Predicting Economic Slowdowns and Clustering Global Economies

Rachel Elizabeth Washington
University of Colorado Boulder
Boulder, USA

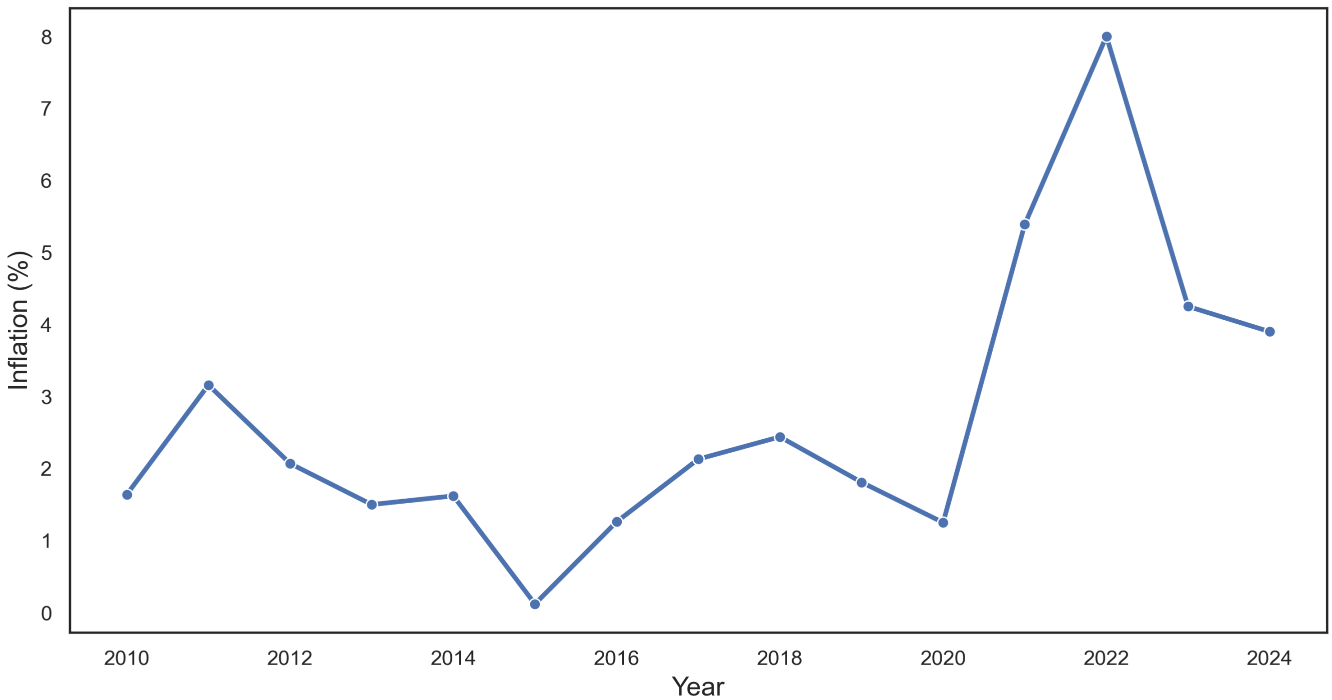

Figure 1: United States inflation rate between 2010 and 2024.

ABSTRACT

Predicting economic slowdowns or recessions is of upmost importance to policy makers to accurately adjust monetary policy. There are several features that economists currently use to predict economic slowdowns including unemployment rate, inflation rates, level of inflation, and GDP. These metrics can provide good data to try and predict economic slowdowns and the probability they can happen. The federal reserve aims for an inflation rate around 2% [2], which signals maximal employment and price stability. However, a high or low inflation rate can lead to significant economic instability such as low unemployment and poor economic growth [3]. The economies of different countries can also vary significantly in economic strength and whether the country experiences an economic slowdown. To better understand economies and predict inflation I will use a two-pronged approach. I will utilize logistic regression and random forest to determine GDP slowdowns based on inflationary trends. I will also use k-means clustering and hierarchical clustering to group countries based on GDP and inflation rates. My findings were that overall hierarchical clustering and logistic regression showed similar trends and provided insights on if economic slowdown was likely in 2025. Random forest and k-means clustering did not show a similar correlation like hierarchical clustering and logistic regression. Furthermore, for several major economies random forest and hierarchical clustering showed significant differences in if an economic slowdown would occur in 2025.

Data Mining Project, University of Colorado Boulder, 2025
.

KEYWORDS

Inflation, GDP, economic slowdown, K-means Clustering, Logistic Regression, Hierarchical Clustering, Random Forest

1 INTRODUCTION

Economic slowdowns significantly impact the global economy. Predicting economic slowdowns utilizing past data can help prevent economic slowdowns. Inflation rates specifically can be great indicators of whether an economic slowdown is imminent. Further my clustering economies together it can be determined what countries are most at risk of an economic slowdown.

Nonlinear relationships between inflation rate and economic growth typically are not captured by traditional macroeconomic models such as the Taylor Rule. Using machine learning techniques, we can better understand relationships between economic slowdown probability and inflation. Previous work has utilized supervised learning to also predict economic slowdown, with specific focus on recession and stock market crash. For example, Malladi utilized supervised learning techniques to determine if they could predict the economic crash due to COVID19 [6]. Bluwtein and coworkers also utilized supervised machine learning techniques to predict economic crashes using credit growth [1]. Similarly, Lekaj also utilized machine learning techniques to understand the relationship between inflation, yield curve, and the next recession [5].

My work expands upon this established body of work. Specifically, my work expands by looking at a global scale and predicting if each individual economy is likely to face an economic slowdown in 2025. Furthermore, I grouped individual global economies together utilizing unsupervised machine learning techniques. Overall I found that logistic regression and hierarchical clustering showed significant agreement in results and grouped countries similarly. This indicates good agreement on if an economic will have an economic slowdown in 2025. However, little agreement was reached between random forest and k-means clustering. However, results from hierarchical clustering and logistic regression show promise. Future work could be expanding the dataset to include more years of economic data to improve the models predictability.

2 LITERATURE REVIEW

Several researchers before me have done projects looking at using supervised machine learning techniques to determine economic slowdowns. Malladi used data to determine the COVID-19 U.S. recession and stock market crash [6]. Malladi states that machine learning algorithms specifically predicted the crash of the S&P 500 2 months before it happened. However, machine learning algorithms did not predict the crash before March 2021. Furthermore, the machine learning techniques have a three times higher false discovery rates of recessions compared to actual crashes. Indicating that significant work still needs to be done in the field. To predict this stock market crash Malladi specifically utilized 13 algorithms in total including – ensemble learning, linear support vector machine, k-nearest neighbors, discriminant analysis, binary decision tree, generalized linear regression model, naïve Bayes, decision trees, logistic regression, gradient boosting, ADABoost, random forest, and XG Boost. Overall, each model provided a 3-month forecast into the economic forecast. The conclusion was that the S&P 500 was the feature that was the greatest predictor of stock market crash.

Other research by Bluwstein and coworkers [1] look specifically how credit growth and the yield curve can be used as predictors in supervised machine learning models to predict recession. Bluwstein and colleagues mention that economic slowdown might not be able to be avoided but they could be mitigated. Furthermore, that machine learning models currently struggle to pick up sudden economic shocks such as the COVID19 economic slowdown. They utilized logistic regression, decision trees, random forests, extremely randomized trees, support vector machines, and artificial neural networks. They found that overall, these machine learning techniques had significant predictive power and overall decision tree-based models had the best predictive power. credit relative to GDP and a flat or inverted yield curve. They also mention that machine learning models are so powerful because they pick up on nonlinearities present in data.

Paz-Marín and coworkers used hierarchical clustering to determine a country's progress to a knowledge economy [4]. They mention that knowledge creation is a key factor in economic advantage in times of economic crisis. First, they built clusters using hierarchical clustering. Then Paz-Marín and their colleagues utilized those clusters to predict if the advancement of countries to a knowledge economy. They identified four main clusters advanced knowledge economies, follower knowledge economies, moderated knowledge economies, and early knowledge economies. Overall, their methodology successfully indicated a knowledge economy with similar classification as the World Bank. Furthermore, this tool could prove useful in several ways. It allows economies that are early in the transition to a knowledge economy to be pinpointed, and successful classification that could aid stakeholders and experts. My work is built upon this previous work but utilizes a two-pronged approach and aims to utilize a single predictor (inflation rate) to determine an economic slowdown.

3 METHOD AND PROPOSED WORK

For this project, two main methods - logistic regression and random forest were utilized to predict if 2025 will have an economic slowdown based previous data. K-means clustering and hierarchical clustering were utilized to cluster different global economies together. Data for this project will be obtained from Kaggle [8], this data source was cross referenced for accuracy to multiple sources. The dataset includes data from 19 different countries and has data starting from 2010 and ending in 2024. All machine learning techniques are implemented using sci-kit learn [7]. The aim of my work is to build upon previous work but utilize a two-pronged approach and a single predictor (inflation rate) to determine if an economic slowdown will occur in 2025 based on a specific inflation rate.

3.1 Logistic Regression and Random Forest

For logistic regression and random forest data was cleaned into individual tables such that each country had its own data frame. Then years were counted as either a slowdown or not based on the percentage of economic growth or GDP growth. If a country had a growth of less than 2% it was considered a year of economic slowdown, otherwise it was seen as a year of economic prosperity or normality. Both if inflation rate was low and if inflation rate was high was tested for predictions. If an economic slowdown would occur if the inflation rate was 1.5% or 3.5% if an economic slowdown would occur. To test the success of the model precision, recall and accuracy will be reported. Data will be split into test and training data using a 0.8/0.2 split. The results will be if the inflation rate was either 1.5% or 3.5% if an economic slowdown will happen for that country and what the overall probability is.

3.2 K-means Clustering

For k-means clustering the approach was to cluster data based on the features of GDP, inflation rate, unemployment rate, and economic growth or GDP rise. Then the data will be scaled using

Table 1: Predicted Economic Slowdowns for 2025 by Logistic Regression if Inflation Rate is High

Country	2025 Inflation Rate (%)	Probability of Slowdown (%)	Predicted Slowdown (1=Yes, 0=No)
USA	3.5	20.08	0
Japan	3.5	94.63	1
Germany	3.5	3.23	0
India	3.5	8.80	0
UK	3.5	67.23	1
Canada	3.5	33.93	0
Russia	3.5	58.76	1
Australia	3.5	19.38	0
France	3.5	65.57	1
South Korea	3.5	16.52	0
Saudi Arabia	3.5	66.67	1
Brazil	3.5	56.42	1
Italy	3.5	81.60	1
Malaysia	3.5	0.73	0
Pakistan	3.5	7.94	0
China	3.5	7.94	0
Bangladesh	3.5	7.94	0
Indonesia	3.5	7.94	0
Turkey	3.5	7.94	0

Table 2: Predicted Economic Slowdowns for 2025 by Random Forest if Inflation Rate is High

Country	2025 Inflation Rate (%)	Probability of Slowdown (%)	Predicted Slowdown (1=Yes, 0=No)
USA	3.5	63.00	1
Japan	3.5	100.00	1
Germany	3.5	0.00	0
India	3.5	0.00	0
UK	3.5	37.00	0
Canada	3.5	3.00	0
Russia	3.5	94.00	1
Australia	3.5	0.00	0
France	3.5	36.00	0
South Korea	3.5	55.00	1
Saudi Arabia	3.5	66.42	1
Brazil	3.5	97.00	1
Italy	3.5	97.00	1
Malaysia	3.5	0.00	0
Pakistan	3.5	0.00	0
China	3.5	0.00	0
Bangladesh	3.5	0.00	0
Indonesia	3.5	0.00	0
Turkey	3.5	0.00	0

sci-kit learn's scaler. After that data will be fit k-means model and clustered into 3 groups.

3.3 Hierarchical Clustering

For hierarchical clustering data was clustered using the following features - based on the features of GDP, inflation rate, unemployment rate, and economic growth or GDP rise. Data was scaled using sci-kit learn's scaler. Model was run and a dendrogram was created (Figure 3).

4 EVALUATION METRICS

To evaluate logistic regression and random forest precision, recall and accuracy were used. These techniques are used to determine if logistic regression and random forest can accurately capture differences between false negatives and false positives. All data was

Table 3: Predicted Economic Slowdowns for 2025 by Logistic Regression if Inflation Rate is Low

Country	2025 Inflation Rate (%)	Probability of Slowdown (%)	Predicted Slowdown (1=Yes, 0=No)
USA	1.5	29.57	0
Japan	1.5	92.72	1
Germany	1.5	9.33	0
India	1.5	9.15	0
UK	1.5	65.69	1
Canada	1.5	47.51	0
Russia	1.5	45.25	0
Australia	1.5	11.40	0
France	1.5	80.77	1
South Korea	1.5	16.74	0
Saudi Arabia	1.5	66.67	1
Brazil	1.5	54.65	1
Italy	1.5	97.44	1
Malaysia	1.5	6.22	0
Pakistan	1.5	7.84	0
China	1.5	7.84	0
Bangladesh	1.5	7.84	0
Indonesia	1.5	7.84	0
Turkey	1.5	7.84	0

Table 4: Predicted Economic Slowdowns for 2025 by Random Forest if Inflation Rate is Low

Country	2025 Inflation Rate (%)	Probability of Slowdown (%)	Predicted Slowdown (1=Yes, 0=No)
USA	1.5	64.00	1
Japan	1.5	100.00	1
Germany	1.5	0.00	0
India	1.5	0.00	0
UK	1.5	28.00	0
Canada	1.5	89.00	1
Russia	1.5	27.00	0
Australia	1.5	1.00	0
France	1.5	33.00	0
South Korea	1.5	0.00	0
Saudi Arabia	1.5	66.42	1
Brazil	1.5	97.00	1
Italy	1.5	100.00	1
Malaysia	1.5	0.00	0
Pakistan	1.5	0.00	0
China	1.5	0.00	0
Bangladesh	1.5	0.00	0
Indonesia	1.5	0.00	0
Turkey	1.5	0.00	0

split 0.8/0.2 into training and testing data respectively. For k-means clustering and hierarchical clustering I will determine if the clusters are spread out or are close together and if meaningful data can be extrapolated from the clusters formed. Furthermore, similarities between the clustering and supervised techniques was examined to further determine economic patterns.

5 RESULTS

5.1 Logistic Regression

Overall logistic regression was able to capture predict if there would be an economic slowdown in 2025 if inflation rate was 3.5% and if i. Furthermore recall, accuracy, precision and F1-score were all 1. This indicates that future work should expand this dataset further to see if accuracy, precision, recall and F1-score would change. However,

Table 5: Classification metrics for Logistic Regression

Metric	Value
Accuracy	1.00
Precision	1.00
Recall	1.00
F1-Score	1.00

Table 6: Classification metrics for Random Forest

Metric	Value
Accuracy	1.00
Precision	1.00
Recall	1.00
F1-Score	1.00

logistic regression had similar results whether the inflation rate was high or low. Some countries did vary significantly like Malaysia and France. But most countries saw marginal differences in probabilities. Indicating that inflation rate did not heavily influence the prediction of economic slowdown in 2025 for the majority or countries. Overall this work can be expanded with a larger dataset. The high recall, accuracy, precision and F1-score were most likely because each country only had 15 data points. Expanding this research to include a larger dataset could provide even more insightful findings and improve predictor power.

5.2 Random Forest

Random forest was able to predict economic slowdown similar to logistic regression. However, random forest gave more definitive predictions. Logistic regression probability was always under 100% or over 0%. However, random forest predicted several countries probability for recession was 100% and 0%; including Italy and Japan for 100% and China, Pakistan, and Turkey as 0%. There were also countries that had much higher or lower predictions than logistic regression. Noticeably, probability of slowdown in France was approximately 80% for logistic regression but was only approximately 30 % for random forest. Similarly the United States had a probability of approximately 30% for logistic regression but was at 64% for random forest. Similar to logistic regression however, inflation rate being high or low seemed to have little impact on the probability for an economic slowdown. Additional research can be done to further these economic predictions and get even better insights on the possibility for economic recession.

5.3 K-means Clustering

Overall k-means Clustering has three distinct groups in Figure 2. Larger economies with greater GDP and overall lower inflation rates were in one cluster. Countries with smaller economies and greater inflation rates were in another cluster. However, looking at the average of clusters across countries the clusters were not very pronounced. Turkey a country on its own was identified as the weakest economy with the highest inflation and lowest GDP. Where the United States was seen as the richest economy with the highest

Table 7: Country Cluster Assignments by K-means Clustering

Country	Assigned Cluster
Turkey	0
Australia	1
South Korea	1
Saudi Arabia	1
Russia	1
Pakistan	1
Malaysia	1
Japan	1
UK	1
Italy	1
India	1
Germany	1
France	1
China	1
Canada	1
Brazil	1
Bangladesh	1
Indonesia	1
USA	2

GDP. All other countries were in the middle. K-means clustering was unable to significantly separate countries into more distinct clusters. I predict it is because Turkey is such a distinct country and the United States is also treated by the algorithm as such a distinct country that further separation cannot occur. Additional cluster groups were attempted however similar results were obtained with additional clusters and resulted in empty cluster buckets.

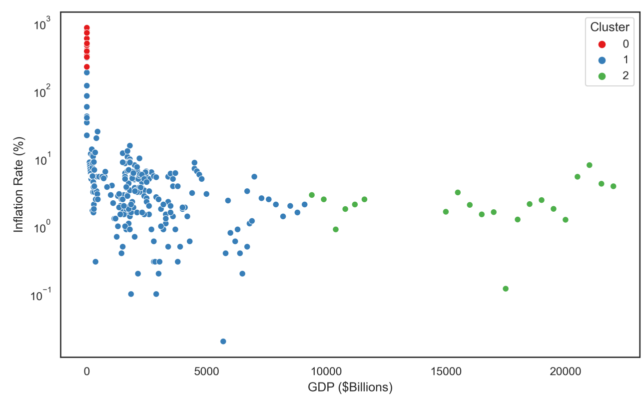

Figure 2: K-means clustering using multiple features and comparing inflation to GDP plotted on a log-scale

5.4 Hierarchical Clustering

Hierarchical clustering showcased some interesting findings and interestingly lined up well with logistic regression. Hierarchical clustering grouped Indonesia, Bangladesh, India and Malaysia together. All these countries had probabilities of slowdowns within 3%

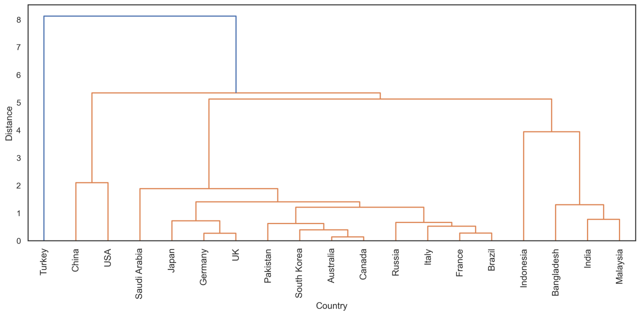

Figure 3: Hierarchical Clustering graphs showing groups of countries that are closely related

of each other for low interest rates. Where, France and Brazil were grouped together they had also had similar interest rates for logistic regression when the interest rate is high. There were many relationships that were similar to logistic regression. If the dataset was larger using hierarchical clustering and logistic regression could prove a useful tool in predicting economic slowdowns in the future.

6 TIMELINE

My timeline was to do my proposal work the first week. The second week was to complete all my calculations and project checkup. The third week was to complete my final report and presentation as well as any additional calculations. My goal to complete this project by the deadline was met. All goals were not met in the proposed timeline but the project has been completed. Future work can be completed to get better predictions with the models utilized here.

7 CONCLUSION

Useful insights on the possibility of economic slowdown in 2025 for different countries was obtained. I successfully utilized logistic regression and random forest to determine the probability an economic slowdown would occur based on a given inflation rate of either 1.5% or 3.5%. However, more data points for each country is needed to gain further insights and more accurate results. I also utilized hierarchical clustering and k-means clustering to further see if countries could be clustered by several economic features and if that would match the supervised learning models I created. Hierarchical clustering did show some similarities showing promise of using a two-pronged approach with logistic regression and hierarchical clustering to predict future economic slowdowns.

8 FUTURE WORK

Future work would be to utilize a larger dataset and expand the calculations. Utilizing the technique of logistic regression and hierarchical clustering show great promise as a two-pronged approach in understanding economies. Logistic regression could be especially useful when it comes to using the single predictor in inflation rate to obtain the probability of recession for a specific global economy.

REFERENCES

[1] Kristina Bluwstein, Marcus Buckmann, Andreas Joseph, Sujit Kapadia, and Özgür Sımşek. 2023. Credit Growth, the Yield Curve and Financial Crisis Prediction: Evidence from a Machine Learning Approach. *Journal of International Economics* 145 (2023), 103773.

[2] Board of Governors of the Federal Reserve System. 2023. What is inflation and how does the Federal Reserve evaluate changes in the rate of inflation? https://www.federalreserve.gov/faqs/economy_14400.htm Accessed: 2025-04-28.

[3] Congressional Research Service. 2024. Overview of U.S. Inflation and the Federal Reserve's Response. https://www.congress.gov/crs-product/R47479 Accessed: 2025-04-28.

[4] Mónica de la Paz-Marín, Pedro Antonio Gutiérrez, and César Hervás-Martínez. 2015. Classification of Countries' Progress Toward a Knowledge Economy Based on Machine Learning Classification Techniques. *Expert Systems with Applications* 42, 1 (2015), 562–572.

[5] D. Lekaj. 2023. The Federal Reserve's Monetary Policy, Inflation, the Inverting Yield Curve, and the Next Recession: A Machine Learning Framework for Evaluating Recession Risk. https://digitalcollection.zhaw.ch/handle/11475/28313. Accessed: 2025-04-28.

[6] R.K. Malladi. 2024. Application of Supervised Machine Learning Techniques to Forecast the COVID-19 US Recession and Stock Market Crash. Forthcoming, Computational Economics.

[7] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

[8] Adil Shamim. 2023. Economic Indicators and Inflation. https://www.kaggle.com/datasets/adilshamim8/economic-indicators-and-inflation. Accessed: 2025-04-26.