

# Trading on Twitter

## Using Social Media Sentiment to Better Financial Predictions

Zinan Chen  
Siyu Liu  
Qiuhaio Chengyong  
Tyler McMurray  
Senbo Zhang



# Our Inspiration:

## Emotions in Finance decision-making

---

**Greed** and **fear** are the two main drivers of the stock market. It turns out that positive and negative emotions in social media messages, such as Twitter, can be used to predict daily changes or trends in stock prices.

Although news certainly affects stock market prices, public sentiment may also play an equally important role. Behavioral finance further proves that financial decision-making is largely driven by emotions. It is reasonable to assume that public sentiment can drive stock market prices like news.



## Our Inspiration

Facts about Twitter

---

Over

**300 Million** Users

**500 Million** Tweets sent per day

**1,01 billion** U.S. dollars

The Twitter logo, consisting of the word "twitter" in a lowercase, sans-serif font, followed by a white bird icon (the Twitter bird) on a blue rectangular background.

twitter

Going through the analysis and utilizing tools like machine learning techniques to predict the stock market might not be fruitful in its power to benefit from the stock market but it might be useful for better understanding relevant concepts, relationships, and possibly be used to help create policies around the stock market.

# More in Depth Down Goal

**Greed** and **fear** are the two main drivers of the stock market. It turns out that positive and negative emotions in social media messages, such as Twitter, can be used to predict daily changes or trends in stock prices.



Step 1:

- Use NLP techniques of various data sources to feature engineer for machine learning and time series



Step 2:

- Merge financial information to the engineered features



Step 3:

- Use all Machine Learning Techniques to predict classification buy or sell



Step 4:

- Finally try to interpret some of the things that might be the reason causing these shifts in states

Let's do it!

# Data Used

---



## Twitter Dataset

The Twitter data provided by a Twitter Intelligence and Analytics Website:  
[Followthechastag.com](http://Followthechastag.com).



## Twitter information

Our dataset includes tweets with cashtag and stock code in 79 days including tweets, likes, share and number of followers



## Yahoo Finance

We also pull the financial information from Yahoo Finance, Nasdaq 100 stocks' financial information during 79 days from 2016/03/28 to 2016/06/15

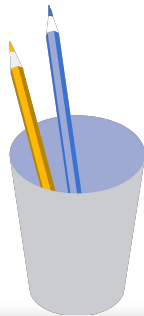
## Best Performance

---

Boosting

**Accuracy**  
**58%**

*You may want to get more  
calculation*



Random Forest

**Accuracy**  
**60%**

*You can try to make money on  
stock*



Selected

Linear Regression

**Accuracy**  
**53%**

*Predict stock buy or sell is hard  
You need a rest*





# Results 1

Best Algorithm:  
Random Forest

Best Window:  
7-8 days

Best Windows:

7-8 Days

	Reference	
Prediction	0	1
0	36	17
1	41	50

Accuracy : 0.5972

95% CI : (0.5123, 0.678)

No Information Rate : 0.5347

P-value [Acc > NIR] : 0.077378

Perform better on  
'Buy' prediction

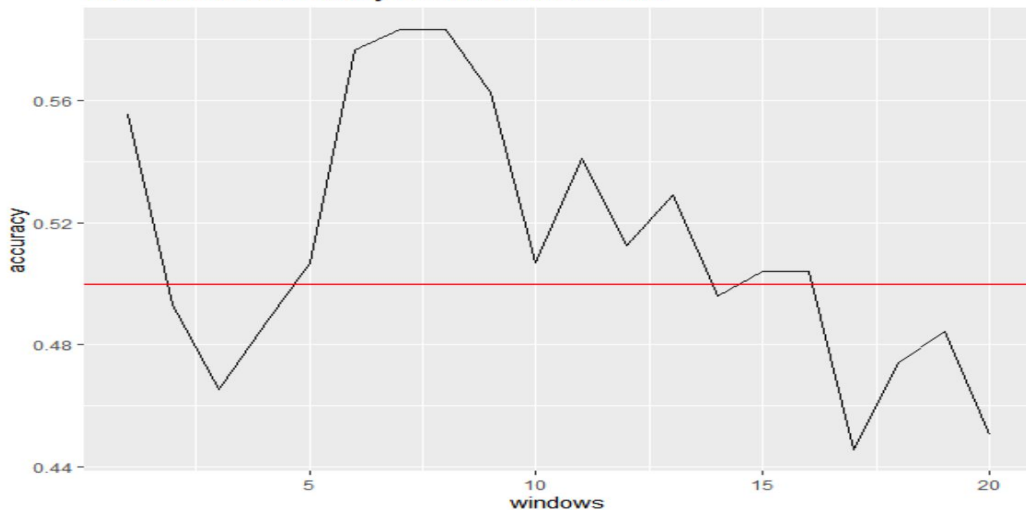
Hit 75%

Significant level 90%

Benchmark: NIR

It performs better  
when predict Increase

Classification Accuracy on Different Windows







# Results 2

Best Window:

54 Days

Name	Gains Or Losses	Beat Winner	Beat Aggregate
Winner	-15.02	No	No
Basket1	-14.14	Yes	No
Basket2	-20.12	No	No
Basket3	-22.58	No	No
Basket4	-10.32	Yes	Yes
Basket5	-15.4	No	No
Basket6	-9.02	Yes	Yes
Basket7	-11.59	Yes	No
Basket8	-9.46	Yes	Yes
Basket9	-8.92	Yes	Yes
Basket10	-14.12	Yes	No

Best Algorithm:

Linear Regression

Winning Basket: Top 5 Predicted Gainers

Tickers included: ULTA, NFLX, MLM, LMT, & ROP

Random Baskets: 10 Baskets  
Randomly put together containing 5 stocks

Used for comparison of winning Basket  
Aggregate Metric:  
Average performance of all stocks in data

The Winning Basket did not beat the market (Aggregate)

70% of Random Baskets beat the Winner

THANK YOU

