CS677 Final Project (NLP)

Sentiment Analysis using VADER on Tweets

Tzupin Kuo

Dec 10, 2019

Sentiment Analysis

- **What**

The process of 'computationally' determining whether a piece of sentence is positive, negative, or neutral (i.e., Polarity). Also known as opinion mining.

- **Why**

A very useful tool for identifying the public sentiments on fields such as Business, Politics, Public Relations, social phenomena, etc.

And ,when it comes to Financial field, it could be used to understand the public sentiments for things such as Company's merging/acquisition, quarterly/annually positive results published, leadership changed for good, etc.

Sentiment Analysis

- **Who**

Researcher, marketer, Analyst, Politician, etc.

- **How**

1. Polarity Detection (positive/negative/neutral)

2. Emotion Detection (happy/sad/angry etc)

3. Intent Detection (e.g. underlying complaint in a product review)

Sentiment Analysis

- **Which**

Rule-based (or lexical) :
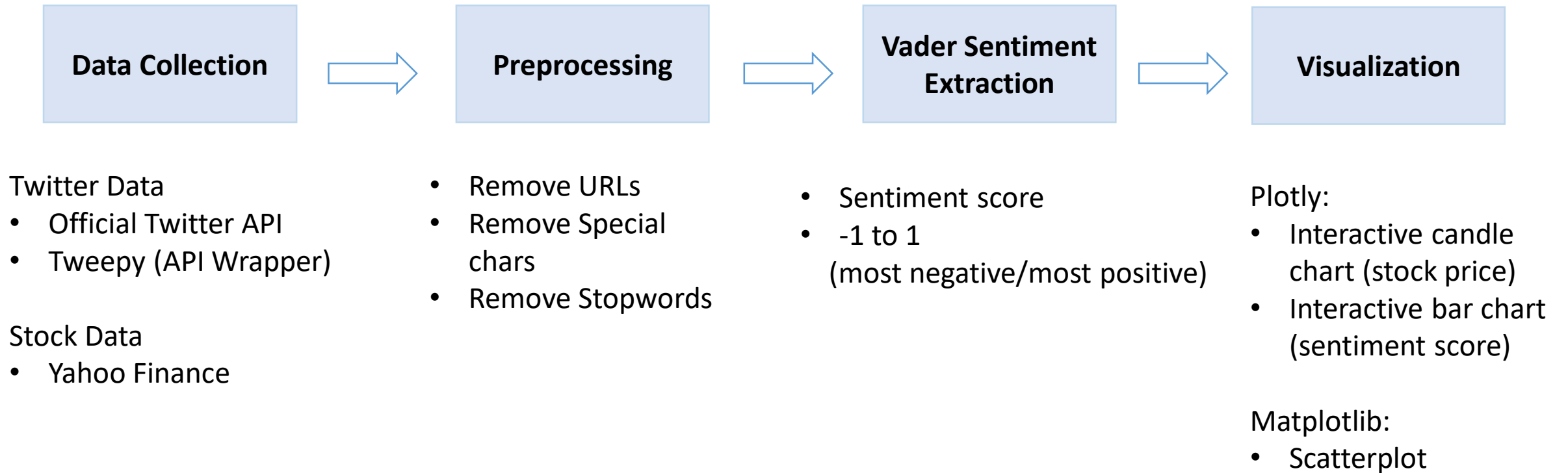
- VADER

- TextBlob

Machine Learning based :

- Standford CoreNLP (RNN)

- LSTM RNN

Extracting tweets' sentiments using VADER

Tweet	Sentiment
Excellent article by @ johngapper # TiffanyAndCo # LVMH # luxurylifestyle # Diamonds	Positive
I am so disappointed. # TiffanyAndCo # FightForFreedomStandWithHongKong	Negative
just like me! But I'm proud of it and I know it will taste amazing! # alltheworldsamage # bakingftw # tiffany # tiffanyandco # tiffanyblue	Positive
Now this is just wrong!!! We need Tiffanyâ€™s at a time like this!!!! # TiffanyAndCo # HurricaneDorainpic.twitter.com	Negative
I have the very best mom and sister in the world They have spoiled me so much for my birthday!! # birthdaygirl # birthdaypresent # tiffanyandco # imsospoiled # mymomisbetterthanyours	Positive

Procedures

Data Collection	→	Preprocessing	→	Vader Sentiment Extraction	→	Visualization

Twitter Data
- Official Twitter API
- Tweepy (API Wrapper)

Stock Data
- Yahoo Finance

- Remove URLs
- Remove Special chars
- Remove Stopwords

- Sentiment score
- -1 to 1 (most negative/most positive)

Plotly:
- Interactive candle chart (stock price)
- Interactive bar chart (sentiment score)

Matplotlib:
- Scatterplot
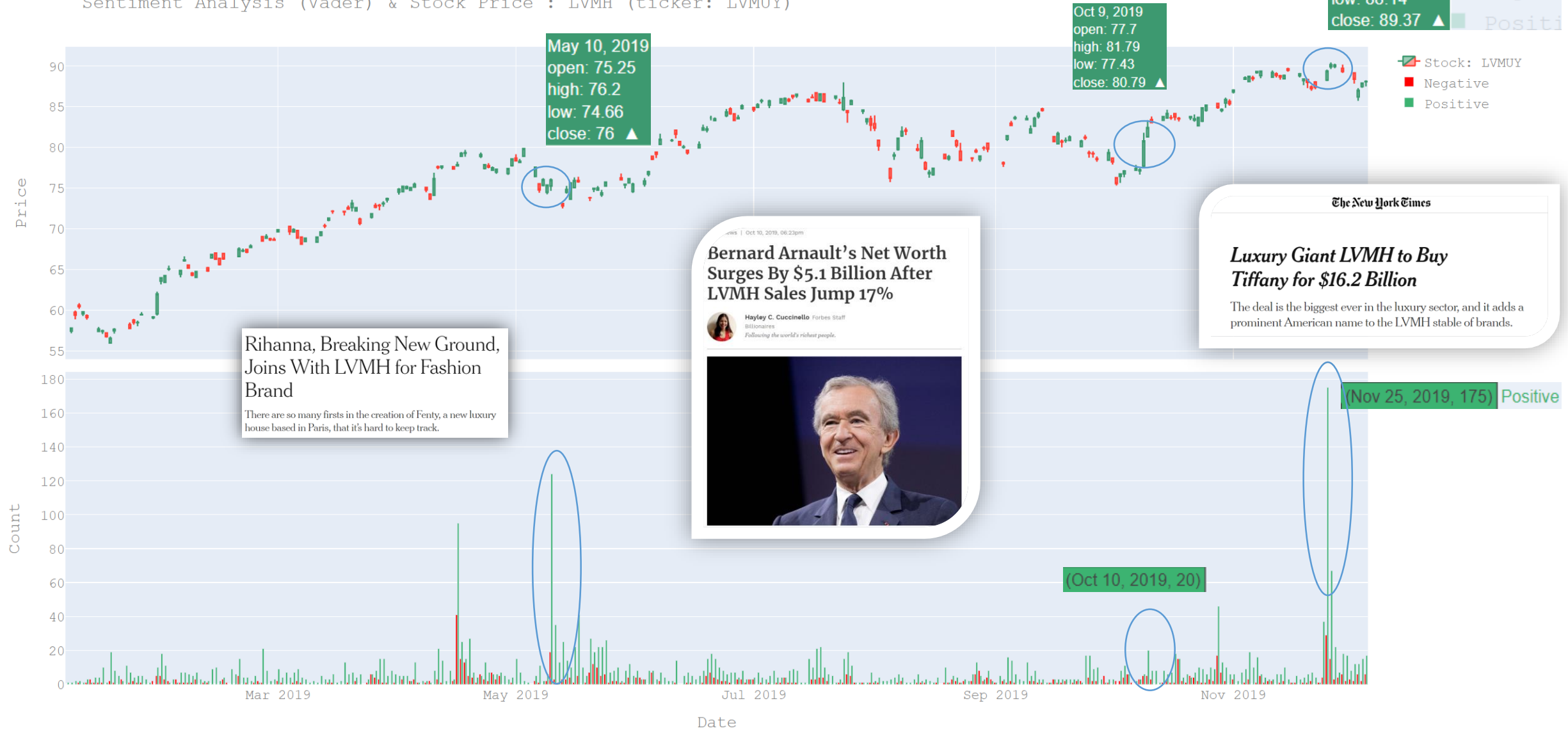
Libraries used : Pandas, Numpy, VaderSentiment, Matplotlib, Plotly

LVMH Moët Hennessy (LVMUY)

Sentiment Analysis (Vader) & Stock Price : LVMH (ticker: LVMUY)

Nov 25, 2019
open: 88.26
high: 89.55
low: 88.14
close: 89.37 ▲

Stock: LVMUY

May 10, 2019
open: 75.25
high: 76.2
low: 74.66
close: 76 ▲

Oct 9, 2019
open: 77.7
high: 81.79
low: 77.43
close: 80.79 ▲

Stock: LVMUY
Negative
Positive

Rihanna, Breaking New Ground, Joins With LVMH for Fashion Brand

There are so many firsts in the creation of Fenty, a new luxury house based in Paris, that it's hard to keep track.

News | Oct 10, 2019, 06:23pm

Bernard Arnault's Net Worth Surges By $5.1 Billion After LVMH Sales Jump 17%

Hayley C. Cuccinello Forbes Staff
Billionaires
Following the world's richest people.

The New York Times

Luxury Giant LVMH to Buy Tiffany for $16.2 Billion

The deal is the biggest ever in the luxury sector, and it adds a prominent American name to the LVMH stable of brands.

(Nov 25, 2019, 175) Positive

(Oct 10, 2019, 20)

Tiffany & Co. (TIF)

Sentiment Analysis (Vader) & Stock Price : Tiffany & Co.
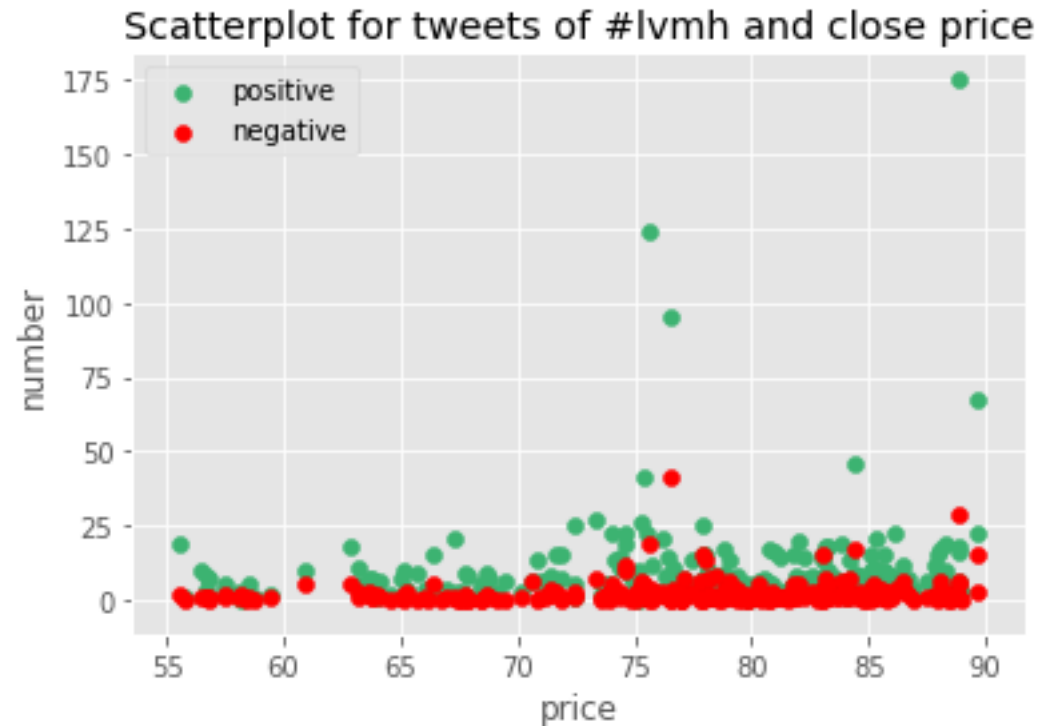
CNN International

Louis Vuitton owner offers to buy jewelry icon Tiffany & Co

Hong Kong (CNN Business) French luxury goods giant LVMH (LVMHF) has offered to buy Tiffany & Co., the iconic New York jeweler known for its little blue ...

1 month ago

Oct 28, 2019
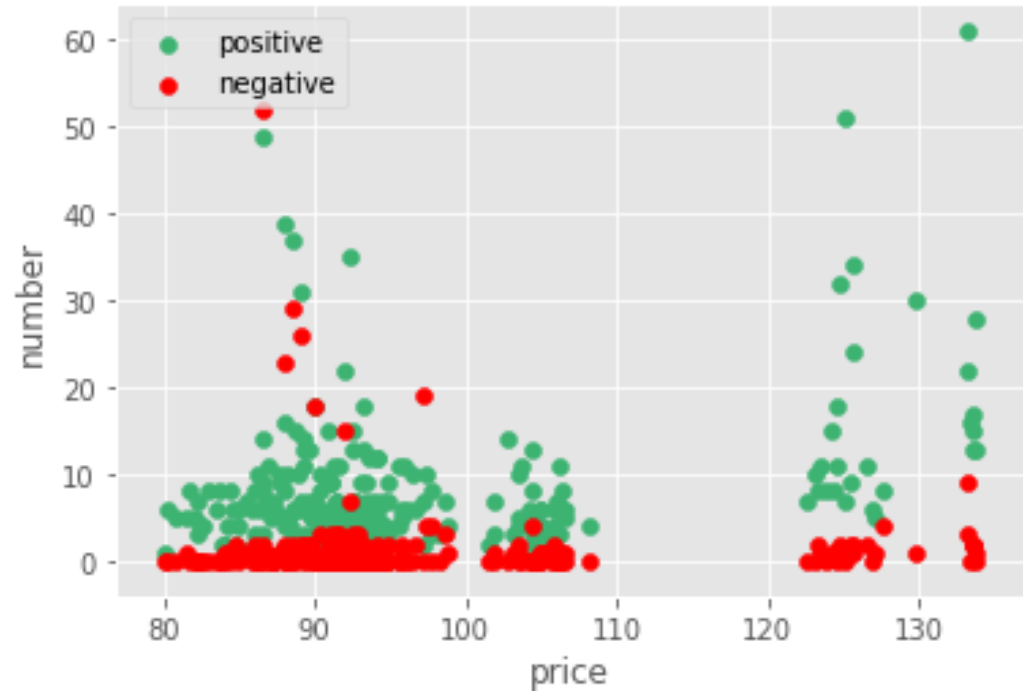open: 127.65
high: 130.4
low: 125.74
close: 129.72 ▲

RETAIL

Tiffany removes ad over Hong Kong controversy

Oct 8, 2019
open: 87.78
high: 88.06
low: 86.44
close: 86.63 ▼

Stock: TIF
Negative
Positive

RED CARPET

Oscars 2019: Lady Gaga's 128-carat Tiffany & Co. diamond was last worn by Audrey Hepburn

PAR EMELINE BLANC

(Oct 8, 2019, 52) Negative

(Oct 28, 2019, 30) Positive

Correlation Analysis - LVMH Moët Hennessy (LVMUY)

Scatterplot for tweets of #lvmh and close price

	v_pos	v_neg	price
v_pos	1.000000	0.773049	0.140242
v_neg	0.773049	1.000000	0.138225
price	**0.140242**	**0.138225**	1.000000

Correlation Analysis - Tiffany & Co. (TIF)

Scatterplot for tweets of #TiffanyAndCo and close price

	v_pos	v_neg	price
v_pos	1.000000	0.523049	0.353015
v_neg	0.523049	1.000000	-0.040759
price	**0.353015**	**-0.040759**	1.000000

Conclusion

- In this analysis, I have shown that a little bit positive correlation exists between rise/fall in stock prices of a company to the public opinions (especially positive tweets) expressed on twitter.

- But other than those 'big news days', it would be hard to tell the effectiveness of public sentiments (through tweets) on stock price fluctuation on those regular days.

- We can set a threshold and only use the "spikes" on positive/negative tweets for stock price prediction.

Future Work

- Use different combination of keywords and assign weights to them for reducing biases.

- Gather data from different sources such as other social medium (Instagram, Facebook, etc), and news headline on forums/new webs. Assign weights to them to reduce the biases.

- Use advanced Machine Learning based approach instead to extract sentiment score (LSTM & BERT) for a better accuracy.

THANK YOU! ANY QUESTION?