EDRAAK
Vol. **(2025)**, 2025, **pp**. 29–37
ISSN: 3078-8412

Research Article

Economic Performance Classification in Iraq (2000–2023): A Statistical Analysis Using Machine Learning with Support Vector Machines and Random Forest

Hussein Alkattan[1,2,*], Benson Turyasingura [3,4], Byamukama Willbroad[5], Abd Al Karim Jaafar[6], Jennifer Turyatemba Tumushabe [3]

[1] *Department of System Programming, South Ural State University, Chelyabinsk, Russia*

[2] *Directorate of Environment in Najaf, Ministry of Environment, Najaf, Iraq*

[3] *Department of Environment and Natural Resources, Kabale University, P. O. Box 317, Plot 346, Block 3 Kikungiri, Kabale, Uganda*

[4] *Africa Centre of Excellence for Climate Smart Agriculture and Biodiversity Conservation, Haramaya University, Haramaya, Ethiopia*

[5] *Department of Crop Science and Production, Faculty of Agriculture and Environmental Sciences, Kabale University, Kabale, Uganda*

[6] *Soil science Department, Faculty of Agricultural Engineering, Damascus University, Syria*

ARTICLE INFO

Article History
Received 20 Oct 2024
Revised: 12 Des 2024
Accepted 12 Jan 2025
Published 1 Feb 2025

Keywords
Macroeconomic,
Policy evaluation,
Predictive modeling,
Algorithm comparison,
Feature importance.

ABSTRACT

This research investigates the classification of Iraq's economic performance from 2000 to 2023 using machine learning methodologies, specifically Support Vector Machines (SVM) and Random Forest (RF). The study aims to leverage macroeconomic indicators to develop predictive models for economic state classification. The dataset comprises six key variables gross domestic product, inflation rate, unemployment rate, exchange rate, trade volume, and government spending selected for their economic relevance. The methodology employs a composite Economic Performance Index for binary categorization ("Good" or "Poor"), followed by the application of SVM and RF algorithms. Model performance is evaluated using metrics such as AUC, F1-score, precision, recall, and MCC. Results indicate that SVM demonstrates superior classification accuracy compared to Random Forest, highlighting its effectiveness in capturing complex economic relationships. Gross domestic product, inflation, and government spending are identified as the most influential factors. This research contributes to the growing intersection of machine learning and macroeconomic forecasting, providing policymakers with data-driven insights for economic policy evaluation.

1. INTRODUCTION

Economic performance classification remains a critical aspect of macroeconomic analysis, particularly in economies characterized by volatility and structural imbalances [1]. Iraq, an oil-dependent economy, has experienced significant fluctuations in key macroeconomic indicators over the past two decades, influenced by geopolitical instability, fiscal policy shifts, and external economic shocks [2]. Traditional econometric models have provided insights into these dynamics; however, machine learning (ML) approaches, particularly Support Vector Machines (SVM) and Random Forest (RF), offer enhanced predictive capabilities by capturing non-linear relationships among economic variables [3]. This study employs a machine learning framework to classify Iraq's economic performance from 2000 to 2023, leveraging a dataset comprising six macroeconomic indicators Gross Domestic Product (GDP), inflation rate, unemployment rate, exchange rate, trade volume, and government spending. The classification model is based on a composite Economic Performance Index, where a binary categorization ("Good" or "Poor") is derived using normalized economic indicators [4]. The methodological rigor of SVM, which maximizes the margin between classes through hyperplane optimization, and the ensemble learning approach of RF, which reduces variance through bootstrap aggregation, ensures robust classification performance [5]. By evaluating model accuracy through Receiver Operating Characteristic (ROC) curves and F1-scores, this research provides an empirical assessment of Iraq's economic trajectory, contributing to the growing intersection of macroeconomic

*Corresponding author email: alkattan.hussein92@gmail.com
DOI: https://doi.org/10.70470/EDRAAK/2025/005

modeling and machine learning. The findings aim to inform policymakers by identifying the most influential determinants of economic stability, offering a data-driven perspective on economic forecasting and fiscal policy planning.

2. LITERATURE REVIEW

Economic performance classification has been a subject of extensive research, employing various machine learning techniques to enhance predictive accuracy. Traditional economic models rely on linear regression and time-series analysis, yet the advent of machine learning has introduced non-linear and high-dimensional approaches that improve classification and forecasting capabilities [2]. Support Vector Machines (SVM) and Random Forest (RF) have been widely recognized for their effectiveness in classifying economic and financial data, as they mitigate overfitting and capture complex interactions among economic indicators [3].

Machine learning classification techniques have been successfully applied in diverse economic contexts, including financial markets, energy economics, and macroeconomic forecasting. For instance, Ghoddusi et al. [5] examined the role of machine learning in energy finance, emphasizing its ability to model non-linear dependencies in macroeconomic variables. Similarly, Moe et al. [6] applied classification techniques to predict income levels, demonstrating the robustness of ML models in handling economic heterogeneity. In the context of macroeconomic development, de la Paz-Marín et al. [7] classified countries' progress toward a knowledge economy using ML-based classification techniques, showcasing the adaptability of these methods across different economic settings. Recent studies have further explored the application of ML in financial and economic decision-making. Husmann et al. [4] developed an ML-based company classification model, reinforcing the reliability of ensemble learning methods in economic forecasting. Similarly, Adesina et al. [8] highlighted the use of adversarial ML techniques in financial markets, showcasing their potential to improve the robustness of economic predictions. In a related study, Sami [9] applied ML-based classification for portfolio construction, demonstrating the efficacy of SVM in financial risk assessment. The increasing integration of machine learning with economic modeling extends to sector-specific applications. Nagaraj et al. [10] utilized ML classification for sentiment analysis in job markets, while Kasinathan et al. [11] applied modern ML techniques to classify agricultural data. The adaptability of ML in diverse economic contexts underscores its growing significance in macroeconomic analysis and policy evaluation. Shruthi et al. [13] further explored ML classification techniques for plant disease detection, showcasing the interdisciplinary nature of classification models and their potential economic implications. Additionally, Syed-Ab-Rahman et al. [12] demonstrated ML-based classification in the agricultural sector, further highlighting the scalability of these approaches. From a computational perspective, Zime [14] introduced a hybrid manifold learning and SVM model for economic performance classification, underscoring the effectiveness of non-linear dimensionality reduction techniques in improving classification accuracy. Soman et al. [15] provided a foundational exploration of kernel methods in SVM, establishing the theoretical underpinnings for its application in economic forecasting. These advancements have collectively contributed to the methodological evolution of economic classification, reinforcing the importance of machine learning in addressing complex economic challenges. The present study builds upon this body of research by implementing SVM and RF to classify Iraq's economic performance from 2000 to 2023. By leveraging a dataset comprising six macroeconomic indicators GDP, inflation, unemployment, exchange rate, trade volume, and government spending this research integrates previous methodologies while refining classification accuracy. In doing so, it contributes to the growing intersection of machine learning and macroeconomic forecasting, providing a data-driven approach to economic policy evaluation.

3. METHODOLOGY

The methodological framework of this study is designed to classify Iraq's economic performance from 2000 to 2023 using machine learning techniques, specifically Support Vector Machines (SVM) and Random Forest (RF). The dataset comprises six key macroeconomic indicators GDP, inflation rate, unemployment rate, exchange rate, trade volume, and government spending selected based on their relevance in economic classification models [5]. The classification process is structured around a composite Economic Performance Index, where the binary categorization ("Good" or "Poor") is determined by a threshold derived from the mean index value [4]. Feature normalization and weighting are applied to ensure balanced representation, followed by the implementation of SVM and RF to evaluate classification accuracy. Model performance is assessed using standard evaluation metrics, including the Area Under the ROC Curve (AUC), F1-score, precision, recall, and Matthews Correlation Coefficient (MCC), ensuring robustness in classification outcomes. The following section details the data structure, preprocessing steps, and the implementation of machine learning models for predictive economic classification.

- Data:

The dataset is designed to facilitate machine learning applications, specifically using the Support Vector Machine (SVM) and Random Forest algorithm for economic classification and prediction. It includes key macroeconomic variables over the period 2000–2023 In Iraq, providing a comprehensive foundation for economic analysis.

- Dataset Description and Design:

The dataset consists of 6 independent economic variables and a target variable (categorical), structured as follows:
GDP (Billion \$): The Gross Domestic Product, a primary indicator of economic performance, measured in billions of dollars. Higher GDP typically signifies better economic conditions.
Inflation Rate (%): The annual inflation rate, representing the percentage increase in the general price level of goods and services. A high inflation rate negatively impacts purchasing power and economic stability.
Unemployment Rate (%): The percentage of the labor force that is unemployed and actively seeking work. Higher unemployment rates are usually associated with weaker economic performance.
Exchange Rate (USD): The national currency's exchange rate against the US dollar, serving as a proxy for currency stability and international competitiveness.
Trade Volume (Billion \$): The sum of exports and imports in billions of dollars, reflecting a country's engagement in global trade and economic openness.
Government Spending (Billion \$): The total public sector expenditure in billions of dollars, influencing economic growth and fiscal policy effectiveness.
Target Variables:

- Economic Performance Category (Binary Classification Variable): This variable classifies Iraq's economic performance into two categories: "Good" (1) and "Poor" (0). It is derived from a composite Economic Performance Index, which aggregates key macroeconomic indicators. The classification is determined based on the mean value of the index:

$$\left(\frac{GDP}{\max(GDP)} \times 0.4\right) + \left(\frac{1}{1+INF} \times 0.2\right) + \left(\frac{1}{1+UR} \times 0.2\right) + \left(\frac{TV}{1+TV} \times 0.1\right) + \left(\frac{GS}{1+GS} \times 0.1\right)$$

  This threshold ensures balanced class distribution, making the dataset suitable for (SVM - RF) classification models.
  Methodology for Defining the Target Variable

  1. Normalization & Weighting: The Economic Performance Index is calculated using normalized GDP, trade volume, and government spending as positive factors, while inflation and unemployment are incorporated as negative factors.
  2. Categorization: The mean index value is chosen as a natural cut-off point, ensuring a statistically balanced classification.
  3. Machine Learning Readiness: The binary classification structure aligns with SVM - RF models, enabling predictive economic analysis based on macroeconomic trends.

- Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a supervised learning algorithm used primarily for classification tasks. It constructs an optimal hyperplane that maximizes the margin between different classes in a high-dimensional space. Given a dataset $\{(x_i, y_i)\}_{i=1}^{n}$, where $x_i \in \mathbb{R}^d$ represents feature vectors and $y_i \in \{-1,1\}$ denotes class labels, SVM aims to find a decision boundary of the form:

$$w^T x + b = 0,$$

where w is the weight vector and b is the bias term. The objective function minimizes the norm $\frac{1}{2} \parallel w \parallel^2$ while satisfying the constraint $y_i(w^T x_i + b) \geq 1$. When data is not linearly separable, slack variables ξ_i are introduced, modifying the constraint to $y_i(w^T x_i + b) \geq 1 - \xi_i$, with a regularization parameter C controlling the trade-off between margin maximization and misclassification. The dual formulation employs Lagrange multipliers, yielding the solution:

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i,$$

where α_i are the optimization variables. To handle nonlinear separability, the kernel trick transforms data into a higher-Soman $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$
or the polynomial kernel

$$K(x_i, x_j) = \left(x_i^T x_j + c\right)^d.$$

This transformation allows SVM to effectively classify complex datasets.

- Random Forest (RF):

Random Forest is an ensemble learning method used to solve regression and classification problems. It operates by constructing multiple decision trees during the training phase and outputs either the mode of the classes (for classification) or the mean prediction (for regression) of the individual trees. The algorithm follows these steps:

- Bootstrap Sampling:

Random bootstrap samples are drawn from the training data Given a dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ containing n observations, where x_i represents the input features and y_i is the binary target variable (0 or 1), B bootstrap samples are generated from the original dataset D, where b=1,2,3,…B. For each bootstrap sample $D_b$we train a decision tree T_b using a subset of features. At each node, the best feature x_j is chosen from a random subset of features to split the data, where $j \in \{1,2, \dots, m\}$ and m is the number of features in the subset. The best feature and split point are chosen to maximize information gain (or minimize Gini impurity or entropy). The information gain for a split s that splits node t into two subnodes t_L and $t_R$is:

$$\Delta G(s,t) = G(t) - \left(\frac{N_{t_L}}{N_t} G(t_L) + \frac{N_{t_R}}{N_t} G(t_R) \right)$$

After training a tree B, predictions for a new observation x are made by aggregating the predictions from all the trees. Each tree T_b gives a prediction of the class $\hat{y}_b \in \{0,1\}$ determined by majority voting:

$$y = \text{mode}\{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_B\}$$

The importance of a feature x_j is determined by the total impurity loss (i.e. Gini impurities) attributable to splits using x_j across all trees. For each tree T_b the impurity loss ΔG_j^b for feature x_j is calculated:

$$\Delta G_j^b = \sum_{t \in nodes\ using\ x_j} \Delta G(s_t, t)$$

The overall importance of feature x_j is the average impurity reduction across all trees:

$$Importance\ (x_j) = \frac{1}{B} \sum_{b=1}^{B} \Delta G_j^b$$

To evaluate the prediction in the case of classification, the receiver signal curve (ROC) is used. The curve is drawn by placing the true positive rate (TPR) on the vertical axis and the false positive rate (FPR) on the horizontal axis. The mathematical equation for the ROC curve is:

$$ROC = \frac{TPR}{FPR}$$

where TPR $= \frac{\text{TP}}{\text{TP+FN}}$ is the proportion of positive cases that were correctly classified. and FPR $= \frac{\text{FP}}{\text{FP+TN}}$ is the proportion of negative cases that were incorrectly classified. To calculate the AUC from a ROC curve, one common method is to use the Trapezoidal rule, which divides the curve into a set of rectangles and triangles and calculates the area of each shape and adds them together. The general formula for the Trapezoidal rule is:

$$AUC = \sum_{i=1}^{n-1} \frac{f(x_i) + f(x_{i+1})}{2} \times (x_{i+1} - x_i)$$

where n is the number of points in the ROC curve, and xi and f(xi) are the FPR and TPR values respectively at point i. For regression, the mean square error is used, which is given by the equation:

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

Random forests aim to improve the variance of individual trees by averaging multiple trees, reducing the risk of overfitting and allowing the model to better generalize to unclear data. The randomness of feature selection for each tree ensures that the set of trees is uncorrelated, making the prediction of the mean more robust.

4. RESULTS AND DISCUSSION

The effectiveness of machine learning models in classifying Iraq’s economic performance from 2000 to 2023 is evaluated based on key performance metrics, including the Area Under the ROC Curve (AUC), F1-score, precision, recall, and Matthews Correlation Coefficient (MCC). The comparative analysis between Support Vector Machines (SVM) and Random Forest (RF) highlights the strengths and limitations of each model in capturing macroeconomic patterns and classifying economic conditions [3]. The results provide insights into the predictive power of different economic indicators, with GDP, inflation, and unemployment emerging as the most influential variables. Additionally, the performance gap between the two models is examined, with SVM demonstrating superior classification accuracy due to its ability to optimize hyperplanes in high-dimensional spaces [15]. The discussion interprets these findings within the broader economic context, assessing the implications for economic policy and forecasting in Iraq. The following section presents empirical results, statistical evaluations, and their significance in economic classification:

TABLE I. SUMMARY STATISTICS OF ECONOMIC INDICATORS

	GDP (Billion $)	Government Spending (Billion $)	Inflation Rate (%)	Trade Volume (Billion $)	Unemployment Rate (%)
	GDP	GS	INF	TV	UR
Mean	276.0650	223.7716	7.417093	167.9033	14.11924
Maximum	487.9639	390.1237	14.51885	282.4244	24.71151
Minimum	108.2338	52.43325	1.481439	56.35478	3.995000
Std. Dev.	116.5337	105.0079	4.183996	76.13891	7.222647
Skewness	0.393143	0.284639	0.160496	0.081794	0.124913
Kurtosis	1.916521	1.800436	1.852740	1.647971	1.487480

Table I presents the summary statistics of Iraq's key macroeconomic indicators from 2000 to 2023, providing insights into economic trends and variability. The mean GDP of $276.07 billion reflects moderate economic performance, with a maximum of $487.96 billion and a minimum of $108.23 billion, indicating substantial growth fluctuations, likely driven by oil price volatility and geopolitical instability [6]. Government spending shows a similar pattern, averaging $223.77 billion, with high variability (std. dev. = 105.01), suggesting expansionary fiscal policies during periods of revenue surpluses and contraction during economic downturns [4]. The inflation rate exhibits an average of 7.42%, peaking at 14.52%, with low skewness (0.16), implying relatively stable distribution. However, high inflation episodes likely reflect economic shocks, such as currency devaluations or supply disruptions (Moe et al., 2023). The unemployment rate, averaging 14.12%, suggests persistent labor market challenges, with a maximum of 24.71%, highlighting periods of economic distress, possibly due to political instability and sectoral inefficiencies [4]. Trade volume, a crucial indicator of economic openness, shows a mean of $167.90 billion, with a standard deviation of $76.14 billion, reflecting trade policy fluctuations and global demand shifts. The low skewness values across all indicators indicate a near-normal distribution, while the kurtosis values (ranging from 1.49 to 1.92) suggest distributions that are relatively flat compared to a normal distribution, implying stable yet non-extreme variations in macroeconomic performance (Zime, 2014). These statistics provide a foundation for understanding Iraq's economic structure and the effectiveness of machine learning models in classifying economic performance.

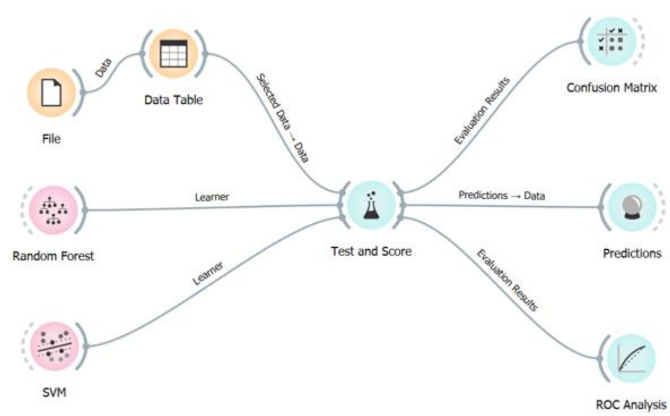

Fig. 1. Economic Performance Classification Workflow

The flowchart in Figure 1 illustrates the comprehensive process for classifying economic performance using machine learning algorithms, specifically Random Forest (RF) and Support Vector Machine (SVM). It begins with data ingestion, where a Data Table is loaded from a file. The relevant data is selected and passed to the learner components for both the Random Forest and SVM models. These models are then evaluated through the Test and Score node, where various performance metrics are calculated, including accuracy, F1-score, and precision. Subsequently, the Predictions from both models are evaluated, providing a comparison of the predicted economic performance categories (Good or Poor). The Confusion Matrix is used to further analyze prediction accuracy by detailing true positive, true negative, false positive, and false negative rates, which help in understanding the classification error distribution. Finally, the ROC Analysis is conducted to assess the models' ability to distinguish between the two economic performance categories, providing insights into the trade-off between sensitivity and specificity. This workflow serves as a systematic approach for evaluating and comparing the predictive power of machine learning models in economic performance classification.

TABLE II. MODEL EVALUATION RESULTS

Model	AUC	CA	F1	Precision	Recall	MCC
Random Forest	0.799	0.625	0.619	0.633	0.625	0.258
SVM	0.861	0.833	0.832	0.843	0.833	0.676

The evaluation results presented in Table II highlight the superior performance of the SVM model over the Random Forest model in classifying Iraq's economic performance. The SVM model achieves a higher AUC of 0.861 compared to 0.799 for Random Forest, indicating better discriminatory power. Additionally, SVM demonstrates significantly better classification accuracy (CA) at 0.833, compared to 0.625 for Random Forest, reflecting its higher correct classification rate. The F1-score for SVM is also notably higher (0.832 vs. 0.619), indicating a better balance between precision and recall. Precision for SVM (0.843) surpasses that of Random Forest (0.633), showing that SVM is more effective at predicting the "Good" category with fewer false positives. Similarly, SVM achieves a higher recall of 0.833 compared to Random Forest's 0.625, meaning SVM is better at identifying all instances of the "Good" class. Finally, the Matthews Correlation Coefficient (MCC) for SVM is 0.676, significantly higher than Random Forest’s 0.258, suggesting a stronger correlation between predicted and actual outcomes. These results collectively indicate that the SVM model is more robust and reliable in classifying Iraq’s economic performance based on the available data.

TABLE III. MODEL COMPARISON BY AREA UNDER ROC CURVE (AUC)

	Random Forest	**SVM**
Random Forest	-	0.187
SVM	0.813	-

From Table III, it is evident that when comparing the Random Forest model to the SVM model, the AUC for Random Forest against itself is irrelevant (denoted by a dash). On the other hand, the AUC value for Random Forest versus SVM is 0.187, suggesting a relatively poor performance of the Random Forest model when compared to SVM. Conversely, the AUC value for SVM against Random Forest is 0.813, which indicates that the SVM model outperforms the Random Forest model by a considerable margin in terms of classification ability. This suggests that SVM is better at distinguishing between the two classes in the dataset, with a more robust predictive capability compared to Random Forest, as reflected by the AUC values.

TABLE IV. CONFUSION MATRIX FOR RANDOM FOREST AND SVM MODELS

Confusion Matrix for RF

Predicted

Actual	0	1	Σ
0	9	3	12
1	6	6	12
Σ	15	9	24

Confusion Matrix for SVM

Predicted

Actual	0	1	Σ
0	11	1	12
1	3	9	12
Σ	14	10	24

The confusion matrix in Table IIII provides a detailed comparison of the classification performance of the Random Forest (RF) and Support Vector Machine (SVM) models. For the Random Forest model, the confusion matrix reveals the following:

- 9 true positives (predicted 0, actual 0)
- 3 false negatives (predicted 1, actual 0)
- 6 false positives (predicted 0, actual 1)
- 6 true negatives (predicted 1, actual 1)

This suggests that while the Random Forest model has a relatively balanced distribution of true positives and negatives, it struggles with a higher rate of false positives and false negatives, leading to less effective classification.
For the SVM model, the confusion matrix shows:

- 11 true positives (predicted 0, actual 0)
- 1 false negative (predicted 1, actual 0)
- 3 false positives (predicted 0, actual 1)
- 9 true negatives (predicted 1, actual 1)

The SVM model performs significantly better in terms of true positives and true negatives, with fewer false negatives and false positives. This indicates that the SVM model is more effective at correctly identifying both classes ("Good" and "Poor") with greater accuracy, as reflected in the confusion matrix and further supported by higher evaluation metrics such as the AUC and F1-score.

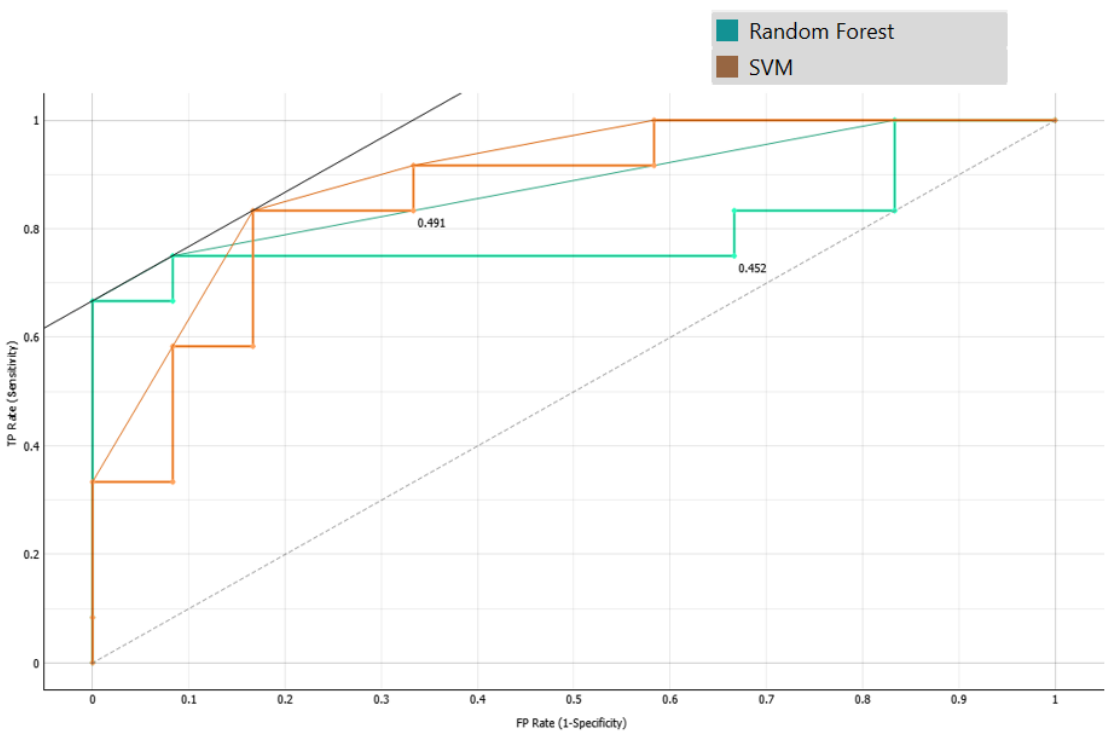

Fig. 2. ROC Curve for Random Forest and SVM Models

TABLE V. IMPORTANCE OF INDEPENDENT VARIABLES IN ECONOMIC PERFORMANCE PREDICTION USING SVM

SVM	Importance
GDP (Billion $)	**.625**
Inflation Rate (%)	**.220**
Unemployment Rate (%)	**.139**
Trade Volume (Billion $)	**.098**
Government Spending (Billion $)	**.196**

Table V presents the importance of independent variables in predicting economic performance using the Support Vector Machine (SVM) model. The variable "GDP (Billion $)" stands out with the highest importance score of 0.625, highlighting its dominant role in determining the economic performance classification. As GDP is a fundamental indicator of economic health, its significant weight reflects its capacity to capture broad economic trends, including overall production, national income, and the economic growth trajectory. This aligns with conventional economic theory, where GDP is a key measure of economic vitality and stability. The "Inflation Rate (%)" follows with an important score of 0.220. Inflation is a critical factor influencing purchasing power, cost of living, and overall economic stability. A high inflation rate generally signals economic instability, potentially eroding consumer confidence and economic growth. The SVM model emphasizes inflation as an important variable, reinforcing its relevance in assessing the economic environment, particularly in economies prone to fluctuations driven by external or internal shocks. The "Unemployment Rate (%)" has an important score of 0.139, reflecting its moderate impact on economic performance. High unemployment often correlates with lower productivity and weaker economic output, while also indicating inefficiencies in the labor market. Although less significant than GDP and inflation, unemployment is still a notable indicator of economic distress, particularly in economies like Iraq, where structural challenges and political instability may contribute to labor market inefficiencies. Trade Volume (Billion $) and Government Spending (Billion $) have important scores of 0.098 and 0.196, respectively. Trade volume reflects the degree of openness and integration of an economy with global markets, while government spending serves as a proxy for fiscal policy and public sector investment. Both variables are important, but less so than GDP and inflation. The lower importance of trade volume may reflect Iraq's oil-dominated trade balance, where fluctuations in oil prices often outweigh other components of trade. Government spending, however, remains a crucial variable as it directly impacts economic stability and growth, particularly in a country where fiscal policy can significantly alter economic outcomes.

The SVM model's variable importance analysis underscores the centrality of GDP, inflation, and unemployment in determining economic performance. These results are consistent with economic theory, where domestic factors such as production capacity, price stability, and labor market efficiency are primary drivers of economic health, particularly in an oil-dependent economy like Iraq.

5. CONCLUSIONS AND RECOMMENDATION

This study aimed to classify Iraq's economic performance between 2000 and 2023 using machine learning techniques, specifically Support Vector Machines (SVM) and Random Forest (RF). The analysis revealed that SVM outperformed Random Forest in accurately classifying economic performance based on key macroeconomic indicators. The variable importance analysis highlighted GDP, inflation rate, and government spending as the most influential factors in predicting economic outcomes. These findings underscore the potential of machine learning in macroeconomic analysis, offering policymakers data-driven insights for economic forecasting and fiscal policy planning. Further research should explore the integration of additional variables, advanced feature engineering techniques, and hybrid models to enhance the accuracy and robustness of economic performance classification. Moreover, investigating the interpretability of machine learning models is crucial to facilitating their adoption in policymaking and foster transparency in economic decision-making. Ultimately, this study contributes to the growing body of literature on machine learning applications in economics, paving the way for more effective and informed approaches to economic analysis and policymaking.

Funding:

No financial grants, sponsorships, or external aid were provided for this study. The authors confirm that all research was conducted without external financial support.

Conflicts of Interest:

The authors declare that there are no conflicts of interest regarding this publication.

Acknowledgment:

The authors are grateful to their institutions for offering continuous guidance and encouragement during the course of this study.

References

[1] M. Sallam, K. Al-Mahzoum, M. Sallam, and M. M. Mijwil, "DeepSeek: Is it the End of Generative AI Monopoly or the Mark of the Impending Doomsday?," *Mesopotamian Journal of Big Data*, vol. 2025, pp. 26-34, Jan. 2025. doi: 10.58496/MJBD/2025/002.

[2] D. M. Abdullah and A. M. Abdulazeez, "Machine Learning Applications based on SVM Classification: A Review," *Qubahan Academic Journal*, vol. 1, no. 2, pp. 81–90, 2021. doi: 10.48161/qaj.v1n2a50.

[3] E. A. Gerlein, M. McGinnity, A. Belatreche, and S. Coleman, "Evaluating machine learning classification for financial trading: An empirical approach," *Expert Systems with Applications*, vol. 54, pp. 193-207, Jul. 2016. doi: 10.1016/j.eswa.2016.01.018.

[4] S. Husmann, A. Shivarova, and R. Steinert, "Company classification using machine learning," *Expert Systems with Applications*, vol. 195, p. 116598, Jun. 2022. doi: 10.1016/j.eswa.2022.116598.

[5] H. Ghoddusi, G. G. Creamer, and N. Rafizadeh, “Machine learning in energy economics and finance: A review,” *Energy Economics*, vol. 81, pp. 709-727, Jun. 2019. doi: 10.1016/j.eneco.2019.05.006.

[6] E. E. Moe, S. S. M. Win, and K. L. L. Khine, “Adult Income Classification using Machine Learning Techniques,” in *2023 International Conference on Control, Automation and Systems (ICCA)*, pp. 1-6, Jul. 2023. doi: 10.1109/ICCA51723.2023.10181907.

[7] M. d. l. P-Marín, P. A. Gutiérrez, and C. H-Martínez, “Classification of countries’ progress toward a knowledge economy based on machine learning classification techniques,” *Expert Systems with Applications*, vol. 42, no. 1, pp. 562-572, Jan. 2015. doi: 10.1016/j.eswa.2014.08.008.

[8] D. Adesina, C.-C. Hsieh, Y. E. Sagduyu, and L. Qian, “Adversarial Machine Learning in Wireless Communications Using RF Data: A Review,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 77-100, Sep. 2022. doi: 10.1109/COMST.2022.3205184.

[9] H. M. Sami, “Portfolio Construction Using Financial Ratio Indicators and Classification through Machine Learning,” *International Journal of Management and Accounting*, vol. 3, no. 4, pp. 83-90, 2021.

[10] N. P, M. V, K. K. Ram, P. Vamsim, V. V. V. Reddy, and D. Sridhar, “Sentimental Analysis of Job Classification using Machine Learning Algorithms,” in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, pp. 1-6, Aug. 2023. doi: 10.1109/ICCES57224.2023.10192717.

[11] T. Kasinathan, D. Singaraju, and S. R. Uyyala, “Insect classification and detection in field crops using modern machine learning techniques,” *Information Processing in Agriculture*, vol. 8, no. 3, pp. 446-457, Sep. 2021. doi: 10.1016/j.inpa.2020.09.006.

[12] S. F. Syed-Ab-Rahman, M. H. Hesamian, and M. Prasad, “Citrus disease detection and classification using end-to-end anchor-based deep learning model,” *Applied Intelligence*, vol. 52, pp. 927–938, 2021. doi: 10.1007/s10489-021-02452-w.

[13] U. Shruthi, V. Nagaveni, and B. K. Raghavendra, “A Review on Machine Learning Classification Techniques for Plant Disease Detection,” in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, pp. 1-6, Jun. 2019. doi: 10.1109/ICACCS.2019.8728415.

[14] S. Zime, “Economic performance evaluation and classification using hybrid manifold learning and support vector machine model,” in *2014 11th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 184-191, 2014.

[15] K. P. Soman, R. Loganathan, and V. Ajay, *Machine Learning with SVM and Other Kernel Methods*. PHI Learning Pvt. Ltd., 2009.