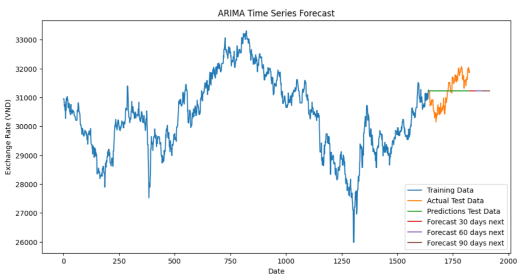


**FIGURE 21.** ARIMA model's result with 9:1 splitting proportion

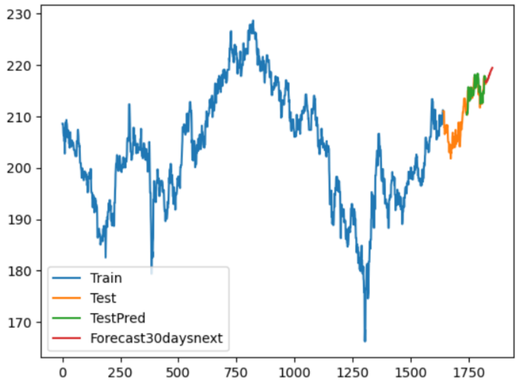


**FIGURE 22.** RNN model's result with 9:1 splitting proportion

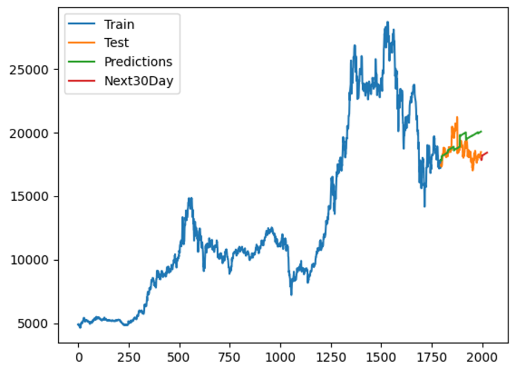


**FIGURE 23.** DLM model's result with 9:1 splitting proportion

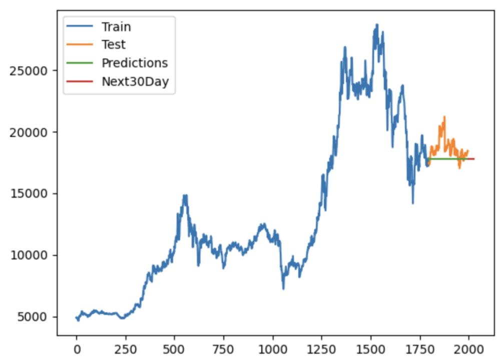


**FIGURE 24.** SES model's result with 9:1 splitting proportion

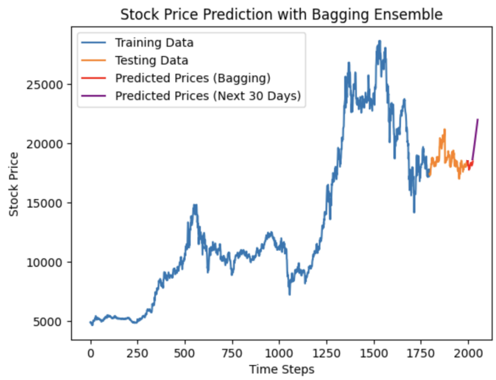


**FIGURE 25.** Bagging-GRU model's result with 9:1 splitting proportion

## D. JPY-VND DATASET

| Model | Train:Test:Validate | RMSE | MAPE (%) | MAE |
|---|---|---|---|---|
| LN | 7:3 | 15.557 | 8.46 | 14.665 |
| | 8:2 | 7.39 | 3.791 | 6.491 |
| | **9:1** | **4.749** | **2.399** | **4.078** |
| ARIMA | 7:3 | 8.645 | 4.284 | 7.6 |
| | 8:2 | 8.343 | 4.368 | 7.469 |
| | **9:1** | **2.749** | **1.264** | **2.173** |
| ETS | 7:3 | 9.371 | 4.698 | 8.327 |
| | 8:2 | 4.414 | 2.253 | 3.876 |
| | **9:1** | **3.989** | **1.688** | **2.92** |
| RNN | 7:3 | 1.592 | 0.663 | 1.1698 |
| | 8:2 | 1.635 | 0.755 | 1.29 |
| | **9:1** | **1.516** | **0.672** | **1.161** |
| GRU | 7:3 | 1.403 | 0.549 | 0.971 |
| | 8:2 | 1.149 | 0.473 | 0.811 |
| | **9:1** | **1.212** | **0.461** | **0.8** |
| Stacking Model | 7:3 | 1928.667 | 17.7 | 10.181 |
| | **8:2** | **6.196** | **4.921** | **2.858** |
| | 9:1 | 11.974 | 10.704 | 6.28 |
| LSTM | 7:3 | 1.795 | 0.791 | 1.39 |
| | 8:2 | 1.468 | 0.658 | 1.126 |
| | **9:1** | **1.31** | **0.536** | **0.931** |
| MLP | **7:3** | **?** | **?** | **?** |
| | 8:2 | ? | ? | ? |
| | 9:1 | ? | ? | ? |

**TABLE 4.** JPY-VND Dataset's Evaluation

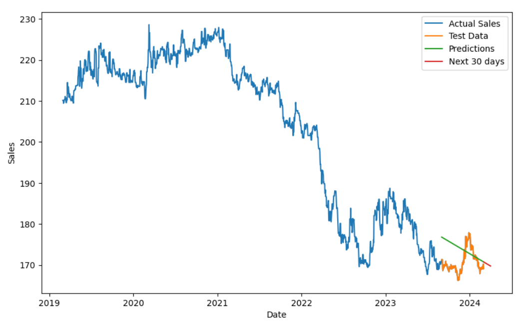


**FIGURE 26.** Linear model's result with 9:1 splitting proportion

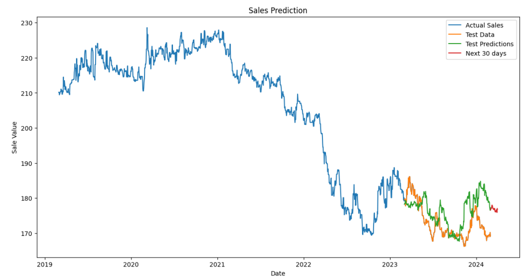


**FIGURE 27.** Stacking model's result with 8:2 splitting proportion

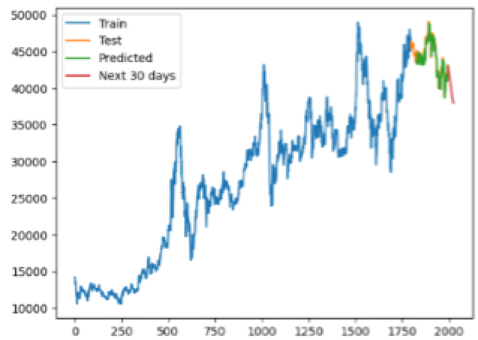


**FIGURE 28.** GRU model's result with 9:1 splitting proportion

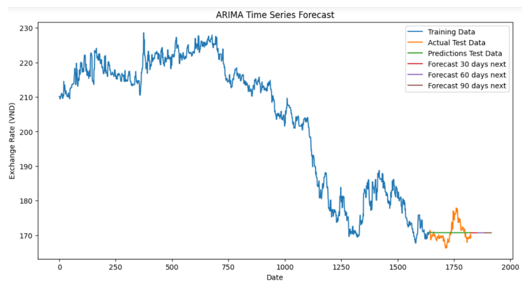


**FIGURE 29.** ARIMA model's result with 9:1 splitting proportion

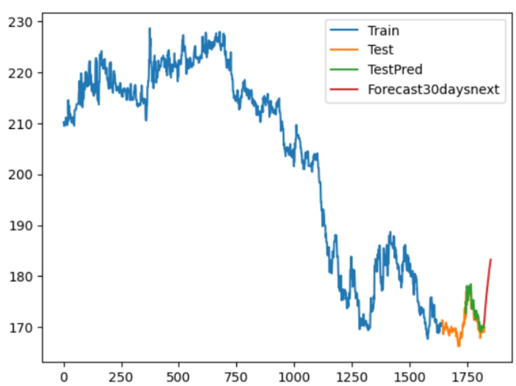


**FIGURE 30.** RNN model's result with 9:1 splitting proportion

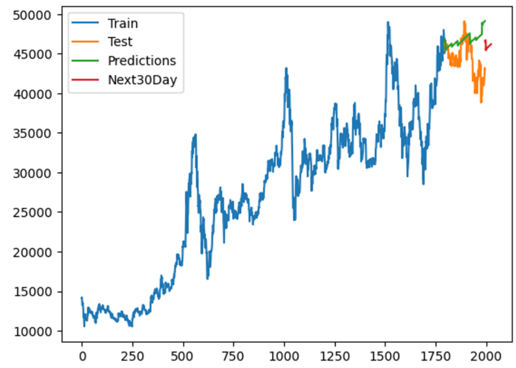


**FIGURE 31.** DLM model's result with 9:1 splitting proportion

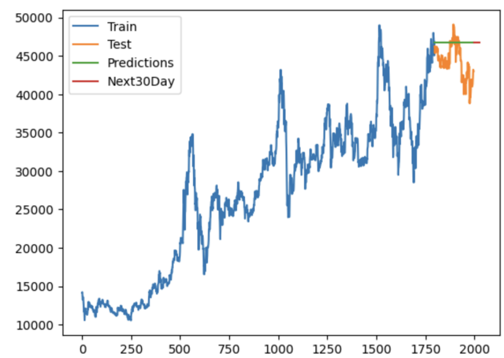


**FIGURE 32.** SES model's result with 9:1 splitting proportion

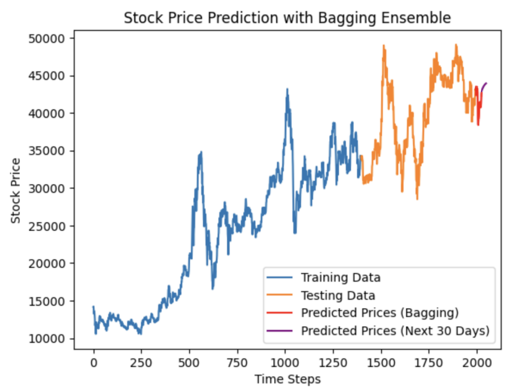


**FIGURE 33.** Bagging-GRU model's result with 7:3 splitting proportion

## VI. CONCLUSION

### A. SUMMARY

In the study, we developed and evaluated several models for forecasting currency price, leveraging different statistical, deep,v and machine learning techniques. The eight models used are Linear Regression, ARIMA, Exponential Smoothing (ETS), Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), Stacking Model and Multi-layer Perceptron (MLP). The assessment and comparison of forecasting methods highlighted that each technique possessed its own advantages and drawbacks. We use metrics like RMSE, MAE and MAPE to evaluate model accuracy. By comparing these evaluation metrics, we determined that <> are well-suited for forecasting currency price. These models performed more accurate future prices than the others.

### B. FUTURE PLANS

The above algorithms have demonstrated promising results in forecasting currency prices. However, it is necessary to enhance the model to achieve greater accuracy and reliability. To accomplish this, several key strategies can be implemented

• Enhancing the accuracy of the model. It includes improving data quality by cleaning and preprocessing data before being used in the model.

## ACKNOWLEDGMENT

First and foremost, we would like to express our sincere gratitude to **Assoc. Prof. Dr. Nguyen Dinh Thuan** and **Mr. Nguyen Minh Nhut** for their exceptional guidance, expertise, and invaluable feedback throughout the research process. Their mentorship and unwavering support have been instrumental in shaping the direction and quality of this study. Their profound knowledge, critical insights, and attention to detail have significantly contributed to the success of this research.

This research would not have been possible without the support and contributions of our mentors. We would like to extend our heartfelt thanks to everyone involved for their invaluable assistance, encouragement, and belief in our research. Thank you all for your invaluable assistance and encouragement.

## REFERENCES

[1] Liu, Pengfei, Ze Wang, Daoqun Liu, Jingyang Wang, and Tiezhu Wang. "A CNN-STLSTM-AM Model for Forecasting USD/RMB Exchange Rate." Journal of Engineering Research 11, no. 2 (June 1, 2023): 100079. https://doi.org/10.1016/j.jer.2023.100079

[2] "Foreign Exchange Currency Rate Prediction Using a GRU-LSTM Hybrid Network - ScienceDirect." Accessed April 14, 2024. https://www.sciencedirect.com/science/article/pii/S2666222120300083sec0022.

[3] Liu, Siyuan, Qiqian Huang, Mingchen Li, and Yunjie Wei. "A New LASSO-BiLSTM-Based Ensemble Learning Approach for Exchange Rate Forecasting." Engineering Applications of Artificial Intelligence 127 (January 1, 2024): 107305. https://doi.org/10.1016/j.engappai.2023.107305.

[4] Zhu, Qimian. "Forecasting the US Dollar/Euro Exchange Rate Based on ARIMA Model." Advances in Economics, Management and Political Sciences 15 (September 13, 2023): 369–78. https://doi.org/10.54254/2754-1169/15/20230951.

[5] Kamruzzaman, Joarder, and Ruhul Sarker. "ANN-Based Forecasting of Foreign Currency Exchange Rates" 3 (January 1, 2004).

[6] "Staying positive: challenges and solutions in using pure multiplicative ETS models | IMA Journal of Management Mathematics | Oxford Academic." Accessed: May 05, 2024. [Online]. Available: https://academic.oup.com/imaman/advance-article/doi/10.1093/imaman/dpad028/7475884

[7] "Exponential Smoothing- Definition, Formula, Methods and Examples," BYJUS. Accessed: May 05, 2024. [Online]. Available: https://byjus.com/maths/exponential-smoothing/

[8] "STAT481581 Introduction to Time Series Analysis.pdf." Accessed: May 15, 2024. [Online]. Available: https://math.unm.edu/ lil/Stat581/8-ets.pdf

[9] M. Rahman, Md. S. Hossain, T.-S. Junaid, M. Forhad, and M. Hossen, “Predicting Prices of Stock Market using Gated Recurrent Units (GRUs) Neural Networks,” vol. 19, pp. 213–222, Jan. 2019.

[10] M. Yurtsever, “Gold Price Forecasting Using LSTM, Bi-LSTM and GRU,” European Journal of Science and Technology, Dec. 2021, doi: 10.31590/ejosat.959405

[11] Buja, A., and Stuetzle, W. ”Observations on bagging”. University of Pennsylvania and University of Washington, Seattle. 2002.

[12] B. M. Henrique, V. A. Sobrero, and H. Kimura, “Comparison Of Fuzzy Time Series And ARIMA”, August 2019. Available:https://www.ijstr.org/final-print/aug2019/Comparison-Of-Fuzzy-Time-Series-And-Arima-Model.pdf. [Accessed 19 June 2023]. 4

[13] Jason Brownlee, “How to Create an ARIMA Model for Time Series Forecasting in Python”, November 18, 2023. Available:https://www.ijstr.org/final-print/aug2019/Comparison-Of-Fuzzy-Time-Series-And-Arima-Model.pdf.

[14] Jason Brownlee, “A Gentle Introduction to SARIMA for Time Series Forecasting in Python”, August 21, 2019.

[15] Alexandra M. Schmidt and Hedibert F. Lopes, ”Dynamic models”, 2019.

[16] Timothy O. Hodson, ”Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not”, 2022, https://doi.org/10.5194/gmd-15-5481-2022.

[17] Priya Pedamkar,”Support Vector Regression”, March 24, 2023. Retrieved from $https : //www.educba.com/support - vector - regression/?fbclid = IwAR0ibzdmqpaaDKq2 - Q4JRcjxQcVt - C7TrHNEc90q_tCSrn8rds9x2AG8Y78$

[18] Seok-Ho Han, Husna Mutahira, Hoon-Seok Jang, "Prediction of Sensor Data in a Greenhouse for Cultivation of Paprika Plants Using a Stacking Ensemble for Smart Farms", Applied Sciences, vol.13, no.18, pp.10464, 2023.