Acadlore Transactions on AI and Machine Learning
https://www.acadlore.com/journals/ATAIML

Comparative Analysis of Machine Learning Models for Predicting Indonesia's GDP Growth

Rossi Passarella[1*], Muhammad Ikhsan Setiawan[1], Zaqqi Yamani[2]

[1] Department of Computer Engineering, Faculty of Computer Science, Universitas Sriwijaya, 30129 Palembang, Indonesia
[2] Department of Information System, Faculty of Computer Science, Universitas Sriwijaya, 30129 Palembang, Indonesia

[*] Correspondence: Rossi Passarella (passarella.rossi@unsri.ac.id)

Received: 05-04-2025 **Revised:** 06-23-2025 **Accepted:** 06-29-2025

Citation: R. Passarella, M. I. Setiawan, and Z. Yamani, "Comparative analysis of machine learning models for predicting Indonesia's GDP growth," *Acadlore Trans. Mach. Learn.*, vol. 4, no. 3, pp. 157–173, 2025. https://doi.org/10.56578/ataiml040302.

© 2025 by the author(s). Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

Abstract: Accurate forecasting of Gross Domestic Product (GDP) growth remains essential for supporting strategic economic policy development, particularly in emerging economies such as Indonesia. In this study, a hybrid predictive framework was constructed by integrating fuzzy logic representations with machine learning algorithms to improve the accuracy and interpretability of GDP growth estimation. Annual macroeconomic data from 1970 to 2023 were utilised, and 19 input features were engineered by combining numerical economic indicators with fuzzy-based linguistic variables, along with a forecast label generated via the Non-Stationary Fuzzy Time Series (NSFTS) method. Six supervised learning models were comparatively assessed, including Random Forest (RF), Support Vector Regression (SVR), eXtreme Gradient Boosting (XGBoost), Huber Regressor, Decision Tree (DT), and Multilayer Perceptron (MLP). Model performance was evaluated using Mean Absolute Error (MAE) and accuracy metrics. Among the tested models, the RF algorithm demonstrated superior performance, achieving the lowest MAE and an accuracy of 99.45% in forecasting GDP growth for 2023. Its robustness in capturing non-linear patterns and short-term economic fluctuations was particularly evident when compared to other models. These findings underscore the RF model's capability to serve as a reliable tool for economic forecasting in data-limited and volatile macroeconomic environments. By enabling more precise GDP growth predictions, the proposed hybrid framework offers a valuable decision-support mechanism for policymakers in Indonesia, contributing to more informed resource allocation, proactive economic intervention, and long-term development planning. The methodological innovation of integrating NSFTS with machine learning extends the frontier of data-driven macroeconomic modelling and provides a replicable template for forecasting applications in other emerging markets.

Keywords: GDP growth; Economic prediction; Machine learning; Comparative analysis; Time series data

1 Introduction

Effective economic policymaking, especially in developing nations like Indonesia, depends on accurate projections of GDP growth. Being the biggest economy in Southeast Asia, Indonesia's stability and course of development have major regional and international consequences. Notwithstanding strong growth potential, the country faces particular economic difficulties, including vulnerability to world commodity price fluctuations, volatility in foreign investment, and the need to control domestic consumption and inflation while aiming at fair development. The dynamic and sometimes nonlinear character of these elements makes exact GDP growth prediction a difficult but important chore for legislators. Particularly in developing countries, traditional econometric models often find it difficult to reflect the natural nonlinear connections and dynamic changes seen in contemporary economies.

Reflecting the whole value of goods and services generated over a given period, GDP is the main indicator of a nation's economic activity. It is absolutely important for determining economic scale and growth, enabling international economic comparisons, and directing government policymaking. Furthermore, GDP usually serves as a gauge of public welfare and offers strategic guidance for choices on investments [1]. Supported by rather consistent inflation and less variation in production, Indonesia's economy has grown rather steadily in recent years. With a 5% increase in GDP right now, Indonesia is demonstrating a post-pandemic economic recovery and strengthening its importance in the world economy [2]. Given Indonesia's status as one of the biggest economies in the world, this success makes it an important topic in studies of the world economy and policies.

Nevertheless, events outside of government control, like natural disasters and world pandemics, can greatly affect economic development, challenging accurate future trend forecasts. By influencing supply chains, lowering consumer demand, and driving company closures or operational scale-backs, the COVID-19 epidemic, for example, seriously disrupted world economies, including Indonesia. As a result, Indonesia's GDP growth rate dropped dramatically in 2020 from that of the previous decade, which makes it much more difficult to forecast future GDP growth rates depending just on historical data [3].

The analysis and prediction of economic growth, especially GDP, is crucial in supporting data-driven economic decision-making processes. Traditional methods of predicting economic growth generally rely on econometric approaches such as time series Autoregressive Integrated Moving Average (ARIMA), vector autoregression (VAR), autoregressive (AR), or structural macroeconomic models. However, recent studies have found that machine learning prediction models are better than traditional econometric models because they can work with complicated data and find patterns that are hard for traditional methods to see [4, 5]. Ansari et al. [6] found that machine learning models, specifically MLP algorithms, were better at predicting GDP in Pakistan than ARIMA methods. This finding highlights the potential of machine learning in capturing complex data patterns and improving the accuracy of forecasting results. Yu [7] used traditional methods and machine learning algorithms to predict GDP in the United States and found that algorithms like gradient boosting and Long Short-Term Memory (LSTM) performed better and gave more accurate results than older methods, especially for short-term predictions. These results underscore the advantages of machine learning-based approaches in capturing the dynamics of fast-changing economic data and improving prediction accuracy in shorter time frames.

While machine learning models have shown a high level of accuracy in predicting GDP growth, there are still a number of challenges that need to be overcome, such as the need for high-quality data and the complexity of the economic system itself. In addition, the models' reliance on historical data may limit their ability to predict unprecedented economic events. Nevertheless, the integration of machine learning in economic forecasting still offers promising prospects for improving the precision and reliability of future GDP predictions [8, 9]. To achieve better prediction accuracy, one strategy that can be applied to this goal is to conduct a comparative evaluation of various machine learning models. Through the process of testing and comparing performance between models, the most optimal method can be identified in the context of predicting the value of GDP growth. Each machine learning model has its own advantages and limitations, which are strongly influenced by the complexity of the data being analysed and the need to interpret the prediction results. Therefore, comparative evaluation between models is a crucial step in identifying the most optimal method for forecasting GDP growth.

Based on this background, this research aims to develop a prediction model for Indonesia's GDP growth using machine learning algorithms with a focus on evaluating the model's performance in handling the complexity of Indonesia's macroeconomic data. Thus, this research is expected to contribute to improving the accuracy of Indonesia's economic growth prediction as well as providing data-based insights for more effective policymaking.

2 Literature Review

This section discusses various studies that apply machine learning approaches to predict GDP growth rates. The main focus is on how the approach is used to identify important patterns in economic data, as well as evaluating the effectiveness of the models in producing accurate predictions. Yoon [10] showed that the development and application of machine learning models, specifically gradient boosting and RF, are able to provide higher accuracy in predicting Japan's real GDP growth compared to traditional methods. A comparative evaluation of the performance of these models with forecasts from leading international institutions such as the International Monetary Fund (IMF) and the Bank of Japan (BOJ) shows both the advantages and limitations of this data-driven approach. One of the main objectives of this approach is to improve the accuracy of GDP growth predictions, which is crucial for policymakers to develop more time- and data-driven economic responses. This research also encourages the application of machine learning techniques in macroeconomics, given the challenges faced in forecasting complex and dynamic economic variables.

From an economic policy perspective, the ability to accurately predict GDP growth is of high relevance. In addition, the study has proven that the use of machine learning models is considered capable of reducing the level of uncertainty in the forecasting process. This approach also offers methodological innovation by utilising advanced algorithms and data-driven analysis, which not only improves the quality of predictions but also contributes to methodological development in economics and statistics. The empirical evidence from this research encourages openness to the utilisation of new technologies and widely available data in the process of economic analysis. The main contribution of this research lies in filling the gap in the literature related to the use of machine learning in GDP forecasting, particularly for the Japanese context, which is still relatively unexplored. As such, this research opens up opportunities for follow-up studies and broadens the understanding of the practical application of machine learning techniques in macroeconomic analysis.

Other studies have proven the effectiveness of this approach in economic forecasting. For example, Hossain et al. [11] developed a prediction model for GDP growth in Bangladesh using various economic indicators such as total investment, inflation rate, and unemployment rate. The study applied feature engineering using Lasso regression and compared the performance of RF and gradient boosting models. The results showed that the RF regressor achieved the highest accuracy, emphasising the potential of machine learning in economic forecasting and the importance of accurate GDP prediction in supporting policymaking.

Utama and Firinda [3] investigated the application of various machine learning algorithms to enhance the accuracy of predicting Indonesia's GDP growth. Utilising a comprehensive dataset of monthly economic and financial indicators from 2013 to 2023, the study employed models such as ElasticNet, RF, XGBoost, and Support Vector Machine (SVM) to analyse the relationship between these indicators and GDP fluctuations. The findings revealed that the models yielded robust performance, with Root Mean Square Error (RMSE) values indicating high accuracy during both normal and pandemic periods. Notably, the nowcasting exercise for Q4 2023 projected an average GDP growth of 5.13%. The research underscores the significance of specific economic indicators, including the Purchase of Durable Goods Index and Consumer Confidence Index (CCI), in influencing GDP outcomes. Through advanced interpretative techniques like Shapley value analysis, the study contributes valuable insights for policymakers in navigating economic challenges and formulating effective strategies.

Ahammad et al. [12] investigated the application of machine learning techniques to predict Bangladesh's GDP over a span of 43 years, utilising a dataset sourced from Wikipedia and the World Bank. The study examined nine key economic indicators, including inflation rate, unemployment, and remittance, to assess their impact on GDP. Various machine learning algorithms, such as K-Nearest Neighbours (KNN), RF, and Adaptive Boosting (AdaBoost), were employed, with KNN demonstrating the highest accuracy of 98.40%. The findings underscore the potential of machine learning to enhance economic forecasting by capturing complex nonlinear relationships within the data. It was concluded that early GDP predictions can facilitate informed policy decisions, highlighting the importance of incorporating advanced analytical methods for future economic assessments in Bangladesh. Sa'adah and Wibowo [13] also predicted GDP in Indonesia, demonstrating the efficacy of deep learning in predicting GDP growth, highlighting the robustness of LSTM and Recurrent Neural Network (RNN) models with 90% accuracy. This poses a challenge to improve accuracy using different approaches.

3 Data and Methods

This section provides a comprehensive description of the origin of the data used, as well as the process of collecting and preparing it in a format suitable for analysis. It also provides an explanation of the data collection process. In the methods section, the study discusses the types of algorithms and approaches used in data processing, including specific techniques used in analysing the data, such as classification and other statistical analyses. It explains the rationale for choosing each method and describes steps in data handling, from pre-processing to preparing for further analysis.

3.1 Data

Time series data obtained from World Bank Data were used in this study. The World Bank is a source of financial and technical assistance for developing countries and, since 2010, has provided open access through a web Application Programming Interface (API) that allows users to download a comprehensive collection of development data from various countries around the world [14]. The dataset used in this study includes 53 rows of data representing the period from 1970 to 2023 and consists of eight independent variables and one dependent variable.

The data obtained from the World Bank were initially available in separate forms for each variable. All variables were then compiled and combined into a unified dataset that forms the basis of analysis in this study. This dataset was chosen for its completeness and large number of data values, which allows for the exploration of relationships between variables and the identification of hidden patterns in the data that may not be detected through conventional approaches [15]. Information about the data used and its description are shown in Table 1.

Table 1. Description of features in the dataset

Data	Description
GDP growth (annual %)	Annual percentage growth rate of GDP
GDP (current LCU)	GDP value
Final consumption expenditure (current LCU)	Household consumption expenditure
Foreign direct investment, net inflows (% of GDP)	Equity direct investment in the economy
Expenses (current LCU)	Government expenses
Exports of goods and services (current LCU)	Value of exports
Imports of goods and services (current LCU)	Value of imports
GDP deflator	A measure of the price level of all goods and services produced in the economy
Real GDP	Real GDP value

The input variables in this study were selected based on the standard formula for calculating GDP, which incorporates expenditure components such as consumption, investment, government spending, exports, and imports. These variables were included to ensure that the machine learning models capture the primary economic drivers influencing GDP growth. Additionally, GDP deflator and real GDP were added to account for inflationary effects and adjusted economic output, further aligning the input features with economic theory.

Figure 1 shows that the distribution of the GDP growth variable in the dataset shows a pattern that tends to be asymmetrical and right-skewed, which indicates that most of the economic growth data is below the average value, while there are a small number of extreme values that are much higher. This tendency is shown by the series of GDP growth values concentrated in the range of 4% to 8%, with the peak of the distribution (mode) occurring at around 6%, while the maximum and minimum values experience a wider spread. The average GDP growth of 5.37% is slightly shifted from the centre of the distribution, indicating that the data has a positive skew. The variation in values is quite large, which is evident by the fact that some extreme values, both negative and positive, are far from the centre.

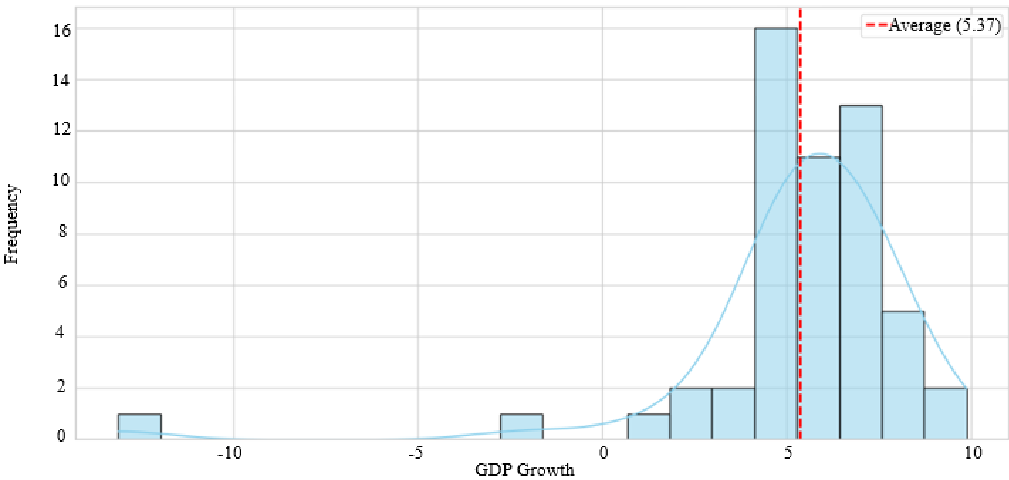

Figure 1. Variable distribution of target GDP growth

Based on the results of the correlation visualisation in the form of a heatmap in Figure 2, it can be seen that most macroeconomic variables, such as GDP, expenditure, expense, export, import, GDP deflator, and real GDP, have a very high correlation with each other, with values close to 1.00. This relationship indicates that these variables move linearly and are closely related, possibly because they are all components of GDP or are calculated using interconnected formulas. For example, GDP has a perfect correlation with expenditure, expenses, and exports, indicating a structural dependency between these variables. However, in terms of the correlation with GDP growth, it can be seen that almost all variables have a weak correlation and even tend to be negative.

The highest correlation value to GDP growth is only 0.25, which comes from the investment variable, while the rest range from -0.20 to 0.28, indicating that the linear relationship between GDP growth and other variables is relatively weak. This indicates that economic growth (GDP growth) is not directly influenced linearly by the absolute value of GDP, exports, or government expenses, but rather influenced by the dynamics of change or a combination of complex and nonlinear variables. This reinforces the rationale for using machine learning models, especially nonlinear models such as MLP and XGBoost, as ordinary linear models may not be able to capture the complex relationships between variables on the target variable GDP growth. In addition, this heatmap also provides important information regarding the potential for very high multi-collinearity between features, which needs to be considered in model selection and interpretation of regression results.

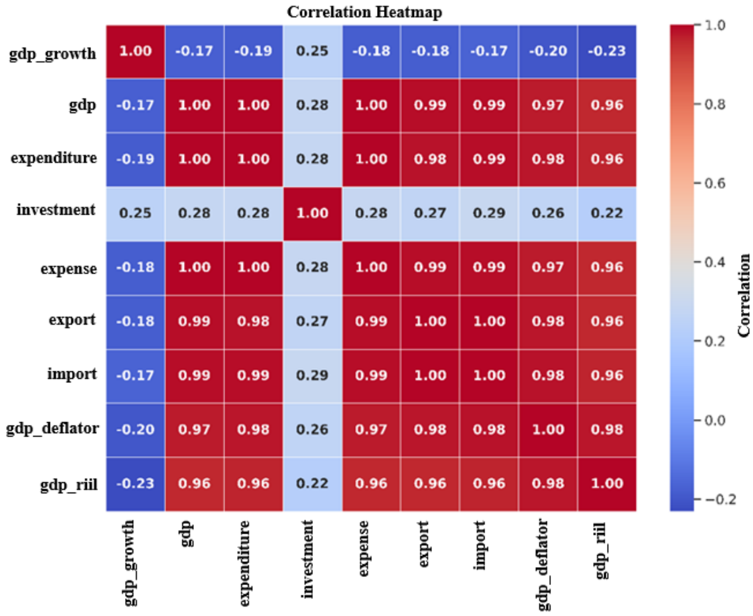

Figure 2. Heatmap of the correlation matrix

3.2 Data Pre-Processing

At this stage, after retrieving data from the World Bank website, a pre-processing step was carried out to improve the quality of the data before further analysis or modelling [16]. The data were initially filtered. The dataset obtained consisted of various countries. As the scope of this study is limited to the Indonesian economy, the dataset was systematically filtered to retain only records pertaining to Indonesia. The data used in this study includes various economic variables that contribute to GDP growth.

3.2.1 Missing value

Before further processing, missing values were checked. As can be seen in Figure 3, the linear interpolation method was applied to the government expense variable to handle data with missing values. This method was chosen because it is able to estimate the missing values based on the trend of the data before and after, thus maintaining the continuity of the data without causing significant distortion in the original pattern of the dataset.

3.2.2 Fuzzification and labelling process

In this study, a hybrid approach was applied that combines the NSFTS method with machine learning algorithms to predict Indonesia's GDP growth for the year 2023. The goal of this approach is to leverage the temporal pattern recognition capabilities of NSFTS while harnessing the strength of machine learning in modelling complex multivariate relationships. The first step involves fuzzifying the historical GDP growth data. The Interquartile Range (IQR) method was employed to divide the GDP growth data into three fuzzy intervals representing low (A0), medium (A1), and high (A2) categories. This process produced nine fuzzy-labelled columns corresponding to each of the original raw features, enabling a categorical representation of their respective value distributions. Both the IQR-based fuzzification and the NSFTS modelling were performed separately for four distinct economic periods: 1970–1997 (pre-crisis), 1998–2007 (crisis & recovery), 2008–2019 (stable), and 2020–2023 (pandemic). This segmentation was designed to account for structural changes and regime shifts in the Indonesian economy over time.

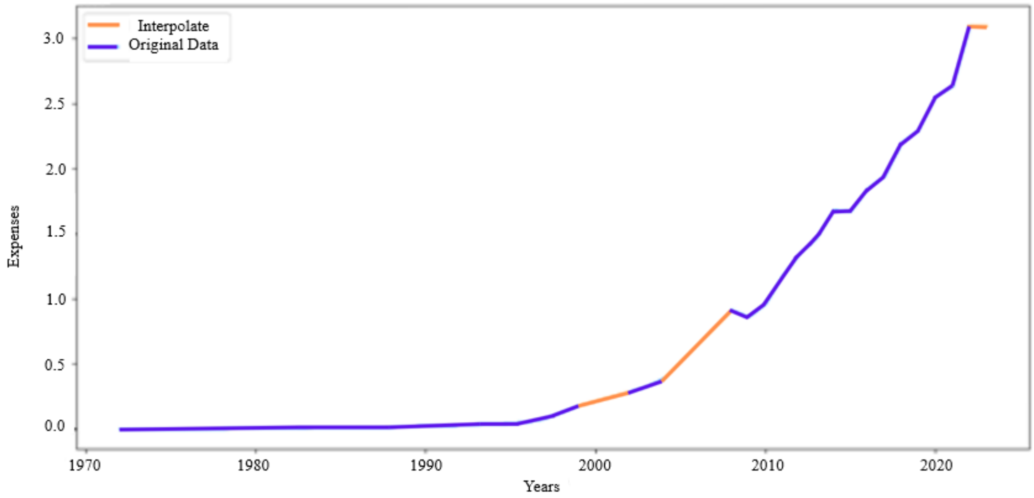

Figure 3. The result of filling in missing values on the government expense variable

3.2.3 Feature integration for machine learning

Following the fuzzification, the NSFTS method was applied to capture temporal dependencies within the GDP growth data:

• Fuzzy Logical Relationship Groups (FLRGs) were constructed based on historical data, segmented into distinct economic periods (pre-crisis, crisis & recovery, stable, and pandemic).

• Using the established FLRGs, predictions of the GDP growth category for the year 2023 were generated.

The predicted fuzzy label from NSFTS was stored in a separate column, which reflects the anticipated fuzzy state of GDP growth based solely on temporal patterns. For the machine learning model, a dataset consisting of 19 input features was constructed, structured as follows:

• Nine raw data columns: These include primary economic indicators such as GDP, expenditure, investment, expense, export, import, GDP deflator, and others.

• Nine IQR-based labelled columns: Each raw feature is accompanied by its fuzzified categorical label derived through the IQR method, representing its respective state (A0, A1, and A2).

• One NSFTS predicted label column: This column provides the NSFTS-based fuzzy prediction for GDP growth, offering an additional temporal insight to the model.

This hybrid feature set is designed to combine numerical strength (via raw data), categorical insight (via IQR-based fuzzification), and temporal pattern learning (via NSFTS). This hybrid representation aims to provide the machine learning model with a more comprehensive understanding of the underlying data structure—leveraging both the precision of numerical variables and the conceptual abstraction afforded by fuzzy logic. The inclusion of fuzzy features is not intended to replace the original inputs, but rather to augment them with interpretable, domain-informed categorical perspectives. This approach is particularly beneficial in economic contexts, where data is often non-stationary, heteroskedastic, and subject to concept drift. Fuzzy logic, with its inherent capacity to model uncertainty and ambiguity, offers a flexible framework for representing such complexity [17]. While this technique is relatively uncommon in time-series feature engineering, its potential to enhance predictive accuracy by uncovering nonlinear and semantically rich patterns justifies its application in this study.

3.2.4 Addressing heterogeneity

An IQR-based labelling approach customised to every variable and segmented historical period was used to fit the unique statistical traits of economic variables. Based on Indonesia's macroeconomic path, the data were split into four economic phases: pre-crisis, crisis & recovery, stable, and epidemic. This approach addresses both cross-variable and temporal heterogeneity by allowing the classification of varying values into "low", "medium", and "high" categories based on their distribution inside each temporal segment.

Principal Component Analysis (PCA) was purposefully excluded in this study, since its homogeneity and linear relationship assumptions contradict the heterogeneous and nonlinear character of macroeconomic indicators for Indonesia. PCA may hide the substantive economic meanings of important variables, including GDP, exports, and investment, each of which has different and understood relevance. Thus, it is advisable to preserve both interpretability and structural relevance by individual labelling depending on distributional properties and historical segmentation.

Specially to handle structural changes and regime shifts in economic data over time, the NSFTS method was used. NSFTS lets the model learn dynamic transition patterns and consider temporal variance in economic behaviour by building fuzzy logical relationships (FLRs) and FLRGs for every time segment independently. Particularly suited for non-stationary datasets, this method does not assume stationary relationships across the whole time span.

3.3 Methods

In Figure 4, it can be seen that this research begins with data collection with data obtained from the World Bank [16]. The data collected include macroeconomic indicators that are believed to influence the value of GDP growth. Once the data were obtained, a pre-processing stage was performed with several steps, including handling missing values using an interpolation method, as well as standardising and transforming the data to ensure a uniform scale. In this study, outliers in the dataset were not manipulated to maintain the original characteristics of the data. To further enhance the model's performance, new pertinent features were developed. The data exploration stage was conducted by visualising the data distribution to understand the patterns in the dataset. In addition, correlation analysis between variables was conducted to identify factors that have a significant relationship with the GDP growth rate.

This work compares the accuracy of six machine learning models—RF, XGBoost, SVR, Huber Regressor, DT, and MLP—in forecasting GDP growth. These models were chosen depending on their different methodological benefits in managing nonlinear and complex economic data:

• RF and XGBoost: These ensemble techniques were selected for their extraordinary adaptability in capturing intricate, nonlinear data patterns without needing rigorous causal relationship assumptions like classical econometric models. Their natural application of cross-validation techniques essentially avoids overfitting and enables exact hyperparameter tuning, producing more consistent and reliable forecasts [10].

• SVR: This model was chosen because of its great generalising ability, especially in macroeconomic forecasting when sample sizes are sometimes small. Its use of an epsilon-insensitive margin and regularisation balances model complexity with predictive accuracy, strengthening it even in noisy data [18].

• Huber Regressor: This model was included to especially handle possible outliers in macroeconomic data. Unlike ordinary least squares, the Huber loss function adaptively transitions its behaviour—quadratic for small errors and linear for large errors—providing superior robustness against extreme deviations without totally discounting their influence, greatly improving model stability and dependability [19].

• DT: Mostly driven by its natural interpretability and capacity to explicitly model nonlinear and hierarchical decision-making structures—often found in economic systems—this model was included. Moreover, its structural character makes DTs naturally more resistant to outlier influence than conventional linear models [20].

• MLP: This model was selected to capture the complex, nonlinear, and sometimes nonstationary patterns unique to macroeconomic time series. Particularly during times of economic stability, MLPs—a type of feedforward artificial neural network—have shown great predictive accuracy in past GDP forecasting studies and are quite well-suited for revealing latent connections between variables [21].

Although LSTM networks and deep learning models are strong for time-series prediction, great thought was given to including them in this work. Usually demanding large data depth and computational resources, LSTM presents difficulties given the limited historical span and consistency of macroeconomic datasets, especially for a developing economy like Indonesia. The present model selection balances predictive performance with practical data needs by giving methods that perform robustly and efficiently with the given data features top priority. Future research should still benefit much from investigating advanced deep learning architectures such as LSTM, particularly as data availability increases or more complex data augmentation methods improve.

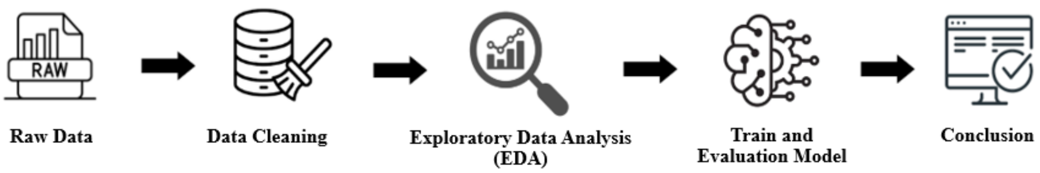

Figure 4. Overview of the research workflow

3.3.1 Data split and standardisation

The dataset was prepared for model training and evaluation following the exhaustive data preparation stages. There were 54 annual observations in the whole dataset, which ran from 1970 to 2023. A chronological split strategy was used to preserve the temporal integrity of the time series data for the aim of model development and evaluation. Designed as the training set—53 observations—the data from 1970 to 2022 were used to teach and maximise the parameters of the several machine learning models. The observation corresponding to the year 2023 was reserved as the test set to evaluate the predictive performance of the trained models on unseen, out-of-sample data, thereby simulating a real-world forecasting scenario.

Data standardisation employing the Z-score normalising technique was used to guarantee that no single feature unduly dominates the learning process due to differences in scale and unit and to improve the convergence and performance of some machine learning algorithms (e.g., SVR and MLP). The Z-score standardising converts the data such that, computed as Eq. (1), its mean is zero and its standard deviation [22] is one:

$$Z = \frac{(x - \mu)}{\sigma} \tag{1}$$

where, x is the original value of the data, μ is the mean of the variable, and σ is the standard deviation.

It is important to underline that the mean (μ) and standard deviation (σ) parameters for standardising came just from the training set (1970–2022). These same criteria then were used to standardise the test set (2023) as well as the training set itself. This method guarantees an objective assessment of the real predictive ability of the model by preventing data leakage from the future (test set) into the training process.

Although a fixed chronological split is a common practice in time series forecasting for its simplicity and direct relevance to real-world out-of-sample prediction, it can be recognised that more dynamic validation strategies, such as rolling window cross-valuation, could offer deeper insights on model robustness and stability across varying economic regimes and structural changes over such a long historical period. Future research should carefully consider the adoption of such dynamic validation strategies to comprehensively assess model performance under evolving economic conditions.

4 Results and Discussion

This section presents the results of experiments conducted to analyse the performance of various machine learning models in predicting GDP growth. The models compared include RF, SVR, XGBoost, Huber Regressor, DT, and MLP. The analysis is based on evaluation metrics such as MAE, RMSE, and R-squared (R^2).

4.1 Modelling and Analysis Training

The machine learning models were trained in this study using past GDP growth data ranging from 1970 to 2022. Grid search was used only to methodically find the optimal parameter combination for every model in hyperparameter optimization. Through cross-valuation, all possible parameter value combinations inside a predefined search space were evaluated by grid search, assessing performance based on metrics such as MAE or R^2. This whole process guarantees that the models not only fit the training data but also generalize robustly to unprocessed data. Although grid search can be computationally costly, particularly in relation to more effective strategies like random search or Bayesian optimization, its methodical and thorough character is absolutely critical for this work. Unlike probabilistic or random methods, grid search ensures that no optimal parameter setting is neglected inside the defined grid, greatly improving reproducibility and transparency—important features of empirical research. Moreover, using different tuning techniques would have required significant reorganizing of the proposed experimental setup and could bring unwanted variability, compromising the consistency needed for strong model comparisons between algorithms. Thus, in the framework of this work, which gives consistent evaluation across several machine learning models top priority for dependable and interpretable hyperparameter optimization, grid search was chosen as the approach. Table 2 lists the particular values best fit for every model.

In Table 3, the performance of each model is shown using the MAE, RMSE, and R^2 evaluation metrics.

As shown in Table 3 and Figure 5, after applying the chronological data split method, the evaluation results of model performance show that there is a notable shift in the evaluation metrics, particularly MAE, RMSE, and R^2. This analysis is essential because the chronological split more accurately reflects real-world forecasting scenarios, especially in time series problems like GDP growth prediction, where historical data is used to project future outcomes.

With a minimum MAE of 1.033 and a rather low RMSE of 2.032, RF showed among the models the best overall performance, indicating strong predictive accuracy. Its R^2 value of 0.116 consistently reduced prediction errors even if it suggests only modest variance explanation. Conversely, XGBoost obtained an R^2 of 0.111 and an MAE of 1.526, lower than those of RF. With MAE values of 1.203 and 1.188, respectively, SVR and Huber Regressor showed similar performance levels and moderate R^2 values, indicating balanced trade-offs between error and variance explanation. With an MAE of 1.273, a high RMSE of 2.144, and a low R^2 of 0.015, the DT model displayed poorer performance, highlighting its tendency toward overfitting and poor generalisation. With the highest MAE (1.643), the highest RMSE (2.181), and a negative R^2 (-0.018), MLP lastly underperformed compared to a basic baseline. RF turned out to be the most dependable model overall since it efficiently combined fuzzy logic-based characteristics with raw data. These results emphasize the need for model selection and validation even if machine learning models show promise for GDP growth prediction, particularly in small and complicated macroeconomic datasets. The absence of key external economic variables in the dataset may partially explain the variance not captured by the proposed model. Further study, including global elements, could help to provide more complete knowledge.

Table 2. Parameters of the machine learning models

Models	Parameters			
RF	n_estimators=50	max_depth = 7	min_samples_split=2	min_samples_leaf=9
XGBoost	early_stopping=True	max_depth = 1	n_estimators=500	learning_rate=0.1
SVR	kernel = linear	C=1	gamma = scale	epsilon=0.01
Huber Regressor	alpha =0.7	epsilon = 1	max_iter=50	tol=1e-5
DT	max_depth = 7	min_samples_split = 2	min_samples_leaf=6	criterion= absolute_error
MLP	early_stopping=true	hidden_layer_sizes=(64,32)	max_iter=200	alpha=0.1

Table 3. Performance of the machine learning models

Models	MAE	RMSE	R^2
RF	1.033	2.032	0.116
XGBoost	1.526	2.037	0.111
SVR	1.203	2.026	0.12
Huber Regressor	1.188	1.977	0.162
DT	1.273	2.144	0.015
MLP	1.643	2.181	-0.018

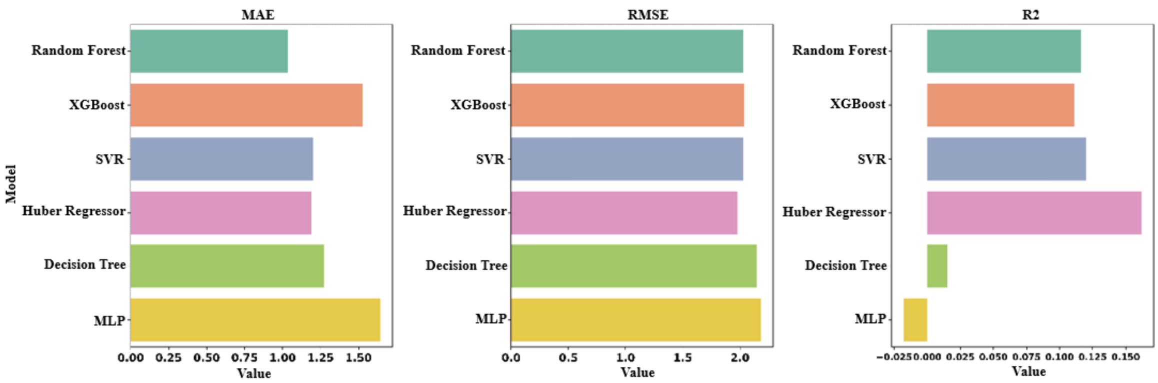

Figure 5. Evaluation metrics

Following the training and assessment of the six machine learning models, as shown in Table 4 and Figure 6, every model was applied to forecast Indonesia's GDP growth for 2023. To gauge the accuracy of every model, the expected values were then matched with the real GDP growth value (5.48106%) for 2023. With the highest accuracy of 99.45%, the RF model showed the best performance based on the results, followed by SVR (97.93%)

and the Huber Regressor (98.92%). These three models indicated their great capacity to fit to current data patterns since they generated forecasts that quite matched the real value.

Table 4. GDP growth prediction results for 2023

Models	Predicted GDP Growth for 2023 (%)	Actual Value of GDP Growth for 2023 (%)	Accuracy
RF	5.076019	5.48106	99.45%
XGBoost	5.927236		82.58%
SVR	5.152487		97.93%
Huber Regressor	4.993409		98.92%
DT	5.692571		87.23%
MLP	5.242731		96.14%

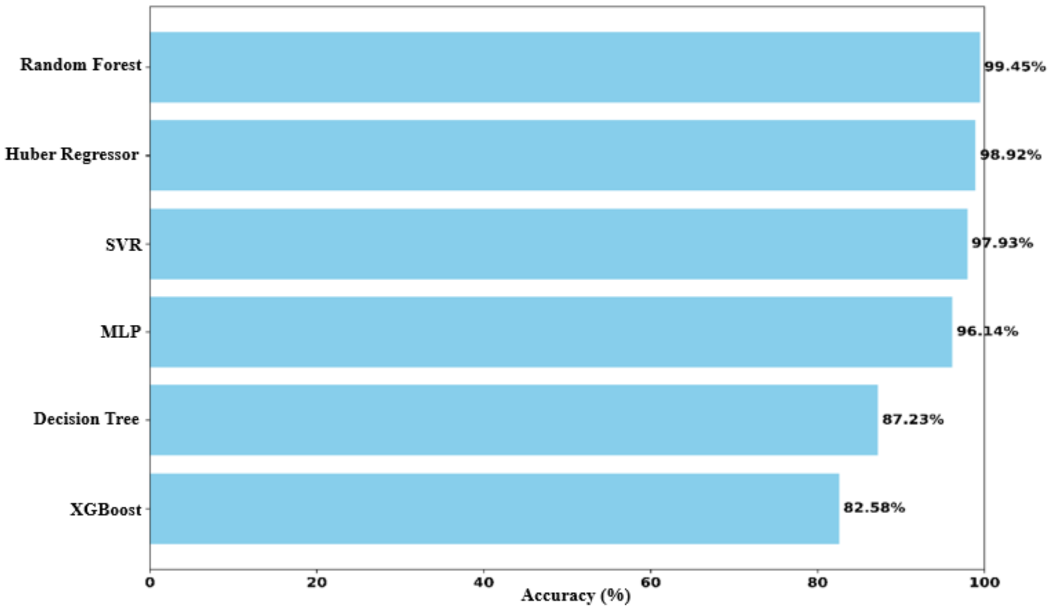

Figure 6. GDP growth prediction results for 2023

However, when these findings are linked to the models' prior performance on the test dataset (based on a chronological train-test split), an interesting observation emerges. The RF model, which achieved an R^2 score of 0.116 during the general evaluation, not only delivered competitive performance in internal validation but also resulted in the most accurate prediction for 2023, reflecting both good generalization and robustness. Similarly, SVR and Huber Regressor, with R^2 scores of 0.12 and 0.162, respectively, also exhibited balanced performance between test evaluation and real-world prediction, further affirming their adaptability.

The XGBoost model, despite recording the R^2 score (0.111) which shows lower values among most of the others during the test evaluation, also showed a lower prediction accuracy of 82.58% for 2023. The DT model, which had a relatively low R^2 score of 0.015, achieved 87.23% accuracy, suggesting moderate effectiveness in capturing short-term trends but weaker general performance. Meanwhile, the MLP model, with the weakest R^2 of -0.018 in the test evaluation, predicted with 96.14% accuracy—indicating an improvement in point-specific prediction despite general underperformance.

The superior performance of the RF model in predicting Indonesia's GDP growth for 2023 can be attributed to its ensemble learning mechanism, which combines multiple DTs to reduce variance and improve prediction stability. Unlike linear models that assume a specific functional form between predictors and the target variable, RF is capable of capturing complex, nonlinear interactions among economic indicators, which are often present in macroeconomic data. This flexibility allows it to model intricate relationships that may not be easily captured by simpler algorithms.

Previous research has highlighted the effectiveness of RF in economic forecasting contexts. For example, Gawthorpe [23] demonstrated that RF outperformed traditional econometric models and institutional forecasts in predicting GDP growth for the Czech economy, achieving high predictive accuracy and robustness. The study emphasizes RF's ability to learn from data without imposing strict economic assumptions, enabling it to capture nonlinear patterns and complex interactions inherent in economic systems. These findings align with the results of this study, further supporting the model's suitability for macroeconomic forecasting.

The limited size of the dataset presents a significant challenge in building predictive models with strong generalization capabilities. Small sample sizes tend to increase the risk of overfitting, particularly in complex models such as RF, XGBoost, or MLP, where the models may learn noise instead of capturing the true underlying relationships between variables. Additionally, limited data variation results in narrow exposure to potential future economic fluctuations, which can reduce the models' accuracy when applied to data outside the training range. This issue should be carefully considered when interpreting the prediction results and using the model outcomes to support economic policy decisions.

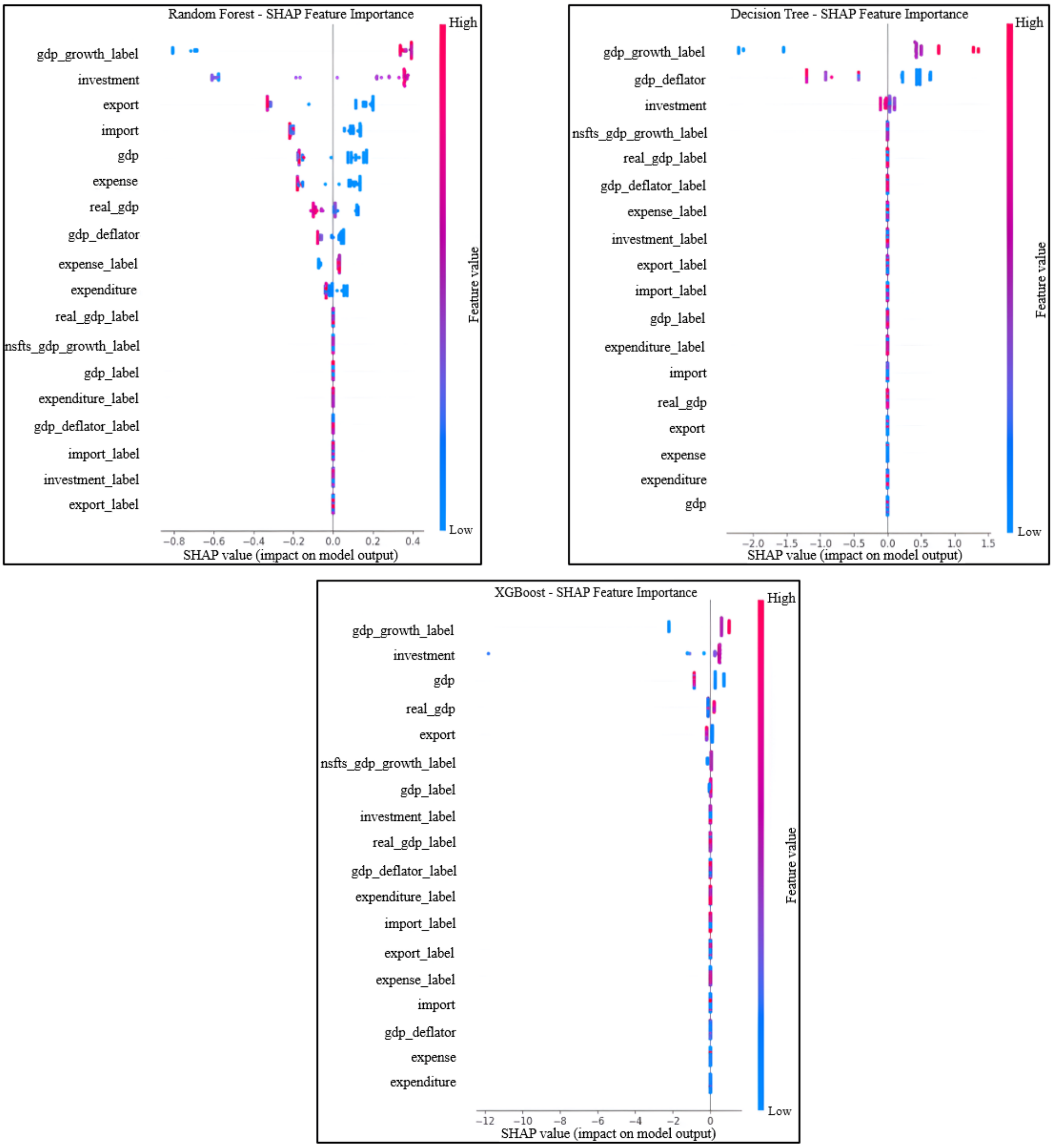

Figure 7. Feature importance of three tree-based models

4.2 Model Interpretability and Feature Importance

The SHapley Additive exPlanations (SHAP) summary plots presented in Figure 7 provide a comprehensive view of feature contributions toward the model output for RF, XGBoost, and DT regressors. Across all three models, the feature gdp_growth_label consistently emerges as the most influential predictor, indicating its strong relevance in capturing temporal dynamics or encoded patterns from previous GDP growth trends. This result validates the effectiveness of incorporating engineered temporal labels (such as fuzzy logic-based or segmented growth indicators) in forecasting tasks.

Investment, export, and import come rather close in relevance for the RF model. This implies that the expected GDP growth is much influenced by macroeconomic activities connected to trade flows and capital building. Features, including gdp, expenditure, and real_gdp also show modest contributions that fit the classical economic theory that states GDP performance is much influenced by production and consumption levels.

With greater SHAP magnitudes than the other two models, the XGBoost model emphasizes investment, GDP, and real_gdp more clearly. XGBoost might thus be more sensitive to these fundamental economic indicators. Furthermore, among the top predictors is nsfts_gdp_growth_label (derived from NSFTS), underscoring the value of advanced time series representations in improving model performance. Though its SHAP value distribution is flatter, the DT model also ranks investment, gdp_growth_label, and gdp_deflator as main drivers. Possibly because of its simpler structure and reduced model complexity compared to ensemble techniques, the model shows less sensitivity to features like export and import.

The SHAP analysis shows generally that including both engineered features (e.g., gdp_growth_label and nsfts_gdp_growth_label) and raw macroeconomic indicators (e.g., investment, gdp, and real_gdp) greatly improves model interpretability and predictive power. The need for temporal and structural patterns in economic forecasting is underlined by the consistent high ranking of label-based features across models. These revelations not only support the design of the feature engineering pipeline but also provide interpretable explanations that might help legislators grasp the underlying reasons for GDP growth projections.

From 1970 to 2023, the overlay plot in Figure 8 shows a consistent upward trend across all main economic variables, signifying consistent economic development. Particularly following the year 2000, which reflects notable macroeconomic expansion, GDP and household final consumption expenditure (i.e., expenditure) show the most notable increases. Particularly in recent years, the widening gap between GDP and expenditure over time points to components like government spending (expense), net exports and imports playing a progressively important role in promoting economic growth. Reflecting rising national productivity, the visualization emphasizes how GDP dominates other economic factors following 2010.

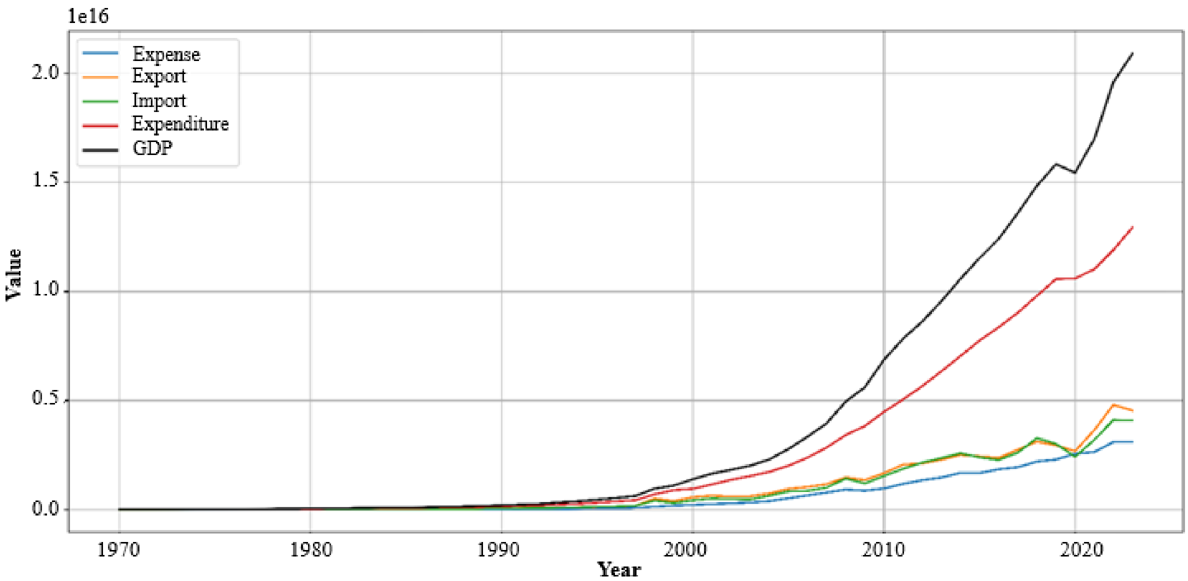

Figure 8. Time series trend

Particularly in GDP, exports, import, and expenditure, a clear fluctuation around 2020 across many variables is seen that corresponds to the worldwide economic disturbance brought about by the COVID-19 epidemic. Their overall trend still continues to be rising. Expense indicates controlled spending relative to GDP increase since it shows the slowest increase among all factors. This visualization also shows possible structural breaks or regime changes around important times, supporting the segmentation of time series data in next modelling techniques, including NSFTS.

4.3 Long-Term Prediction

The long-term prediction of GDP growth was conducted to evaluate the model's capability in capturing underlying economic trends beyond short-term fluctuations. While short-term forecasts are valuable for immediate policy and investment decisions, long-term projections offer strategic insights that are essential for macroeconomic planning, sustainable development, and fiscal policymaking. By extending the forecasting horizon beyond 2023, this analysis aims to assess the model's generalizability, robustness, and its ability to learn from historical patterns to anticipate future economic performance. Additionally, long-term prediction serves to identify potential shifts in growth trajectories under stable or evolving economic conditions, thereby enhancing the practical relevance and applicability of the proposed forecasting framework in real-world economic scenarios.

Table 5. Long-term prediction

Models	Forecasting GDP Growth (%)				
	2024	**2025**	**2026**	**2027**	**2028**
RF	5.155807	5.499683	5.573028	5.574672	5.409397
XGBoost	2.861613	3.672897	3.913201	4.201386	3.33348
SVR	4.790935	6.815759	7.411722	6.726914	5.325227
Huber Regressor	4.542968	6.475857	7.193313	6.83682	5.171086
DT	6.099917	6.099917	6.099917	4.975471	5.030874
MLP	14.813575	13.206728	11.958854	12.52878	11.117736

Based on Table 5 displaying the GDP growth forecasts from 2024 to 2028, a comparative analysis of six regression models—RF, XGBoost, SVR, Huber Regressor, DT, and MLP—can be conducted. The RF model yields relatively stable and realistic predictions, ranging from 5.15% to 5.57%, indicating moderate growth and robustness in handling projected input data. Similarly, the Huber Regressor and SVR models exhibit consistent predictions, albeit with slightly greater fluctuations and magnitudes. In contrast, the XGBoost model produces lower forecast values across all years, starting from 2.86% in 2024 and gradually increasing to 3.33% in 2028, suggesting a more conservative estimation behaviour. On the other hand, the MLP model displays significantly higher predicted values compared to all other models, beginning at 14.81% in 2024 and slightly decreasing to 11.12% by 2028. This pattern may indicate overfitting or high sensitivity. The DT model exhibits an unusual pattern, maintaining identical forecasts for the first three years (6.10%) before a notable drop in the subsequent years, suggesting a lack of adaptability to changing inputs.

The forecasting process was conducted in multiple stages due to the unavailability of actual feature data beyond 2023. Initially, the values of the independent variables for the period from 2024 to 2028 were predicted using the ARIMA model. Each regression model was then retrained iteratively for each forecast year. For instance, to predict GDP growth in 2024, the model was trained on historical data from 1970 to 2023. The prediction for 2024 was subsequently appended to the training data for predicting 2025, and this recursive process continued annually up to 2028. Importantly, the hyperparameters for each model remained unchanged throughout this process to maintain consistency and focus evaluation on each model's adaptability to progressively forecasted input data.

This recursive forecasting approach, often referred to as iterative or step-wise forecasting, mirrors real-world scenarios in multi-step time series predictions. The results presented in Table 5 provide valuable insights into each model's ability to maintain stability and predictive accuracy over extended horizons when relying on forecasted inputs rather than actual observed data.

4.4 Scenario-Based Analysis

To assess the robustness and sensitivity of the model in various macroeconomic conditions, several hypothetical scenarios were constructed for the year 2023 by modifying key economic variables. These scenarios are designed to simulate extreme yet plausible shifts in economic conditions.

As shown in Table 6, the economic crisis scenario reflects a downturn characterized by reductions in export (-30%), investment (-25%), and household expenditure (-20%). The government stimulus scenario simulates fiscal expansion, where government expense is increased by 40% and household expenditure by 20%. In the hyperinflation scenario, a doubling of the GDP deflator is assumed, along with increased government spending (+30%) and household consumption (+10%) to reflect inflationary pressure. The massive investment drop scenario models a sharp contraction in economic activity through a significant drop in investment (-70%), household expenditure (-50%), and GDP (-30%). Finally, the strong domestic scenario simulates accelerated domestic demand with substantial increases in investment (+60%), household expenditure (+80%), and import levels (+20%). These variations enable a comprehensive evaluation of the models' predictive performance under diverse economic conditions.

Table 6. Scenario description

Scenario	Variables		
Economic crisis	Export * 0.7	Investment * 0.75	Expenditure * 0.8
Government stimulus	Expense * 1.4	Expenditure * 1.2	
Hyperinflation	Expense * 1.3	Expenditure * 1.1	GDP deflator * 2
Massive investment drop	Investment * 1.6	Expenditure * 0.5	GDP * 0.7
Strong domestic	Investment * 1.6	Expenditure * 1.8	Import * 1.2

Table 7. Model prediction on scenario

Models	Predicted GDP Growth for 2023 (Without Scenario)	Model Prediction Based on Each Scenario (%)					Actual Value of GDP Growth for 2023
		Economic Crisis	Government Stimulus	Hyper-inflation	Massive Investment Drop	Strong Domestic	
RF	5.076019	5.076019	5.076019	5.076019	4.601401	5.076019	5.048106
XGBoost	5.927236	5.964660	5.927236	5.927236	5.801558	5.913652	
SVR	5.152487	6.182950	4.851175	0.612603	2.012889	5.951224	
Huber Regressor	4.993409	5.850339	5.338409	2.020886	2.166722	6.314512	
DT	5.692571	5.692571	5.692571	5.692571	5.692571	5.692571	
MLP	5.242731	3.629533	6.095713	7.509294	3.538266	6.713164	

Table 8. Accuracy of model prediction on scenario

Models	Accuracy of Predicted GDP Growth for 2023 (Without Scenario)	Accuracy Prediction Based on Each Scenario				
		Economic Crisis	Government Stimulus	Hyper-inflation	Massive Investment Drop	Strong Domestic
RF	99.45%	99.45%	99.45%	99.45%	91.15%	99.45%
XGBoost	82.58%	81.84%	82.58%	82.58%	85.07%	82.85%
SVR	97.93%	77.52%	96.1%	12.14%	39.87%	82.11%
Huber Regressor	98.92%	84.11%	94.25%	40.03%	42.92%	74.91%
DT	87.23%	87.23%	87.23%	87.23%	87.23%	87.23%
MLP	96.14%	71.9%	79.25%	51.25%	70.09%	67.02%

Table 7 and Table 8 present the GDP growth predictions for 2023 generated by six different machine learning models—RF, XGBoost, SVR, Huber Regressor, DT, and MLP—under both baseline and various hypothetical economic scenarios. The actual GDP growth value for 2023 was 5.048106%. Among all models, RF consistently demonstrated the highest and most stable predictive accuracy across all scenarios, maintaining performance above 91%, with a peak accuracy of 99.45% under the baseline, government stimulus, and strong domestic scenarios. XGBoost also showed moderate resilience in the face of economic shocks, although its accuracy declined notably under the hyperinflation scenario (12.14%). SVR achieved the highest accuracy under baseline conditions (97.93%) but exhibited significant sensitivity to extreme scenarios, with its accuracy dropping to 12.14% and 39.87% under hyperinflation and massive investment drop, respectively. Huber Regressor displayed fluctuating performance, ranging from relatively high accuracy under government stimulus (94.25%) to severe degradation under hyperinflation and massive investment drop scenarios, with accuracies of 40.03% and 42.92%, respectively. The DT model maintained consistent but moderate accuracy levels across all scenarios (87.23%), while MLP performed well under baseline conditions (96.14%) but showed substantial drops under stress, including 71.99% in economic crisis scenario and 67.02% in strong domestic scenario. These results emphasize the superior robustness of ensemble models—particularly RF—in sustaining reliable predictions across varying macroeconomic conditions.

4.5 LSTM Model Analysis

The LSTM model was constructed to analyse temporal patterns in economic growth indicators, employing a rigorous time-series validation framework. The dataset was partitioned into training (1980–2012), validation (2008–2012), and test sets (2013–2022) to ensure chronological integrity. Feature scaling was performed using StandardScaler, while a sequence length of three years was adopted to capture medium-term dependencies. Using grid search, the architecture comprised a single LSTM layer (32 units, tanh activation) followed by a dense output layer, optimized with Adam (learning rate=0.01) and early stopping (patience=10, monitoring validation loss). Model efficacy was assessed through MAE, RMSE, and R^2 metrics on strictly unseen test data, adhering to best practices in machine learning to prevent data leakage. Hyperparameters were tuned iteratively, prioritizing generalization over in-sample fit. This methodology addresses common pitfalls in economic forecasting, particularly non-stationarity and limited data samples, through disciplined temporal partitioning and regularization techniques.

Table 9. LSTM evaluation metrics

Model	MAE	RMSE	R^2
LSTM	1.829	3.083	-0.542

The performance of the LSTM model in predicting Indonesia's GDP growth for 2023 is summarized in Table 9 and Table 10. The model achieved a predicted GDP growth of 5.743966%, compared to the actual value of 5.048106%, resulting in an absolute accuracy of 86.22%. Despite this seemingly moderate predictive accuracy, the evaluation metrics reveal notable shortcomings in the model's performance. Specifically, MAE and RMSE were relatively high at 1.829 and 3.083, respectively. Most notably, the model produced a negative R^2 score of -0.542, indicating that LSTM failed to generalize effectively to the underlying pattern of the data and performed worse than a simple mean-based prediction. This suggests that while the model approximated the 2023 value with some closeness, it struggled to learn a consistent and reliable mapping from input features to GDP growth values over the entire dataset. The poor R^2 also highlights the LSTM model's potential sensitivity to the relatively small dataset and the non-stationarity characteristics of macroeconomic indicators, making it less suitable without further tuning or data augmentation.

Table 10. LSTM prediction of GDP growth for 2023

Model	Predicted GDP Growth for 2023 (%)	Actual Value of GDP Growth for 2023 (%)	Accuracy
LSTM	5.743966	5.048106	86.22%

5 Conclusion

This work effectively shows a strong hybrid forecasting method for estimating Indonesia's GDP growth by combining fuzzy logic with several machine learning techniques. Using a complete set of 19 input features—including new fuzzy-based representations and the NSFTS forecast label derived from annual macroeconomic data (1970–2023)—this study offers an insightful analysis of the predictive powers of several algorithms. With the lowest MAE and the best accuracy in forecasting Indonesia's GDP growth for 2023 (99.45%), the RF model regularly emerged among the six evaluated models as the best performer. This better performance emphasizes RF's capacity to efficiently capture short-term economic fluctuations inherent in macroeconomic time series data as well as complex, non-linear relationships. The results confirm the great promise of advanced machine learning methods in producing very accurate and dependable economic forecasts, providing a vital data-driven tool for Indonesian legislators to guide strategic economic planning and enable proactive decision-making and timely interventions to promote sustainable national development and stability.

5.1 Limitations

This study has several limitations despite its strong framework and encouraging findings that offer rich ground for future research. First of all, this study mainly emphasizes domestic economic factors. Although thorough, this strategy may not be able to adequately reflect the major influence of outside economic shocks or world trends clearly affecting the course of a country. Second, although the models show great prediction accuracy—especially for 2023—the rather low R^2 values for some models in the training phase point to unresolved variance in the historical time series data. This can suggest the natural complexity and variability in economic time series that the present model designs might not have completely addressed. Thirdly, although annual statistics gave a consistent historical picture, economic fluctuations often occur at higher frequencies, which could provide a more exact understanding of short-term dynamics. Fourthly, the dataset from 1970 to 2023, although comprehensive, may not fully represent the dynamic changes in Indonesia's economic and structural landscape. A fixed data division approach may not fully capture model performance across different economic periods. Finally, economic variables often have heterogeneous data formation processes that change over time, and ignoring this heterogeneity can affect model reliability.

5.2 Future Works

Building on the insights and limitations of this study, future studies could investigate several interesting paths to improve the generalisability and predictive ability of the suggested framework. Incorporating important outside economic elements, including international trade policies, global economic trends, key commodity prices (e.g., oil, coal, and palm oil relevant to Indonesia), and exchange rate swings, is critical in the next step. By considering more general macroeconomic effects, including these factors, could greatly increase model resilience and predictive capability. Moreover, tackling the complexity of data heterogeneity and dynamic data creation processes will be crucial; this could entail looking at more advanced time-series modelling techniques such as dynamic sampling methods (e.g., rolling window cross-validation) to evaluate model performance over different economic conditions. Alternatively, future research may investigate time-varying parameter models or regime-switching models (e.g., Markov-switching models) to explicitly identify various economic regimes or adaptive learning algorithms that change model parameters as data distributions change. These approaches would improve prediction generalisation and help to grasp how models behave under changing market conditions. Furthermore, it would be interesting to look at advanced deep learning architectures especially intended for complex, heterogeneous time-series data, such as Transformers, using longer and richer historical feature sets. Lastly, future research could evaluate the framework's generalisability and uncover country-specific economic characteristics influencing predictive accuracy by means of a comparison between its performance in estimating GDP growth across other emerging economies with different economic structures.

Author Contributions

Conceptualization, R.P.; methodology, R.P and M.I.S.; software, M.I.S; writing—original draft preparation, M.I.S.; writing—review and editing, R.P.; formal analysis, Z.Y; investigation, Z.Y.

Data Availability

The data that supporting our research results are deposited in Github, which does not issue DOIs. The data can be accessed at https://github.com/Ikhsan030904/GDP_Growth.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] C. M. Bătuşaru, A. Baboş, and V. Pîrvuţ, "Tracking economic growth: A correlation analyse on the determinants of gross domestic product," *Int. Conf. Knowl.-Based Organ.*, vol. 29, no. 2, pp. 24–29, 2023. https://doi.org/10.2478/kbo-2023-0031

[2] K. S. and M. Jegadeeshwaran, "Macro-economic analysis between the Indian economy and Indonesian economy," *Int. J. Multidiscip. Res.*, vol. 5, no. 1, pp. 1–10, 2023. https://doi.org/10.36948/ijfmr.2023.v05i01.1495

[3] G. Utama and N. Firinda, "Nowcasting GDP with machine learning: The case of Indonesia," *Research Square Preprint*, 2024. https://doi.org/10.21203/rs.3.rs-3903568/v1

[4] N. D. Muchisha, N. Tamara, A. Andriansyah, and A. M. Soleh, "Nowcasting Indonesia's GDP growth using machine learning algorithms," *Indones. J. Stat. Its Appl.*, vol. 5, no. 2, pp. 355–368, 2021. https://doi.org/10.29244/ijsa.v5i2p355-368

[5] G. Cicceri, G. Inserra, and M. Limosani, "A machine learning approach to forecast economic recessions-An Italian case study," *Mathematics*, vol. 8, no. 2, p. 241, 2020. https://doi.org/10.3390/math8020241

[6] G. A. Ansari, S. S. Bhat, and S. Bhoite, "A comparative analysis and prediction of diabetes using machine learning approaches," in *Proceedings of Fifth Doctoral Symposium on Computational Intelligence*. Singapore: Springer, 2024, pp. 211–223. https://doi.org/10.1007/978-981-97-6318-4_16

[7] M. Yu, "Machine learning for economic forecast based on past 30 years and Covid-19 pandemic data," *Adv. Econ. Manag. Polit. Sci.*, vol. 80, pp. 129–138, 2024. https://doi.org/10.54254/2754-1169/80/20241897

[8] A. R. Subian, D. A. Mulkan, H. H. Ahmady, and F. Kartiasih, "Comparison methods of machine learning and deep learning to forecast the GDP of Indonesia," *Sistemasi*, vol. 13, no. 1, pp. 149–166, 2024. https://doi.org/10.32520/stmsi.v13i1.3445

[9] C. X. Luo and Y. T. Xue, "Affecting factors and forecasting trends of American GDP," *BCP Bus. Manag.*, vol. 40, pp. 343–350, 2023. https://doi.org/10.54691/bcpbm.v40i.4399

[10] J. Yoon, "Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach," *Comput. Econ.*, vol. 57, pp. 247–265, 2021. https://doi.org/10.1007/s10614-020-10054-w

[11] A. Hossain, M. Hossen, M. M. Hasan, and A. Sattar, "GDP growth prediction of Bangladesh using machine learning algorithm," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, 2021, pp. 812–817. https://doi.org/10.1109/ICICV50876.2021.9388593

[12] M. S. Ahammad, S. A. Sinthia, M. Hossain, M. M. Ahmed, M. N. A. Ikram, and N. A. A. Asif, "Machine learning for GDP forecasting: Enhancing economic projections in Bangladesh," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 2024, pp. 1–5. https://doi.org/10.1109/ICCCNT61001.2024.10725136

[13] S. Sa'adah and M. S. Wibowo, "Prediction of Gross Domestic Product (GDP) in Indonesia using deep learning algorithm," in *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Yogyakarta, Indonesia, 2020, pp. 32–36. https://doi.org/10.1109/ISRITI51436.2020.9315519

[14] World Bank Group, "The World Bank policy on access to information (English)." http://documents.worldbank.org/curated/en/843001468166481505

[15] R. Estrada-Lavilla and J. Ruiz-Navarro, "Method for and analysis of early-stage firm growth patterns using World Bank Data," *Sustainability*, vol. 16, no. 4, p. 1450, 2024. https://doi.org/10.3390/su16041450

[16] World Bank Group, "GDP growth (annual %) - Indonesia." https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?end=2024&locations=ID&start=2000&view=chart

[17] P. C. de Lima e Silva, C. A. Severiano, M. A. Alves, R. Silva, M. W. Cohen, and F. G. Guimarães, "Forecasting in non-stationary environments with fuzzy time series," *Appl. Soft Comput.*, vol. 97, p. 106825, 2020. https://doi.org/10.1016/j.asoc.2020.106825

[18] J. M. Valente and S. Maldonado, "SVR-FFS: A novel forward feature selection approach for high-frequency time series forecasting using support vector regression," *Expert Syst. Appl.*, vol. 160, p. 113729, 2020. https://doi.org/10.1016/j.eswa.2020.113729

[19] B. Z. Chen, W. J. Zhai, and Z. Y. Huang, "Low-rank elastic-net regularized multivariate Huber regression model," *Appl. Math. Model.*, vol. 87, pp. 571–583, 2020. https://doi.org/10.1016/j.apm.2020.05.012

[20] B. Soybilgen and E. Yazgan, "Nowcasting US GDP using tree-based ensemble models and dynamic factors," *Comput. Econ.*, vol. 57, pp. 387–417, 2021. https://doi.org/10.1007/s10614-020-10083-5

[21] Y. Q. Yang, X. C. Xu, J. F. Ge, and Y. Xu, "Machine learning for economic forecasting: An application to China's GDP growth," *arXiv preprint, arXiv:2407.03595*, 2024. https://doi.org/10.48550/arXiv.2407.03595

[22] M. Z. Al-Faiz, A. A. Ibrahim, and S. M. Hadi, "The effect of Z-Score standardization (normalization) on binary input due the speed of learning in back-propagation neural network," *Iraqi J. Inf. Commun. Technol.*, vol. 1, no. 3, pp. 42–48, 2019. https://doi.org/10.31987/ijict.1.3.41

[23] K. Gawthorpe, "Random forest as a model for Czech forecasting," *Prague Econ. Pap.*, vol. 30, no. 3, pp. 336–357, 2021. https://doi.org/10.18267/j.pep.765