Predicting stock price by analyzing influential figures twitter feeds

For the stock market, tweets are fairly important, but some are more important than others!!

CNN tech

BUSINESS   CULTURE   GADGETS   FUTURE   STARTUPS   CNNMONEY

Snapchat stock loses $1.3 billion after Kylie Jenner tweet

by Kaya Yurieff   @kyurieff

February 23, 2018: 12:11 PM ET

Recommend 9.7K

Social Surge - What's Trending

China is the big wild card in Trump's Iran decision

Iran fears send US oil above $70 for first time since 2014

Apple's game-changing iMac turns 20 years old
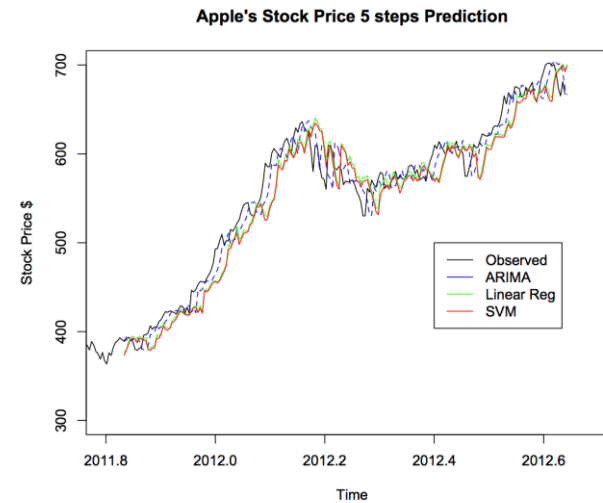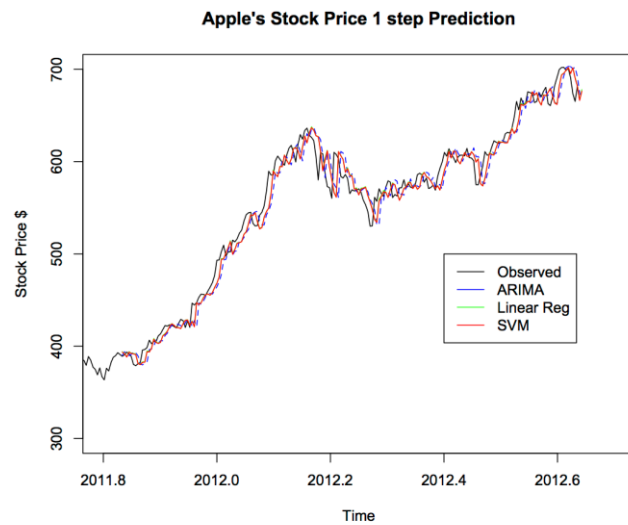
Investing
More from CNNMoney

PACIFIC: The Disney-Comcast War Gets Hot

Snap stock drops after Kylie Jenner tweet

Snapchat stock loses $1.3 billion after Kylie Jenner tweet

- There are many predictive models for the stock market – mostly insensitive to twitter feeds

Apple's Stock Price 1 step Prediction

Apple's Stock Price 5 steps Prediction

https://www.cse.unr.edu/~harryt/CS773C/Project/

How can we correct prediction models based on significant tweets?!

Proposed Algorithm

```
┌─────────────────────┐          ┌─────────────────────────┐
│   Stocks clustering  │ ········ │    Sector, Capital,     │
│                      │          │  Volatility, and others │
└─────────────────────┘          └─────────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌──────────────────────────────┐
│ Select influencers in│ ······· │ President of the US, Elon Musk, Kim │
│  each stock cluster  │          │  Kardashian, CEOs in general  │
└─────────────────────┘          └──────────────────────────────┘
          │
          ▼
┌──────────────────────────────┐
│ Text mining of the history of│
│ tweets for topic-specific words│
└──────────────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌──────────────────────────────┐
│ Correlation analysis │ ······· │  frequency of topic-specifics for │
│                      │          │    associated stock cluster   │
└─────────────────────┘          └──────────────────────────────┘
```
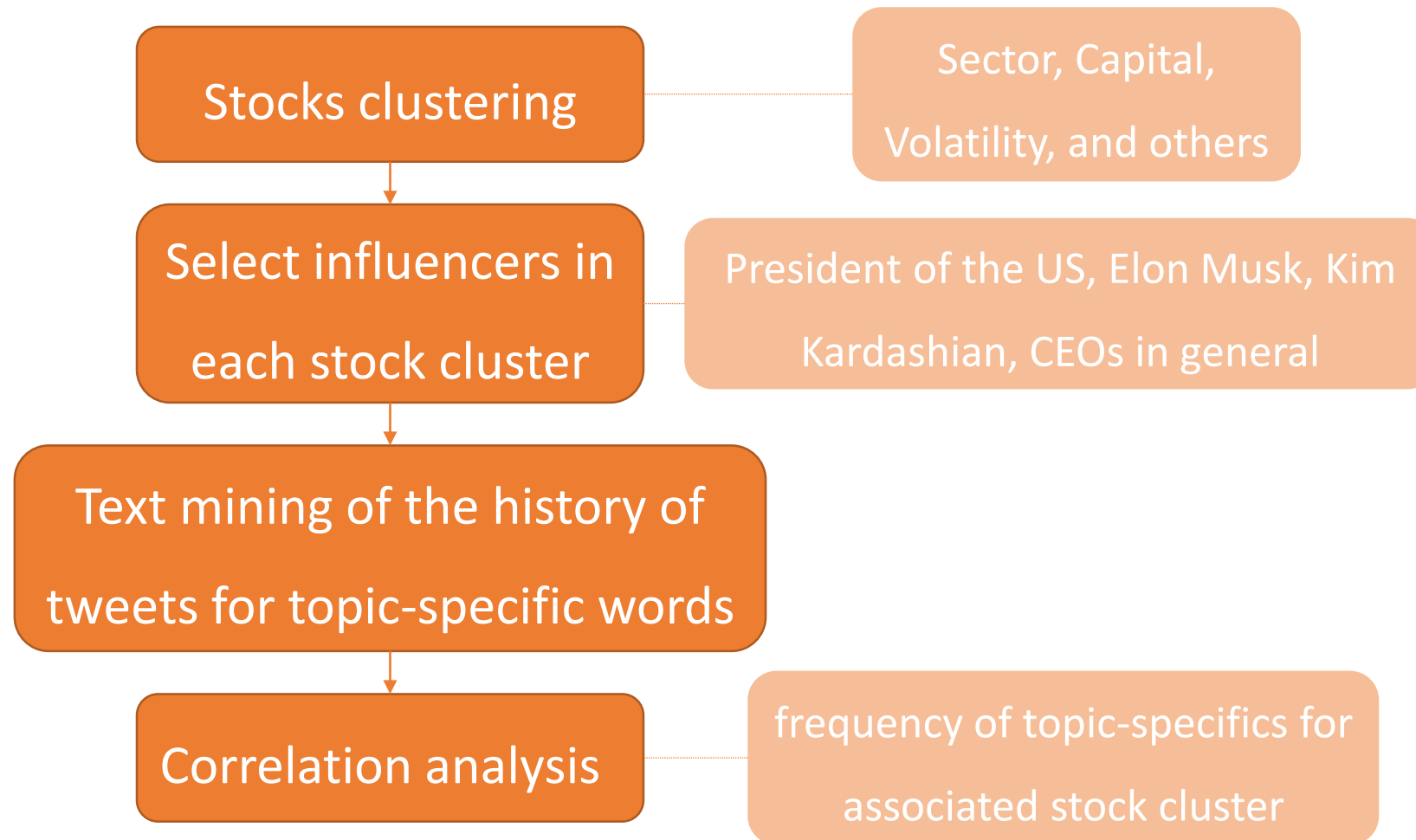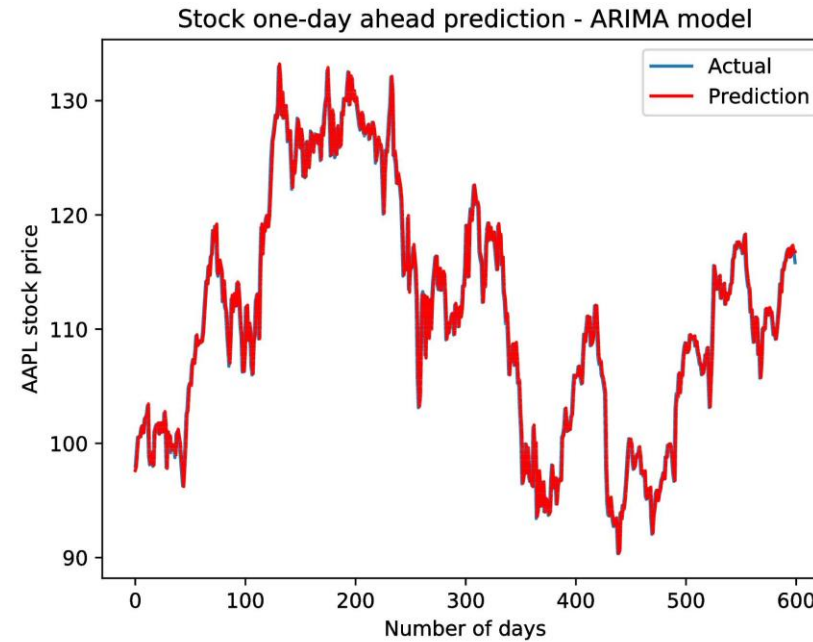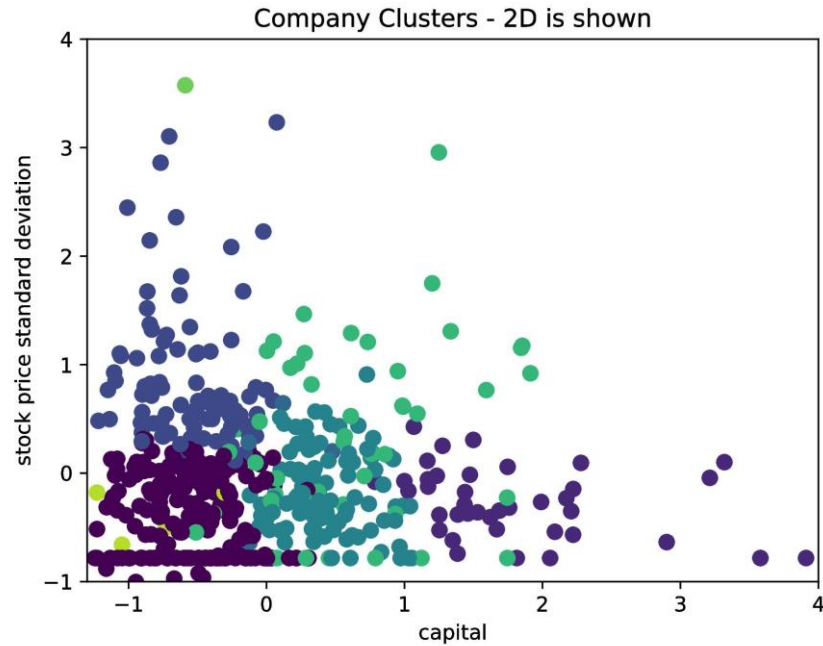
Data Source

- 100 MB data for S&P 500 companies historical prices with fundamental data

- For twitter feeds:

1. Tweepy

2. Twitter Capture and Analysis Toolset – Github

3. TAGS

4. Specific character's database:

- http://www.trumptwitterarchive.com/

Analysis

Company Clusters - 2D is shown

Stock one-day ahead prediction - ARIMA model

Next Step:
- Regression between binary word vector and daily price change for different clusters
- Predictive model updating – ARIMA , Markov Chain, DNN