THE GRADUATE CENTER, CITY UNIVERSITY OF NEW YORK

Predicting Economic Recessions using Machine Learning

Advanced Data Analysis, Spring 2020

Sheridan Kamal

5/22/2020

Using the supervised machine learning methods logistic regression and decision tree classification, we will determine the best method to predict economic recessions. To do this, we will be using train_test_split to create a train and test set and we will be using a combination of supervised and unsupervised machine learning methods to determine the best model to predict recessions. The best prediction model is the supervised machine learning model logistic regression using the default parameters with the scaled data and it had an accuracy of 93.7% and ROC AUC of 0.98. It would be best to use the supervised machine learning models over the unsupervised models for this type of analysis since the unsupervised models all have low scores.

Introduction

Due to the pandemic caused by the novel Coronavirus (COVID-19), you may have noticed that the stock markets have been turbulent as a result of fears that the pandemic will have on different industries and trade. These fears have caused financial markets to tumble at an alarming rate and at times triggering the "circuit breakers" to halt trading. The triggering of these "circuit breakers" seems to occur more frequently with the duration of the pandemic. It is safe to say if the pandemic is prolonged then the negative impacts on the financial markets and the economy as a whole could push us into a recession, which we have not been in for 11 years. It is precisely these thoughts which drove me to research predicting recessions using machine learning methods.

In finance, a central idea is “the best indicator of future performance is past performance”, which is why many forecasting models are heavily dependent on historical data. In this paper, I will be using the supervised machine learning methods logistic regression and decision tree classification to predict recessions. By using machine learning methods we may be able to get a sense of where the market is heading even if we cannot predict a recession, which will be useful for the Federal Reserve and policy makers to know because if the market is heading towards a recession they may be able to provide a boost to the economy and in turn the financial markets. This research is also useful for traders because if it looks like we may be heading towards a recession then they may want to shift their money from equities into other financial instruments.

Related Work

Azhar Iqbal and Kyle Bowman (2018)[1] decided to use various machine learning models and statistical data mining to determine if this would improve recession prediction accuracy. In their paper, they decided to use gradient boosting, random forest, data mining (logit/probit), and benchmark-probit models for their analysis. For each model, they generated a ROC curve for the in-sample and out-sample data as well as the AUC for the in-sample and out-sample data. They determined that the machine learning models (gradient boosting and random forest models) provided more accurate results than the statistical data mining models.

Although I will be creating machine learning models as in the research mentioned above, I will be using logistic regression and decision tree classification models rather than gradient boosting and random forest models as a way to expand on their research. I also am using accuracy scores as well as AUC scores for my model evaluation rather than relying completely on the ROC curve and AUC scores. I will also be using the unsupervised machine learning methods k-Means clustering, Agglomerate clustering, and DBSCAN clustering to determine if unsupervised machine learning methods would be better for this type of analysis instead of the supervised machine learning methods.

---

[1] Iqbal, Azhar, and Kyle Bowman. "Can Machine Learning Improve Recession Prediction Accuracy?" *Journal of Applied Economics and Business*, vol. 6, no. 4, Dec. 2018, pp. 16–34.

Data

The dataset that I will be using is monthly data from a combination of economic and financial data. Since the features I need for my dataset are not conveniently included in a downloadable dataset, I downloaded each feature separately and combined them together into one data frame in Python using an inner join. The best place to pull each economic feature from is the FRED (Federal Reserve Economic Data from the Federal Reserve Bank of St. Louis) because I can be sure the data is reliable and can be sure of the quality of the data. To pull the economic data, I will be using Quandl, which also has the added bonus of automatically calculating selected transformations or different data frequencies if I chose to do so. The financial feature is downloadable from Yahoo! Finance so I can download the dataset and create any transformed variables in Excel and import the dataset as a CSV in Python. I will also create Recession labels from a list of start and end dates found from the NBER (National Bureau of Economic Research). After creating the data frame of the features (including the transformed features) and the recession labels, I created a correlation heat map to determine which features will make up my final dataset. The goal was to choose features that were not highly correlated with each other, but were still correlated with the target feature.

By visualizing the data, I was able to see that the features are not normally distributed so feature scaling may be required. I was also able to see that my target feature is also not equally distributed as there are more “no recession” labels than there are “recession” labels.

For the supervised machine learning models, I will be running the models on both the original data and on the data that I scaled using StandardScaler. For the unsupervised machine learning models, I will be running the models on both the scaled data and on the scaled data that has been transformed and projected to 3 components after PCA has been performed.

Figure 1. Correlation heat map for final dataset feature selection.

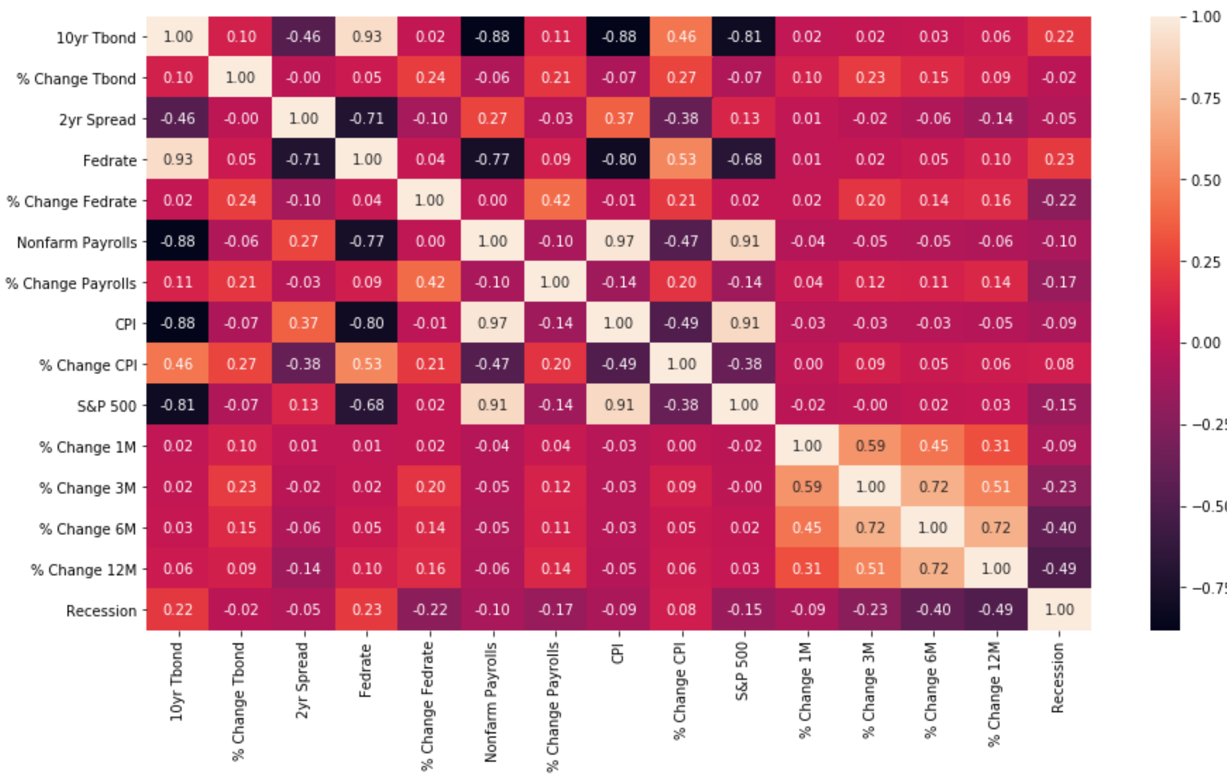


Table 1. Summary statistics table for the chosen 6 features from the original 14 features, 06-01-1976 to 04-01-2020.

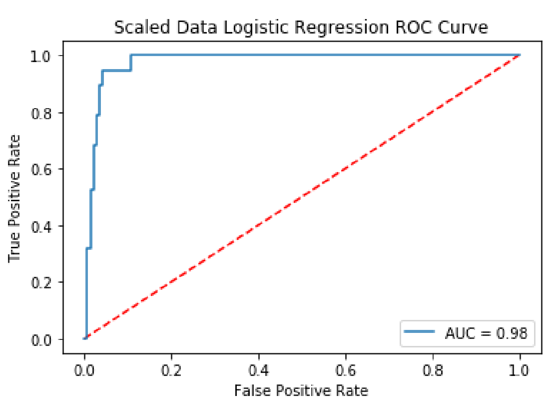
Variable	Count	Mean	STD	Min	Max
10yr Bond	527	6.177287	3.25887	0.66	15.32
2yr Spread	527	0.931366	0.912993	-2.13	2.83
Fedrate	527	4.890266	4.039102	0.05	19.1
% Change Payrolls	527	0.000975	0.006252	-0.135484	0.012404
CPI	527	163.281934	57.315716	56.7	259.05
% Change 12M (S&P 500 Index)	527	0.078462	0.152531	-0.593415	0.42489

Results

Table 2. The results of the supervised machine learning models using the hold-out validation method.

	Accuracy	ROC AUC
Logistic Regression		
Unscaled	88%	0.99
Tuned Unscaled	89.2%	0.91
Scaled	93.7%	0.98
Tuned Scaled	93.7%	0.98
Decision Tree Classification		
	Accuracy	ROC AUC
Unscaled	89.2%	0.55
Tuned Unscaled	92.1%	0.5
Scaled	92.4%	0.71
Tuned Scaled	93%	0.87

Table 2 summarizes the results of the supervised models using the hold-out validation method. As we can see, the best prediction model is the logistic regression model using scaled data because it has a higher accuracy than the logistic regression model with the unscaled data even though it has a higher ROC AUC (by only 0.01). Since the tuned logistic regression model using scaled data selected the same parameters as the default logistic regression model using scaled data, the second best prediction model is the tuned decision tree classification model using scaled data with the following tuned parameters: max_depth = 1, max_features = 6, min_samples_leaf = 2, and the criterion = gini.



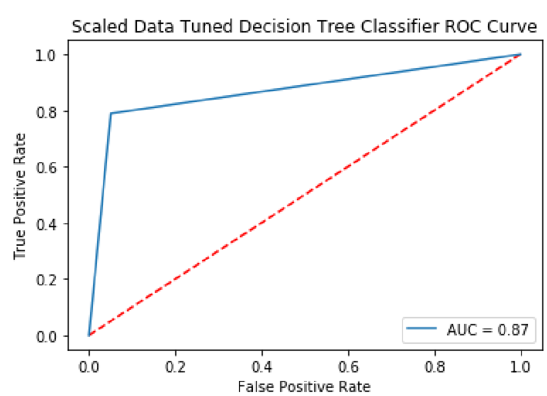


Table 3. The results of the logistic regression model using PCA transformed scaled data.

Logistic Regression after PCA (n_components = 3)	
Accuracy	ROC AUC
93.7%	0.97

After performing PCA analysis on the scaled data, it was determined that the optimal number of components that captures at least 75% of the variance in the data is 3 so we fit and transformed the data into 3 components and determined that the most important features were 'Fedrate', '% Change Payrolls', and '2yr Spread'. Using the PCA transformed data we re-ran the logistic regression model to determine if PCA would improve the accuracy and/or ROC AUC score and found that although the accuracy remained the same the ROC AUC score decreased.

Table 4. The results of the unsupervised machine learning models using the hold-out validation method.

	k-Means Clustering	
	Adjusted Rand Score	Silhouette Score
Scaled (k = 15)	0.03	0.34
Projected (k = 9)	0.07	0.38
	Agglomerate Clustering	
	Adjusted Rand Score	Silhouette Score
Scaled	0.02	0.29
Projected	0.02	0.35
	DBSCAN Clustering	
	Adjusted Rand Score	Silhouette Score
Scaled	-0.12	-0.13
Projected	0.5	0.12

Table 4 summarizes the results of the unsupervised models using the hold-out validation method. The best prediction model is the k-Means clustering model with k = 9 using the PCA transformed scaled data. As all the scores are low, unsupervised machine learning is not the best choice for this analysis.

Figure 2. Unsupervised clustering models using PCA transformed scaled data and scored using adjusted rand score.

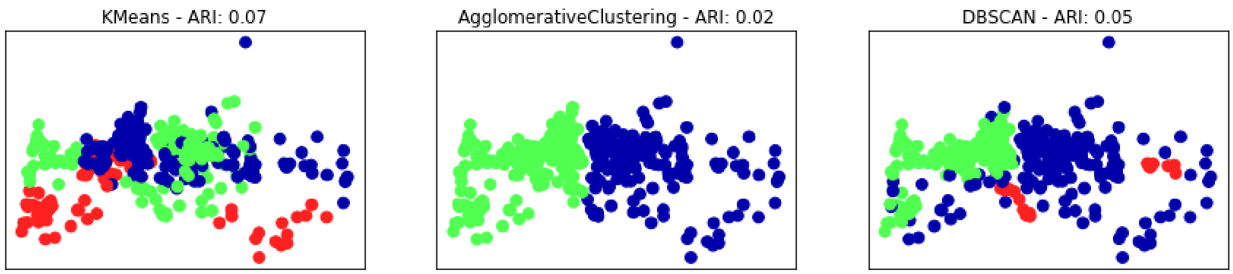
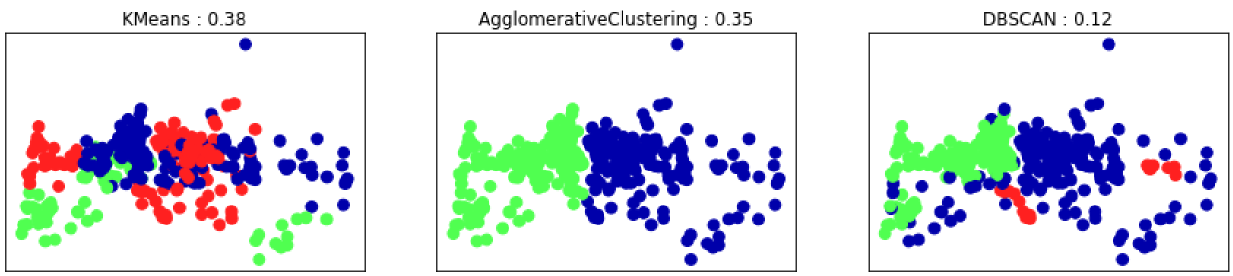


Figure 3. Unsupervised clustering models using PCA transformed scaled data and scored using silhouette score.



Conclusion

The purpose of this paper was to use machine learning methods to predict economic recessions. It was determined that the best model to predict economic recessions is the logistic regression model using scaled data. The second best model to predict economic recessions is the decision tree classification regression model using scaled data.

Through the results of our models, it is safe to say that the best type of machine learning method to use to predict economic recessions is supervised machine learning, most specifically the logistic regression model, as the unsupervised machine learning models all have low scores.

If I had to do everything over, I would have focused more on the supervised machine learning models by adding a random forest classification model to compare the other supervised models to since the unsupervised models are inconclusive and I would have created a dataset with more economic and financial features such as housing features.