



DataX

About Me:

TwitTrade

Project Overview

Yicheng Chen, Tinashe Giyavha, Frank Mei, Dan Shen

Requirements List Format

Top Technical Requirements

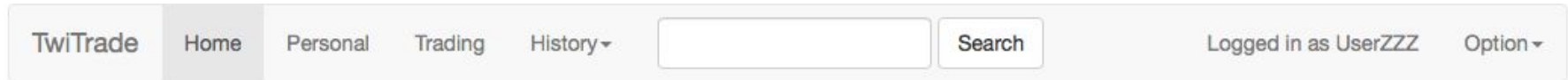
1. Twitter API
2. Stock API
3. Sentiment analysis
4. Prediction models
5. Website
6. Mobile Application
7. Database - user info
8. Search keywords reference
9. Web scraping

Requirements List Format

Top User Requirements

1. Users want a graphic UI
2. Section for trending stocks
3. Personal interests section
4. Email notification for constant updates
5. Users would also want a mobile app
6. Ability to buy/sell securities

Place an illustration of your Final User Interface below



Welcome to TwiTrade.

Use our Twitter-based stock price prediction and make money!

Trending Tweets

The following tweets are trending.

Related companies and industries:

XXXXXXXXXXXX

YYYYYYYYYYY

[View details »](#)



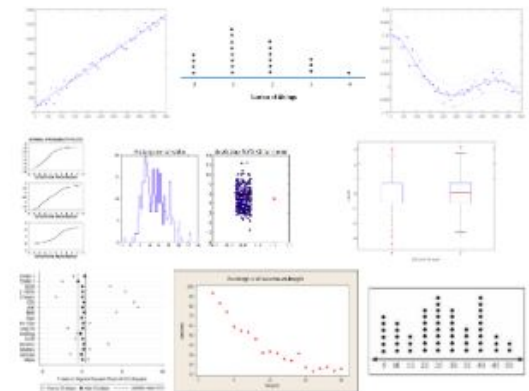
Historical Stock Prices

[View details »](#)



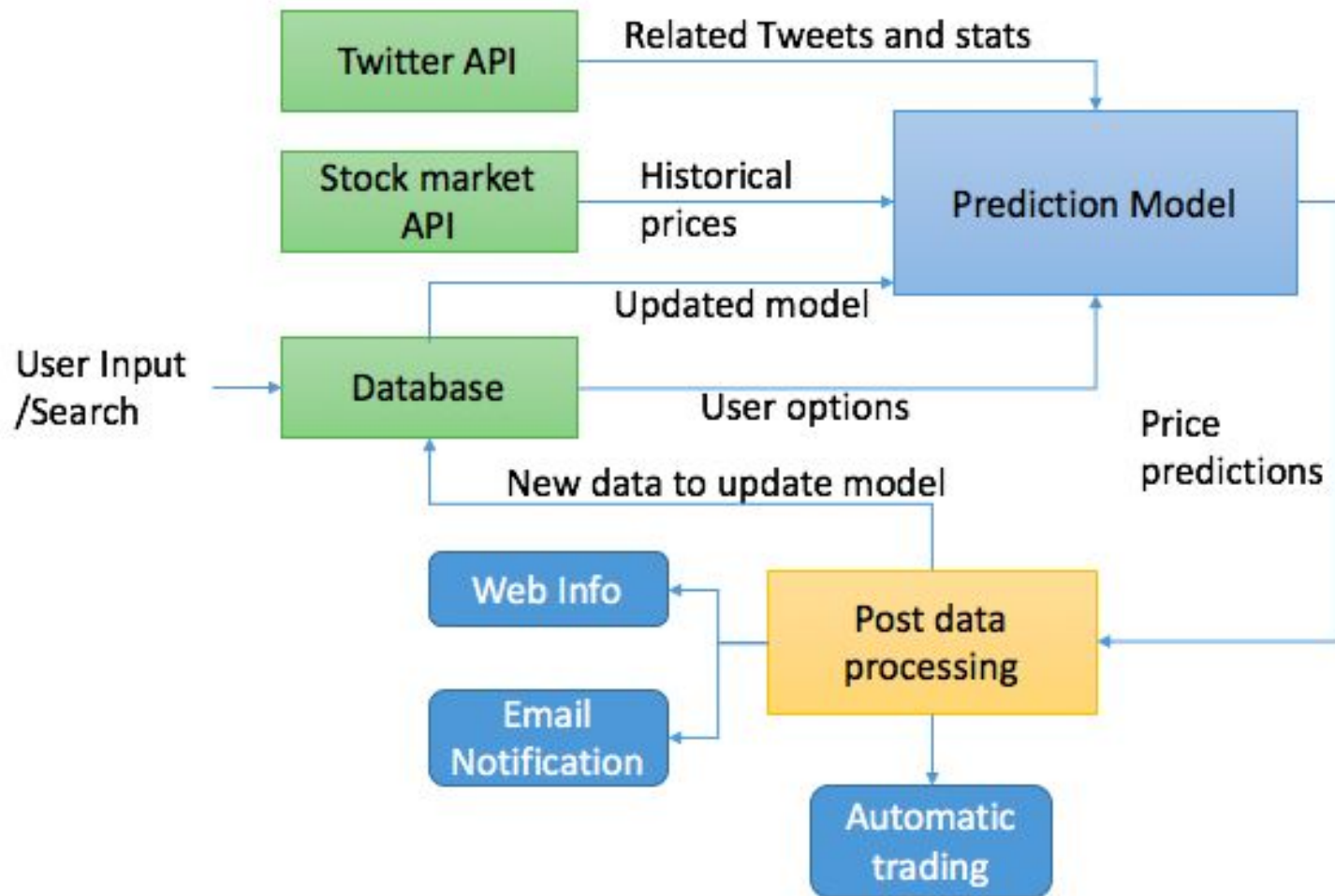
Statistics and Predictions

[View details »](#)



Data X

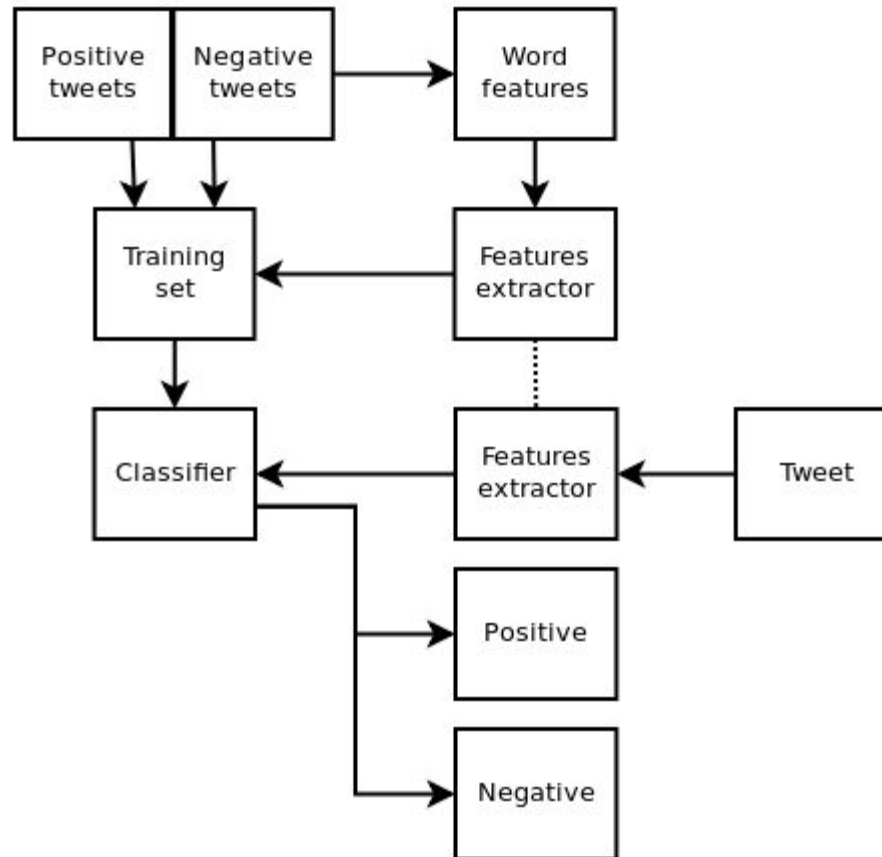
Data Model



Week 7 Agile Sprint Cycle

Stock API and training data acquisition -- Dan
Website UI, Twitter data acquisition -- Yicheng
Sentiment analysis coding -- Tinashe
Prediction model research and building -- Frank

Sentiment analysis



Sentiment analysis - code snippets

```
In [1]: # List containing positive tweets
pos_tweets = [('I love this car', 'positive'),
               ('This view is amazing', 'positive'),
               ('I feel great this morning', 'positive'),
               ('I am so excited about the concert', 'positive'),
               ('He is my best friend', 'positive')]
```

```
In [17]: def extract_features(document):
          document_words = set(document)
          features = {}
          for word in document_words:
              features['contains(%s)' % word] = (word in document_words)
          return features
```

```
In [38]: tweet = 'Larry is my friend'
          print (classifier.classify(extract_features(tweet.split())))
```

positive

Potential User Survey

Basic Info Questions:

- 1) Gender:
 - a) Male
 - b) Female
- 2) Age:
 - a) <21
 - b) 21-30
 - c) 31-40
 - d) 41-50
 - e) >50
- 3) Do you invest in stocks/securities?
 - a) Yes
 - b) No
- 4) How often do you invest?
 - a) More than twice a week
 - b) Once a week
 - c) Once a month
 - d) Less than once a month
- 5) Which segment(s) of the market are you most interested in?
- 6) What would you like to know more about the market to make better investments?
- 7) Do you think Twitter provides information about the moving trends of the market?

Product Preference Questions:

We are building a product that can provide trading recommendations according to tweets on Twitter.

- 1) What platform do you prefer?
 - a) Mobile app
 - b) Website
 - c) Both
- 2) What do you expect this product to provide? (multiple)
 - a) Overall social sentiment on Twitter towards a specific company
 - b) Recommendations on buy/sell for a specific stock
 - c) Analysis on popular/trending companies and industries on the market
 - d) Option to trade stocks directly on the site
- 3) What do you want in the interface? (multiple)
 - a) Graphical UI on predicting trends for stocks
 - b) Personalizable tracking section
 - c) Email notification setting
- 4) What will you most likely use this product for your investment decisions?
 - a) I will not use it for my investment decisions. This is just something for fun.
 - b) I will take it into account for investment decisions. It provides new perspective and insights to the market.
 - c) I will make some investments based on its recommendations.
 - d) I will fully rely on this to make my investment.
- 5) What else would you like to see in this product?



Twitter API

Twitter provides API

Python: Tweepy

```
# set authorization
auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
```

```
# create API object
api = tweepy.API(auth)
```

Twitter API -- Home timeline feeds

```
# get some tweets from home timeline
public_tweets = api.home_timeline(count=5)
for tweet in public_tweets:
    print(tweet.text, '\n')
```

.@POTUS: "Our current immigration system costs America's taxpayers many billions of dollars a year." #TrumpAddress... <https://t.co/DAhPKMPJTd>

.@POTUS: "Nations around the world, like Canada, Australia and many others, have a merit-based immigration system."... <https://t.co/KlgzRfuC5A>

"They wanted me to ride one [a Harley Davidson], and I said 'no thank you,'" Trump says to laughs.

.@POTUS: "Protecting our workers also means reforming our system of legal immigration." #TrumpAddress... <https://t.co/07YXtXvCXR>

President Trump: "My economic team is developing historic tax reform that will reduce the tax rate on our companies" <https://t.co/9P4OcR9dbi>



Twitter API -- Recent statuses of certain user

```
# get N most recent statuses of a user
recent = api.user_timeline('realDonaldTrump', count=5)
for i, t in enumerate(recent):
    print(i, t.created_at, t.text, '\n')
```

```
0 2017-03-01 01:42:16 #JointSession #MAGA🇺🇸
https://t.co/RD06Jt2pip
```

```
1 2017-03-01 01:30:49 Join me live at 9:00 P.M.
#JointAddress https://t.co/J882zbyVkJ https://t.co/gTtK3vJmkU
```

```
2 2017-02-28 02:43:03 I will be interviewed on @foxandfriends at 6:00 A.M.
. Enjoy!
```

```
3 2017-02-27 21:48:21 GOP now viewed more favorably than Dems, in Trump e
ra (per NBC/WSJ poll) via @HotlineJosh:
https://t.co/Soebxtg2Zd
```

```
4 2017-02-27 17:06:23 Great meeting with CEOs of leading U.S. health insu
rance companies who provide great healthcare to the American peo... https://t.co/60jigi6ffs
```



Twitter API -- real time

```
# real time listening of Twitter
class listener(StreamListener):

    def on_data(self, data):
        j = json.loads(data) # convert the json to dictionary
        print('#####', j['text'], '\n')
        return(True)

    def on_error(self, status):
        print(status)

# streaming API
twitterStream = Stream(auth, listener())
twitterStream.filter(track=["#Trump"])
```

```
##### RT @ASE: #Trump lies his way through trying to defend why
he is focusing on Muslims to counter terrorism threat in America. #Join
tSession
```

```
##### (#TrumpAddress #JointAddress) #Trump on immigration: "Our
current #immigration system costs American taxpayers billions of dollar
s a year."
```

```
##### Ditzzy Democrats Wear White In A Nod To The Women's Suffrag
e Movement, Because Trump... https://t.co/wj6NSshxei #Trump #TCOT
```

```
##### I was thinking the same thing. #ISIS #Trump #Dems # https://t.co/hggAqUHUom
```

```
##### RT @UniteBlueSC: #Trump Claims everyone wants what he is s
aying, yet protests are overwhelming and he has no election mandate...
```


Stock price API

Historical stock price of 'YHOO'

	Date	Close price
0	2016-03-11	33.810001
1	2016-03-10	32.82
2	2016-03-09	33.509998
3	2016-03-08	32.93
4	2016-03-07	33.959999
5	2016-03-04	33.860001

Historical stock price of 'GOOG'

	Date	Close price
0	2016-03-11	744.869995
1	2016-03-10	732.169983
2	2016-03-09	725.409973
3	2016-03-08	713.530029
4	2016-03-07	712.799988
5	2016-03-04	730.219971



Prediction Model

X--Twitter Variables:

- content of the tweet, sentiment
- post time, mean frequency between posts
- verified account or not
- # of likes
- # of retweets

Y--Stock Price

