chasing the Bitcoin

BACHELOR´S PROJECT PRESENTATION

AUTHOR: **VLAD HUTUPASU**
STUDENT NO: **2572229**
COORDINATOR: **PROF. JACOPO URBANI**
FACULTY: **COMPUTER SCIENCE**

VU UNIVERSITY AMSTERDAM

agenda

- research
- introduction to Bitcoin
- approach
- results
- future work

research question

What is the accuracy with which one can predict the direction of Bitcoin price movement?

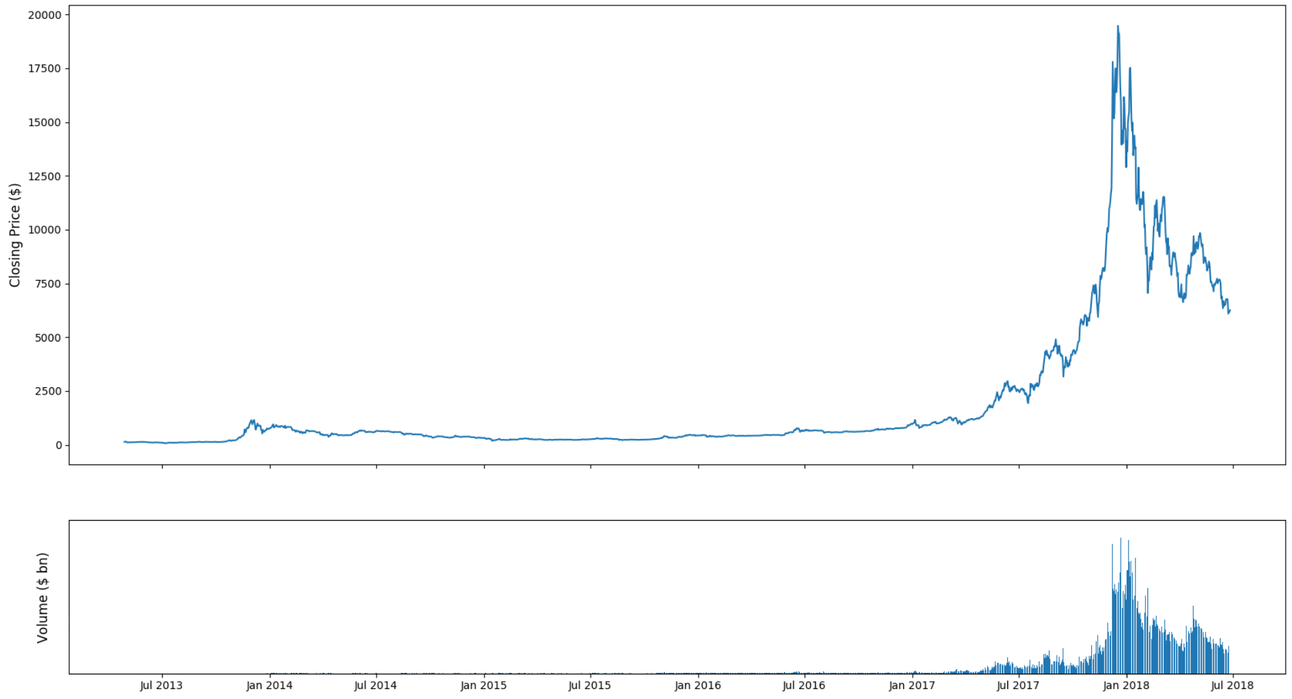
motivation

Nowadays, we live in a fast-paced world where technology is developing each and every second of the day. I do not know what the world will look like in 10, 20 or 30 years. Bitcoin gives me a sense of empowerment: with a currency that not even Governments can control, I can speculate and I can be rich from it!

1. Introduction to Bitcoin

Bitcoin is the world's first cryptocurrency, a form of electronic cash. It is the first decentralized digital currency: the system works without a central bank or single administrator.
(Investopedia, 2018)

2. Evolution



currently trading at:

USD | EUR | GBP | JPY | RUR

COINTELEGRAPH
The Future of Money

$ 5 899.90

DAY	- 3.6%
WEEK	- 6.39%
MONTH	- 17.17%

Open 24H	**$ 6 119.65**
High 24H	**$ 6 364.61**
Low 24H	**$ 5 801.28**
Last Price	**$ 5 899.90**

Total	**Ƀ 17 100 675.00**
Mkt. Cap	**$ 101B**
Vol 24H	**Ƀ 65 912.23**
Vol.24H	**$ 393 643 655.54**

3. Approach

3.1 What do people say?

- Bitcoin discussions on reddit (/r/Bitcoin/) & (/r/BitcoinMarkets/)
- Wait a second... I could extract all these comments and create a sentiment analysis

Will that work?

The Sentiment Intensity Analyzer

How does the SIA work?

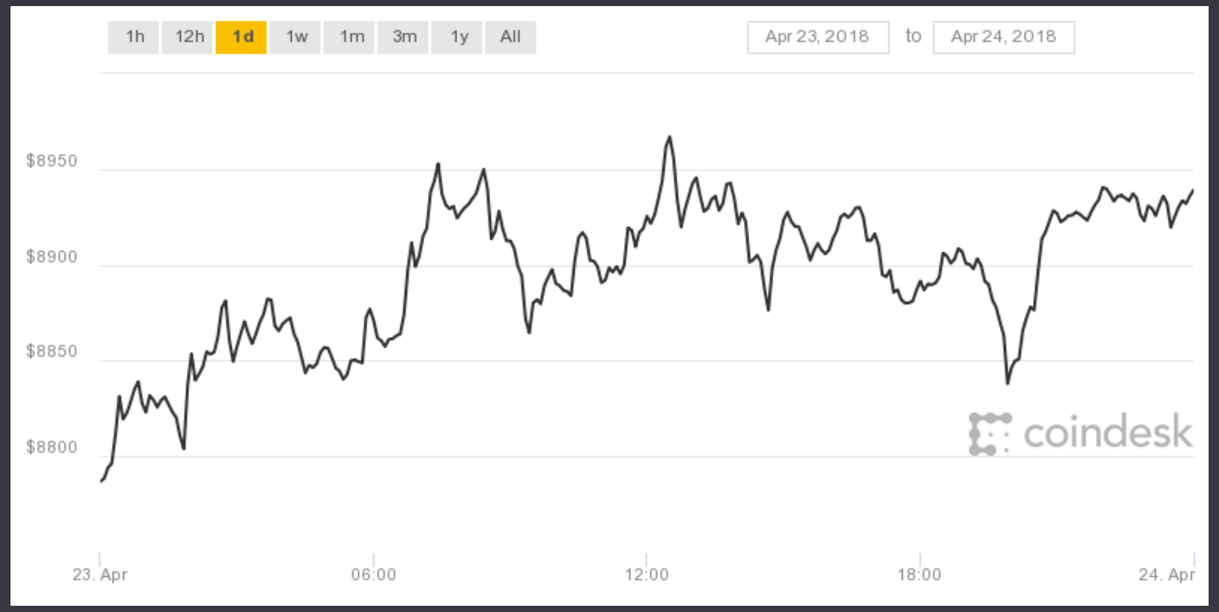
- Take each sentence/comment
- Tokenize it
- Evaluate each word
- Result represented as compound:

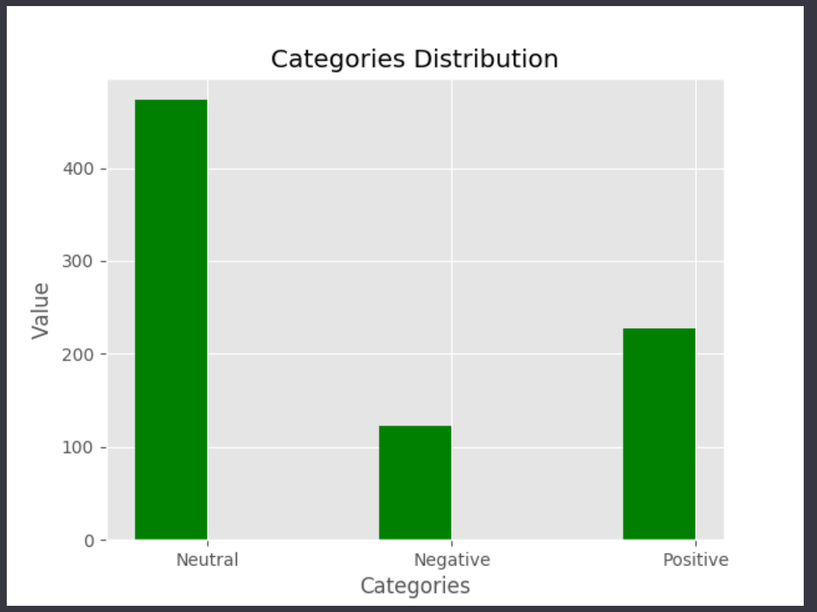
>0.2 positive
<0.2 negative
else neutral

Extract: Scrape them all!

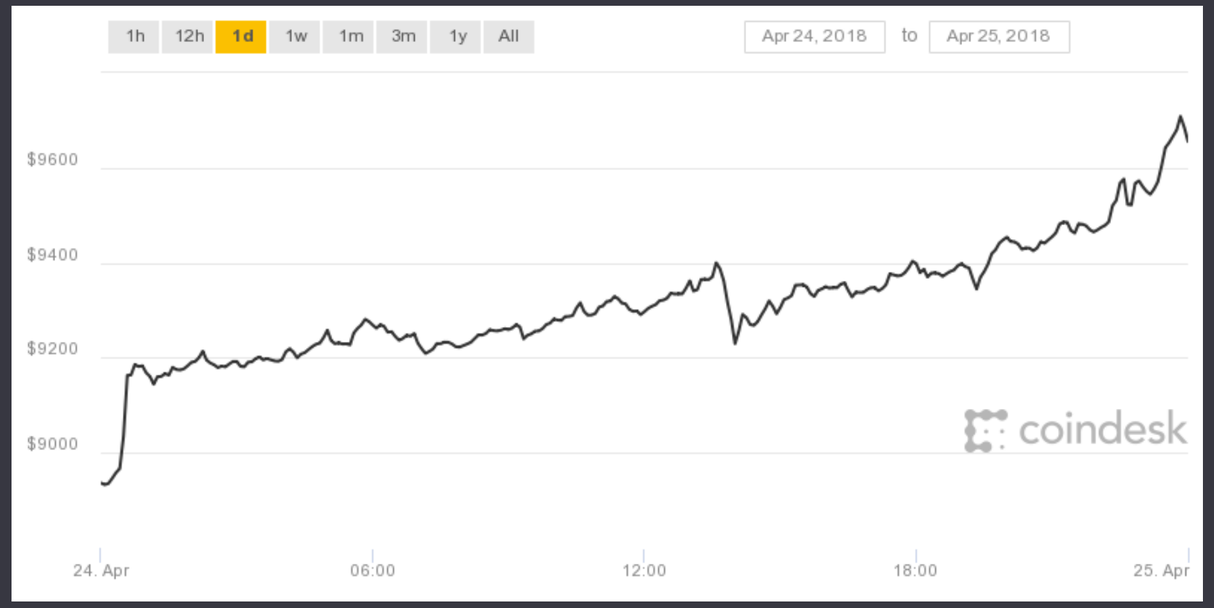
- Reddit API Wrapper - PRAW
- Get all topics and comments for the specific day and fetch them to the SIA algorithm

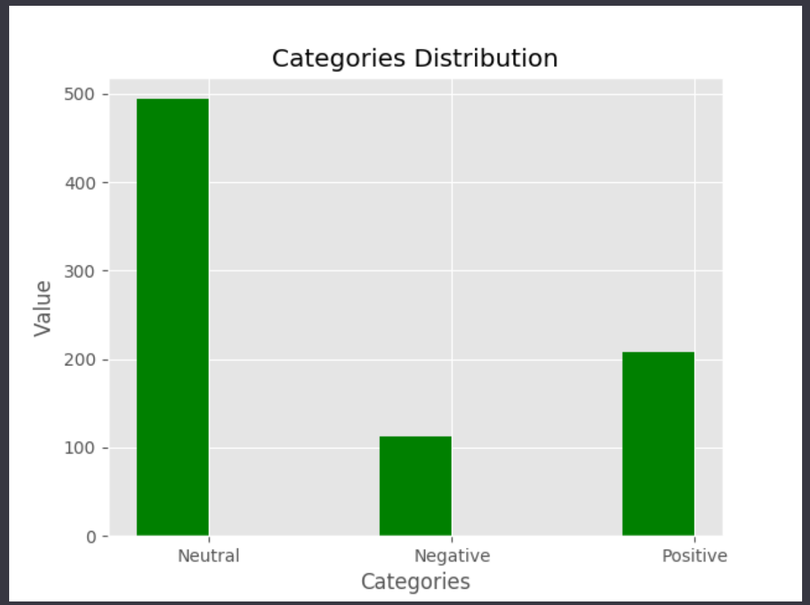
First day 23-24 April





Second day 24-25 April





Wow! It works...

How about testing it a few more times before getting rich?

After testing it for one week, I realized that there will always be more positive comments than negative.. unless something highly negative occurs unexpectedly.

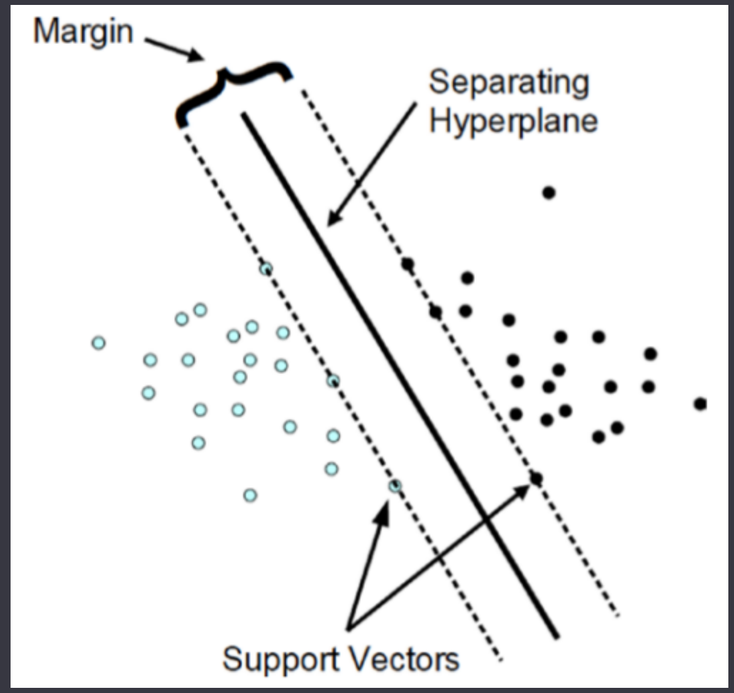
Should I listen to everyone?

- Restrict the chosen topics only to Daily Discussions
- This way avoiding redundant/unhelpful comments

3.2 Support Vector Machine

- Machine learning algorithm which can be used for either classification or regression challenges
- Builds a model in a n-dimensional space in which different classes are separated by a gap (hyperplane)
- The "n" is given by the number of features

Prerequisites: train & test data, features.



Gathering the features

- Extract historical data about bitcoin (https://coinmarketcap.com):

Date	Open*	High	Low	Close**	Volume	Market Cap
2018-06-26	6253.55	6290.16	6093.67	6093.67	3279760000	107033000000
2018-06-25	6171.97	6327.37	6119.68	6249.18	5500810000	105625000000
2018-06-24	6164.28	6223.78	5826.41	6173.23	4566910000	105480000000

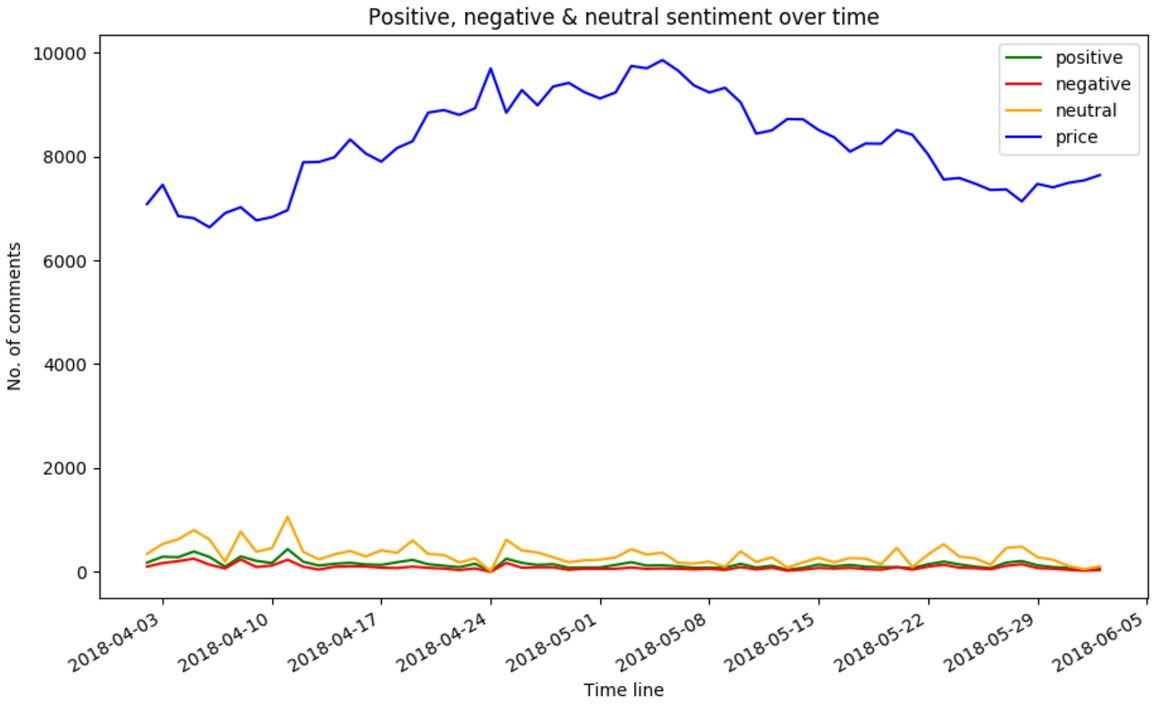
- Find relations between the extracted values and set price movement:

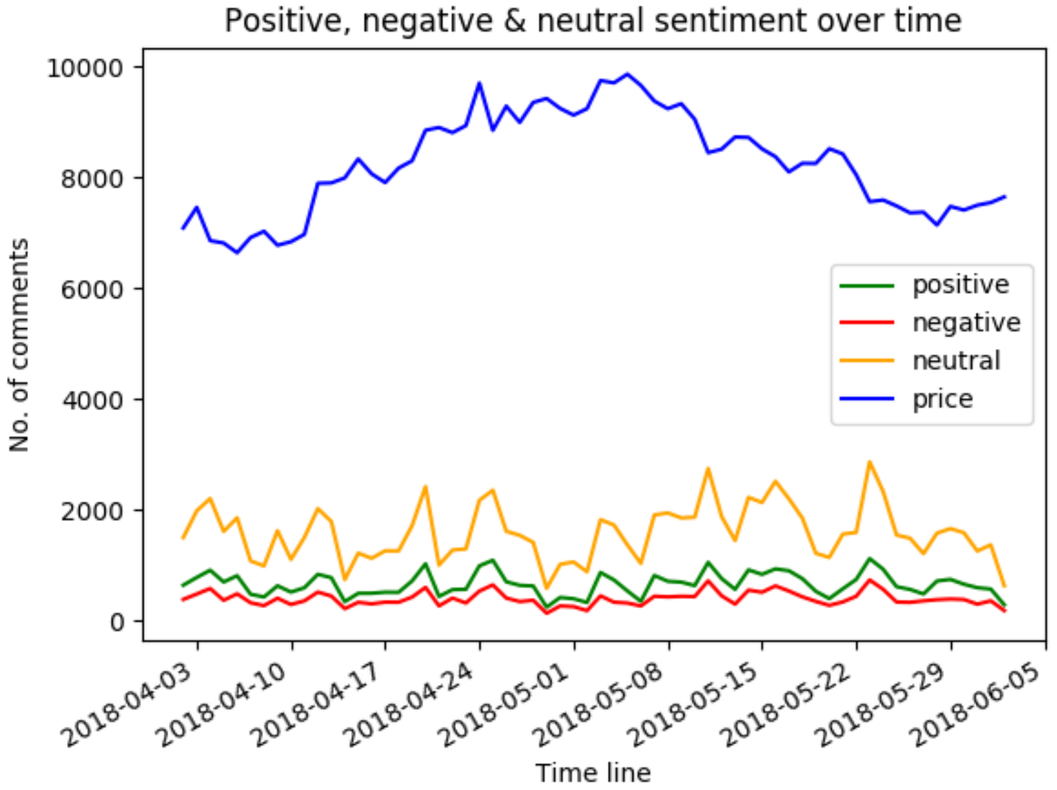
Date	Close**	Volume	Close Off High	Day Diff	Volatility	Movement
2018-05-06	9654.80	7222280000	2017.098697	193.503303	0.048235	Down
2018-05-07	9373.01	7394020000	3484.988027	282.676061	0.045027	Down
2018-05-08	9234.82	7415870000	3608.573646	155.689184	0.035709	Up

- Add SIA results:

Date	Close**	Volume	Close Off High	Day Diff	Volatility	SIA	SIA2	Movement
2018-04-02 00:00:00	7083.80	4333440000	-6759.384114	-349.079455	0.046588	102.024613	123.003195	Up
2018-04-03 00:00:00	7456.11	5499700000	-6735.521867	-498.221693	0.064550	92.340951	122.367101	Down
2018-04-04 00:00:00	6853.84	4936000000	8499.699700	808.123480	0.089319	89.750542	68.100358	Down

SIA in April & May





Calculations

DAY DIFF = $\frac{Open - Close}{Open} * 10000$

CLOSE OFF HIGH = $\frac{2 * (High - Close)}{(High - Low) - 1} * 10000$

VOLATILITY = $\frac{High - Low}{Open}$

SIA = $\frac{(Positive - Negative)}{Total} * 1000$

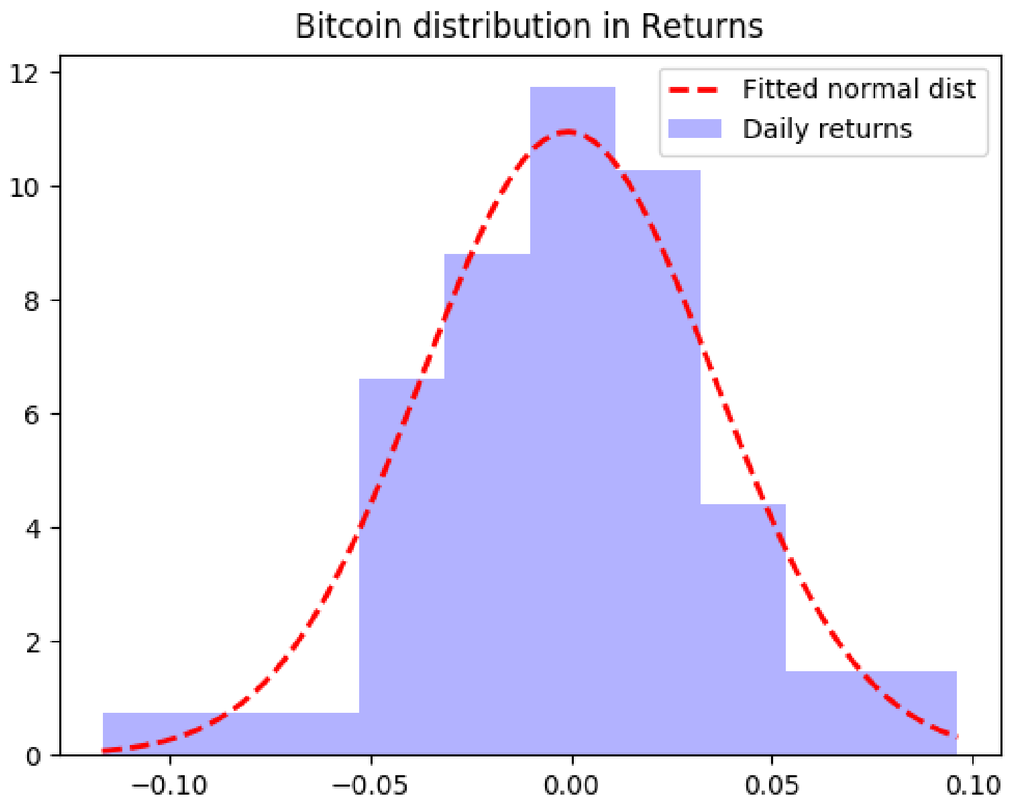
4. Results

Choosing the right features

- Values calculated upon Bitcoin historical data: Day Diff, Close Off High & Volatility
- Daily sentiment given by SIA

After conducting multiple tests, the winner combination was given by:

[Day Diff, SIA, Volatility]
60.00% Accuracy



Up or Down?

Is that enough?

Taking into consideration past events, predict the percentage movement for the following day.

Important to know when to take the risk.

5. Future work

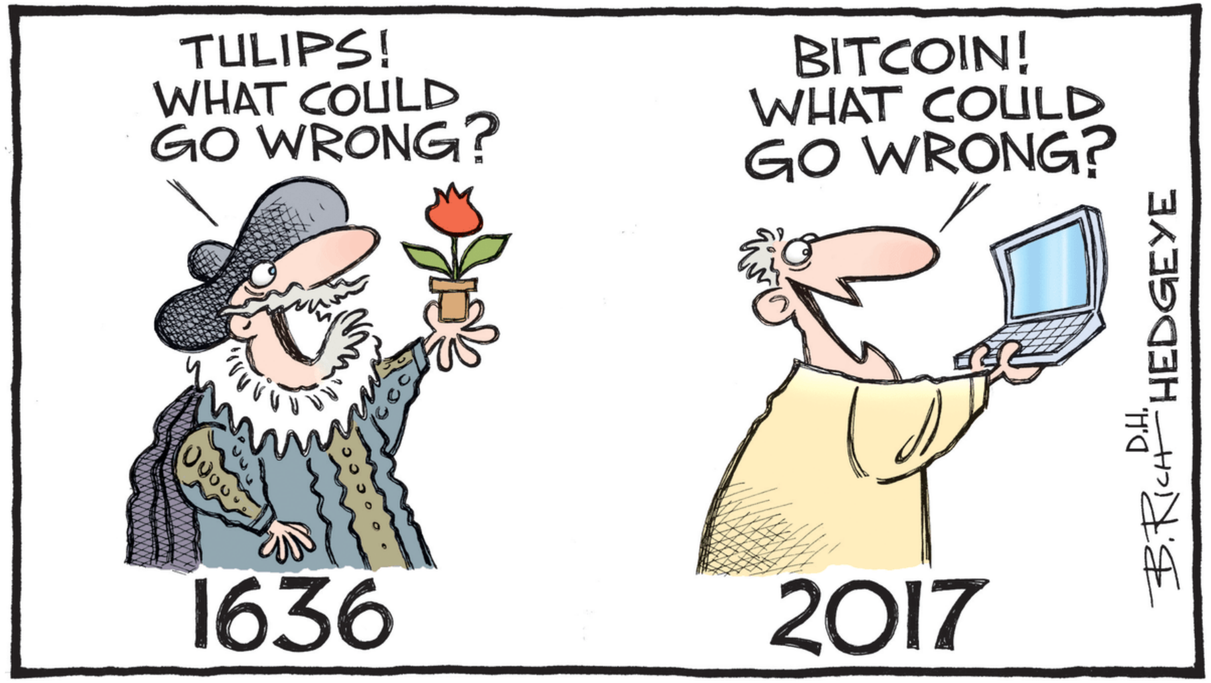
Implement an automated bot that takes as input a certain amount of money and simulates trading Bitcoin based on my program's predictions

Java vs. Python

GUI?

When to take the risk?

What accuracy should it listen to?


TULIPS!
WHAT COULD
GO WRONG?
1636
BITCOIN!
WHAT COULD
GO WRONG?
2017
B. Rich
D.H.
HEDGEYE

thank you!

recap

Bitcoin - world's first cryptocurrency

Strong evolution & volatility

Analyze what people say - 'SIA'

SVM as a machine learning algorithm

60% accuracy

Take the risk?