Sentiment analysis of Federal Open Market Committee (FOMC) minutes: are they leading indicators of a recession similar to traditional Metrics?

Patrick Long

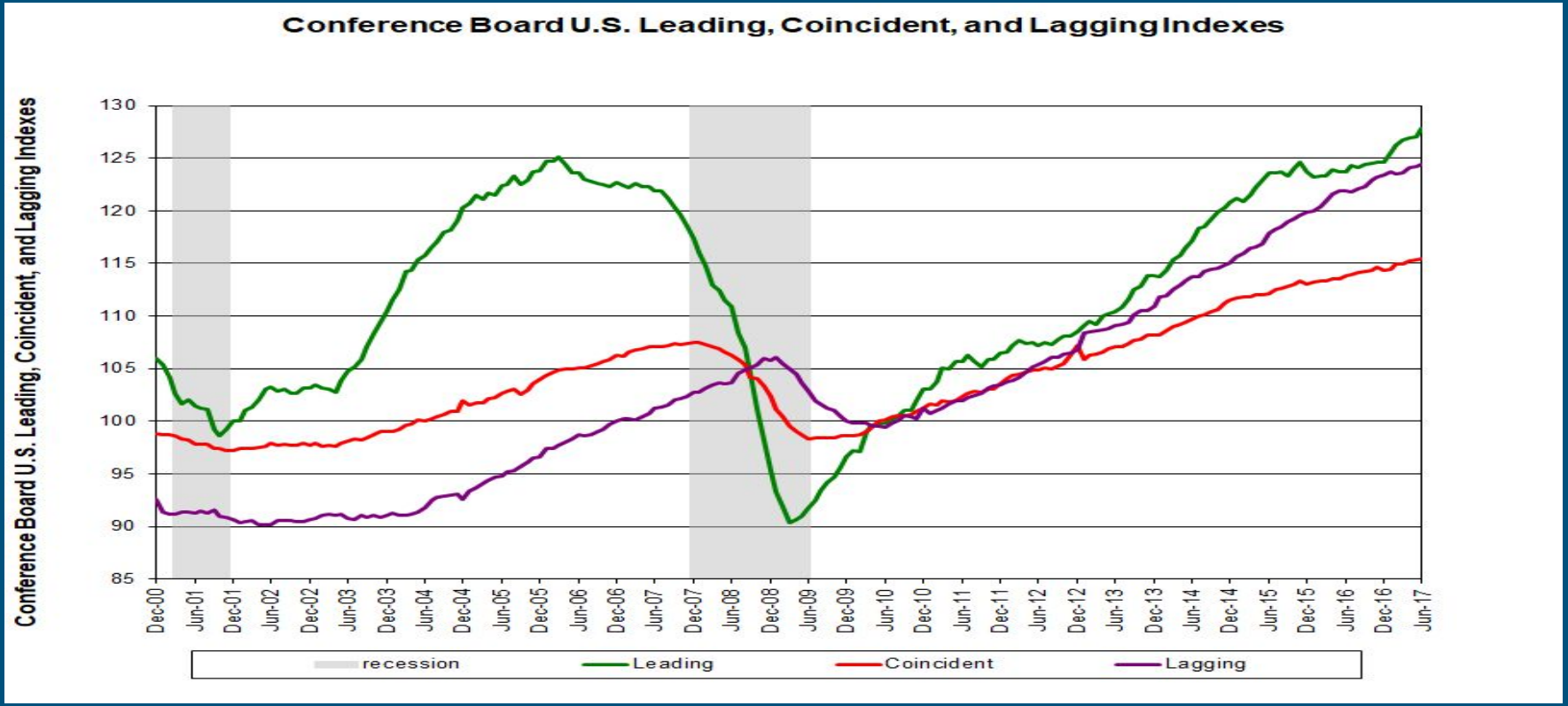
What’s in the Presentation?

1. Recap of the Project Overview and EDA
2. Developing a target Variable and Final EDA
3. Model Comparisons and Evaluations
4. Conclusions
5. To-do List - What’s next?

Recap of the Project

- Dataset: FOMC Meeting Minutes

- Use NLP to assess whether FOMC Meeting Minutes can act as a leading indicator of a recession, i.e., trends downward typically 12 to 18 months before a recession starts and upward similarly early for an economic expansion.



Source: Conference Board https://www.conference-board.org/us/

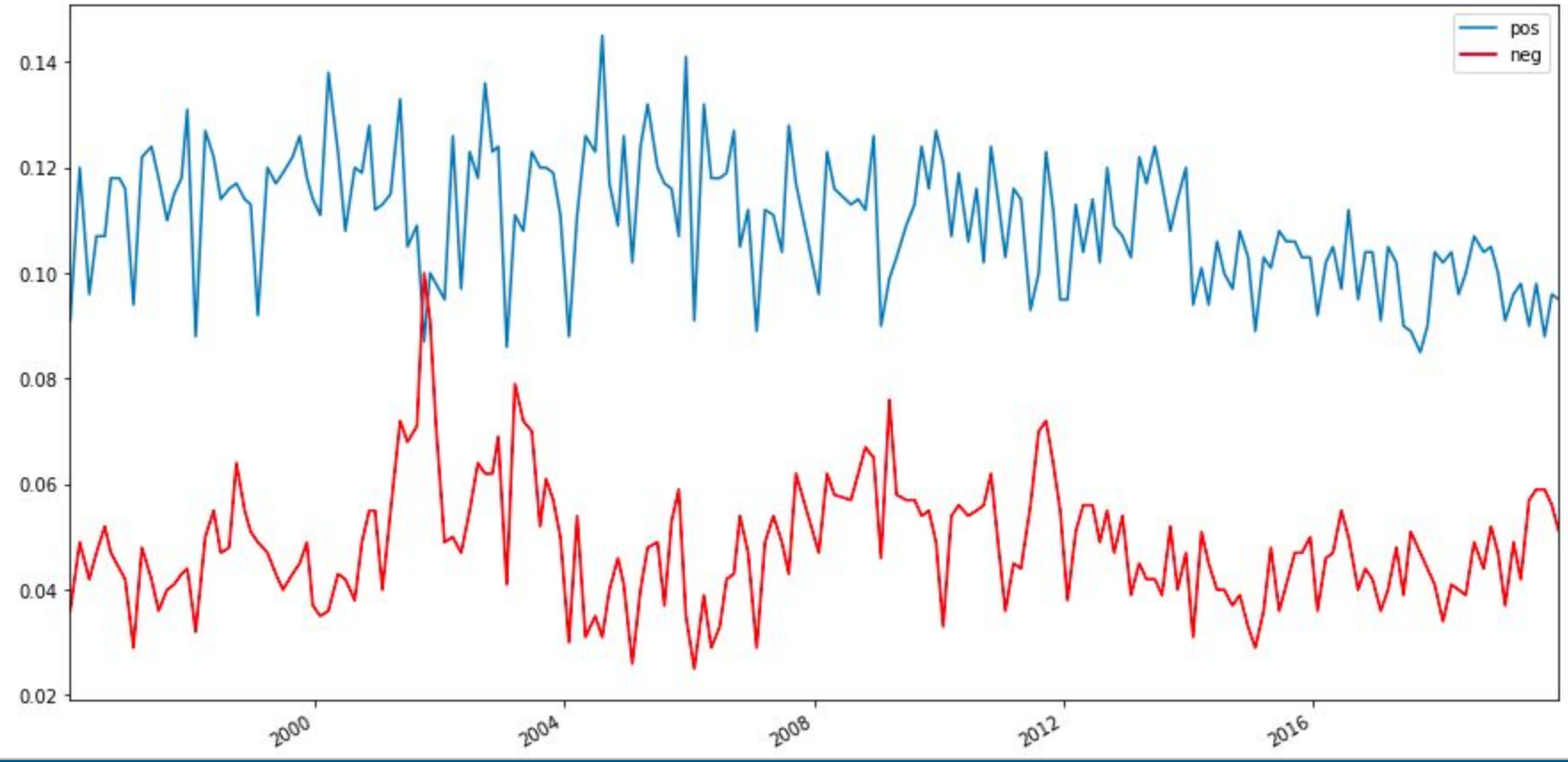
Recap of Dataset Preparation and EDA

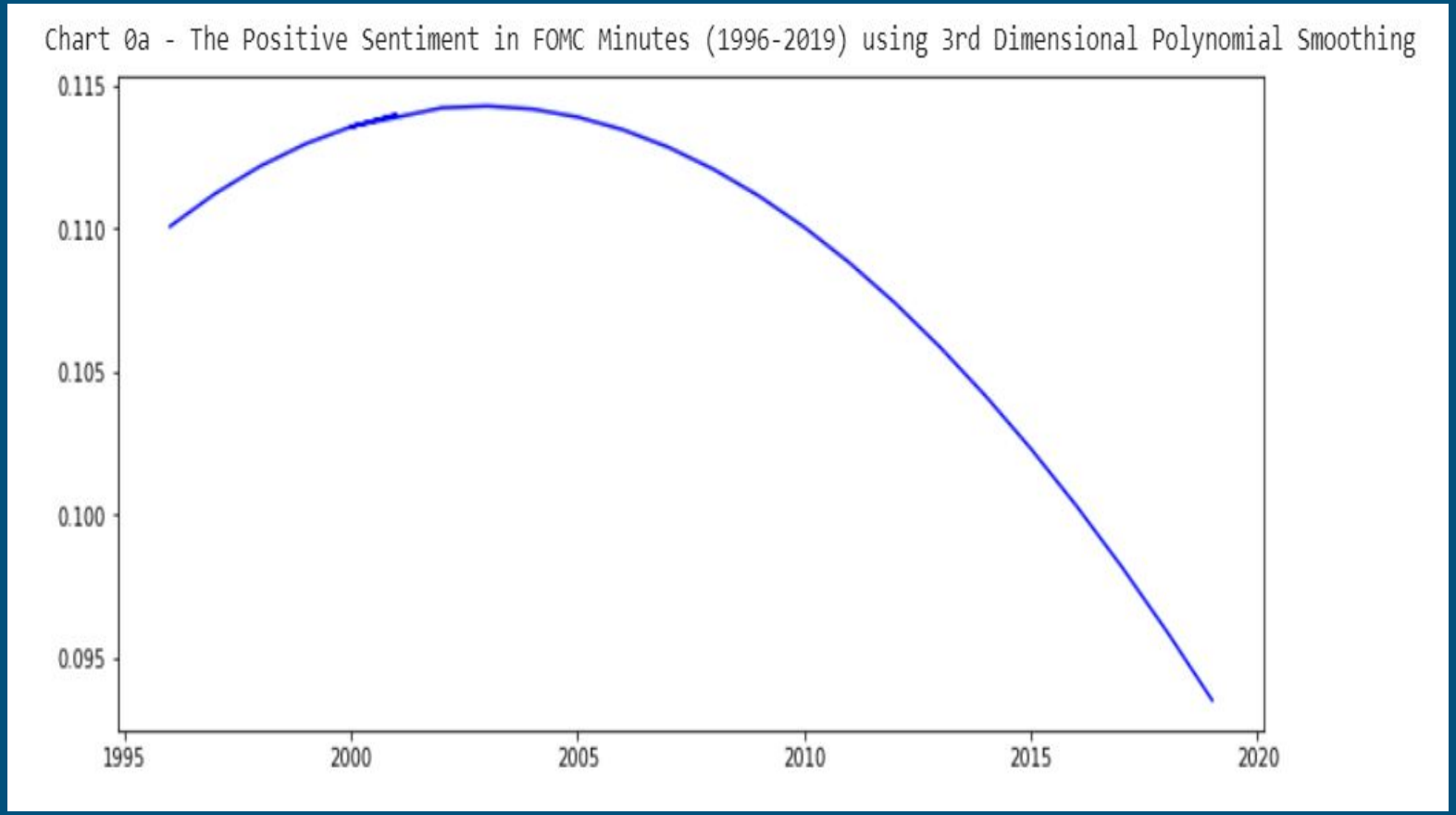
- Dataset: FOMC Meeting Minutes text

- Use: using VADER Sentiment Library generated corresponding scores that assess percentages of meeting minutes that are positive, negative and neutral.

- EDA: Reviewed these scores over time

Chart 0 - The Positive, Negative and Neutral Sentiment in FOMC Minutes (1996-2019)

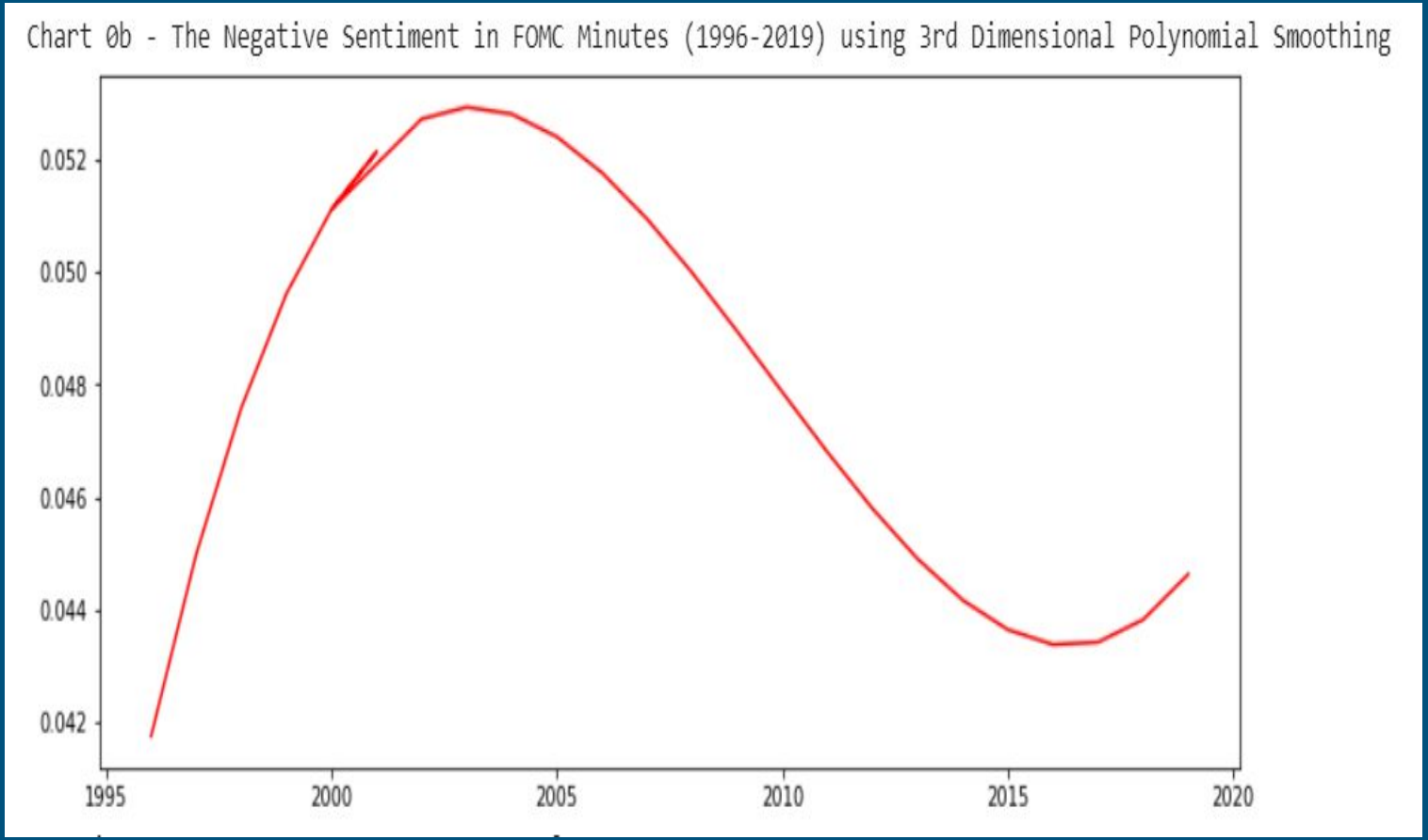




Positive Sentiment

Year

Negative Sentiment

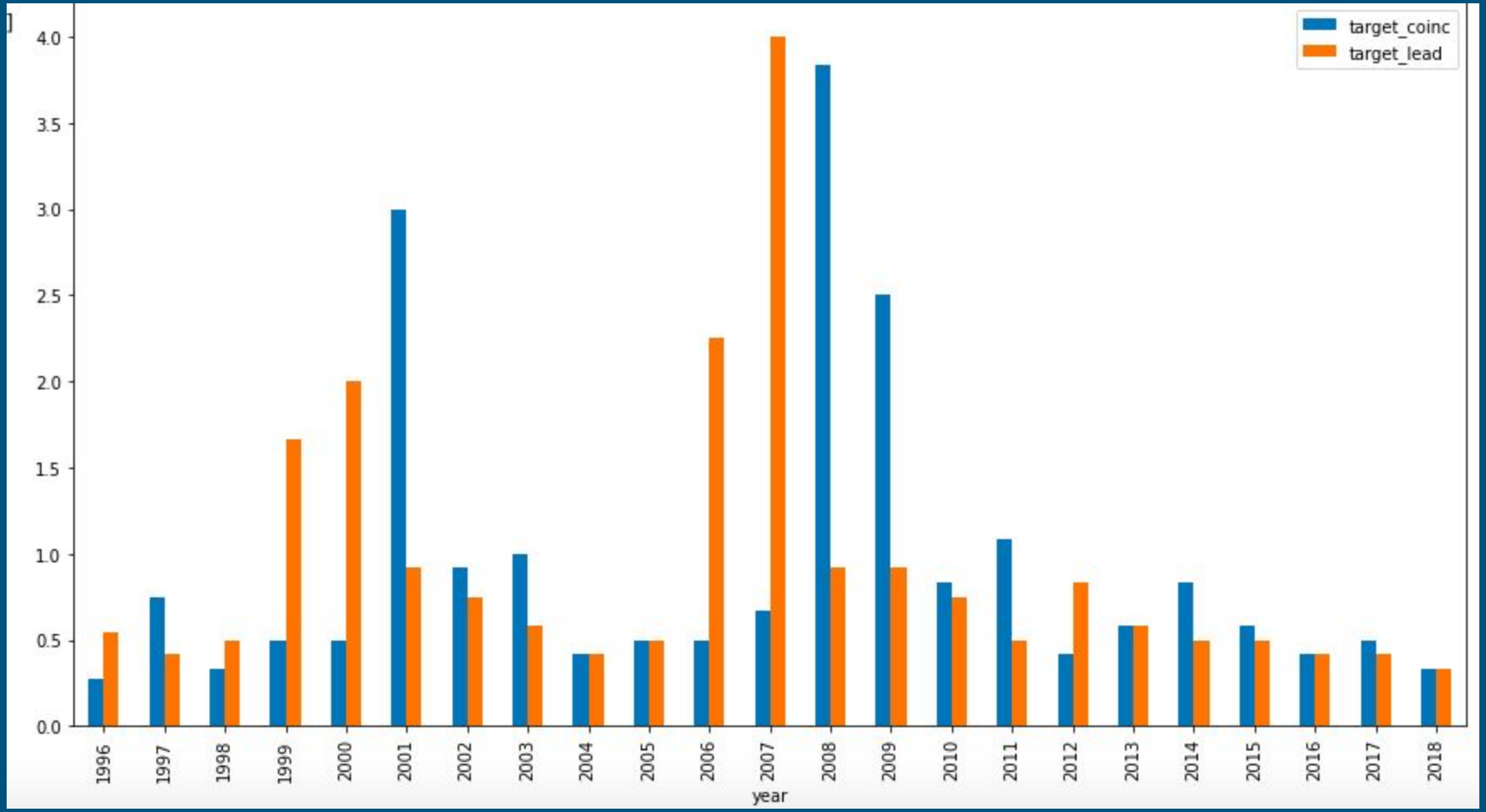


Year

Creating Target Variables
Recipe:

1. Is it a recession or not?
2. Did the Real GDP Decrease this Quarter?
3. Did the Unemployment Rate Decrease this month?
4. Did the Employment Level Decrease this month?

Based on:


target_coinc
target_lead
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
year

Code

```
%%time
!pip install vaderSentiment
import matplotlib.pyplot as plt
import vaderSentiment
from vaderSentiment.vaderSentiment im
import pandas as pd
import re
import requests
from bs4 import BeautifulSoup
from urllib.request import urlopen
import datetime

calendar_url = 'https://www.federalre
recent_socket = urlopen(calendar_url)
```

Leading Indicator as Target Variable

	model / parameters	trainingImportance	testingImportance
0	Gaussian Naive Bayes	0.523810	0.567568
1	MN Naive Bayes (alpha=0.1)	0.489796	0.486486
2	MN Naive Bayes (alpha=0.01)	0.489796	0.486486
3	5-NN (brute)	0.605442	0.405405
4	5-NN (ball-tree)	0.619048	0.432432
5	Logistic Regression (c=1 - defult)	0.489796	0.486486
6	Logistic Regression - (c=100)	0.496599	0.486486
7	Logistic Regression (c=0.1)	0.489796	0.486486
8	Logistic Regression - (c=1000)	0.489796	0.567568
9	Logistic Regression - (c=0.0001)	0.489796	0.486486
10	Logistic Regression (c=0.1)	0.489796	0.486486
11	Logistic Regression - (c=0.001)	0.489796	0.486486
12	LinearSVC (c=1 - defult)	0.489796	0.486486
13	LinearSVC (c=0.1)	0.489796	0.486486
14	LinearSVC (c=100)	0.517007	0.486486

Leading Indicator as Target Variable

Gradient Boosting Classification --- Results Table

	parameters	trainingImportance	testingImportance
0	default parameters	0.979592	0.459459
1	learning_rate=0.01	0.741497	0.405405
2	learning_rate=0.001	0.523810	0.486486
3	max_depth=10	0.986395	0.405405
4	max_depth=5	0.986395	0.405405
5	max_depth=10, learning_rate=0.01	0.972789	0.405405

Leading Indicator as Target Variable

Random Forest Classification --- Results Table

	parameters	trainingImportance	testingImportance
0	default parameters	0.980769	0.461538
1	n_estimators=5	0.916667	0.589744
2	n_estimators=5, max_depth=2	0.487179	0.538462
3	n_estimators=5, max_depth=8	0.801282	0.461538
4	n_estimators=5, max_depth=16	0.903846	0.589744
5	n_estimators=10, random_state=1	0.935897	0.435897
6	n_estimators=10, max_depth=8	0.839744	0.512821
7	n_estimators=12, max_depth=8	0.846154	0.487179
8	n_estimators=100, max_depth=25	0.980769	0.461538

Coincident Indicator as Target Variable

	model / parameters	trainingImportance	testingImportance
0	Gaussian Naive Bayes	0.467949	0.461538
1	MN Naive Bayes (alpha=0.1)	0.435897	0.487179
2	MN Naive Bayes (alpha=0.01)	0.435897	0.487179
3	5-NN (brute)	0.576923	0.564103
4	5-NN (ball-tree)	0.576923	0.512821
5	Logistic Regression (c=1 - defult)	0.435897	0.487179
6	Logistic Regression - (c=100)	0.455128	0.461538
7	Logistic Regression (c=0.1)	0.435897	0.487179
8	Logistic Regression - (c=1000)	0.442308	0.487179
9	Logistic Regression - (c=0.0001)	0.435897	0.487179
10	Logistic Regression (c=0.1)	0.435897	0.487179
11	Logistic Regression - (c=0.001)	0.435897	0.487179
12	LinearSVC (c=1 - defult)	0.435897	0.487179
13	LinearSVC (c=0.1)	0.435897	0.487179
14	LinearSVC (c=100)	0.461538	0.512821

Coincident Indicator as Target Variable

Random Forest Classification --- Results Table

	parameters	trainingImportance	testingImportance
0	default parameters	0.986395	0.486486
1	n_estimators=5	0.918367	0.351351
2	n_estimators=5, max_depth=2	0.598639	0.513514
3	n_estimators=5, max_depth=8	0.802721	0.351351
4	n_estimators=5, max_depth=16	0.918367	0.351351
5	n_estimators=10, random_state=1	0.938776	0.459459
6	n_estimators=10, max_depth=8	0.816327	0.486486
7	n_estimators=12, max_depth=8	0.843537	0.459459
8	n_estimators=100, max_depth=25	0.986395	0.486486
9	default parameters	0.980769	0.461538
10	n_estimators=5	0.916667	0.589744
11	n_estimators=5, max_depth=2	0.487179	0.538462
12	n_estimators=5, max_depth=8	0.801282	0.461538

Coincident Indicator as Target Variable

Gradient Boosting Classification --- Results Table

	parameters	trainingImportance	testingImportance
0	default parameters	0.979592	0.459459
1	learning_rate=0.01	0.741497	0.405405
2	learning_rate=0.001	0.523810	0.486486
3	max_depth=10	0.986395	0.405405
4	max_depth=5	0.986395	0.405405
5	max_depth=10, learning_rate=0.01	0.972789	0.405405
6	default parameters	0.974359	0.487179
7	learning_rate=0.01	0.782051	0.512821
8	learning_rate=0.001	0.500000	0.538462
9	max_depth=10	0.980769	0.512821
10	max_depth=5	0.980769	0.512821
11	max_depth=10, learning_rate=0.01	0.980769	0.538462
12	default parameters	0.974359	0.487179

What's Next?

1. Finalize Analysis
2. More Parameter tuning on classification models that were most promising
3. Determine if there's evidence to support the initial hypothesis that FOMC Minute Sentiments can serve as a leading indicator of a recession
4. What could be improved and done differently?

Sentiment analysis of Federal Open Market Committee (FOMC) minutes: are they leading indicators of a recession similar to traditional Metrics?

Patrick Long