Evaluating the Predictive Power of Random Forest Regression on Economic Growth

Ojo Gabriel O[1]✉. & Ayodele Emmanuel[2]
[1]Department of Mathematics & Statistics, Federal Polytechnic, Ilaro, Nigeria
[2]Department of Computer Science, Federal Polytechnic, Ilaro, Nigeria.
✉: gabriel.ojo@federalpolyilaro.edu.ng

ABSTRACT
This paper examines the probabilities of using the Random Forest regression model to predict the Logarithm of Gross Domestic Product (LGDP), offering a valuable resource in economic forecasting. This model demonstrated outstanding predictive accuracy, with a Mean Squared Error (MSE) of 0.0045, Root Mean Squared Error (RMSE) of 0.0674, and Mean Absolute Error (MAE) of 0.0565. These metrics confirm the model's effectiveness in capturing economic trends. Additionally, an R-squared value of 0.9430 indicates that the model explains 94.30 per cent of the variation in LGDP, highlighting its strength and reliability in economic prediction. To provide a comprehensive assessment, the Random Forest model is compared with the traditional Ordinary Least Squares (OLS) regression, which shows higher error rates (MSE: 0.0089, RMSE: 0.0943, MAE: 0.0782) and a lower R-squared of 0.8975, demonstrating the superiority of the Random Forest approach in data prediction. The Labour Growth Rate (LAGR) emerges as the most influential predictor of economic growth, followed by the Service Sector Index (LSER) and the Industrial Output Index (LIND). It is important to note that agriculture has long been a major determinant of economic performance, further reinforcing the significance of the agricultural sector in national economic outcomes. These findings suggest that policymakers should consider targeted investments and supportive policies in agriculture to promote environmentally sustainable economic growth. The Random Forest regression model proves to be an effective and widely applicable tool for economic forecasting, serving as a foundation for policy planning and decision-making. This study offers a valuable framework for decision-makers to manage economic uncertainties and foster long-term growth by leveraging the method’s capacity to handle complex, non-linear relationships in high-dimensional data.

KEYWORDS: Random Forest Regression, logarithm of gross domestic, Mean Absolute Error, R-squared, Economic Growth, Ordinary Least Squares.

1.0 INTRODUCTION

The Nigerian economy is a mixed middle income rapid growth economy due to the efforts of its manufacturing industry beside its financial industry and service industry and its communications industry and technology industry together with its entertainment industry. The African economy is the second-largest in 2023 and takes 27th and 39th places according to the levels of purchasing power parity and nominal GDP, respectively (CBN, 2023). In 2019, the economic power of Nigeria was evident since it was growing due to enhanced oil industry performance and improvement in agricultural production (OPEC, 2019). The country is an oil-producing country, and this has resulted into reliance on petroleum income which forms the foundation of government budget accounts and exports of the same to other countries. Nigeria has continued to be hamstrung by a lack of adequate political stability as well as lack of adequate infrastructure hence slow development in its economy. Oil is a significant factor in determining the Nigerian GDP advances but also exposes the Nigerian economy to the global change in the price of petroleum products. The project initiated to take the economy to a level it no longer depends on oil has been fluctuating in its performance. The various sectors such as manufacturing and services and telecommunication and agriculture are now contributing more to the GDP numbers of Nigeria in spite of current problems they are facing due to poor infrastructure and fluctuating government policies. The petroleum business is a major source of revenue to the government although it gives little percentages towards the national economic system. There is the future of economy of Nigeria basing on creation of a new national economy outside petroleum that will be centered with industrial development and service sectors and agricultural programs. Economic growth of Nigeria is dependent on closing infrastructure gaps and additional training of human capital and developing similar policy systems. The influence of the external factors such as the criteria of the world trade as well as the unstable petroleum prices and geopolitical situation will define the development of the economy of Nigeria (Micheal & Abiodun, 2024). The history indicates that the agricultural sector has sustained stability of the Nigerian economy despite the oil boom that the country faced. This economic activity (agricultural sector) fulfils its inherent economic importance in satisfying food productions and raw materials as well as employing large number of employees in addition to the national income and sustenance. The sector contributes largely in domestic and international market trade and generates marketable agricultural assets and brings about economic development by undertaking savings activities. To the Nigerian government, agriculture is viewed as its major strategy of ensuring that the nation is well fed with food and eradicating unemployment. Moreover, in the case of most developing countries, agricultural development is critical towards their social-economic development and progress as it forms a foundation towards attaining economic prosperity. Agriculture is already contributing to the economy of many developing nations in the world, among them is Nigeria. In particular, the share of agriculture in the Gross Domestic Product (GDP) of Nigeria is more than 25 percent, followed by 32 percent of the service sector, 11 percent of manufacturing, and 30 percent of agriculture (Sertoğlu et al., 2022).

Adofu & Tijani (2020), unraveled manufacturing as the process of creating goods that are to be consumed or sold by applying machine, tools, work as well as the utilization of chemicals or biologic process. It entails the development of raw or semi products into finished goods, pitting together high technologies and human skills. The industrialization; growth of industries; is so much dependent on the technological development of the method of production in the modern economy. This has been an evolution to a new system, i.e one that is modern and is governed by mass production as compared to a conventional system with a low output (Ayodele & Falokun, 2023). The service sector in Nigeria is very wide and spreading ranging in areas of banking, retail and wholesale trade, real estate, tourism, telecommunications, information and communication technology, entertainment, and education. Service industry is the most rapidly developing part of the economies of the world, serving as the key instrument of improving the employment rates and serving an active contribution to evaluation of gross domestic product (GDP) of a large number of countries.

In Nigeria, there has been a significant change in the economic situation to the point that services take over almost every aspect of life and the economy with almost 60 percent of GDP in the national economy and around 33 percent in the employment sector in 2015 (Adetokunbo & Edioye, 2020). The agricultural, services and the industry sectors in Nigeria influence the GDP of the country in a great way. Different experts in this industry have questioned the role played by agricultural practices on economic development indicators. The study is meant to provide the legitimacy regarding considering agriculture as a growth fuel. Relying on the information about 85 countries, Lavorel et al. (2019) examined how GDP per capita was interconnected with agricultural value added per worker. The results of their study encompassed a lot of conflicting information. According to the researchers, they found proof indicating that agricultural value added contributes to the economic growth pattern in developing economies and at the same time economic relationships between them are not felt in the developed economies. The tree-based algorithm is a tree-based data-driven approach through which researchers can make conclusions and predictions across different fields of economics and finances. Athey et al. (2019) interpret the traditional kernel weighting function with a new approach to weight based on RFR that is adaptive. This study by Gu et al. (2020) and Ng (2021) illustrates that RFR can be used to predict returns on stock market using firm-specific and common variables and economic recessions respectively. It has been studied that RFR performs very efficiently in the high-dimensional setting since it has the feature of identifying significant predictors but emerging studies advise to restrict the initial number of predictors prior to the application of any forecast method. RFR has been used in healthcare as part of disease prediction and risk stratification challenges, where electronic health records (EHRs) and genomic data can be used to further support personalized medicine (Rajkomar et al., 2019). RFR is especially suitable in consolidating a variety of data modalities and enhancing predictive applications in healthcare due to the fact that it can accommodate heterogeneity of data sources and nonlinear relationships.

In the transformation of the Nigerian economy, Olusoji (2019) looked into Service sector as a potential in the transformation of the country concerning some macroeconomic variables. They noted that agricultural sector contributes more to the economy of Nigeria, as compared to service sector. The manufacturing industry has a higher productivity than other industries do. Valeria et al., (2024) made a study of how the containment measures, the rate of infection, cumulative deaths, and vaccination affect the growth of GDP. The research results indicated that Regression Tree and Random Forest models were able to reasonably predict GDP growth, Covid-19 induced mortality, and the reproduction rate of the virus, which further influenced the magnitude of the ML models applying regressions on various outcomes.

The research by Abdulloh et al., (2024) scrutinised the impact of interest rates and inflation in economic growth in Indonesia and in addition, compare the level of performance of machine learning models using Random Forest and XGBoost in the analysis of inflation effects. The deduction made in their results revealed that Random Forest is superior to XGBoost in the prediction of the effects of inflation on economic growth rates, inflation, and economic growth. Meher et al. (2024) explored how well the Random Forest model performed in making price predictions of stock prices of the four largest solar energy firms in India by considering data of the daily stock prices. The predictive model achieved high levels of accuracy through high levels of R 2 (0.9928 0.9939) and lower values of MAE in all the firms considered in the study. This showed that it was dependable in exact forecasting of stock prices.

Senoussi (2021) offered an analysis with the aid of Random Forest (RF) in accordance with the growth model of Barro to examine the non-linear affiliation amid inflation variability and economic growth as well as inflation levels. A researched study finds that intense inflation and low inflation threatens to create negative effects on an economy hence affecting key central banks inflation- targeting policies. Aysan et al. (2024) applied the random forest methods in analyzing risk management in Islamic banking since they used survey data. Large enterprises have more stable feedback in banks because they are more confident and have enough financial resources. Therefore, the research shows a positive correlation between the level of economic development (with high GDP and low inflation and interest rates) and predictability of risk-related evaluation and the fact the risk of credit portfolios decreased and the issues of funding terrorism and cybersecurity were one of the concerns because of the better regulation and investments. Anesti et al. (2021) evaluate the predictive capacity of large survey panels on GDP development in the UK utilizing the linear regularization and machine learning models, such as Random Forests, Support Vector Regressions (SVR) and Neural Networks. According to their results, Ridge and Partial Least Squares models outstand among other linear toolsets, SVR is the most accurate model in short-term forecasting, and Neural Networks and Random Forests could be better in long-term ones. Consequently, the purpose of the research study is to determine the correctness and predictive power of Random Forest Regression model in predicting the economic growth of the Nigeria, along with the comparison of a conventional regression technique to legitimize the correctness of this model. superior predictive power.

2.0 MATERIALS AND METHOD

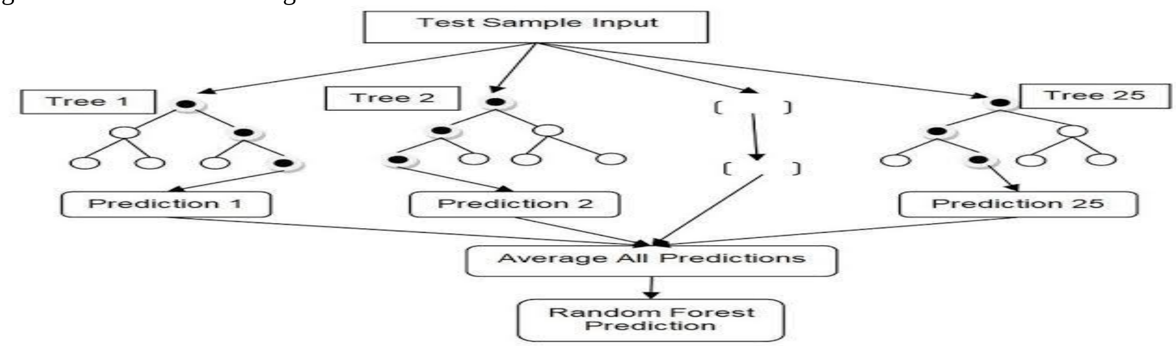
School statistics will be used in this study based on time series data between 2002 and 2021 and the data is provided by the Central Bank of Nigeria (CBN) statistical database accessed in the National Bureau of Statistics (NBS) Gross Domestic Product (GDP) template. This paper uses the Random Forest Regression method to measure the relative relevance of the economic sectors (Agriculture, Industry and Services) in interpreting the macroeconomic variables and future GDP as an indicator of economic development by using R-studio package. In order to curb possible pitfalls of Random Forest, the performance of an Ordinary Least Squares (OLS) regression, an old methodology, is provided as a comparative measure to the predictive capabilities of Random Forest.

2.1 Model Identification and Estimation

2.1.1 Random Forest Regression

The Random Forest Regression Model generates lots of decision trees based on different parts of data as revealed by [Liaw, A., & Wiener, M. (2002); Breiman (2001)] and Breiman (2001). The decision trees ensemble works best by using bootstrapping methods of feature randomization processes and offers more precise and non-overfitting predictions. In the random forest system, the predictions harvested decision tree are chosen until the end prediction is arrived through the marriage of most common forecasts made by separate trees. Because of this approach, one decision tree is not overfitted. The sampled data collected is based on the classification of the outcome of the classification of data subjected to a simple majority vote.

Figure 1: Random Forest Regression Tre



e

Source: Pangarkar et al., 2020

2.1.2 Ordinary Least Squares (OLS) Regression

The Ordinary Least Squares (OLS) regression is a conventional linear regressional model that approximates its relationship between the dependent variable (LGDP) and independent variables (LAGR, LSER, LIND) based on a minimized sum of squared residuals. OLS model model supposes the linearity, independence, homoscedasticity, and unpredictability of errors. OLS is widely used, in spite of the possibility of its violation of the assumptions in the economic data, because it is simple, and its interpretation is straightforward. It is estimated the same data used to estimate Random Forest model should be used to estimate the OLS model so that a balanced evaluation can be made. Model:

$LGDP = \beta_0 + \beta_1 LAGR + \beta_2 LSER + \beta_3 LIND + \varepsilon$

β0 is the intercept; 1 indicate the coefficient of LAGR, 2, the coefficient of LSER and 3, the coefficient of LIND and 1 is the error term. The evaluation of the OLS model uses the same metrics as Random Forest (they are Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared).

2.1.3 Performance Metrics

1. **Mean Squared Error (MSE)**

It is applied as an index of measurement of the model performance and assists in measuring the predicting error of the regression models and gives an indication of the reliability and a measure of performance.

$MSE = (1/n)\, \Sigma(y_i - \hat{y}_i)^2$

2. **Root Mean Squared Error (RMSE)**

It is a metric that is used widely to determine whether the prediction of a model is accurate or not. It is the square root of the mean of the squared value of the difference between the forecasted and the actual values.

$RMSE = \sqrt{[(1/n)\, \Sigma(y_i - \hat{y}_i)^2]}$

3. **Mean Absolute Error (MAE)**

It is one of the measures of the accuracy of a regression model. Through an average error degree between forecast and recorded data. Smaller MAE means high performance where the predicted value is near to the actual one.

$MAE = (1/n)\, \Sigma|y_i - \hat{y}_i|$

4. **R-Squared (R^2)**

It is a measure to test the wellness of explaining variance of dependent variable in the regression models over the mean of the dependent variable. R-squared is a number between 0 and 1 with the larger number representing a superior match between the model and the data.

$R^2 = 1 - (\Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2)$

Where; n represents the number of the observations, y i is the true value, y i is the forecast value, y is the average of the true values, Σ(y i - y i) 2 is the residual errors of squares, and Σ(y i - y) 2 is the total sum of squares.

3.0 RESULTS

Table 1: Summary of Random Forest Regression Model Performance Metrics

METRICS	VALUE
Type of Random Forest	Regression
Number of Trees	500
Variables Tried per Split	1
Mean of Squared Residuals	0.006398178
Tree with Minimum MSE	228
Percentage of Variance Explained	94.85%

Source: Author's computation using RStudio

Table 1 indicates that the Random Forest regression model has performed remarkably well in the ability to predict LGDP with the value of mean of squared residuals (MSR) = 0.006398178 which denotes that the model has settled closely in terms of forecasting and the value. This model explores 94.85 percent of the variance in the LGDP and this clearly shows that the predictor variables are capable of representing variability in the model. This degree of accuracy and variance explained highlights how solid and effective the Random Forest model is in detecting significant patterns within the data making us more confident in the strength of forecasting the data using the model. It is important to note here that the model attains the minimum mean squared error at the 228th tree implying the most optimal trade off between the complexity and accuracy. This is despite there being 500 trees in the model, whereby the error level becomes stabilized after the 228 th tree meaning that more trees in a model do not make much difference in terms of accuracy. Such a result will be a good rule of thumb in future modeling activities because it suggests that optimal performance could be achieved using a relatively small number of trees, (say about 228 trees) and this will make progress much less time consuming and computationally expensive.

Figure 1: Random Forest Model Convergence Plot

The above figure proves successful in illustrating that Random Forest model rapidly approaches a level of diminishing returns at which the extra trees would not essentially enhance the performance of the model. This can be supported with the fact that the error rate was stabilized. The initial steep decay in error is a relative indication of the capability of the model to fit significant data structures with relative less trees. The stable and reliable nature of the model is strengthened by the plateau since it establishes that the predictions will not vary regardless of whether more trees are involved. Thus, these empirical findings conclude on the efficiency and effectiveness of the model in foreseeing the expansion of the economy of any nation.

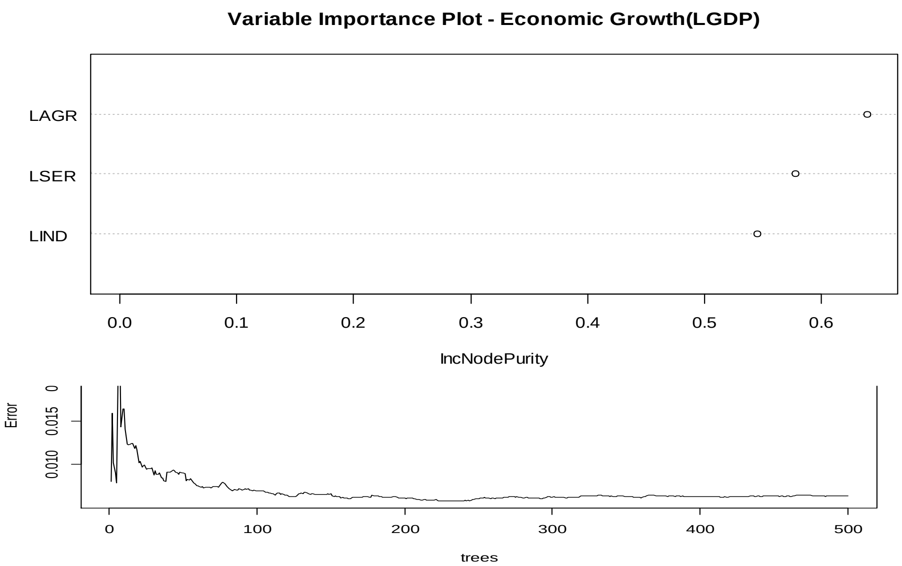


Figure 2: Variable Importance Plot for Economic Growth (LGDP)

As it is seen in figure 2, LAGR is the main determinant of economic growth (LGDP) in the data set, then LSER and LIND. The huge disparity on the measure of increase on the node purity of LAGR compared with the other variables indicates how instrumental the agricultural sector is in shaping the economic growth.

Table 2: Variable Importance Based on Increase in Node Purity

Riable	IncNodePurity
LAGR	0.6396
LIND	0.5451
LSER	0.5782

Source: Author's computation using RStudio

The Table 2 shows that LAGR is the biggest determinant of the prediction of economic growth (LGDP) then LSER and LIND. The great significance of LAGR is an indication that agricultural performance is of great influence to the economic growth and there is a call to policies and investment in agricultural growth. The comparison reveals that LSER and LIND are various sectors which contribute significant economical contributions to the growth of a country. The industrial sector has the least delusion on economic growth than the agricultural and service sector according to the LIND indicator analysis.

Performance Metrics	Random Forest Value	OLS Value
Mean Squared Error (MSE)	0.0045	0.0089
Root Mean Squared Error (RMSE)	0.0674	0.0943
Mean Absolute Error (MAE)	0.0565	0.0782
R-Square (R^2)	0.9430	0.8975

Table 3: Comparative Predictive Performance Metrics
Source: Author's computation using RStudio

Table 3 shows the predictive measures of the two models Random forest and OLS regression. The random forest model has given a better result with its low Mean Squared Error (MSE) of 0.0045, Root Mean Squared Error (RMSE) of 0.0674 and Mean Absolute Error (MAE) of 0.0565. In comparison, the OLS model has larger errors and has an MSE, RMSE and MAE of 0.0089,0.0943 and 0.0782 respectively. The Random Forest model also has a higher R squared of 0.9430 which explains how 94.30 percent of the variance in LGDP is explained as compared to the R squared of 0.8975 in the OLS model which explains how 89.75 percent of the variance in LGDP is explained. These criteria bring out the high level of accuracy and capability of the Random Forest model to reflect non-linear and complex relationships in the data compared to the linear OLS model. The minimal error rates and good explanatory strength of Random Forest substantiate its application as an efficient instrument of economic forecasting.

3.1 Discussion of Results

Table 1 indicate the measures evaluation of Random Forest regression, which predicts Logarithm of Gross Domestic Product (LGDP). There is a very high agreement between the imitation of the actual LGDP values and the value of the Mean of Squared Residuals value of 0.006398178. LGDP variance that model explains is 94.85 percent. Most of the LGDP underlying factors are incorporated well using the model. The number of trees, with the minimum value of Mean Squared Error, is at the 228 th tree and the model efficiency is optimal. The maximum accuracy in the model is at 228 trees and the use of 500 trees does not give a better performance which indicates the fitting of model simplicity with successful operation. Based on the available research recommendation, a predictive modeling technique should be selected on the basis of about 228 trees since it injects an optimal level of accuracy as well as uses minimal computational resources. The character of the convergence of the model presented in Figure 1 was illustrated by the Random Forest model and its behavior. The model with a view of having the greatest accuracy takes only the first 228 trees to reign and this shows that more trees do not have a great impact on the model in terms of its performance outcome. The model proves the capability of data structure capturing since it shows accelerated rate of error reduction achieved by employing a small number of trees. The plateau phase on the model shows that it is even reliable since it shows consistency in performance as more trees are added to the structure.

Table 3 shows the comparison of the predictive performance of Random Forest and OLS regression models. It can be seen that Random Forest perform much better than OLS on any of the metrics, with improving MSE (0.0045 vs. 0.0089), RMSE (0.0674 vs. 0.0943), and MAE (0.0565 vs. 0.0782), as well as higher R-squared (0.9430 vs. 0.8975). This implies that Random Forest provides a superior way of describing the non-linear and complex connections existing in the economic data, whereas OLS, which has its limitations towards a non-linear representation (being linear in nature), is not the best fit. Superiority

of Random Forest further goes to emphasize that it serves well when working with high-dimensional and non-linear economic data in order to supply more accurate LGDP forecasts.
Table 2 and Figure 2 indicates the predictive value of numerous variables towards economical development forecast. Through the analysis, it has been revealed that LAGR, the agricultural industry propels LGDP in its peak high capacity where LSER, the services sector ranks second and LIND, the industrial sector ranks third. The high impact of the LAGR on the economy is the main argument providing the vital role of the LAGR in the sphere of the economy growth as the loss in agricultural performance seriously influences general financial results. The figures show that the influence of the industrial sector is not so great as the influence of the agricultural and services sectors, but LIND and LSER contribute to economic growth to a great extent. The findings of the research point to the fact that the properly orchestrated investment in the agricultural development will have a potent impact to influence the economic growth. The predictive performance results in the Random Forest model are determined based on the following summary (in Table 3). The low MSE of 0.0045, RMSE of 0.0674 and MAE of 0.0565 are indicators of predictive reliability of the model as it gives accurate forecasts of LGDP. R-squared is 0.9430, thus letting it know that the model is able to explain 94.30 percent of variance in LGDP and 5.70 percent of variability in the rest of the data. The introduction of the model is accurate in terms of the results it displays in its strength of stability and operations excellence needed as an instrumental tool in economic forecasting. The model proves valuable in LGDP forecast because it has a high explanatory power with small error levels and this qualifies it as an effective tool in formulation and evaluation of economic policies especially in comparison to other conventional analysis tools such as OLS.

4.0 CONCLUSION

Random Forest regression model is also very good in predicting the Logarithm of Gross Domestic Product (LGDP) by producing a Mean Squared Residuals of 0.006398178 with 94.85 percent variance explanation as compared to the Ordinary Least Squares (OLS) regression model which produces a higher MSE (0.0089) and R-squared (0.8975). Random Forest model proves to be competent in identifying real variation in LGDP as the measures of error (RMSE: 0.0674, MAE: 0.0565) are lower than those of the OLS (RMSE: 0.0943, MAE: 0.0782). It was observed that a number of trees of 228 in the Random Forest model achieved the most efficient property dealing with optimal performance per unit of complexity and efficiency. The effectiveness of the model is determined by the plot of convergence of the model and it indicates that very few trees could be used to make accurate predictions. The variable importance graph illustrates that the agriculture sector (LAGR) continues to be the primary source of economic growth as the services (LSER) and the industry (LIND) sectors come second, thus laying emphasis on the critical role played by agriculture. This comparison proves that the Random Forest method is more appropriate to complex and non-linear non-linear relationships in economic data which makes it an outstanding method in predicting any economic phenomenon. The information on drivers of economic growth obtained as a result of the present research can become a perfect ground of economic policy creation and long-term research and strategic development.

5.0 REFERENCES

Abdulloh, F F., Afrig, A., Majid, R., & Fetrus, J.H. (2024). Performance Analysis of Machine Learning Algorithms Using the Ensemble Method in Predicting the Impact of Inflation on Indonesia's Economic Growth. Int. J. Inform. Visualization, 8(4). pp. 2316-2323.

Adetokunbo, A., & Edioye, A. (2020). Response of economic growth to the dynamics of the service sector in Nigeria. Future Business Journal, 6(1), 27.

Adofu, I., & Tijani, Y. (2020). Manufacturing sector and economic growth in Nigeria (pp. 1–6).

Anesti, N., Eleni, K. & George, K. (2021). Forecasting UK GDP growth with large survey panels (Staff Working Paper No.923). Bank of England.

Athey, S. J., Tibshirani, R., & Wager, S. (2019). Generalized random forests. The Annals of Statistics, 47, 1148–1178.

Ayodele, G. O., & Falokun, M. (2023). The Nigerian economy: Structure and pattern of development. JODAD.

Aysan, A. F., Bekir, S. C., & Ibrahim, M.U. (2024). Predictive Power of Random Forests in Analyzing

Risk Management in Islamic Banking, Journal of Risk and Financial Management 17:104. Pp.2 – 19.
Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5 – 32.

Central Bank of Nigeria (2023). Central Bank of Nigeria statistical bulletin (Vol. 28).

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33, 2223–2273.

Lavorel, S., Storkey, J., Bardgett, R. D., Bello, F., & Harrington, R. (2019). A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. Journal of Vegetation Science, 24(5), 942–948.

Liaw, A., & Wiener, M. (2002). Classification and regression by Random Forest. R News, 2(3), 18–22.
Meher, B.K., Abhishek, A., Sunil, K., Ramona, B & Manohar, S. (2024). Effectiveness of Random Forest Model in Predicting Stock Prices of Solar Energy Companies in India, International Journal of Energy Economics and Policy, Vol 14, Issue 2. Pp. 426 – 434.

Micheal, A., & Abiodun, O. (2024). Impact of agricultural variables on gross domestic product (GDP) in Nigeria: A SARIMA approach. African Journal of Agricultural Science and Food Research, 14(2).

NG, S. (2021). Boosting recessions. Canadian Journal of Economics/Revue Canadienne d'Économique, 47, 1–34.

Olusoji, M. A. (2019). Human capital investment and economic growth in Nigeria: The role of education and health manager. 15(1), 190–201.

OPEC. (2019). Organization of the Petroleum Exporting Countries Annual Statistical Bulletin (ABS), 54th edition.

Pangarkar, D.J., Rajesh, S., Amita, S., & Madhu, S. (2020). Assessment of the Different Machine Learning Models for Prediction of Cluster Bean (Cyamopsis tetragonoloba L. Taub.) Yield. Advances in Research. 21(9): 98-105, ISSN:2348-0394.

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., & Dean, J. (2019). Scalable and accurate deep learning for electronic health records. NPJ Digital Medicine, 1(1), 1–10.

Senoussi, H. (2021). Inflation and Inflation Uncertainty in Growth Model of Barro: An Application of Random Forest Method, International Econometric Review (IER), 13(1), 4 - 23.

Sertoğlu, K., Acar, M., & Ünlü, K. (2022). The contribution of the agricultural sector to economic growth in Nigeria. International Journal of Economics and Financial Issues, 7(1), 547–552.