Economic Growth and Fiscal Policy in Peru: Prediction Using Machine Learning Models

Fidel Huanco Ramos[1], Yesenia Valentin Ccori[2], Henry Shuta Lloclla[3],
Martha Yucra Sotomayor[4], Ilda Mamani Uchasara[5]
Universidad De San Martín De Porres, Arequipa, Perú[1]
Universidad Nacional De San Antonio Abad De Cusco, Cusco, Perú[2]
Universidad Nacional Del Altiplano, Puno, Perú[3, 4]
Universidad Peruana Unión, Juliaca, Peru[5]

***Abstract*—The empirical literature presents several indicators related to fiscal policy and economic growth. The paper aims to predict Peru's economic growth using fiscal policy variables. For this purpose, open data from the Central Reserve Bank of Peru was used, data preprocessing and the study used Python programming through Google Colab to evaluate eight machine learning models. Metrics such as Root Mean Square Error (RMSE), Mean absolute error (MAE), Mean square error (MSE), and Coefficient of Determination (R^2) were used to measure their performance. In addition, SHapley Additive exPlanations (SHAP) was applied to interpret the importance of macroeconomic variables. The results show that the K-Nearest Neighbors (KNN) model obtained the best performance, with an R^2 of 0.972 and low prediction errors. In the same way, important variables in fiscal policy such as Net Debt, Liabilities, and Interest on External Debt were identified. In conclusion, the study shows that KNN and Ensemble Bagging are highly effective models for predicting Peru's economic growth.**

Keywords—Machine learning; predictive models; fiscal policy; economic growth

I. Introduction

The relationship between fiscal policy and economic growth has been the subject of study for decades, constituting an important topic in contemporary economics. Fiscal policy, understood as the set of decisions related to taxation and public spending, has a direct influence on economic activity and social welfare. According to Barro [1], efficient management of fiscal policy can stimulate long-term economic growth, while an inappropriate approach can lead to imbalances and recessions. Furthermore, according to Caprioli [2], considered a closed production economy, fiscal authority, infinitely live agents, and the absence of capital, provides a simplified framework for analysing the effects of fiscal policies in an economy.

The government has to match an exogenous stream of public spending through proportional taxes on labor income and government bonds. It pursues optimal taxation, given the initial amount of debt, it chooses policy instruments to maximize consumer welfare. However, the analysis of the impact of these policies has become more complex due to increased volatility in markets and non-linear interrelationships between economic variables [3], [4].

Given the characteristic non-linear behavior of many economic variables, new methodologies based on artificial neural networks have been explored since the 1990s [5]. They therefore offer a significant advantage by allowing the modelling of both linear and non-linear relationships between the input and output variables of a system. In this perspective, it has proven valuable in highly volatile environments, such as financial markets, where economic variables often exhibit complex and non-linear patterns [6]. In this sense, the technology contributes to dynamic macroeconomic forecasting [7], under the approach of scientific, reliable, and complete historical statistics.

Machine learning and artificial intelligence have opened up new possibilities for economic analysis. Models such as K-Nearest Neighbors (KNN), Ensemble Bagging, and Random Forest allow large volumes of data to be processed and non-linear interactions to be modelled with greater accuracy. In addition, the incorporation of explanatory techniques such as SHAP (SHapley Additive exPlanations) facilitates the interpretation of models, offering a deeper understanding of the factors driving economic growth and has positioned itself as an innovative tool for managing large volumes of data and uncovering hidden patterns [8], [9]. This methodology has proven to be effective in a number of fields, including economics, where it can provide more accurate and adaptive predictions for public policy formulation [10], [11]. In addition, studies suggest the use of machine learning algorithms to improve the ability to forecast and predict the impact of fiscal policies on economic growth, overcoming the limitations of conventional econometric models [12], [13].

The main objective of the study is to predict Peru's economic growth using fiscal policy variables. To this end, it seeks to rigorously evaluate how the main fiscal policy variables, such as public spending, tax collection, and the fiscal deficit, influence the country's economic development. Furthermore, this study not only seeks to contribute to the understanding of the relationship between fiscal policy and economic growth in Peru, but also to offer practical recommendations that can guide fiscal decision-making to promote sustainable economic development through the use of machine learning.

II. Literature Review

A. Artificial Intelligence Theory

The application of artificial intelligence (AI) techniques in the prediction of fiscal policies and economic growth has experienced exponential growth, consolidating itself as a highly relevant field of research [14], [15]. AI is used to analyze large volumes of data, identifying patterns and trends that enable a country's decision-making. Machine learning algorithms can analyze historical data such as tax collection, public spending, and economic indicators to predict how these factors will behave under different scenarios. According to Kaliuzhniak [16], this research explores the use of AI, neural networks, and machine learning in economic forecasting, highlighting the importance of making appropriate decisions in social management. Similarly, Aliyev [17], discusses how AI, machine learning and data mining are applied in economic modelling and prediction of economic growth and fiscal policy in the context of the Industrial.

B. Theory of Fiscal Policy

Fiscal policy, combined with monetary policy, has long been one of the main instruments used by governments to intervene and influence economic activity [18]. Thus, depending on economic conditions, public expenditures and taxes are used to influence the economy in the direction of expansion or contraction [19]. However, the views and theories on the effectiveness of this fiscal instrument have often been contradictory [20].

Thus, John Maynard Keynes provided a theoretical foundation for the use of fiscal policy, demonstrating that public expenditures and taxes are effective tools for regulating economic cycles [21]. According to this theory, insufficient aggregate demand is the primary cause of economic recessions. Therefore, increasing public expenditures or encouraging private spending through tax cuts enhances purchasing power and, consequently, consumption, stimulating short-term economic growth [22], [23]. In fact, for Keynesians, there is a relationship between the level of expenditure and national income (and, consequently, employment), as an increase in expenditure stimulates household consumption and encourages producers to expand their production to meet additional demand, thereby creating jobs. According to Xin Li [24], fiscal policies have been and will continue to be an essential component in mitigating the effects of the pandemic on the Chinese economy. However, their impact will be gradual and a balance between fiscal stimulus and pandemic control measures will need to be maintained.

C. Economic Growth Theory

It is the sustained increase in the production of goods and services in an economy over time. It is generally measured through real Gross Domestic Product (GDP), which represents the total value of final goods and services produced in a country during a given period, adjusted for inflation [25] . Likewise, the Solow-Swan model is an exogenous economic growth model, while the endogenous growth theory explains economic growth from factors internal to a country [26]. The latter emphasizes the importance of capital accumulation and technological progress as engines of economic growth, without relying exclusively on external factors such as foreign investment or international trade [27]. Similarly, the Harrod-Domar model is an economic model that explains how the rate of investment influences productive capacity and aggregate demand [28]. It focuses on investment as the main driver of economic growth, highlighting the relationship between savings, investment, and GDP growth [29]. In addition, the British economist John Maynard Keynes said that capital formation is a fundamental element of growth analysis, and his argument is simple. Investment will increase employment levels. According to the principle of effective demand, investment is a variable that reduces the gap between the level of income or the level of production and the level of consumption [30].

D. Limitations and Contributions of this Study

While the literature demonstrates significant advances in the application of artificial intelligence and economic models for the analysis of economic growth and fiscal policy, limitations persist that warrant the development of new approaches [31]. Most studies focus on international contexts or developed economies, neglecting particularities of developing countries such as Peru. In addition, many studies omit important fiscal policy variables, which limits the accuracy and applicability of predictive models. Likewise, there are few explanatory techniques that allow interpreting the results of predictive models, which limits their usefulness in governmental decision-making. Faced with these gaps, the present study proposes an approach that incorporates multiple machine learning algorithms, together with the use of SHAP as an explanatory technique, applied specifically to the Peruvian context [32]. This methodology not only improves the ability to predict economic growth, but also makes it possible to interpret the relative impact of fiscal variables, providing valuable tools for the formulation of more effective public policies.

III. MATERIALS AND METHODOLOGY

The materials used in the research include open data from the Central Reserve Bank of Peru (BCRP) for the period 1990-2023, consisting of annual time series on fiscal policy and economic growth. Implemented with the following libraries: Numpy, Pandas, Matplotlib, Seaborn, SHAP, and Scikit-learn in colab Google Python. Based on four processes:

A. Data and Variables of the Study

The data for the study were obtained from the Banco Central de Reserva del Peru (BCRP), since the data are reliable and consist of a set of economic resources of the central government, allocated to various sectors of the executive branch in order to meet their objectives, such as the maintenance of infrastructure, the provision of public services and the payment of debts, among others. Public spending, tax revenues, public investment and public debt are fundamental components of fiscal policy, which directly influence economic growth. It is also divided into various categories and subcategories that allow a detailed breakdown of the use of resources. In this case, the following variables are identified: Central Government Non-Financial Expenditure.

Central Government Remuneration

Central Government Goods and Services

Central Government Transfers

Central Government Capital Expenditure

Central Government Gross Capital Formation

Other Central Government Capital Expenditure

Central Government Total Interest

Interest on Central Government Domestic Debt

Interest on Central Government External Debt

Total Central Government Expenditure

Non-Financial Expenses of the Rest of the Central Government

Non-Financial Current Expenditure of Rest of Central Government

Capital Expenditure of Rest of the Central Government

Interest of Rest of the Central Government

Tax Revenue of Rest of the Central Government

Income Tax

Wealth Tax

Tax on Exports

Tax on Imports

General Sales Tax - IGV

IGV - Internal

IGV - Imports

Selective Consumption Tax - ISC

ISC - Fuels

ISC - Others

Other Tax Revenues

Refunds

Non-Tax Revenues

Total Central Government Current Revenues

Public Sector - Public Investment

Assets

Liabilities

Net Debt

Gross Domestic Product (Real GDP)

B. Data Pre-processing

It ensures that Machine Learning models work with appropriate and well-structured data. To do this, it starts with loading the dataset from an Excel file stored in Google drive, followed by the separation of the predictor features (X) and the target variable (y). Subsequently, the dataset is split into training (80 per cent) and testing (20 per cent) using train_test_split to ensure adequate generalization of the model. To standardize the features, StandardScaler is applied, fitting the transformation with the training data and applying it to the test set, ensuring that all variables have mean zero and standard deviation one. This process improves the numerical stability and performance of Machine Learning models by eliminating inconsistent scaling in the data.

C. Training of Machine Learning Models

The selection of machine learning models for the analysis focused on those capable of predicting various economic variables, based on a set of fiscal policy indicators that include public spending, tax revenue, public investment, and public debt. The models are presented below:

Gaussian Process Regression (GPR) is a powerful statistical technique with a Bayesian approach, used in machine learning and data analysis to predict unknown values from observed data [33]. Its ability to model uncertainty and capture complex relationships makes it ideal for a wide range of applications [34].

Given a training data set {X, y}, the model assumes that the output values y follow a joint Gaussian distribution:

$$y \sim N(\mu(X), K(X,X) + \sigma^2 I) \tag{1}$$

where, (μ(X) is the mean function, usually assumed to be 0, K(X, X) is the kernel-based covariance matrix and $\sigma^2 I$ I is the noise variance (I is the identity matrix).

Then, it is to predict a new point X*, the conditional distribution is:

$$\hat{y}^* = K(X^*, X)K(X,X)^{-1} \tag{2}$$

where, $K(X^*, X)$ represents the covariance between the new points X* and the training data, $K(X,X)^{-1}$ y is the estimate of the values based on the previous data.

Fine Trees are a type of supervised learning model based on decision trees, where the tree is allowed to grow to a considerable maximum depth (in this case, max_depth=10) to capture complex relationships without overfitting. It is based on recursive feature space partitioning using impurity reduction. This model is useful in time series, economic forecasting, pattern recognition, and tabular data analysis [34].

$$f(X) = \sum_{i=1}^{N} c_i 1(X \epsilon R_i) \tag{3}$$

where, N is the total number of terminal regions.

Random Forest is a machine learning algorithm based on a set of decision trees that improves the accuracy and stability of predictions by reducing overfitting [35]. In this study, a Random Forest machine learning algorithm is used to estimate economic growth in Peru, which is considered a function approximation (regression) problem.

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(X) \tag{4}$$

where, $\hat{y}$ represents the predicted economic growth, N is the total number of trees in the forest and $T_i(X)$ is the prediction of $i - \acute{e}simo$

The Linear Support Vector Machine (SVM) is a machine learning model that seeks to find an optimal hyper plane to separate data into different classes (classification) or make numerical predictions (regression) [36]. In regression (Linear SVR), instead of classifying, the model finds a hyperplane that minimizes the error within a margin ϵ, ignoring small deviations and penalizing larger errors [37]. Its main advantage is its ability to handle high-dimensional data and avoid overfitting by regularizing its parameters.

$$f(X) = w^T X + b \tag{5}$$

where, *w* and *b* are the parameters that are tuned to minimize the error. To avoid overfitting, Linear SVR employs an insensitive loss function ϵ and a hyperparameter C that controls the penalty for out-of-range errors. This model is useful in economic forecasting, time series, and financial analysis problems, as it offers a robust solution to data variability.

The MLPRegressor (Artificial Neural Network for Regression) model is a multi-layered artificial neural network used for regression tasks. It consists of an input layer, one or more hidden layers, and an output layer, where each neuron applies a nonlinear transformation to the data using activation functions such as ReLU or sigmoid.

Given an input set X, the three-layer (input, hidden, and output) neural network is defined as:

$$h^l = f(W^l h^{(l-1)} + b^{(l)}) \tag{6}$$

where, h^l is the activation of layer *l*, W^l is the weight matrix of layer l, b^l is the bias vector and f(.) is the activation function.

For the output layer in regression, the final prediction is:

$$\hat{y} = W^L h^{(L-1)} + b^{(L)} \tag{7}$$

where, ***L*** is the final layer.

The KNeighborsRegressor is a machine learning model based on the k-Nearest Neighbors (KNN) algorithm, used for regression tasks [9], [38]. Instead of learning an explicit feature during training, it stores the data and predicts the value of a new instance by calculating the average of the k nearest neighbors in the feature space [39].

For an input X, the prediction is defined as:

$$\hat{y} = \frac{1}{k}\sum_{i=1}^{k} y_i \tag{8}$$

where, k is the number of neighbors y y_i are the target values of the k nearest neighbors. The model uses a distance metric (by default, Euclidean) to find the nearest neighbors and may weight their contribution according to closeness.

The BaggingRegressor is an ensemble learning model based on the Bootstrap Aggregating (Bagging) technique, which improves the accuracy and stability of regression models by reducing the variance [40].

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \tag{9}$$

$\hat{y}$: Final prediction of the Bagging model.

B: Total number of base models (estimators).

$f_b(x)$: Base model prediction bbb for an input x

$\sum_{b=1}^{B} f_b(x)$: Suma de las predicciones de todos los modelos base

$\frac{1}{B}$: Average of the predictions.

Under this approach it helps to reduce the variance of the base model, making it more robust and stable compared to a single regression model.

A "Coarse Decision Tree" refers to a decision tree with limited depth, meaning it has a small number of levels or splits [41].

$$\hat{y}(x) = \sum_{m=1}^{M} c_m . 1(x \epsilon R_m) \tag{10}$$

$\hat{y}(x)$: Decision tree prediction for an input *x.*

M: Total number of regions (or leaf nodes) created by the tree. R_m: *m-th* region into which the feature space is divided.

c_m: Output value for the region R_m (usually the average of the training values in that region).

$1(x \epsilon R_m)$: Indicator function, which is *1* if *x* belongs to the region R_m and 0 otherwise.

D. Model Evaluation

i. Root Mean Square Error (RMSE)

RMSE is a widely used metric that measures the average magnitude of inter-predicted errors y_i and observed $\bar{y}_i$ values [42]. The accuracy of the model is assessed comprehensively, with lower values indicating better performance [43].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{11}$$

where, ***N*** is the total number of observations.

ii. Mean absolute error (MAE)

MAE quantifies the mean absolute difference between predicted and predicted$\bar{y}_i$ and observed y_i values, providing a measure of the accuracy of the model without considering the direction of the errors [44].

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}_i| \tag{12}$$

where, N is the total number of observations.

iii. Mean square error (MSE)

MSE measures the average of the squares of the inter-predicted errors $\bar{y}_i$ and observed y_i values, providing a measure of the model's accuracy that emphasizes errors larger than MAE [45].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \tag{13}$$

iv. Coefficient of Determination (R^2)

Often denoted as R^2, assesses the proportion of the variance in the dependent variable (economic growth) that is predictable from the independent variables (public expenditure). It ranges between 0 and 1, where higher values indicate better explanatory power [46], [47].

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{14}$$

where, N is the total number of observations y_i is the observed value, $\bar{y}_i$ es el valor, Foreseen $\bar{y}_i$ is the mean of the observed values.

These performance indices jointly assess the predictive accuracy of the model, highlighting different aspects of the prediction errors

E. Interpretation of Results

SHAP is used to interpret the importance of variables and understand their impact on predictions [48]. That is, it is a technique based on cooperative game theory that is used to explain the predictions of machine learning models. It aims to assign an importance value to each variable in a prediction, indicating how much each variable contributes to the final outcome of the model.

IV. RESULTS

Table I presents the performance metrics of the evaluated models. It is observed that the K-Nearest Neighbors (KNN) model obtained the best performance, with an R^2 of 0.972 and low errors (RMSE: 507.67, MAE: 479.79, MSE: 2.577), followed closely by Ensemble Bagging, which presents the best performance, with significantly low error values (RMSE, MAE, and MSE) and a coefficient of determination R^2 of 0.971 respectively, indicating excellent predictive capacity. The Fine Trees and Decision Tree-Coarse models also shows good performance with R^2 of 0.70 and 0.86, respectively. On the contrary, Gaussian Process Regression, Support Vector Machine - Linear, Neural Network, and Random Forest presented poor results, with negative R^2 values, which indicates that these models failed to correctly capture the relationship in the data and have poor predictive capacity, as seen in Table I. Therefore, the K-Nearest Neighbors (KNN) model is selected as the best model due to its higher coefficient of determination and better performance in error metrics.

Fig. 1 presents four bar charts comparing the performance of different models using statistical metrics such as RMSE, MAE, MSE, and R^2. One specific model is observed to have high RMSE, MAE, and MSE values, indicating that its predictions are highly inaccurate compared to the others. Furthermore, the R^2 chart shows a significant negative value, suggesting that the model performs worse than a simple mean of the data. These results highlight the importance of choosing models with lower error and greater explanatory power to ensure more accurate predictions in the analysis of economic growth and fiscal policy in Peru.

TABLE I. PERFORMANCE METRICS (R^2, RMSE, MSE, MAE) OF EIGHT MACHINE LEARNING ALGORITHMS IN ECONOMIC GROWTH PREDICTION

Algorithms	**RMSE**	**MAE**	**MSE**	**R2**
Gaussian Process Regression	7112.721647	5947.319059	5.059081 e+07	-4.450982
Fine Trees	1651.050643	1244.274174	2.725968 e+06	0.706287
Random Forest	640502.801882	242370.227261	4.102438 e+11	-44201.331238
Support Vector Machine - Linear	4494.461921	3850.421271	2.020019 e+07	-1.176499
Neural Network	93653.049841	68484.708929	8.770894 e+09	-944.032962
K-Nearest Neighbors	507.668979	479.793237	2.577278 e+05	0.972231
Ensemble bagging	514.511294	417.417693	2.647219 e+05	0.971477
Decision Tree-Coarse	1117.733454	884.153401	1.249328 e+06	0.865389

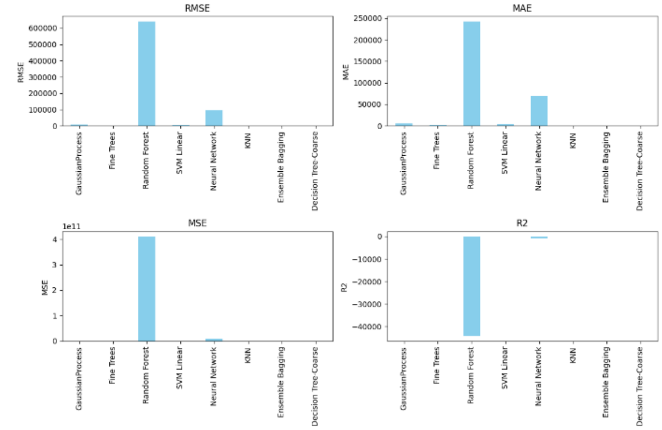

Fig. 1. Evaluating the performance of machine learning models.

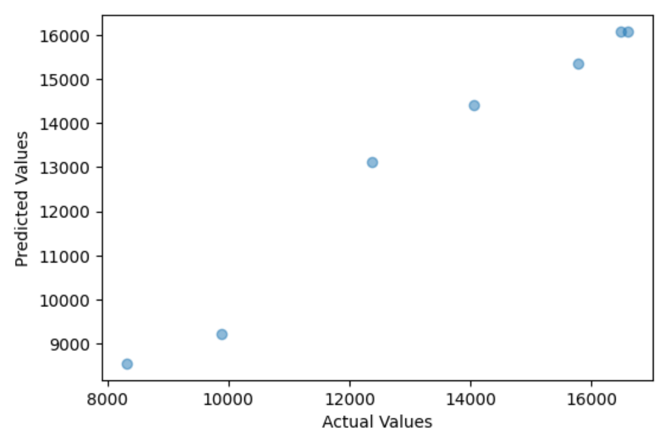

Fig. 2. Comparison of actual and predicted values.

Fig. 2 presents a scatter plot comparing actual values with values predicted by a K-Nearest Neighbors (KNN) model, showing a positive linear relationship where the points cluster around an upward trend, suggesting that as actual values increase, predicted values also increase. Therefore, the model used has a good predictive capability.

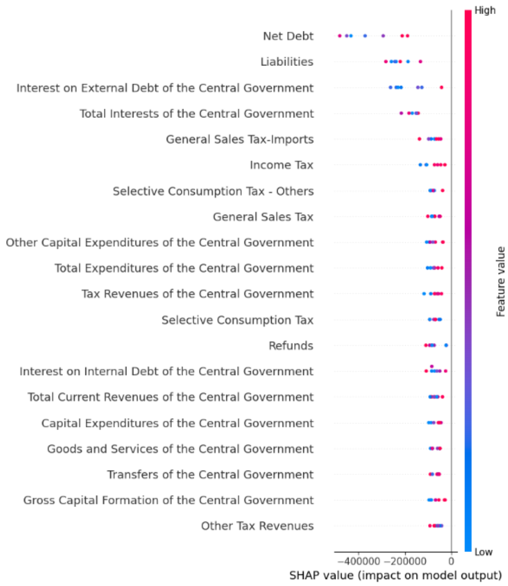

Fig. 3. SHAP Analysis of the importance of variables.

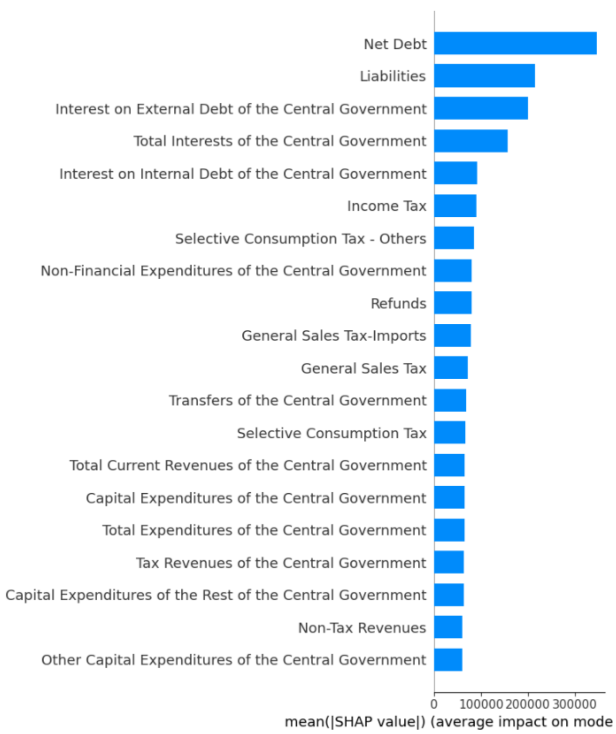

Fig. 4. Average importance of the characteristics according to SHAP values.

The importance of variables in predictive growth models using SHAP values. Fig. 3, shows the greatest influence of a variable on a specific prediction at each point, with colors indicating the characteristic's value (blue for low values and red for high values). Variables such as "Net Debt," "Liabilities," and "Interest on Central Government External Debt" are observed to have an influence on the model's predictions, as they have high SHAP values, indicating their strong contribution to the results. Fig. 4, presents the summary of the average importance of each variable in the model, highlighting again that "Net Debt" and "Liabilities" are the most influential factors. Therefore, these visualizations confirm that variables related to debt and tax revenue play an important role in predicting economic growth, suggesting that policies that affect these indicators can directly impact economic projections.

Fig. 5. Breakdown analysis of the results of fiscal policy variables.

Fig. 5 provides a detailed breakdown of the importance of fiscal policy variables in predicting economic growth outcomes. Similarly, SHAP explains and shows how different variables influence the model's prediction. The prediction base is approximately 2.52 million, but the factors represented in blue have decreased the final prediction to 9,206.99. Public investment has a slight positive influence, while net debt, liabilities, central government interest, income tax, and central government gross capital formation have negative effects on the prediction, suggesting that these factors are associated with a decline in economic growth. This reinforces economic theory, indicating that high levels of debt and tax burden can limit the dynamism of real GDP, highlighting the importance of a balanced fiscal policy in fostering sustainable economic development.

V. Discussion

The results obtained in this study demonstrate that the K-Nearest Neighbors (KNN) model and Ensemble Bagging are highly effective in predicting Peru's economic growth using fiscal policy variables. In particular, the KNN model achieved a coefficient of determination (R^2) of 0.972. These models showed high prediction accuracy, with low errors in the RMSE, MAE, and MSE metrics. Furthermore, the variables with the greatest influence on the prediction, according to SHAP, were Net Debt, Liabilities, and Interest on External Debt, suggesting that these variables play a relevant role in the country's economic dynamics.

Thus, these findings underscore the importance of fiscal policy management for economic growth. Previous studies have highlighted that government debt and fiscal policy can have positive or negative effects depending on their administration [1], [4]. The identification of net debt and liabilities as important variables confirms the importance of prudent public debt management to avoid adverse impacts on the economy. Furthermore, such management is essential for promoting sustainable economic growth in Peru.

It is worth noting that the results are consistent with previous research, such as the usefulness of Machine Learning models in macroeconomic prediction [49]. The superiority of the KNN model over other models, such as Random Forest and Neural Networks, agrees with studies that highlight the effectiveness of models based on economic time series [50]. Furthermore, the application of SHAP for model interpretation reinforces the need for explanatory tools in predictive economics, as they allow results to be more accessible, understandable, and ultimately, more useful for informed and responsible decision-making [51].

Despite the promising results, this study has some limitations. First, the data used come exclusively from the

Central Reserve Bank of Peru, which may limit the generalizability of the findings to other economies with different fiscal policy structures. Furthermore, social or environmental variables, such as income inequality or climate change, which could influence economic growth, were not included [52]. Finally, the study focused on annual data, which may not capture short-term economic dynamics.

Consequently, future research could focus on integrating high-frequency data and incorporating international financial indicators to improve the models' predictive capacity. Likewise, the use of hybrid approaches that combine econometric techniques with deep learning could provide better results in economic prediction [53].

VI. Conclusion

This study analyzes the application of machine learning models to predict Peru's economic growth using fiscal policy variables with a high degree of accuracy, especially with the KNN and Ensemble Bagging models. These models have proven to be highly effective and robust for economic analysis, as they are able to capture nonlinear relationships between fiscal variables and economic growth, outperforming traditional approaches such as Random Forest and neural networks.

Furthermore, the SHAP analysis identified the most influential variables, providing valuable information for fiscal policy decision-making. In particular, it highlights the importance of Net Debt, Liabilities, and Interest on External Debt as key factors affecting economic growth.

Finally, the K-Nearest Neighbors (KNN) model was selected as the best for its performance and robustness. As a result, the study validates the use of machine learning models for economic forecasting and suggests that prudent fiscal management, focused on debt and liability control, positively influences Peru's economic growth. Furthermore, this work contributes significantly to the field of macroeconomic forecasting.

Acknowledgment

We would like to thank all the individuals and institutions that contributed to this study on the evaluation of machine learning models for predicting fiscal policy and economic growth in Peru (1990–2023). We especially acknowledge the support of our colleagues, mentors, and experts, whose valuable perspectives were key to this research.

References

[1] Barro R.J, "Government Spending in a Simple Model of Endogenous Growth," Journal of Political Economy, vol. 98, no. 5, 1990.

[2] F. Caprioli, "Optimal fiscal policy under learning," J Econ Dyn Control, vol. 58, pp. 101–124, Sep. 2015, doi: 10.1016/J.JEDC.2015.05.008.

[3] A. J. Auerbach and Y. Gorodnichenko, "Fiscal Multipliers in Recession and Expansion," Fiscal Policy after the Financial Crisis, pp. 63–98, Jan. 2013, doi: 10.7208/CHICAGO/9780226018584.003.0003.

[4] O. Blanchard and R. Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," Q J Econ, vol. 117, no. 4, pp. 1329–1368, Nov. 2002, doi: 10.1162/003355302320935043.

[5] A. W. De Barros Silva et al., "Methodology Based on Artificial Neural Networks for Hourly Forecasting of PV Plants Generation," IEEE Latin America Transactions, vol. 20, no. 4, pp. 659–668, Apr. 2022, doi: 10.1109/TLA.2022.9675472.

[6] F. Villada, N. Muñoz, and E. García, "Application of Artificial Neural Networks to Price Forecasting in the Stock Exchange Market," Información Tecnológica, vol. 23, no. 4, pp. 11–20, 2012, doi: 10.4067/S0718-07642012000400003.

[7] D. Fantozzi and A. Muscarnera, "A News-Based Policy Index for Italy: Expectations and Fiscal Policy," Italian Economic Journal 2025, pp. 1–45, Mar. 2025, doi: 10.1007/S40797-025-00320-X.

[8] X. Dou, W. Chen, L. Zhu, Y. Bai, Y. Li, and X. Wu, "Machine Learning for Smart Cities: A Comprehensive Review of Applications and Opportunities," International Journal of Advanced Computer Science and Applications, vol. 14, no. 9, pp. 999–1016, 2023, doi: 10.14569/IJACSA.2023.01409104.

[9] F. Huanco-Ramos and A. Apaza-Tarqui, "Modelo de identificación de Covid-19 usando técnicas de Deep Learning a partir de imágenes Rayos X de tórax de pulmones," Aug. 11, 2023. doi: 10.18687/LACCEI2024.1.1.1491.

[10] T. Hastie, R. Tibshirani, and J. Friedman, "Springer Series in Statistics The Elements of Statistical Learning - Data Mining, Inference, and Prediction," vol. 2nd, p. undefined-undefined, 2009, Accessed: Oct. 20, 2024. [Online]. Available: https://www.mendeley.com/catalogue/c0d5cf64-c1d7-3a92-8862-df1797c83829/

[11] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, vol. 13-17-August-2016, pp. 785–794, Aug. 2016, doi: 10.1145/2939672.2939785/SUPPL_FILE/KDD2016_CHEN_BOOSTING_SYSTEM_01-ACM.MP4.

[12] H. R. Varian, "Big data: New tricks for econometrics," Journal of Economic Perspectives, vol. 28, no. 2, 2014, doi: 10.1257/jep.28.2.3.

[13] S. Athey and G. W. Imbens, "Machine Learning Methods That Economists Should Know about," Annu Rev Econom, vol. 11, no. Volume 11, 2019, pp. 685–725, Aug. 2019, doi: 10.1146/ANNUREV-ECONOMICS-080217-053433/CITE/REFWORKS.

[14] F. H.-R. Alejandro Apaza-Tarqui, "Models of automatic recognition of collaborative emotions in rural secondary education institutions in Puno, 2024," Journal of International Crisis and Risk Communication Research , pp. 3306–3325, Jan. 2025, Accessed: Feb. 18, 2025. [Online]. Available: https://jicrcr.com/index.php/jicrcr/article/view/2615

[15] J. Balfer and J. Bajorath, "Systematic Artifacts in Support Vector Regression-Based Compound Potency Prediction Revealed by Statistical and Activity Landscape Analysis," PLoS One, vol. 10, no. 3, p. e0119301, Mar. 2015, doi: 10.1371/JOURNAL.PONE.0119301.

[16] K. A, "Artificial intelligence application in the forecasting of economic model," Artif Intell, vol. 28, no. AI.2023.28(2)), pp. 88–93, Sep. 2023, doi: 10.15407/JAI2023.02.088.

[17] A. Aliyev, "Conceptual basis of Development and Application of Artificial Intelligence Technologies in Forecasting Economic Processes," Artificial societies, vol. 18, no. S1, p. 0, 2023, doi: 10.18254/S207751800028599-7.

[18] J. Király, B. Csontó, L. Jankovics, and K. Mérő, "Monetary, Macroprudential, and Fiscal Policy," Contributions to Economics, pp. 255–324, 2022, doi: 10.1007/978-3-030-93963-2_6.

[19] L. Ayala, A. Herrero, and J. Martinez-Vazquez, "Welfare benefits in highly decentralized fiscal systems: Evidence on interregional mimicking," Papers in Regional Science, vol. 100, no. 5, pp. 1178–1208, Oct. 2021, doi: 10.1111/PIRS.12605.

[20] I. Khanchaoui, S. El Aboudi, and A. El Moudden, "Empirical investigation on the impact of public expenditures on inclusive economic growth in Morocco: Application of the autoregressive distributed lag approach," International Journal of Advanced Computer Science and Applications, vol. 11, no. 4, pp. 171–177, 2020, doi: 10.14569/IJACSA.2020.0110423.

[21] V. Glassner and M. Keune, "La crisis y la política social: el papel de los convenios colectivos," Revista Internacional del Trabajo, vol. 131, no. 4, pp. 385–413, Dec. 2012, doi: 10.1111/J.1564-9148.2012.00154.X.

[22] J. A. Schumpeter and J. M. Keynes, "The General Theory of Employment, Interest and Money.," J Am Stat Assoc, vol. 31, no. 196, 1936, doi: 10.2307/2278703.

[23] J. M. Keynes, “The general theory of employment, interest, and money,” The General Theory of Employment, Interest, and Money, pp. 1–404, Jul. 2018, doi: 10.1007/978-3-319-70344-2.

[24] X. Li, “Analysis of economic forecasting in the post-epidemic era: evidence from China,” Scientific Reports 2023 13:1, vol. 13, no. 1, pp. 1–9, Feb. 2023, doi: 10.1038/s41598-022-19011-z.

[25] I. Rosales, J. Avitia, J. Ramirez, and E. Urbina, “Local productive systems within the perspective of the circular economy,” Universidad Ciencia y Tecnología, vol. 25, no. 111, pp. 57–66, Dec. 2021, doi: 10.47460/UCT.V25I111.516.

[26] N. P. Lien, “How Does Governance Modify the Relationship between Public Finance and Economic Growth: A Global Analysis,” VNU Journal of Science: Economics and Business, vol. 34, no. 5E, Dec. 2018, doi: 10.25073/2588-1108/VNUEAB.4165.

[27] A. Bellocchi, G. Travaglini, and B. Vitali, “How capital intensity affects technical progress: An empirical analysis for 17 advanced economies,” Metroeconomica, vol. 74, no. 3, pp. 606–631, Jul. 2023, doi: 10.1111/MECA.12421.

[28] C. Le-Van and B. Tran-Nam, “Comparing the Harrod-Domar, Solow and Ramsey growth models and their implications for economic policies,” Fulbright Review of Economics and Policy, vol. 3, no. 2, pp. 167–183, Dec. 2023, doi: 10.1108/FREP-06-2023-0022.

[29] W. Eltis, “Harrod–Domar Growth Model,” The New Palgrave Dictionary of Economics, pp. 1–5, 1987, doi: 10.1057/978-1-349-95121-5_1267-1.

[30] J. Keynes, “Teoría general de la ocupación el interés y el dinero,” p. undefined-undefined, 1936, Accessed: Dec. 29, 2024. [Online]. Available: https://www.mendeley.com/catalogue/7956acf4-be9f-3fbe-809e-290dffba9288/

[31] M. Basheer, V. Nechifor, A. Calzadilla, C. Ringler, D. Hulme, and J. J. Harou, “Balancing national economic policy outcomes for sustainable development,” Nature Communications 2022 13:1, vol. 13, no. 1, pp. 1–13, Aug. 2022, doi: 10.1038/s41467-022-32415-9.

[32] J. Yuan and S. Liu, “A double machine learning model for measuring the impact of the Made in China 2025 strategy on green economic growth,” Scientific Reports 2024 14:1, vol. 14, no. 1, pp. 1–16, May 2024, doi: 10.1038/s41598-024-62916-0.

[33] Y. Chen and X. Yao, “Predicting the academic achievement of students using black hole optimization and Gaussian process regression,” Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 1–15, Mar. 2025, doi: 10.1038/s41598-025-86261-y.

[34] S. Kularathne, A. Perera, N. Rathnayake, U. Rathnayake, and Y. Hoshino, “Analyzing the impact of socioeconomic indicators on gender inequality in Sri Lanka: A machine learning-based approach,” PLoS One, vol. 19, no. 12, p. e0312395, Dec. 2024, doi: 10.1371/JOURNAL.PONE.0312395.

[35] S. K. Jain and A. K. Gupta, “Application of Random Forest Regression with Hyper-parameters Tuning to Estimate Reference Evapotranspiration,” International Journal of Advanced Computer Science and Applications, vol. 13, no. 5, pp. 742–750, 2022, doi: 10.14569/IJACSA.2022.0130585.

[36] R. O. Tachibana, N. Oosugi, and K. Okanoya, “Semi-Automatic Classification of Birdsong Elements Using a Linear Support Vector Machine,” PLoS One, vol. 9, no. 3, p. e92584, Mar. 2014, doi: 10.1371/JOURNAL.PONE.0092584.

[37] R. K. Mishra, S. Urolagin, J. Angel, A. Jothi, N. Nawaz, and H. Ramkissoon, “Machine Learning based Forecasting Systems for Worldwide International Tourists Arrival,” IJACSA) International Journal of Advanced Computer Science and Applications, vol. 12, no. 11, p. 2021, Accessed: Apr. 18, 2025. [Online]. Available: www.ijacsa.thesai.org

[38] I. Malashin, D. Martysyuk, V. Tynchenko, V. Nelyub, A. Borodulin, and A. Galinovsky, “Mechanical Testing of Selective-Laser-Sintered Polyamide PA2200 Details: Analysis of Tensile Properties via Finite Element Method and Machine Learning Approaches,” Polymers (Basel), vol. 16, no. 6, Mar. 2024, doi: 10.3390/POLYM16060737.

[39] S. Sánchez-Herrero, L. Calvet, and A. A. Juan, “Machine Learning Models for Predicting Personalized Tacrolimus Stable Dosages in Pediatric Renal Transplant Patients,” BioMedInformatics, vol. 3, no. 4, pp. 926–947, Dec. 2023, doi: 10.3390/BIOMEDINFORMATICS3040057.

[40] G. Ravindiran et al., “Impact of air pollutants on climate change and prediction of air quality index using machine learning models,” Environ Res, vol. 239, Dec. 2023, doi: 10.1016/J.ENVRES.2023.117354.

[41] J. Ma, K. Du, F. Zheng, L. Zhang, and Z. Sun, “A segmentation method for processing greenhouse vegetable foliar disease symptom images,” Information Processing in Agriculture, vol. 6, no. 2, pp. 216–223, Jun. 2019, doi: 10.1016/J.INPA.2018.08.010/A_SEGMENTATION_METHOD_FOR_PROCESSING_GREENHOUSE_VEGETABLE_FOLIAR_DISEASE_SYMPTOM_IMAGES.PDF.

[42] N. Rathnayake, U. Rathnayake, I. Chathuranika, T. L. Dang, and Y. Hoshino, “Projected Water Levels and Identified Future Floods: A Comparative Analysis for Mahaweli River, Sri Lanka,” IEEE Access, vol. 11, pp. 8920–8937, 2023, doi: 10.1109/ACCESS.2023.3238717.

[43] D. K. Tiwari, K. R. Singh, and V. Kumar, “Forecasting of water quality parameters of Sandia station in Narmada basin, Central India, using AI techniques,” Journal of Water and Climate Change, vol. 15, no. 3, pp. 1172–1183, 2024, doi: 10.2166/WCC.2024.520/1368237/JWC2024520.PDF.

[44] L. I. Mampitiya, R. Nalmi, and N. Rathnayake, “Classification of Human Emotions using Ensemble Classifier by Analysing EEG Signals,” Proceedings - 2021 IEEE 3rd International Conference on Cognitive Machine Intelligence, CogMI 2021, pp. 71–77, 2021, doi: 10.1109/COGMI52975.2021.00018.

[45] M. Herath, T. Jayathilaka, H. M. Azamathulla, V. Mandala, N. Rathnayake, and U. Rathnayake, “Sensitivity Analysis of Parameters Affecting Wetland Water Levels: A Study of Flood Detention Basin, Colombo, Sri Lanka,” Sensors, vol. 23, no. 7, Apr. 2023, doi: 10.3390/S23073680.

[46] H. P. Piepho, “An adjusted coefficient of determination (R2) for generalized linear mixed models in one go,” Biometrical Journal, vol. 65, no. 7, Oct. 2023, doi: 10.1002/BIMJ.202200290.

[47] N. Rathnayake, U. Rathnayake, I. Chathuranika, T. L. Dang, and Y. Hoshino, “Cascaded-ANFIS to simulate nonlinear rainfall–runoff relationship,” Appl Soft Comput, vol. 147, Nov. 2023, doi: 10.1016/J.ASOC.2023.110722.

[48] Y. Zhang, “Multi-Factors Analysis Using Visualizations and SHAP: Comprehensive Case Analysis of Tennis Results Forecasting,” International Journal of Advanced Computer Science and Applications, vol. 16, no. 1, pp. 143–152, Feb. 2025, doi: 10.14569/IJACSA.2025.0160114.

[49] A. Batz, D. F. D’Croz-Barón, C. J. Vega Pérez, and C. A. Ojeda-Sanchez, “Integrating machine learning into business and management in the age of artificial intelligence,” Humanities and Social Sciences Communications 2025 12:1, vol. 12, no. 1, pp. 1–20, Mar. 2025, doi: 10.1057/s41599-025-04361-6.

[50] R. Giacomini and B. Rossi, “Forecasting in macroeconomics,” Handbook of Research Methods and Applications in Empirical Macroeconomics, Sep. 2013, doi: 10.4337/9780857931023.00024.

[51] J. K. MA Wynne, “One-size-fits-all monetary policy: Europe and the US,” Federal Res Bank Dallas Econom Lett, vol. 7, p. 9, 2012.

[52] Isubalew Daba Ayana, Wondaferahu Mulugeta Demissie, W. Mulugeta, and Atnafu Gebremeskel Sore, “Fiscal policy and economic growth in Sub-Saharan Africa: Do governance indicators matter?,” PLoS One, vol. 18, no. 11, p. e0293188, Nov. 2023, doi: 10.1371/JOURNAL.PONE.0293188.

[53] E. F. Mumbuli, E. N. Abelly, M. M. Mgimba, and A. E. Twum, “Towards Precision Economics: Unveiling GDP Patterns Using Integrated Deep Learning Techniques,” Comput Econ, pp. 1–29, Jan. 2025, doi: 10.1007/S10614-025-10863-X/TABLES/4.