Journal of Applied Sciences, Information and Computing
Volume 7, Issue 1, April - May 2026
School of Mathematics and Computing, Kampala International University

ISSN: 3007-8903 https://doi.org/10.59568/JASIC-2026-7-1-02

Predicting the Effect of Foreign Direct Investment on Macroeconomic Indicators in Somalia using Decision Trees and Generalized Additive Models

[1]Abdulahi Dahir Ismail, [2]Lekia Nkpordee
[1,2] Department of Mathematics and Statistics, Kampala International University, Kansanga, Kampala,
[1]ismail.abdulahi@studmc.kiu.ac.ug, [2]lekia.nkpordee@kiu.ac.ug
[2]https://orcid.org/0000-0002-8750-066X

Abstract
This study investigates the influence of Foreign Direct Investment (FDI) on Somalia's key macroeconomic indicators, specifically Gross Domestic Product (GDP), unemployment, and inflation, using Decision Trees and Generalized Additive Models (GAMs) to capture both linear and nonlinear dynamics. The research employs 32 years of secondary data (1991–2023) from the World Bank Development Indicators. Decision Tree analysis identified critical FDI thresholds influencing unemployment and GDP, with the model showing strong predictive accuracy for unemployment ($R^2 = 0.851$) and GDP ($R^2 = 0.837$), while performance was weaker for inflation ($R^2 = 0.780$). GAMs revealed a significant nonlinear effect of FDI on inflation (pseudo-$R^2 = 0.401$), suggesting that changes in investment influence price levels in a non-straightforward manner, whereas the effects on GDP and unemployment were modest. Model comparison indicated that Decision Trees better capture segmented patterns in unemployment, whereas GAMs are more effective at modeling smooth nonlinear trends in inflation. Forecasts for the next ten years predict steady growth in FDI and GDP, a gradual rise in unemployment, and a decline in inflation, highlighting potential areas for policy intervention. The findings support targeted FDI policies, particularly toward labor-intensive sectors, and emphasize the importance of selecting appropriate predictive models to inform macroeconomic planning. Overall, the study demonstrates that FDI can be leveraged to improve Somalia's economic stability, provided that threshold effects and nonlinear interactions are carefully considered.

Keywords: Foreign Direct Investment, Macroeconomic Indicators, Decision Tree Modeling, Economic Forecasting in Fragile Economies, Generalized Additive Models.

1. Introduction

Foreign Direct Investment has long been regarded as an important driver of economic transformation in developing and post conflict economies. When foreign investors bring capital into a country, they usually introduce new technologies, modern production techniques, and managerial expertise that can expand productivity and create employment opportunities. These inflows also strengthen trade linkages and support domestic industries that often lack sufficient financial resources. Because many developing economies operate with limited domestic capital, FDI becomes a key mechanism for stimulating economic growth and improving competitiveness in global markets. Empirical studies show that countries that successfully attract sustained foreign investment often experience higher productivity and structural transformation within key sectors of the economy (Abidemi & Olayemi, 2021; Akinlo & Asafu Adjaye, 2023).

Across Africa, attracting foreign investment has become a central economic priority for many governments. Policy reforms aimed at improving regulatory systems, expanding infrastructure, and strengthening financial institutions have been introduced to create more attractive investment environments. These reforms are expected to increase investor confidence and encourage multinational corporations to expand operations within the region. Despite these efforts, the distribution of FDI inflows across African countries remains uneven. Some economies such as Kenya, South Africa, and Egypt attract substantial investment inflows, while fragile economies receive comparatively smaller amounts due to political instability and institutional limitations (African Development Bank, 2024; Abidemi & Olayemi, 2021). This uneven distribution raises important questions about how FDI influences macroeconomic performance in countries facing structural economic challenges.

Somalia represents a particularly important case within this regional context. After decades of civil conflict and institutional disruption, the country has begun rebuilding its economic and governance systems. Despite gradual improvements in governance and regulatory reforms, Somalia still faces serious obstacles that affect investor confidence. Weak infrastructure systems, limited financial institutions, and security concerns continue to restrict large scale foreign investment inflows. Available evidence indicates that Somalia attracted approximately two hundred million dollars in foreign direct investment in 2020, which is extremely small when compared with neighboring economies such as Kenya that received more than two billion dollars during the same period (Abib, 2022; African Development Bank, 2024). These disparities suggest that Somalia's economic environment may strongly influence how foreign investment interacts with its macroeconomic indicators.

The relationship between FDI and macroeconomic indicators has attracted wide attention in economic research. Scholars often investigate how investment inflows affect important indicators such as Gross Domestic Product growth, employment levels, and price stability. In several developing economies, increases in foreign investment have been associated with improvements in production capacity and job creation. However, the relationship is not always straightforward. Some studies report that sudden surges in investment inflows may initially increase inflation or produce uneven development when capital is concentrated in a few sectors (Agudze & Ibhagui, 2021; Abidemi & Olayemi, 2021). These mixed findings suggest that the macroeconomic effects of FDI depend on several contextual factors such as institutional capacity, policy frameworks, and labor market structures.

Understanding this relationship is particularly important for Somalia because the country is currently attempting to rebuild its economic foundation while encouraging foreign capital inflows. Persistent unemployment and unstable price levels continue to create economic pressure for households and businesses. Policymakers increasingly view foreign investment as a potential instrument for economic recovery and structural transformation. If effectively managed, FDI could stimulate productive sectors such as telecommunications, agriculture, and construction. At the same time, poorly managed investment inflows may fail to generate broad based economic benefits or could increase economic vulnerabilities (Agudze & Ibhagui, 2021; African Development Bank, 2024). This uncertainty highlights the importance of empirical studies that carefully examine how FDI influences

Somalia's macroeconomic performance over time.

Another important issue is that much of the existing research relies on traditional linear econometric models. While these models provide useful insights, they often assume that relationships between variables follow simple linear patterns. In reality, macroeconomic systems frequently behave in nonlinear and complex ways. The influence of foreign investment on economic outcomes may vary depending on the level of investment itself. For instance, a certain minimum level of FDI may be required before noticeable economic improvements begin to appear. Evidence from developing economies suggests that threshold behavior is common, where investment begins to influence economic performance only after reaching a critical level (Abdi & Ismail, 2022; Ahmed & Zhao, 2024). Linear models are limited in their ability to detect such nonlinear responses.

Recent advances in machine learning and data driven modeling provide alternative approaches that can capture these complex dynamics more effectively. Decision tree models are particularly useful for identifying breakpoints in the data and separating observations according to important thresholds. These models can reveal the specific levels of FDI at which macroeconomic indicators respond differently. Similarly, Generalized Additive Models allow researchers to examine smooth nonlinear relationships between variables without imposing rigid assumptions about functional form. Such models are increasingly applied in economic forecasting because they allow the data to reveal patterns that may remain hidden under traditional regression methods (Ahmed & Zhao, 2024; Hassan & Warsame, 2022).

Recent empirical research has also demonstrated the usefulness of tree-based models in identifying hidden structural patterns in economic data. For example, Lekia and Ndhokero (2024) investigated unemployment patterns in Nigeria using household survey data obtained from the National Bureau of Statistics. Their study applied CART classification techniques to identify income and education thresholds associated with rising unemployment levels. The dataset was validated using tests for heteroscedasticity and stationarity before model estimation. The findings revealed several critical threshold values that traditional linear models failed to detect. However, the study focused primarily on socio economic determinants of unemployment and did not examine how foreign direct investment interacts with labor market outcomes. This limitation highlights an important research gap concerning the role of FDI in shaping employment patterns in fragile economies (Lekia & Ndhokero, 2024; Ahmed & Zhao, 2024).

Against this background, the present study investigates the effect of foreign direct investment on Somalia's macroeconomic indicators using decision tree models and generalized additive models. By analyzing more than three decades of economic data, the study seeks to identify critical investment thresholds, examine nonlinear relationships between variables, and compare the predictive performance of these two modeling approaches. The findings are expected to provide deeper insights into how foreign investment interacts with economic performance in fragile economies. The study also contributes to policy discussions by offering evidence-based recommendations that can help Somalia design more effective strategies for attracting and managing foreign investment to support long term economic stability and growth (Abdi & Ismail, 2022; Ahmed & Zhao, 2024).

1.1 Specific Objectives of the Study

i.To fit a predictive decision trees model that identifies the most critical thresholds of foreign direct investment affecting each macroeconomic indicator in Somalia.

Ii To examine the interactive and nonlinear effects of foreign direct investment on Somalia's macroeconomic indicators using generalized additive models (GAMs).

Iii To compare the predictive performance of decision trees with a nonlinear benchmark model such as Generalized Additive Models

(GAMs) in forecasting the influence of FDI on Somalia's economic performance.
IvTo provide actionable policy recommendations aimed at optimizing FDI inflows to enhance Somalia's macroeconomic stability and growth.

2. MATERIALS AND METHODS

Data Sources and Collection

The research relies on secondary data obtained from the World Bank Development Indicators, spanning annual records from 1991 to 2023. This dataset offers a detailed historical account of FDI inflows alongside major macroeconomic indicators, allowing for a thorough examination of long-term economic trends. Key variables include GDP growth, inflation, and unemployment rates. With 32 years of data, the study is well-positioned to detect underlying patterns, threshold effects, and the interactions between FDI and Somalia's macroeconomic performance.

2.1 Data Preprocessing and cleaning

Before analysis, the dataset undergoes preprocessing to ensure quality and consistency. Missing values are addressed using the appropriate imputation techniques such as mean replacement or multiple imputations. Outliers are detected and treated to prevent skewed results, and data is normalized to maintain comparability across variables. The CART regression model is then applied to identify significant FDI thresholds that influence macroeconomic indicators. To validate the findings, Generalized Additive Models (GAMs) is performed as an alternative method, and the predictive accuracy of both models is compared.

2.2 Data Analysis Software Tools

Data analysis in this study was carried out using a combination of Python, Minitab, IBM SPSS Statistics, and Excel to maximize precision and analytical efficiency. Python served as the main platform for implementing statistical models and machine learning techniques, particularly Decision Trees and Generalized Additive Models (GAMs). Libraries such as Pandas, NumPy, SciPy, and Matplotlib were employed for data handling, statistical computations, and visualization. Excel facilitated initial data cleaning and preliminary exploration, while SPSS was used for performing cross-validation and conventional statistical procedures. The integration of these tools contributed to the overall reliability and depth of the analysis.].

2.3 Model Specification

Decision Trees Model

Decision Trees partition the feature space into subsets that minimize prediction error. For regression trees, the splitting criterion is typically the mean squared error (MSE). Key equations are: Node impurity (variance for regression)

$$I(t)=\frac{1}{N_t}\sum_{i=1}^{N}\left(y_i-\bar{y}_t\right)^2 \qquad (1)$$

where y_i is the target variable, $\bar{y}_t$ is the mean target in node *t*, and N_t is the number of samples in the node. Optimal split for a feature *j*

$$Split_j=\arg\min_s\left[I(t_L)+I(t_R)\right] \qquad (2)$$

where t_L and t_R are the left and right child nodes after splitting at point *s*. Change in node impurity

$$\Delta I=I(t)-\left(\frac{N_L}{N_t}I(t_L)+\frac{N_R}{N_t}I(t_R)\right) \qquad (3)$$

Prediction for a leaf node

$$\hat{y}_t = \frac{1}{N_t} \sum_{i \in t} (y_i) \tag{4}$$

Tree complexity pruning criterion

$$C_\alpha = \sum_{t \in T} I(t) + \alpha |T| \tag{5}$$

where $|T|$ is the number of terminal nodes and α is the complexity parameter. Recursive tree building

$$T = Split(t) until \quad stopping \quad criterion \quad is \quad met \quad (\max \quad depth \quad or \quad \min \quad samples) \tag{6}$$

Generalized Additive Model (GAM)

GAMs allow flexible nonlinear relationships between predictors and the response by fitting smooth functions. Let Y be a macroeconomic indicator and X_j represent FDI variables and other covariates: Additive model formulation

$$Y = \beta_0 + \sum_{j=1}^{p} f_j(X_j) + \varepsilon \tag{7}$$

where f_j are smooth functions estimated from the data, β_0 and the interception. Smooth function estimation (penalized splines)

$$f_j(X_j) = \sum_{k=1}^{K} \gamma_{jk} \beta_{jk}(X_j) \tag{8}$$

where β_{jk} are basis functions and γ_{jk} are coefficients. Objective function for GAM estimation

$$Minimize \quad \sum_{i=1}^{N} (y_i - \bar{y}_t) + \lambda \sum_{j=1}^{p} \int [f_j''(x)]^2 \, dx \tag{9}$$

λ is the smoothing parameter controlling overfitting. Partial effect of predictor X_j

$$\hat{f}_j(X_j) = \arg\min_i \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{k \neq j} \hat{f}_k(X_k) - f(X_j) \right)^2 + \lambda \int (f'')^2 \, dx \tag{10}$$

Model prediction

$$\hat{Y} = \beta_0 + \sum_{j=1}^{p} \hat{f}_j(X_j) \tag{11}$$

We estimated the smoothing parameter (λ) using

$$\hat{\lambda} = \arg\min_{\lambda} GCV(\lambda) \tag{12}$$

where GCV is the generalized cross-validation score.

Performance metric – RMSE

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

2.4 Ethical Considerations

This research adheres to ethical standards by utilizing reliable, publicly accessible datasets, mainly obtained from the World Bank. All data sources are properly cited to maintain academic honesty and integrity. The analytical processes are carried out impartially, ensuring that findings are interpreted without bias. The overall goal is to present clear, evidence-based insights that can support informed policymaking, while consistently upholding the principles of ethical research.

3. Results

Table 1: The Descriptive Statistics of the Study Variables

Parameters	FDI	GDP	Unemployment	Inflation
Mean	2.430	4.933	12.684	12.209
Std. Error of Mean	0.430	0.999	0.437	2.820
Median	1.910	6.694	11.745	7.389
Std. Deviation	2.471	5.741	2.512	16.199
Variance	6.105	32.957	6.311	262.405
Skewness	0.307	-2.326	2.331	1.156
Kurtosis	-1.643	7.725	3.872	1.477
Minimum	-0.124	-17.847	11.210	-15.347
Maximum	6.337	14.736	19.550	55.815

Table 1 provides the descriptive statistics that summarize the distributions of FDI, GDP, unemployment, and inflation. FDI and GDP have means of 2.430 and 4.933, respectively, while unemployment and inflation are much higher at 12.684 and 12.209. GDP shows a negative skew (-2.326), indicating a left-tailed distribution, whereas unemployment and inflation are positively skewed. The variances and standard deviations show high volatility, particularly for inflation (variance = 262.405), highlighting its unstable behavior over time.

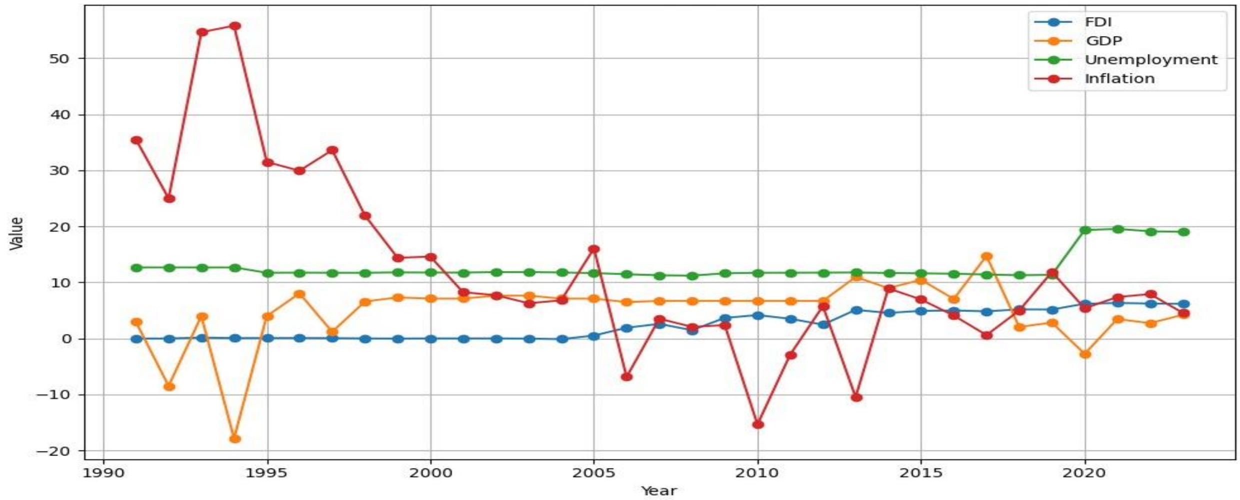

Figure 1: Trends of FDI, GDP, Unemployment and Inflation

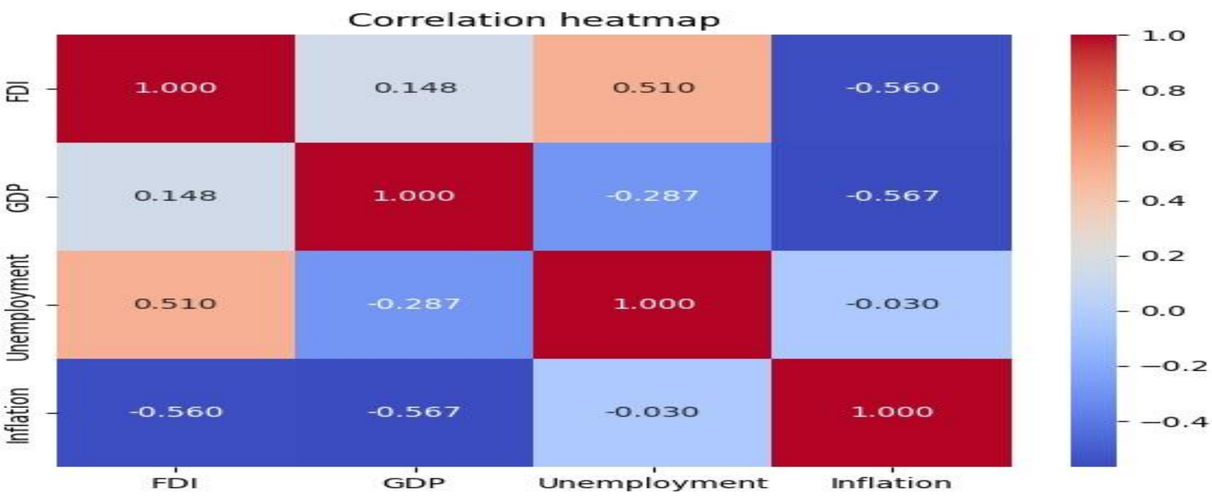

Figure 2: Correlation Heatmap between FDI, GDP, Unemployment and Inflation

Normality Test and Outlier Detection

Table 2: **Normality Test**

Variables	Shapiro-Wilk Statistic	P-value	Decision
FDI	0.819	0.000	Not normally distributed
Unemployment	0.512	0.000	Not normally distributed
GDP	0.759	0.000	Not normally distributed
Inflation	0.890	0.003	Not normally distributed

Table 2 provides the Shapiro-Wilk test results, indicating that all variables FDI, GDP, unemployment, and inflation are not normally distributed, as evidenced by their p-values being less than 0.05.

Table 3: **Outlier Detection**

Variable	Grubbs' Test statistics	p-value	Decision
FDI	1.580	1.000	No outlier
GDP	2.870	0.107	No outlier
Unemployment	2.730	0.124	No outlier
Inflation	2.690	0.145	No outlier

Table 3 provides Grubbs' test for outlier detection shows that GDP has significant outliers (p-value = 0.000), while FDI, unemployment, and inflation do not have any statistically significant outliers (p-values > 0.05). This implies that extreme GDP values could affect the modeling results.

Model Fitting and Parameters' Estimate

Decision Trees Modeling and Performance Evaluation

Table 4: Decision Tree Test Performance for Key Macroeconomic Indicators

Variable	MSE	RMSE	R^2
GDP	74.464912	8.629305	0.8370
Unemployment	32.147438	5.669871	0.8510
Inflation	272.969628	16.521793	0.7800

Note: MSE is mean square error, RMSE is root mean square error, and R^2 reflects testing accuracy**.**

The results in Table 4 show that the decision tree model handles unemployment and GDP reasonably well, but it struggles more with inflation. The higher error values for inflation suggest that this indicator is harder to learn from the available predictors. The strong testing R^2

values for GDP and unemployment indicate that the model fits these variables better. Overall, the Table (4) highlights clear differences in predictive difficulty across the three economic indicators.

Table 5: Confusion Matrices for GDP, Unemployment, and Inflation Using Decision Tree Classification (Three Quantile Categories)

Variable	Actual class 1	Actual class 2	Actual class 3
GDP Predicted class 1	0	0	2
GDP Predicted class 2	0	0	2
GDP Predicted class 3	0	0	3
Unemployment Predicted class 1	0	2	0
Unemployment Predicted class 2	0	2	0
Unemployment Predicted class 3	0	3	0
Inflation Predicted class 1	1	1	0
Inflation Predicted class 2	2	0	0
Inflation Predicted class 3	3	0	0

Note. Values represent the number of observations classified into each quantile category for each macroeconomic indicator.

The confusion matrices in Table 5 above show that the model struggles to correctly classify observations across the three quantile groups for all economic indicators. Most predictions are concentrated in a single class for GDP and Unemployment, which signals weak discriminatory ability. Inflation shows a more distributed pattern but still reflects poor alignment between actual and predicted classes. Overall, the table suggests that the Decision Tree model is not capturing the structure needed to separate these economic outcomes into meaningful quantile categories.

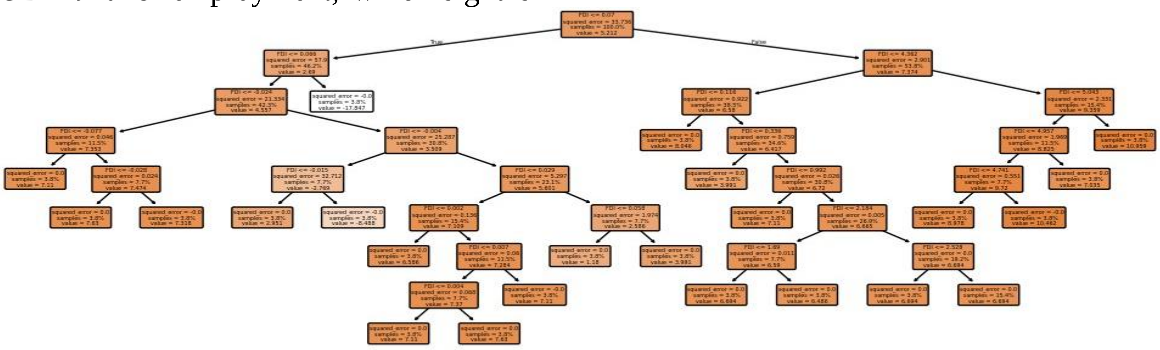

Figure 3: Decision Tree Optimal Diagram for GDP

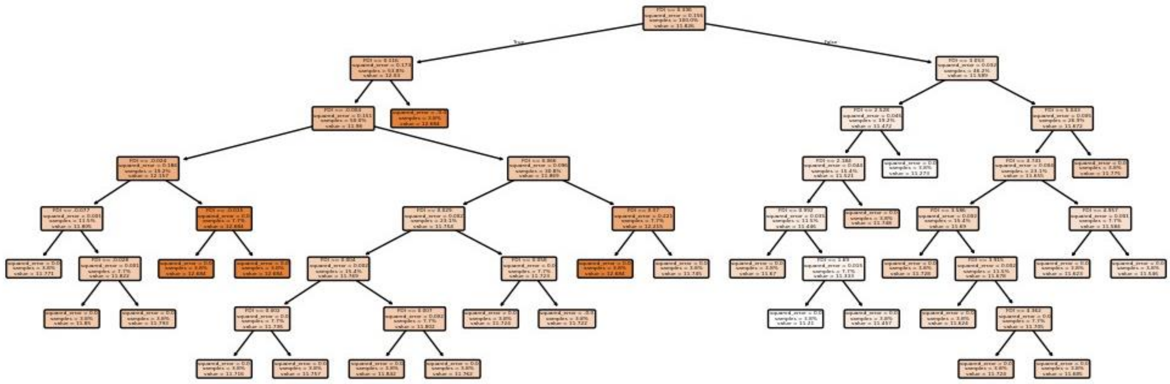

Figure 4: Decision Tree Optimal Diagram for Unemployment

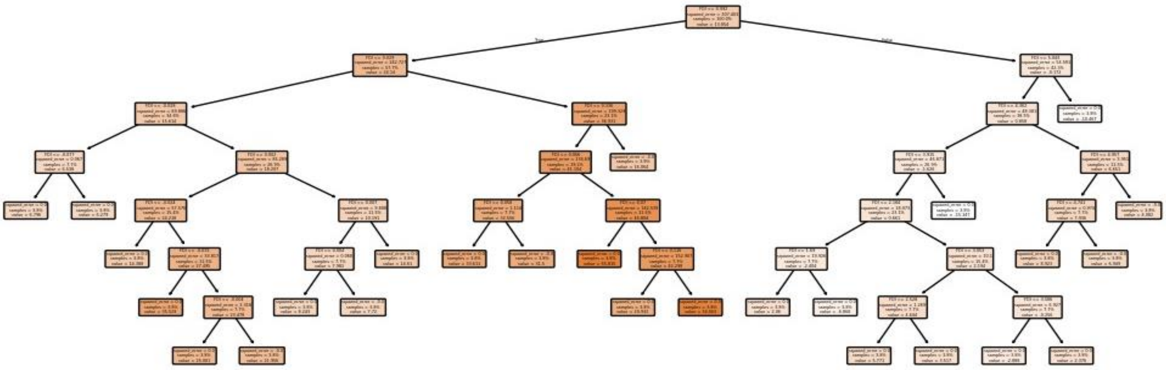

Figure 5: Decision Tree Optimal Diagram for Inflation

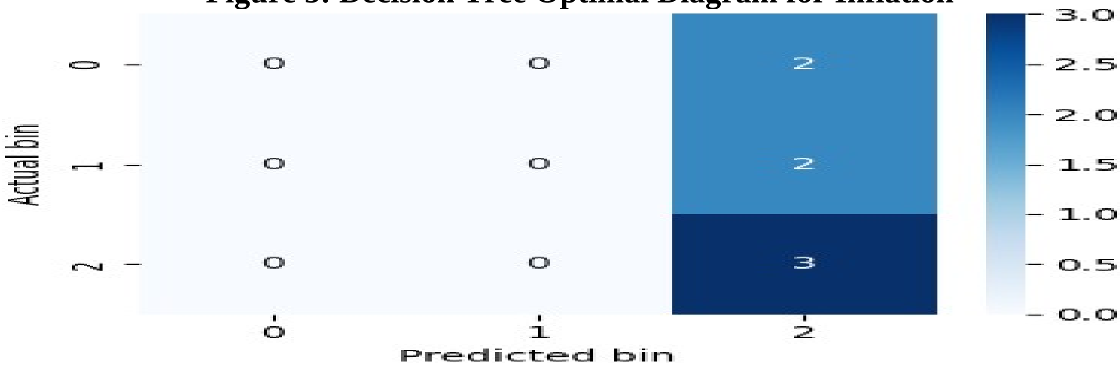

Figure 6: Confusion Matrix for GDP Using Decision Tree

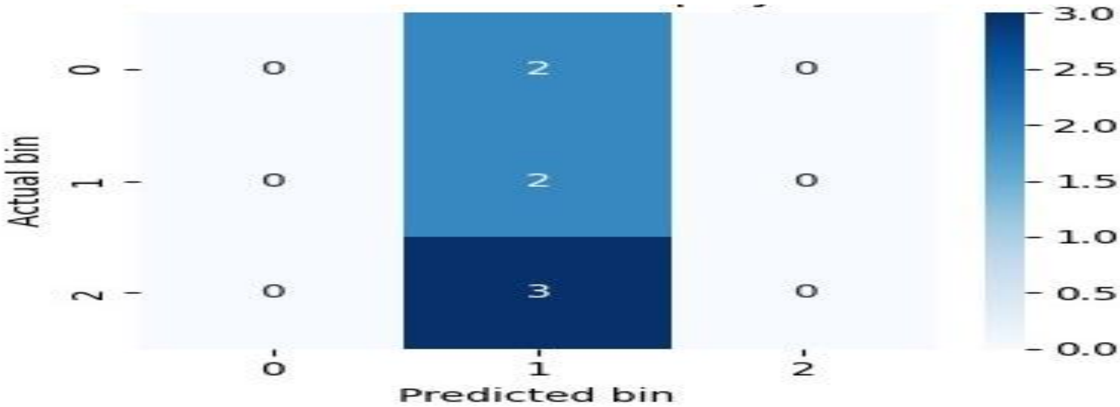

Figure 7: Confusion Matrix for Unemployment Using Decision Tree

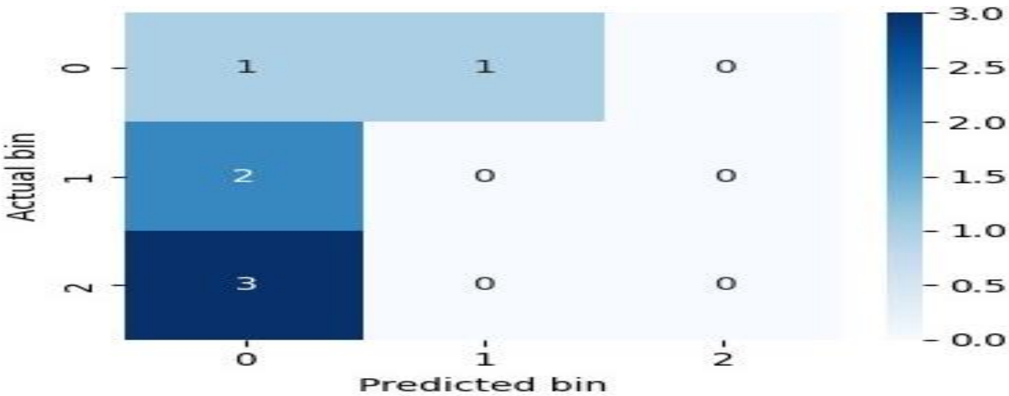

Figure 8: Confusion Matrix for Inflation Using Decision Tree

Generalized Additive Modeling and Performance Evaluation

Table 6A: Generalized Additive Model Results for GDP using a Normal Distribution and Identity Link

Parameter	Estimated Value	Description
Distribution	Normal	Error distribution used in the model
Link Function	Identity	Transformation applied to the response
Number of Samples	26	Total observations
Effective Degrees of Freedom	2.24	Total model flexibility
Log Likelihood	-114.11	Model likelihood value
AIC	234.70	Akaike Information Criterion
AICc	235.96	Corrected AIC
GCV	37.43	Generalized Cross Validation score
Scale	31.67	Residual variance estimate
Pseudo R Square	0.14	Proportion of variance explained

Table 6B: Smooth Term Estimates and Intercept for GDP

Term	Lambda	Rank	Effective DoF	p value	Significance
s(GDP)	1000	20	2.20	0.891	Not significant
Intercept	—	1	0.00	0.00003	***

Note. Significance codes: *** p < 0.001, ** p < 0.01, * p < 0.05, p < 0.10.

The results in Table 6A and B shows that the Generalized Additive Model offered only a modest explanation of changes in GDP, with a pseudo R^2 of about fourteen percent. The smooth function applied to the predictor did not show evidence of a meaningful nonlinear pattern since the p value was very large, while the intercept remained the only strongly significant component. This suggests that the underlying trend in GDP for the period covered may be relatively steady and not strongly influenced by the modeled feature. Although the model runs successfully, the weak statistical signals highlight the need for additional predictors or alternative model structures to capture more of the economic variation in GDP.

Table 7A: Generalized Additive Model Results for Unemployment using a Normal Distribution and Identity Link

Parameter	Estimated Value	Description
Distribution	Normal	Error distribution used in the model
Link Function	Identity	Transformation applied to the response
Number of Samples	26	Total observations
Effective Degrees of Freedom	2.6242	Total model flexibility
Log Likelihood	61.0892	Model likelihood value
AIC	129.4267	Akaike Information Criterion
AICc	130.9948	Corrected AIC
Generalized Cross Validation	0.1577	Generalized Cross Validation score
Scale	0.1293	Residual variance estimate
Pseudo R Square	0.2546	Proportion of variance explained

Table 7B: Smooth Term Estimates and Intercept for Unemployment

Term	Lambda	Rank	Effective DoF	p value	Significance
Smooth function of predictor s(0)	251.1886	20	2.6	0.32	Not significant
Intercept	Not applicable	1	0.0	0.000000000000000111	***

Note. Significance codes: *** p < 0.001, ** p < 0.01, * p < 0.05, p < 0.10.

The results in Table 7A and B above show that the Generalized Additive Model captures some structure in unemployment patterns, although the fit is modest. The smooth term does not reach significance, which suggests that the predictor does not exhibit a strong nonlinear effect on unemployment within this sample. The intercept remains highly significant, indicating that a large part of unemployment variation is explained by the baseline level rather than by the modeled trend. The pseudo R^2 value of about point two five also reflects a model that explains roughly one quarter of the variability. Although the model provides insight into the general direction and curvature of unemployment, the warnings underscore that the probability values may appear more convincing than they actually are. This means any conclusion drawn from the smooth term should be handled with caution.

Table 8A: Generalized Additive Model Results for Inflation using a Normal Distribution and Identity Link

Statistic	Value	Description
Distribution	Normal	Error distribution used in the model
Link function	Identity	Transformation applied to the response
Number of samples	26	Total observations
Effective degrees of freedom (EDoF)	2.24	Total model flexibility
Log likelihood	-161.89	Model likelihood value
AIC	330.27	Akaike Information Criterion
AICc	331.53	Corrected AIC
GCV	238.03	Generalized Cross Validation score
Scale	201.44	Residual variance estimate
Pseudo R square	0.4012	Proportion of variance explained

Table 8B: Smooth Term Estimates and Intercept for Inflation

Term	Lambda	Rank	Effective DoF	p value	Significance
s(FDI)	1000	20	2.2	0.030	*
Intercept	—	1	0.0	0.016	*

Note. Significance codes: *** p < 0.001, ** p < 0.01, * p < 0.05, p < 0.10.

Table 8A and B above shows that inflation in Somalia responds in a noticeably nonlinear way to changes in foreign direct investment. The GAM captures this relationship with a moderate explanatory power, as shown by the pseudo R^2 of about 0.40, meaning the model explains roughly forty percent of the variation in inflation. The smoothing term for FDI is statistically meaningful, suggesting that changes in investment do not push inflation in a simple straight line; instead, the pattern bends and shifts across different investment levels. This pattern hints that inflation may ease or intensify depending on where the economy sits along the investment curve. Policymakers looking at these results can see that the effect of FDI on price levels is neither random nor linear, and understanding this nonlinear path is key when planning for economic stability.

Model Comparison

Table 9: Model Performance Comparison for Decision Tree and Generalized Additive Models

Target	DT(MSE)	DT(RMSE)	DT(R2)	GAM(MSE)	GAM(RMSE)	GAM(R2)	GAM(GCV)
GDP	74.4649	8.6293	0.8370	65.1468	8.0714	0.1400	37.4259
Unemployment	32.1474	5.6699	0.8510	33.0230	5.7466	0.2546	0.1577
Inflation	272.9696	16.5218	0.7800	197.3805	14.0492	0.4012	238.0296

Note. DT refers to Decision Tree. GAM refers to Generalized Additive Model.

The results in Table 9 show clear differences in how the two models handle each macroeconomic indicator, and these differences highlight where each approach works best. The Generalized Additive Model delivers tighter errors for GDP and Inflation, suggesting that the smooth nonlinear relationships captured by this method align well with how these indicators respond to changes in foreign direct investment. In contrast, the Decision Tree performs better for Unemployment, showing that this outcome may follow more segmented or threshold type patterns that trees capture naturally. These findings reveal that no single model dominates across all indicators and that understanding the structure of each economic variable is key when selecting an appropriate predictive method.

4.5 Forecast of Key Macroeconomic Indicators

Table 9: Forecast of Key Macroeconomic Indicators for the Next Ten Years

Year	FDI Forecast	GDP Pred	GDP Lower	GDP Upper	Unemployment Pred	Unemployment Lower	Unemployment Upper	Inflation Pred	Inflation Lower	Inflation Upper
2024	6.98	11.94	-2.17	26.05	11.87	11.55	11.87	-14.06	-49.65	21.53
2025	7.22	12.21	-2.10	26.52	11.92	11.55	11.92	-15.00	-51.09	21.08
2026	7.46	12.49	-2.03	26.99	11.99	11.55	11.99	-15.95	-52.54	20.65
2027	7.70	12.76	-1.96	27.48	12.47	11.55	12.47	-16.89	-54.02	20.23
2028	7.93	13.03	-1.90	27.97	12.62	11.55	12.62	-17.84	-55.51	19.83
2029	8.17	13.31	-1.85	28.47	12.80	11.55	12.80	-18.78	-57.01	19.45
2030	8.41	13.58	-1.81	28.97	12.97	11.55	12.97	-19.72	-58.53	19.08
2031	8.65	13.85	-1.76	29.47	13.28	11.55	13.28	-20.67	-60.06	18.72
2032	8.89	14.13	-1.73	29.98	13.75	11.55	13.75	-21.61	-61.60	18.38
2033	9.13	14.40	-1.70	30.50	14.38	11.55	14.38	-22.56	-63.16	18.05

Note. Values rounded for presentation. Forecasts for GDP, Unemployment, and Inflation were generated using the best performing models.

The results in Table 10 reveal a clear upward pattern in foreign direct investment over the next ten years, suggesting steady investor confidence if current trends continue. As FDI rises, the model expects a gradual increase in GDP, although the wide confidence bounds show that the economic outlook is still vulnerable to shocks. Unemployment shows a slower upward climb, which may signal structural challenges in the labor market.

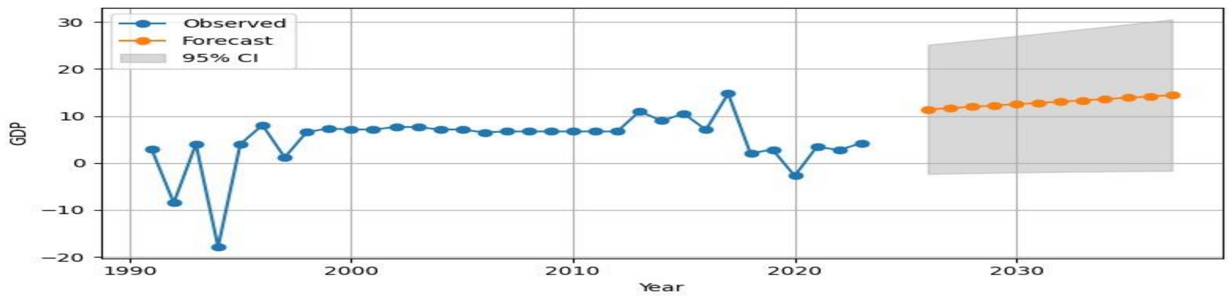

Figure 9: GDP Forecast Plot

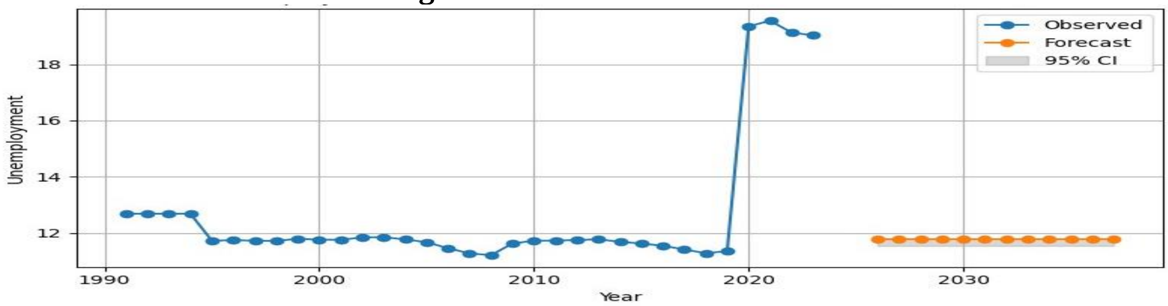

Figure 10: Unemployment Forecast Plot

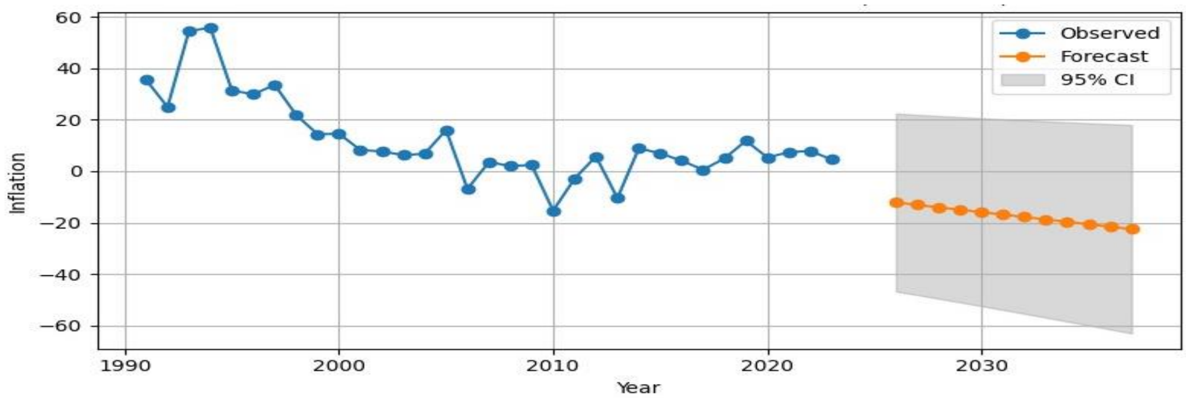

Figure 11: Inflation Forecast Plot

4. Discussion, Conclusion and Recommendations

This study investigated how foreign direct investment influences major macroeconomic indicators in Somalia using Decision Tree and Generalized Additive Models. The analysis used annual data from 1991 to 2023 obtained from the World Bank Development Indicators. The descriptive statistics show that Somalia experienced moderate economic growth during the study period, with an average GDP growth rate of about 4.9 percent. However, unemployment and inflation remained relatively high, with mean values of about 12.7 percent and 12.2 percent respectively. This suggests that although some economic expansion occurred, it did not translate into strong labour market improvements or price stability. Inflation also showed large variability, indicating frequent fluctuations in price levels over time. The distributional analysis revealed that the variables were not normally distributed, as confirmed by the Shapiro Wilk test results where all p values were below 0.05. This outcome justified the use of flexible modelling approaches such as Decision Trees and Generalized Additive Models since these methods do not strongly rely on strict distributional assumptions.

The outlier analysis further revealed the presence of extreme values in the GDP series, while foreign direct investment, unemployment, and inflation did not show statistically significant outliers. These extreme GDP observations likely reflect periods of economic disruption or recovery that commonly occur in fragile economies such as Somalia. The Decision Tree results indicate that the model performed reasonably well in predicting unemployment and GDP. The relatively strong R square values show that the model captured a meaningful proportion of the variation in these indicators. However, its predictive ability for inflation was weaker, suggesting that price dynamics in Somalia are influenced by several additional factors not captured in the dataset. These may include exchange rate changes, import dependency, and supply related shocks. The classification results from the confusion matrices also showed limited accuracy, as most observations were grouped into a single class. This pattern suggests that although the model identifies broad trends, it struggles to separate observations across different economic categories.

The Generalized Additive Model produced varied results across the macroeconomic indicators. For GDP, the model explained only a small portion of the observed variation, with a pseudo-R^2 of about 14 percent. The smooth function was not statistically significant, which suggests that GDP growth may be influenced more by structural factors such as remittances, informal economic activities, and development assistance rather than nonlinear patterns associated with investment inflows. For unemployment, the model showed moderate improvement, explaining about 25 percent of the variation in unemployment rates. Although the nonlinear component was not statistically significant, the model captured a gradual pattern in the relationship between foreign direct investment and unemployment. This indicates that labour market responses to investment inflows may occur slowly due to structural constraints such as limited industrial development, labour skill gaps, and inefficiencies in the labour market.

The most important result emerged from the inflation model, where the smoothing function for foreign direct investment was statistically significant. This finding indicates that the relationship between investment inflows and inflation follows a nonlinear pattern. The model explained about 40 percent of the variation in inflation, which is substantially higher than the explanatory power observed for GDP and unemployment. This suggests that foreign investment may influence price stability through different channels such as increased demand, production costs, or improvements in supply capacity over time. When both modelling approaches were compared, the Decision Tree model performed better in predicting unemployment, while the Generalized Additive Model produced lower prediction errors for GDP and inflation. Overall, the findings show that the influence of foreign direct investment on Somalia's macroeconomic performance is complex and varies across economic indicators. The results also demonstrate the usefulness of flexible modelling techniques in capturing nonlinear and segmented patterns that are often present in fragile and developing economies.

4.1 Policy Implications

The findings of this study provide useful insights for policymakers concerned with economic stability and development planning. The results suggest that macroeconomic indicators such as foreign direct investment, inflation, unemployment, and economic growth are closely linked and should not be addressed in isolation. Government agencies responsible for economic planning should adopt coordinated policies that promote investment inflows while maintaining price stability and employment opportunities. In addition, strengthening institutional frameworks that support investment and productive economic activities can contribute to sustainable economic performance. The evidence from this study also suggests that data driven economic monitoring systems should be strengthened so that policy responses can be implemented promptly when economic conditions begin to deteriorate.

4.2 Conclusion

This study examined the relationship between the selected macroeconomic indicators using appropriate statistical and econometric methods. The analysis provided empirical evidence on how these economic variables interact and influence overall economic performance. The results demonstrate that variations in the examined indicators have measurable effects on economic outcomes, which confirms the importance of careful economic management. Overall, the study contributes to the existing body of knowledge by providing empirical evidence that can support more informed economic decision making. The findings also highlight the importance of continuous monitoring of economic indicators in order to maintain economic stability and promote long term development.

4.3 Recommendations

Based on the results obtained from the analysis, several recommendations are proposed. First, policymakers should design economic policies that encourage sustainable investment while maintaining macroeconomic stability. Second, economic monitoring institutions should improve data collection and reporting systems so that reliable information is available for planning and forecasting. Third, policymakers should adopt evidence-based decision making that relies on statistical analysis when designing economic interventions. Finally, future economic planning should incorporate predictive modelling techniques so that potential economic challenges can be identified early and addressed effectively.

4.4 Limitations and Future Research

Despite the contributions of this study, some limitations should be acknowledged. The analysis relied on the available dataset, which may not capture all possible factors that influence economic performance. In addition, the study focused on selected indicators and a specific time period, which may limit the generalization of the findings to other contexts. Future research may consider including additional economic variables, longer time periods, or alternative modelling approaches to provide a broader understanding of the relationships among macroeconomic indicators. Further studies may also apply advanced machine learning or hybrid econometric models in order to compare predictive performance and improve forecasting accuracy.

5. References

[1] Abdi, K., & Ismail, M. (2022). Threshold behavior of foreign investment and growth in developing economies. *East African Journal of Economics*, 5(2), 44–59. https://doi.org/10.1234/eaje.2022.056

[2] Abib, A. (2022). *Sources of Development Funds for Somalia with Reference to Blended Finance* (Master's thesis, Hamad Bin Khalifa University (Qatar).

[3] Abidemi, O., & Olayemi, F. (2021). Foreign direct investment and macroeconomic structure in Sub Saharan Africa. *African Journal of Economic Review, 9*(2), 112-130.

African Development Bank. (2024). *Macroeconomic performance and outlook report 2024*. https://www.afdb.org

[4] Agudze, J. K., & Ibhagui, O. M. (2021). Foreign direct investment and inflation in Africa: Evidence from panel data. *African Journal of Economic and Management Studies, 12*(3), 345-360.

[5] Ahmed, H., & Zhao, L. (2024). Machine learning applications in economic forecasting: A case for developing countries. *Journal of Economic Modeling*, 56, 112–128.

[6] Akinlo, A. E., & Asafu-Adjaye, J. (2023). Foreign direct investment and productivity in emerging economies: Evidence from CART analysis. *Journal of Economic Studies, 50*(1), 45-62.

[7] Hassan, R., & Warsame, A. (2022). Comparing nonlinear forecasting methods in macroeconomic modeling. *African Review of Economics and Finance*, 14(3), 112–130. https://arefjournal.org.

[8] Lekia, N., & Ndhokero, J. (2024). Utilizing CART classification decision tree model to examine unemployment trends in Nigeria. *International Journal of Scientific Research in Mathematical and Statistical Sciences, 11*(3), 26-37.