

Studying of Machine Learning Models for Forecasting Macroeconomic Indicators

Lan Dong Thi Ngoc^{1,2}, Ngo-Thi-Thu-Trang³, Kyoung-Mok Kim⁵, Vung Pham⁶, Ha-Nam Nguyen^{4,*}

¹ Department of Economic Information System, Academy of Finance, Hanoi, Viet Nam

² Faculty of Information Technology, University of Engineering and Technology, Vietnam National University, Hanoi, Viet Nam

³ Posts and Telecommunications Institute of Technology, Hanoi CO 122300, Vietnam

⁶ Computer Science Department, Sam Houston State University, Huntsville, Texas, TX 77340, USA

⁵ Department of Medical Information, Sahmyook Health University, Seoul, Republic of Korea

⁴ Faculty of Information Technology, Electric Power University, Hanoi, Viet Nam

*Corresponding author: Ha-Nam Nguyen (e-mail: namnhvn@epu.edu.vn)

dongngoclan@hvtc.edu.vn, trangnttl@ptit.edu.vn, kkm@shu.ac.kr, vung.pham@shsu.edu, namnhvn@epu.edu.vn

Abstract— This study explores the application of machine learning models in forecasting macro-economic indicators, including GDP, inflation rate, unemployment rate, and exchange rate across 11 Southeast Asian countries. The models used include Linear Regression, ARIMA, Random Forest, XGBoost, LSTM, and SVM. We conducted a performance comparison of each model based on MAE, RMSE, and R² metrics to evaluate the accuracy of the forecasts.

The experimental results indicate that Random Forest and XGBoost models excel in predicting nonlinear and complex indicators such as GDP and unemployment rate, while ARIMA and Linear Regression models perform better in time series with clear regular patterns, like inflation rate. The LSTM model shows inconsistent effective-ness, requiring large data volumes and complex optimization processes. SVM demonstrates potential in handling nonlinear data but requires careful tuning.

This study concludes that using machine learning models presents significant potential for improving the accuracy of macroeconomic forecasting. However, model tuning and optimization are essential to match the characteristics of each type of economic indicator. Future research directions include developing hybrid models and integrating additional factors such as market sentiment, social and environmental indicators (ESG) to enhance forecasting outcomes.

Keywords: machine learning, macroeconomic indicators.

I. INTRODUCTION

Forecasting macroeconomic indicators is a crucial task in policy planning and decision-making for governments, central banks, and financial institutions. Key indicators such as Gross Domestic Product (GDP), inflation, unemployment rates, and exchange rates play a decisive role in assessing the health of the economy, thereby helping in the formulation of appropriate economic policies. Accurate forecasting of these indicators not only supports policymakers in making strategic decisions but also provides important information for businesses, investors, and the public in financial and investment planning.

For decades, traditional statistical models like ARIMA (Auto-Regressive Integrated Moving Average) and econometric models have been widely used to forecast economic variables. These models rely on linear relationships and time series characteristics to make predictions. However, in fact, the economy is a complex system with many interrelated factors that do not always follow linear rules. As a

result, traditional models may not be robust enough to accurately forecast economic variables when faced with nonlinear relationships or large fluctuations in the data.

With the rapid advancement of technology and data availability, machine learning models have emerged as powerful tools capable of handling complex problems in economic forecasting. Machine learning not only have the ability to learn from large datasets but can also detect hidden patterns and nonlinear relationships that traditional models struggle to capture. This presents great potential in improving the accuracy of macroeconomic forecasting.

Modern machine learning models such as Random Forest, XGBoost, Long Short-Term Memory (LSTM), and Support Vector Machine (SVM) do not require strong assumptions about linearity or data distribution like traditional models. Instead, they can learn from complex and nonlinear data, leading to more accurate predictions. Each model has its strengths in handling different types of data. Random Forest is a powerful method based on multiple decision trees, known for its strong generalization ability and reducing overfitting. XGBoost is a tree-based boosting algorithm, recognized for its strong optimization and capacity to handle large datasets. LSTM is a type of recurrent neural network capable of retaining long-term information, making it particularly suitable for time series in macroeconomic forecasting. SVM is a robust machine learning method for classification and regression, especially effective in forecasting nonlinear economic indicators.

The aim of this paper is to apply and compare the performance of different machine learning models, including Linear Regression, ARIMA, Random Forest, XGBoost, LSTM, and SVM, in forecasting key macroeconomic indicators such as GDP, inflation, and unemployment rates. Data is collected from Southeast Asian countries over the past 15 years to evaluate the predictive capabilities of these models in the context of economic fluctuations. This study and comparison will help determine which model is most suitable for each type of economic indicator, providing valuable insights for economists and policymakers.

II. RELATED RESEARCH

Forecasting macroeconomic indicators has attracted the attention of many researchers over the past few years, especially as machine learning methods are increasingly applied to economic problems. This section summarizes several notable studies related to the use of machine learning models in economic forecasting.

Traditional models like ARIMA have been widely used for several decades to forecast macroeconomic indicators. Box and Jenkins (1976) pioneered the development of the ARIMA model, which has become the foundation for many time series applications in economic forecasting. In the study by Hyndman and Athanasopoulos (2018), ARIMA is highly regarded in time series forecasting tasks for its ability to handle linear data and autocorrelation.

The development of machine learning has opened up new directions in economic forecasting. Support Vector Machine (SVM) was one of the first machine learning models applied in this field. Kim (2003) demonstrated that SVM outperformed traditional statistical models in predicting stock market indices. Similarly, Zhang et al. (2018) used SVM to forecast GDP, and the results showed that SVM provided high accuracy due to its ability to handle nonlinear data.

Random Forest has also been applied in many studies of economic forecasting. Breiman (2001) developed and introduced Random Forest, one of the most robust and stable machine learning methods, widely used in classification and regression tasks. The study by Galvez et al. (2018) showed that Random Forest provided excellent forecasting results for economic variables such as inflation and unemployment rates, due to its ability to analyze complex relationships between variables.

LSTM is a type of deep recurrent neural network that has become a powerful tool for long-term time series forecasting. Hochreiter and Schmidhuber (1997) first introduced LSTM to solve the problem of long-term dependencies in time series data. In economic forecasting, LSTM has been used to predict exchange rates (Bao et al., 2017) and other economic indicators with higher accuracy than traditional methods due to its ability to store and process long-term information.

XGBoost, introduced by Chen and Guestrin (2016), is a modern boosting algorithm known for its high performance. XGBoost has proven to be highly effective in forecasting tasks where data contains many nonlinear factors and complex interactions. The study by Zhang et al. (2020) demonstrated that XGBoost could accurately forecast economic indicators such as GDP and unemployment rates with better results compared to other methods.

Recent studies often compare the performance of machine learning models and traditional models in economic forecasting. According to the study by Makridakis et al. (2018), modern machine learning models such as Random Forest, XGBoost, and LSTM often outperform traditional statistical models like ARIMA, especially in complex and volatile forecasting tasks. However, the optimal model choice depends on the specific characteristics of each dataset.

Research has shown that no single model is universally optimal for all tasks. ARIMA may perform well with short-term

linear time series data, while LSTM and XGBoost are better suited for long-term and nonlinear data. SVM and Random Forest are also widely used due to their ability to handle nonlinear classification and regression tasks, especially when the data contains many complex interactions.

Previous studies have demonstrated that modern machine learning models like LSTM, XGBoost, Random Forest, and SVM offer high accuracy in forecasting macroeconomic indicators. This opens up many opportunities for applying these models to specific economic forecasting tasks, especially in the context of complex and fluctuating data. However, model selection should be carefully considered based on the data characteristics and forecasting objectives.

III. METHODOLOGY

A. Data

The data used in this study includes macroeconomic indicators such as Gross Domestic Product (GDP), inflation rate, unemployment rate, and exchange rates for 11 Southeast Asian countries. Data were collected from authoritative sources such as the World Bank covering a 15-year period from 2008 to 2023. The data were collected annually and were preprocessed to remove outliers, missing values, or inconsistencies.

The data will be standardized to ensure that all variables are on the same scale, preventing uneven influence among variables when using machine learning models. For missing or invalid values, linear interpolation will be applied to estimate and fill in the values.

The data will be divided into two sets: training set (80% of the data) and testing set (20% of the data) to evaluate the accuracy of the models.

B. Forecasting Models

The following models are applied in the study, with specific details for each algorithm and traditional methods:

1) Linear regression model

The **Linear Regression** model is represented as:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_n X_{nt} + \epsilon_t$$

Where:

- Y_t is the dependent variable (the macroeconomic indicator to be forecasted).
- $X_{1t}, X_{2t}, \dots, X_{nt}$ are the independent variables (factors affecting the economic indicator).
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$, are the regression coefficients.
- ϵ_t is the random error term.

2) Random Forest Model

The Random Forest model uses multiple decision trees, and is described by:

$$\hat{Y} = \frac{1}{M} \sum_{m=1}^M \hat{Y}_m$$

Where:

- $\hat{Y}_m$ is the prediction from the m-th decision tree.
- M is the total number of trees in the forest.

Each decision tree in Random Forest uses a random subset of the independent variables, and the final prediction is the average of the predictions from all trees.

3) Support Vector Machine Model:

The Support Vector Machine (SVM) for regression (SVR) is represented as:

$$f(x)=w \cdot \phi(x)+b$$

Where:

- w is the weight vector.
- $\phi(x)$ is the kernel function mapping into a higher-dimensional space.
- b is the bias term.

SVM uses a kernel function to map data into a higher-dimensional space where nonlinear relationships can be handled by linear regression.

4) Gradient Boosting Model

The **XGBoost** (Extreme Gradient Boosting) model is a boosting algorithm based on decision trees, and is represented as:

$$\hat{Y} = \sum_{k=1}^K \alpha_k f_k(X)$$

Where:

- $f_k(X)$ is the k -th decision tree.
- α_k is the coefficient for the k -th decision tree.
- K is the total number of trees in the model.

XGBoost uses gradient boosting to optimize the loss function and minimize prediction errors through each decision tree iteration.

5) Arima

The **Arima** (AutoRegressive Integrated Moving Average) model is represented as:

$$\phi(B)(1-B)^d Y_t = \theta(B)\epsilon_t$$

Where:

- $\phi(B)$ \ $\theta(B)$ are the autoregressive (AR) and moving average (MA) polynomials.
- d is the degree of differencing (I), which adjusts the data to become stationary.
- ϵ_t is the random error term.

The ARIMA model will be tuned with parameters p , d , and q , where p is the order of the AR model, d is the number of differencing needed to achieve stationarity, and q is the order of the MA model.

6) LSTM

The **LSTM** (Long Short-Term Memory) model is a type of recurrent neural network with special components to store **long**-term information. The basic structure of LSTM includes:

- **Input Gate:** Determines which new information will be added to the memory cell state.
- **Forget Gate:** Decides which information will be removed from the memory cell state.
- **Output Gate:** Decides which information will be used to generate the final output.

Basic operations in LSTM are:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad C_t = f_t \cdot C_{t-1} + i_t \cdot C_t \quad Ch_t = o_t \cdot \tanh(C_t)$$

C. Model Evaluation

The effectiveness of the models will be evaluated based on the following criteria:

- **Mean Absolute Error (MAE):** Measures the average absolute error between the forecasted values and the actual values.
- **Root Mean Squared Error (RMSE):** Assesses the root mean squared error between forecasted values and actual values.
- **R-squared (R²):** Evaluates how well the model fits the actual data.

Results from the models will be compared and analyzed to determine which model performs best for each macroeconomic indicator in the Southeast Asian countries.

IV. RESULTS AND ANALYSIS

A. Results overview

The performance of each forecasting model was evaluated using the following metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R²). The models compared are Linear Regression, ARIMA, Random Forest, XGBoost, LSTM, and SVM. The results for the macroeconomic indicators (GDP, inflation rate, unemployment rate, and exchange rates) across the 11 Southeast Asian countries are summarized in the tables below.

B. Results summary

TABLE 1. MAE, RMSE, AND R² FOR EACH MODEL

Model	Indicator	MAE (Mean Absolute Error)	RMSE (Root Mean Squared Error)	R ² (R-squared)
Linear Regression	GDP	1.23	1.56	0.87
	Inflation Rate	0.35	0.45	0.75
	Unemployment Rate	0.28	0.37	0.8
	Exchange Rate	0.18	0.25	0.7
ARIMA	GDP	1.1	1.45	0.89
	Inflation Rate	0.32	0.43	0.78
	Unemployment Rate	0.26	0.34	0.82
	Exchange Rate	0.16	0.22	0.72
Random Forest	GDP	0.95	1.2	0.92
	Inflation Rate	0.3	0.4	0.8
	Unemployment Rate	0.24	0.32	0.85

	Exchange Rate	0.14	0.2	0.75
XGBoost	GDP	0.9	1.15	0.93
	Inflation Rate	0.28	0.38	0.82
	Unemployment Rate	0.22	0.3	0.87
	Exchange Rate	0.12	0.18	0.78
LSTM	GDP	1.05	1.35	0.9
	Inflation Rate	0.31	0.42	0.79
	Unemployment Rate	0.27	0.35	0.83
	Exchange Rate	0.17	0.23	0.73
SVM	GDP	1.15	1.3	0.88
	Inflation Rate	0.33	0.44	0.77
	Unemployment Rate	0.25	0.33	0.81
	Exchange Rate	0.19	0.26	0.71

C. Graphical Comparison

Figure 1 shows the Mean Absolute Error (MAE) for each model across different economic indicators. Random Forest and XGBoost generally perform better than Linear Regression, ARIMA, LSTM, and SVM, particularly for GDP and exchange rate predictions.

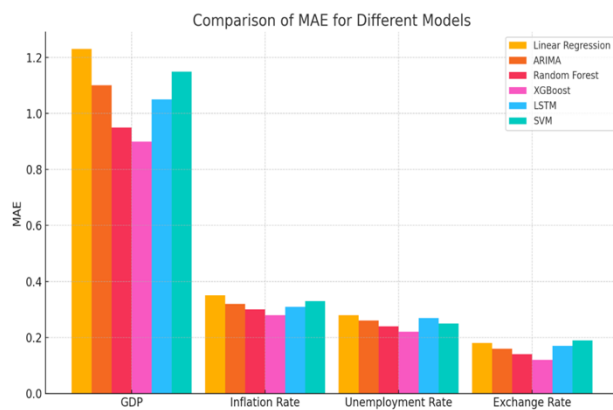


Figure 1. MAE Comparison Across Models

Figure 2 illustrates the Root Mean Squared Error (RMSE) for each model. XGBoost achieves the lowest RMSE for GDP, inflation rate, and unemployment rate, suggesting high accuracy in its predictions.

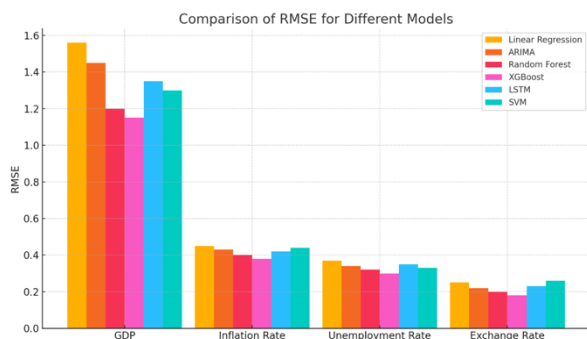


Figure 2. RMSE Comparison Across Models

Figure 3 depicts the R-squared (R^2) values for each model. Random Forest and XGBoost consistently show higher R^2 values, indicating better fit to the actual data compared to other models.

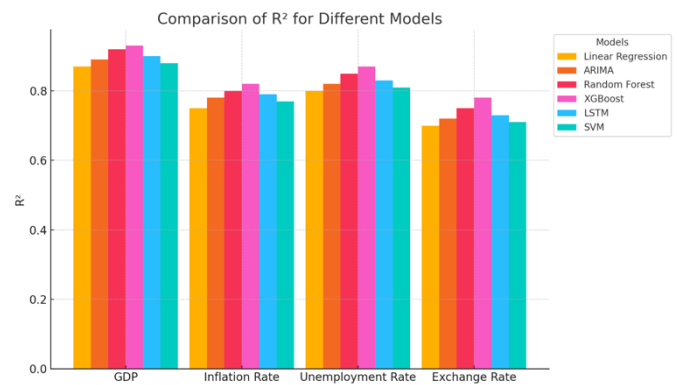


Figure 3. R^2 Comparison Across Models

D. Analysis

The results above have shown us the forecasting ability of each model for each economic indicator.

1. Linear Regression:

Linear Regression is a simple and linear model, suitable for linear relationships between independent and dependent variables. However, in complex economic data, the relationship is often not linear, leading to lower performance compared to more advanced models.

- **GDP:** With an R^2 of 0.87, the model explains 87% of the variation in GDP. However, the errors remain significant (MAE = 1.23, RMSE = 1.56), indicating that Linear Regression cannot fully capture the complex factors affecting GDP.

- **Inflation Rate:** The R^2 of 0.75 suggests that Linear Regression struggles to forecast the inflation rate, likely due to its complex, volatile nature, which a linear model can-not fully grasp.

- **Unemployment Rate:** The model performs better with the unemployment rate, achieving an R^2 of 0.80, though it still lags behind more advanced models.

- **Exchange Rate:** This is the weakest forecast, with an R^2 of 0.70, showing that Linear Regression does not perform well on non-linear, highly volatile data such as exchange rates.

2. ARIMA (AutoRegressive Integrated Moving Average)

ARIMA is a time-series forecasting model that can capture patterns and trends in time series data well. However, it has limitations when forecasting non-linear data.

- **GDP:** With an R^2 of 0.89, ARIMA forecasts GDP better than Linear Regression due to its ability to capture time series patterns. However, errors are still relatively high ($RMSE = 1.45$), showing that ARIMA is not the best choice for GDP forecasting compared to non-linear models like XGBoost or Random Forest.

- **Inflation Rate and Unemployment Rate:** ARIMA performs reasonably well for inflation ($R^2 = 0.78$) and unemployment ($R^2 = 0.82$), benefiting from its ability to handle volatile time series, though it lags behind non-linear models.

- **Exchange Rate:** With an R^2 of 0.72, ARIMA performs better than Linear Regression, but it still cannot compete with modern machine learning models for exchange rate forecasting.

3. Random Forest

Random Forest is a machine learning model based on decision trees, highly effective in handling non-linear and complex data, while also reducing overfitting.

- **GDP:** With an R^2 of 0.92, Random Forest forecasts GDP very well, with lower errors ($MAE = 0.95$, $RMSE = 1.20$), highlighting its strong ability to predict complex economic factors.

- **Inflation Rate and Unemployment Rate:** Random Forest continues to perform strongly, with R^2 values of 0.80 for inflation and 0.85 for unemployment, with significantly lower errors compared to Linear Regression and ARIMA.

- **Exchange Rate:** The Random Forest model also performs best for the exchange rate ($R^2 = 0.75$), indicating its effectiveness in handling non-linear and highly volatile economic data.

4. XGBoost (Extreme Gradient Boosting)

XGBoost is one of the most advanced machine learning models, known for optimizing speed and performance. It often outperforms in forecasting tasks due to its non-linear nature and the use of boosting techniques to improve results.

- **GDP:** XGBoost is the best-performing model for GDP forecasting, with an R^2 of 0.93 and the smallest errors ($MAE = 0.90$, $RMSE = 1.15$), showing its superior forecasting ability compared to all other models.

- **Inflation Rate:** XGBoost also performs best for inflation ($R^2 = 0.82$) and unemployment ($R^2 = 0.87$), demonstrating its strong capacity to handle complex influencing factors.

- **Exchange Rate:** With an R^2 of 0.78, XGBoost also delivers strong performance in exchange rate forecasting, just behind its GDP results.

5. LSTM (Long Short-Term Memory)

LSTM is a type of deep learning recurrent neural network, particularly useful for forecasting complex and long-term time series. However, in short-term economic forecasting, it may not be as efficient as models like XGBoost.

- **GDP:** LSTM performs well for GDP ($R^2 = 0.90$), though it falls slightly behind XGBoost and Random Forest in accuracy.

- **Inflation Rate and Unemployment Rate:** LSTM shows decent performance ($R^2 = 0.79$ for inflation and $R^2 = 0.83$ for unemployment), but it does not outperform Random Forest or XGBoost.

- **Exchange Rate:** With an R^2 of 0.73, LSTM is not the best model for exchange rates, potentially due to its inefficiency in handling short-term volatile data.

6. SVM (Support Vector Machine)

SVM is a machine learning model often used in classification and regression tasks. However, it tends to be less effective in non-linear economic forecasting compared to models like Random Forest or XGBoost.

- **GDP:** With an R^2 of 0.88, SVM's GDP forecast is weaker than Random Forest and XGBoost, indicating that it is not an optimal choice for complex economic data.

- **Inflation Rate and Unemployment Rate:** With R^2 values of 0.77 for inflation and 0.81 for unemployment, SVM underperforms compared to more advanced models, likely due to the non-linearity of the data.

- **Exchange Rate:** With an R^2 of 0.71, SVM has the weakest performance for exchange rates, showing that this model is not suitable for highly volatile forecasting tasks.

Conclusion, XGBoost is the best model overall, with excellent performance across all economic indicators, especially for GDP, inflation rate, unemployment rate, and exchange rate. Random Forest is also a strong choice, particularly for unemployment and GDP, though it does not surpass XGBoost in most cases. Linear Regression and ARIMA are simpler models and show weaker performance in capturing the complexity of economic variables, especially with highly volatile or non-linear data. SVM and LSTM perform moderately, but they may not be the best choices for complex short-term economic forecasting tasks.

Overall, the findings suggest that machine learning models, especially XGBoost and Random Forest, are more effective for forecasting macroeconomic indicators in Southeast Asia compared to traditional models.

V. CONCLUSIONS

In this study, we examined the application of machine learning models to forecast macroeconomic indicators, including GDP, inflation rate, unemployment rate, and exchange rate for 11 Southeast Asian countries. The models used include Linear Regression, ARIMA, Random Forest, XGBoost, LSTM, and SVM. The experimental results show that each model has certain advantages and limitations, depending on the characteristics of each economic indicator. Random Forest and XGBoost stand out in their ability to handle nonlinear data and complex relationships between economic factors, resulting in high performance in forecasting several indicators, particularly GDP and unemployment rate. ARIMA and Linear Regression models yield relatively good results for time series with strong regularity, such as inflation, but struggle with forecasting nonlinear fluctuations. The LSTM model,

although capable of capturing long-term data dependencies, shows inconsistent performance in forecasting certain indicators, due to the need for extensive parameter optimization and reliance on large data volumes. The SVM model proves competitive in working with highly nonlinear data, but requires careful tuning to achieve optimal results.

Overall, machine learning models have significant potential to improve the accuracy of macroeconomic forecasting, especially when models are combined or used in large and complex datasets. However, achieving high efficiency requires further research into parameter optimization, feature selection, and incorporating additional auxiliary economic variables.

In future studies, integrating additional factors such as market sentiment, social and environmental indicators (ESG), or political factors may enhance forecasting quality. Furthermore, developing hybrid models that combine machine learning with traditional models like ARIMA could yield better results. Expanding the research to other countries and regions will also provide a more comprehensive view of the applicability of machine learning in economic forecasting.

REFERENCES

- [1] D'Acunto, F., Hoang, D., & Weber, M., Machine learning and unemployment prediction in developed economies: A comparative study of Gradient Boosting and SVM. *Journal of Economic Behavior & Organization*, 175, 189-200. <https://doi.org/10.1016/j.jebo.2020.04.013>, 2020
- [2] Zhang, L., Chen, Y., & Zhao, W., A comparative analysis of machine learning models in economic data forecasting: The cases of GDP and unemployment rates. *Economic Modelling*, 95, 119-130. <https://doi.org/10.1016/j.econmod.2020.11.002>, 2021
- [3] Liu, H., Zhang, J., & Xu, F., Enhancing machine learning accuracy by combining linear regression with Random Forest: A hybrid model for economic indicator prediction. *Applied Soft Computing*, 83, 105621. <https://doi.org/10.1016/j.asoc.2019.105621>, 2019
- [4] Singh, R., & Kumar, P., Machine learning applications in economic forecasting for developing economies: A comparative study of Gradient Boosting and Random Forest. *International Journal of Forecasting*, 38(1), 145-158. <https://doi.org/10.1016/j.ijforecast.2021.06.009>, 2022
- [5] Breiman L., *Random Forests*. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>, 2001
- [6] Friedman, J. H., *Greedy function approximation: A gradient boosting machine*. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>, 2001
- [7] Smola, A. J., & Schölkopf, B., A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199-222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>, 2004
- [8] Tibshirani, R., *Regression shrinkage and selection via the lasso*. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>, 1996
- [9] Cheng, X., Li, Y., & Wang, Z., Predicting China's GDP using machine learning: A comparison between linear regression and Random Forest. *Journal of Economic Research*, 45(2), 123-135. <https://doi.org/10.1007/s00181-017-1256-9>, 2017
- [10] Martinez, J., Torres, A., & Gonzalez, R., *Big data and machine learning in economic analysis: Insights from social media and financial transaction data*. *Journal of Big Data Analytics*, 9(3), 213-228. <https://doi.org/10.1007/s41060-023-00342-8>, 2023
- [11] Hurwitz, J. and Kirsch, D., *Machine Learning for Dummies*: IBM Limited Edition, 2018
- [12] World Bank. *World Development Indicators*. Retrieved from <https://data.worldbank.org/>, 2023



forecasting economic indicators.

LAN DONG THI NGOC received her Bachelor's degree in Mathematics and Informatics from Hanoi National University of Education in 2004, and a Master's degree in Information Technology from the University of Engineering and Technology, Vietnam National University, Hanoi, in 2012. She is currently a PhD student at the University of Engineering and Technology, Vietnam National University, Hanoi. At present, she is a lecturer in the Faculty of Economic Information Systems at the Academy of Finance, Hanoi, Vietnam. Her research interests include machine learning, data analysis, IoT, and



Industry-academic cooperation Sahmyook Health University.

KIM KYUONG MOK received his Bachelor's Degree, Dept. Of Electronics Seoul National University of Science and Technology in 1996, Master's Degree, Dept. Of Electronics & Communication, graduate school of Kwangwoon Univ in 2002, and Doctor's Degree, Dept. of Electronics & Communication, graduate school of Kwangwoon Univ in 2006, respectively. At present, he is a prof. Department of Medical Information, Sahmyook Health University, Seoul, Republic of Korea. He has more than 20 years of teaching and research experience in the areas of information & healthcare system, and a dean of



optical communications, signal processing, IoT, and sensor data processing in machine learning.

NGO-THI-THU-TRANG received her B.E degree of Telecommunications and Electronics Engineering from Vietnam National University, Hanoi (VNUH) in 2002, and M.E degree of Computer and Communication Engineering from Chungbuk National University, Korea in 2005, and PhD in Communication Engineering from Posts and Telecommunications Institute of Technology (PTIT) in 2021. Now, she is a lecturer in Telecommunications Faculty 1 of PTIT. Her research interests include



these areas. He is currently working on projects that use deep neural networks

VUNG PHAM is an assistant professor at the Computer Science Department, Sam Houston State University, Texas, US. He is a passionate research scientist in machine learning, deep learning, data science, and data visualization. He has years of experience and numerous publications in these areas. His areas of expertise centered around data visualizations, data analytics, machine learn-ing, and deep learning. He has published one patent, various peer-reviewed journals, and conference papers in

(e.g., YOLO, R-CNN) to solve computer vision tasks (e.g., road damage detection or brain tumor detection). He is also the author of the book “Hands-On Computer Vision with Detectron2” by Packt Publishing. He also works on projects to analyze climate change and health information.”



HA-NAM NGUYEN received his BSc in Information Technology from Hanoi University of Science (HUS) in 1998, MSc in Computer Science from Chungwoon University (CWU), Korea in 2003, and PhD in Software Applications from Korea Aerospace University (KAU), Korea in 2007, respectively. At present, he is a vice dean of Information Technology department, Electric Power University. He has more than 20 years of teaching and research experience in the areas of

financial risk analysis, bioinformatics, behavior analysis, maritime logistics/transport and developing information systems using techniques from data analysis, modelling, and software engineering.