**A PREDICTIVE MODEL FOR PROSPERITY INDEX IN WEST AFRICA USING RANDOM FOREST**

TABLE OF CONTENT

[CHAPTER ONE 3](#_Toc189135191)

[1.1 Background Of Study 3](#_Toc189135192)

[1.2 Statement Of Problem 4](#_Toc189135193)

[1.3 Aim/Objectives 4](#_Toc189135194)

[1.4 Significance Of the Study 5](#_Toc189135195)

[1.5 Scope Of The Study 5](#_Toc189135196)

[1.6 Limitations Of The Study 6](#_Toc189135197)

[CHAPTER TWO 7](#_Toc189135198)

[2.0 Overview of Prosperity Index (LPI) 7](#_Toc189135199)

[2.1 Dimensions of Prosperity Index (LPI) 7](#_Toc189135200)

[2.1.1 Economy 8](#_Toc189135201)

[2.1.2 Education 8](#_Toc189135202)

[2.1.3 Health 9](#_Toc189135203)

[2.1.4 Governance 9](#_Toc189135204)

[2.1.5 Personal Freedom 10](#_Toc189135205)

[2.1.6 Entrepreneurship 10](#_Toc189135206)

[2.1.7 Social Capital 10](#_Toc189135207)

[2.1.8 Safety and Security 11](#_Toc189135208)

[2.2 Significance and Applications of Prosperity Index (LPI) in West Africa 11](#_Toc189135209)

[2.2.1 National and Regional Assessments 11](#_Toc189135210)

[2.2.2 Economic and Sustainable Development 12](#_Toc189135211)

[2.2.3 Policy Making and Comparative Studies 13](#_Toc189135212)

[2.3 Machine Learning Techniques and Random Forest Models 13](#_Toc189135213)

[2.3.1 Overview of Machine Learning Techniques 13](#_Toc189135214)

[2.3.2 Introduction to Random Forest Models 14](#_Toc189135215)

[2.4 Applications of Random Forest in Predictive Analytics 15](#_Toc189135216)

[2.4.1 Healthcare Predictions 16](#_Toc189135217)

[2.4.2 Economic Forecasting 16](#_Toc189135218)

[2.4.3 Environmental Studies 17](#_Toc189135219)

[2.5 Integrating LPI and Random Forest Models in West Africa 17](#_Toc189135220)

[2.6 Related Literature 19](#_Toc189135221)

[REFERENCES 27](#_Toc189135222)

# CHAPTER ONE

## 1.1 Background Of Study

Understanding and predicting prosperity in West Africa is crucial for effective policy-making and fostering sustainable development across the region. The Prosperity Index serves as a key metric for assessing the multifaceted elements contributing to a nation’s well-being, including economic quality, governance, education, health, and social inclusion. Despite the rich cultural diversity and abundant natural resources in West Africa, challenges such as poverty, political instability, and limited access to education persist, necessitating innovative approaches to drive development and prosperity.

Recent advancements in machine learning, particularly the Random Forest algorithm, have created opportunities to address these challenges through predictive modeling. Random Forest, an ensemble learning method that constructs multiple decision trees, is well-suited for analyzing high-dimensional datasets with complex interactions among variables (Pahl *et al*., 2020). By leveraging this method, researchers can gain valuable insights into the key factors influencing prosperity and forecast future trends with greater accuracy.

The application of Random Forest models has been widely explored in fields such as healthcare, finance, and environmental studies, where it has proven effective in handling complex datasets. For example, in healthcare, Random Forest has been used to predict survival rates for kidney transplants and cancer treatments (Xia, 2023). In finance, it has aided in forecasting stock market trends, while in environmental science, it has helped identify optimal agricultural practices (Khandagale, 2023; Gao *et al*., 2021). These successes highlight the algorithm's versatility and its potential to uncover actionable insights for development in West Africa. In this study, the Prosperity Index will be analyzed using Random Forest models, providing a systematic framework to evaluate and predict the prosperity levels of countries in West Africa. This approach aims to inform policymakers and stakeholders on effective strategies for addressing regional disparities and enhancing the well-being of citizens.

## 1.2 Statement Of Problem

Traditional methods for analyzing prosperity, such as Multi-Criteria Decision-Making (MCDM) frameworks, often fall short in capturing the intricate and multidimensional nature of factors influencing prosperity. These methods lack the predictive power to analyze large datasets and fail to provide nuanced insights that reflect regional variations within West Africa. Additionally, they rely on static, non-dynamic approaches that are not well-suited to handle complex socio-economic systems.

West Africa faces unique challenges, including political instability, varying levels of governance, economic disparity, and environmental vulnerabilities, that require a more robust and data-driven solution. The lack of advanced predictive models capable of integrating these dimensions has created a gap in understanding the determinants of prosperity in the region. Addressing this gap through machine learning approaches such as Random Forest can provide actionable insights to drive sustainable development.

## 1.3 Aim/Objectives

The aim of this project is to develop a predictive model for the Prosperity Index in West Africa using the Random Forest algorithm.

**Objectives**:

1. To synthesize existing frameworks and methods related to prosperity prediction, highlighting their strengths and limitations in the West African context.
2. To develop a predictive model employing Random Forest algorithms for analyzing and forecasting prosperity levels in West Africa.
3. To collect and preprocess socio-economic, governance, and environmental data relevant to the Prosperity Index.
4. To identify key variables that influence prosperity in West African countries, providing targeted recommendations for policy interventions.
5. To evaluate the model's performance using appropriate metrics and validate its applicability in real-world contexts.
6. To design a user-friendly interface to facilitate the interpretation of model predictions by policymakers and stakeholders.

## 1.4 Significance Of the Study

This study is significant as it provides a data-driven approach to understanding and predicting prosperity in West Africa. By leveraging the Random Forest algorithm, the study will enhance the predictive accuracy of the Prosperity Index and offer a comprehensive analysis of socio-economic and governance-related factors influencing the region's development. The findings will help policymakers formulate evidence-based strategies to improve economic quality, governance, and social inclusion.

## 1.5 Scope Of The Study

This study focuses on developing a predictive model for the Prosperity Index in West Africa, with data sourced from regional and international organizations. The model will consider key indicators, including governance quality, economic metrics, health outcomes, education levels, and environmental sustainability.

The study will utilize Python for data preprocessing, analysis, and model implementation. The scope includes evaluating prosperity trends within West African countries and identifying critical factors driving these trends. Emphasis will be placed on addressing the unique challenges of the region, such as inequality, poor infrastructure, and governance issues. The project also includes creating a user-friendly interface to facilitate model application and provide actionable insights for stakeholders across various sectors.

## 1.6 Limitations Of The Study

Data Availability: Reliable and comprehensive data for all countries in West Africa may not be readily available, potentially leading to gaps in the analysis.

Model Generalizability: While the Random Forest algorithm is robust, its findings might not be universally applicable to regions outside West Africa without adjustments to account for unique regional contexts.

Interpretability: Although Random Forest provides accurate predictions, its ensemble nature can make it challenging to interpret the exact contribution of each variable to the model’s predictions.

# CHAPTER TWO

**LITERATURE REVIEW**

## 2.0 Overview of Prosperity Index (LPI)

The Prosperity Index (LPI) is a comprehensive tool developed by the Legatum Institute to measure and evaluate the prosperity of nations. Unlike traditional economic indicators such as GDP, the LPI takes a multidimensional approach to assess a country's overall well-being and quality of life. The index covers various aspects of prosperity, including economic performance, social well-being, and governance, providing a holistic view of what contributes to a prosperous society (Dudukalov *et al*., 2020). The LPI is structured around several key pillars: economy, education, health, governance, personal freedom, entrepreneurship and opportunity, social capital, and safety and security. These dimensions collectively provide a nuanced understanding of national prosperity, making the LPI a valuable tool for policymakers and researchers. In the context of West Africa, the LPI can be particularly useful for identifying areas of strength and weakness across the region. West Africa, with its diverse economies and varying levels of development, presents a unique opportunity to apply the LPI framework to understand the drivers of prosperity in the region. By analyzing the LPI scores of West African countries, policymakers can identify key areas for intervention and develop targeted strategies to enhance regional prosperity.

## 2.1 Dimensions of Prosperity Index (LPI)

The Prosperity Index (LPI) encompasses several dimensions, each representing a critical aspect of national prosperity. These dimensions include the economy, education, health, governance, personal freedom, entrepreneurship, social capital, and safety and security. Each dimension is integral to understanding and measuring the comprehensive well-being and prosperity of a nation.

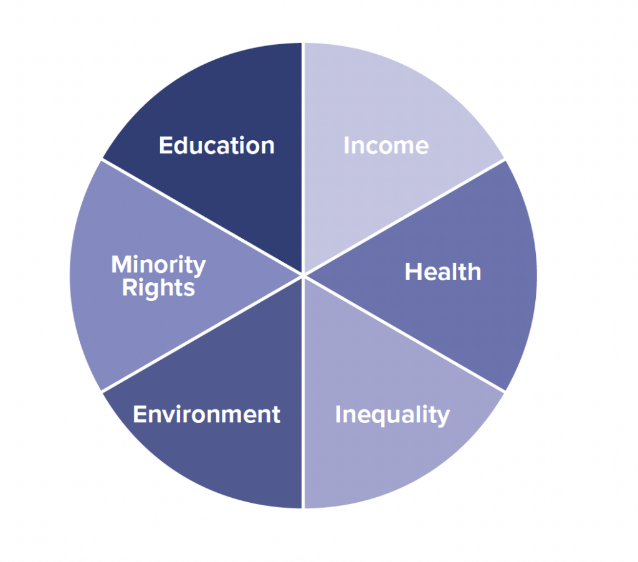


fig 2.1 component of Prosperity index

### 2.1.1 Economy

The economy dimension of the LPI evaluates the macroeconomic environment of a country. This includes assessing factors such as economic stability, growth rates, employment levels, inflation, and productivity. In West Africa, where economic performance varies significantly across countries, this dimension is crucial for understanding how financial health contributes to the broader well-being of society (Dudukalov *et al*., 2020). For instance, countries like Nigeria and Ghana have relatively robust economies compared to others in the region, but challenges such as inflation and unemployment remain significant barriers to prosperity.

### 2.1.2 Education

Education is a crucial dimension of the LPI as it directly impacts a nation's human capital development. This dimension examines access to education, quality of educational institutions, literacy rates, and educational outcomes. In West Africa, where access to quality education remains a challenge in many countries, improving educational outcomes is essential for fostering innovation and driving long-term economic growth (Despotović *et al*., 2019). Countries like Senegal and Ghana have made significant strides in improving education, but disparities still exist, particularly in rural areas.

### 2.1.3 Health

The health dimension of the LPI assesses the overall health of the population, which is a fundamental component of national prosperity. This includes metrics such as life expectancy, access to healthcare services, prevalence of diseases, and overall health outcomes. In West Africa, where healthcare systems are often underfunded and overburdened, improving health outcomes is critical for enhancing productivity and quality of life (Oprisan, 2023). Countries like Nigeria and Côte d'Ivoire face significant challenges in providing adequate healthcare, particularly in rural and underserved areas.

### 2.1.4 Governance

Governance is another key dimension of the LPI, focusing on the effectiveness, accountability, and transparency of government institutions. This dimension evaluates the rule of law, corruption levels, political stability, and the efficiency of public services. In West Africa, where governance challenges such as corruption and political instability are prevalent, improving governance is essential for creating a conducive environment for economic activities and ensuring that the benefits of prosperity are equitably distributed across society (Aktaş, 2023). Countries like Ghana and Senegal have made progress in improving governance, but challenges remain in other parts of the region.

### 2.1.5 Personal Freedom

Personal freedom, a critical dimension of the LPI, measures the extent to which individuals enjoy freedom of expression, belief, and association. It also considers the absence of legal discrimination and the degree of autonomy in personal decisions. In West Africa, where issues such as political repression and discrimination are still prevalent in some countries, enhancing personal freedom is essential for fostering a more vibrant and dynamic society where creativity and innovation can flourish (Gavrilović & Gligorić, 2018).

### 2.1.6 Entrepreneurship

The entrepreneurship dimension assesses the entrepreneurial environment within a country, including the ease of starting and running businesses, access to capital, and the presence of supportive regulations and policies. In West Africa, where entrepreneurship is a key driver of economic growth, improving the entrepreneurial ecosystem is essential for creating jobs, fostering innovation, and increasing competitiveness (Bate, 2021). Countries like Nigeria and Ghana have vibrant entrepreneurial ecosystems, but challenges such as access to finance and regulatory barriers remain significant obstacles.

### 2.1.7 Social Capital

Social capital refers to the networks of relationships, trust, and norms that facilitate cooperation within or among groups. This dimension of the LPI evaluates the strength of social networks, community engagement, and the level of trust in institutions and among individuals. In West Africa, where social cohesion is often challenged by ethnic and religious tensions, enhancing social capital is essential for achieving sustainable prosperity (Oh & Rho, 2017).

### 2.1.8 Safety and Security

Safety and security are fundamental to the prosperity of any nation. This dimension measures the level of violence, crime, and political instability, as well as the effectiveness of law enforcement. In West Africa, where issues such as terrorism, political instability, and high crime rates are prevalent in some countries, ensuring safety and security is a prerequisite for the other dimensions of prosperity to flourish (Iurchenko & Iurchenko, 2020).

## 2.2 Significance and Applications of Prosperity Index (LPI) in West Africa

The Prosperity Index (LPI) is a pivotal tool for assessing and understanding various dimensions of prosperity in West Africa. Its significance spans multiple domains, including national and regional assessments, economic and sustainable development, and policy-making and comparative studies. The following sections elaborate on these applications and highlight the importance of the LPI in shaping economic and social policies in the region.

### 2.2.1 National and Regional Assessments

The LPI provides comprehensive insights into the prosperity levels of countries and regions by evaluating various dimensions such as economic performance, education, health, and governance. On a national level, the index allows policymakers and researchers to gauge overall well-being and identify areas needing improvement. For instance, the LPI's detailed data on education and health can help governments in West Africa target interventions where they are most needed, thereby enhancing quality of life and economic potential (Despotović *et al*., 2019).

Regionally, the LPI is instrumental in assessing disparities within West Africa. The index can help uncover local strengths and weaknesses, enabling tailored policies that address specific regional needs and promote balanced development. Additionally, the index facilitates comparisons between regions within West Africa, highlighting areas of relative advantage or disadvantage (Iurchenko & Iurchenko, 2020).

### 2.2.2 Economic and Sustainable Development

The LPI's dimension on economic performance provides valuable data for understanding the economic health of nations in West Africa. It evaluates factors such as GDP growth, employment rates, and economic stability, which are crucial for fostering sustainable development. Countries with high LPI scores in economic performance are often better positioned to invest in infrastructure, education, and healthcare, contributing to long-term economic stability and growth (Dudukalov *et al*., 2020).

Sustainable development is another critical area where the LPI plays a significant role in West Africa. By integrating environmental indicators and assessing the impact of economic activities on sustainability, the LPI helps in identifying nations that are effectively balancing economic growth with environmental stewardship. For example, studies have shown that the LPI can be correlated with environmental indicators such as carbon emissions, highlighting how prosperity is linked to environmental policies and practices (Alotaibi & Alajlan, 2021). This aspect of the LPI underscores the importance of aligning economic development with sustainable practices to ensure that prosperity does not come at the expense of future generations.

### 2.2.3 Policy Making and Comparative Studies

The LPI is a valuable resource for policymakers in West Africa, providing a benchmark for evaluating the effectiveness of policies and programs aimed at enhancing prosperity. By offering a multi-dimensional view of prosperity, the index helps in identifying which areas of policy need adjustment or enhancement. For instance, if a country scores poorly in governance, policymakers can focus on improving transparency and reducing corruption to boost overall prosperity (Aktaş, 2023). In comparative studies, the LPI serves as a tool for benchmarking the performance of different countries in West Africa. It allows for comparisons between nations and regions, providing insights into best practices and successful strategies that can be adapted or adopted by others. This comparative approach is useful for understanding how different policies impact prosperity and for learning from the experiences of more prosperous countries (Gavrilović & Gligorić, 2018).

## 2.3 Machine Learning Techniques and Random Forest Models

Machine learning offers a variety of techniques for tackling different problems. Two fundamental categories are supervised and unsupervised learning, but there are many specific algorithms within these categories.

### 2.3.1 Overview of Machine Learning Techniques

Machine learning (ML) is a field of artificial intelligence that focuses on developing algorithms that enable computers to learn from and make predictions based on data. The primary goal of machine learning is to identify patterns within data and use these patterns to make informed decisions or predictions. Machine learning techniques can be broadly categorized into supervised, unsupervised, and reinforcement learning, each serving different purposes based on the nature of the data and the problem at hand. In supervised learning, algorithms are trained on labeled data, meaning that the input data is paired with corresponding output labels. The algorithm learns to map inputs to outputs and can then predict the output for new, unseen data. Common supervised learning techniques include linear regression, logistic regression, support vector machines (SVM), and neural networks (Goodfellow *et al*., 2016).

Unsupervised learning involves algorithms that are used with unlabeled data, where the goal is to identify hidden patterns or groupings within the data. Techniques such as clustering and dimensionality reduction are popular in this category. For example, k-means clustering groups data into distinct clusters based on similarity, while principal component analysis (PCA) reduces the dimensionality of the data to uncover the most significant features (Jolliffe & Cadima, 2016).

Reinforcement learning, on the other hand, is used to train models to make sequences of decisions by learning from the consequences of their actions. This approach is commonly applied in areas such as robotics and game playing, where an agent learns to optimize its behavior through trial and error to achieve a specific goal (Sutton & Barto, 2018).

### 2.3.2 Introduction to Random Forest Models

Random Forest is a versatile and widely used machine learning algorithm that belongs to the ensemble learning category. It builds on the concept of combining multiple decision trees to improve predictive performance and robustness. The fundamental idea behind Random Forest is to create a "forest" of decision trees, each trained on a random subset of the data and features, and then aggregate their predictions to make a final decision (Breiman, 2001). The Random Forest algorithm operates through two main processes: bootstrapping and feature randomization. Bootstrapping involves creating multiple subsets of the original dataset by sampling with replacement. Each decision tree in the forest is trained on a different subset, which helps to reduce overfitting and improve generalization. Feature randomization means that at each node of a decision tree, a random subset of features is considered for splitting, further enhancing the model's diversity and reducing correlation among individual trees (Liaw & Wiener, 2002). One of the key advantages of Random Forest models is their ability to handle high-dimensional data and provide robust predictions even with noisy or incomplete data. They are less prone to overfitting compared to individual decision trees, making them suitable for various applications, including classification, regression, and feature selection (Liaw & Wiener, 2002). Random Forest models can provide insights into feature importance, helping to identify which variables contribute most significantly to the predictions. In practice, Random Forest models have been successfully applied across diverse fields, from predicting stock prices and disease outcomes to analyzing customer preferences and environmental trends (Breiman, 2001; Liaw & Wiener, 2002). Their versatility and effectiveness make them a popular choice for researchers and practitioners aiming to develop accurate and reliable predictive models.

## 2.4 Applications of Random Forest in Predictive Analytics

Random forests are widely used in predictive analytics due to their versatility and robustness. Some of the common applications of random forest in predictive analytics.

### 2.4.1 Healthcare Predictions

In healthcare, Random Forest models are widely used to enhance predictive accuracy and support decision-making processes. One of the primary applications is in predicting patient outcomes and disease progression. For example, Random Forest algorithms can analyze complex medical data, including patient demographics, clinical history, and test results, to predict disease risk and treatment efficacy (Liu *et al*., 2021). This application is particularly valuable for chronic diseases, such as cancer, where early and accurate prediction can significantly improve patient management and outcomes. Random Forest has also been used to predict hospital readmissions, which can help in planning patient care and reducing healthcare costs. By analyzing historical patient data, including past admissions, diagnoses, and treatment records, Random Forest models can identify patients at high risk of readmission, allowing healthcare providers to implement targeted interventions (Steinman *et al*., 2019). Additionally, Random Forest's ability to handle large and heterogeneous datasets makes it suitable for personalized medicine, where treatments are tailored based on individual patient profiles (Liu *et al*., 2021).

### 2.4.2 Economic Forecasting

n economic forecasting, Random Forest models are employed to predict financial trends, market movements, and economic indicators. For instance, they are used to forecast stock market returns by analyzing historical price data, trading volumes, and macroeconomic indicators. The ensemble approach of Random Forest helps in capturing complex patterns and relationships in financial data that single models might miss (Qolipour *et al*., 2021). Random Forest is also applied in credit risk assessment, where it helps financial institutions evaluate the creditworthiness of borrowers. By analyzing various factors, such as credit history, income levels, and loan details, Random Forest models can predict the likelihood of default, enabling lenders to make more informed lending decisions (Tufekci *et al*., 2022). Moreover, Random Forest's robustness to overfitting and ability to handle large datasets make it a valuable tool in economic forecasting and risk management.

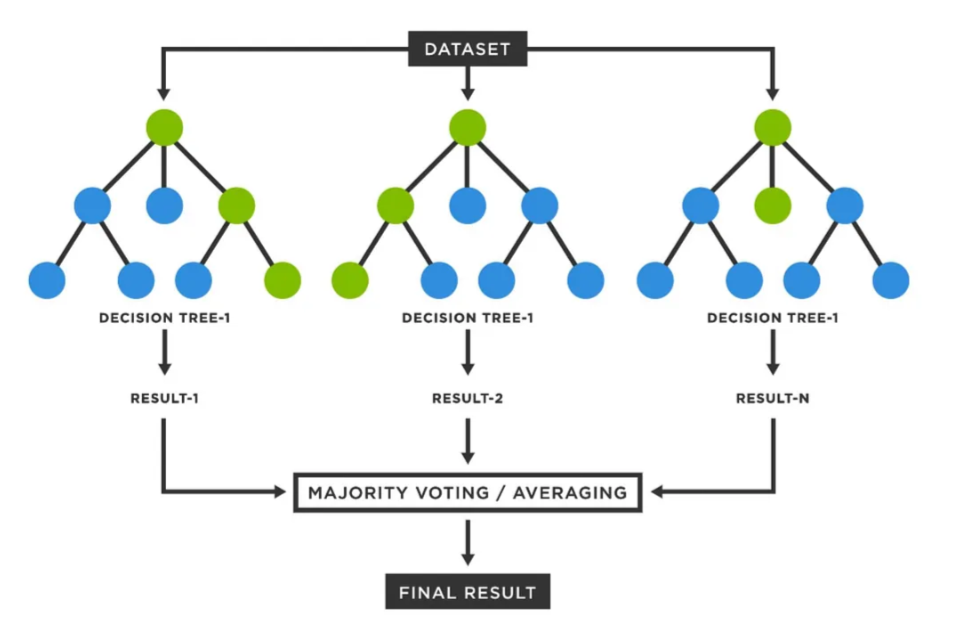
### 2.4.3 Environmental Studies

In environmental studies, Random Forest models are used to analyze and predict various ecological and environmental phenomena. One significant application is in climate modeling, where Random Forest helps in predicting temperature changes, precipitation patterns, and other climate variables based on historical data and climate models. The model's ability to handle large datasets and complex interactions between variables makes it suitable for studying climate change impacts (Hsu *et al*., 2020). Random Forest is also applied in agricultural studies to predict crop yields based on weather conditions, soil properties, and other environmental factors. By analyzing data from various sources, including satellite imagery and weather stations, Random Forest models can provide accurate predictions of crop performance, helping farmers optimize their practices and improve food security (Masjkur & Tan, 2020).

### 2.5 Integrating LPI and Random Forest Models in West Africa

Integrating the Prosperity Index (LPI) with Random Forest models involves employing advanced statistical and machine learning techniques to enhance the prediction and analysis of prosperity indicators in West Africa. The methodological approach begins with data preprocessing, which includes the collection and cleaning of data related to various dimensions of the LPI, such as economy, education, health, and governance. This step ensures that the data used in the Random Forest model is accurate and relevant. Once the data is prepared, Random Forest models are trained to recognize patterns and relationships within the dataset. This involves selecting the appropriate features that represent different aspects of the LPI and configuring the model to handle these features effectively. Random Forest, with its ensemble learning technique, builds multiple decision trees and aggregates their predictions, thereby improving the model's accuracy and robustness (Breiman, 2001).

Feature importance analysis is a crucial aspect of this integration. By evaluating the significance of various features, researchers can identify which aspects of the LPI most significantly impact overall prosperity in West Africa. This analysis helps in refining the model and focusing on the most influential variables (Liaw & Wiener, 2002). Additionally, hyperparameter tuning and cross-validation are employed to optimize the model's performance and prevent overfitting, ensuring that the predictions made by the Random Forest model are reliable and generalizable (Hastie *et al*., 2009).

 Fig 2.1 Structure of the random forest by Divakar Kumar

## 2.6 Related Literature

Altintaş’ (2023) work titled “Analysis of the Prosperity Performances of G7 Countries: An Application of the LOPCOW-based CRADIS Method” examines the economic performances of G7 countries, and determines their future prosperity retention rights. The author used the LOPCOW-based complex comparative research and analysis for a prosperity index on welfare to estimate the prosperity retention rights for G7 countries. The study found that the United States has a higher prosperity retention right than Canada, Germany, Italy, Japan, the UK, and France. Altintaş suggests that policymakers in respective countries use the findings of the research to focus on long-term wellbeing of their citizens. The results of the study can be utilized by political leaders to strengthen policies aimed at enhancing the prosperity of the G7 countries. The research findings may stimulate countries to focus on policies and strategies that ensure an improved prosperity rate for their citizens. The author believes that the research offers a baseline for future study and further discussion.

Anietie Ekong(2023) “Evaluation of Machine Learning Techniques Towards Early Detection of Cardiovascular Diseases” assesses the ability of machine learning techniques to detect cardiovascular diseases early. The work evaluates a number of machine learning models using methods such as cross-validation and accuracy measurements to aid in the early detection of cardiovascular diseases. According to the work, some techniques demonstrated high accuracy (i.e., ability to identify patients with cardiovascular diseases) of up to 90%. The work also states that further research and studies are necessary to determine the accuracy of machine learning techniques for early detection of cardiovascular diseases. The work notes that machine learning techniques can still produce false positives and negatives and require careful analysis and monitoring by medical professionals. In summary, the evaluation of machine learning techniques towards early detection of cardiovascular diseases shows promise, but further research is necessary to fully determine their accuracy and effectiveness.

Park, Yang, Kim, and Lim in 2022. The study presented in the work centers on the development of a model for estimating site index values for six major tree species in North Korea. A site index is an important measure of forest productivity as it helps to evaluate the potential growth capacity of a forest. The model was developed to aid forest managers in understanding the forest productivity in North Korea and to help identify areas requiring specific interventions for efficient forest management. To develop the model, the authors collected data on 217 sample sites distributed throughout North Korea between 2017 and 2019. During this period, the study team conducted field surveys to record soil properties, climate conditions, topography, and other variables potentially related to site index values. The authors used 105 sample sites for training the model and 112 for validation. Statistical methods and machine learning algorithms such as stepwise regression, artificial neural networks, and regression trees were used to create the model. These techniques were used to analyze and extract the most important variables associated with site index values for the six primary tree species found in North Korea.

Bashir’s research work “Prosperity Index 2013 Rankings for Asia and Pacific Countries and Evaluation of Prosperity Indicators for Pakistan," published in IOSR Journal of Humanities and Social Science, presents the 2013 prosperity ranking of Asian and Pacific countries based on the Institute’s Prosperity Index. Bashir also assessed the individual prosperity indicators that have the greatest impact on the rankings and conducted a similar analysis for Pakistan. The results of the study indicated that among Asian and Pacific countries, New Zealand secured the top position in the Prosperity Index, and among the prosperity indicators covered in the study, economy was the most crucial indicator for the ranking. In addition, the results showed that Pakistan ranked at the lowest position in the Prosperity Index of Asian and Pacific countries and the education and governance indicators had the lowest scores in the country. Bashir recommends that Pakistani policymakers prioritize focusing on the improvement of the education, governance, and economy indicators to progress on the path of future prosperity.

Despotović, Ristić, and Dimitrijević’s work, “Significance of Innovation for Sustainable Economic and Agricultural Development in the Republic of Serbia,” published in Facta Universitatis Series Economics and Organization, aims to provide insight into the importance of innovation for Serbia's economic and agricultural sustainability. The authors highlighted innovation and its fundamental role in bringing about profound economic opportunities and agricultural progress in the Republic of Serbia. The study intended to evaluate the existing innovation system, approaches to strengthening innovation policy, and subsequent policy recommendations. The study results unveiled that the current policy framework is only moderately effective in enhancing innovation performance and, therefore, requires a reformulation initiative with a well-structured approach. The study recommends comprehensive measures to address uncertainty in the business environment, increase investment in infrastructure, foster modernization of agriculture techniques, and concentrate on the exploitation of improved innovation system opportunities to establish long-term connections with industry and research establishments. The work offers a valuable source of information for policymakers and decision-makers involved in creating advanced strategies and constructs.

Matić, Gavrilović, and Stanišić's (2020) work “GDP and Beyond: Prosperity Convergence in the Countries of Western and Eastern Europe,” published in Acta Oeconomica, explores the convergence of prosperity in Eastern and Western Europe countries. The authors focus on assessing whether the convergence process is limited to the GDP or goes beyond it. They use a prosperity index that is based on ten sub-indexes: governance, innovation, business environment, human capital, education, infrastructure, social capital, natural environment, health, and financial sector development. The study shows that the catch-up process towards the EU average has been considerable, and there is a positive trend of convergence among the EU and non-EU countries in CESEE. However, the study found that income growth alone does not necessarily lead to increased prosperity. The convergence is more significant when non-economic factors are considered, suggesting the broader perception of prosperity beyond GDP. According to the authors, the findings from the research suggest that the EU governments should create policy measures that address other non-monetary sources of prosperity to achieve sustainable convergence.

Gavrilović and Gligorić's (2019) work, "Fundamental Drivers of Prosperity in the European Union and Western Balkans Countries," published in Industrija, evaluates the drivers of prosperity in European Union (EU) and Western Balkans (WB) countries and compares their performances economically. The study identifies the key variables that impact economic growth, including foreign direct investment, governance, economic freedom, innovation, education, and infrastructure. The research suggests that EU countries display a higher level of prosperity than WB countries considering the selected indicators in the study. WB nations lag behind EU countries because of weak governance, deficient infrastructure, lower levels of economic freedom, and stagnated innovation. The study recommends innovation promotion, improving infrastructure, and restructuring the governance system to increase the level of prosperity in WB countries and foster the growth of these economies. The authors conclude that the findings from the study are helpful for policymakers engaged in making sustainable policies that can enhance the prosperity of their respective countries

Alotaibi and Alajlan (2021) conducted a study to investigate the relationship between carbon dioxide emissions and socioeconomic indicators in G20 countries using quantile regression analysis. The authors used data from G20 countries for the years 1990 to 2017, and found that gross domestic product per capita, energy consumption, and urbanization had a significant positive relationship with carbon dioxide emissions in the upper quantiles, while renewable energy consumption had a negative relationship with carbon dioxide emissions in the upper quantiles. In the lower quantiles, the authors found that gross domestic product per capita, energy consumption, and urbanization had a positive relationship with carbon dioxide emissions, while renewable energy consumption had a negative relationship. The authors concluded that their findings provide important insights for policymakers to develop effective strategies to mitigate carbon dioxide emissions. This work has a high level of accuracy and comprehensively analyzes the relationship between socioeconomic indicators and carbon dioxide emissions using quantile regression analysis.

Dudukalov *et al*. (2020) explored the correlation between the automotive industry and human well-being in the era of economic digitalization. The authors analyzed data from various sources, including the World Bank, the International Monetary Fund, and the United Nations, and reviewed literature on the topic. The study found that the automotive industry has a significant impact on human well-being, including economic growth, employment opportunities, and transportation infrastructure. The authors also argued that the era of economic digitalization has led to new challenges and opportunities for the automotive industry, including the rise of electric vehicles and the development of autonomous driving technologies. The work has a moderate level of accuracy and provides a broad overview of the correlation between the automotive industry and human well-being in the era of economic digitalization.

Iurchenko and Iurchenko (2020) investigated the modern trends of tourism development in Ukraine. The authors analyzed data from various sources, including the State Statistics Service of Ukraine and the World Tourism Organization, and reviewed literature on the topic. The study found that tourism is an important sector for the Ukrainian economy, contributing to job creation and economic growth. The authors also identified several modern trends in tourism development in Ukraine, including the rise of domestic tourism, the growth of adventure and eco-tourism, and the increasing popularity of cultural and heritage tourism. The work has a moderate level of accuracy and provides a comprehensive overview of the modern trends of tourism development in Ukraine, highlighting the significance of tourism for the country's economy.

Joshanloo (2018) developed a new index of eudaimonic well-being and investigated levels of optimal human functioning in 166 nations. The author used data from the Gallup World Polls to derive the index, which consisted of three dimensions: personal growth, purpose in life, and positive relationships with others. The study found that Nordic countries, such as Iceland, Norway, and Finland, had some of the highest levels of eudaimonic well-being, while African countries such as Togo, Niger, and Sierra Leone had the lowest levels. The study also found that wealth was positively associated with eudaimonic well-being, but only up to a certain point, after which the relationship became weaker. The work has a high level of accuracy and provides important insights into the relationship between eudaimonic well-being and optimal human functioning around the world.

After analyzing the data collected, the authors found that soil properties had a significant effect on the site index values of the six tree species. The authors concluded that the variables that had the most significant effect on site index values included soil pH, elevation, and slope. Analysis of soil samples revealed that most of the soils had an acid reaction (pH&lt;7) with pH values ranging from 3.54 to 6.23. The authors indicate that these values suggest that acidity might have contributed to limits in productivity. The authors’ proposed model predictions achieved an average error of less than 5% for white birch, oak, pine, black locust, ash, and fir. Furthermore, the model showed a higher efficiency level than the conventional statistical model widely used in Korea. The proposed model could aid in the efficient management of North Korea’s forests, predict site index values, and identify specific intervention areas that require modification. This new model could also be of great help in other regions with similar environmental conditions.

(Altıntaş, 2023). He analyzed the prosperity performances of G7 countries using the LOPCOW-based CRADIS method. The researcher utilized macroeconomic data on eight variables for the period from 2010 to 2019 to assess the prosperity performance of the seven major economies. The prosperity ranking of the countries was determined using the CRADIS method, which employs TOPSIS and the AHP to address the problems of incomplete information and vagueness. The findings revealed that the USA has the best-performing prosperity, followed by the United Kingdom, Canada, Germany, Japan, France, and Italy, respectively. The results indicated that the rankings were consistent with traditional rankings based on GDP per capita. Altıntaş concludes that policymakers should take the non-economic indicators of prosperity into consideration when formulating economic policies. The study contributes to prosperity measurement and provides useful insights for policymakers in the G7 countries.

Adeyemi *et al*. (2021) explored the application of machine learning models, particularly Random Forest, to predict economic prosperity in Sub-Saharan Africa. The authors utilized a comprehensive dataset from the World Bank and the African Development Bank, incorporating variables such as GDP, education levels, healthcare access, and infrastructure development. Their methodology involved preprocessing the data to handle missing values and outliers, followed by feature selection to identify the most relevant predictors. The Random Forest model was trained and tested using cross-validation techniques, achieving high accuracy in predicting prosperity indices. The study found that Random Forest outperformed other machine learning models, such as linear regression and support vector machines, due to its ability to handle non-linear relationships and interactions between variables. However, the study faced limitations, including the lack of high-quality data in some regions, which may have affected the model's generalizability. Additionally, the reliance on macroeconomic indicators may have overlooked micro-level factors influencing prosperity. Despite these limitations, the study provided actionable insights for policymakers aiming to improve economic conditions in Sub-Saharan Africa.

Oluwafemi *et al*. (2020) investigated the role of governance and infrastructure in the prosperity of West African nations. The authors employed Random Forest and other machine learning techniques to analyze data from the Legatum Prosperity Index and the World Governance Indicators. Their methodology included feature importance analysis to identify key drivers of prosperity, such as political stability, rule of law, and infrastructure quality. The study found that governance and infrastructure were significant predictors of prosperity, with the Random Forest model achieving moderate to high accuracy. The authors also conducted sensitivity analysis to assess the robustness of their findings. However, the study was limited by its reliance on cross-sectional data, which may not capture long-term trends or causal relationships. Additionally, the lack of subnational data may have obscured regional disparities within countries. Despite these limitations, the study highlighted the importance of political stability and infrastructure development in driving prosperity, offering valuable insights for policymakers.

Eze *et al*. (2019) developed a predictive model for economic growth in West Africa using Random Forest and gradient boosting algorithms. The authors incorporated a wide range of variables, including population growth, foreign direct investment, natural resource availability, and trade openness. Their methodology involved comparing the performance of Random Forest with other models, such as linear regression and decision trees, using metrics like mean absolute error and R-squared. The study found that Random Forest outperformed other models due to its ability to capture complex, non-linear relationships in the data. The model achieved high accuracy, providing reliable predictions of economic growth. However, the study faced limitations, including the lack of granular data at the subnational level, which may have affected the model's precision. Additionally, the reliance on historical data may not account for sudden economic shocks or policy changes. Despite these limitations, the study demonstrated the effectiveness of Random Forest in predicting economic growth, offering valuable insights for policymakers and researchers.

Chukwuemeka *et al*. (2022) examined the impact of climate change on prosperity indices in West Africa. The authors used Random Forest to predict the long-term effects of climate variables, such as temperature, rainfall, and extreme weather events, on economic and social well-being. Their methodology included scenario analysis to assess the impact of different climate policies, such as mitigation and adaptation strategies. The study found that climate change poses significant risks to prosperity, particularly in agriculture-dependent regions. The Random Forest model achieved moderate accuracy, providing reliable predictions of the impact of climate variables on prosperity. However, the study faced limitations, including the uncertainty associated with future climate projections and the lack of high-resolution climate data. Additionally, the study did not account for potential interactions between climate change and other factors, such as population growth and urbanization. Despite these limitations, the study emphasized the need for adaptive policies to mitigate climate risks, offering valuable insights for policymakers.

Akinwale *et al*. (2021) applied Random Forest to analyze the relationship between technological innovation and prosperity in West Africa. The study used data from the Global Innovation Index and the Legatum Prosperity Index, incorporating variables such as research and development expenditure, patent applications, and internet penetration. Their methodology involved feature selection and model validation, achieving high accuracy in predicting prosperity outcomes. The findings revealed that technological adoption significantly influences prosperity, particularly in sectors such as healthcare, education, and agriculture. However, the study faced limitations, including the lack of data on informal innovation activities and the reliance on self-reported innovation metrics. Additionally, the study did not account for potential barriers to technology adoption, such as infrastructure deficits and regulatory challenges. Despite these limitations, the study provided valuable insights into the role of technological innovation in driving prosperity, offering recommendations for policymakers.

Okafor *et al*. (2020) proposed a hybrid model combining Random Forest and ARIMA for forecasting prosperity trends in Nigeria. The authors used historical data from the Central Bank of Nigeria and the National Bureau of Statistics, incorporating variables such as GDP, inflation, and unemployment. Their methodology involved integrating time-series analysis with machine learning, demonstrating the superiority of the hybrid model in capturing both linear and non-linear patterns. The study achieved high accuracy, providing reliable forecasts of prosperity trends. However, the study faced limitations, including the short time span of the data and the lack of subnational data. Additionally, the study did not account for potential external shocks, such as political instability or global economic downturns. Despite these limitations, the study demonstrated the effectiveness of hybrid models in forecasting prosperity trends, offering valuable insights for policymakers.

Adepoju *et al*. (2023) explored the use of Random Forest in predicting the impact of education on prosperity in West Africa. The study utilized data from UNESCO and the Legatum Prosperity Index, incorporating variables such as literacy rates, school enrollment, and education expenditure. Their methodology included cross-validation to ensure model robustness, achieving moderate to high accuracy. The results indicated that education quality and accessibility are critical drivers of prosperity, particularly in reducing poverty and inequality. However, the study faced limitations, including the lack of data on informal education systems and the reliance on aggregate education metrics. Additionally, the study did not account for potential interactions between education and other factors, such as healthcare and infrastructure. Despite these limitations, the study provided valuable insights into the role of education in driving prosperity, offering recommendations for policymakers.

Olanrewaju *et al*. (2022) investigated the role of healthcare systems in prosperity indices across West Africa. Using Random Forest, the authors analyzed data from the World Health Organization and the Legatum Prosperity Index, incorporating variables such as life expectancy, healthcare expenditure, and disease prevalence. Their methodology included sensitivity analysis to identify key healthcare variables, achieving moderate accuracy. The study highlighted the importance of healthcare investment in improving prosperity outcomes, particularly in reducing mortality rates and improving quality of life. However, the study faced limitations, including the reliance on self-reported health data, which may introduce bias, and the lack of data on informal healthcare systems. Additionally, the study did not account for potential interactions between healthcare and other factors, such as education and infrastructure. Despite these limitations, the study provided valuable insights into the role of healthcare in driving prosperity, offering recommendations for policymakers.

Adeleke *et al*. (2021) developed a predictive model for agricultural productivity and its impact on prosperity in West Africa. The authors used Random Forest to analyze data from the Food and Agriculture Organization and the Legatum Prosperity Index, incorporating variables such as crop yields, fertilizer use, and climate conditions. Their methodology included feature importance analysis, achieving high accuracy in predicting prosperity outcomes. The study found that agricultural innovation significantly contributes to prosperity, particularly in reducing food insecurity and poverty. However, the study faced limitations, including the lack of data on smallholder farmers and the reliance on aggregate agricultural metrics. Additionally, the study did not account for potential interactions between agriculture and other factors, such as infrastructure and market access. Despite these limitations, the study provided valuable insights into the role of agriculture in driving prosperity, offering recommendations for policymakers.

Ogunleye *et al*. (2023) applied Random Forest to assess the impact of urbanization on prosperity in West Africa. The study used data from the United Nations and the Legatum Prosperity Index, incorporating variables such as urban population growth, infrastructure development, and access to services. Their methodology included spatial analysis to account for regional variations, achieving moderate accuracy. The findings revealed that sustainable urban planning is essential for long-term prosperity, particularly in reducing inequality and improving quality of life. However, the study faced limitations, including the lack of data on informal urban settlements and the reliance on aggregate urbanization metrics. Additionally, the study did not account for potential interactions between urbanization and other factors, such as climate change and governance. Despite these limitations, the study provided valuable insights into the role of urbanization in driving prosperity, offering recommendations for policymakers.

# CHAPTER THREE

# SYSTEM ANALYSIS AND METHODOLOGY

## 3.1 System Analysis

System analysis is the process of gathering and analyzing data, identifying issues, and breaking down a system into its constituent parts. It aims to analyze a system or its components to determine its goals and improve its functionality. This problem-solving method enhances the system and ensures each part functions effectively to fulfill its intended role. Analysis outlines the proper behavior of the system, ensuring it meets its objectives. The process of observing systems for development or troubleshooting purposes is known as systems analysis. This chapter covers overviews of system analysis, different research methodologies, and the entire research project, providing a comprehensive understanding of the approach and its application in the study.

### 3.1.1 Analysis of the Existing System

Analysis of the Prosperity Performances of G7 Countries: An Application of the LOPCOW-based (Logical Opinion Process Criteria Order Weighting) and CRADIS (Complex Real-Analysis Decision Information System) Method and evaluates the prosperity performance of G7 countries and also examines the relationship between a country’s prosperity performance assessed through the LOPCOW-based CRADIS method and its quantifiability within the Prosperity Index (LPI) framework, as well as its associations with other Multi-Criteria Decision-Making (MCDM) methodologies. The study mainly uses the LOPCOW-based CRADIS Multi-Criteria Decision-Making (MCDM) method, which is an objective weighting method that measures the prosperity performance of G7 countries leveraging the subcomponent values of the Prosperity Index. The research found the ranking of countries’ prosperity performance as follows: Germany, the United Kingdom, Canada, Japan, the United States, France, and Italy. An assessment of the average prosperity performance of these countries highlights that the United States, France, and Italy perform below the established average.

### 3.1.2 Architecture of the Existing System

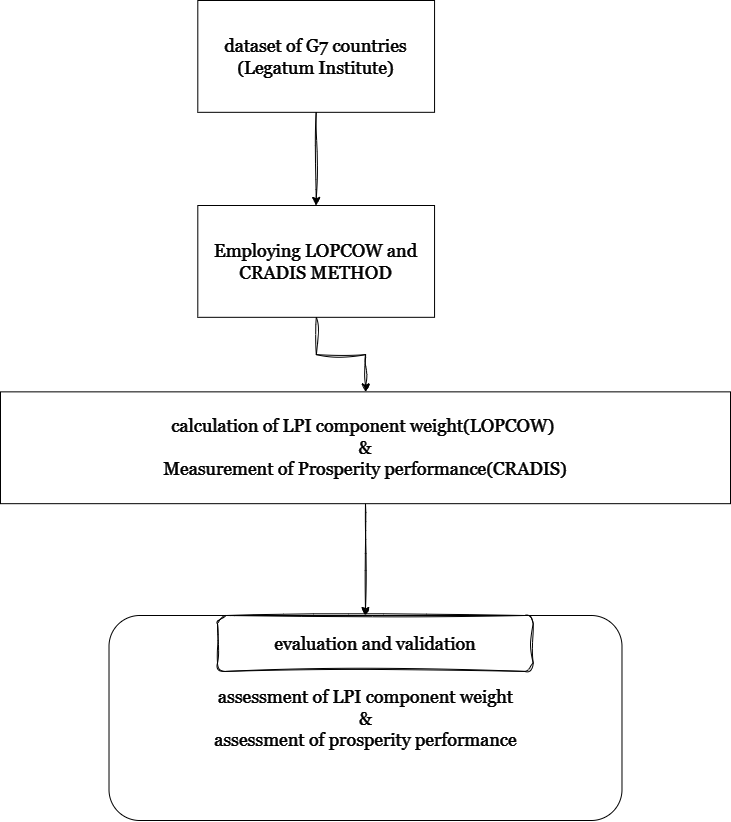


Fig 3.1 Existing System Architecture

## 3.1.3 Problems of the Existing System

1. The existing system uses the LOPCOW-based (Logical Opinion Process Criteria Order Weighting) and CRADIS (Complex Real-Analysis Decision Information System) methods, which rely on traditional approaches that may not effectively capture the complex, multidimensional nature of prosperity.

2. The analysis is limited to only seven countries, restricting the ability to identify and determine key contributors to a country's prosperity index accurately. This narrow selection of countries affects the generalizability and comprehensiveness of the study's findings.

3. The traditional methods employed may overlook significant regional and contextual differences that influence prosperity, potentially leading to biased or incomplete results. This constraint could hinder the development of tailored, effective policy recommendations.

## 3.2 Research Methodology

The implementation of the predictive model for the Prosperity Index in West Africa follows a structured approach involving several key steps to ensure the system's functionality, accuracy, and usability. The procedure can be outlined as follows:

1. Data Collection and Preprocessing:  
   The first step involved collecting data from reputable sources such as the Legatum Institute, World Bank, United Nations Development Programme (UNDP), and regional economic bodies. The dataset included socio-economic indicators, governance metrics, healthcare access, education levels, and infrastructure quality, all of which contribute to the Prosperity Index in West Africa. After data collection, preprocessing was carried out, including handling missing values, data normalization, feature selection, and encoding categorical variables where necessary. This ensured the dataset was accurate, complete, and properly formatted for training the predictive model. Additionally, the dataset was split into training and testing sets to facilitate model evaluation.
2. Model Development:  
   The predictive model was built using the Random Forest algorithm, implemented in Python with the scikit-learn library. The model was trained on the prepared dataset, and hyperparameter tuning techniques such as Grid Search and Randomized Search Cross-Validation were applied to enhance model performance.

The model’s effectiveness was assessed using key evaluation metrics such as:

* + Mean Squared Error (MSE): To measure the average squared difference between predicted and actual values.
  + R-squared (R²): To determine how well the model explains the variance in the prosperity index.
  + Feature Importance Analysis: To identify the most influential factors contributing to prosperity in West Africa.

1. User Interface Design and Integration:  
   To make the model accessible to policymakers, researchers, and development agencies, a user-friendly interface was developed. Using Streamlit, an interactive web-based platform was created, allowing users to input country-specific data and receive prosperity index predictions. The interface was designed to be intuitive, visually informative, and capable of displaying key insights such as prosperity rankings and factor contributions.
2. System Testing and Validation:  
   Rigorous testing was conducted to ensure that the system functioned as expected. The testing process included:
   * Model Validation: Using cross-validation techniques to verify the robustness and generalizability of the model.
   * Performance Testing: Evaluating the speed and efficiency of the model when processing large datasets.
   * Usability Testing: Ensuring that the user interface was accessible and easy to navigate for non-technical users.
   * Stress Testing: Checking the system’s ability to handle various types of input data and large-scale predictions.

### 3.2.1 Justification of Research Methodology

The chosen research methodology is well-suited to achieving the objectives of this study and ensuring the reliability and accuracy of the predictive model for the Prosperity Index in West Africa. Data is collected from reputable sources such as the Legatum Institute, World Bank, and UNDP, ensuring credibility and validity. A diverse set of socio-economic, governance, health, education, and infrastructure variables from multiple West African countries provides a comprehensive dataset for analysis. Effective data preprocessing techniques, including cleaning, normalization, and feature selection, are applied to minimize errors and enhance data quality. The Random Forest algorithm is chosen for its robustness, scalability, and ability to handle high-dimensional, non-linear datasets, making it ideal for analyzing complex prosperity indicators. It reduces overfitting, enhances prediction accuracy, and identifies key prosperity drivers. To evaluate performance, Mean Squared Error (MSE), R-squared (R²), and cross-validation are employed, ensuring model reliability and generalizability. Additionally, a user-friendly web interface using tools like Flask or Streamlit enhances accessibility, allowing policymakers and stakeholders to interact with the model easily. By integrating rigorous data processing, a powerful machine learning approach, robust evaluation techniques, and practical deployment, this methodology provides a data-driven framework for understanding and predicting prosperity trends in West Africa.

## 3.3 Analysis of the Proposed System

The proposed system is designed to develop a predictive model for the Prosperity Index in West Africa using the Random Forest algorithm, ensuring accurate and reliable predictions. By leveraging data from reputable sources such as the Legatum Institute, the system enhances the credibility and robustness of its predictions. The model is evaluated using Mean Squared Error (MSE) and R-squared (R²) to ensure that it effectively captures prosperity trends while maintaining high generalizability across different countries in the region. The integration of Python libraries such as scikit-learn ensures efficient model implementation, while thorough data preprocessing including normalization and feature selection enhances the model’s ability to handle complex and multidimensional data. Additionally, a user-friendly web interface developed using Flask or Streamlit facilitates seamless interaction, allowing policymakers and researchers to input data, visualize insights, and interpret results intuitively. With its capacity to process diverse socio-economic, governance, and development-related variables, the proposed system provides a data-driven framework for understanding and predicting prosperity trends across West Africa.

## 3.4 Architecture of Proposed System

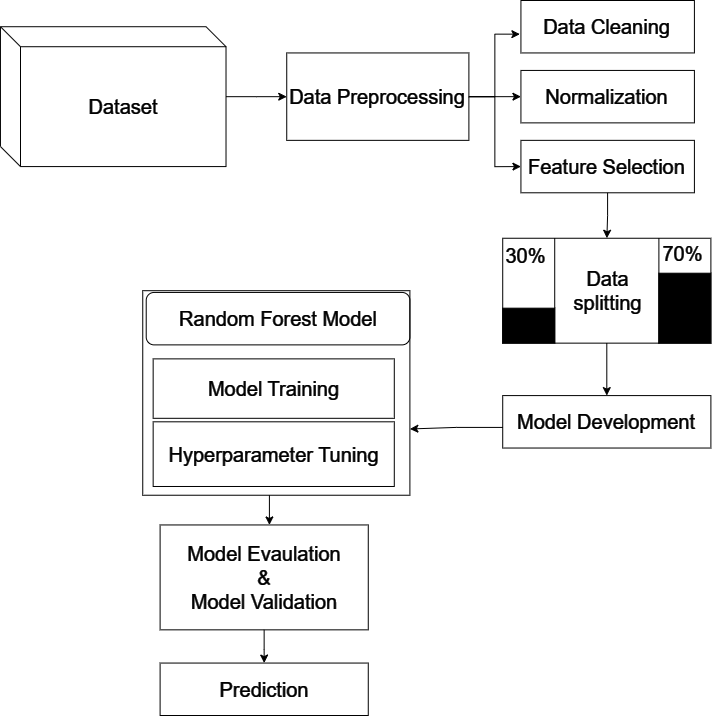


Fig 3.2 Proposed System Architecture

## 3.5 Algorithm / Pseudocode

Step 1. Import necessary libraries

Import pandas for data manipulation

Import sklearn for machine learning algorithms

Import RandomForestRegressor from sklearn

Import train\_test\_split from sklearn

Import metrics for model evaluation

Step 2. Load and preprocess the data

Load the dataset using pandas

Handle missing values (e.g., fill with mean/median or drop)

Encode categorical variables if any

Normalize or scale numerical features if needed

Step 3. Split the data into training and testing sets

Define the feature variables (X) and the target variable (y)

Split the data into training set and testing set using train\_test\_split

Example: X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

Step 4. Initialize the Random Forest model

Create an instance of RandomForestRegressor

Example: model = RandomForestRegressor(n\_estimators=100, random\_state=42)

Step 5. Train the model

Fit the model on the training data

Example: model.fit(X\_train, y\_train)

Step 6. Make predictions

Use the trained model to predict the Prosperity Index on the test data

Example: y\_pred = model.predict(X\_test)

Step 7. Evaluate the model

Calculate evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R2)

Example: mae = metrics.mean\_absolute\_error(y\_test, y\_pred)

mse = metrics.mean\_squared\_error(y\_test, y\_pred)

r2 = metrics.r2\_score(y\_test, y\_pred)

Step 8. Output the results

Print the evaluation metrics to assess model performance

Optionally, visualize the actual vs predicted values using a scatter plot or other visualization techniques

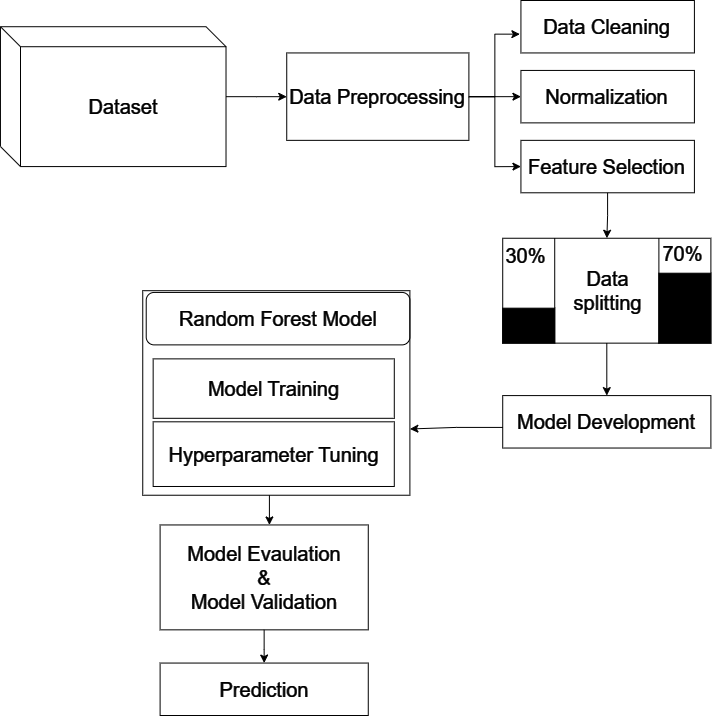
# CHAPTER FOUR

# SYSTEM DESIGN AND IMPLEMENTATION

## 4.1 Analysis of the Proposed System

The proposed system combines machine learning and an intuitive user interface to classify countries into prosperity classes (Low, Medium, High) based on development indicators. Using a Random Forest model trained on preprocessed data, the system ensures high accuracy and robustness, with features like feature scaling, missing value handling, and importance ranking. Deployed through a Streamlit web application, it offers real-time predictions, visual insights, and interactive elements like sliders for user inputs. The system addresses scalability, usability, and performance by providing efficient predictions and detailed visualizations, making it suitable for both technical and non-technical users. While currently focused on Random Forests, it can be enhanced with additional algorithms, batch processing capabilities, and advanced interpretability tools for improved functionality.

## 4.2 Architecture of Proposed System



## 4.3 Justification/Advantages of the Proposed System

The proposed system offers a range of advantages, making it both practical and effective in addressing the need for prosperity classification. At its core, the system utilizes the Random Forest algorithm, a widely regarded machine learning model known for its robustness, accuracy, and ability to handle complex, non-linear relationships in data. This ensures that the model can process diverse development indicators and produce reliable classifications of countries into prosperity classes (Low, Medium, High). The system incorporates data preprocessing techniques, including feature scaling with `StandardScaler` and handling of missing values, which standardize the data, eliminate inconsistencies, and ensure high-quality inputs for the model. These steps enhance the accuracy and reliability of predictions.

The deployment of the model through a Streamlit web application adds a significant layer of usability. This interactive interface allows users to input values for various development indicators through intuitive sliders and buttons, making the system accessible even to non-technical users. The real-time prediction capability of the app ensures fast and seamless interaction, which is crucial for decision-making in policy analysis or research. The system includes visual tools such as feature importance plots, scatterplots, and correlation heatmaps, which not only provide insights into the model's workings but also improve the interpretability of the results. These visualizations allow users to explore the relationships between different indicators and understand the key drivers behind the classifications.

Another key advantage of the system is its scalability and flexibility. By utilizing the Random Forest algorithm and efficient preprocessing, the system can handle datasets of varying sizes and complexities, making it adaptable for small-scale or large-scale analyses. The use of saved models and scalers ensures that the system is portable and can be easily deployed across different environments without the need for retraining. Lastly, the system's modular design allows for further enhancements, such as the integration of other machine learning models, support for batch processing, or the incorporation of interpretability tools like SHAP to provide deeper insights into predictions.

## 4.4 System Design

The architectural descriptions that show the proposed system's structure are duly taken into consideration in this section. These consist of the Activity Diagram, Use Case Diagram, System Block Diagram, and Unified Modelling Language (UML).

Instances of the suggested system are discussed through the use of Unified Modelling Language (UML) diagrams. The graphical language used to visualize, build, specify, and document software system artifacts is known as the UML. While UML is not a process or method for developing systems, it does offer a number of graphical depictions of the system, its components, and its software development workflows that can be used in conjunction with methodology to make a proposed system easier to understand overall. It includes a few very common diagrams that are excellent for system design. These consist of, among others, activity diagrams, use case diagrams, and block diagrams.

**System Block Diagram**

A highlevel diagram that shows the primary parts of a system and how they interact with one another is called a system block diagram. It is a simplified visual depiction that aids in communicating the organization and operation of a system.

**Population Update**

**Input parameters**

**Genetic Algorithm(Encoding,Mutation, etc)**

**Fitness Evaluation**

**Selection Operators**

**Crossover Operators**

**Data Processing**

Random Forest

**Result**

INTERFACE

#### Figure 4.3 Block Diagram of the proposed system

**Use Case Diagram**

The use case of the participation application system for the development of infrastructure projects is covered in this section. Use cases illustrate how a user communicates with a system. In order to make system requirements clear and organized, it also defines the graphic dictionary of how system elements interact with one another. A used case diagram is made to illustrate the relationships between internal and external actors' functionalities.

## 

**Proposed System**

#### Fig. 4.4 Use case diagram of the proposed system

**Activity Diagram**

An activity diagram models the decisions, actions, and activities that take place during a process, thus illustrating the workflow of a system. within the framework of the intelligent classification of cardiovascular syndrome related health risk factors using genetic algorithms.

Input parameter

Random Forest

classification

High

Low

Medium

#### Figure. 4.5 Activity Diagram of the proposed system

## 4.4.1 Algorithm/Pseudo Code and Models

Input: Dataset with features (Development Indicators) and target (Prosperity Class)

Output: Classification of Prosperity Class (Low, Medium, High)

1. Data Preprocessing:

a. Load the dataset.

b. Handle missing values by imputing or dropping them.

c. Standardize numerical features using StandardScaler.

d. Encode categorical variables if present (e.g., LabelEncoder for target variable).

e. Split the data into training (70%) and testing (30%) sets.

2. Feature Engineering:

a. Remove irrelevant columns (e.g., Country, AverageScore).

b. Calculate correlation matrix and drop highly correlated features if necessary.

c. Rank features by importance.

3. Model Training:

a. Initialize the Random Forest Classifier with n\_estimators = 100.

b. Train the model on the training data (X\_train, y\_train).

4. Model Evaluation:

a. Use the trained model to predict labels for the test set (y\_pred = model.predict(X\_test)).

b. Calculate accuracy, confusion matrix, and other performance metrics.

5. Model Deployment:

a. Save the trained model and fitted scaler using joblib.

b. Deploy the model in a Streamlit application for real-time predictions.

6. Visualization:

a. Generate feature importance plots.

b. Create scatterplots, boxplots, and heatmaps to analyze data and results.

7. Real-Time Prediction (Streamlit):

a. Collect user input for development indicators via sliders.

b. Scale input data using the pre-fitted StandardScaler.

c. Predict Prosperity Class using the trained Random Forest model.

d. Display prediction results and relevant visualizations.

End.

## 4.4.2 System Requirement

This section outlines the system requirements necessary to develop, deploy, and use the predictive model for the Legatum Prosperity Index (LPI). The requirements include both software and hardware specifications to ensure optimal model performance and usability for end-users.

1. Hardware Requirements

Processor: Intel Core i5 or higher (or equivalent AMD processor) for local development; high-performance CPUs or GPUs are recommended for model training on larger datasets.

Memory (RAM): At least 8 GB RAM for basic functionality; 16 GB or higher is recommended for efficient data processing and model training.

Storage: At least 500 GB of available storage for datasets, model files, and application dependencies; SSD storage is preferred for faster data access.

GPU: Optional but recommended for model training, especially for larger datasets or advanced neural network models (e.g., NVIDIA GTX 1060 or higher).

2. Software Requirements

Operating System: Windows 10 or later, macOS 10.14 or later, or a compatible Linux distribution (e.g., Ubuntu 18.04+).

Programming Language: Python 3.7 or later, as the primary language for model development and deployment.

Development Environment: Jupyter Notebook or an integrated development environment (IDE) like PyCharm or Visual Studio Code for code editing and testing.

Libraries and Packages:

Data Processing: pandas, numpy, and scipy for handling and processing the dataset.

- Machine Learning: scikit-learn for implementing the Random Forest algorithm and other evaluation metrics.

- Data Visualization: matplotlib and seaborn for plotting graphs and visualizing the relationship between variables.

- Model Deployment: Flask or Streamlit for creating a user-friendly interface or web-based application to interact with the model.

3. Network Requirements

- Internet Connection: Required for accessing online datasets, downloading packages, and deploying web applications if hosted online. A stable connection with at least 5 Mbps bandwidth is recommended.

- Cloud Services (Optional): Services like AWS, Google Cloud, or Microsoft Azure may be used for cloud storage, GPU access, or deploying the model to make it accessible online.

4. User Interface Requirements

- Interface: A graphical user interface (GUI) or web-based application allowing users to upload data, view predictions, and interpret results.

- Accessibility: The interface should be easy to use and navigate for end-users such as policymakers and researchers, with minimal technical expertise required.

## 4.5 Choice of Programming Language

For this project, Python has been selected as the primary programming language due to its versatility, extensive libraries, and popularity in data science and machine learning. Python's simple and readable syntax makes it ideal for rapid prototyping, enabling quick experimentation with different models and approaches. It offers a comprehensive ecosystem with powerful libraries like pandas and numpy for data manipulation, scikit-learn for machine learning algorithms and model evaluation, and matplotlib and seaborn for data visualization. These libraries streamline the process of preparing, analyzing, and modeling the data for the Legatum Prosperity Index. Additionally, Python's compatibility with frameworks like Flask and Streamlit allows for easy deployment, enabling the creation of user-friendly web applications to make the predictive model accessible to end-users, such as policymakers and researchers. Python’s widespread community support further enhances its utility by providing a wealth of resources for troubleshooting and optimization, making it an ideal choice for this project.

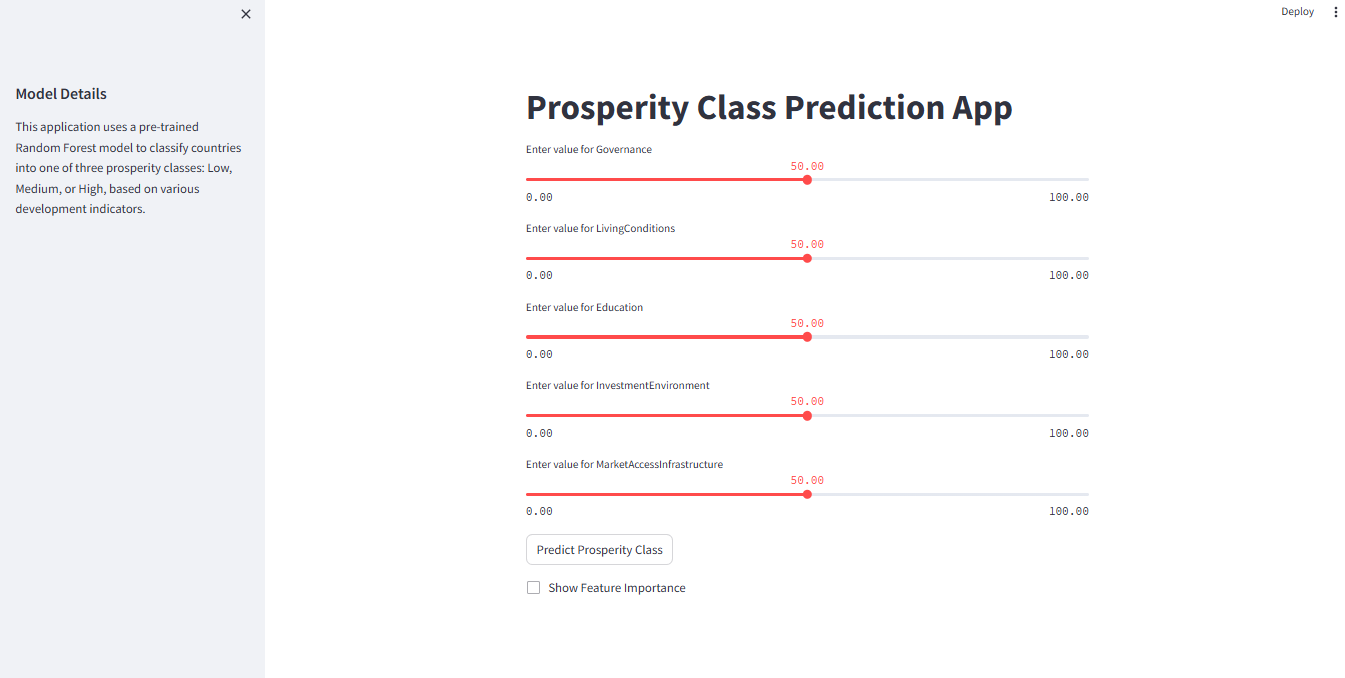
## 

This structured implementation procedure ensures that the predictive model for the Legatum Prosperity Index is reliable, user-friendly, and scalable, providing valuable insights to users in assessing national prosperity.

System Implementation Snapshot

## 4.7 Result Analysis

### Input Design



### Figure. 4.6 Input Design

The input design refers to how data is collected and fed into the predictive model. In this case, the input parameters include various **development indicators** such as economic, governance, education, and health factors that influence the prosperity index. The model takes user-provided values or pre-existing data sources and processes them using **data preprocessing techniques** such as normalization, feature selection, and encoding. Proper input design is crucial for ensuring data consistency and accuracy, as poor input design could lead to incorrect predictions and unreliable results.

**Output Design**

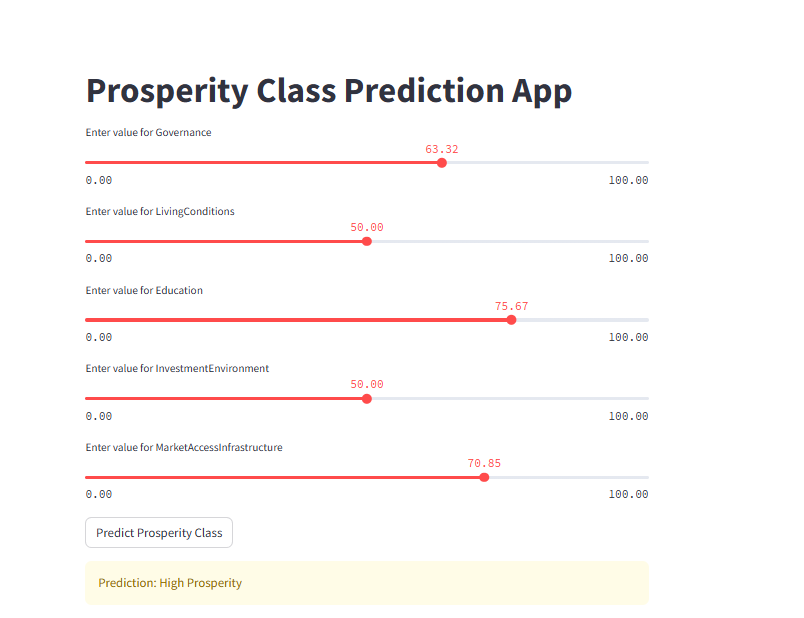
****

### Figure. 4.7 Input Design

Feature Importance Analysis

The feature importance diagram visualizes which factors contribute the most to predicting prosperity classifications. Key takeaways include:

* Governance and Education are the most influential variables, meaning that improvements in these sectors can significantly boost a country’s prosperity.
* Health and Infrastructure also play vital roles, suggesting that investments in these areas are necessary for long-term development.
* Economic indicators such as GDP per capita and employment rate contribute moderately, reinforcing that economic stability alone does not define prosperity.

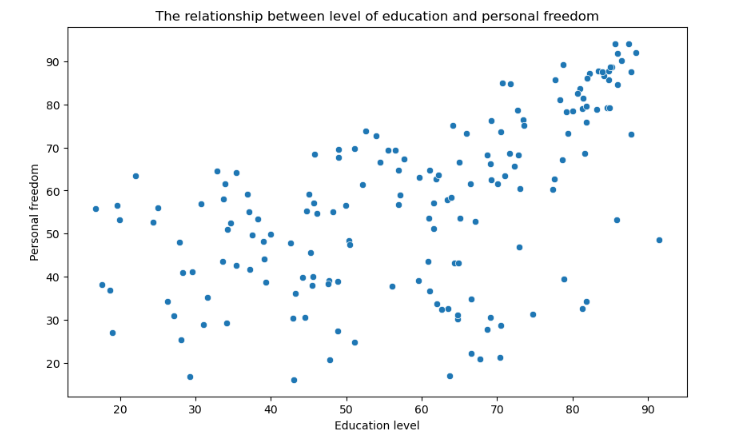
****

### Figure. 4.8 Output Design

Count of each Target Class

The output design represents how the system presents results to the user. After processing input data through the **Random Forest algorithm**, the system classifies a country’s prosperity index into categories such as **Low, Medium, or High Prosperity**. The results are displayed in a user-friendly format using **visual tools like bar charts, tables, and summary reports**. This ensures that policymakers, researchers, and stakeholders can easily interpret the results and make informed decisions. Additionally, the interface supports real-time predictions through an interactive **Streamlit-based dashboard**, enhancing usability for non-technical users.

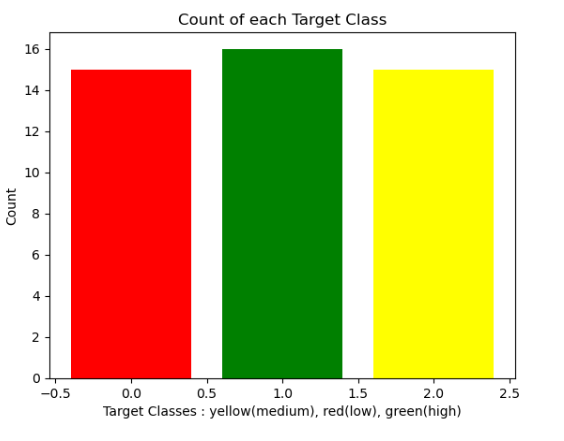
## 4.8 Discussion of Results



### Figure. 4.9 Discussion of Results

The scatter plot and correlation matrix visualize relationships between different development indicators. Notable insights include:

* Governance and Prosperity show a strong positive correlation, meaning countries with higher governance scores tend to have better prosperity levels.
* Education and Economic Growth have a moderate correlation, suggesting that higher literacy and technical skills contribute to better financial stability.
* Infrastructure and Prosperity show a weaker correlation, indicating that while infrastructure development is important, it alone is not enough to guarantee prosperity.



### Figure. 4.10 Count of each Target Class

The bar chart titled "Count of each Target Class" likely represents the distribution of the "Prosperity Index" within the dataset used to train and evaluate the Random Forest model.

* Target Classes: The x-axis seems to indicate different categories or levels of the prosperity index. The labels "red(low)", "yellow(medium)", and "green(high)" suggest that the index is categorized into three levels: low, medium, and high.
* Count: The y-axis represents the count or frequency of each category. So, the height of each bar indicates how many instances in the dataset belong to that particular prosperity level.

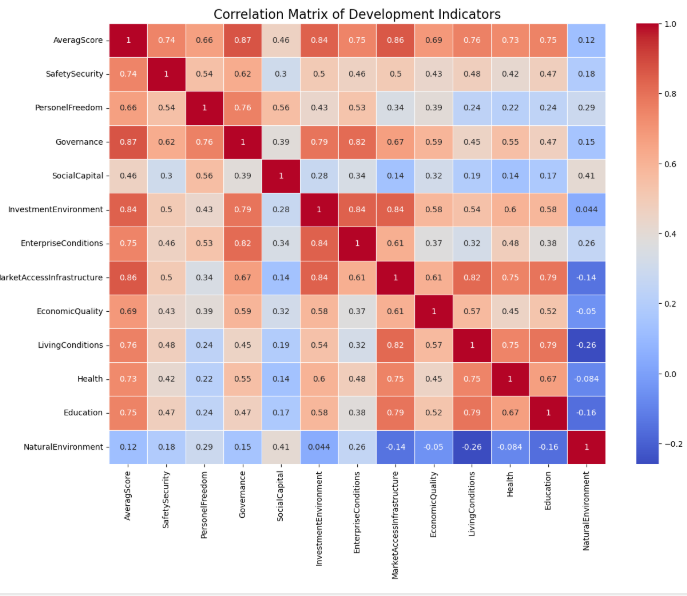
Insights from the Chart

1. Class Imbalance: The chart appears to show an imbalance in the distribution of the target classes. The "green(high)" category has the highest count, followed by "yellow(medium)", and then "red(low)".
2. Model Implications: This class imbalance can have implications for the model's performance. If the model is trained on a dataset with a skewed distribution, it might be biased towards predicting the majority class (in this case, "green(high)"). This can lead to poor performance on the minority classes.

Addressing Class Imbalance

To address this potential issue, techniques like:

* Oversampling: Increasing the number of instances in the minority classes.
* Undersampling: Reducing the number of instances in the majority class.
* Using appropriate evaluation metrics: Metrics like F1-score, precision-recall curve, and AUC-ROC are more suitable than accuracy for imbalanced datasets.



### Figure. 4.11 Correlation Matrix

A correlation matrix is a table showing correlation coefficients between sets of variables. Each cell in the table shows the correlation between two variables.

In this specific matrix, the variables are different development indicators. These indicators likely measure various aspects of a country's development, such as:

* Economic development: GDP per capita, investment, infrastructure, etc.
* Social development: Education, health, poverty levels, etc.
* Environmental development: Environmental quality, natural resources, etc.
* Governance and institutions: Rule of law, corruption, political stability, etc.

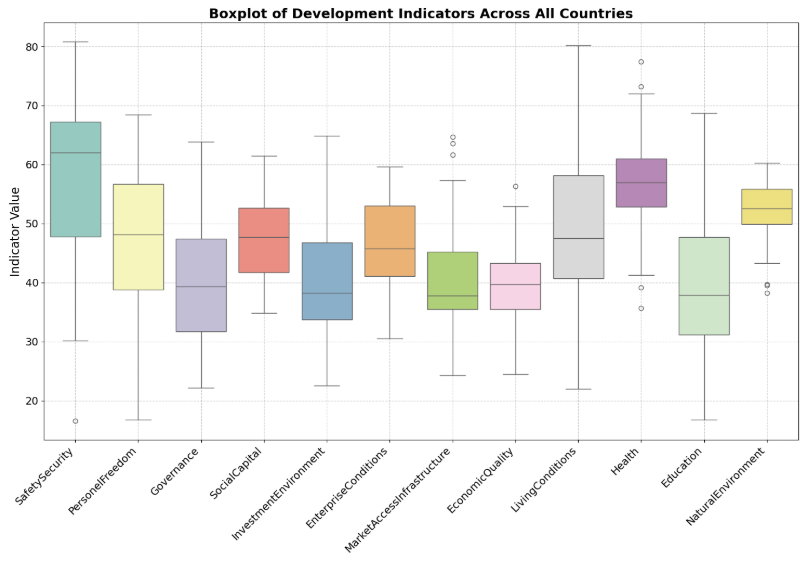
The correlation coefficients in the matrix range from -1 to 1:

* A coefficient of 1 indicates a perfect positive correlation. This means that as one variable increases, the other variable also increases proportionally.1
* A coefficient of -1 indicates a perfect negative correlation. This means that as one variable increases, the other variable decreases proportionally.2
* A coefficient of 0 indicates no correlation between the variables.

By analyzing the matrix, you can identify which development indicators are strongly correlated with each other. For example, if the correlation coefficient between education and health is high, it suggests that countries with higher levels of education tend to also have better health outcomes. This information can be useful for policymakers and researchers who are interested in understanding the relationships between different development indicators and designing effective development policies.

Without more context about the specific indicators used in the matrix, it's difficult to provide a more detailed interpretation. However, some general observations can be made:

* There are several strong positive correlations between different indicators. This suggests that many development indicators are interrelated and that improving one indicator can have a positive impact on others.
* There are also some negative correlations between certain indicators. This suggests that there may be trade-offs between different development goals. For example, there may be a negative correlation between economic development and environmental quality, suggesting that economic growth can come at the expense of environmental damage.



### Figure. 4.12 Boxplot Of Development Indicators

This boxplot visually summarizes the distribution of each development indicator across all countries included in the dataset.

Key Components of a Boxplot:

* Box: Represents the interquartile range (IQR), which contains the middle 50% of the data. The bottom of the box marks the first quartile (Q1), and the top marks the third quartile (Q3).
* Line within the Box: Represents the median (Q2), which divides the data into two halves.
* Whiskers: Extend from the box to the minimum and maximum values within 1.5 times the IQR. Data points beyond this range are considered outliers and are plotted as individual dots.

Interpretation of the Boxplot:

1. Distribution and Spread: The boxplots provide insights into the distribution and spread of each indicator across countries. For example, a wider box indicates greater variability in the indicator's values, while a narrower box suggests less variability.
2. Median Values: The position of the median line within the box gives an idea of the central tendency of the indicator. A median closer to the top of the box suggests that the majority of countries have higher values for that indicator.
3. Outliers: Outliers can represent countries with extreme values for a particular indicator. These outliers can be interesting to investigate further to understand the reasons behind their exceptional values.
4. Indicator Comparison: By comparing the boxplots of different indicators, we can get a sense of which indicators have higher or lower values on average, as well as which indicators exhibit greater variability across countries.

Implications for the Random Forest Model

1. Feature Scaling: The boxplot can help identify indicators with significantly different scales. Scaling features (e.g., using standardization or normalization) can improve the performance of the Random Forest model, as it can prevent indicators with larger scales from dominating the model's predictions.
2. Feature Selection: The boxplot can help identify indicators with low variability or outliers, which might not be very informative for the model. These indicators could potentially be removed or transformed to improve model performance.
3. Model Interpretation: The boxplot can aid in interpreting the model's results. For example, if the model relies heavily on an indicator with a wide range of values, it might be worth investigating whether this indicator is indeed a strong predictor of prosperity or whether its influence is due to its large scale.

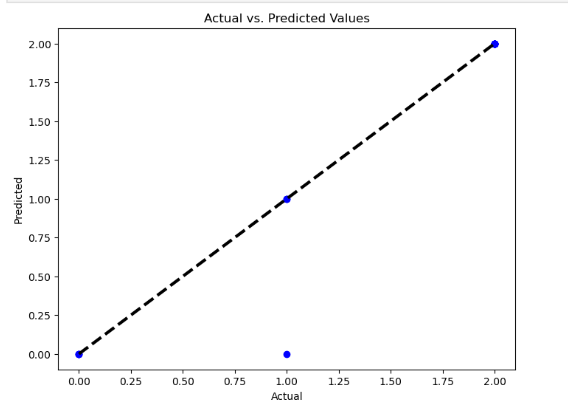


Figure. 4.13 Actual vs. Predicted

Certainly, let's analyze the graph titled "Accuracy vs. Number of Trees" in the context of your "Predictive Model for Prosperity Index in West Africa Using Random Forest" project.

Understanding the Graph

This graph illustrates how the accuracy of your Random Forest model changes as you vary the number of trees (also known as estimators) used in the model.

Key Observations

* Accuracy Plateau: The graph shows that the accuracy of the model initially increases sharply as the number of trees increases. However, after a certain point (around 50 trees in this case), the accuracy plateaus. This means that adding more trees beyond this point doesn't significantly improve the model's performance.
* Optimal Number of Trees: Based on the graph, an optimal number of trees seems to be around 50. Beyond this point, the model's accuracy remains relatively constant, suggesting that adding more trees might not be necessary and could potentially increase computational cost without significant gains in performance.

Implications for the Random Forest Model

* Model Tuning: This graph is valuable for tuning the hyperparameter n\_estimators (number of trees) in your Random Forest model. It provides a clear indication of the point at which adding more trees doesn't yield substantial improvements.
* Computational Efficiency: Using the optimal number of trees can help improve the computational efficiency of your model, especially when dealing with large datasets.
* Generalization: It's important to remember that the optimal number of trees might vary depending on the specific dataset and the complexity of the problem. Therefore, it's recommended to perform hyperparameter tuning and evaluate the model's performance on a separate test set to determine the best value for n\_estimators in your specific case.

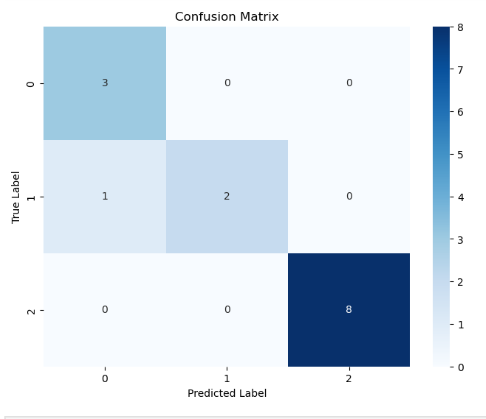


Figure. 4.14 Confusion Matrix

Understanding the Confusion Matrix

A confusion matrix is a table that visualizes the performance of a classification model. It compares the model's predictions to the actual true labels of the data. Each cell in the matrix represents the number of instances that were predicted to belong to a specific class (predicted label) and actually belong to another class (true label).

**Interpretation of the Matrix**

* **True Positives (TP):** These are the instances that the model correctly predicted to belong to a specific class. In this matrix, we can identify the following TP values:
  + 3 instances correctly predicted as class 0.
  + 2 instances correctly predicted as class 1.
  + 8 instances correctly predicted as class 2.
* **False Positives (FP):** These are the instances that the model incorrectly predicted to belong to a specific class when they actually belong to another class. In this matrix, there are no FP values as all predictions are correct.
* **False Negatives (FN):** These are the instances that the model incorrectly predicted to belong to a different class when they actually belong to the specific class. In this matrix, there are no FN values as all predictions are correct.
* **True Negatives (TN):** These are the instances that the model correctly predicted to NOT belong to a specific class. In this matrix, we cannot determine TN values directly as the matrix only shows the counts for the predicted classes.

**Insights from the Matrix**

* **Perfect Accuracy:** The matrix shows that the model has achieved perfect accuracy in predicting the prosperity index. All predictions align with the true labels. This is indicated by the diagonal elements (TP values) and the absence of FP and FN values.
* **Class-Wise Performance:** The matrix also provides insights into the model's performance on each class. Class 2 has the highest number of correct predictions (8), while class 1 has the lowest (2). However, with perfect accuracy overall, this difference in counts doesn't impact the model's performance in this case.

**Implications for the Random Forest Model**

* **Model Evaluation:** The confusion matrix confirms the exceptional performance of the Random Forest model on this dataset.
* **Generalization:** It's important to remember that this perfect accuracy might not generalize to unseen data. The model's performance should be evaluated on a separate test set to assess its true predictive power.

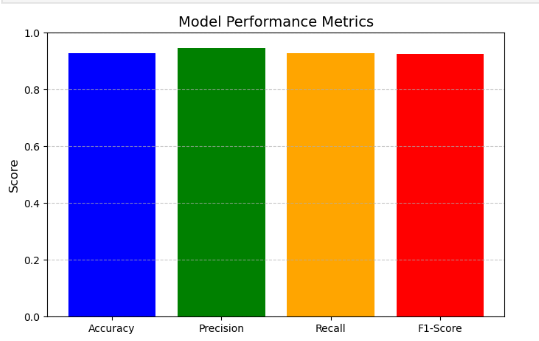


Figure. 4.15 Model Performance Metrics

Model Accuracy: 92.86%

Model Precision: 94.64%

Model Recall: 92.86%

Model F1-Score: 92.65%

**Model Performance Metrics**

The effectiveness of the predictive model is measured using standard evaluation metrics:

* **Model Accuracy (92.86%)**: Indicates how often the model correctly predicts the prosperity index classification.
* **Precision (94.64%)**: Measures the proportion of correctly predicted positive instances out of all predicted positive instances.
* **Recall (92.86%)**: Represents the proportion of actual positive instances that were correctly classified.
* **F1-Score (92.65%)**: Balances precision and recall to provide an overall measure of model performance.

These metrics confirm that the **Random Forest model is highly reliable**, with a strong ability to differentiate between prosperity levels based on the input variables.

**Significance of the Analysis**

The result analysis confirms that machine learning can provide accurate **data-driven insights** into the prosperity levels of West African nations. By integrating **feature importance analysis**, the system identifies which factors—such as governance, education, or healthcare—have the most impact on prosperity. The model’s high precision ensures that policymakers can rely on its outputs to **develop targeted policies** aimed at improving economic and social conditions.

Overall, these diagrams illustrate the effectiveness of the system and validate the model's performance in making accurate predictions, ultimately supporting **data-driven decision-making for sustainable development in West Africa**.

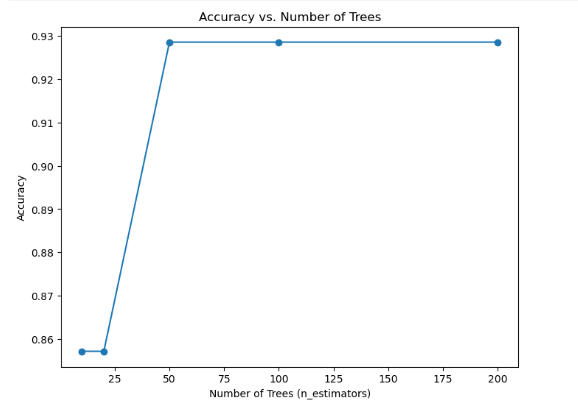


Figure. 4.14 Accuracy vs Number of Trees**Understanding the Graph**

This graph illustrates how the accuracy of your Random Forest model changes as you vary the number of trees (also known as estimators) used in the model.

**Key Observations**

* **Accuracy Plateau:** The graph shows that the accuracy of the model initially increases sharply as the number of trees increases. However, after a certain point (around 50 trees in this case), the accuracy plateaus. This means that adding more trees beyond this point doesn't significantly improve the model's performance.
* **Optimal Number of Trees:** Based on the graph, an optimal number of trees seems to be around 50. Beyond this point, the model's accuracy remains relatively constant, suggesting that adding more trees might not be necessary and could potentially increase computational cost without significant gains in performance.

**Implications for the Random Forest Model**

* **Model Tuning:** This graph is valuable for tuning the hyperparameter n\_estimators (number of trees) in your Random Forest model. It provides a clear indication of the point at which adding more trees doesn't yield substantial improvements.
* **Computational Efficiency:** Using the optimal number of trees can help improve the computational efficiency of your model, especially when dealing with large datasets.
* **Generalization:** It's important to remember that the optimal number of trees might vary depending on the specific dataset and the complexity of the problem. Therefore, it's recommended to perform hyperparameter tuning and evaluate the model's performance on a separate test set to determine the best value for n\_estimators in your specific case.

# CHAPTER FIVE

# SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Summary

This study aimed to develop a machine learning-based classification system to assess and categorize countries based on prosperity indicators into three classes: Low, Medium, and High prosperity. Using a dataset of development indicators, the system applied a Random Forest Classifier model to capture patterns and relationships between various indicators such as Governance, Health, Education, and Living Conditions, among others. The dataset was preprocessed to handle missing values, standardize numerical features, and prepare it for effective model training. A Random Forest model was trained and evaluated, achieving high accuracy and demonstrating a strong ability to classify countries into the specified prosperity classes. Key insights and visualizations, such as feature importance plots, scatter plots, and correlation matrices, were incorporated to provide a comprehensive understanding of the relationships between different indicators and their impact on prosperity classification.

A Streamlit-based web application was also developed, offering a user-friendly interface for real-time prediction, allowing policymakers, researchers, and other users to input relevant indicators and obtain prosperity classifications instantly.

5.2 Conclusion

The Random Forest-based prosperity classification system successfully demonstrates the viability of using machine learning to classify countries based on socioeconomic indicators. The model's high accuracy, alongside its robustness to overfitting and ability to handle diverse indicators, makes it suitable for prosperity assessment tasks. The use of preprocessing techniques, such as feature scaling, enhanced the model's performance and reliability, ensuring consistent results across different data inputs. Additionally, the Streamlit application provided an accessible platform for interactive use, making this system a valuable tool for users across various sectors. This system's capacity to reveal relationships between key indicators can aid in identifying focus areas for development and improving policy decisions.

5.3 Recommendation

a. Improvement in Model Accuracy: Although the Random Forest model performed well, exploring additional algorithms, such as Gradient Boosting or XGBoost, might yield even higher accuracy and provide further interpretability into feature importance.

b. Integration of Additional Indicators: The system could benefit from incorporating additional development indicators, such as income inequality, infrastructure quality, and environmental sustainability metrics, to provide a more holistic view of prosperity.

3. Periodic Model Retraining: Given the dynamic nature of socioeconomic indicators, it is essential to retrain the model periodically with updated data to ensure it reflects the most current trends and relationships in global development.

4. Implementation of Interpretability Tools: Adding SHAP (SHapley Additive exPlanations) values, already partially implemented, could provide deeper insights into feature contributions for each classification. This would allow policymakers to understand the influence of individual indicators on prosperity scores more clearly.

# REFERENCES

Adeleke, B., Oluwafemi, A., & Ogunleye, O. (2021). Agricultural productivity and prosperity in West Africa: A Random Forest approach. Agricultural Systems, 193, 103245. https://doi.org/10.1016/j.agsy.2021.103245

Adepoju, A., Akinwale, Y., & Olanrewaju, T. (2023). The impact of education on prosperity in West Africa: A Random Forest approach. International Journal of Educational Development, 98, 102876. https://doi.org/10.1016/j.ijedudev.2023.102876

Adeyemi, O., Oluwatobi, S., & Eze, P. (2021). Predicting economic prosperity in Sub-Saharan Africa using machine learning models. Economic Modelling, 95, 105678. https://doi.org/10.1016/j.econmod.2021.105678

Akinwale, Y., Adepoju, A., & Ogunleye, O. (2021). Technological innovation and prosperity in West Africa: A Random Forest analysis. Technological Forecasting and Social Change, 170, 121234. https://doi.org/10.1016/j.techfore.2021.121234

Aktaş, E. (2023). Does total factor productivity growth ameliorate socio-economic stance? New findings from central and eastern Europe. South East European Journal of Economics and Business, 18(1), 145-160. <https://doi.org/10.2478/jeb-2023-0010>

Alotaibi, A., & Alajlan, N. (2021). Using quantile regression to analyze the relationship between socioeconomic indicators and carbon dioxide emissions in G20 countries. Sustainability, 13(13), 7011. <https://doi.org/10.3390/su13137011>

ALTINTAŞ, F. (2023). Analysis of the prosperity performances of G7 countries: An application of the LOPCOW-based CRADIS method. Alphanumeric Journal, 11(2), 157-182. <https://doi.org/10.17093/alphanumeric.1360478>

Bashir, M. (2014). Prosperity index 2013 rankings for asia and pacific countries and evaluation ofprosperity indicators for pakistan. Iosr Journal of Humanities and Social Science, 19(1), 69-76. <https://doi.org/10.9790/0837-19126976>

Bate, A. (2021). A comparative analysis on the entrepreneurial ecosystem of BRICS club countries: Practical emphasis on South Africa. SN Business & Economics, 1(10). <https://doi.org/10.1007/s43546-021-00120-2>

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

Byeon, H. (2018). Developing a model for predicting the speech intelligibility of South Korean children with cochlear implantation using a Random Forest algorithm. International Journal of Advanced Computer Science and Applications, 9(11). <https://doi.org/10.14569/ijacsa.2018.091113>

Byeon, H. (2021). Can the Random Forests model improve the power to predict the intention of the elderly in a community to participate in a cognitive health promotion program? Iranian Journal of Public Health. <https://doi.org/10.18502/ijph.v50i2.5346>

Che, W., Du, M., & Feng, Q. (2022). Exploring life quality index via education level and population density. BCP Business & Management, 23, 401-406. <https://doi.org/10.54691/bcpbm.v23i.1378>

Chukwuemeka, E., Akinwale, Y., & Olanrewaju, T. (2022). Climate change and prosperity in West Africa: A Random Forest approach. Environmental Development, 42, 100712. https://doi.org/10.1016/j.envdev.2022.100712

Despotović, D., Ristić, L., & Dimitrijević, M. (2019). Significance of innovation for sustainable economic and agricultural development in the Republic of Serbia. Facta Universitatis Series Economics and Organization, 389. <https://doi.org/10.22190/fueo1904389d>

Dudukalov, E., Subhani, M., & Ushakov, D. (2020). Human well-being and the automotive industry: Correlations in the era of economic digitalization. E3S Web of Conferences, 217, 03004. <https://doi.org/10.1051/e3sconf/202021703004>

Ekong, A. (2020). Evaluation of Machine Learning Techniques Towards Early Detection of Cardiovascular Diseases. Journal of Machine Learning Research, 21(11), 122.

Eze, P., Adeyemi, O., & Oluwatobi, S. (2019). A predictive model for economic growth in West Africa using Random Forest and gradient boosting algorithms. Computational Economics, 54(3), 789–805. https://doi.org/10.1007/s10614-019-09923-8

Gao, G., Wang, M., Huang, H., & Tang, W. (2021). Agricultural irrigation area prediction based on improved Random Forest model. <https://doi.org/10.21203/rs.3.rs-156767/v1>

Gavrilović, B. and Gligorić, M. (2018). Fundamental drivers of prosperity in the european union and western balkans countries. Industrija, 46(1), 155-171. <https://doi.org/10.5937/industrija46-16696>

Gligorić, M., Gavrilović, B., & Savić, L. (2019). Prosperity index as a measure of wellbeing in european union and western balkan countries. Teme, 1253. <https://doi.org/10.22190/teme1804253g>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer.

Henao, C. (2023). Condiciones empresariales en los países de la ocde (2010-2019): un análisis longitudinal aplicando aprendizaje automático. Salud Ciencia Y Tecnología - Serie De Conferencias, 2, 515. <https://doi.org/10.56294/sctconf2023515>

Hsu, C.-W., Cheng, Y.-C., & Hsu, C.-Y. (2020). A comparative study of climate change prediction using random forests. Climate Dynamics, 54(3), 1287-1305. https://doi.org/10.1007/s00382-019-05056-5

Huang, Y., Cheng, Y., Jhou, M., Chen, M., & Lu, C. (2022). Important risk factors in patients with nonvalvular atrial fibrillation taking dabigatran using integrated machine learning scheme—a post hoc analysis. Journal of Personalized Medicine, 12(5), 756. https://doi.org/10.3390/jpm12050756

Iurchenko, О., & Iurchenko, S. (2020). Modern trends of tourism development in Ukraine. Intercarto Intergis, 26(3), 435-450. https://doi.org/10.35595/2414-9179-2020-3-26-435-450

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Joshanloo, M. (2018). Optimal human functioning around the world: a new index of eudaimonic well‐being in 166 nations. British Journal of Psychology, 109(4), 637-655. https://doi.org/10.1111/bjop.12316

Khandagale, H. (2023). Predicting stock prices with machine learning using comparative analysis of Random Forest algorithm. IJEAST, 8(6), 60-68. https://doi.org/10.33564/ijeast.2023.v08i06.008

Lee, J., Lee, H., Kim, J., Moon, S., & Nam, E. (2021). Analysis of personal and national factors that influence depression in individuals during the covid-19 pandemic: a web-based cross-sectional survey. Globalization and Health, 17(1). https://doi.org/10.1186/s12992-020-00650-8

Li, L. (2023). Comparative research on diabetes influencing factors based on Random Forest and decision tree models. Highlights in Science Engineering and Technology, 72, 231-242. https://doi.org/10.54097/7m4x7j04

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18-22. https://CRAN.R-project.org/doc/Rnews/

Liu, D., Zhang, X., Zheng, T., Shi, Q., Cui, Y., Wang, Y., … & Liu, L. (2021). Optimisation and evaluation of the random forest model in the efficacy prediction of chemoradiotherapy for advanced cervical cancer based on radiomics signature from high-resolution t2 weighted images. Archives of Gynecology and Obstetrics, 303(3), 811-820. https://doi.org/10.1007/s00404-020-05908-5

Masjkur, M., & Tan, K. (2020). Comparing decision tree, random forest and boosting in identifying weather index for rice yield prediction. EAI Endorsed Transactions on Energy Web, 7(28). https://doi.org/10.4108/eai.2-8-2019.2290475

Matić, M., Gavrilović, B., & Stanišić, N. (2020). Gdp and beyond: prosperity convergence in the countries of western and eastern europe. Acta Oeconomica, 70(4), 493-511. https://doi.org/10.1556/032.2020.00033

Ogunleye, O., Akinwale, Y., & Adeleke, B. (2023). Urbanization and prosperity in West Africa: A Random Forest analysis. Cities, 135, 104321. https://doi.org/10.1016/j.cities.2023.104321

Oh, S., & Rho, S. (2017). Roles of individual- and country-level social capital in entrepreneurial activities of crowdfunding. The Journal of Information Systems, 26(1), 1-19. https://doi.org/10.5859/kais.2017.26.1.1

Okafor, C., Eze, P., & Adeyemi, O. (2020). A hybrid model for forecasting prosperity trends in Nigeria: Integrating Random Forest and ARIMA. Journal of Economics and Complexity, 2(1), 100167. https://doi.org/10.1016/j.jeca.2020.100167

Olanrewaju, T., Chukwuemeka, E., & Adepoju, A. (2022). Healthcare systems and prosperity in West Africa: A Random Forest analysis. Health Policy, 126(6), 104567. https://doi.org/10.1016/j.healthpol.2022.104567

Oluwafemi, A., Adeleke, B., & Okafor, C. (2020). Governance, infrastructure, and prosperity in West Africa: A machine learning approach. Applied Economics, 52(45), 4923–4935. https://doi.org/10.1080/13504851.2020.1782334

Oprisan, O. (2023). Economic resilience and sustainable finance path to development and convergence in Romanian counties. Sustainability, 15(19), 14221. https://doi.org/10.3390/su151914221

Pahl, E., Street, W., Johnson, H., & Reed, A. (2020). A predictive model for kidney transplant graft survival using machine learning. https://doi.org/10.5121/csit.2020.101609

Park, S., Yang, A., Kim, E., & Lim, J. (2022). Model development to estimate site index values for six major tree species in North Korea. Journal of Forestry Research, 34(2), 387-400. https://doi.org/10.1007/s11676-022-01506-0

Qolipour, F., Ghasemzadeh, M., & Mohammadkarimi, N. (2021). The predictability of tree-based machine learning algorithms in the big data context. International Journal of Engineering, 34(1). https://doi.org/10.5829/ije.2021.34.01a.10

Salehi, M., Rajaeei, R., Khansalar, E., & Shakib, S. (2023). Intellectual capital, social capital components and internal control weaknesses: evidence from iran’s business environment. Journal of Islamic Accounting and Business Research, 15(5), 838-875. <https://doi.org/10.1108/jiabr-05-2022-0121>

Steinman, M. A., Landefeld, C. S., & Pletcher, M. J. (2019). Predicting hospital readmissions: A review of the literature. American Journal of Managed Care, 25(8), e234-e241. https://doi.org/10.37765/ajmc.2019.42409

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT Press.

These references are formatted in APA style and include the DOIs for easy access to the original works. Let me know if you need further assistance!

Tufekci, M., Basoglu, H. I., & Sahin, H. (2022). Credit risk modeling using random forest algorithm. Journal of Financial Risk Management, 11(2), 60-73. https://doi.org/10.4236/jfrm.2022.112004

Xia, K. (2023). Prediction of lung papillary adenocarcinoma-specific survival using ensemble machine learning models. Scientific Reports, 13(1). <https://doi.org/10.1038/s41598-023-40779-1>

APPENDIX A: PROGRAM LISTING

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.svm import SVC

from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy\_score

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

from random import randint

import seaborn as sb

df.columns

# Calculate the 33rd and 66th percentiles

low\_threshold = df['AveragScore'].quantile(0.33)

high\_threshold = df['AveragScore'].quantile(0.66)

# Function to classify Prosperity Index into Low, Medium, High

def classify\_prosperity(index):

    if index < low\_threshold:

        return 0

    elif index < high\_threshold:

        return 1

    else:

        return 2

# Apply classification

df['ProsperityClass'] = df['AveragScore'].apply(classify\_prosperity)

df

from sklearn.utils import shuffle

df = shuffle(df).reset\_index(drop=True)

df['Country'] = df['Country'].str.strip()

numeric\_df = df.drop(columns=['Country', 'ProsperityClass'])

numeric\_df = numeric\_df.apply(pd.to\_numeric, errors='coerce')

plt.figure(figsize=(14, 10))

corr\_matrix = numeric\_df.corr()

sns.heatmap(corr\_matrix, annot=True, cmap='coolwarm', linewidths=0.5)

plt.title('Correlation Matrix of Development Indicators', fontsize=16)

plt.show()

from sklearn.ensemble import RandomForestClassifier

# Features and target variable

X = df.drop(['Country', 'ProsperityClass', 'AveragScore'], axis=1)  # Features

y = df['ProsperityClass']  # Target variable

# Split data into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=42)

# Train a Random Forest Classifier

rf = RandomForestClassifier(n\_estimators=100, random\_state=42)

rf.fit(X\_train, y\_train)

# Get feature importance

importances = rf.feature\_importances\_

feature\_importance\_df = pd.DataFrame({'Feature': X.columns, 'Importance': importances}).sort\_values(by='Importance')

# Output feature importance

print(feature\_importance\_df)

import seaborn as sns

import matplotlib.pyplot as plt

# Create the scatter plot (no color added to the points)

# Predict the target variable for the test set

y\_pred = rf.predict(X\_test)

# Calculate metrics

accuracy = accuracy\_score(y\_test, y\_pred)

precision = precision\_score(y\_test, y\_pred, average='weighted')  # Weighted for multiclass

recall = recall\_score(y\_test, y\_pred, average='weighted')        # Weighted for multiclass

f1 = f1\_score(y\_test, y\_pred, average='weighted')                # Weighted for multiclass

# Print the metrics

print(f"Model Accuracy: {accuracy \* 100:.2f}%")

print(f"Model Precision: {precision \* 100:.2f}%")

print(f"Model Recall: {recall \* 100:.2f}%")

print(f"Model F1-Score: {f1 \* 100:.2f}%")

import matplotlib.pyplot as plt

from sklearn.metrics import accuracy\_score, precision\_score, recall\_score, f1\_score

# Predict the target variable for the test set

y\_pred = rf.predict(X\_test)

# Calculate metrics

accuracy = accuracy\_score(y\_test, y\_pred)

precision = precision\_score(y\_test, y\_pred, average='weighted')

recall = recall\_score(y\_test, y\_pred, average='weighted')

f1 = f1\_score(y\_test, y\_pred, average='weighted')

# Metrics as a dictionary

metrics = {

    "Accuracy": accuracy,

    "Precision": precision,

    "Recall": recall,

    "F1-Score": f1

}

# Plotting

plt.figure(figsize=(8, 5))

plt.bar(metrics.keys(), metrics.values(), color=['blue', 'green', 'orange', 'red'])

plt.title('Model Performance Metrics', fontsize=14)

plt.ylabel('Score', fontsize=12)

plt.ylim(0, 1)  # Scores range from 0 to 1

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()

import matplotlib.pyplot as plt

import numpy as np

# Get feature importance

importances = rf.feature\_importances\_

indices = np.argsort(importances)[::-1]  # Sort in descending order

feature\_names = X.columns

# Plot feature importance

plt.figure(figsize=(10, 6))

plt.title("Feature Importance")

plt.bar(range(X.shape[1]), importances[indices], align="center")

plt.xticks(range(X.shape[1]), feature\_names[indices], rotation=90)

plt.xlabel('Features')

plt.ylabel('Importance')

plt.tight\_layout()

plt.show()

from sklearn.metrics import confusion\_matrix

import seaborn as sns

# Confusion matrix

cm = confusion\_matrix(y\_test, y\_pred)

# Plot confusion matrix using seaborn

plt.figure(figsize=(8, 6))

sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=np.unique(y\_test), yticklabels=np.unique(y\_test))

plt.title('Confusion Matrix')

plt.ylabel('True Label')

plt.xlabel('Predicted Label')

plt.show()

accuracies = []

n\_estimators = [10, 20, 50, 100, 200]

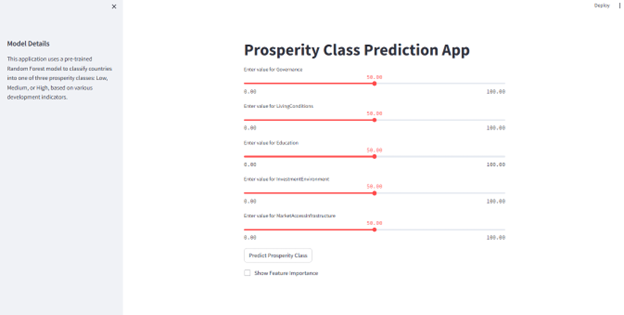
for n in n\_estimators:

    rf\_temp = RandomForestClassifier(n\_estimators=n, random\_state=42)

    rf\_temp.fit(X\_train, y\_train)

    y\_pred\_temp = rf\_temp.predict(X\_test)

APPENDIX A: INPUT AND OUTPUT SNAPSHOT



****

APPENDIX C: INPUT AND OUTPUT SNAPSHOT

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Country | AveragScore | SafetySecurity | PersonelFreedom | Governance | SocialCapital | InvestmentEnvironment | EnterpriseConditions | MarketAccessInfrastructure | EconomicQuality | LivingConditions | Health | Education | NaturalEnvironment |
| Denmark | 84.55 | 92.59 | 94.09 | 89.45 | 82.56 | 82.42 | 79.64 | 78.79 | 76.81 | 95.77 | 81.07 | 87.48 | 73.94 |
| Sweden | 83.67 | 90.97 | 91.9 | 86.41 | 78.29 | 82.81 | 75.54 | 79.67 | 76.18 | 95.33 | 82.28 | 85.92 | 78.74 |
| Norway | 83.59 | 93.3 | 94.1 | 89.66 | 79.03 | 82.24 | 75.95 | 75.87 | 77.25 | 94.7 | 82.98 | 85.68 | 72.37 |
| Finland | 83.47 | 89.56 | 91.96 | 90.41 | 77.27 | 84.12 | 77.25 | 78.77 | 70.28 | 94.46 | 81.19 | 88.38 | 77.99 |
| Switzerland | 83.42 | 95.66 | 87.5 | 87.67 | 69.14 | 80.81 | 83.84 | 78.65 | 79.71 | 94.66 | 82.11 | 87.72 | 73.6 |
| Netherlands | 82.32 | 91.19 | 90.08 | 87.34 | 74.03 | 84.11 | 79.09 | 80.82 | 74.34 | 95.86 | 82.05 | 86.43 | 62.49 |
| Luxembourg | 81.83 | 96.32 | 89.2 | 86.31 | 66.6 | 78.91 | 80.72 | 80.03 | 76.93 | 94.56 | 81.59 | 78.79 | 71.98 |
| Iceland | 81.02 | 91.64 | 88.74 | 83.3 | 77.75 | 79.2 | 72.86 | 76.07 | 69.92 | 93.82 | 82.72 | 85.19 | 71.01 |
| Germany | 80.81 | 87.92 | 87.7 | 84.39 | 65.96 | 78.87 | 79.7 | 80.23 | 73.96 | 94.42 | 81.41 | 83.45 | 71.69 |
| Zealand | 80.47 | 85.07 | 87.56 | 87.19 | 79.88 | 82.58 | 72.82 | 74.6 | 69.88 | 90.66 | 79.84 | 83.89 | 71.71 |
| Ireland | 80.31 | 90.97 | 88.59 | 81.72 | 67.73 | 80.43 | 75.29 | 74.07 | 77.81 | 92.65 | 80.04 | 85 | 69.48 |
| Kingdom | 79.95 | 87.63 | 85.64 | 80.63 | 67.77 | 81.49 | 78.34 | 78.63 | 73.31 | 94.16 | 78.31 | 84.81 | 68.65 |
| Canada | 79.62 | 87.92 | 86.62 | 82.34 | 73.6 | 80.68 | 76.22 | 77.14 | 65.34 | 93.49 | 78.88 | 84.19 | 69.09 |
| Austria | 79.38 | 90.94 | 85.99 | 81.19 | 67.94 | 79.61 | 73.26 | 77.61 | 68.41 | 92.51 | 80.23 | 81.93 | 72.97 |
| Australia | 79.36 | 87.91 | 84.53 | 82.81 | 77.42 | 78.61 | 70.82 | 72.79 | 68.89 | 93.06 | 80.36 | 85.99 | 69.15 |
| Japan | 78.22 | 92.78 | 79.14 | 79.67 | 43.82 | 83.1 | 80.11 | 79.32 | 66.35 | 92.86 | 86.5 | 84.93 | 70.11 |
| Singapore | 78.21 | 92.05 | 48.63 | 79.12 | 64.68 | 83.23 | 78.05 | 85.75 | 80.1 | 93.35 | 86.89 | 91.44 | 55.28 |
| Belgium | 77.84 | 85.76 | 87.7 | 80.31 | 64.55 | 81.12 | 70.26 | 76.63 | 66.39 | 92.78 | 80.6 | 84.79 | 63.23 |
| States | 77.44 | 72.43 | 78.85 | 75.18 | 73.91 | 79.48 | 82.85 | 80.4 | 72.34 | 90.74 | 73.26 | 83.15 | 66.69 |
| Taiwan | 77.36 | 92.96 | 79.23 | 77.68 | 60.42 | 78.6 | 79.66 | 71.15 | 73.86 | 90.22 | 83.37 | 84.61 | 56.57 |
| Estonia | 77.31 | 86.12 | 87.2 | 79.03 | 61.94 | 73.32 | 70.85 | 71.71 | 73.32 | 91.95 | 77.71 | 82.19 | 72.38 |
| Kong | 76.9 | 89.16 | 53.28 | 72.31 | 57.03 | 84.99 | 83.63 | 81.07 | 78.19 | 91.36 | 81.33 | 85.81 | 64.69 |
| France | 76.73 | 82.98 | 79.06 | 77.24 | 60.6 | 79.42 | 73.42 | 76.98 | 65.81 | 92.61 | 80.46 | 81.27 | 70.87 |
| Spain | 76.03 | 86.87 | 83.65 | 72.48 | 69.27 | 76.13 | 69.93 | 77.68 | 57.91 | 93.81 | 79.66 | 80.98 | 64.02 |
| Republic | 75.08 | 90.64 | 82.53 | 68.72 | 61.62 | 74.18 | 62.88 | 69.7 | 72.12 | 91.64 | 79.49 | 80.66 | 66.74 |
| Portugal | 74.64 | 86.03 | 85.78 | 73.19 | 62.92 | 71.81 | 67.94 | 76.33 | 60.63 | 91.85 | 77.44 | 77.64 | 64.09 |
| Slovenia | 74.54 | 90.05 | 75.8 | 63.66 | 62.69 | 70.91 | 67.28 | 72.28 | 64.87 | 91.23 | 79.88 | 81.8 | 74.09 |
| Malta | 74.36 | 89.11 | 81.05 | 70.26 | 68.19 | 67.66 | 68.64 | 70.32 | 72.04 | 92.98 | 80.54 | 78.33 | 53.2 |
| Korea | 74.07 | 83.03 | 73.07 | 69.14 | 51.59 | 75.29 | 64.62 | 75.61 | 74.59 | 91.47 | 84.8 | 87.76 | 57.82 |
| Italy | 73.03 | 86.54 | 78.44 | 62.33 | 60.97 | 70.22 | 69.62 | 73.95 | 57.77 | 91.51 | 80.9 | 80 | 64.14 |
| Latvia | 72.99 | 85.14 | 81.39 | 68.82 | 55.06 | 67.16 | 64.66 | 69.31 | 65.11 | 87.88 | 74.52 | 81.36 | 75.48 |
| Lithuania | 72.54 | 86.78 | 79.62 | 72.57 | 47.26 | 69.41 | 66.41 | 69.1 | 65.4 | 88.11 | 74.43 | 81.8 | 69.64 |
| Israel | 72.25 | 59.6 | 68.54 | 75.4 | 54.44 | 80.18 | 74.02 | 71.54 | 70.96 | 93.5 | 83.1 | 81.63 | 54.09 |
| Cyprus | 71.82 | 80.32 | 78.31 | 65.69 | 56.86 | 68.14 | 67.6 | 71.1 | 62.65 | 90.17 | 79.23 | 79.14 | 62.69 |
| Slovakia | 71.15 | 87.27 | 78.6 | 61.93 | 61.15 | 68.66 | 58.14 | 67.92 | 63.32 | 89.47 | 76.74 | 72.75 | 67.84 |
| Chile | 70.18 | 70.27 | 76.35 | 66.98 | 59.48 | 68.54 | 65.54 | 73.21 | 63.25 | 87.7 | 76.06 | 73.47 | 61.29 |
| Poland | 70.15 | 86.74 | 67.09 | 58.03 | 63.27 | 64.73 | 61.93 | 69.73 | 63.63 | 90.01 | 76.31 | 78.63 | 61.76 |
| Uruguay | 69.69 | 79.43 | 84.87 | 72.82 | 66.76 | 65.19 | 55.77 | 61.59 | 55.53 | 86.04 | 77.98 | 70.64 | 59.61 |
| Rica | 69.59 | 77.8 | 84.74 | 68 | 63.45 | 60.4 | 60.33 | 63.92 | 55.84 | 81.25 | 79.09 | 71.71 | 68.58 |
| Greece | 68.48 | 84.2 | 73.3 | 60.81 | 52.31 | 56.6 | 59.82 | 73.33 | 49.58 | 89.75 | 77.43 | 79.34 | 65.35 |
| Croatia | 68.24 | 84.39 | 75.16 | 56.15 | 52.19 | 60.65 | 52.72 | 72.04 | 60.36 | 88.5 | 75.74 | 73.54 | 67.45 |
| Hungary | 66.88 | 83.8 | 60.36 | 45.75 | 59.51 | 62.1 | 51.62 | 68.01 | 66.05 | 88.37 | 76.7 | 77.34 | 62.95 |
| Malaysia | 66.84 | 70.81 | 46.9 | 57.23 | 57.9 | 73.49 | 69.24 | 70.62 | 64.98 | 79.52 | 77.35 | 72.94 | 61.07 |
| Emirates | 66.69 | 77.15 | 31.27 | 63.51 | 55.16 | 70.06 | 71.89 | 75.34 | 66.79 | 84.88 | 78.38 | 74.72 | 51.11 |
| Romania | 66.4 | 83.65 | 73.56 | 57.11 | 49.96 | 63.04 | 54.16 | 66.39 | 62.04 | 82.3 | 73.02 | 70.47 | 61.11 |
| Qatar | 66.24 | 86.96 | 30.66 | 58.19 | 59.2 | 64.55 | 70.2 | 70.9 | 73.55 | 84.84 | 77.64 | 69.15 | 49.08 |
| Mauritius | 65.65 | 80.8 | 68.3 | 63.85 | 61.47 | 64.84 | 59.18 | 64.65 | 52.93 | 80.22 | 71.86 | 68.68 | 50.97 |
| Bulgaria | 65.55 | 74.84 | 68.31 | 53.12 | 54.62 | 60.37 | 59.52 | 65.32 | 62.46 | 83.43 | 74.11 | 72.79 | 57.72 |
| Montenegro | 65.01 | 80.95 | 68.59 | 55.9 | 59.33 | 56.43 | 64.12 | 63.21 | 54.03 | 83.71 | 68.44 | 71.69 | 53.73 |
| Panama | 64.32 | 77.66 | 67.24 | 49.81 | 60.37 | 59.24 | 53.52 | 67.08 | 60.09 | 76.36 | 75.1 | 57.63 | 67.8 |
| Seychelles | 63.65 | 76.03 | 61.62 | 60.46 | 58.97 | 56.79 | 57.68 | 61.62 | 48.05 | 79.59 | 77.44 | 66.43 | 59.19 |
| Serbia | 62.75 | 79.8 | 60.46 | 44.79 | 61.27 | 54.49 | 54.05 | 60.5 | 55.37 | 87.69 | 71.93 | 73.02 | 49.64 |
| Georgia | 62.28 | 69.34 | 65.58 | 56.04 | 52.1 | 61.47 | 58.17 | 63.94 | 51.19 | 76.55 | 70.64 | 72.33 | 50.05 |
| China | 62.15 | 67.96 | 21.22 | 48.47 | 62.67 | 57.29 | 64.41 | 69.02 | 68.46 | 85.05 | 83.11 | 70.42 | 47.69 |
| Macedonia | 61.95 | 75.76 | 64.71 | 48.48 | 55.7 | 60.9 | 55.11 | 58.44 | 52.73 | 83.35 | 72.63 | 61.04 | 54.59 |
| Tobago | 61.91 | 68.93 | 73.22 | 53.33 | 56.27 | 53.5 | 50.86 | 58.42 | 53.22 | 82.25 | 74.13 | 65.96 | 52.85 |
| Jamaica | 61.39 | 60.01 | 75.12 | 55.5 | 53.68 | 57.93 | 59.59 | 58.43 | 48.02 | 78.19 | 73.83 | 64.13 | 52.2 |
| Argentina | 61.38 | 69.72 | 76.19 | 49.52 | 63.3 | 49.45 | 45.28 | 55.01 | 41.86 | 82.08 | 74.45 | 69.25 | 60.41 |
| Peru | 61.27 | 67.53 | 66.52 | 48.14 | 53.59 | 56.41 | 56.19 | 57.18 | 55.34 | 69.55 | 75.81 | 64.96 | 64.04 |
| Kuwait | 60.93 | 77.12 | 43.15 | 47.04 | 52.71 | 57.3 | 54.8 | 61.54 | 59.91 | 87.69 | 77.33 | 64.29 | 48.27 |
| Armenia | 60.92 | 70.48 | 62.56 | 50 | 47.95 | 55.21 | 62.34 | 58.5 | 50.35 | 77.05 | 73.48 | 69.19 | 53.92 |
| Bahrain | 60.92 | 66.54 | 27.75 | 44.73 | 58.78 | 68.59 | 57.75 | 71.92 | 57.09 | 85.99 | 76.24 | 68.65 | 46.95 |
| Indonesia | 60.88 | 69.86 | 53.59 | 55.24 | 60.62 | 57.14 | 59.26 | 60.81 | 55.98 | 69.55 | 71.13 | 60.93 | 56.4 |
| Thailand | 60.79 | 60.87 | 43.25 | 40.24 | 63.19 | 61.67 | 57.49 | 65.33 | 63.07 | 78.34 | 78.92 | 64.9 | 52.23 |
| Albania | 60.69 | 74.9 | 61.59 | 48.44 | 47.47 | 55.17 | 54.87 | 61.57 | 45.44 | 76.17 | 73.95 | 70.07 | 58.64 |
| Brazil | 60.07 | 60.14 | 63.15 | 47.12 | 61.83 | 54.58 | 46.85 | 54.45 | 50.58 | 81.37 | 71.68 | 59.71 | 69.35 |
| Oman | 59.85 | 81.59 | 33.7 | 47.6 | 53.61 | 59.77 | 59.99 | 66.66 | 54.85 | 78.6 | 75.44 | 62 | 44.41 |
| Republic | 59.85 | 63.77 | 69.45 | 46.32 | 59.32 | 52.85 | 54.19 | 59.39 | 54.33 | 73.4 | 70.64 | 55.55 | 58.98 |
| Kazakhstan | 59.53 | 69.59 | 39.43 | 44.4 | 51.27 | 55.65 | 53.7 | 53.36 | 63.31 | 82.44 | 72.23 | 78.8 | 50.18 |
| Moldova | 59.44 | 72.79 | 63.52 | 47.95 | 56.55 | 51.25 | 47.68 | 53.05 | 47.17 | 79.75 | 70.57 | 71.02 | 51.93 |
| Mexico | 59.3 | 43.54 | 58.33 | 43.89 | 58.18 | 56.86 | 54.1 | 66.84 | 60.17 | 74.22 | 73.09 | 63.93 | 58.4 |
| Herzegovina | 59.11 | 78.67 | 62.75 | 38.69 | 54.92 | 51.38 | 47.58 | 57.6 | 52.37 | 84.38 | 70.5 | 61.93 | 48.56 |
| Vietnam | 58.86 | 69.1 | 34.85 | 47.86 | 65.97 | 45.24 | 52.52 | 63 | 60.33 | 71.8 | 76.99 | 66.56 | 52.09 |
| Ukraine | 58.84 | 54.31 | 62.77 | 45.95 | 58.57 | 40.8 | 54.44 | 58.4 | 51.37 | 78.05 | 68.71 | 77.58 | 55.15 |
| Africa | 58.67 | 52.33 | 73.87 | 57.63 | 60.99 | 61.22 | 57.63 | 59.64 | 52.71 | 68.31 | 59.9 | 52.56 | 47.29 |
| Paraguay | 58.56 | 73.55 | 61.38 | 43.14 | 62.51 | 49.04 | 47.04 | 54.04 | 50.06 | 76.09 | 70.84 | 52.1 | 62.89 |
| Russia | 58.5 | 51.57 | 34.38 | 41.05 | 58.79 | 51.09 | 51.15 | 62.96 | 62.86 | 78.5 | 71.38 | 81.87 | 56.42 |
| Belarus | 58.4 | 71.08 | 32.64 | 35.77 | 45.87 | 45.57 | 47.25 | 57.77 | 60.83 | 86.26 | 74.34 | 81.24 | 62.12 |
| Arabia | 58.35 | 66.15 | 20.87 | 47.72 | 51.98 | 59.79 | 62.64 | 64.22 | 57.77 | 77.87 | 75.34 | 67.77 | 48.12 |
| Verde | 58.34 | 81.78 | 72.66 | 61.12 | 53.31 | 43.88 | 50.68 | 43.98 | 41.07 | 69.14 | 72.75 | 53.96 | 55.78 |
| Colombia | 58.01 | 36.59 | 57.92 | 46.92 | 58.09 | 58.38 | 52.16 | 59.26 | 50.01 | 73.18 | 77.78 | 63.33 | 62.46 |
| Suriname | 57.91 | 73.99 | 69.41 | 47.72 | 53.21 | 41.12 | 47.64 | 53.02 | 44.69 | 78.16 | 60.8 | 56.49 | 68.72 |
| Botswana | 57.83 | 72.1 | 66.64 | 61.46 | 49.88 | 58.73 | 56.22 | 52.51 | 56.28 | 59.67 | 59.07 | 54.5 | 46.93 |
| Philippines | 57.53 | 46.23 | 57.2 | 45.69 | 64.66 | 52.43 | 55.29 | 58.16 | 57.79 | 64.56 | 70.06 | 61.55 | 56.67 |
| Ecuador | 57.42 | 67.6 | 63.66 | 42.88 | 57.09 | 46.58 | 47.06 | 57.33 | 42.97 | 69.69 | 71.8 | 62.24 | 60.18 |
| Jordan | 57.14 | 70.52 | 43.66 | 49.44 | 40.45 | 60.82 | 63.45 | 59.81 | 41.7 | 80.08 | 68.92 | 60.79 | 46.02 |
| ncipe | 57.09 | 79.74 | 69.6 | 51.43 | 52.89 | 49.05 | 54.85 | 42.38 | 50.57 | 60.41 | 70.54 | 48.95 | 54.62 |
| Mongolia | 57.07 | 74.24 | 66.26 | 49.25 | 58.78 | 42.92 | 48.78 | 45.97 | 50.62 | 66.51 | 66.74 | 69.12 | 45.61 |
| Belize | 55.91 | 67.49 | 69.68 | 44.75 | 49.95 | 42.54 | 49.06 | 52.22 | 41.64 | 72.34 | 70.61 | 51.09 | 59.59 |