Predicting Recessions Using Economic Indicators and Logistic Regression

1. Introduction

A recession is defined as a period of economic decline in GDP, employment, and trade, lasting for at least six months. Recession can lead to significant reductions in consumer and business spending and overall economic activity. Recessions are a natural part of the economic cycle, but they can have harsh impacts on societies, affecting employment, stability, and more (IMF).

Predicting recessions allows policymakers to implement measures to mitigate economic downturns. Understanding the early signs of a recession allows governments and central banks to take proactive steps to stabilize the economy, such as adjusting interest rates, or enacting regulatory changes (Investopedia). Predicting recession can help businesses make informed decisions regarding investments resource allocation as well. Accurate recession predictions enable companies to anticipate changes in market conditions, helping them to manage risks more effectively and seize opportunities when economic conditions improve (Investopedia). Investors can adjust their portfolios to manage risks and capitalize on potential opportunities (NBER). Investors can adopt strategies that protect their assets during downturns, such as shifting to safer investments or diversifying their portfolios to reduce exposure to riskier sectors by predicting recessions.

2. Data Collection

We utilized the Federal Reserve Economic Data (FRED) database to collect various economic indicators that are potential predictors of a recession. FRED offers a comprehensive repository of economic data, providing reliable and timely information that is essential for economic analysis and forecasting. Using FRED ensures that our data is gathered from a well-maintained platform, making our research credible.

During our research, we used a variety of economic indicators to train and test our model, however the main indicators to predict recession turned out to be manufacturing purchasing managers' index (PMI), the federal funds rate, building permits, and the unemployment rate.

PMI Growth indicates the health of the manufacturing sector. The PMI is a key indicator of manufacturing activity, reflecting changes in new orders, production levels, supplier deliveries, and inventory levels. Growth in PMI suggests expanding manufacturing activity, which is typically associated with economic growth. The Federal Funds Rate reflects monetary policy, and influences interest rates across the economy. Lagged data on the federal funds rate helps to capture the delayed effects of monetary policy changes on economic activity. Building Permits indicate construction activity. Building permits are a leading indicator of future construction activity and economic expansion. An increase in building permits typically signals confidence in the economy and expectations of future growth. The Unemployment Rate reflects labor market conditions. The unemployment rate is a critical measure of economic health, with rising unemployment often preceding or accompanying economic downturns. Lagged and growth data on unemployment provide insights into the dynamics of the labor market and its impact on the broader economy.

3. Methodology

We used a logistic regression model to accurately and confidently predict a recession. A logistic regression model estimates the probability of an outcome based on one or more predictor variables. It is well-suited for this study as we aim to predict the binary outcome of whether a recession will occur.

Our approach involved feature engineering to create lagged variables and growth rates, data processing to handle missing values and scale predictors, and model training and evaluation. Feature engineering captured complex relationships between these indicators by creating new features that better represent economic conditions. We first collected and cleaned the data, created lagged variables and calculated growth rates, and removed highly correlated predictors to minimize multicollinearity. Next, we created and trained a logistic regression model with our selected predictors and evaluated the model's performance using test data. Finally, we identified the most important predictors based on feature importance, trained a simplified model using only significant predictors, and evaluated its performance.

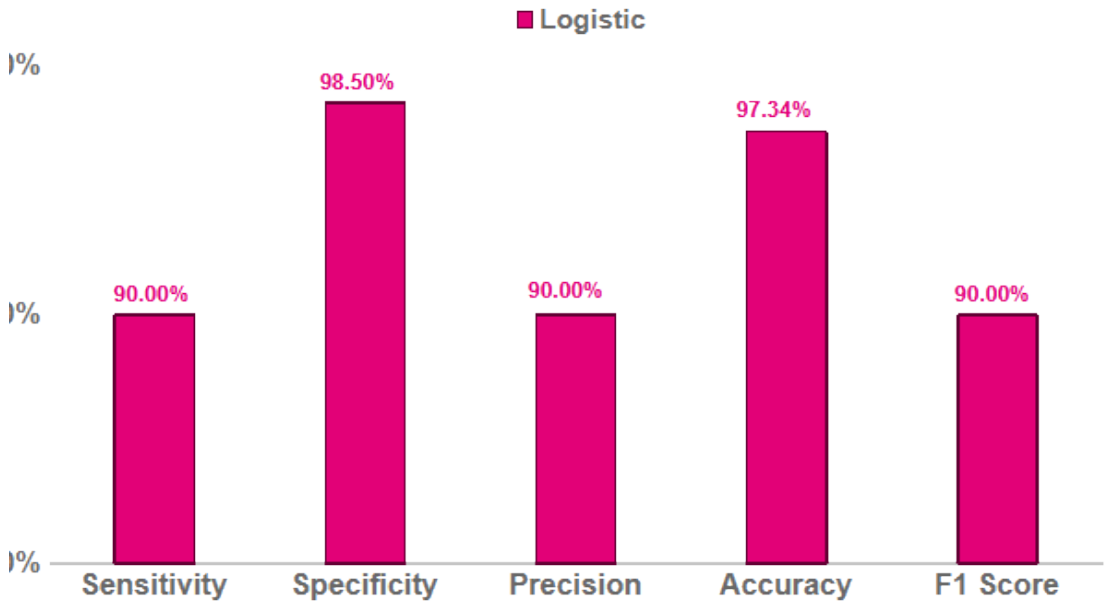
The software R Studio was the IDE used for this analysis. The R Studio packages used in this analysis were dplyr, caret, pROC, and writexl. "dpylr" makes it easy to clean, transform, and manipulate data, which helps us manipulate large sets of data, like the data used in the study. "dyplr" allows us to focus more on our analysis instead of data preparation. "caret" was used to build and train a predictive model that offers a variety of tasks: Data splitting, model

training, and performance evaluation. The “pROC” package was used to assess the model's accuracy between recession and non-recession periods. “pROC” generates ROC curves and calculates the area under the curve (AUC), providing a clear indicator of the model's performance. FRED API key was utilized to read the data from FRED database directly. “writexl” was used for writing Excel files. The writexl package exports results, such as predicted recession periods or model performance metrics, back into Excel for reporting or further analysis.

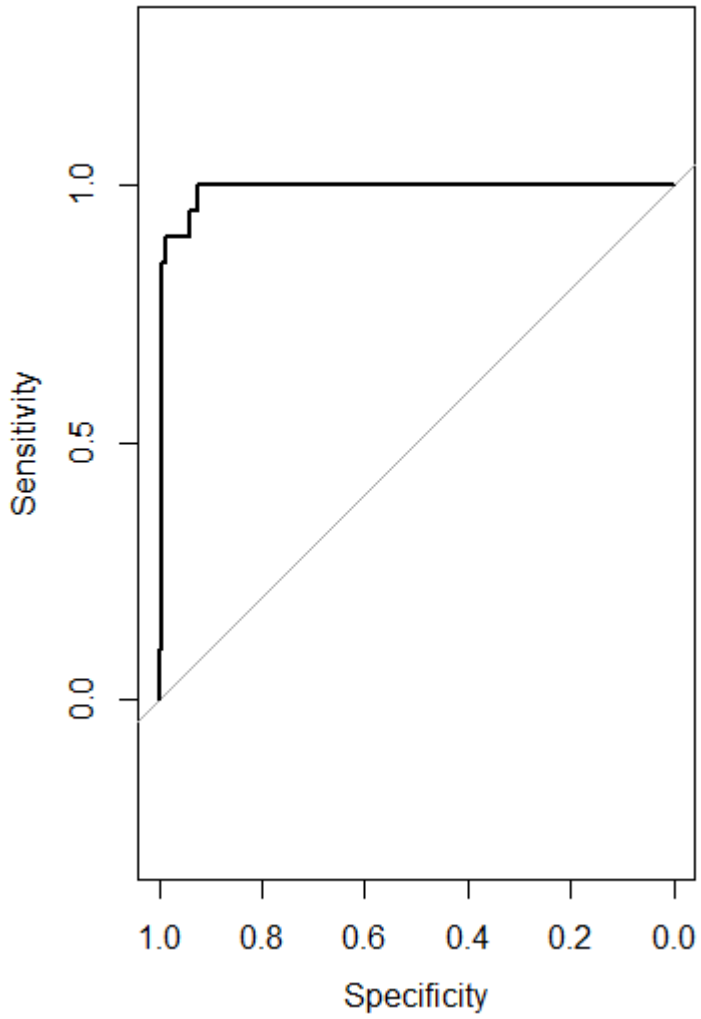
4. Results

The initial model showed high accuracy but potential over-fitting due to a large number of predictors.

		Predicted	
		No Recession (0)	Recession (1)
Actual	No Recession (0)	131 (TN)	2 (FP)
	Recession (1)	2 (FN)	18 (TP)



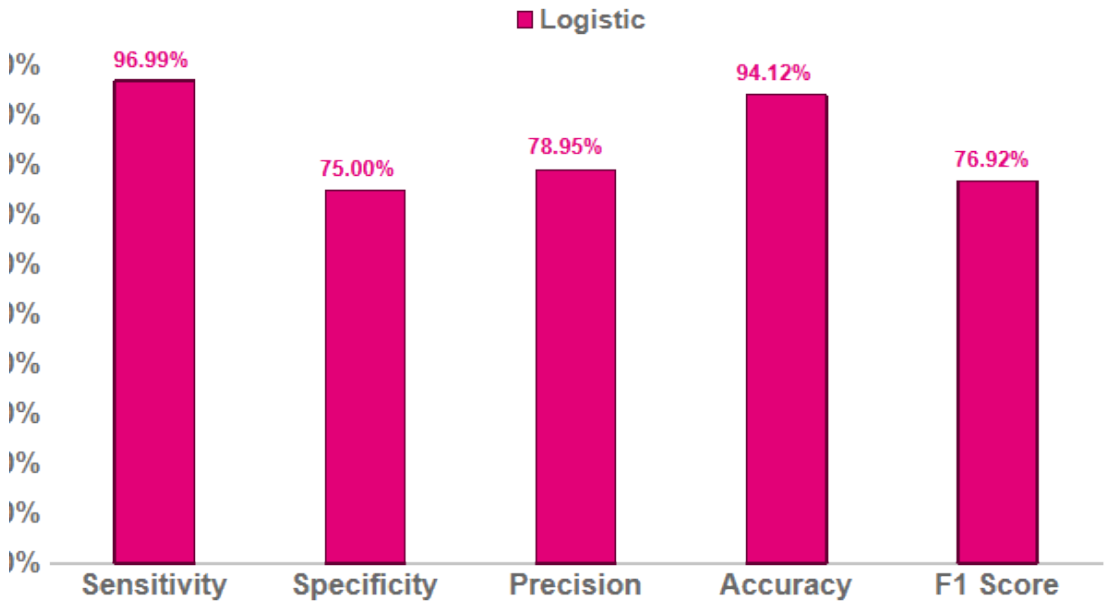
Variables	Estimate	Std. Error	z value	Pr(>|z|)
Manufacturing_PMI_Growth	-616.36	110.81	-5.56	0.000
Federal_Funds_Rate_Lag3	0.57	0.15	3.84	0.000
Building_Permits_Lag1	0.00	0.00	-3.73	0.000
Unemployment_Rate_Growth	27.03	9.70	2.79	0.005
Unemployment_Rate_Lag1	-0.72	0.28	-2.56	0.011
Manufacturing_PMI_Lag3	0.00	0.00	-2.38	0.017
Building_Permits_Growth	-11.95	5.15	-2.32	0.020
(Intercept)	19.88	12.76	1.56	0.119
M2_Money_Stock	0.00	0.00	-1.47	0.141
Industrial_Production_Growth	-60.36	49.77	-1.21	0.225
Personal_Income_Growth	-41.54	35.37	-1.17	0.240
Labor_Force_Participation_Rate_Growth	96.83	90.04	1.08	0.282
Personal_Consumption_Expenditures_Growth	-37.47	42.93	-0.87	0.383
CPI_Growth	92.30	136.42	0.68	0.499
Federal_Funds_Rate_Growth	-1.18	2.15	-0.55	0.583
Housing_Starts_Growth	1.36	3.12	0.44	0.662
Producer_Price_Index_Growth	-15.04	38.90	-0.39	0.699
M2_Money_Stock_Growth	20.63	83.04	0.25	0.804
Core_PCE_Price_Index_Growth	-12.66	199.64	-0.06	0.949
Commercial_Lending_Growth	-1.81	35.51	-0.05	0.959
Labor_Force_Participation_Rate_Lag1	-0.01	0.27	-0.03	0.979
Industrial_Production	0.00	0.04	0.01	0.994



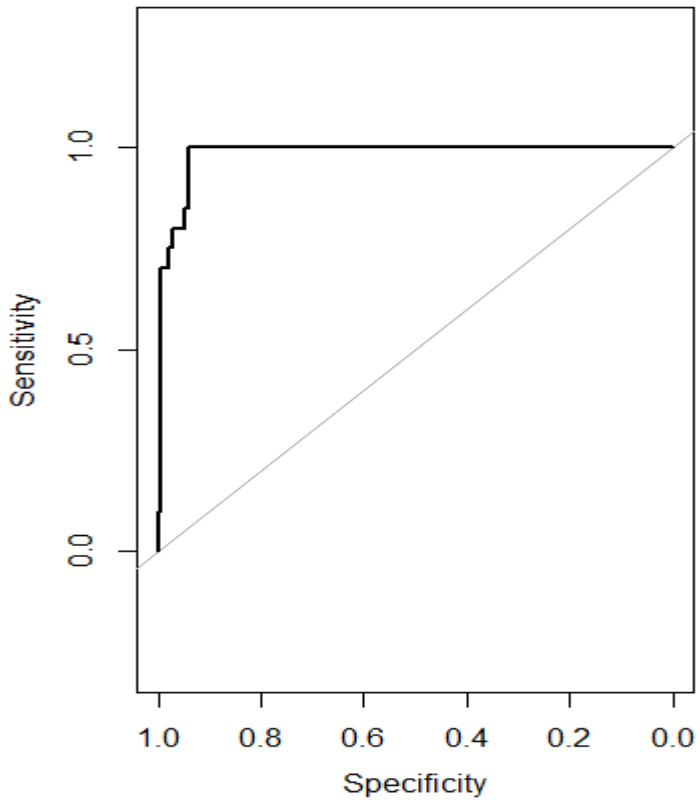
The simplified model utilized the top 7 variables based on variable importance and the variables with p value less than 0.05 in initial model.

Variables	Variable Importance
Manufacturing_PMI_Growth	5.562
Federal_Funds_Rate_Lag3	3.845
Building_Permits_Lag1	3.729
Unemployment_Rate_Growth	2.786
Unemployment_Rate_Lag1	2.556
Manufacturing_PMI_Lag3	2.382
Building_Permits_Growth	2.322
M2_Money_Stock	1.472
Industrial_Production_Growth	1.213
Personal_Income_Growth	1.175
Labor_Force_Participation_Rate_Growth	1.075
Personal_Consumption_Expenditures_Growth	0.873
CPI_Growth	0.677
Federal_Funds_Rate_Growth	0.549
Housing_Starts_Growth	0.437
Producer_Price_Index_Growth	0.387
M2_Money_Stock_Growth	0.248
Core_PCE_Price_Index_Growth	0.063
Commercial_Lending_Growth	0.051
Labor_Force_Participation_Rate_Lag1	0.026
Industrial_Production	0.008

Logistic Confusion Matrix		Predicted	
		No Recession (0)	Recession (1)
Actual	No Recession (0)	129 (TN)	4 (FP)
	Recession (1)	5 (FN)	15 (TP)



Variables	Estimate	Std. Error	z value	Pr(>|z|)
Manufacturing_PMI_Growth	-670.86	93.69	-7.16	0.0000
Building_Permits_Lag1	-0.01	0.00	-5.37	0.0000
Federal_Funds_Rate_Lag3	0.46	0.11	4.26	0.0000
Unemployment_Rate_Lag1	-0.67	0.19	-3.53	0.0004
Building_Permits_Growth	-12.53	4.14	-3.02	0.0025
(Intercept)	9.83	3.30	2.98	0.0029
Unemployment_Rate_Growth	21.62	8.33	2.60	0.0094
Manufacturing_PMI_Lag3	0.00	0.00	-1.89	0.0585



5. Insights & Learnings

The new model retained high accuracy with fewer predictors, indicating robustness. The AUC value of 0.9812 demonstrated excellent discriminatory power.

Key predictors include the growth rate of Manufacturing PMI and the lagged Building Permits, Federal Funds Rate, and Unemployment Rate.

Simplifying the model improved interpretability without sacrificing performance.

6. Conclusion & Future Work

Our study demonstrates that a logistic regression model can effectively predict recessions using key economic indicators such as the Manufacturing PMI, Federal Funds Rate, Building Permits, and Unemployment Rate. The simplified model, which retained only the most significant predictors, showed strong accuracy with an AUC value of 0.9812, confirming its robustness and predictive power. This model's ability to maintain high performance with fewer variables highlights its practical utility for early recession prediction, offering valuable insights for policymakers, businesses, and investors to make informed decisions and prepare for potential economic downturns.

Next Steps

- Model Validation: Further validation of the model using more recent data to ensure its performance holds in different economic conditions.
- Incorporating Additional Indicators: Explore additional leading indicators or other data sources to refine the model further and enhance its predictive accuracy.
- Real-Time Predictions: Implement the model in a real-time environment to track its predictive capability as new economic data is released, providing ongoing recession risk assessment.
- Broader Applications: Investigate the model's applicability across different industries or regions to understand how specific economic factors influence recessionary risks in various contexts.

Sources

- Investopedia. “Recession.” *Investopedia*, 20 Apr. 2023, https://www.investopedia.com/terms/r/recession.asp.
- International Monetary Fund. “Recession.” *Finance & Development*, IMF, https://www.imf.org/en/Publications/fandd/issues/Series/Back-to-Basics/Recession.
- National Bureau of Economic Research. “Business Cycle Dating.” *NBER*, https://www.nber.org/research/business-cycle-dating.