PROJECT UPDATE

Gauging Market Reaction Based on Tweet Sentiment

Why?

- I would like to see if I it is possible to create an application to test if public sentiment expressed on Twitter has an impact on a publicly traded company's value
  - If it turns out public sentiment has an impact, how much of an impact does it have? Is it positive or negative?
  - My current assumptions are that some companies' share prices are influenced heavily by public social media sentiment (especially celebrities) while others are remain somewhat stable with regards to media influence.

Update:
I still want to move forward with previous assumptions, I believe it will be possible to gain some insight using text mining

Data

- My plan involves gathering and analyzing two sets of data with corresponding dates:
  1. Stock market historical data for a certain company
  2. Historical tweet data which corresponds with the market history

Tools and Process

REPLACE

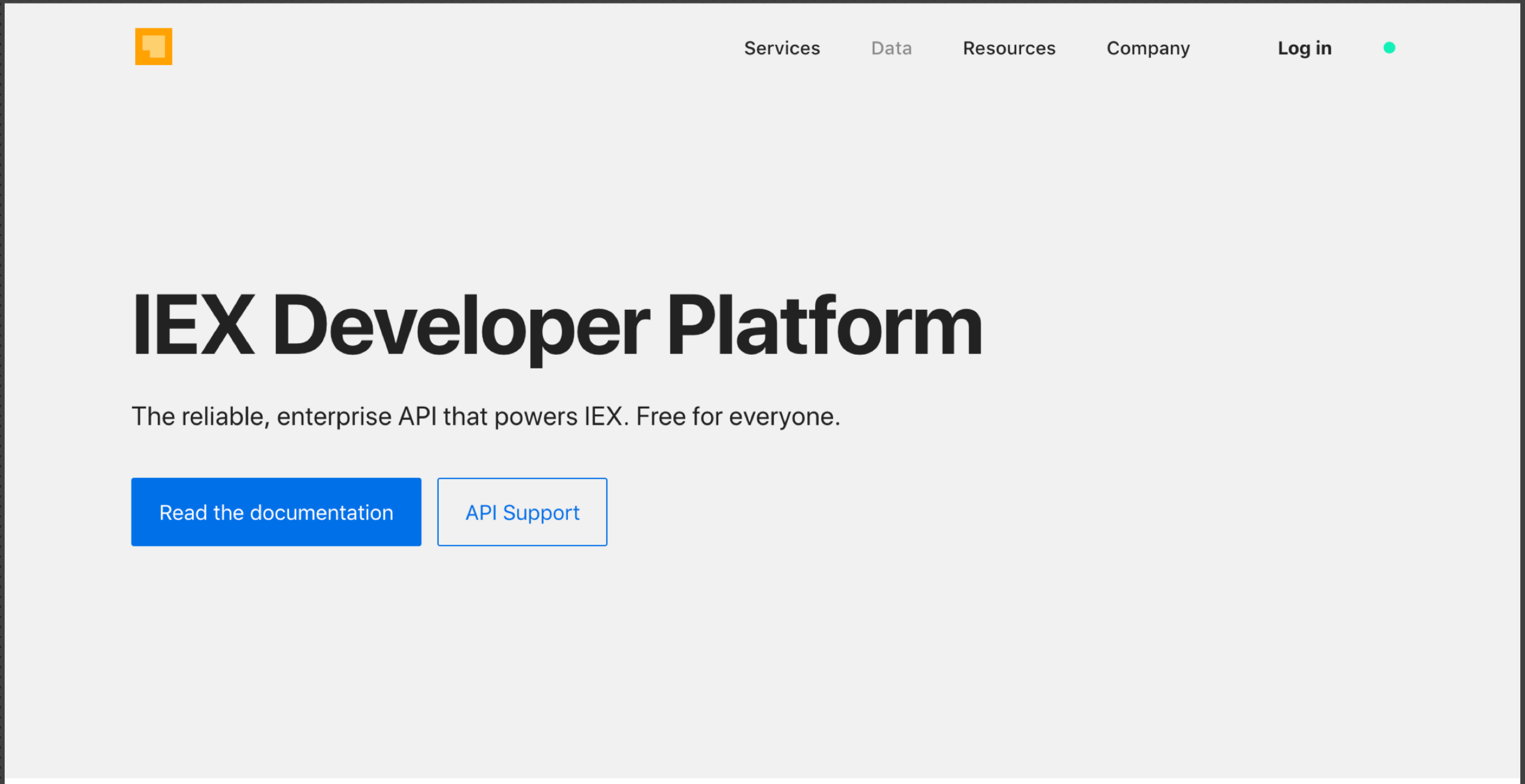
- APIs:
  - AlphaVantage – More granular finance data
  - YahooFinance- For finance data
  - TwitterAPI- For gathering tweets
- Algorithms:
  - To predict sentiment I will primarily use SVM and MNB models
  - I also will try Bernoulli which I feel could yield better results with binary input given the small size of each tweet sample (max 260 characters).
  - Finally, I will do simple linear regression comparing sentiment scores to stock returns and close prices.

Additional Goals

- In addition to stock price/returns I would also like to potentially look at reported revenue for a company compared to sentiment tweets about it.
  - EX: Look at a sample of tweets measuring sentiment compared to past quarter revenue trying to predict financial performance for the quarter.
- I am also interested in doing some topic modeling on the tweets to see what kind of topics influence stock price.

Updated Methodologies

- New API finance platform for access to faster financial information (IEX Trading).



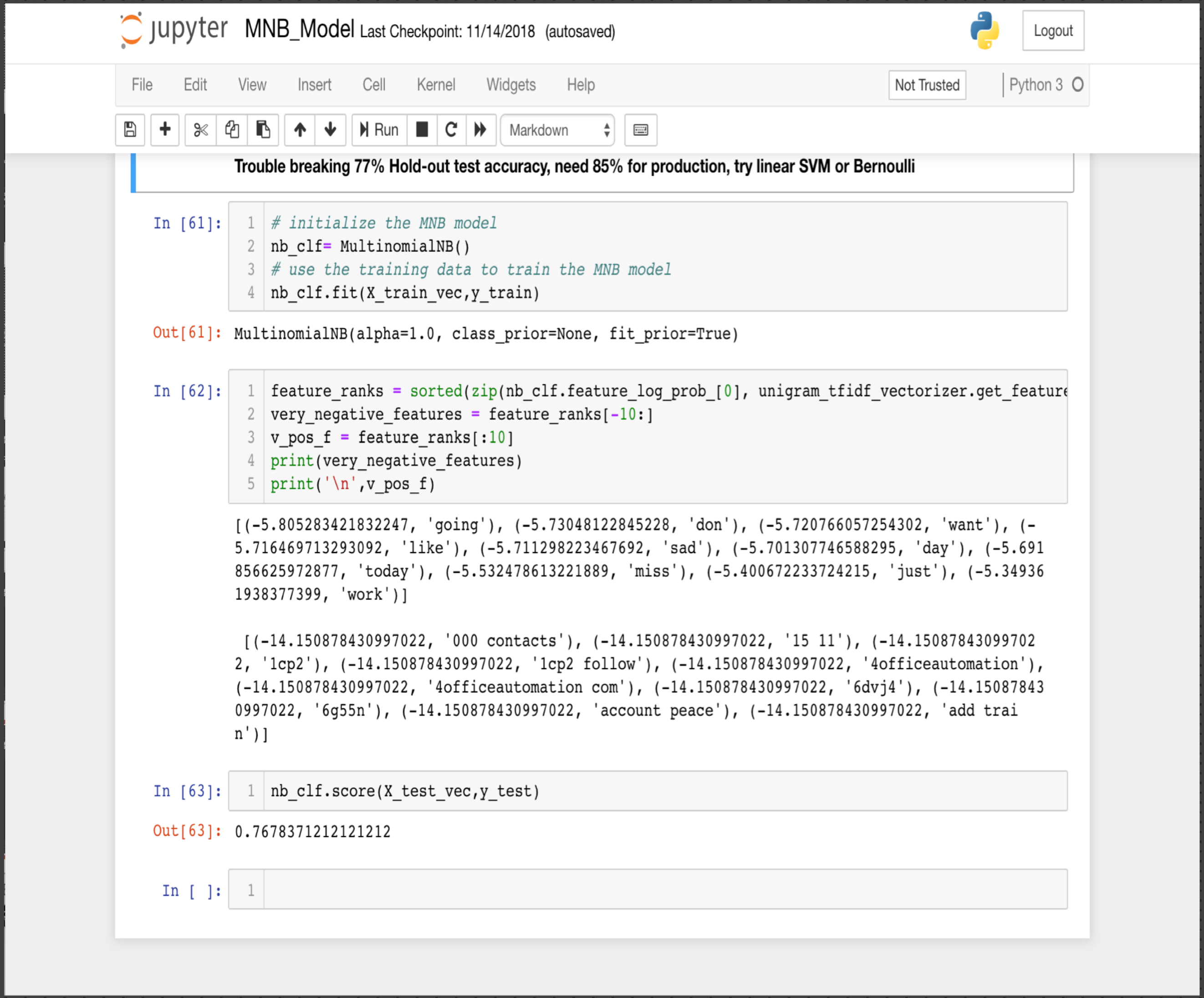
Updated Methodologies

- Moved to OOP python
- Benefits include: easier to request new data.
- Development still done in jupyter.

```
#########################################
# IMPORTS
#########################################

import fix_yahoo_finance as yf
from passwords import API_KEY, API_SECRET_KEY, ACCESS_TOKEN, ACCESS_TOKEN_SECRET

#########################################
# STOCK CLASS
#########################################

class Stock:

    # init requires ticker to create new stock class
    def __init__ (self,ticker):
        self.ticker = ticker

    # takes a start and end date range in YYYY-MM-DD format-
    #  and outputs a pandas dataframe
    def getDataRange(self,start,end):
        try:
            self.data = yf.download(self.ticker,start,end)
        except ValueError:
            print("\n!!! Error getting data- make sure date range is valid !!!\n")
        else:
```

Updated Methodologies

- Developed both MNB and SVM models using 1.4 Million Kaggle tweets.
- Still tuning model parameters
  - MNB Accuracy: 76.78%
  - SVM Accuracy: 78.39%



jupyter MNB_Model Last Checkpoint: 11/14/2018 (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3

Run Markdown

Trouble breaking 77% Hold-out test accuracy, need 85% for production, try linear SVM or Bernoulli

In [61]:
```
# initialize the MNB model
nb_clf= MultinomialNB()
# use the training data to train the MNB model
nb_clf.fit(X_train_vec,y_train)
```
Out[61]: MultinomialNB(alpha=1.0, class_prior=None, fit_prior=True)

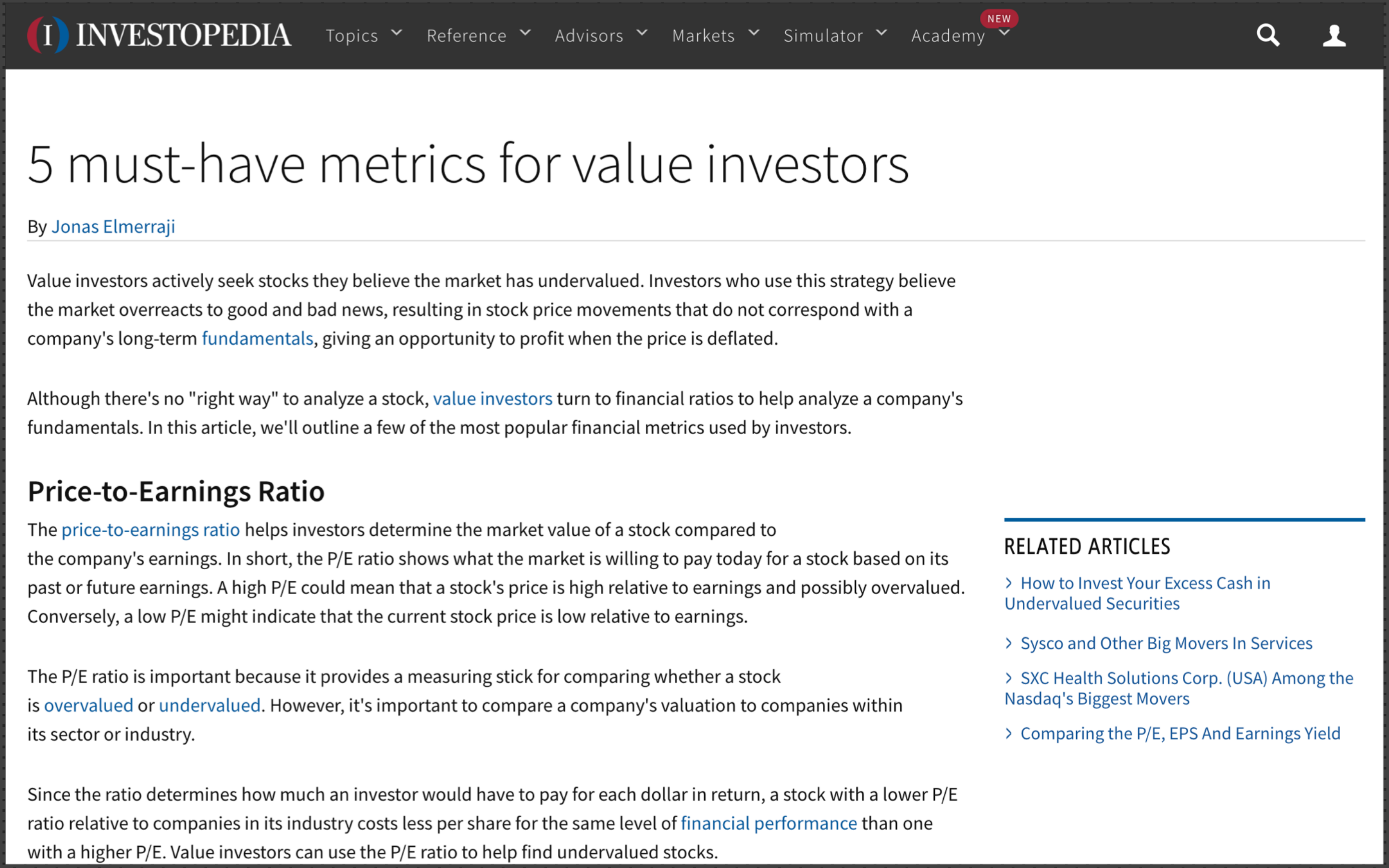
In [62]:
```
feature_ranks = sorted(zip(nb_clf.feature_log_prob_[0], unigram_tfidf_vectorizer.get_feature
very_negative_features = feature_ranks[-10:]
v_pos_f = feature_ranks[:10]
print(very_negative_features)
print('\n',v_pos_f)
```
```
[(-5.805283421832247, 'going'), (-5.73048122845228, 'don'), (-5.720766057254302, 'want'), (-
5.716469713293092, 'like'), (-5.711298223467692, 'sad'), (-5.701307746588295, 'day'), (-5.691
856625972877, 'today'), (-5.532478613221889, 'miss'), (-5.400672233724215, 'just'), (-5.34936
1938377399, 'work')]

 [(-14.150878430997022, '000 contacts'), (-14.150878430997022, '15 11'), (-14.15087843099702
2, '1cp2'), (-14.150878430997022, '1cp2 follow'), (-14.150878430997022, '4officeautomation'),
(-14.150878430997022, '4officeautomation com'), (-14.150878430997022, '6dvj4'), (-14.15087843
0997022, '6g55n'), (-14.150878430997022, 'account peace'), (-14.150878430997022, 'add trai
n')]
```

In [63]:
```
nb_clf.score(X_test_vec,y_test)
```
Out[63]: 0.7678371212121212

In [ ]:

Updated Methodologies

- Updated metrics for correlating sentiment with
- Still learning as a value investor. Potential limitations of domain knowledge



INVESTOPEDIA Topics Reference Advisors Markets Simulator Academy NEW

5 must-have metrics for value investors

By Jonas Elmerraji

Value investors actively seek stocks they believe the market has undervalued. Investors who use this strategy believe the market overreacts to good and bad news, resulting in stock price movements that do not correspond with a company's long-term fundamentals, giving an opportunity to profit when the price is deflated.

Although there's no "right way" to analyze a stock, value investors turn to financial ratios to help analyze a company's fundamentals. In this article, we'll outline a few of the most popular financial metrics used by investors.

Price-to-Earnings Ratio

The price-to-earnings ratio helps investors determine the market value of a stock compared to the company's earnings. In short, the P/E ratio shows what the market is willing to pay today for a stock based on its past or future earnings. A high P/E could mean that a stock's price is high relative to earnings and possibly overvalued. Conversely, a low P/E might indicate that the current stock price is low relative to earnings.

The P/E ratio is important because it provides a measuring stick for comparing whether a stock is overvalued or undervalued. However, it's important to compare a company's valuation to companies within its sector or industry.

Since the ratio determines how much an investor would have to pay for each dollar in return, a stock with a lower P/E ratio relative to companies in its industry costs less per share for the same level of financial performance than one with a higher P/E. Value investors can use the P/E ratio to help find undervalued stocks.

RELATED ARTICLES

- How to Invest Your Excess Cash in Undervalued Securities
- Sysco and Other Big Movers In Services
- SXC Health Solutions Corp. (USA) Among the Nasdaq's Biggest Movers
- Comparing the P/E, EPS And Earnings Yield

Questions?
Comments?
Suggestions?