Project Report

Identifying drivers of financial recession in the US

Romain Fernex

Contents

1 Executive Summary:

1.1 Structure of this Paper:

The introduction provides background information on the topic and sets the stage for the rest of the paper. The literature review summarizes and critically evaluates existing research on the topic. The research methodology section describes the methods used to collect and analyse the data we got from kaggle. The data analysis and results sections present and interpret the findings of the study. Lastly, in the discussion section, we expose the limitations of our model and discuss how it could be improved.

1.2 Research Question:

To what extent do leading macroeconomic indicators predict economic downturns in the United States from 1996 to 2020 ?

1.3 Introduction:

This study examines the predictive power of macroeconomic indicators for identifying market recessions in the United States. Economic recessions, defined as sustained periods of declining economic activity, generate substantial ripple effects throughout the financial system, affecting asset valuations, corporate performance, and societal well-being. While the devastating impacts of recessions are well-documented, the ability to anticipate these downturns remains a critical challenge for both policymakers and market participants. This research contributes to the existing literature by analyzing the effectiveness of leading macroeconomic indicators in forecasting U.S. business cycle contractions. The significance of this investigation is threefold: first, it enhances our understanding of recession prediction mechanisms; second, it provides policymakers with empirical evidence to support preemptive stabilization measures; and third, it offers market participants analytical tools for strategic decision-making. By examining historical patterns of economic indicators preceding past recessions, this study aims to identify reliable early warning signals that could improve economic forecasting and risk management practices.

1.4 Literature Review:

The current economic and geopolitics conjuncture has raised concerns all across the world regarding the potential arrival of a new recession. Though these fears have gained in intensity in the recent years, they are in no way new, especially when looking at crises such as the 2008 market crash. Therefore, analysts throughout the world have tried to identify potential predictors for such downturns in the business cycle in order to cushion their impact. Currently, in the US, half of the fifty states have already witnessed a slowing down of their economies which prompted institutions such as the Saint Louis Federal Reserve Bank (2022) to warn against the risk of an incoming recession in the coming months. In it latest report the San Francisco Federal Bank (2022) backed this warning, this time by highlighting the rise in the unemployment rate which is often regarded as a reliable signal of a downturn. Interestingly, the report also specified that unemployment rate is the most reliable when looking at near term prediction and tends to lose in relevance to the benefit of the market yield curve as the time horizon increases. Apart from these first two indicators, significant attention has also been given to indicators closely linked to the real economy such as consumption and firm's performance indexes. These are indeed often tied to recession, as shocks tend to follow decrease in spendings and investments due to ailing confidence from consumers and investors.

Most of these insights stem from economists who have already identified a wide array of different indicators that perform well in predicting recession and which we have decided to focus on. Firstly, as regards financial indicators, studies have shown that the S&P 500 composite stock price index is especially meaningful (A. Estrella & S.Michkin, 1998) for short term forecasts. Then, following the 2001 US recession, others have instead focused on monetary and labor indicators, such as the aforementioned unemployment rate, and asserted that they are close to financial indicators in terms of prediction quality (Mehdi Mostaghimi, 2006). Lastly we needed to find an appropriate dependent variable that would faithfully tell us when periods of recession took place. To this end we turned to the recession index published by the National Bureau of

Economic Research (NBER) which has been used in most studies in this field and has already been used to identify the best recession indicators for different time horizons (Weiling Liu & Emmanuel Moench, 2016)

1.5 Research Methodology:

We use a dataset found on Kaggle, leveraging data from Yahoo, the NBER, quandl and macrotrends. We more specifically base our analysis on the Macrodata dataset which maps the level of a variety of macroeconomic indicators to the state of the US economy the corresponding year. The state of the US economy is a binary variable with the economy being either "normal" or in "recession". We take "normal" as the base level in our analysis as described later in this paper. The data is collected monthly and covers a period that spans from 1996 to 2020. Indicators present in the dataset can be categorized as follows : Output and Income, Labour Market, Real Estate, Consumption and Expenditures, Money and Credit, Interest and Exchange Rates, Prices, Stock Market". Each of these groups contains several variables such as Real Personal Income, Civilian Employment, New Private Housing Permits, Real Personal Consumption Expenditures, Money Stock, Effective Federal Funds Rate, Finished Goods, and the 500 S&P's Common Stock Price Index.

For this study we take the state of the economy variable as our dependent variable. As regards independent variable, we carry out dimensionality reduction using two separate methods : a PCA analysis and a stepwise VIF analysis. In the case of the stepwise VIF analysis (more about this below) we retained the following 9 indicators :

- UNRATE : Unemployment rate
- HOUST : Housing Starts (Total New Privately Owned)
- ISRATIO : Inventories to Sales Ratio
- T1YFFM : Spread between 1-year Treasury rate and the Federal Funds rate
- AAAFM : Spread between highest-rated corporate bonds and Federal Funds rate
- EXUSUK : U.S. / U.K. Foreign Exchange Rate
- NASDAQ composite index
- S&P P/E ratio
- S&P Dividend yield

For the analysis a multiple logistic regression is being used. The multiple logistic regression is a statistical method used to analyse the relationship between multiple independent variables (also known as predictor variables) and a binary dependent variable. The goal of this method is to model the probability of the occurrence of the dependent variable as a function of the independent variables. The logistic regression model uses a logistic function to model the probability of the occurrence of the dependent variable, with the function output ranging between 0 and 1 . The independent variables are represented by coefficients (or beta values) in the model, which are estimated through the optimization process. The coefficients indicate the strength and direction of the relationship between the independent variables and the dependent variable.

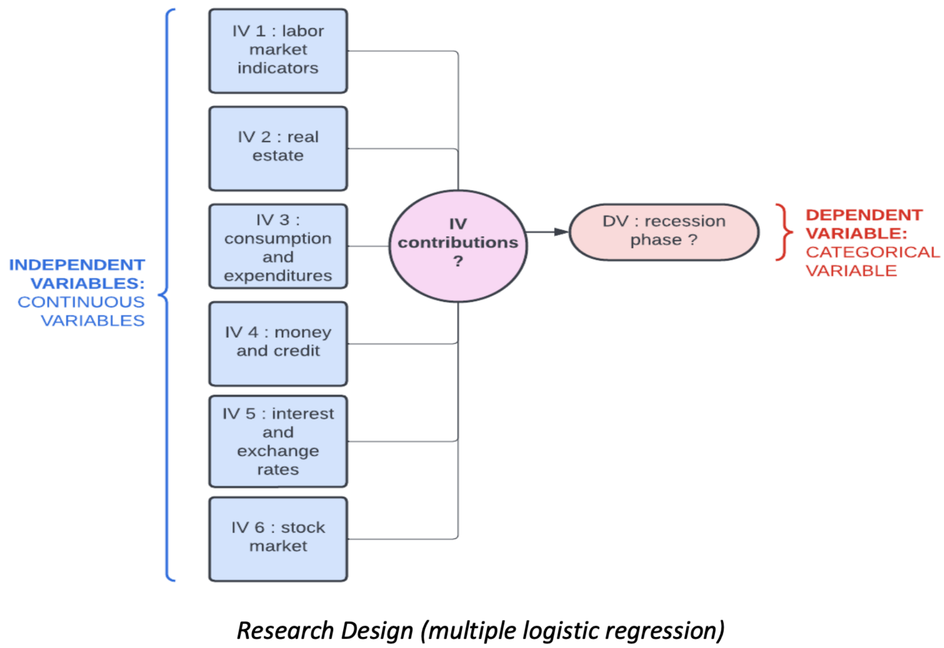

Figure 1: Research Design

We then perform a multiple variable logistic regression to identify which indicators are significant associated with the odds of entering a recession period. Finally, we compare the performance of the logistic regression model as a predictor depending on the method used for dimensionality reduction (PCA method compared to the stepwise VIF method)

2 Exploratory Data Analysis

Prior to analysis, we conducted a comprehensive data quality assessment using Python's DataFrame methods. The .describe() function provided summary statistics including measures of central tendency and dispersion, while .info() revealed the data structure, variable types, and completeness of observations. This preliminary examination confirmed the dataset's integrity, with all variables properly formatted and no missing values.

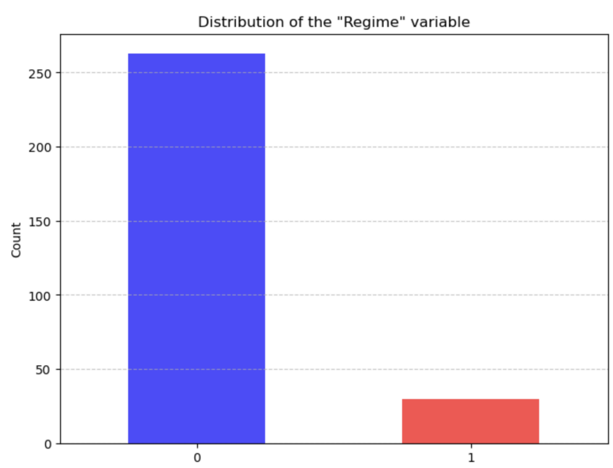

Figure 2: Distribution of the 'Regime' variable

Looking at the distribution of the binary variable for the state of the economy, we notice that there is a strong imbalance between the number of recession periods compared to the number of periods without a recession. We account for this by defining sampling weights during our analysis.

3 Scaling

After using the .describe() method as well as performing a visual inspection of the data using matplotlib we decided to scale the variables in our dataset in order to make the results easier to interpret. To this end, we relied on sklearn.prepocessing StandardScaler which scales the data to ensure that all variables follow a standard normal distribution. While the dataset still remains sensitive to outliers, scaling the data to the unit variance is vital in order to reduce risks of misclassification error and improve the accuracy rate of our model. Therefore, we limit the risk of a variable contributing more than another due only to a difference in scale. Here is an overview of the dataset before and after the rescaling:

	RPI	INDPRO	CE16OV	UNRATE	PAYEMS	USGOOD	USTPU
0	8909.327	74.6841	125125.0	5.6	118316.0	23196.0	23947.0
1	8983.863	75.8344	125639.0	5.5	118739.0	23280.0	23988.0
2	9015.588	75.7631	125862.0	5.5	118993.0	23276.0	24030.0
3	9039.466	76.4562	125994.0	5.6	119158.0	23316.0	24043.0
4	9078.928	77.0161	126244.0	5.6	119486.0	23358.0	24137.0

Figure 3: Original dataset

	RPI	INDPRO	CE16OV	UNRATE	PAYEMS	USGOOD	USTPU
0	-1.848982	-2.874032	-2.131115	-0.080942	-2.175660	0.894624	-2.005193
1	-1.815403	-2.727781	-2.066145	-0.134477	-2.120532	0.933683	-1.963169
2	-1.801111	-2.736846	-2.037958	-0.134477	-2.087429	0.931823	-1.920120
3	-1.790354	-2.648725	-2.021273	-0.080942	-2.065926	0.950423	-1.906795
4	-1.772576	-2.577538	-1.989673	-0.080942	-2.023179	0.969953	-1.810447

Figure 4: Scaled dataset

4 Dimensionality reduction methods

4.1 Method 1 : Principal Component Analysis:

We begin by identify which We thus carried out a Principal Component Analysis after dividing our dependent variables in two main categories :

- Financial indicators: S&P PE ratio, S&P dividend yield, US/UK Foreign exchange rate...
- Non-financial indicators : civilian unemployment rate, real personal income, new orders for consumer goods...

Using the PCA, we restricted the number of relevant variables to 4 for financial indicators and 5 for non financial indicators, giving us a total of 9 dependent variables. This enabled us to improve the model quality, as will be shown below. However this also made interpreting our results more difficult as we cannot know what the variables created by the PCA exactly describe.

4.2 Method 2 : Using VIF coefficients (stepwise)

VIF coefficients are used To test for multicollinearity but they can also be used for dimensionality reduction. Compared to the PCA method, it is more interpretable as it keeps original features, and removes the ones causing multicollinearity. That said it potentially capturing less information than the PCA method which is why we compare the effectiveness of these methods later in our analysis. We decided to take a stepwise approach using statsmodels' variance_inflaction_factor function. The program performs a series of VIF analysis and removes the variable with the highest VIF score until all variables remaining have a VIF score that is strictly inferior to 4.

```
Final VIF scores:
         Variable       VIF
0           const  1.000000
1          UNRATE  2.752493
2           HOUST  3.227608
3         ISRATIO  2.539421
4          T1YFFM  1.599237
5          AAAFFM  2.783558
6          EXUSUK  3.518797
7          NASDAQ  2.643204
8             P/E  1.655014
9  Dividend Yield  2.147862
```

Figure 5: Relevant Independent Variables obtained through stepwise VIF method

5 Analysis process and results

5.1 Test and Train split

After performing the exploratory data analysis (EDA), we split the data set into a training and testing set using the formula ' train_test_split.' For the training size, we changed the proportion from 80% to 60% of our dataset so as to get more exploitable results for the out-of-sample ROC curve.
This training set is used to generate the logistic regression model which is fitted to the dataset to ensure a sufficient level of in-sample fit. Then, testing set is used to access the quality of the regression model and control for out-of-sample fit. The variables which passed the previous multicollinearity test (or those generated after the PCA in the case of the PCA approach) are used as 'x' and the dataset containing the values of the dependent variable is used as 'y'.
The data sizes for each are listed below :

- Observation in train dataset : 175
- Observation in test dataset :118

5.2 Results:

5.2.1 Multiple logistic regression : process summary

We use the LogisticRegression function of the sklearn package to run our multiple variable logistic regression. This enables us to add sampling weights that help correct the imbalance present for the dependent variable. The model is first fitted using the training set and the .fit function of sklearn. Then we compute the p-values and standard deviations to draw conclusions on the significance of the coefficients we determined for each predicting variable. We also compute odds ratio using np.exp as they are easier to interpret. Lastly, we compute predicted values for the test set with the .predict_ proba function of the sklearn package which will be used to get the ROC curve.

5.2.2 Results : Multiple Logistic Regression (VIF method)

VIF Logistic Regression Results:

Variable	Coefficient	Std Error	p-value	Odds Ratio
UNRATE	0.1691	0.1069	0.1137	0.6875
HOUST	0.3578	0.1305	0.0061	0.4578
ISRATIO	0.9534	0.1110	0.0000	0.2483
T1YFFM	-1.2093	0.0946	0.0000	2.9508
AAAFFM	1.3715	0.1249	0.0000	1.8311
EXUSUK	1.6687	0.1389	0.0000	0.9855
NASDAQ	1.3867	0.1163	0.0000	0.5194
P/E	0.8619	0.0885	0.0000	0.9127
Dividend Yield	0.7305	0.1029	0.0000	1.3074

Figure 6: Logit regression results (VIF)

Looking at the *p*-value, we see that all variables except the the unemployment rate have a significant impact on the odds of the US experiencing a recession. The variables with the strongest impact on odds of recession are as follows :

- with a one unit increase in the inventories to sales ratio (ISRATIO), the odds of recession decrease by over 80%
- Similarly, with a one unit increase in the number of new residential construction projects (HOUST) the odds of recession are nearly halved.
- Conversely, with a one percentage point increase in the spread between 1-year Treasury rate and the Federal Funds rate (T1YFFM), the odds of a recession skyrocket by nearly 200%

5.2.3 Results : Multiple Logistic Regression (PCA method)

While the PCA method is leads to less transparent results, there are some noticeable differences between the results we obtain with this method and those we got earlier : On average, financial variable appear to a have a significantly greater impact on the odds of recession than non financial ones. This runs counter to the previous results where the variables with the highest explanatory power were non financial (HOUST..).

PCA Logistic Regression Results:

Variable	Coefficient	Std Error	p-value	Odds Ratio
FIN_PC1	-0.3748	0.1659	0.0239	0.6875
FIN_PC2	-0.7813	0.1852	0.0000	0.4578
FIN_PC3	-1.3933	0.0901	0.0000	0.2483
FIN_PC4	1.0821	0.1083	0.0000	2.9508
FIN_PC5	0.6049	0.1151	0.0000	1.8311
NFIN_PC1	-0.0146	0.1181	0.9017	0.9855
NFIN_PC2	-0.6551	0.0993	0.0000	0.5194
NFIN_PC3	-0.0913	0.0722	0.2060	0.9127
NFIN_PC4	0.2680	0.1781	0.1323	1.3074

Figure 7: Logit regression results (PCA)

Looking at p-values we get the following results :

Group	Significant Variables	Positively Correlated
Financial Variables	4/5	3/5
Non-financial Variables	1/4	3/4

The predominance of financial variable is consistent with most of the existing scientific literature. While it is difficult to determine what each variable corresponds to, we observe that most variables contribute negatively to the odds of recession. This echoes the findings we got using the VIF method. Finally, there significantly fewer relevant variables when using the PCA method compared to the VIF method, although this difference is not easily interpretable.

5.3 Comparison of model performance : Out-of-sample ROC curves

After creating these two models, we tried to compare their respective performance in order to evaluate which one was the most efficient predictor of an incoming recession. To this end, we plotted the out-of-sample Receiver Operating Chatacteristic (ROC) curve for each model. This yielded the following graph :

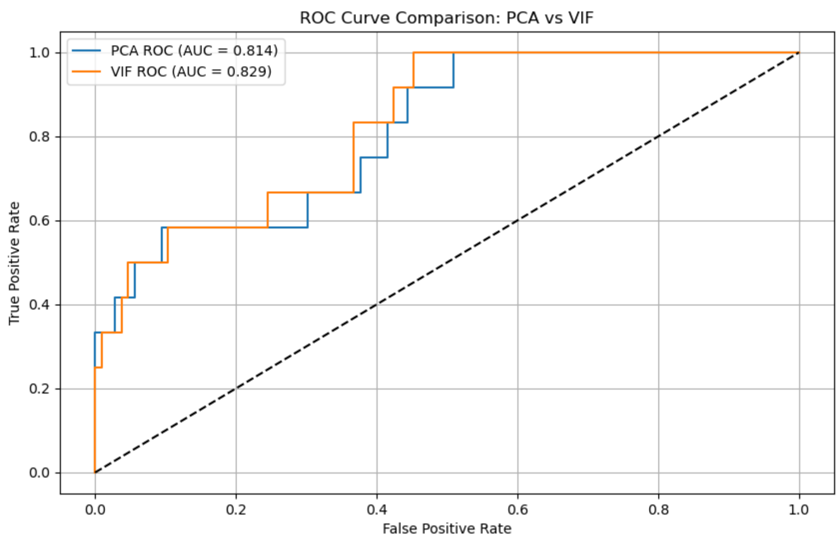

Figure 8: ROC curves

At first glance both models seem to perform fairly well, as shown by decent AUC levels that lie around 0.8. No clear distinction can be made between the two models performance-wise, at least based on the AUCs. Nonetheless, looking at the confusion matrixes tells a different story :

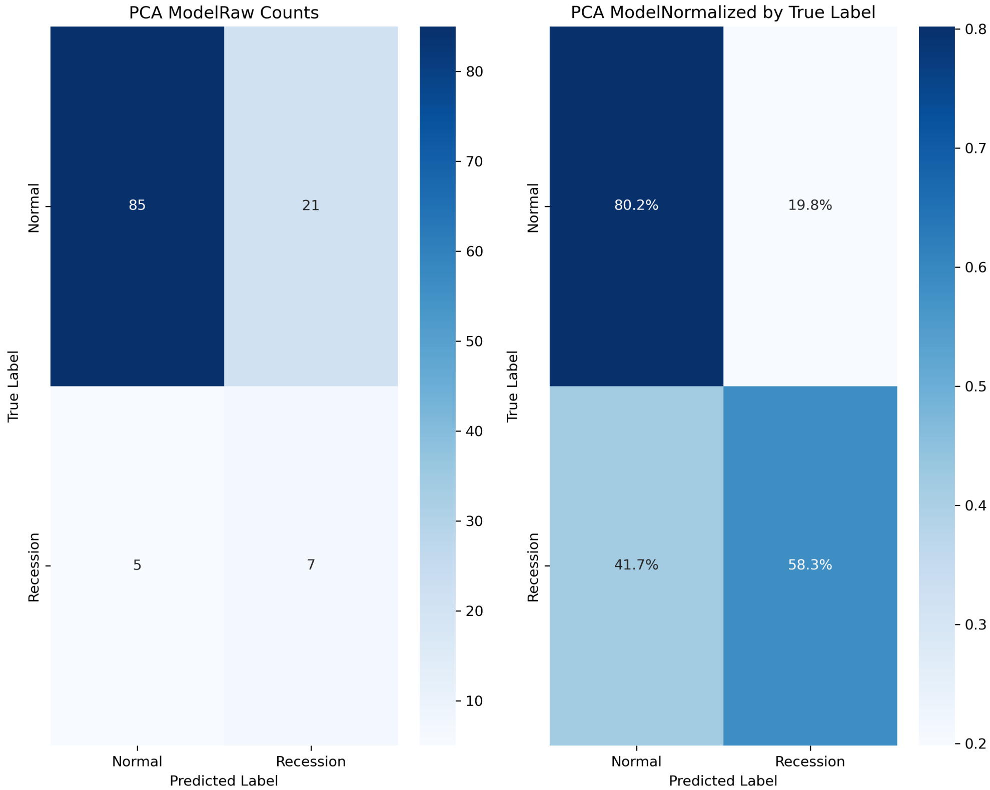

Figure 9: PCA Confusion Matrix

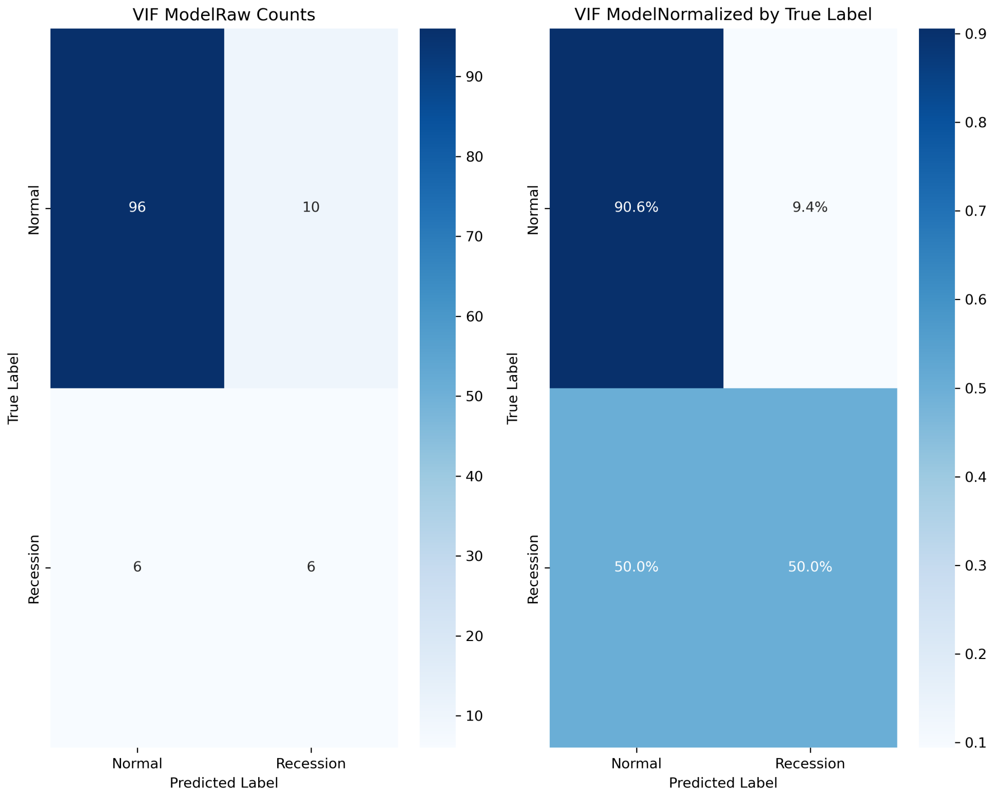

Figure 10: VIF Confusion Matrix

Both models visibly suffer from very poor performances in terms of recall and precision, despite satisfying accuracy levels when identifying normal periods. Some reasons for these remarkably lackluster performance could include :

- a failure of the selected variables to sufficiently explain our variable of interest (i.e the odds of having a recession). Observed R-squared are fairly low which signifies a lack of information
- the imbalance between the number of recession periods compared to the number of normal periods. The logistic regression model is notably weak to these gaps and our weighting most likely did not offset this effect completely. The weighted metrics we computed seem to point in that direction as well (see below)

```
PCA Metrics:
   Metric    Score
 Accuracy 0.779661
Precision 0.873823
   Recall 0.779661
 F1 Score 0.814735
AUC Score 0.814465
VIF Metrics:
   Metric    Score
 Accuracy 0.864407
Precision 0.883599
   Recall 0.864407
 F1 Score 0.872788
AUC Score 0.828616
```

Figure 11: Weighted metrics

6 Conclusion : Limitations and Potential Improvements

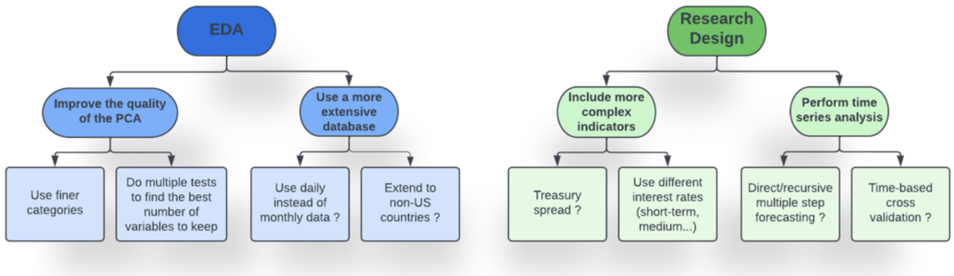

Figure 12: improvement

6.1 Exploratory Data Analysis

As mentioned earlier when discussing the principal component analysis approach, we only separated the independent variables in two main categories: financial and non-financial indicators. This made interpreting the results more complicated as we had less visibility on what each variables created by the PCA actually portrayed. A finer categorization might involve segmenting our datasets between employment, consumption, monetary, financial market indicators etc... This would enable us to better understand which specific factors have the highest impact on the probability of a recession happening. Another important limitation of our study lies in the relatively small database that we had access to. For instance, we had to restrain the size of our training set significantly in order to improve the quality of our ROC curves which suffered from a lack of datapoints in the testing set. This undoubtedly caused a loss of quality in terms of in-sample-fit and altered the quality of our model. Thus, finding a dataset with daily instead of monthly reporting would most likely improve our study significantly.

6.2 Research Design and Techniques Used

A core limitation of our study is linked to the dependent variables themselves. Most modern studies that cover this issue tend to include more complex and diverse factors such as interest rate spread or, most importantly, the yield curve. This last indicator has been proven to consistently outperform other indicators when it comes to predicting recession for a 1 to 2 years time horizon (Estrella & Mishkin, 1996). Therefore, a more detailed mathematical model might be needed in order to take these other variables into account in an efficient way. In the medium term, our model would also gain using with other more advanced classification methods such as adaptive or gradient boosting. Looking at the ROC curve obtained with each of these methods might give us insight regarding which lead would be the most promising to maximize the efficiency of our model based on the existing dataset. Finally our analysis is cross-sectional and not longitudinal. This implies that the conclusions derived from our study are more about the potential association between the variables we used rather than about their potential as predictors. While undeniably useful, a better designed analysis would rely primarily on time series analysis techniques. In the case of our study, direct multistep forecast strategy could be an interesting lead. Indeed, this method involves creating several separate models for different time horizons. This would enable us to pinpoint which variables are the most relevant predictors in the short, medium or long term.

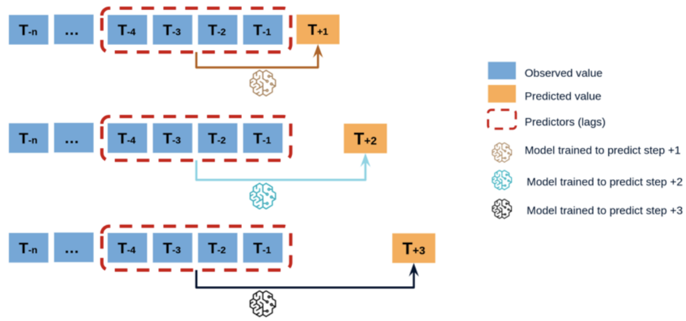

Figure 13: multistep forecast

Description of the direct multi-step forecast method

7 References:

7.1 Newspaper reports

- Link to the Fortune article by Philipp CARLSSON-SZLEZAK and Paul SWARTZ
- Link to the Reuters article by Michael S. DERBY

7.2 Papers

- Estrella, A., & Mishkin, Predicting US recessions: Financial variables as leading indicators. The Review of Economics and Statistics, 80(1), 45-61, 1998
- Mehdi Mostaghimi, Predicting US 2001 Recession, composite leading economic indicators, structural change and monetary policy, The Singa pore Economic Review, vol. 51, N**3**, 343-363, 2006
- W.Liu & Emmanuel Moench, What predicts US recessions ?, International journal of forecasting, n°32, 1138-1150, 2016
- Estrella, A., & Mishkin, An overview of using the yield curve as a forecasting tool, Current Issues in Economics and Finance (2) 7, June 1996

7.3 Datasets

- Link to the kaggle page to download the dataset used in this study