A Keynesian Approach to U.S. GDP Growth Using Key Determinants

Kaleb Jordan

2025-01-10

An Investigation of Key Determinants of U.S. GDP Growth: A Keynesian Approach

Abstract

This paper examines the key determinants of U.S. Gross Domestic Product (GDP) growth using a combination of Keynesian economic theory and empirical modeling. By leveraging both traditional econometric techniques (i.e., stepwise regression) and advanced machine learning models (i.e., Random Forest), this study aims to provide an understanding of the factors influencing GDP growth. In particular, we focus on variables such as consumption, investment, government spending, and net exports—central to Keynesian economic theory—while validating the predictions using a robust machine learning approach. The results suggest that government expenditures and Total Credit to Private Non-Financial Sector play significant roles in driving GDP growth, with the Random Forest model providing more insightful predictive capabilities in comparison to the conventional stepwise regression model.

Introduction

Economic growth is a fundamental objective for policymakers, economists, and institutions around the world. To enhance the effectiveness of fiscal and monetary policies, it is critical to understand the underlying factors that drive economic growth. For over a century, the Keynesian economic framework has provided a solid foundation for understanding aggregate demand and its components, notably consumption, investment, government spending, and net exports. In this study, we investigate the relationship between GDP growth and these key determinants, utilizing a combination of theoretical insight and empirical modeling techniques.

Given the complexities of economic systems, we use two complementary approaches to modeling the relationships between the variables:

1. **Linear Regression**: model that examines the linear relationships in line with traditional Keynesian economic theory.
2. **Random Forest**: a non-linear machine learning method, to validate and extend our findings by capturing intricate interactions and non-linear relationships.

Theoretical Framework

Keynesian Economic Theory

Keynesian economic theory emphasizes the role of aggregate demand in determining the overall level of economic activity. GDP is the sum of consumption (C), investment (I), government spending (G), and net exports (X-M). Mathematically:

Where:

- **Consumption (C)**: Driven by consumer sentiment and household income, consumption is often the largest component of GDP.
- **Investment (I)**: Reflects business sentiment and the amount of credit available to firms, with potential lag effects on the economy.
- **Government Spending (G)**: Comprising public expenditures aimed at stimulating demand through fiscal policies, such as infrastructure investments.
- **Net Exports (X-M)**: Representing the trade balance, this component reflects external demand for domestic goods and services.

Keynesian Dynamics

In practice, these components interact within a dynamic system affected by external forces, such as changes in government policy, interest rates, or global economic conditions. The role of these exogenous factors has often been a subject of debate, especially when considering how external factors influence these domestic economic components.

- **Government Expenditures (G)** can directly inject demand into the economy through stimulus packages or public services.
- **Interest Rates** can influence **Investment (I)** by altering borrowing costs.
- **Sentiment Indices** (such as **Consumer Sentiment** and **Business Sentiment**) provide information on the willingness of households and firms to spend or invest, thereby influencing **C** and **I**.
- **Net Exports (X-M)** are impacted by exchange rates, trade policies, and global demand for domestic goods and services.

While the basic Keynesian model treats these relationships as linear and static, modern data-driven methods allow for more flexibility in modeling potentially complex dynamics.

Data Description

This study uses quarterly U.S. data from 2001 to 2023, gathered from Federal Reserve Bank of St. Louis. The data set consists of:

- **GDP (In Billions):** The target variable representing the U.S. Gross Domestic Product.
- **Consumer Sentiment:** A proxy for consumption, captured through surveys that measure consumer confidence.
- **Business Sentiment:** Reflected by business survey indicators, directly influencing private investment decisions.
- **Total Credit to Private Non-Financial Sector:** An indicator of financial conditions, representing the amount of credit available to businesses.
- **Capacity Utilization:** A measure of the economy's ability to produce, serving as a gauge of economic efficiency.

- **Government Expenditures:** A measure of total government spending at the federal level, including all government investments and expenditures.
- **Net Exports:** The difference between exports and imports, which reflects international trade dynamics.

The time series data are structured on a quarterly basis, ensuring consistency for time series modeling. All data were cleaned, transformed to a uniform time frequency, and merged into a single data set.

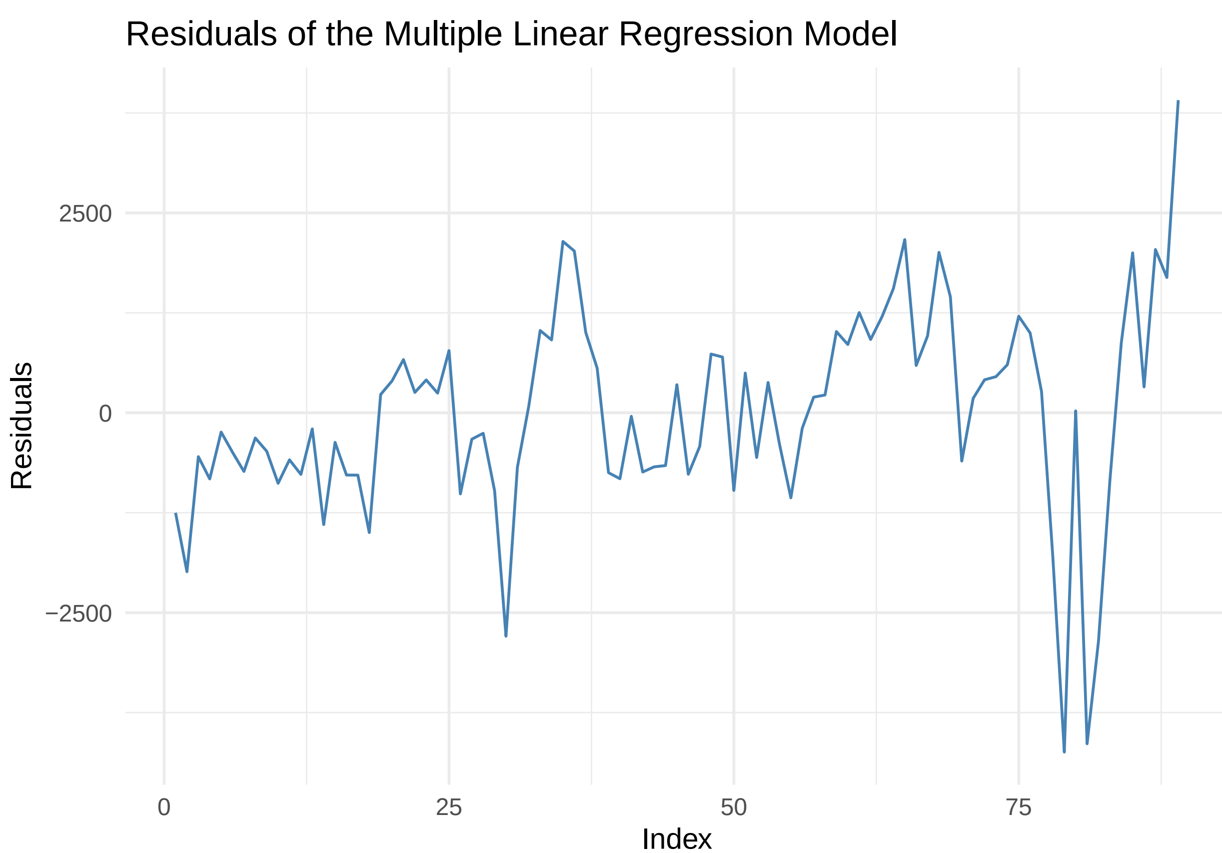
Methodology

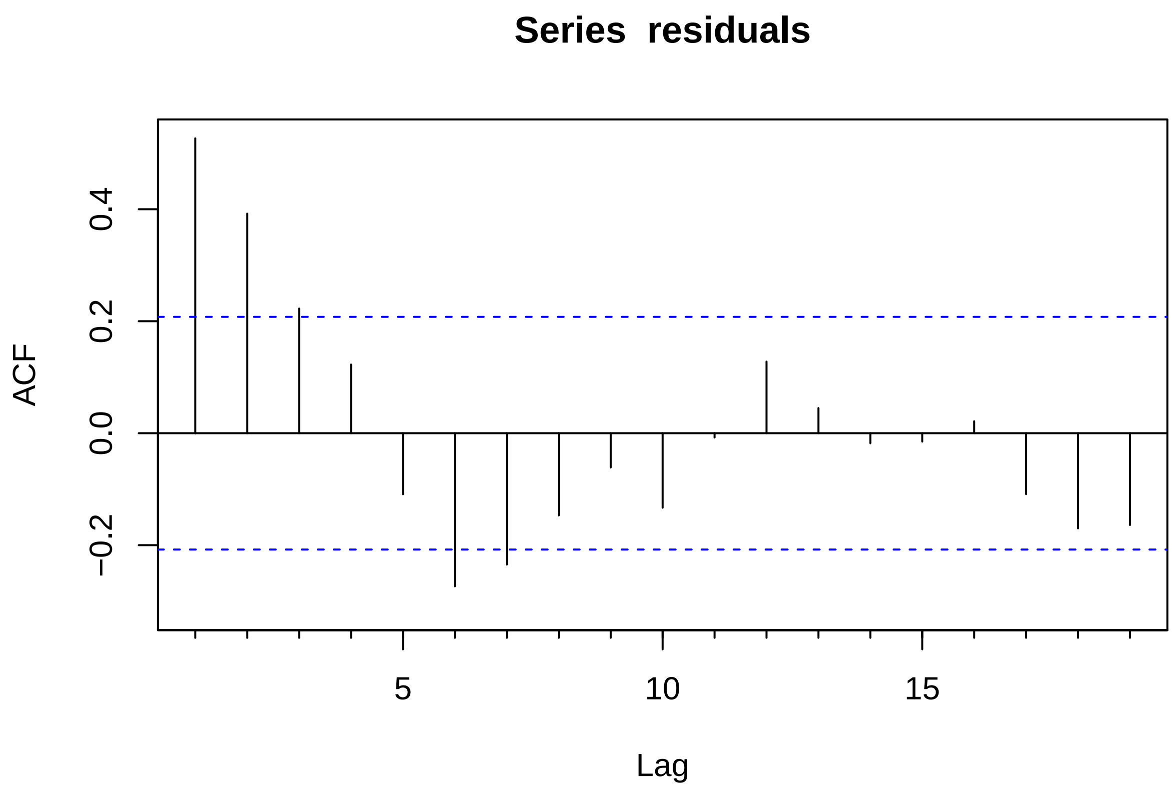
Multiple Linear Regression

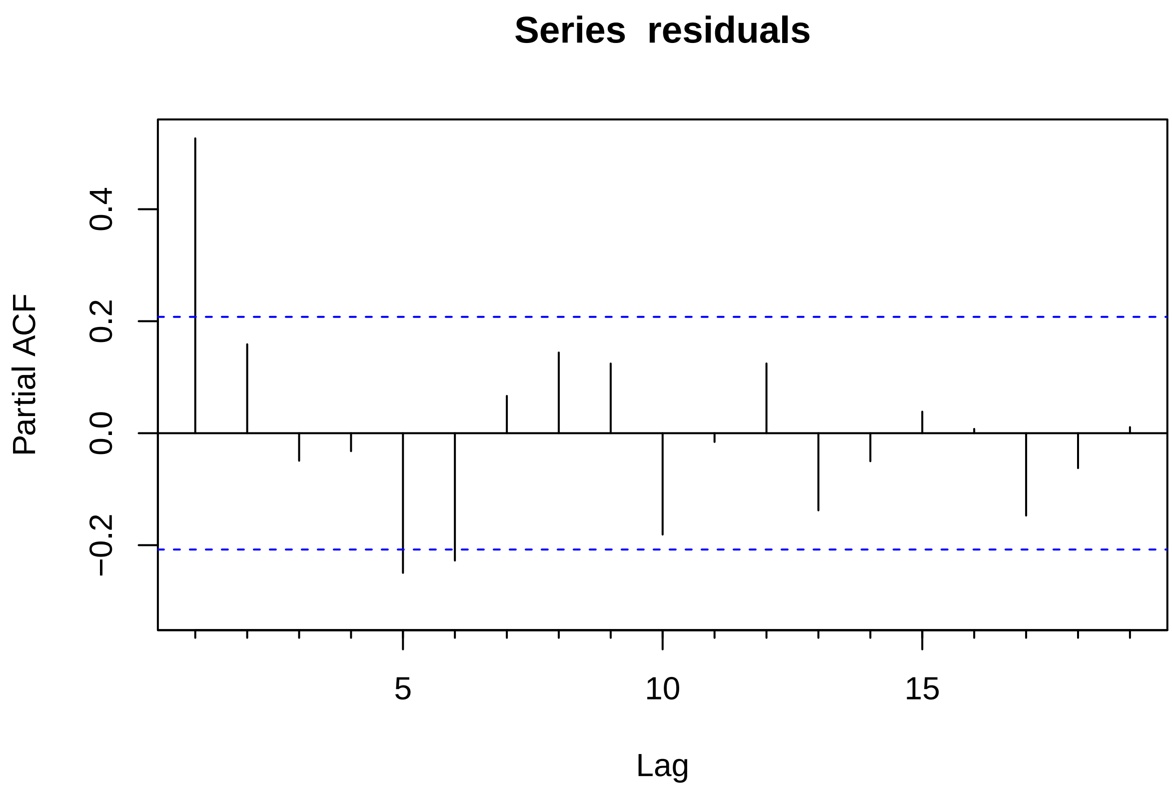
I first approached the modeling task by fitting a multiple linear regression model. This method assumes that the relationships between GDP and the exogenous variables are linear:

The model is developed with the following steps:

- **Stepwise Selection:** A stepwise approach is used to refine the model, iteratively adding or removing predictors based on their significance.
- **Diagnostic Tests:** To verify model assumptions, residual analysis is performed to check for autocorrelation and heteroscedasticity.




Series residuals
ACF
0.4
0.2
0.0
−0.2
5
10
15
Lag



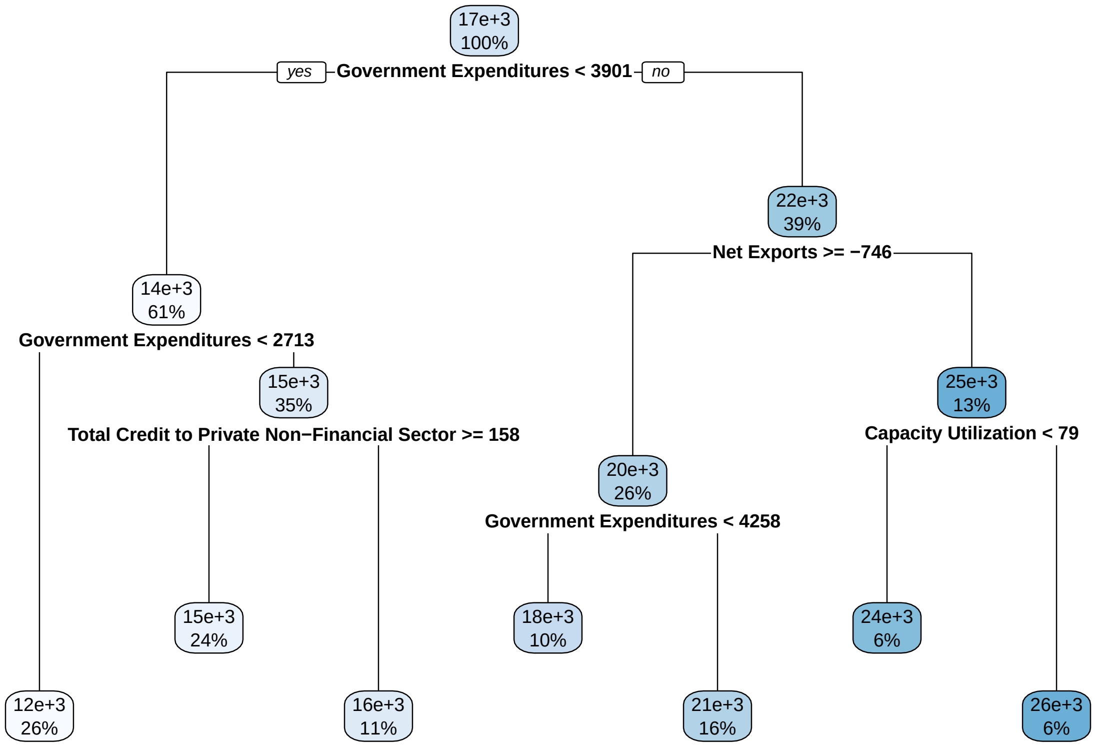
Random Forest

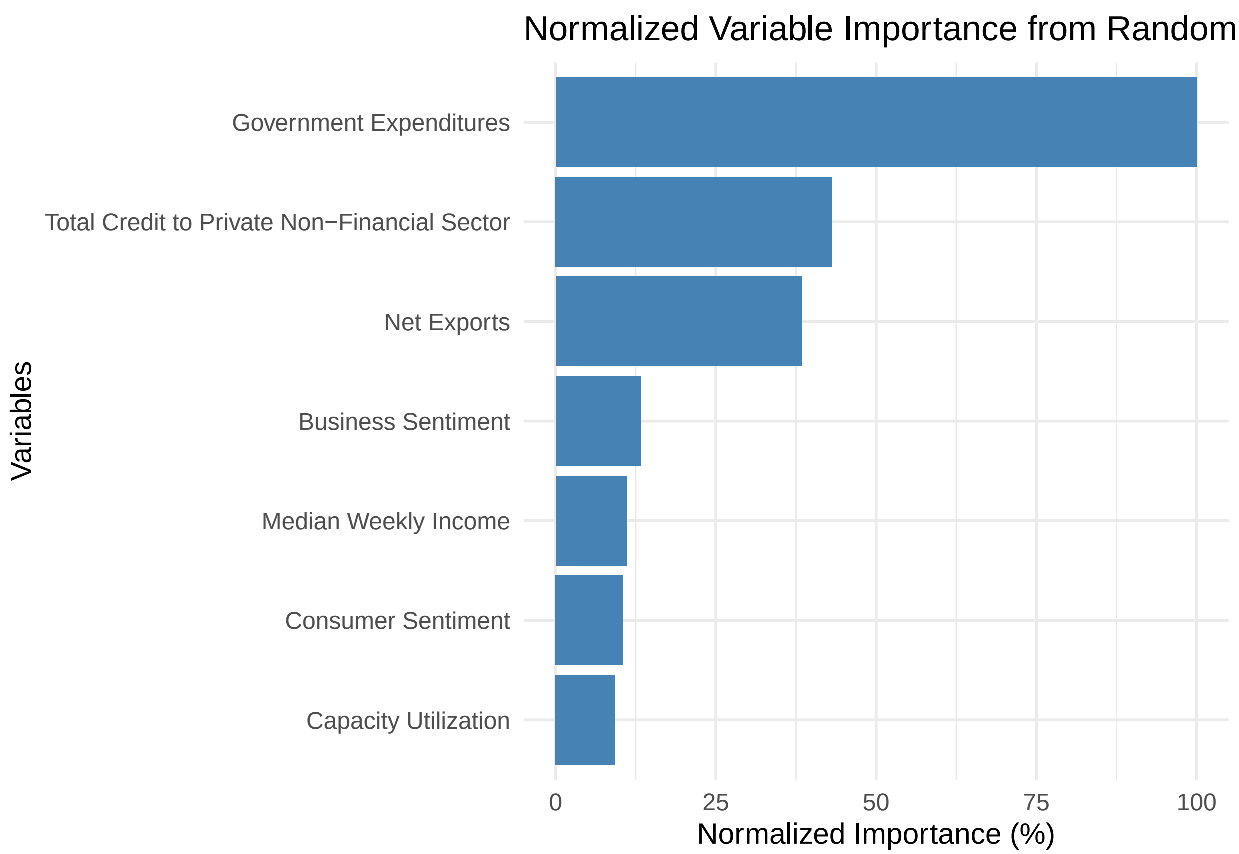
While the linear model provides insights, it may miss non-linear interactions between the predictors and GDP growth. To capture such complexities, a Random Forest model is applied.

Random Forest is a non-parametric ensemble machine learning method, which builds multiple decision trees based on random subsets of the predictors. Each tree contributes to the final prediction, which is aggregated by averaging the outputs of all trees. This approach is robust and able to model non-linear and interaction effects between variables. We train the Random Forest model on 70% of the data and test it on the remaining 30%, providing both predictive accuracy and variable importance estimates.

Key steps in the RF modeling process include:

- **Train-Test Split:** 70% of the data is used for training, and 30% is held back for testing.
- **Hyperparameter Tuning:** The number of trees (ntree) and the number of variables considered at each split (mtry) are set based on experimentation to minimize overfitting and maximize predictive accuracy.


17e+3
100%
yes
Government Expenditures < 3901
no
14e+3
61%
Government Expenditures < 2713
15e+3
35%
Total Credit to Private Non−Financial Sector >= 158
22e+3
39%
Net Exports >= −746
20e+3
26%
Government Expenditures < 4258
25e+3
13%
Capacity Utilization < 79
12e+3
26%
15e+3
24%
16e+3
11%
18e+3
10%
21e+3
16%
24e+3
6%
26e+3
6%



Results

Regression Model Results

The initial linear regression model provided an R-squared of **0.90**, indicating a strong relationship between GDP and the selected predictors. Following stepwise selection, variables such as **Capacity Utilization**, **Government Expenditures**, and **Median Weekly Income** remain significant drivers of GDP growth, whereas **Business Sentiment** showed no significant effect and was removed from the model.

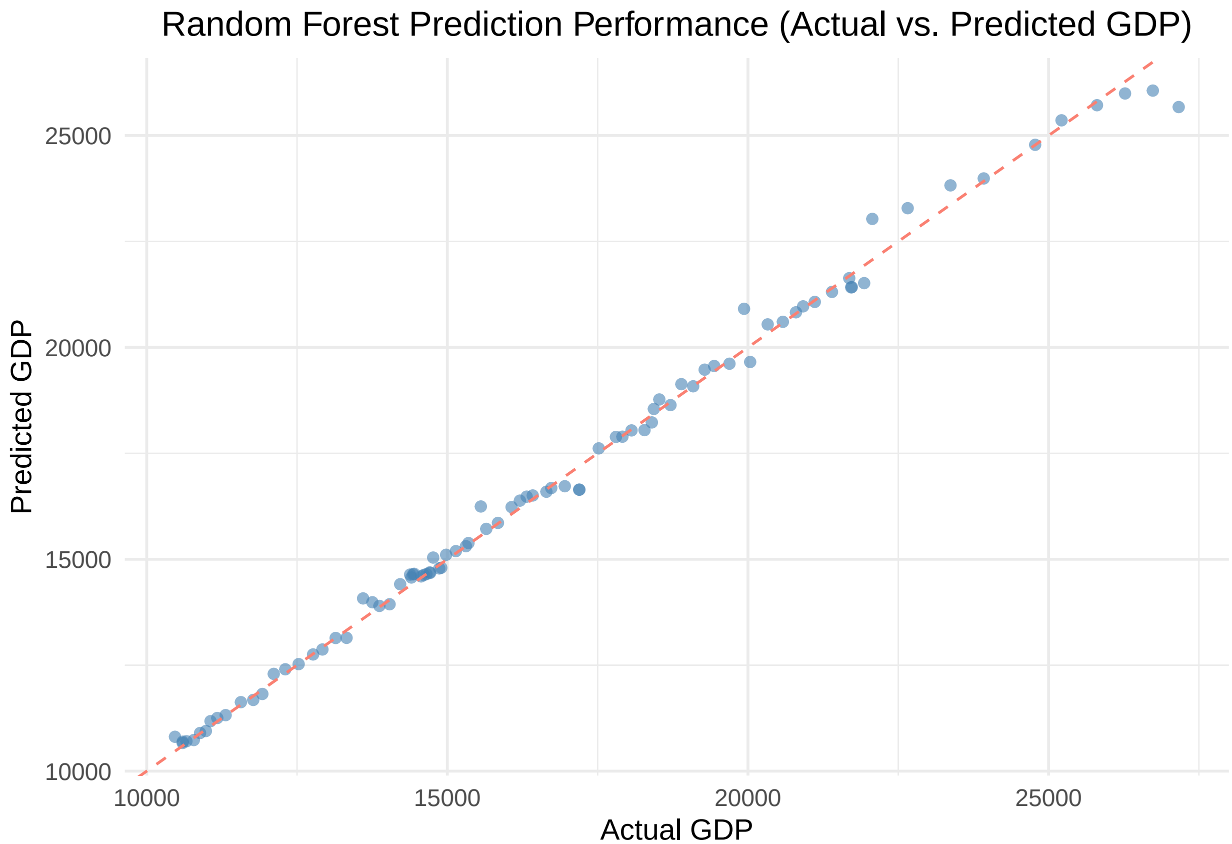
Residual Analysis: Residuals from the model were analyzed to ensure the adequacy of the fit. The ACF and PACF plots did not show significant autocorrelation, and the residuals displayed no obvious pattern, indicating that the model assumptions were reasonable.

Random Forest Model Results

The **Random Forest model** resulted in an **R-squared of 0.9791** significantly outperforming the regression model in terms of predictive accuracy. This indicates that the Random Forest model explains 97.91% of the variation in GDP.

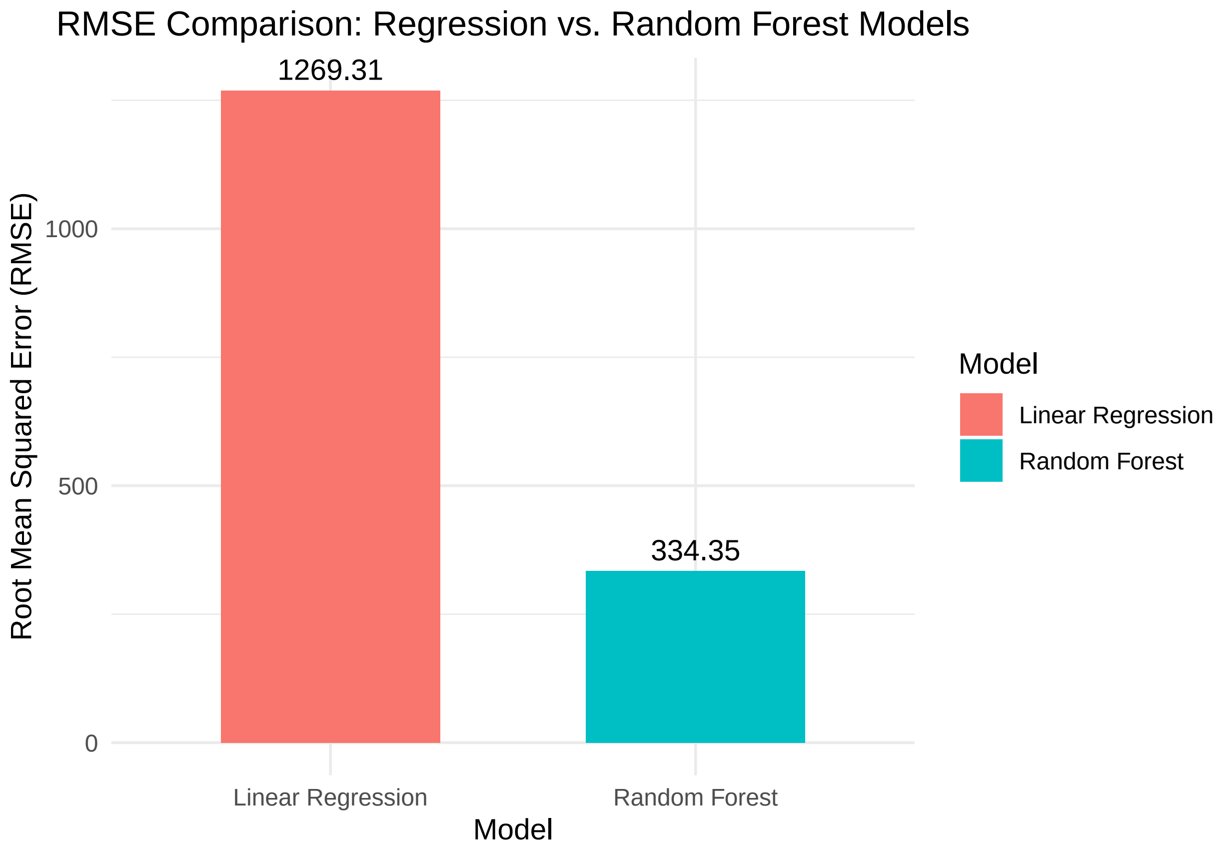
- **RMSE (Root Mean Squared Error):** 334.4, reflecting the typical deviation from the actual GDP values.
- **Variable Importance:** The Random Forest model confirmed that **Government Expenditures** and **Total Credit to Private Non-Financial Sector** were the most significant variables in predicting GDP, in line with the results of the regression model.

Variable Importance: The Random Forest technique also quantified the importance of each predictor, providing further insight into how different factors contribute to economic growth.



Model Comparison

Both models agree that **Capacity Utilization** and **Government Expenditures** are key drivers of GDP growth. However, the Random Forest model provides superior predictive performance, likely due to its ability to model complex, non-linear relationships and interactions that the linear regression model could not capture.



Discussion

The results of this study support the Keynesian framework, showing that **government spending** is the strongest determinants of GDP growth. This is consistent with Keynesian principles, where government expenditure acts as an essential source of demand, and efficient use of productive capacity can drive output.

Despite the linearity of the regression model, the **Random Forest** model provides an additional advantage by capturing complex, non-linear relationships, demonstrating the complementary value of machine learning techniques in macroeconomic forecasting.

Conclusion

This study successfully identifies the key determinants of U.S. GDP growth, showing that **capacity utilization** and **government expenditures** play crucial roles in driving economic output. Both traditional and machine learning approaches support these findings, with Random Forest offering superior predictive capability. These results provide valuable insights for policymakers, indicating that focusing on improving government spending and efficient utilization of capacity can significantly impact GDP growth. Future research may benefit from exploring other factors, such as interest rates and inflation, and further enhancing the Random Forest model through hyperparameter optimization and model ensembles.

Sources

- **Gross Domestic Product (GDP)**
  Federal Reserve Bank of St. Louis. (2024). *Gross Domestic Product (GDP)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/GDP
- **Consumer Sentiment (UMCSENT)**
  Federal Reserve Bank of St. Louis. (2024). *Consumer Sentiment (UMCSENT)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*UMCSENT*
- **Federal Funds Effective Rate (FEDFUNDS)**
  Federal Reserve Bank of St. Louis. (2024). *Federal Funds Effective Rate (FEDFUNDS)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*FEDFUNDS*
- **Business Sentiment (MACROBUS)**
  Federal Reserve Bank of St. Louis. (2024). *Business Sentiment (MACROBUS)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*MACROBUS*
- **Total Credit to Private Non-Financial Sector (QUSPAM770A)**
  Federal Reserve Bank of St. Louis. (2024). *Total Credit to Private Non-Financial Sector (QUSPAM770A)*[Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*QUSPAM770A*
- **Capacity Utilization (TCU)**
  Federal Reserve Bank of St. Louis. (2024). *Capacity Utilization (TCU)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/TCU
- **Government Expenditures (FGEXPND)**
  Federal Reserve Bank of St. Louis. (2024). *Government Expenditures (FGEXPND)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*FGEXPND*
- **Government Balance (M318501Q027NBEA)**
  Federal Reserve Bank of St. Louis. (2024). *Government Balance (M318501Q027NBEA)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/*M318501Q027NBEA*
- **Net Exports (NETEXP)**
  Federal Reserve Bank of St. Louis. (2024). *Net Exports (NETEXP)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/NETEXP