Abaad Iktissadia Review Vol: (14) N°(01) year: 2024 pages(391-410)

Random Forest Modeling for Forecasting Economic Growth: A Machine Learning Approach in the Case of Algeria (1964-2022)

Malika Messaoudi*
Hassiba Benbouali University of Chlef,
Algeria
m.messaoudi@univ-chlef.dz

submitted: 29/04/2024 **Accepted:20/06/2024** **published:30/06/2024**

Abstract

This research aims to explore the efficacy of machine learning techniques, specifically Random Forest modeling, in forecasting economic growth. The research problem lies in the challenge of accurately predicting economic trends, which is crucial for effective policy formulation and decision-making. The study follows a structured methodology comprising data collection, preprocessing, feature selection, model training, and validation. Results demonstrate the effectiveness of Random Forest modeling in capturing the intricate patterns of economic data and outperforming traditional forecasting methods. This approach offers promising prospects for enhancing the accuracy and reliability of economic growth forecasts, thereby facilitating informed decision-making processes in various sectors.

Keywords: Economic forecasting, Time series analysis, Machine learning (ML), Random Forest (RF), Tree decision.

JEL Classification Codes: C50, C52, C55, C63

***Correspondent author**

Introduction

Economic forecasting, particularly in predicting GDP growth, holds paramount importance for policymakers, investors, and businesses. Traditional econometric models have served as the cornerstone of economic forecasting. However, the emergence of machine learning (ML) techniques has heralded a new era of heightened accuracy and reliability in economic predictions. ML techniques, with their potent ensemble learning methods, have garnered widespread acclaim across diverse domains, particularly within economics.

Through ongoing refinement efforts by both researchers and practitioners, ML algorithms have evolved to address myriad challenges encountered in predictive modeling. These challenges include mitigating overfitting, ensuring model interpretability, and enhancing computational efficiency. Leveraging their versatility in handling varied data types, capturing nonlinear relationships, and processing large-scale datasets, ML algorithms have found extensive utility across multifarious domains, including finance, healthcare, and marketing.

In the realm of economic modeling, ML techniques have demonstrated remarkable versatility and effectiveness in capturing the complex relationships and nonlinearities inherent in economic data. Their applications span from nowcasting GDP to predicting financial market trends and identifying key drivers of economic performance. One such exemplary technique is the Random Forest (RF). RF constructs a multitude of decision trees independently and combines their predictions through an averaging or voting mechanism. This ensemble method harnesses the diversity of individual trees to mitigate overfitting and bolster model robustness. RF has garnered popularity owing to its simplicity, scalability, and resistance to overfitting, rendering it suitable for a wide array of machine learning tasks. Moreover, RF's adeptness in handling high-dimensional data, missing values, and categorical variables further amplifies its utility in practical applications. Despite its apparent simplicity, RF consistently delivers competitive performance and remains a staple algorithm in the machine learning toolkit.

Research problem

The study endeavors to address the following question:

How does the machine learning technique based on Random Forest modeling accurately predict economic growth?

Study hypotheses

This study relies on two fundamental hypotheses:

- There exist nonlinear relationships among various economic variables, and therefore, machine learning techniques may better represent these relationships than traditional models.
- The advancement of technology in the field of artificial intelligence facilitates the process of collecting, processing, and analyzing economic data more quickly and efficiently, thereby enhancing the feasibility of using machine learning techniques for economic growth prediction.

Objectives of the study

The study aims to:

- Understand the concept of machine learning and the fundamentals of Random Forest modeling.
- Explore how machine learning and Random Forest modeling are applied in economic growth prediction.
- Analyze the data used in the process of economic forecasting.
- Clarify the algorithms used in machine learning techniques based on Random Forest modeling.
- Review the practical applications of this technique in economic growth analysis.

Study methodology

The study adopted both deductive and inductive methodologies according to the research requirements:

- Deductive methodology: Utilized in the theoretical processing of the study, relying on descriptive tool.
- Inductive methodology: The inductive approach was employed in the quantitative analysis of economics, where the machine learning technique based on Random Forest modeling is used to predict economic growth. Given the large volume of data, deep learning techniques could also be considered to potentially enhance prediction accuracy further.

The dataset was partitioned into training, validation, and testing phases to ensure unbiasedness. Cross-validation techniques were employed to fine-tune the model and evaluate its performance effectively.

Literature review

In recent years, the utilization of machine learning (ML) techniques in economic forecasting has attracted significant attention, with a burgeoning body of literature investigating its potential and limitations.

- (Periklis & al. 2022) delve into the pivotal task of forecasting the unemployment rate in the Euro Area using Machine Learning. The authors employ three ML methodologies: decision trees (DT), random forests (RF), and support vector machines (SVM), alongside an elastic-net logistic regression (logit) model derived from econometrics. The results show that the optimal RF model outperforms the other models by reaching a full-dataset forecasting accuracy of 88.5% and 85.4% on the out-of-sample.

- Furthermore, (Sermpinis et al., 2014) [1]present work focusing on the utilization of a hybrid machine learning technique named genetic support vector regression (GSVR) for forecasting inflation and unemployment. The authors leverage genetic support vector regression, which amalgamates support vector regression (SVR) with genetic algorithms (GA) for parameter optimization.
In the realm of forecasting GDP growth, numerous researchers have explored the application of machine learning methodologies, underscoring the advantages of ensemble methods in handling nonlinearities and capturing intricate patterns in economic data, thereby showcasing superior performance compared to traditional forecasting models. In terms of our results, GA-SVR outperforms all benchmark models and provides evidence on which macroeconomic variables can be relevant predictors of US inflation and unemployment in the specific period under study.

-(Yoon, 2021) investigates the prediction of real GDP growth using ML approaches, demonstrating their effectiveness in capturing the nuanced dynamics of economic variables. Their research incorporates a novel feature engineering approach and model architecture optimization to enhance the accuracy of GDP growth predictions.
The results of this paper show that for the 2001–2018 period, the forecasts by the gradient boosting model and random forest model are more accurate than the benchmark forecasts. Between the gradient boosting and random forest models, the gradient boosting model turns out to be more accurate. This study encourages increasing the use of machine learning models in macroeconomic forecasting.

Despite these advancements, challenges persist in adopting ML techniques for economic forecasting, encompassing issues such as data availability, model interpretability, and robustness to structural breaks. Nevertheless, the growing body of empirical evidence suggests that ML-based approaches hold promise in enhancing the accuracy and reliability of economic predictions.

The study plan

In the theoretical aspect, the study delves into the realm of artificial intelligence, with a focus on machine learning and related branches. It comprehensively discusses the fundamentals of predictive modeling, elucidates the pivotal role of machine learning within this context, and explores advanced prediction techniques, notably the Random Forest (RF) and Gradient Boosting Machine (GBM) techniques.

Transitioning to the applied aspect, the study outlines the methodology employed, emphasizing data preparation as a crucial step. It delineates the modeling method, which includes the RF Nowcasting and RF forecasting models. Subsequently, the evaluation of these models is conducted, followed by a thorough examination of the results and their analysis. Additionally, sensitivity analysis is performed to gauge the robustness of the models.

Firstly : Theoretical side

The theoretical aspect involves exploring relevant concepts of the machine learning technique, aiding in understanding the theoretical background of the subject and framing the research.

1. Machine learning and artificial intelligence

Machine learning is a subset of artificial intelligence (AI) that focuses on creating systems that learn or improve performance based on the data they consume (Velidi, G. 2022, p41). Artificial intelligence is a broad term referring to systems or devices that simulate human intelligence. Machine learning and artificial intelligence are often discussed together and the terms are sometimes used interchangeably, but they do not mean the same thing. It is important to note that while all machine learning techniques are considered artificial intelligence, not all artificial intelligence represents machine learning. Machine learning is ubiquitous in business realms; interacting with banks, shopping online, or utilizing statistical models, machine learning algorithms significantly enhance efficiency, seamlessness, and security.

2. The role of machine learning

Machine learning algorithms play a crucial role in predictive modeling using big data. These algorithms can autonomously learn and improve through experience without being explicitly programmed. They can handle vast amounts of data, identify complex patterns, and make accurate predictions (Dash, et al. 2016, p 42).

3. Predictive modeling and machine learning using big data

Predictive modeling and machine learning have become indispensable tools in the field of big data analysis, enabling the enhancement of decision-making processes due to vast quantities of data.

3.1. Understanding predictive modeling

Predictive modeling involves using historical data to make forecasts or predictions about future outcomes. It encompasses a range of statistical techniques and machine learning algorithms (Katris, C. 2019, p 673) capable of identifying patterns, trends, and interrelationships within large datasets. By analyzing these patterns, predictive models can generate accurate forecasts, enabling organizations and governments to mitigate risks.

3.2.Advanced prediction techniques in machine learning

There are many machine learning techniques, but two important types based decision tree are metioned: Gradient Boosting Machines (GBM) and Random Forests (RF) (Saurabh Ghosh et al, 2023, p 33).

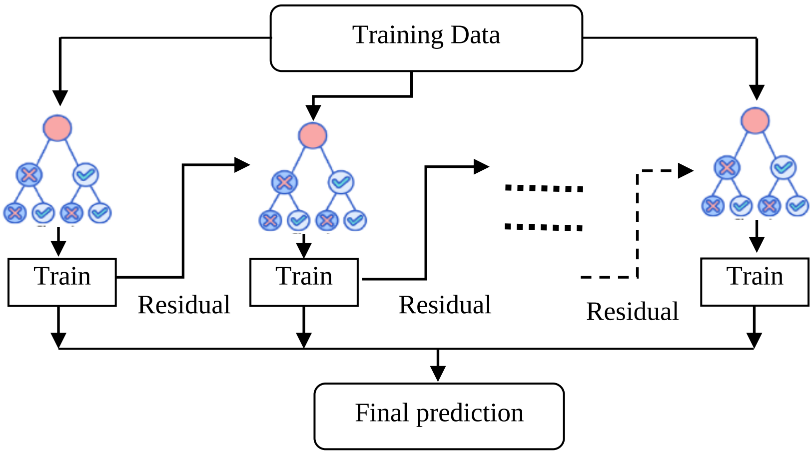
3.2.1. Gradient Boosting Machine (GBM) technique

The Gradient Boosting Machine (GBM) operates by sequentially combining the multiple weaknesses of decision trees, where at each step, weaknesses (errors) are corrected. The GBM adapts new trees repeatedly with the model, gradually reducing errors and enhancing overall prediction accuracy (see Figure 1) (Friedman, J. 2001, p 1189).

3.2.2.Random Forest (RF) technique

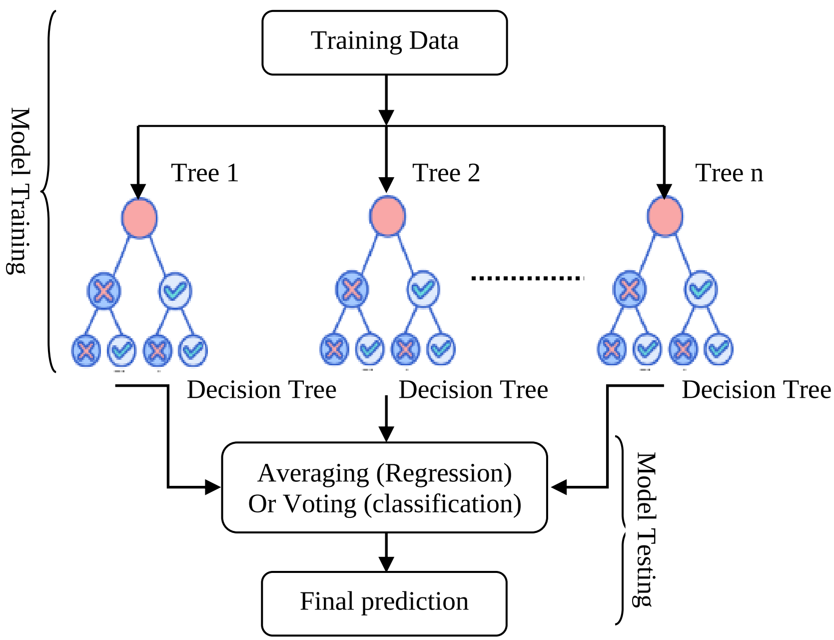
The Random Forest (RF) technique is considered a powerful tool in machine learning, operating by creating numerous decision trees using random subsets of the data (see Figure 2). (Breiman, et al, 2001, p 5) Each tree considers different features to make predictions independently of the other trees, akin to having a diverse set of opinions. Then, the Random Forest aggregates all these opinions to make a final prediction, aiding in more accurate predictions (Trevor Hastie, et al, 2009).

Figure N⁰01: GBM modeling technique



Source: Prepared by the researcher

Figure N⁰02: RF modeling technique



Source: Prepared by the researcher

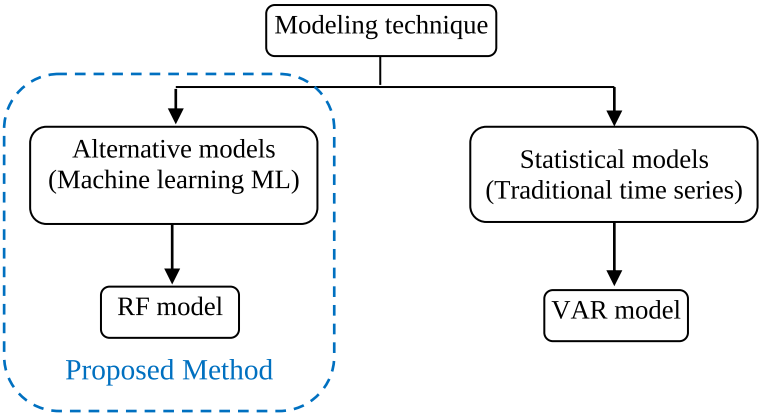
Secondly: Applied aspect

In the applied aspect, the (RF) algorithm is utilized to forecast (GDP) growth using economic scenario data.

1.Methodology

Traditional time series models such as AutoRegressive (AR) models are popular in predicting time series, relying on past observations to guide future trends. However, the emergence of machine learning techniques has spurred interest in exploring alternative methodologies. Machine learning models offer a unique advantage in illuminating nonlinear relationships and complex interactions within data, making them highly suitable for predictive tasks where traditional linear models may fall short (see Figure 3). Through a detailed analysis of machine learning models for economic forecasting, this study seeks to provide insights into the relative strengths and weaknesses of the (RF) model by evaluating its performance across diverse metrics and real-world datasets. Furthermore, to gain a deeper understanding of the capabilities and limitations of the proposed approach, sensitivity analysis is conducted to assess the resilience of machine learning models to fluctuations in input parameters.

Figure N⁰03: Proposed modeling technique

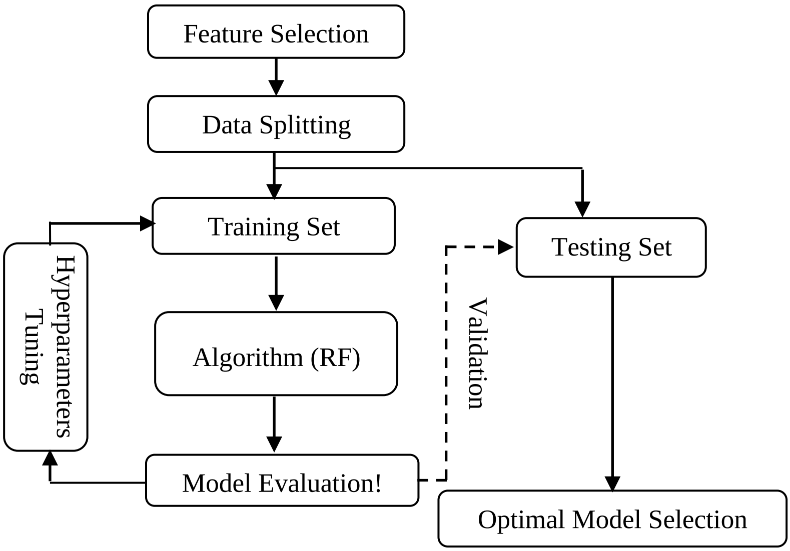


Source: Prepared by the researcher.

A detailed explanation is provided on how to use the (RF) algorithm to forecast (GDP) growth using a set of traditional economic indicators and economic scenario data. The study begins by identifying the most influential factors on GDP growth.

Various variables are carefully analyzed, and statistical techniques are employed to identify key factors. These selected features become the input variables in our model. Subsequently, the dataset is divided into training and testing sets using cross-validation methods. This critical step assists us in fine-tuning the model settings and effectively evaluating its performance (see Figure 4).

Figure N°04: Machine learning workflow



Source: Prepared by the researcher

2. Data preparation

Investigating the relationship between Algeria's Government Expenditure, Inflation, and GDP growth offers critical insights into the effectiveness of the country's fiscal policies and their broader implications for economic stability and growth. Algeria's government plays a central role in allocating funds, influencing various sectors of the economy, and affecting employment, investment, and overall economic activity. Analyzing how changes in Government Expenditure correlate with inflation rates and GDP growth provides valuable indicators of the government's fiscal policy effectiveness. By assessing the government's ability to balance spending priorities while promoting economic stability and growth, policymakers and economists can better understand Algeria's fiscal landscape and make informed policy decisions to support sustainable development. The investigation leverages a comprehensive dataset comprising three key macroeconomic variables (GDP, inflation rate, and government expenditure), measured quarterly from March 1964 through March 2022, sourced from data published by the World Bank. This dataset provides a reliable foundation for analyzing the dynamics of Algeria's monetary policy and economic performance over the specified period, enabling rigorous examination of the interplay between monetary policy instruments, inflationary trends, and GDP fluctuations.

Before proceeding with modeling, it is imperative to ensure that the collected historical data undergoes thorough cleaning and preprocessing. Any missing values should be handled appropriately, and outliers may need to be addressed to maintain the integrity of the dataset. Additionally, standardizing or normalizing the features in the dataset is recommended to ensure that they are on a similar scale, which can facilitate the convergence and performance of the predicted and forecasted model. The process begins by collecting the historical data (government expenditure and inflation rate) denoted by G and I, respectively, as well as economic scenarios data (G_DIFF and I_DIFF). These datasets are combined into a feature matrix X where each row represents a historical observation and each column represents a feature (Shams, M.*et al.* 2024).

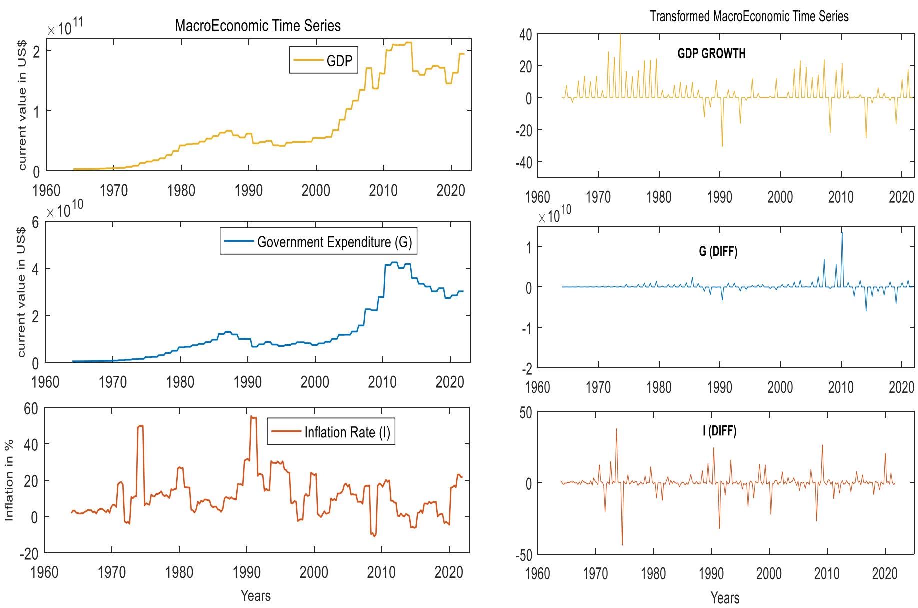
$$X = \begin{bmatrix} G_1 & I_1 & G\_DIFF_1 & I\_DIFF_1 \\ G_2 & I_2 & G\_DIFF_2 & I\_DIFF_1 \\ \vdots & \vdots & \vdots & \vdots \\ G_N & I_N & G\_DIFF_N & I\_DIFF_N \end{bmatrix} \tag{1}$$

Where:
- Government expenditure (G): Total amount of money spent by the government on goods and services for current consumption purposes.
- Inflation rate (I): The rate at which the general level of prices for goods and services is rising.
- G_DIFF: The difference in government expenditure scenario compared to a baseline scenario (average historical Data).
- I_DIFF: The difference in the inflation rate and the central bank's target rate scenario compared to a baseline scenario.

By examining the time series, it seems evident that all series exhibit nonstationarity and possess vastly different scales. Since modeling with non-stationary variables poses a challenge, it is imperative to transform each series appropriately. As illustrated in Figure 5, this transformation renders all series stationary while maintaining their distinct scales.

To ensure the accurate fitting of a model to the data, all variables must exhibit stationarity. To assess stationarity, the Augmented Dickey-Fuller (ADF) test for unit root nonstationarity is conducted. The ADF test results (Refer to Annexe N°01) indicate that the null hypothesis for GDP, Inflation rate (I), and government expenditure (G) is not rejected. However, the ADF test rejects the null hypothesis for all transformed series: GDP Growth, I_DIFF, and G_DIFF, allowing us to include these series in the autoregressive (AR) model. Analysis of the AR coefficients suggests that the growth of GDP in previous periods does not significantly influence current GDP growth, except for a statistically significant coefficient observed for the fourth lag. Additionally, the series becomes stationary after differencing, as confirmed by the (ADF) test results, with all tests showing statistically significant outcomes ($p < 0.001$). However, upon thorough examination, the residuals and autocorrelation exhibit heavy-tailed distributions, as outlined in the Appendix. These findings suggest that while the fitted model may be dependable for forecasting purposes within expected parameter ranges, its reliability may diminish when analyzing extreme tail behavior in the data. This limitation underscores the importance of adopting more sophisticated modeling approaches, such as machine learning techniques, to overcome such challenges.

Figure N⁰05: Transformed MacroEconomic Time Series Variables



Source: Prepared by the researcher based on Matlab software.

3.Modeling method

In this section, the RF nowcasting model and the RF forecasting model are examined.

3.1.RF nowcasting model

Random Forest is an ensemble learning method that operates by constructing a multitude of decision trees during training and outputting the mode of the average prediction (regression) of the individual trees. Each decision tree is built using a random subset of the training data and a random subset of the features. The prediction of (RF) model can be formulated by the equation (2) (Box, G. E.et al, 2015):

$$\hat{Y} = \frac{1}{N}\sum_{i=1}^{N} f_i(X) \tag{2}$$

Where: $\hat{Y}$ represents the predicted outcome (predicted GDP growth), N is the number of decision trees in the forest, and $f_i(X)$ represents the prediction of the i-th decision tree (Richardson, A, et al. 2021, p941).

3.2.RF forecasting model

For forecasting future GDP growth, the economic scenarios data for the forecasted period is aligned with the format and structure of the historical data used during model training. The trained RF model is then applied to predict GDP growth for the upcoming periods. Additionally, uncertainty associated with the forecasted GDP growth is evaluated by generating prediction intervals or confidence intervals, providing valuable insights into the range of possible outcomes.
When forecasting future GDP growth, the feature matrix X is expanded to encompass the economic scenarios data for the forecasted period. Subsequently, the trained RF model is utilized to predict GDP growth for the forthcoming periods. The equation for forecasting can be interpreted as follows (Khaidem, L., et al, 2016):

$$\hat{Y}_{future} = \frac{1}{N}\sum_{i=1}^{N} f_i(X_{future}) \tag{3}$$

Where:

$\hat{Y}_{future}$: represents the predicted outcome (forecasted GDP growth) $f_i(X_{future})$: represents the prediction of the i-th decision tree for the data X_{future}.

4.Model evaluation

To evaluate the performance of the RF model, various evaluation metrics are computed, including the Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics are defined as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i) \tag{4}$$

$$MAE = \frac{i}{N}\sum_{i=1}^{N}\left|(Y_i - \hat{Y}_i)\right| \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left|(Y_i - \hat{Y}_i)\right|}{\sum_{i=1}^{N}\left|(Y_i - \bar{Y}_i)\right|} \tag{6}$$

Where:

N : is the number of observations.

$\bar{Y}_i$: is the mean of the actual values of Y_i .

5.Results and analysis

In this section, the expected outcomes of employing random forests (RF) forecasting techniques are delineated. Furthermore, the significance of conducting sensitivity analysis to assess the resilience of the proposed method to fluctuations in input hyperparameters is examined.

Table.1 collectively depict various aspects of the (RF) method's performance in GDP growth nowcasting. The RF method demonstrated exceptional predictive performance with a mean absolute error (MAE) of 0.7900, mean squared error (MSE) of 3.9025, and an R-squared value of 0.9444 (Refer to Annexe N°02).

Table N°01: Performance metrics comparison

Modeling method	(MAE)(1)	(MSE)(2)	R^{2}(3)
VAR	2.5322	43.6876	0.1924
RF	**0.7900**	**3.9025**	**0.9444**

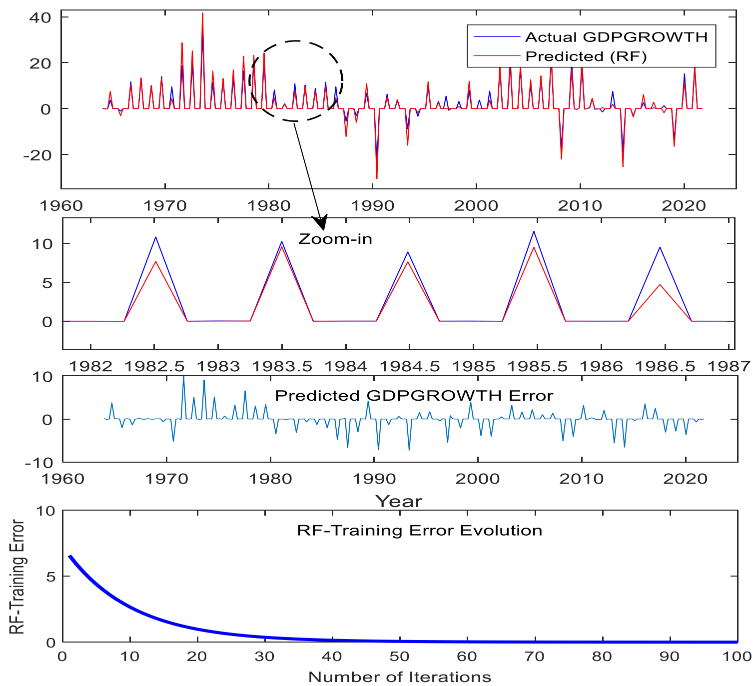
Source: Prepared by the researcher based on Matlab software.

Notes:(1),(2) : The MAE and MSE represents the average absolute and squared difference between the predicted and actual values of GDP growth. A lower MAE indicates better accuracy, and in this case, the Random Forest models outperform the AR model, with the RF achieving the lowest MAE. (3): R-squared indicates the proportion of variance in the GDP growth data explained by the model. Higher R-squared values suggest better generalization ability. In this comparison, again Random Forest models exhibit significantly higher R-squared values compared to AR model, indicating their ability to capture a larger portion of the variability in the data and providing better predictions.

Figure.6 provides a closer examination of the forecasted values during a specific period, offering insights into the accuracy and precision of RF model's predictions. Furthermore, it offers insights into the error of the predicted GDP growth solely for the RF method, indicating its robustness in capturing and predicting GDP growth dynamics.

These results showcase RF's superior accuracy in capturing complex relationships between economic indicators such as Government Expenditure and Inflation Rate, and GDP growth. Notably, RF exhibited lower errors and higher explanatory power, suggesting its effectiveness in forecasting economic outcomes.

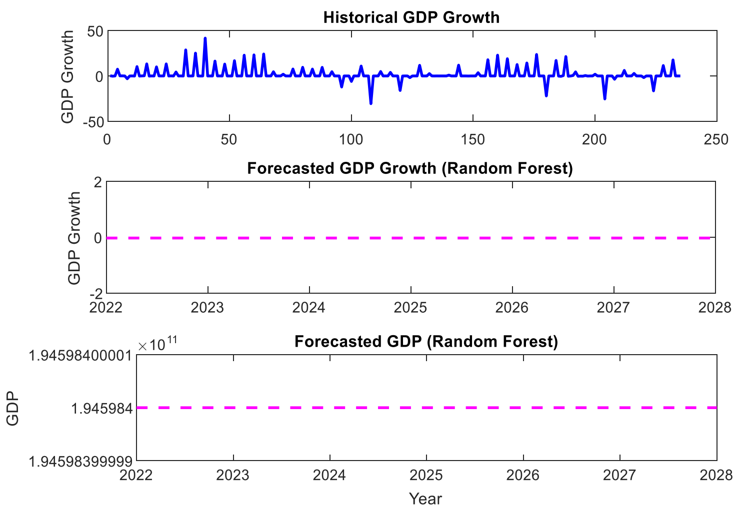
Figure N°06:Predicted GDP growth, and training error evolution with



Source: Prepared by the researcher based on Matlab software.

Finally, Figure 7 displays the forecasted results obtained using the RF method, further emphasizing its exceptional predictive performance and suitability for GDP growth forecasting. These results underscore RF as the preferred choice for GDP growth forecasting, offering robust predictions even in dynamic economic scenarios. However, it's essential to acknowledge that RF's performance may vary based on dataset characteristics and parameter settings, necessitating a careful examination of the sensitivity to hyperparameters to verify its applicability across diverse economic contexts and periods.

Figure N°07:Forecasted GDP/GDP growth using RF model

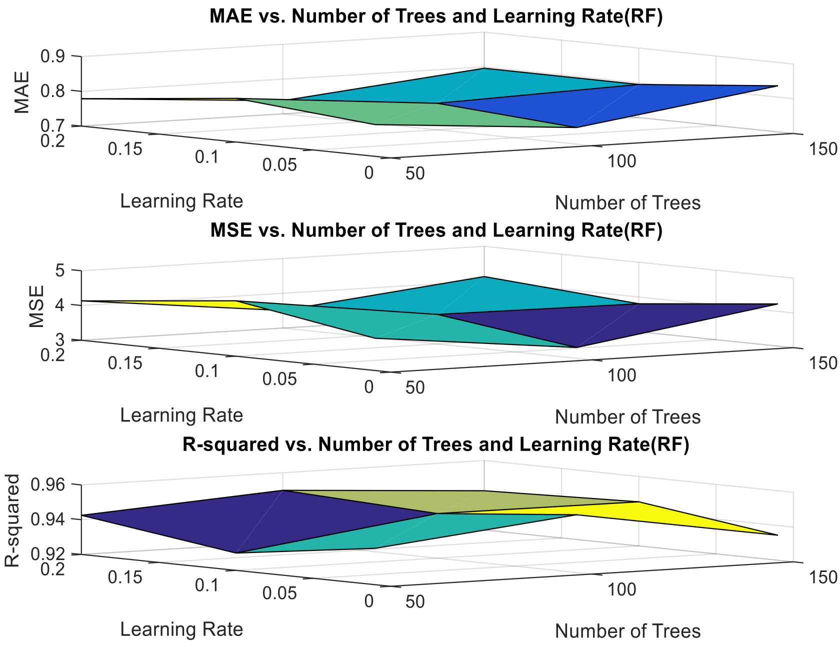


Source: Prepared by the researcher based on Matlab software.

6.Sensitivity analysis

In this section, a sensitivity analysis is conducted by systematically varying the hyperparameters (number of trees, learning rate, and maximum tree depth) of the Random Forest model. The evaluation metrics (MAE, MSE, R-squared) are calculated for each combination of hyperparameters to identify the optimal settings for minimizing MAE, minimizing MSE, and maximizing R-squared. These evaluation metrics provide valuable insights into the performance of the proposed alternative forecasting technique. Figure 8 illustrates the sensitivity analysis results of the RF model. This surface figure offers a visual representation of how changes in hyperparameters impact the performance metrics of each model. By examining the contours and gradients of the surfaces, the optimal hyperparameters are identified for each model across different evaluation metrics (see Tab.2). For instance, when minimizing Mean Absolute Error (MAE), the RF model achieved a lower value (0.7672). Similarly, for minimizing Mean Squared Error (MSE), the RF model obtained a lower value (3.4599). Additionally, when maximizing R-squared, the RF model achieved a slightly higher value (0.9450). These findings suggest that the RF model demonstrates greater robustness and effectiveness across various hyperparameter configurations, indicating its superior performance in sensitivity analysis. Moreover, these results underscore the efficacy of ensemble-based techniques like Random Forest for economic forecasting tasks.

Figure N°08:Sensitivity analysis result of RF model



Source: Prepared by the researcher based on Matlab software.

These hyperparameters [(1),(2)] are identified as the best combination for minimizing the absolute and squared difference between the predicted and actual values of GDP growth. A lower MAE and MSE suggests that the model's predictions are, on average, closer to the true values (better model performance in terms of reducing prediction errors). [(3)]: These hyperparameters maximize the coefficient of determination (R-squared), indicating the proportion of variance in the GDP growth data explained by the model. Higher R-squared values suggest that the model captures a larger portion of the variability in the data and provides better predictions.

Table N°02: Optimal hyperparameters selection for evaluation metrics comparison

Hyperparameters (RF method)		Number of Trees	Learning Rate	Max Tree Depth
Minimizing (MAE)(1)	**0.7672**	**50**	**0.10**	**3**
Minimizing (MSE) (2)	**3.4599**	**50**	**0.10**	**3**
Maximizing R^2 (3)	**0.9433**	**50**	**0.01**	**3**

Source: Prepared by the researcher based on Matlab software.

Conclusion

This study has addressed enhancing understanding of forecasting methodologies and facilitating informed decision-making processes in economic analysis and policy formulation through utilizing machine learning methodologies to predict GDP growth using the RF technique.

Results

-The study demonstrated that the (RF) technique provides high accuracy in economic growth forecasting.
- The results indicated the technique's ability to uncover non-linear relationships within economic data, enhancing its capability to predict changes in the economy.
- Thanks to its capability to handle large amounts of data, the Random Forest technique effectively leveraged big data to improve the accuracy of economic forecasts.
-The study showed that machine learning technology enables rapid and efficient forecasts of economic changes, facilitating strategic decision-making for the government.
-The technique demonstrated greater ability to predict future economic changes, enhancing the effectiveness of economic risk planning and management.
- Sensitivity analyses revealed the proposed model's ability to enhance the accuracy and reliability of economic forecasts, underscoring the importance of adopting modern methodologies in predicting increasingly complex economic phenomena.

Hypothesis testing

- Non-linear relationships exist among various economic variables, thus machine learning techniques may offer the ability to represent these relationships better than traditional models. Because machine learning models provide unique benefits due to their ability to detect and understand non-linear relationships and complex interactions within data.
- The advancement of technology in the field of artificial intelligence facilitates the process of collecting, processing, and analyzing economic data faster and more efficiently. This enhances the possibility of using machine learning techniques to forecast economic growth, as evidenced by the exceptional accuracy of (RF) in capturing complex relationships among economic indicators.

Recommendations

- Adopting (RF) technique as a vital analytical tool in economic planning processes and vital economic decision-making.
-Enhancing the quality and organization of available economic data structures to ensure full utilization of machine learning capabilities.

- Strengthening training and development in the field of machine learning technology among economic experts and analysts to ensure a comprehensive understanding of the benefits and applications of this technology.

List of references

1. Periklis, G., Papadimitriou, T., & Sofianos, E. (2022). Forecasting unemployment in the euro area with machine learning. *Journal of Forecasting*, 41(3), 551–566. https://doi.org/10.1002/for.2824
2. Sermpinis, G., Stasinakis, C., Theofilatos, K., & Karathanasopoulos, A. (2014). Inflation and unemployment forecasting with genetic support vector regression. *Journal of Forecasting*, 33(6), 471–487. https://doi.org/10.1002/for.2296
3. Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: gradient boosting and random forest approach. *Computational Economics,57*(1), 247–265.
4. Velidi, G. (2022). GDP prediction for countries using machine learning models. *Journal of Emerging Strategies in New Economics,1*(1), 41–49.

5. Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. The Journal of Finance and Data Science, 2(1), 42–57.

6. Katris, C. (2019). Prediction of unemployment rates with time series and machine learning techniques. *Computational Economics*, 55(2), 673–706.
7. Saurabh Ghosh & Abhishek Ranjan, (2023). "A Machine Learning Approach To Gdp Nowcasting: An Emerging Market Experience," *Bulletin of Monetary Economics and Banking*, Bank Indonesia, vol. 26(Special I), pages 33-54.
8. Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*(5), 1189–1232. http://www.jstor.org/stable/2699986.
9. Breiman, L. Random Forests. *Machine Learning* 45, 5–32 (2001).
10. https://doi.org/10.1023/A:1010933404324
11. Shams, M.Y., Tarek, Z., El-kenawy, ES.M. *et al.* Predicting Gross Domestic Product (GDP) using a PC-LSTM-RNN model in urban profiling areas. *Comput.Urban Sci.* 4, 3 (2024). https://doi.org/10.1007/s43762-024-00116-2
12. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. Time Series Analysis: Forecasting and Control (2015). *John Wiley & Sons*.
13. Richardson, A., Mulder, T. V. F., & Vehbi, T. (2021). Nowcasting GDP Using Machine Learning Algorithms: A real time assessment. *International Journal of Forecasting*, 37, 941-948.
14. Khaidem, L., Saha, S., & Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. arXiv preprint arXiv:1605.00003.

Annexes

Annexe .1 : ADF Test

ADF Table = 5×8 table

	h	pValue	stat	cValue	Lags	Alpha	Model	Test
Test 1	true	0.001	-15.265	-1.9421	0	0.05	{'AR'}	{'T1'}
Test 2	true	0.001	-10.747	-1.9421	1	0.05	{'AR'}	{'T1'}
Test 3	true	0.001	-8.7299	-1.9421	2	0.05	{'AR'}	{'T1'}
Test 4	true	0.001	-5.5778	-1.9421	3	0.05	{'AR'}	{'T1'}
Test 5	true	0.001	-5.2066	-1.9421	4	0.05	{'AR'}	{'T1'}

Annexe .2 : AR-Stationary 3-Dimensional VAR(4) Model

Effective Sample Size: 231
Number of Estimated Parameters: 39
LogLikelihood: -6617.1
AIC: 13312.2
BIC: 13446.4

	Value	StandardError	TStatistic	PValue
Constant(1)	1.3297	0.6549	2.0304	0.042314
Constant(2)	2.5242	0.65199	3.8716	0.00010814
Constant(3)	0	1.2758e+08	0	1
AR{1}(1,1)	-0.030827	0.070286	-0.43859	0.66096
….	……	……	……	……
AR{1}(3,3)	-0.00030511	0.076014	-0.0040138	0.9968
AR{2}(1,1)	-0.03096	0.070289	-0.44047	0.65959
…..	…….	…….	…….	……
AR{2}(3,3)	-4.1474e-07	0.076014	-5.456e-06	1
AR{3}(1,1)	-0.031134	0.070301	-0.44287	0.65786
……	……	….....	……	……
AR{3}(3,3)	0.00084454	0.076015	0.01111	0.99114
AR{4}(1,1)	0.53539	0.06971	7.6803	1.586e-14
…...	……	…….	…….	……
AR{4}(3,3)	0.29006	0.075872	3.823	0.0001318