Using Machine Learning Methods to Assess the Impacts of Twitter Sentiment on a Company's Stock Trend

Arturo Valera, Pipi Gao, Sarah Kim

Project Objective

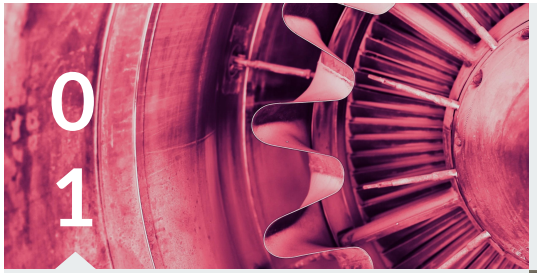
To what extent does Twitter reflect a company's value?

Does a CEO of a company actively use Twitter space to promote their brand values?

Does Twitter accurately capture the public perception of a company?

Is Twitter sentiment correlated with the stock market trend?

Approaches



Sentiment Analysis on Twitter

Twitter Posts & Comments

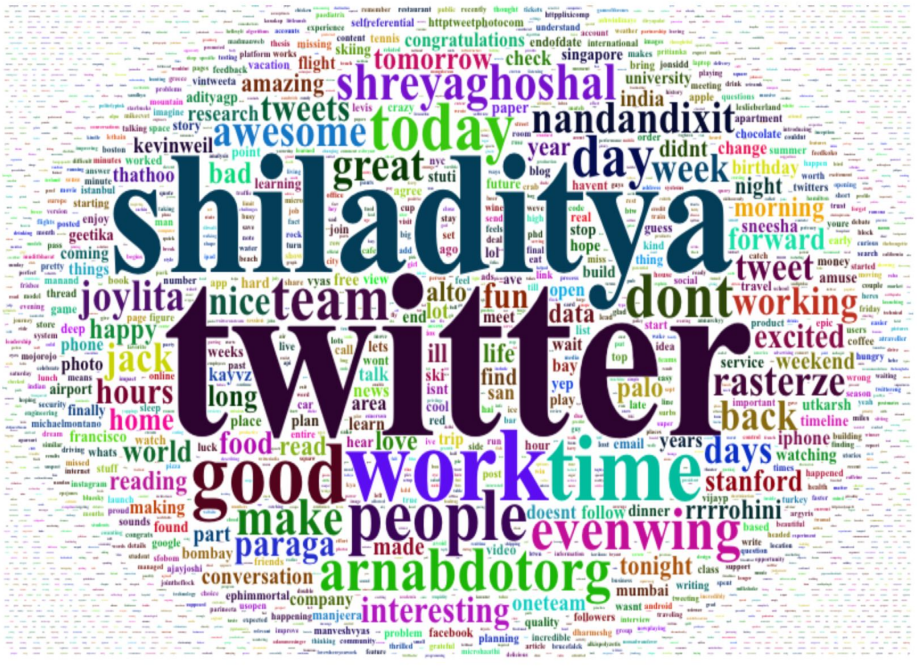
Predicting Stock trend with Sentiments

Data Source 1

- Twitter API v2
- Real-time source of twitter data
  - get_timeline() -> CEO accounts
  - search_tweets() -> hashtags
- Limitations
  - search_tweets() Only returns data from the past 6-9 days
  - GetOldTweets package no longer works since 2020

Twitter Posts Wordclouds & Comments from the CEO’s

year:131


shiladitya
twitter
good
work
time
people
make
evenwing
arnabdotorg
shreyaghoshal
today
nandandixit
awesome
great
tweets
day
week
dont
tweet
working
rasterze
back
days
stanford
rrrrohini
tonight
interesting
oneteam
paraga
conversation
company
manjeera
manvesh
vyas
joylita
nice
team
jack
hours
home
food
reading
making
found
part
bombay
ajayjoshi
ephimmortal
usopen
happening
quality
problem
facebook
planning
incredible
dharmesh
group
employing
thrilled
grateful
introhustle
adobeinc
followers
wasnt
android
trending
mumbai
writing
class
support
based
beautiful
health
location
question
write
apply
opportunity
turkey
faster
happened
timeline
utkarsh
watching
iphone
building
finding
stories
times
email
years
hour
post
hear
love
trip
learn
plan
place
entire
car
cool
red
luck
francisco
watch
world
whats
missed
internet
stuff
instagram
launch
proud
making
sounds
found
google
student
afobom
ajayjoshi
friends
video
tennis
endofdate
international
congratulations
tomorrow
check
singapore
makes
bring
jonsidd
university
india
apple
helicberland
apartment
chocolate
change
summer
feedbacks
birthday
night
twitters
worth
morning
sneesha
forward
money
amuse
coffee
strezdlcr
search
collexis
excited
weekend
hungry
service
coffee
idea
product
epic
users
start
wait
bay
top
list
yep
play
palo
late
important
based
dinner
follow
doesnt
vijaya
turkey
faster
happened
health
location
question
write
apply
opportunity
writing
mumbai
class
support
wasnt
android
trending
followers
interview
manveshvyas
problem
facebook
planning
incredible
dharmesh
group
employing
thrilled
grateful
introhustle
adobeinc
quality
happening
manjeera
ephimmortal
usopen
company
conversation
paraga
bombay
part
making
found
google
student
afobom
ajayjoshi
friends
video
made
arnabdotorg
evenwing
people
make
good
work
time
days
stanford
rrrrohini
tonight
interesting
oneteam
rasterze
back
utkarsh
timeline
watching
iphone

conversation_id	child_tweet_id	tweet_text
1465050134274785282	1465052303006830594	@tim_cook can you help me with money? 🙏😔
1465050134274785282	1465050583937605633	@tim_cook Please sir please give me only 4 million dollars please I request you 😭😭😭
1465050134274785282	1465050548789387267	@tim_cook I am gay gay life is too much difficult in india Plus Poverty my last hope please I request you 😭😭😭
1465050134274785282	1465050448465829893	@tim_cook Please sir please give me only 4 million dollars please I request you 😭😭😭 my father work in construction line and his age 60 years old and I am just 19 year old please I request you 😭😭😭

conversation_id	child_tweet_id	tweet_text
1465050134274785282	1465056353391628293	@tim_cook It's so heart warming to see you reach out to other diversities, as little as it may seems (just a tweet) it makes such a big happiness to so many who appreciate your words. Thank you!
1463236596505841665	1463282956659793927	@tim_cook Loving my MacBook Pro 14”. It amazes me each time I use it. Love Apple. Happy thanksgiving sir
1463236596505841665	1463236916527214601	@tim_cook @Apple you deserve a medal for the democratisation of performance, energy efficiency and opportunity
1463236596505841665	1463236683470626822	@tim_cook Fantastic as always brother 💗🏆

Sentiment Analysis

Sentiment Score

- Text mining: recent tweets under company stock ticker hashtags #AAPL
- Package: sentimentr
- Method: augmented dictionary lookup

```
sentiment("I haven't been sad in a long time.")
 element_id sentence_id word_count sentiment
          1           1          8 0.1767767
```

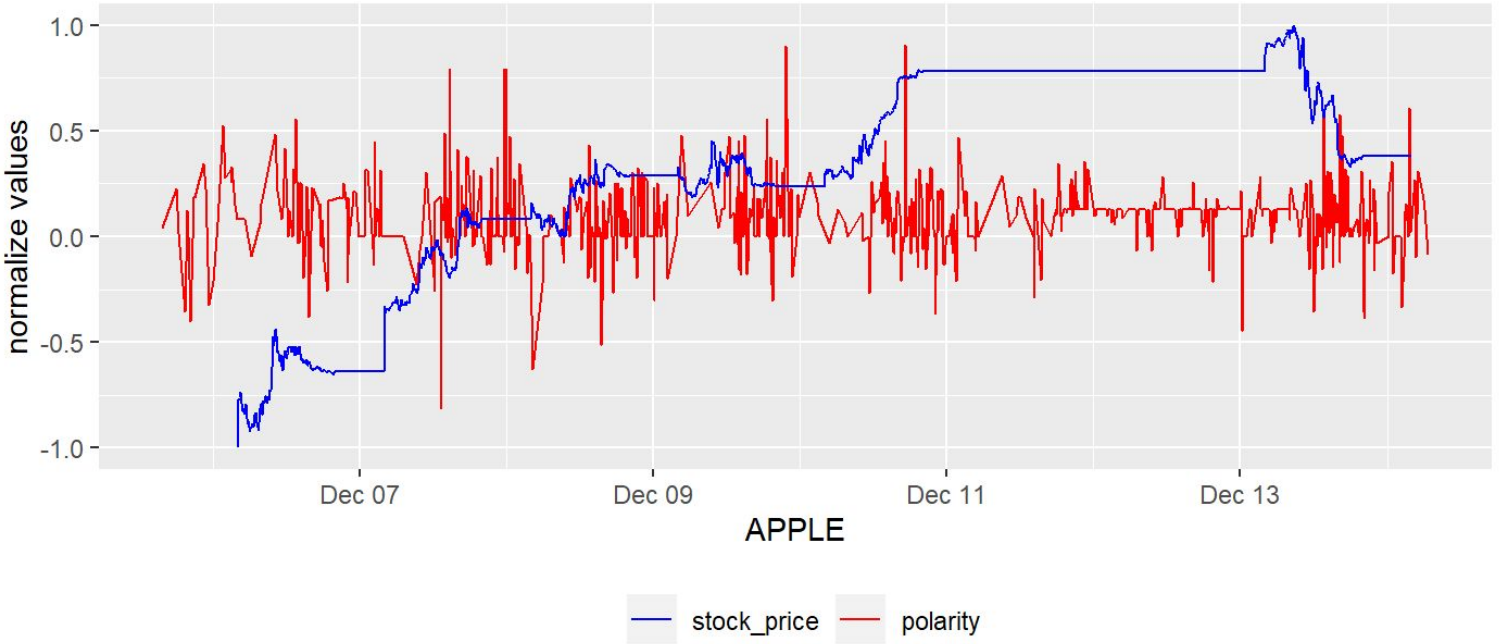
text	sentiment
Dear $AAPL. I love you. That is all. #AAPL	0.41666667
Intraday Technical Analysis on Apple Inc. Read More 👉 https://t.co/D8rGp5ylku #A...	0.00000000
$AAPL doing its best to defend the $160 support line. #AAPL	0.39528471
Thanks papa Elon for given our investors massive gains Yes we trade for yo to get ...	0.8379567

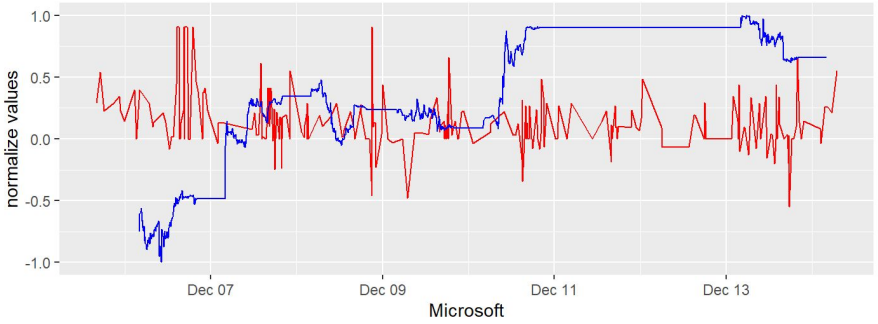
Data Source 2

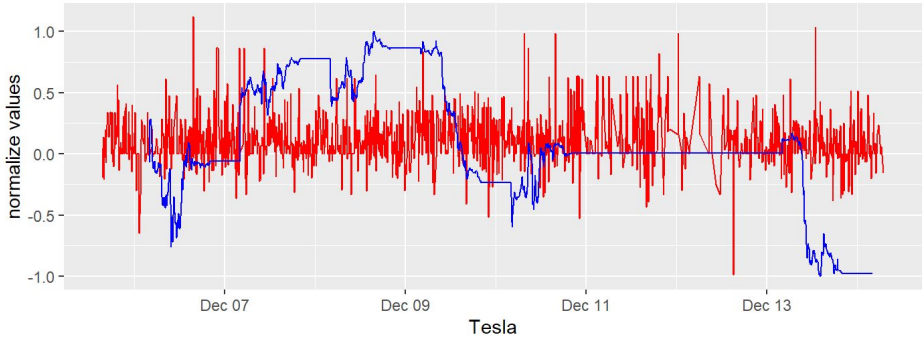
+

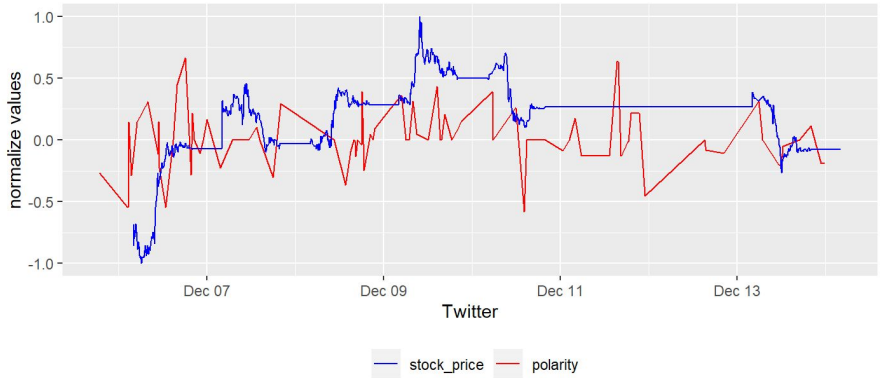
- **Alpha Vantage**
  - Free, time series stock data api
  - Open, close value of given time interval
- **Data Wrangling**
  - Get stock values in 5 min-intervals
  - Round tweet time to the nearest 5 min
  - full_join!

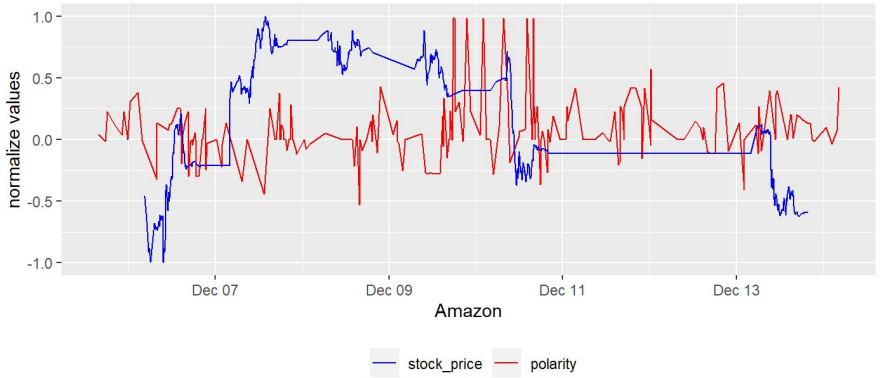
Plots


1.0
0.5
0.0
-0.5
-1.0
normalize values
Dec 07
Dec 09
Dec 11
Dec 13
APPLE
stock_price
polarity


normalize values
1.0
0.5
0.0
-0.5
-1.0
Dec 07
Dec 09
Dec 11
Dec 13
Microsoft


normalize values
1.0
0.5
0.0
-0.5
-1.0
Dec 07
Dec 09
Dec 11
Dec 13
Tesla


normalize values
1.0
0.5
0.0
-0.5
-1.0
Dec 07
Dec 09
Dec 11
Dec 13
Twitter


normalize values
1.0
0.5
0.0
-0.5
-1.0
Dec 07
Dec 09
Dec 11
Dec 13
Amazon

stock_price
polarity

stock_price
polarity

Twitter Sentiment and Stock Price

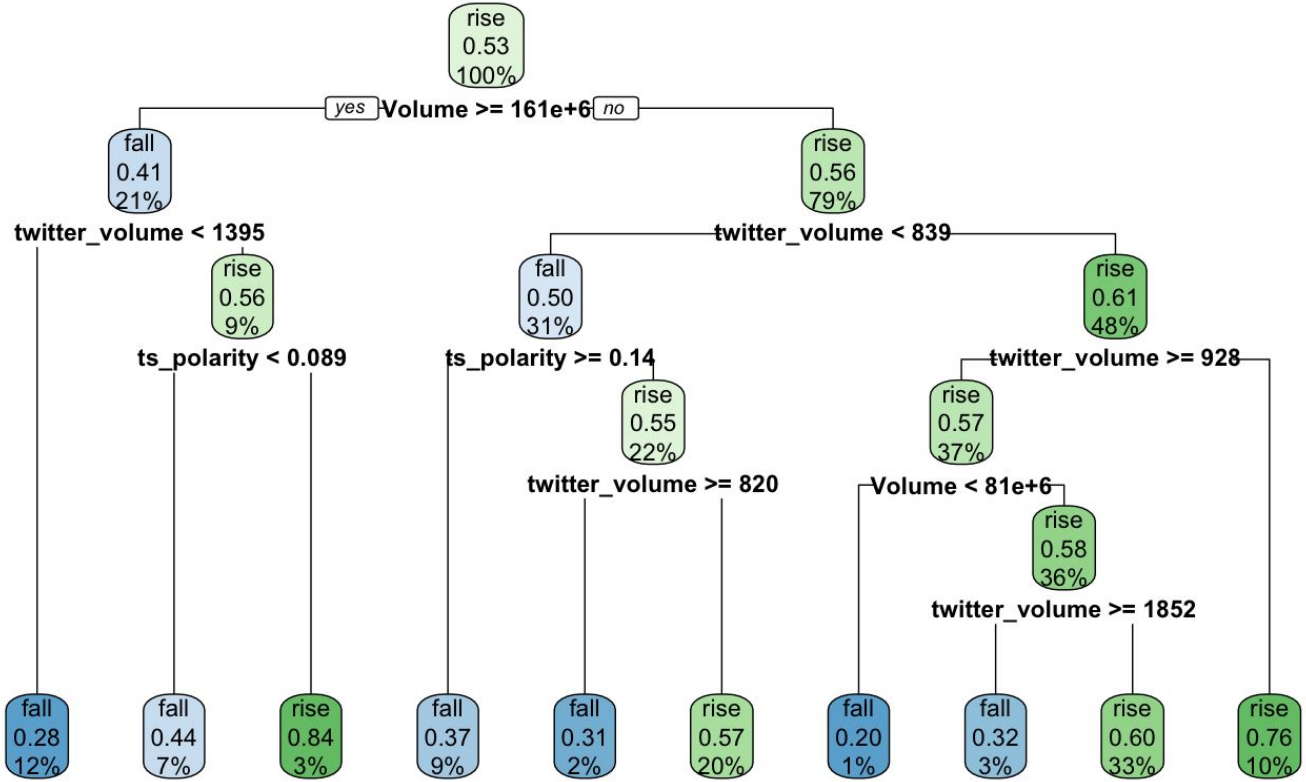
Proposed solution

Dataset

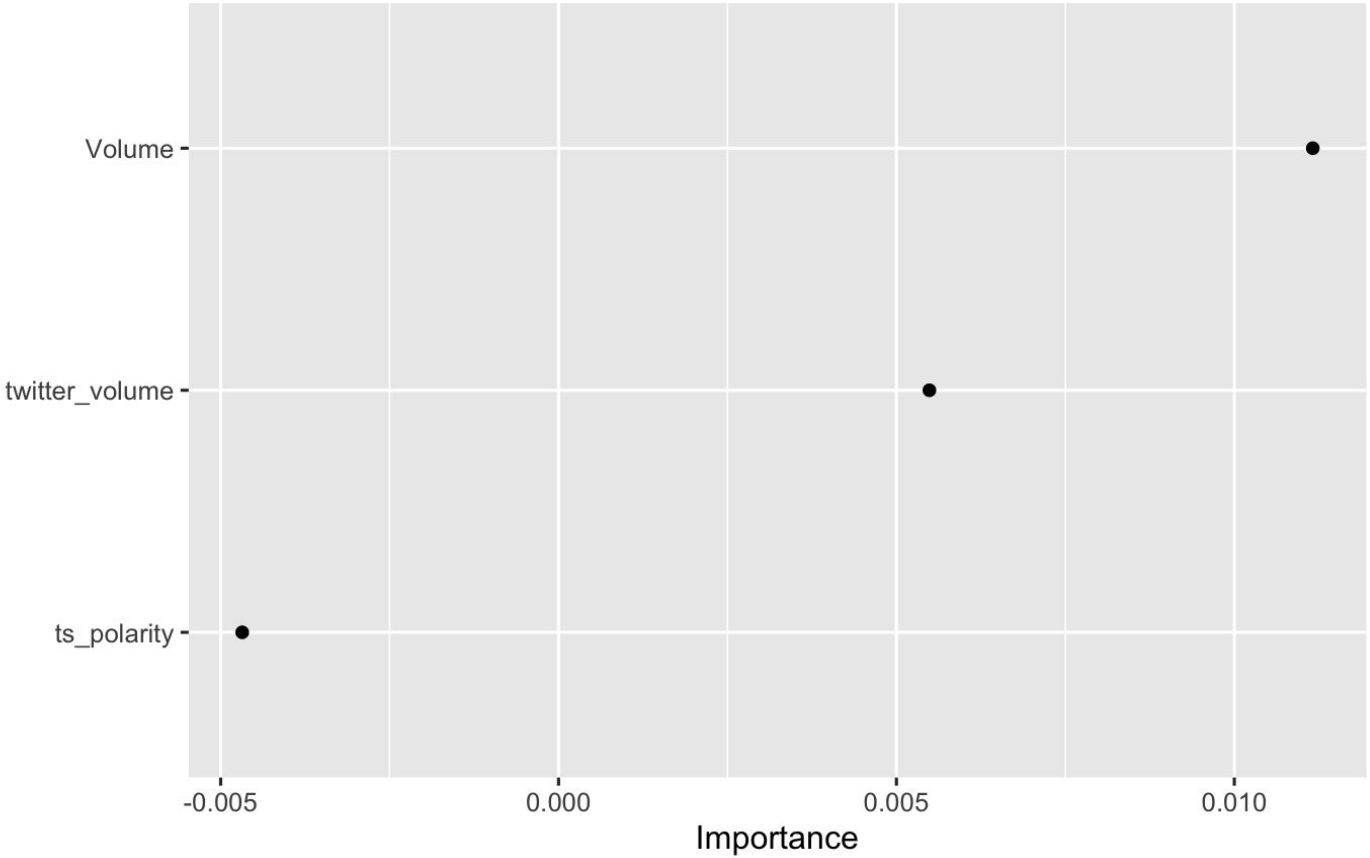
- AAPL stock data from 2016 to 2019
  - date
  - price
  - volume
  - twitter sentiment

- Converted outcome variable (stock price) to trend (rising or falling)

CART


rise
0.53
100%
yes
Volume >= 161e+6
no
fall
0.41
21%
rise
0.56
79%
twitter_volume < 1395
twitter_volume < 839
rise
0.56
9%
fall
0.50
31%
rise
0.61
48%
ts_polarity < 0.089
ts_polarity >= 0.14
twitter_volume >= 928
rise
0.55
22%
rise
0.57
37%
twitter_volume >= 820
Volume < 81e+6
rise
0.58
36%
twitter_volume >= 1852
fall
0.28
12%
fall
0.44
7%
rise
0.84
3%
fall
0.37
9%
fall
0.31
2%
rise
0.57
20%
fall
0.20
1%
fall
0.32
3%
rise
0.60
33%
rise
0.76
10%

RANDOM FOREST



SVM RESULTS

Model	Accuracy
Linear SVM	54%
Polynomial SVM	55%
RBF SVM	56%

Other models

CART	56.2%
Random Forest	57%

Final Model Test Accuracy = 44.6%

Limitations and Ethical Considerations

- Twitter Sentiment is limited to predict the stock trend of different companies
  - Twitter API is limited
  - Twitter sentiment can be used for different things
  - Consider using different text sources for stock forecasting

Thank you.