Relationship Between Trump's tweets and Stock Returns

PROJECT 5

Group members:
Elise Nguyen
Natalie Williams
Yue Liang
Luyao Sun

TWITTER

“ China has been taking advantage of the united states on trade for many years...there will be great and fast economic retaliation against china if our farmers ranchers and or industrial workers are targeted. ”

-- Donald Trump

September 18th, 2018

CONTENT

1

Natural Language Processing(NLP) Algorithms

- LSTM
- Textblob

2

Indices/ ETFs Examined

- S&P 500
- FXI
- SSE

3

Evaluation of NLP Algorithms

Conduct Manually Evaluation of the two NLP Algorithms

4

Regression Results

Run the Linear Regression on the Financial Market and the Sentiment Analysis Result.

Natural Language Processing(NLP) Algorithms

Long Short Term Memory (LSTM) Model

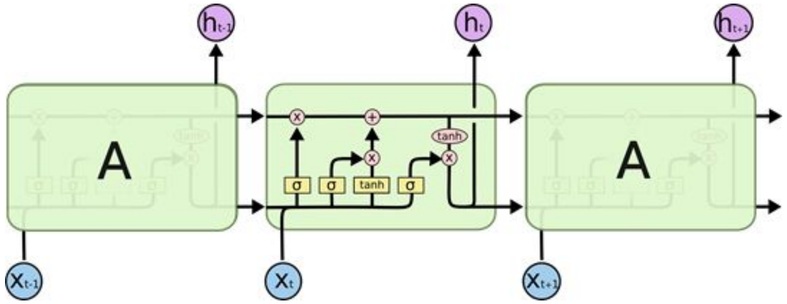
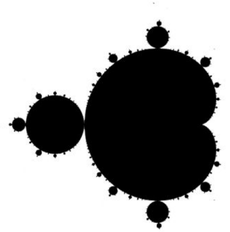


Image Source: Christopher Olah, "Understanding LSTM Networks"

- Recurrent neural network
- Differs from other neural networks: can process sequences of data as opposed to individual data points
- Information learned from previous data is used as context for new data
- Unique ability to de-emphasize or emphasize knowledge gained from past inputs

Textblob Model



For NLP Preprocessing

TextBlob

- Rate the sentiment of tweets as a comparison to the LSTM Model
  - Returns probability of tweet being negative or positive
  - -1 -> negative tweet, +1-> positive tweet
- Includes subjectivity as a result
  - Subjectivity: how much a statement is rooted in personal opinion or emotion rather than facts
  - Represents subjectivity as a float within [0,1]

Indices & ETFs

S&P 500: index that measures the stock performance of 500 large companies listed on stock exchanges in the United States

- Average annual return: 10% (6% after inflation)
- Traded 9.30 a.m. to 4 p.m. EST

SSE: is a stock exchange based in the city of Shanghai, China. It is operating independently in mainland China.

- Traded at the Shanghai Stock Exchange
- Traded 9.30 a.m. to 4 p.m. GMT+8

FXI: track the investment results of the FTSE China 50 Index composed of large-capitalization Chinese equities.

- International Investors Allowed
- Traded on the Hong Kong Stock Exchange
- Traded 9.30 a.m. to 4 p.m. EST

Evaluation of NLP Algorithms

- Two team members read through and manually categorized 100 tweets into positive (1), negative (-1), and neutral (0)
- Compared to algorithms outcome:
  - Average % accuracy for LSTM: 46%
  - Average % accuracy for TextBlob: 59%

Linear Regression Results

- Regressed Close-Open and Open-Close returns of S&P, SSE, FXI against LSTM and Textblob sentiment outcome
- Calculated Point Biserial Correlation: equivalent to the Pearson correlation, used when one variable is dichotomous

Most significant result: SSE vs. Textblob: Close-Open

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0006	0.001	0.702	0.484	-0.001	0.002
C(Positive)[T.1]	-0.0023	0.001	-2.159	0.033	-0.004	-0.000

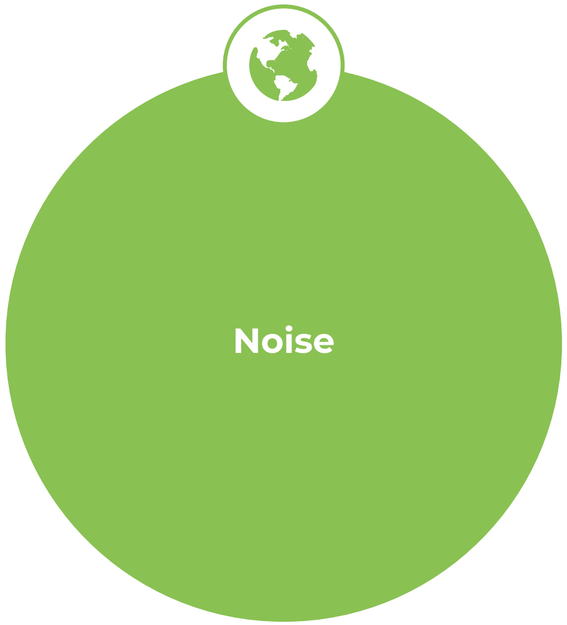
Compared to S&P vs. Textblob: Close-Open

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0002	0.000	0.486	0.627	-0.001	0.001
C(Positive)[T.1]	-0.0001	0.001	-0.249	0.803	-0.001	0.001

Reasons for Low Correlation & Significance

Trump tweets
don't have details on trade deal
➔Hard for stocks movement to reflect the sentiment

Trump tweets may be 'positive' for United States, but negative about China/ other countries
➔Hard for NLP algorithms to categorize into + or - sentiments → not great accuracy



Trump tweets about topics unrelated to China/ Trade war

Ways to Improve

- ★ Improve NLP algorithms: take into account complex sentiment (positive to the US but negative to China for example)
- ★ Expand context to include news headlines or press conference releases, i.e. details of a new tariff deal may be published and markets fluctuate before Trump tweets about the deal
- ★ Experiment with different time window for regression

“ Apple prices may increase because of the massive tariffs we may be imposing on china... make your products in the united states [with a tax incentive] instead of china... ”

-- Donald Trump

September 8th, 2018

References

- Yuan, K., Liu, G., Wu, J., & Xiong, H. (2020). Dancing with Trump in the Stock Market: A Deep Information Echoing Model. ACM Transactions on Intelligent Systems and Technology (TIST), 11(5), 1-22.
- Bhoi, A., Pujari, S. P., & Balabantaray, R. C. (2020). A deep learning-based social media text analysis framework for disaster resource management. Social Network Analysis and Mining, 10(1), 1-14.
- Quantilia (2018) The trade war effect on commodities https://www.quantilia.com/the-trade-war-effect-on-commodities-2/
- https://medium.com/@outside2SDs/an-overview-of-correlation-measures-between-categorical-and-continuous-variables-4c7f85610365
- https://en.wikipedia.org/wiki/Point-biserial_correlation_coefficient
- https://topforeignstocks.com/foreign-adrs-list/the-full-list-of-chinese-adrs/

Project 5 Presentation

Thank You

Group members:
Elise Nguyen
Natalie Williams
Yue Liang
Luyao Sun