JOURNAL OF CULTURAL ANALYSIS AND SOCIAL CHANGE

Journal of Cultural Analysis and Social Change,
2025, 10(3), 1003-1011
ISSN: 2589-1316

Machine Learning-Based Modeling of Economic Growth and Governance Quality: The MENA region

Nadia Farjallah[1*]

[1]*LaREMFiC Laboratory, IHEC, University of Sousse Tunisia; Email: nadia.farjallah@ihecso.u-sousse.tn*

***Corresponding Author:** nadia.farjallah@ihecso.u-sousse.tn

Citation: Farjallah, N. (2025). Machine Learning-Based Modeling of Economic Growth and Governance Quality: The MENA region. *Journal of Cultural Analysis and Social Change*, *10*(3), 1003–1011. https://doi.org/10.64753/jcasc.v10i3.2536

Published: November 28, 2025

ABSTRACT

This study investigates the predictive performance of various machine-learning models in forecasting economic growth across the MENA region. Four approaches were compared: Ordinary Least Squares (OLS), Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Regression (SVR). The dataset was divided into training (70%) and testing (30%) subsets to assess the robustness and generalization capacity of the models. Model accuracy was evaluated using the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). The results indicate that the Random Forest model provides the highest predictive accuracy (MAPE = 0.0192), outperforming traditional econometric approaches. Variable importance analysis highlights that corruption, government effectiveness; political stability, regulatory quality, and rule of law significantly influence economic growth. These findings confirm the relevance of non-parametric methods in capturing complex and nonlinear relationships between governance indicators and economic performance. Moreover, the results emphasize the crucial role of institutional quality as a structural determinant of growth, consistent with institutional and endogenous growth theories. The study concludes that machine learning models, particularly ensemble methods, offer robust and complementary tools for economic forecasting and policy analysis in emerging economies.

Keywords: Random Forest, Machine learning, economic growth, Government effectiveness, Corruption, Rule of law.

INTRODUCTION

The concept of good governance has been the subject of extensive debate and elaboration in contemporary scientific literature (Agere, 2000; Graham et al., 2003a; Armstrong et al., 2005; Andrews, 2008; Bovaird & Löffler, 2009). Regarded as one of the most effective management models in the public sector (Armstrong et al., 2005), good governance is based on a set of fundamental principles aimed at ensuring efficient, fair, and transparent administration of public affairs. Understanding the determinants of economic growth remains a central focus of contemporary economic theory. The neoclassical growth model, initially developed by Solow (1956) and Swan (1956), asserts that economic growth is driven by capital accumulation, labor input, and exogenous technological progress, thereby implying conditional convergence across countries. However, the persistent disparities in growth performance among nations have underscored the limitations of this framework. In response, endogenous growth models—pioneered by Romer (1986) and Lucas (1988) have highlighted the crucial roles of innovation, human capital, and public policies in fostering technological advancement. These contributions emphasize that economic growth is an endogenous and self-sustaining process shaped by policy choices and institutional dynamics. This theoretical shift has led to growing interest in institutions and their role in shaping economic performance. According to North (1990), institutions—both formal and informal—structure economic incentives, reduce transaction costs, and shape development trajectories. This perspective, further consolidated by Acemoglu, Johnson, and Robinson (2001, 2005), distinguishes between *inclusive institutions*, which foster economic participation and protect property rights, and *extractive institutions*, which concentrate power and hinder growth. Within this framework, the quality of governance emerges as a key determinant of sustainable development. The seminal works of Kaufmann, Kraay, and Zoido-Lobatón (1999, 2002) conceptualized governance through six fundamental dimensions: political stability, government effectiveness, regulatory quality, rule of law, control of corruption, and voice and accountability. These dimensions directly influence trust, legal security, and the predictability of public policies—key conditions for investment and productivity (Barro, 1996; Alesina & Perotti, 1996). In the Middle East and North Africa (MENA) region, where resource dependence and fragile institutions often characterize economies, governance represents a strategic lever for promoting economic diversification and enhancing resilience to external shocks (Ben Ali & Sassi, 2016; Baliamoune-Lutz, 2019).

However, traditional econometric approaches, which rely on linearity and homogeneity of effects, often fail to capture the complex interactions between governance and economic growth. The relationships between these variables may be nonlinear, asymmetric, and context-dependent, particularly across different institutional environments. In this regard, machine learning methods offer new analytical perspectives. Models such as Random Forests (Breiman, 2001), Support Vector Regression (SVR), and Gradient Boosting Machines (GBM) enable the exploration of multidimensional and nonlinear relationships while identifying the most influential variables driving economic growth. Accordingly, this study adopts both an empirical and methodological approach aimed at assessing the role of governance dimensions in predicting the Gross Domestic Product (GDP) of MENA countries using machine-learning techniques. By comparing the predictive performance of parametric and non-parametric models, the study seeks to identify the most relevant institutional determinants and to deepen the understanding of the relationship between governance and economic development. Ultimately, this research contributes to the literature on institutional growth by combining the theoretical framework of institutional economics with the empirical tools of artificial intelligence. It underscores the need for hybrid approaches capable of capturing the structural complexity of emerging economies, while opening new perspectives for the design of more effective public policies grounded in robust empirical evidence. Alongside traditional econometric and stochastic approaches, methods derived from the algorithmic culture offer new perspectives for economic analysis and growth prediction (Huang et al., 2014). However, to our knowledge, no empirical study has yet systematically explored the predictive factors of economic growth in the MENA region using modern data mining and machine learning techniques. This gap forms the starting point of the present work. Despite the economic and institutional reforms undertaken in several MENA countries, economic growth remains uneven and sometimes unstable. To what extent does the quality of governance influence economic growth in the MENA region, and what are its main measurable determinants as identified through modern machine learning techniques?

Effective governance, characterized by transparency, accountability, and institutional stability, exerts a significant positive effect on economic growth in the MENA countries. The use of machine learning algorithms allows for better identification and prioritization of the predictive factors underlying this complex relationship. The main objective of this study is to identify the key determinants of economic growth in the MENA countries and to assess their relative importance as predictors through the application of machine learning algorithms. The remainder of the paper is structured as follows: Section 2 presents the adopted methodology and the models used; Section 3 describes the data and preprocessing techniques; Section 4 presents and interprets the modeling results, highlighting predictive accuracy and variable importance; finally, Section 5 concludes by summarizing the main findings and outlining the economic and policy implications of the study.

METHODOLOGY

General Model: Economic growth Y_i is modeled as a function of governance variables $X_i = (x_{i1}, \dots, x_{ip})$:

$$Y_i = f(X_i) + \varepsilon_i \quad \text{With i=1,…, n} \qquad (1)$$

Or $\varepsilon_i \to N(0, \sigma_\varepsilon^2)$

Linear Model (OLS) :

$$\mathrm{Y_i} = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \qquad (2)$$

$$\varepsilon_i \to N(0, \sigma^2)$$

Estimation by Least Squares:

$$\hat{\beta} = \arg \underbrace{\min}_{\beta} \sum_{i=1}^{n} (Y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 \qquad (3)$$

Boosted Tree (GBM): The model is built iteratively

$$\hat{Y}_i^0 = argmin_\gamma \sum_{i=1}^{n} L(Y_i, \gamma) \quad (4)$$

For m=1,…, M

1. Compute the residuals: $r_i^m = -\left[\frac{dL(Y_i, \hat{Y}_i^{m-1})}{d\hat{Y}_i^{m-1}}\right]$
2. Build a tree T_m to approximate r_i^m
3. Update:
$\hat{Y}_i^m = \hat{Y}_i^{m-1} + \vartheta T_m(X_i)$ (5)

- M : Number of trees
- ϑ : Learning rate (0.01-0.1)
- L : Loss function

Radom Forest (R F): Construction of an ensemble of B decision trees $\{T_1, \dots, T_B\}$, each decision tree is built on a bootstrap sample. The final prediction is:
$\hat{Y}_i^{RF} = \frac{1}{B}\sum_{b=1}^{B} T_b(X_i)$ (6)
The primary steps are:
1. Sample n observations with replacement (bootstrap)
2. For each node, m variables are randomly chosen from the p available variables.
3. Split on the variable that maximally reduces impurity (variance).
4. Repeat this process for all trees and calculate the mean prediction.

And:

- B: number of trees (typically 500–1000).
- M: number of variables to consider at each split.

Support Vector Regression (SVR): is to identify a function that best approximates the relationship between the input variables and the target variable.
F(X) $=\omega^T \varphi(X) + b$ (7)
The first step consists of solving the following optimization problem:
$\underbrace{min}_{\omega,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\varepsilon_i + \varepsilon_i^*)$ (8)

Under the following constraints: $\begin{cases} Y_i - \omega^T\varphi(X_i) - b \leq \varepsilon_i + \varepsilon_i^* \\ \omega^T\varphi(X_i) + b - Y_i \leq \varepsilon_i + \varepsilon_i^* \\ \varepsilon_i, \varepsilon_i^* \geq 0 \end{cases}$ (9)

And :

- $\varphi(X_i)$: The kernel function is used to map the data into a higher-dimensional space, thereby facilitating a more linear relationship between the variables.
- C : This parameter represents the regularization term, balancing the model's complexity against the allowable level of error.
- $\varepsilon_i et\ \varepsilon_i^*$: Slack variables are introduced to manage data points that fall outside the margin.ε.

Testing and Evaluating - Cross-Validation

For these methods, we employed k-fold cross-validation to randomly divide the data into *k* mutually exclusive subsets for training and testing. The model is trained on *k–1* folds and tested on the remaining fold. Delen et al. (2012) demonstrated that a single random partition can produce heterogeneous subsets, potentially resulting in biased estimates. To address this, we performed five rounds of 10-fold cross-validation on the entire dataset. In each round, the model was trained on all folds except one, which served as the test subset for that round. The final performance was obtained by averaging the results across the five rounds. Olson and Delen (2008) noted that stratified cross-validation reduces bias compared to standard cross-validation. According to Delen et al. (2012), overall accuracy is calculated as the mean of the individual *k* accuracy measurements.
$CV = \frac{1}{K}\sum_{i=1}^{K} A_i$ (10)

Model Evaluation Metrics

* MSE measures the average squared difference between the predicted and actual values, providing an indication of the model's accuracy. Specifically, for each model, the MSE is calculated as follows:

$MSE_j = \frac{1}{n_{test}}\sum_{i=1}^{n}(y_i - \hat{y}_{i,j})$ (11)

Where:

- n_{test} : The number of observations in the test set.
- y_i : denotes the observed value.
- $\hat{y}_{i,j}$: The predicted value.

* MAPE measures the average of the absolute differences between the actual values and the predicted values, expressed as a percentage of the actual values. The formula is:

$$\text{MAPE}=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{Y_t-\hat{Y}_t}{Y_t}\right)*100 \quad (12)$$

Where:
n: total number of observations.
Y_t: Actual value at time t.
$\hat{Y}_t$: predicted value at time t.

DATASET

The empirical study focuses on countries (Appendix 1) from the MENA region (Middle East and North Africa) and covers the period from 2014 to 2024. Table 1 presents the variables used in the model, their abbreviations, and the data sources, in order to provide a clearer understanding of the relationship between these factors and income inequality. The relationship between economic growth (LGDP), corruption (CORR), government effectiveness (EG), political stability and absence of violence/terrorism (SP), regulatory quality (QR), and rule of law (ED) is expressed as follows:

LGDP=f (CORR, EG, SP, QR, ED,) (13)

Table 2 summarizes the statistics of the variables under study. The descriptive analysis reveals that all governance indicators exhibit negative mean values, reflecting relatively weak institutional performance within the sample. The degree of dispersion varies moderately across variables, with political stability (SP) showing the highest variability. Furthermore, the Shapiro–Wilk normality test indicates that all variables have p-values below 0.05, leading to the rejection of the null hypothesis of normality. These findings suggest that the data deviate from a normal distribution, thereby supporting the use of non-parametric or machine learning methods, such as the Random Forest model, for a more robust analysis. Moreover, machine learning algorithms are capable of uncovering deeply hidden patterns in large datasets involving multiple types of input variables that are not necessarily normally distributed (Delen et al., 2012; Sharda et al., 2017).

Table 1. Description of the Data.

Variables	**Abbreviation**	**Sources**
Economic growth GDP (constant 2010 US$)	LGPD	https://donnees.banquemondiale.org/
Corruption	CORR	https://donnees.banquemondiale.org/
Government effectiveness	EG	https://donnees.banquemondiale.org/
Political stability and absence of violence/terrorism	SP	https://donnees.banquemondiale.org/
Regulatory quality	OR	https://donnees.banquemondiale.org/
rule of law	ED	https://donnees.banquemondiale.org/

Table 2. Descriptive Statistics.

Variables	**Observations**	**Mean**	**Min**	**Max**	**Std Dev**	**Shapiro-Wilk stat**	**P-value**
LGPD	450	10.9148	10.1146	11.8316	0.4341	0.9567	0.0000*
CORR	450	-0.1806	-2.3076	1.5671	0.7703	0.9886	0.0010*
EG	450	-0.1806	-2.3076	1.5092	0.8034	0.9886	0.0014*
SP	450	-0.5841	-3.1808	1.2236	1.0354	0.9756	0.0000*
OR	450	-0.2139	-2.0921	1.2789	0.7944	0.9607	0.0000*
ED	450	-0.9366	-2.0503	0.7866	0.6105	0.9477	0.0000*
VR	450-0.2838		-2.3470	1.3167	0.8688	0.9706	0.0000*

Note: * represent significance levels of 5%.

RESULTS AND DISCUSSION

To predict economic growth in the MENA region, four machine learning models were developed: Random Forest, Support Vector Machine (SVM), Boosted Tree (Gradient Boosting Machine, GBM), and Multiple Linear Regression. The dataset was divided into two subsets—70% for model training and 30% for testing—following the standard validation procedure. This division allows for assessing the generalization capacity of the models while reducing the risk of overfitting. Model performance was evaluated using two accuracy indicators: Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE).

The comparative evaluation (Table 3) of econometric and machine learning models—Ordinary Least Squares (OLS), Random Forest, GBM, and Support Vector Regression (SVR)—reveals significant differences in predictive performance. The OLS model achieved the lowest mean squared error (MSE = 0.5326), reflecting a strong overall fit and internal consistency. However, the Random Forest model recorded the lowest mean absolute percentage error (MAPE = 0.0192), suggesting superior relative accuracy in predicting GDP. This advantage can be attributed to the model's non-parametric and robust nature, capable of capturing nonlinear relationships and complex interactions among explanatory variables (CORR, EG, SP, RQ, and ED).

The GBM model, based on an iterative ensemble of boosted trees, produced intermediate results (MSE = 0.6144; MAPE = 0.0217), indicating satisfactory yet slightly weaker performance than the Random Forest, possibly due to partial overfitting or heightened sensitivity to hyper parameter selection. Conversely, the SVR model yielded the least accurate results (MSE = 0.8047), likely stemming from an unsuitable kernel choice or difficulty in generalizing structural variations within macroeconomic data. Overall, these findings highlight the increasing relevance of machine learning approaches in economic modeling, particularly when dealing with complex and multidimensional datasets. The Random Forest model emerged as the most accurate and stable approach for GDP prediction in the MENA context, confirming its ability to handle nonlinearity, heteroskedasticity, and inter-variable correlations simultaneously.

Figure 1 illustrates the comparison between training and testing datasets for the main model variables—corruption (CORR), government effectiveness (EG), political stability (SP), regulatory quality (RQ), rule of law (ED), and gross domestic product (GDP). This representation verifies the statistical and structural coherence between the two subsets used in the modeling process. The graph shows strong similarity in dynamics and variation amplitudes between the two datasets. The trajectories of the variables display comparable trends, indicating a homogeneous and balanced data distribution. This suggests that the data-splitting procedure was properly executed, avoiding any selection bias that could undermine external validity. Moreover, the temporal coherence between the subsets indicates no significant structural breaks or data drift. The observed fluctuations in governance variables (CORR, EG, SP, RQ, ED) remain stable across samples, reinforcing the robustness of the supervised learning model. For the key economic variable (GDP), although slight variability is noted in the test set, it remains within the training range, reflecting continuity in macroeconomic trends during the study period.

Methodologically, this coherence validates the 70/30 partitioning process, ensuring both reliability and generalization capacity. The model was exposed to a representative sample of the economic and institutional reality, limiting overfitting and improving predictive accuracy in the test phase. This stability confirms a solid empirical foundation for evaluating forecasting performance. Hence, the use of non-parametric models proves to be a robust and complementary alternative to traditional econometric methods, offering deeper empirical insights into the determinants of economic performance.

The Random Forest model stands out as the most reliable and precise, owing to its ability to capture nonlinear and interactive relationships between governance variables and economic growth. As emphasized by Hamza and Larocque (2005), ensemble models such as Random Forest effectively reduce prediction variance while maintaining low bias, enhancing the reliability of estimates.

The variable importance analysis (Table 4, Figure 2) derived from the Random Forest model demonstrates that all institutional and governance indicators significantly influence economic growth. Among them, corruption (CORR) appears as the most influential factor (%IncMSE = 41.37; IncNodePurity = 39.65). This finding confirms that corruption represents a major obstacle to economic development by undermining public policy efficiency and discouraging private investment. As noted by Mauro (1995) and Tanzi and Davoodi (1997), corruption diverts resources toward unproductive activities, reduces infrastructure quality, and erodes institutional trust. Theoretically, this relationship aligns with North's (1990) institutional framework, which posits that institutional quality shapes economic incentives and market performance.

Government effectiveness (EG), with a %IncMSE of 24.92, underscores the importance of administrative capacity and governance quality. Effective governance ensures policy coherence, macroeconomic stability, and efficient resource allocation—factors conducive to growth (Kaufmann & Kraay, 2002). Within the framework of the new institutional economics, economic performance depends on the state's ability to build trust and reduce information asymmetries, thereby fostering long-term productivity and investment.

Political stability and absence of violence/terrorism (SP), though slightly less influential (%IncMSE = 22.88), remain crucial institutional components. Economic literature—particularly Barro (1996) and Alesina and Perotti (1996)—highlights that stable political environments enhance policy predictability, reduce country risk, and attract foreign direct investment. Conversely, political instability and internal conflicts hinder growth by heightening uncertainty and disrupting economic activity.

Regulatory quality (RQ) also plays a vital role in fostering a competitive economic environment. Clear, fair, and consistent regulation supports entrepreneurship, stimulates innovation, and promotes private sector participation in economic development (Kaufmann, Kraay & Zoido-Lobatón, 1999). This aligns with endogenous growth theory (Romer, 1986; Lucas, 1988), which emphasizes that public and institutional policies shape human and technological capital accumulation.

Finally, the rule of law (ED), with a substantial contribution (%IncMSE = 32.87), reinforces the institutional framework's strength. The rule of law guarantees contract enforcement, property rights protection, and legal predictability—key elements of secure economic transactions. According to Acemoglu, Johnson, and Robinson (2001, 2005), countries with robust legal institutions tend to experience sustained growth, as confidence in the legal system lowers transaction costs and encourages investment.

Taken together, these empirical results confirm that institutional quality and good governance are fundamental determinants of economic performance. Consistent with Rodrik, Subramanian, and Trebbi (2004), institutional strength—through transparency, political stability, regulatory quality, and rule of law—emerges as a structural lever for fostering sustainable and inclusive growth.

Therefore, these results should be considered a preliminary step in an ongoing research effort. Further analyses integrating structural variables (investment, trade openness, human capital) and temporal dynamics (catch-up effects or exogenous shocks) are required to enhance the model's explanatory power. Incorporating hybrid approaches combining artificial intelligence with structural econometrics could also improve the causal understanding of the identified relationships. In summary, this research underscores the potential of machine learning methods for economic forecasting while emphasizing the need for continued methodological and empirical refinement to generalize findings across MENA economies.

CONCLUSIONS

This study contributes to the growing literature on the intersection between governance and economic growth by applying advanced machine learning techniques to the MENA region. The comparative analysis of four predictive models—Ordinary Least Squares (OLS), Random Forest, Gradient Boosting Machine (GBM), and Support Vector Regression (SVR)—demonstrates that non-parametric approaches, particularly ensemble-based algorithms, outperform traditional econometric techniques in terms of accuracy and robustness. Among them, the Random Forest model exhibited the lowest Mean Absolute Percentage Error (MAPE), underscoring its superior capacity to handle nonlinearities, variable interactions, and structural heterogeneity across countries.

The findings highlight that corruption remains the most critical constraint on economic growth, followed by rule of law, government effectiveness, regulatory quality, and political stability. These results are consistent with the institutionalist framework of North (1990) and the empirical evidence of Acemoglu, Johnson, and Robinson (2001, 2005), reinforcing the view that strong, transparent, and accountable institutions are essential for long-term economic resilience. In the MENA context, where many economies depend heavily on natural resources and exhibit institutional fragility, governance reform represents a strategic lever for promoting diversification and sustainable development. From a methodological perspective, this research illustrates the potential of machine learning models to enrich traditional economic analysis by uncovering hidden patterns and nonlinear relationships that standard econometric models often overlook. The integration of artificial intelligence techniques into economic modeling offers a complementary and adaptive framework, capable of managing multidimensional data and providing more reliable forecasts.

However, this study is not without limitations. First, the analysis relies on cross-country panel data that may conceal unobserved heterogeneity and country-specific dynamics. Second, the predictive models used, while powerful, remain essentially correlational and do not fully address causality between governance and growth. Third, the study focuses exclusively on macro-institutional indicators and excludes potentially relevant structural variables such as investment rates, trade openness, education, or demographic trends, which could refine the explanatory framework. Furthermore, the static nature of the models prevents capturing temporal dependencies, feedback effects, and dynamic adjustments inherent to economic systems.

Future research could overcome these limitations by adopting hybrid modeling strategies that combine machine learning algorithms with structural econometric approaches, such as panel cointegration or vector autoregression (VAR) frameworks. Incorporating temporal and spatial dimensions, as well as disaggregated sectoral data, could further enhance the interpretability and policy relevance of the findings. In addition, the exploration of causal inference techniques—such as causal forests or Bayesian networks—would enable a deeper understanding of the mechanisms through which governance influences economic performance.

In conclusion, this study underscores both the empirical and methodological significance of integrating machine learning with institutional economics. By demonstrating the predictive power and interpretive value of ensemble models such as Random Forests, it opens new avenues for evidence-based policymaking in emerging economies. Strengthening governance quality and institutional effectiveness thus remains a cornerstone for achieving inclusive and sustainable growth in the MENA region, while data-driven analytical tools hold great promise for guiding future economic strategies and public policy design.

Table 3. Comparison of forecasts of economic growth (GDP)

	MSE	**MAPE**
Random Forest	**0.6497**	**0.0192**
Support Vector Machine (SVR)	0.8047	0.0211
Boosted TREE (GBM)	0,6144	0.0217
Linear Regression (OLS)	0.5326	0.0216

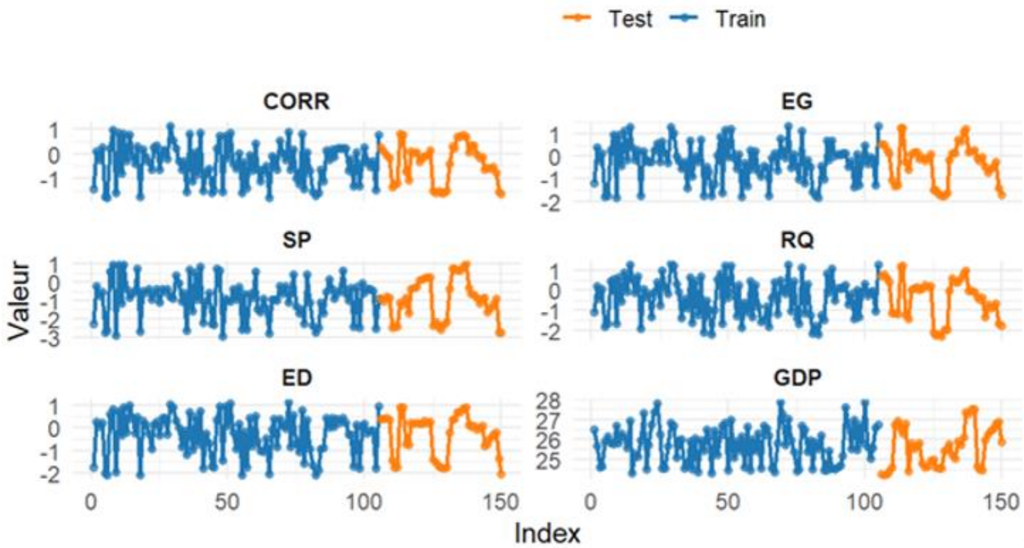

Figure 1. Comparison of training and test data

Table 4. Variable Importance in the Random Forest Model.

	%IncMSE Inc	**NodePurity**
CORR	41.37076	39.64506
EG	24.91943	21.34495
SP	22.88230	21.30756
RO	29.86353	26.06052

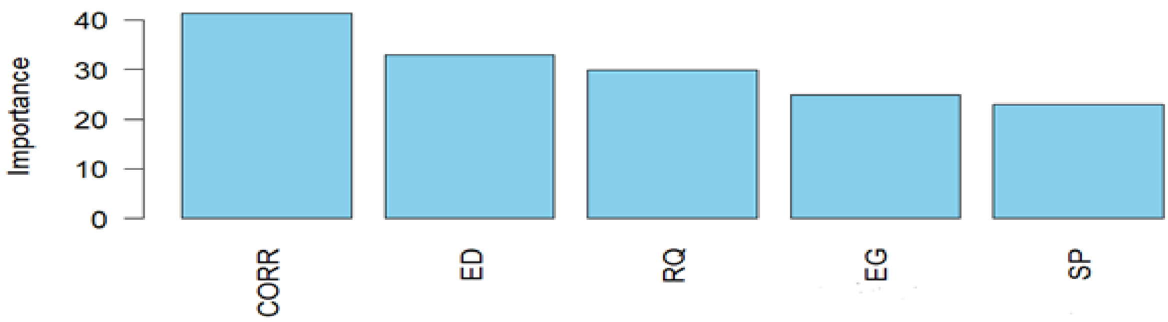

Figure 2. Importance of variables.

REFERENCES

Alesina, A., & Perotti, R. (1996). Income distribution, political instability, and investment. European Economic Review, 40(6), 1203–1228. https://doi.org/10.1016/0014-2921 (95)00030-5

Andrews, M. (2008). The good governance agenda: Beyond indicators without theory. Oxford Development Studies, 36(4), 379–407. https://doi.org/10.1080/13600810802455120

Armstrong, E., & others. (2005). Integrity, transparency and accountability in public administration: Recent trends, regional and international developments and emerging issues. United Nations, Department of Economic and Social Affairs.

Baliamoune-Lutz, M. (2019). Institutions, governance, and economic development in the MENA region. In M. Baliamoune-Lutz (Ed.), Sustainable Development in the Middle East and North Africa (pp. 55–79). Routledge.

Barro, R. J. (1996). Democracy and growth. Journal of Economic Growth, 1(1), 1–27. https://doi.org/10.1007/BF00163340.

Ben Ali, M. S., & Sassi, S. (2016). Institutional quality, trust, and economic growth in the MENA countries: Evidence from panel data. Journal of Business and Economic Development, 1(1), 21–31.

Bovaird, T., & Löffler, E. (2009). Public Management and Governance (2nd ed.). Routledge.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Delen, D., Walker, G., & Kadam, A. (2012). Predicting breast cancer survivability: A comparison of three data mining methods. Artificial Intelligence in Medicine, 34(2), 113–127. https://doi.org/10.1016/j.artmed.2004.07.002.

Graham, J., Amos, B., & Plumptre, T. (2003a). Principles for good governance in the 21st century. Policy Brief No. 15, Institute on Governance, Ottawa.

Hamza, M., & Larocque, D. (2005). An empirical comparison of ensemble methods based on classification trees. Journal of Statistical Computation and Simulation, 75(8), 629–643. https://doi.org/10.1080/00949650410001675995

Huang, C., Tzeng, G. H., & Ong, C. S. (2014). Machine learning approaches to modeling regional economic growth. Expert Systems with Applications, 41(5), 2478–2492. https://doi.org/10.1016/j.eswa.2013.09.027.

Kaufmann, D., Kraay, A., & Zoido-Lobatón, P. (1999). Governance matters. World Bank Policy Research Working Paper No. 2196.

Kaufmann, D., Kraay, A., & Zoido-Lobatón, P. (2002). Governance matters II: Updated indicators for 2000–01. World Bank Policy Research Working Paper No. 2772.

Lucas, R. E. Jr. (1988). On the mechanics of economic development. Journal of Monetary Economics, 22(1), 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

Mauro, P. (1995). Corruption and growth. The Quarterly Journal of Economics, 110(3), 681–712. https://doi.org/10.2307/2946696

North, D. C. (1990). Institutions, institutional change and economic performance. Cambridge University Press.

Olson, D. L., & Delen, D. (2008). Advanced data mining techniques. Springer. https://doi.org/10.1007/978-3-540-76917-0

Romer, P. M. (1986). Increasing returns and long-run growth. Journal of Political Economy, 94(5), 1002–1037. https://doi.org/10.1086/261420

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. Journal of Economic Growth, 9(2), 131–165. https://doi.org/10.1023/B:JOEG.0000031425.72248.85.

Sharda, R., Delen, D., & Turban, E. (2017). Business Intelligence, Analytics, and Data Science: A Managerial Perspective (4th ed.). Pearson.

Solow, R. M. (1956). A contribution to the theory of economic growth. The Quarterly Journal of Economics, 70(1), 65–94. https://doi.org/10.2307/1884513

Swan, T. W. (1956). Economic growth and capital accumulation. Economic Record, 32(2), 334–361. https://doi.org/10.1111/j.1475-4932.1956.tb00434.

Tanzi, V., & Davoodi, H. (1997). Corruption, public investment, and growth. IMF Working Paper No. 97/139.

Appendix 1. The studied MENA countries.

United Arab Emirates	**Bahrain**
Algeria	Egypt
Iran	Iraq
Israel	Jordan
Kuwait	Lebanon
Libya	Morocco
Oman	Qatar
Saudi Arabia	Syrian Arab Republic
Tunisia	Yemen

