International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 12, Issue 5, May-2024, Available online at: www.ijaresm.com

Economic Recession Analysis and Prediction

Vansh Ashok Setpal[1], Saumya Sanjay Padmane[2], Pooja Hemant Gangurde[3], Rishabh Prakash Goray[4], Dr. T.Rajani Mangala[5]

[1,2,3,4]Department of Electronics and Comp Sci./VES Institute of Technology/Mumbai/Maharashtra/India
[5]Department of Electronics and Comp Sci./Professor/VES Institute of Technology/Mumbai/Maharashtra/India

---***************---

ABSTRACT

This study embarks on a thorough exploration of economic recessions, employing a comprehensive array of statistical tools and predictive modeling techniques to unravel the multifaceted complexities and implications inherent in these downturns. With a central objective of understanding the underlying causes, assessing the wide-ranging impacts, and exploring potential remedies, the project aims to provide invaluable insights for policymakers, economists, and businesses alike. Through a meticulous analysis of diverse economic indicators including Gross Domestic Product (GDP), unemployment rates, inflation, and consumer spending, coupled with the application of advanced statistical methodologies such as time series analysis, leading indicators, and regression models, the research endeavors to uncover the intricate patterns, causal relationships and dynamics that drive recessions. Furthermore, by integrating predictive modeling utilizing the Random Forest algorithm, the study seeks to forecast the likelihood of recessions in the years 2024, 2025, and 2026, thereby enhancing our understanding of future economic trends and potential downturns. Through this interdisciplinary approach, the project aspires to offer a holistic view of how economies respond to recessions and how policymakers can effectively intervene to mitigate their adverse effects. By providing evidence-based insights and actionable recommendations, the research aims to contribute to informed decision-making in economic policy, strategic planning, and business resilience, ultimately fostering sustainable economic growth and stability in the face of challenging economic environments.

Keywords— Random Forest, predictive modeling, recessions

INTRODUCTION

Economic recessions, characterized by a significant decline in economic activity, have far-reaching implications for societies, governments, and businesses. Understanding the causes, impacts, and potential remedies of recessions is crucial for policymakers and economists alike. In this study, we delve into the depths of economic recession analysis by harnessing the power of statistical tools. Over the years, statistical methods have proven to be indispensable for comprehending complex economic phenomena and providing evidence-based insights. From time series analysis to leading indicators and regression models, each tool serves a specific purpose in uncovering the intricacies of economic fluctuations. This analysis not only helps in gauging the severity of economic downturns but also assists in formulating informed policies to mitigate their adverse effects.

In this project, a range of statistical tools and their applications in the context of economic recession analysis are explored by employing various statistical techniques. We aim to shed light on the underlying patterns, factors, and dynamics that drive recessions. By combining these techniques, we aim to present a holistic view of how economies respond to recessions and how policymakers can effectively respond. The analysis will be based on a diverse set of economic indicators, such as Gross Domestic Product (GDP), unemployment rates, inflation, consumer spending, and more. Through rigorous statistical examination, we will attempt to discern the relationship between these variables, their potential casualties, and how they contribute to the onset and propagation of recessions.

Additionally, we have integrated predictive modeling using the Random Forest algorithm to forecast the possibility of recessions in the years 2024, 2025, and 2026. By leveraging machine learning techniques, we aim to enhance our understanding of future economic trends and potential downturns. Through this predictive analysis, we strive to provide valuable insights for policymakers and businesses to prepare and mitigate the impacts of potential recessions.

LITERATURE REVIEW

Sheridan Kamal(2021) investigates how supervised machine-learning methods can be used to forecast economic downturns. The study employs techniques like logistic regression, decision tree classifier, k-nearest neighbor classifier, and support vector classifier. The study's goal is to identify the most efficient model for predicting economic downturns by utilizing techniques such as train-test splitting and time-series cross-validation on both scaled and unscaled datasets. The precision-recall area under the curve (PR AUC) metric is used to assess the models' performance in accurately detecting recessions and reducing false alarms. The study also investigates how parameter adjustment affects the models' efficiency. Out of the various models examined, the optimized Support Vector Classifier that was trained on scaled data stood out as the top performer, obtaining a PR AUC score of 0.83. This discovery emphasizes the considerable promise of supervised machine learning, especially support vector classifiers, in enhancing forecasts of economic downturns. The findings of the research are particularly beneficial for policymakers, financial institutions, and investors, offering them improved resources to manage economic uncertainties effectively.

Terrence Zhang(2021) explores how machine learning techniques can be used to predict economic recessions. The research outlines the methodology, focusing on various supervised machine learning algorithms, including logistic regression, decision trees, k-nearest neighbors, and support vector machines. It assesses these algorithms' effectiveness in forecasting economic downturns, likely using performance metrics such as precision, recall, or the area under the curve. Additionally, the study examines the broader implications of its findings for policymakers, financial institutions, and investors. By providing a detailed analysis of machine learning's potential in enhancing recession predictions, the study contributes valuable insights to the field of economic forecasting methodologies, offering a new perspective on how advanced data techniques can aid in anticipating economic changes.

Jefferey Chen(2019) explores the impact of timely alternative data sources, like credit card transactions and search query trends, combined with advanced machine learning (ML) techniques on predicting economic indicators. The research evaluates which data and ML technique combinations most accurately forecast national economic activity. By examining various model specifications, data sources, and variable selection methods, the study assesses the influence of each on predictive accuracy. Results indicate that ensemble methods such as Random Forests are particularly effective, potentially reducing Personal Consumption Expenditure (PCE) revisions by an average of 12% compared to traditional national accounting methods. Despite the timeliness of alternative data, the study finds that conventional data, including employment figures and lagged dependent variables, still hold substantial predictive power. This highlights the necessity of balancing timeliness with the inherent value of traditional data to enhance economic forecasting accuracy.

Parag Verma(2021) examines the extensive impact of the COVID-19 pandemic on global political, social, economic, religious, and financial systems. The pandemic has led to considerable human suffering and economic turmoil, resulting in over 4.6 million screenings, millions of infections, and widespread closures of borders, businesses, and schools. This disruption has severely affected major economies, causing stock market crashes, significant declines in tax revenues, and a sharp downturn in global economic growth. The study analyzes the relationship between COVID-19 and economic growth, with a focus on GDP and key financial indicators such as the S&P 500, crude oil, gold, silver, natural gas, and treasury bonds. Using regression models and data from sources like Yahoo Finance, the IMF, and the Johns Hopkins COVID-19 map, the study finds a moderated positive correlation between the pandemic and economic variables. The research aims to assist policymakers, business strategists, and investors in navigating the crisis and forecasting future economic and stock market trends, thereby improving decision-making during these uncertain times.

ALGORITHM

Step 1- Data Collection and Preprocessing:
This is the first stage where data is gathered from various sources like FRED. The data is then formatted and cleaned to ensure its accuracy and consistency. This involves removing outliers, correcting errors, and ensuring consistency in the format of the data.

Step 2- Exploratory Data Analysis (EDA):
In this stage, we get a sense of the data by performing initial investigations. This involves tasks such as calculating summary statistics, creating visualizations of the data, and identifying patterns and trends.

Step 3- Analysis:
This stage involves performing a more in-depth analysis of the data using statistical methods and modeling techniques. This type of analysis helped us to find an accurate model for further prediction.

Step 4- Predictive Modeling:
The goal of the project is to make predictions about future events, a predictive model isbuilt using the data. This model then is used for future prediction.

Step 5- Evaluation and Validation:
Once a model is built or analysis is complete, it needs to be evaluated to assess its effectiveness. This may involve using techniques such as cross-validation to ensure that the model or analysis generalizes well to unseen data.

Step 6- Insights and Recommendations:
Insights and Recommendations: Finally, the results of the data collection and processing process are summarized and presented. This involves creating reports, visualizations, or dashboards that communicate the insights gained from the data.

WORKING

Analysis:
In this project, a combination of economic and financial data was utilized. Since the necessary features were not readily available in a single downloadable dataset from a variety of sources, each feature was individually downloaded and combined into a single dataframe using a Jupyter notebook with Python 3. To ensure data reliability and timeliness, economic data was sourced from FRED (Federal Reserve Economic Data from the Federal Reserve Bank of St. Louis) using a Python package called Quandl. Additionally, the S&P 500 index data was obtained from Yahoo! Finance, processed in Excel, and imported as a CSV file into the project's Jupyter notebook.

Once the dataframe was created, a correlation heatmap was generated to identify features for the final dataset. The aim was to select features that were not highly correlated with each other but still showed correlation with the target feature. Six features were chosen: the 10-Year Treasury Constant Maturity Rate, the difference between the 10-Year and 2-Year Treasury Constant Maturity Rates, the 1-Month Percentage Change of the Effective Federal Funds Rate, the 1-Month Percentage Change of All Employees, the Consumer Price Index for All Urban Consumers, and the 1-Year Percentage Change of the S&P 500 Index Price. These features covered the period from June 1, 1976, to March 1, 2021, resulting in 539 data points.

Data exploration revealed a class imbalance issue, with recession labels present in only 12.987% of the data, highlighting the need to address this imbalance during machine learning model interpretation. The dataset was then split into training (70%) and testing (30%) sets using the train_test_split function, covering the periods from June 1, 1976, to October 1, 2007, for training, and from November 1, 2007, to March 1, 2021, for testing. Cross-validation using TimeSeriesSplit was employed with GridSearchCV to optimize model parameters.

Four types of models (logistic regression, decision tree classifier, k-nearest neighbor classifier, and support vector classifier) were trained in four variations: unscaled/default, scaled/default, unscaled/tuned, and scaled/tuned. For the support vector classifier models, specific parameters were adjusted for reproducibility and probability calculation. After training the models, accuracy, classification reports, and confusion matrices were generated. Precision-recall curves (PR curves) and receiver operating characteristic (ROC) curves were plotted to evaluate model performance based on PR AUC and ROC AUC scores. This comprehensive analysis provided insights into the models' predictive capabilities for economic recessions.

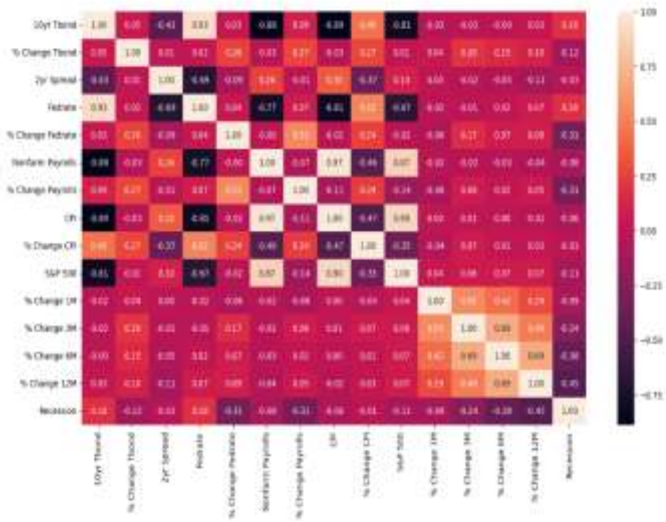

Figure 4.1: Heatmap Visualizing Correlations Between Economic Indicators and Potential Recessionary Triggers

This heat map depicts the correlation between the various economic factors and their potential influence on recessions. The color scale uses red for positive correlations, blue for negative correlations, and color intensity represents the strength of the correlation. Noticeable features include a strong positive correlation between changes in nonfarm payrolls and changes in the Consumer Price Index (CPI).This suggests that periods of increasing employment might coincide with rising inflation. Additionally, a strong negative correlation is observed between changes in stock prices and changes in unemployment rates. This implies that a decline in stock market performance might be linked with rising unemployment.

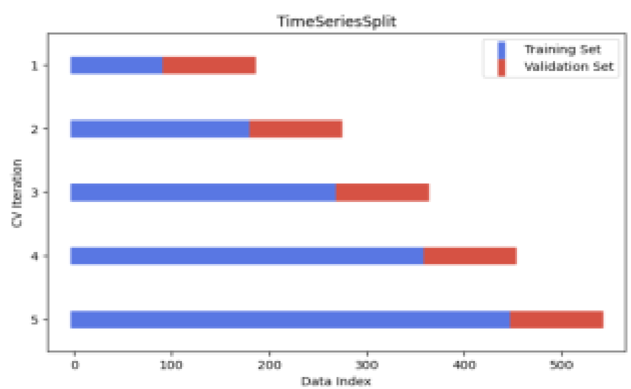

Figure 4.2: Graph of Time Series Split

The graph shows the performance of machine learning model over time. The number of samples processed for training and validation sets are plotted on the y-axis, while the x-axis represents the cross-validation iterations. This suggests the model is being evaluated using cross-validation techniques.

Prediction:

Prediction for 2023:

In this analysis, a classification model using the C5.0 and random forest algorithms was developed to predict economic recessions based on monthly data on recessions and yield curve inversions from 1975 to 2022. Using the Federal Reserve Economic Data (FRED) database, months with recessions in the United States were identified with an indicator variable. Yield curve data, reflecting the difference between the 10-year and 2-year treasury notes, was used to detect inversions, with an indicator variable set to 1 for months with at least one inversion. Rolling statistics were calculated using the rollmax() function to indicate a recession within the next 6, 12, or 18 months. The models were trained on this data, incorporating the monthly average yield and inversion indicator variables. The models' predictions were then compared to actual recession data up to 2022. Predictions from the C5.0 algorithm were shown in red, while those from the random forest algorithm were in blue. Both models showed spikes in predicted recession probabilities around September 2019 and March 2022. Looking forward, both models indicated a probability of recession within the next 18 months starting from April 2022, suggesting the need for continued observation to verify these predictions.

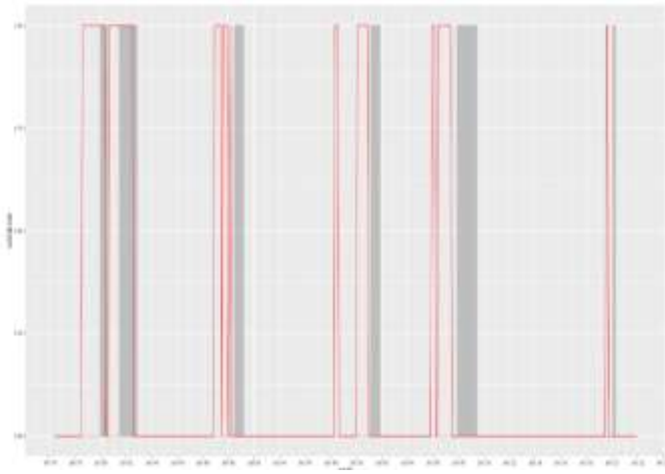

Figure 4.3: Prediction of Recession by Yield Curve Inversion

This Figure displays Recession, shaded in gray, indicate periods of economic contraction, while yield curve inversions, marked in red dashes, highlight instances where short-term interest rates exceed long-term rates, often preceding recessions.

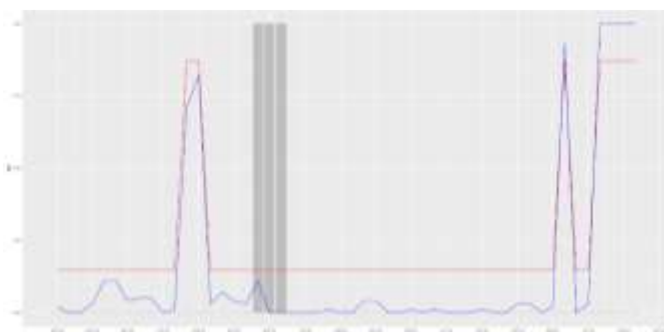

Figure 4.4: Prediction of Recession for the year 2023

Predictions from the C5.0 algorithm are represented in red and the ones from the random forest are represented in blue in the figure 5.4. Note how both models' predictions spiked around September 2019 and then again around March 2022. We can look at the estimated probabilities of a recession up to December of 2022. Notice that both models are flashing for a probability of a recession within the next 18 months starting from April of 2022. That means a recession is called anytime between then and October 2023.

Prediction for 2024 and 2025:
For predicting recessions in 2024 and 2025, historical data up to March 2023 was used to train a Random Forest classifier model. This data included monthly US recession indicators ('USRECM') and corresponding dates from 1854 to March 2023. After training, the model was used to forecast recession indicators for 2024 and 2025. For 2024, the model predicted no recession, whereas for 2025, it identified specific months as potential recession periods. These predictions were visualized using bar graphs, with the x-axis representing predicted outcomes and the y-axis representing the count of predictions. The identified recession months in 2025 provided insights into potential economic trends for the coming years.

International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 12, Issue 5, May-2024, Available online at: www.ijaresm.com

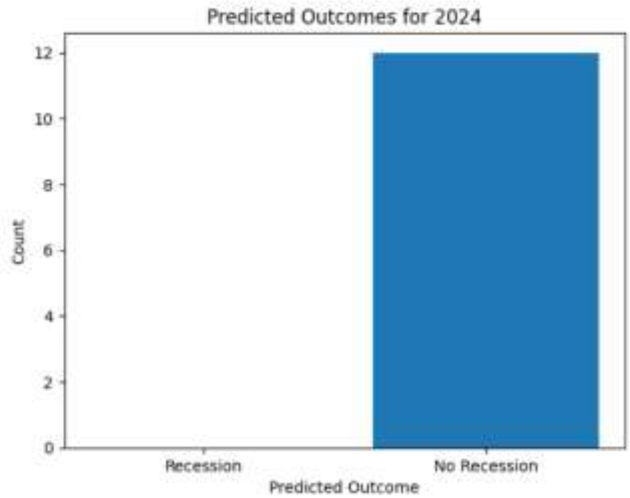

Figure 4.5: Prediction of Recession for the Year 2024

The graph depicts the number of months without a recession on the y-axis, illustrating a stable economic outlook for 2024. This visualization contrasts recessionary periods and non-recessionary periods corresponding to there is no recession 2024 as per prediction model.

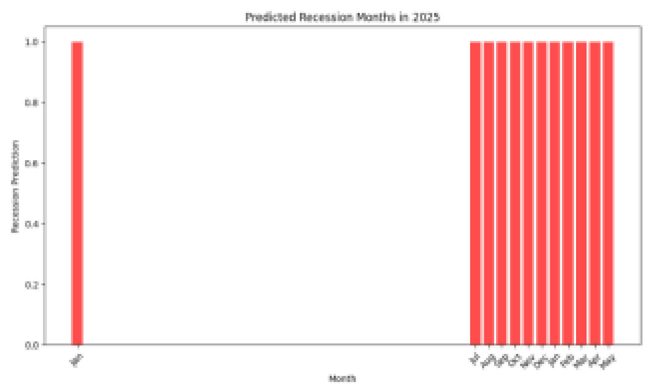

Figure 4.6: Prediction of Recession for the Year 2025-2026

Predicted Recession months in 2025.The y-axis shows the probability (from 0 to 1) of a recession in the figure 4.6. The x-axis shows the months of 2025. The line graph shows a steady increase in the probability of a recession throughout 2025, with the highest probability in December 2025 at around 80% till May 2026.

CONCLUSION

In light of recent events, such as the COVID-19 pandemic, it's evident that economic forecasting, while highly valuable, can still be susceptible to unforeseen external factors. The pandemic, which began in February 2020 and persists to this day, serves as a stark reminder of the challenges inherent in predicting economic recessions with absolute certainty.Despite the sophistication of machine learning algorithms and statistical tools utilized in this study, the unpredictable nature of global events necessitates caution when interpreting recession predictions. While models like the Decision Tree trained on scaled data exhibit promising performance, achieving a high PR AUC score, it's essential to acknowledge the inherent uncertainty in economic forecasting. Nonetheless, by continuously refining and updating these models with the latest economic data, we can better equip ourselves to navigate future economic challenges and foster resilience in the face of uncertainty. The project undertook an in-depth analysis of US economic recession indicators spanning from January 1854 to March 2023, aiming to predict potential recession periods for the years 2024 and 2025. Leveraging machine learning techniques, particularly the Random Forest classifier model, we trained and evaluated predictive models based on historical recession data.

1. Our findings revealed a stable economic outlook for 2024, with the model forecasting no recession during that year.

2. However, for 2025, the model identified specific months that could potentially witness economic downturns. These insights provide valuable foresight for policymakers, businesses, and investors to anticipate and prepare for potential economic challenges.

3. Additionally, the visualizations produced throughout the analysis facilitated a clear understanding of the model's predictions, enhancing the interpretability of the results.
Moving forward, continued refinement and validation of predictive models using updated data will be essential to ensure the reliability and accuracy of future economic forecasts.

REFERENCES

[1] Kamal, Sheridan, “An Analysis of Machine Learning Techniques for Economic Recession Prediction” (2021).CUNY Academic Works.
[2] Zhang, Terrence. “Recession Prediction Using Machine Learning.” Medium, Towards Data Science, 11 Jan. 2021,
[3] Chen, J. et al. “Off to the Races: A Comparison of Machine Learning and Alternative Data for Predicting Economic Indicators.” (2019).
[4] Ron Martin, “The local geographies of the financial crisis: from the housing bubble to economic recession and beyond”, Journal of Economic Geography, Volume 11, Issue 4, July 2011.
[5] Yılmaz Akyüz, “Global Economic Prospects:The Recession May Be Over But Where Next?”, South Centre, 2010.
[6] Parag Verma, Ankur Dumka, Anuj Bhardwaj, Alaknanda Ashok, Mukesh Chandra Kestwal, Praveen Kumar, “A Statistical Analysis of Impact of COVID19 on the Global Economy and Stock Index Returns”, January 2021.
[7] Brownlee, Jason. “ROC Curves and Precision-Recall Curves for Imbalanced Classification.”Machine Learning Mastery, Machine Learning Mastery, Sept.2020
[8] Nyman, Rickard, and Paul Ormerod. “Predicting Economic Recessions Using Machine Learning Algorithms.” ArXiv.org, ArXiv: General Finance, January 2017