Sentiment analysis of tweets and their effects on the Stock Market

Hrishikesh Mahajan*
Department of Computer Science
Binghamton University - State University of New York
hmahaja1@binghamton.edu

Jainil Parikh†
Department of Computer Science
Binghamton University - State University of New York
jparikh1@binghamton.edu

ABSTRACT

In this paper, we apply sentiment analysis and machine learning principles to find the correlation between 'public sentiment' and 'Stock Market price'. In order to explore this relationship, we conducted a study of over 1,81,200 tweets filtered from over 54 million tweets collected over 3 weeks from the Twitter's one-percent stream and over 37,000 rows of Stock Market price for the top 30 performers in the stock market. The findings suggest that there is evidence of correlation between the general mood of the public and investment behavior in the short term; however, the relationship is not yet determined as statistically significant. There is also evidence of causation between public sentiment and the stock market movements, in terms of the relationship between MOOD and the daily closing price. Overall, these results show promise for using sentiment analysis on Twitter data for forecasting market movements. We also performed a time-series prediction for predicting the stock market prices and found the result to be promising by obtaining about 70 percent accuracy.

ACM Reference Format:
Hrishikesh Mahajan and Jainil Parikh. 2020. Sentiment analysis of tweets and their effects on the Stock Market. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Twitter has a large audience potential. Currently it attracts an estimated average of 271 million users every month. It is therefore not surprising that the rich and continuous mass of data made available by these platforms is being harnessed with the purpose of studying individual and group behavior as well as global patterns especially in regards to sentiment towards brands, products, events, recent news and social and political issues. Twitter effect has been shown to be particularly relevant to experiential media products (e.g., movies, music, and electronic games); these are generally the products for which 'instant' success is essential.

Twitter-based models can then be built to aggregate the opinions of the collective population. They can be used to predict future trends while gaining useful insights into individual behavior. Thanks to the availability of an application programming interface (API), which stores tweets that may be accessed by researchers, and its convenient features such as filtering by variables like hashtags and keywords, Twitter has encouraged researchers to take an interest and explore its potential beyond that of a social network. Since the conception of Twitter in 2006 studies into properties of Twitter have grown in popularity and can be classified in one of the following streams; structural, content or sentiment. When evaluating the structural properties of Twitter as a social network studies have focused on user influentially. Content analysis studies have focused on analyzing the content, virality and motivations of tweets. Sentiment analysis studies have focused on using Twitter chatter sentiment for predicting behavior arguing that although each tweet represents individual opinion, an aggregate sample should provide an accurate representation of public mood.

Stock markets is a highly volatile and dynamic platform where events happening in real time have a significant capability to alter the course of the market. Since we have discussed how Twitter has the potential to sway public opinion, we were interested to find out if tweets, its popularity among the community and the sentiment of the tweets could affect the stocks of trending companies.

Motivated by the prospect of a forecasting relationship between tweets and the market performance, we planned to investigate the impact of user-generated content defined as collection of tweets based on recording user sentiments towards the trending performers in the stock market specifically focusing on the top 30 stock symbols. The research questions we are aiming to answer are:

- Does co-relation exist between public sentiment and stock market movement?
- Does the relationship of causation exist between public sentiment regarding stock market movement?
- Can we predict the stock market movement based on a real-time data? If so, how accurate can those predictions be?

In order to explore the potential relationship between the polarised sentiments, we identify key news events such as the inclusion of Tesla in the S & P 500 index as well as the news of The Federal Aviation Administration, which lifted its ban on Boeing's 737 Max on November 18th allowing the plane to return to the skies

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

after being grounded for more than 20 months following crashes in Indonesia and Ethiopia that killed 346 people. During these news, there were a considerable amount of tweets which led to a significant change in the stock market prices of both the companies. Our work was able to successfully link these events and a reveal a relationship among them.

The rest of the paper is organized as follows: Section 2 will look at existing literature in order to illustrate the motivation behind the study. Section 3 will describe the datasets used, Section 4 will illustrate the methodology followed, Section 5 will provide us with the results of our findings. Finally, Section 6 will conclude with a summary of the findings, key limitations and future research implications.

2 BACKGROUND AND RELATED WORK

Twitter is characterized by the real-time transmission of product quality information and reviews, and thus it enables feedback. The receiver of such information can potentially be a very large group, and not just an individual or a small group.[1] Furthermore, the brevity of Twitter content is a unique element that is not typical for other types of word-of-mouth (WOM) communication, but nevertheless contributes to concise evaluations that are perceived as unequivocal. Microblogging is also recognized to hold huge potential for the successful implementation of many other related organization and management practices. Microblogs are a form of social communication whereby users can express their interests and attitudes in short posts, which are instantly distributed to other users via mobile phones, instant message, and the web.[7] At its most basic, Twitter is a communication tool that allows users to post 140-character messages (tweets) to all those who have opted to follow them. However, regardless of its simple exterior, with its numerous features this straightforward platform has proved itself to be incredibly valuable to businesses.[8] In order to determine whether there is a relationship between public mood and stock market performance in the context of this study, prior literature will be used to determine the following:

- The existence of stock market related tweets on Twitter and their implications
- The assumptions about the existence of a relationship between tweets from real-time events and stock market performance
- The ability to use Twitter data to predict stock market movements

2.1 Twitter as a communication platform

Several studies have identified Twitter as a social media platform used primarily for communication and spreading information. Whilst studying user motivations Java [9] observed that users participate in communities which share similar interests, some users taking on the role of information providers and others information seekers. Nonetheless, a more recent study by Smith[13] supports the findings of Java [9] concluding that users of Twitter, compared to Facebook and YouTube, are the least likely to use the platform for self-promotion, instead most use it to engage in discussions and disseminate news. Cha[6]found that news stories consistently receive a high level of retweets over a range of topics, however not all events or news obtain the same amount of attention or level of virality.[5] One explanation for the variance in popularity is a difference in emotional arousal or the content of the post.[10] Authors found that out of all sampled categories political hashtags and therefore discussions were the most persistent on Twitter but, this does not necessarily signify that the increase in the number of tweets is due to new users joining the discussion.

2.2 Relationship between stock market news and market performance

Although it is generally accepted that stock market prices are largely driven by new information and follow a random pattern, many studies have tried to predict stock market behavior using external stimuli on the basis of behavioral economics that emphasizes the important role of emotion in decision making.[11] Niederhoffer[12] observes that world news had a 'discernible influence' on the movement of the market basing their conclusion on larger returns (S&P 500) following world events, as opposed to normal days.

3 DESCRIPTION OF THE DATASETS

In this section the data collection methods will be outlined.

3.1 Twitter Data Collection

We use a python code with access to the official Twitter API.[2] As the aim of this paper is to investigate Stock Market tweet buzz only, sampled tweets need to be filtered based on keywords in the text and hashtags fields. For every tweet that comes in, we are storing the tweet-ID and the timestamp of the tweet. We then pass the tweet to our Twitter Filter which extracts relevant tweets to our project using a list of keywords to match in the tweet data. If present, the function stores them into the Database. Else, discards the data. Figure2 shows a representation of our Twitter's filtered data

3.2 Stock Market Data collection

In order to calculate the daily stock movement, We have implemented a Web crawler to feed us updated data regularly. This crawler is scheduled to fetch us data of the top 30 Stock Market companies every 10 minutes from the popular stock market website investing.com[3] The opening and closing values will be recorded. In addition to this, Name, Last, High, Low, Change, Change Percent and Time attributes will also be collected. Figure 3 shows a sample of data collected.

4 METHODOLOGY

We start by building a Data collection pipeline, for both, the Twitter and the Stock Market. We have described the details of the implementation in the Dataset section. The following subsections describe the methodology followed over the course of the project.

4.1 Data Cleaning for Twitter

As the aim of this paper is to investigate Stock Market Twitter buzz only, sampled tweets needed to be filtered by text and hashtags. Although the hashtags were directly provided, there were at times

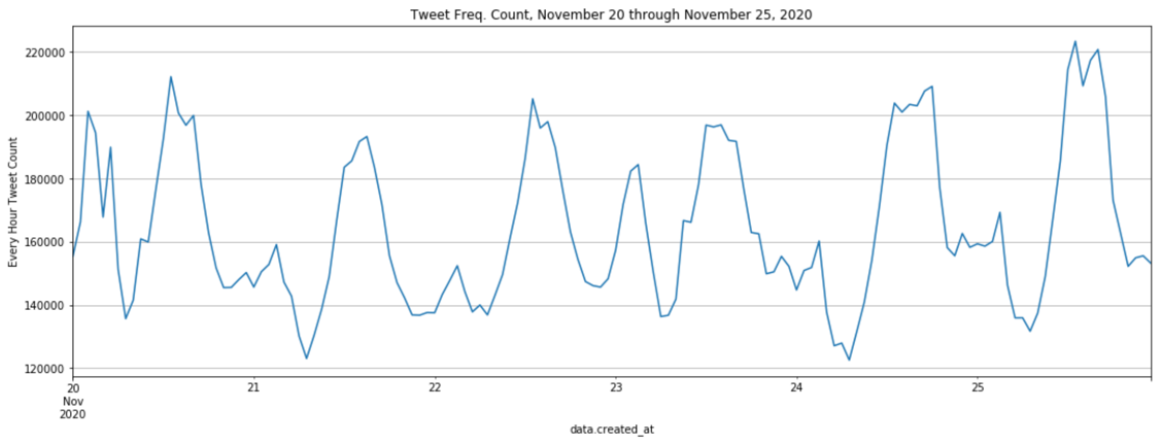

Figure 1: All Tweets collected

id	text	created_At	hashtags	Like_count	Quote_count	Reply_count	Retweet_count
5554976459018254	i just be minding my business then here comes oomf posting an ugly man on my tl	2020-11-08 21:45:08	NaN	0	0	0	0
5555005814943744	I made a onlyfans to help y'all stack y'all bread, manage y'all assets, and build your wealth the right way. Credit, Real Estate, Businesses, all that.	2020-11-08 21:45:15	NaN	0	0	0	0
5555953706655746	Ppl annoying & nosey. Like what other ppl business do for you?!	2020-11-08 21:49:01	NaN	0	0	0	0
5556058564091904	17 business days... my pussy cries itself to sleep every night... its cries echo through the walls... its gonna be okay girlie, hang in there 😔	2020-11-08 21:49:26	NaN	1	0	0	0
5556448630280192	🔔 LIMITED TIME ONLY 🔔\n\nRefer 2 friends and get 7 FREE STOCKS, like Apple, Google, Amazon, Netflix, and more...\n \nLINK: https://t.co/1FEgOGWvrq\n\n#apple #tesla ***aapl***tsla ***zm***amzn ***spce***goog #google spy #stocks #stockmarket #finance #earnings #amazon #bitcoin btc #robinhood https://t.co/Vb1T18iJlg	2020-11-08 21:50:59	apple	0	0	0	1

Figure 2: Twitter Data - Filtered

index	Name	Last	High	Low	Chg	Chg-percentage	Vol.	Time
19	JPMorgan	114.73	115.65	114.10	-0.83	-0.72	7.72M	2020-11-20 13:59:58
20	General Electric	9.69	9.83	9.60	0.03	0.31	53.64M	2020-11-20 13:59:57
25	Iridium	32.39	32.61	31.90	0.03	0.09	357.51K	2020-11-20 13:59:51
12	Crowdstrike Holdings	146.23	146.91	143.39	2.47	1.72	1.63M	2020-11-20 13:59:47
1	Apple	117.97	118.76	117.85	-0.67	-0.56	43.71M	2020-11-20 13:59:39

Figure 3: Stock Market Data

multiple or none hashtags included in the tweets. This meant the process met with several issues:
1. Hashtags variable was left blank.
2. Hashtags had multiple tags
3. Hashtags were included with camel case or as a joint word like 'stockmarket'.

Tweets with issue 1 were extracted by looking at the text from the tweets. Tweets with issue 2 were denormalized and an indivdual record pertaining to each hashtag was obtained. Tweets with issue 3 were difficult to segregate. So we added popular hashtags in our keywords dictionary in order to pick such tweets.

4.2 Sentiment Analysis of Twitter Data

In order to extract the sentiment of every tweet a lexicon based sentiment classifier, VADER was used. VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media.[4]. This is a very promising tool which gives us a compund score. The compound score is computed by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between -1 (most extreme negative) and +1 (most extreme positive). This is the most useful metric if you want a single unidimensional measure of sentiment for a given sentence. Calling it a 'normalized, weighted composite score' is accurate.

4.3 Identifying Real time events

In order to correlate the potential relationship of the tweets with the market prices, we had to look for data which had a consequential impact on both of our datasets, twitter and stock market. In order to do so, we identified stocks with the maximum change percentages in a fixed time interval. Symbols of such stock were then looked in the twitter data for a possible topics of discussion or searched on the news for any new updates. We managed to identify 2 such real-time events to support our claim of the paper. One was the inclusion of Tesla in S&P 500 index which led to a big spike in Tesla's stock price. The second one was about the Federal Aviation authority granting permission to airlines to fly the grounded Boeing 737's after a massive crash which led to death of over 364 fliers on board.

4.4 Proof using Statistical variables

For both the events we identified from our dataset for analysis, we used to variables CHANGE and MOOD in order to statistically bolster our findings by the algorithms.

- Independent Variable - MOOD the MOOD for a given day t is defined as:

$$MOOD_t = \frac{(VPOS_t - VNEG_t)}{VTOTAL_t} \tag{1}$$

- Dependent Variable - CHANGE The stock movement at a day t is defined as the normalized change in CLOSE from the previous day, which can be expressed as:

$$CHANGE_t = \frac{CLOSE_t - CLOSE_t - 1}{CLOSE_t - 1} \tag{2}$$

4.5 Price Prediction

One of our research question is to identify how much a tweet can sway people's opinion. With the data available to us from both the sources, Twitter and Stock market, we plan to identify dependency between the variables of the dataset as a whole and train a Machine Learning model to predict the price. Since we have timestamps, we have an advantage of separating training data from test data. We will identify time slots of data and mark a timestamp as divider between test and train data. We plan to have multiple sets of such data in order to train the model on multiple sets of data and find a better fit. We use tensorflow's Time-series prediction to predict the 'Last' attribute value for the symobols 'TSLA' and 'AAPL'.

5 RESULTS

This section depicts the results of the experiments proposed in the Methodology section.

5.1 Sentiment Analysis

The VADER[4] gave us a compound score as well as a split score into 3 categories namely positive, negative and neutral. The following table describes a sample of the scores we obtained.

Compound	positive	negative	neutral
-0.5106	0.000	0.171	0.829
0.8481	0.317	0.000	0.683
-0.3164	0.153	0.294	0.552
-0.4939	0.101	0.171	0.728
0.8555	0.268	0.054	0.678

The compound score is computed by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between -1 (most extreme negative) and +1 (most extreme positive). This is the most useful metric if you want a single uni-dimensional measure of sentiment for a given sentence. Calling it a 'normalized, weighted composite score' is accurate. It is also useful for researchers who would like to set standardized thresholds for classifying sentences as either positive, neutral, or negative. Typical threshold values are:

- positive sentiment: compound score >= 0.05
- neutral sentiment: (compound score > -0.05) and (compound score < 0.05)
- negative sentiment: compound score <= -0.05

The pos, neu, and neg scores are ratios for proportions of text that fall in each category (so these should all add up to be 1... or close to it with float operation). These are the most useful metrics if you want multidimensional measures of sentiment for a given sentence. Figure 5 gives us the output of Tesla's sentiment analysis after the surge in the stock market price. and the figure 5 gives us the output of the sentiment scores after the news announcement.

5.2 Statistical Proofs

We calculate the MOOD and CHANGE variables for both the events in order to verify if the findings of the VADER were accurate enough.

STOCK	VPOS	VNEG	MOOD	CHANGE
TSLA	31	6	0.676	-0.101
BA	3	4	-0.14	0.03213

Here as we can see, the positive mood score on TSLA, implies that the most of the tweets encountered were postive and hence have potray a postive influence on stock market price. The negative change value indicates the closing price on the next day higher than the previous day.
The opposite is the case for Boeing. We see a negative mood score

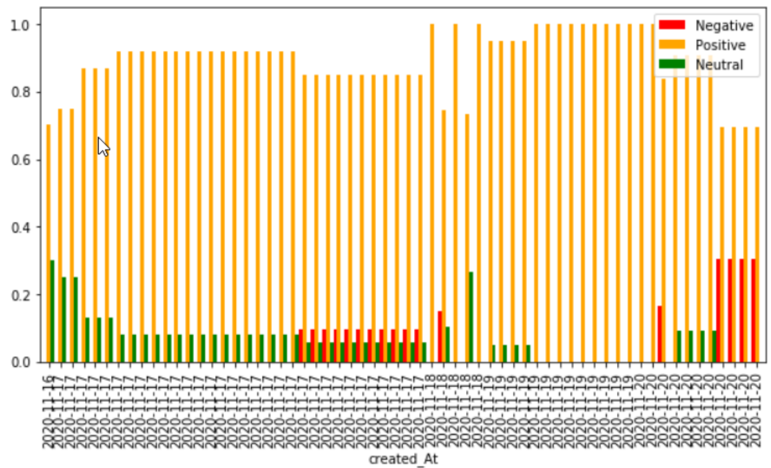

Figure 4: Tesla Sentiment Score

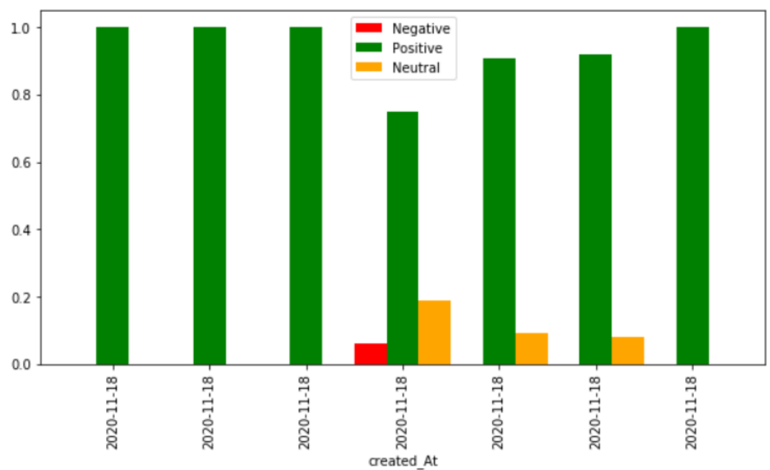

Figure 5: Boeing Sentiment Score

which conveys that the negative tweets were influential in affecting the stock market price. The positive change value indicates the closing value on that day lesser than the day before.

5.3 Time Series Prediction

To answer our RQ3, we used a tensorflow algorithm built by Facebook named prophet. Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.

The input to Prophet is always a dataframe with two columns: ds and y. The ds (datestamp) column should be of a format expected by Pandas, ideally YYYY-MM-DD for a date or YYYY-MM-DD HH:MM:SS for a timestamp. The y column must be numeric, and represents the measurement we wish to forecast.

The Figure 6 represents a graph of predicted Tesla's stock market prices for the week.

The Figure 7 represents a graph of predicted Apple's stock market prices for the week.

6 CONCLUSION

Motivated by the prospect of a forecasting relationship between Twiiter and market performance and the prospect of evaluating Twitter as a social media tool in a market context, data mining methods using the Twitter API allowed for a preliminary study to be carried out investigating the relationship between mood and stock market performance. Lexicon based sentiment analytics allowed for data sentiment to be classified in order to measure the general public mood in relation to the real-time events and stock market indicators such as volume of trades, market closing price and the average daily change in price were collected to track the market movements around the same time. A range of correlation

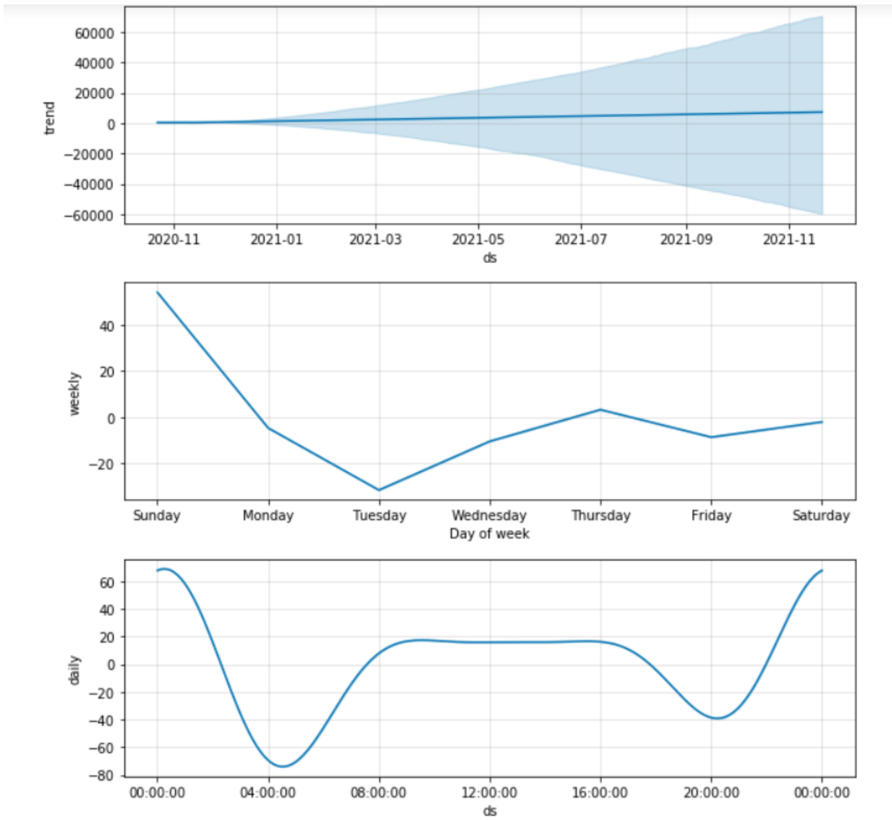

Figure 6: Tesla Price Prediction

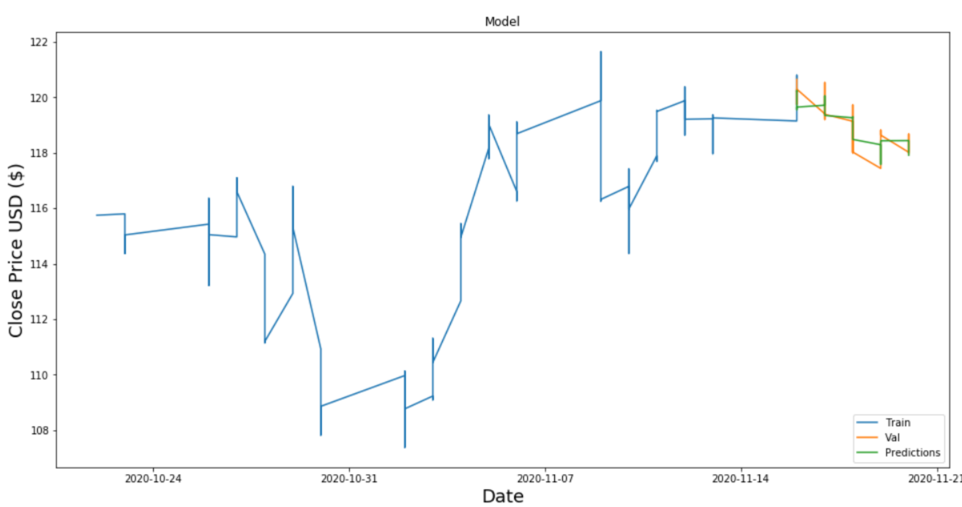

Figure 7: Apple Price Prediction

and regression based tests determined the existence or lack thereof of relationships between the variables in context and allowed this investigation to explore the influence of general public opinions on investment decision making. Although the sample size prevented this study from acquiring statistically significant results in regards to the relationship between Twitter chatter and stock market movements, the following trends can be observed: 1. MOOD and CLOSE show potential of a strong causation relationship in this sample, almost at the significance level. 2. There are promising correlations between CHANGE and MOOD with various time lags highlighting the possibility of using Twitter chatter as a forecasting tool for future Stock Market performance.

In order to conclude, the original research questions must be reflected upon. In terms of Question 1, observations in this study have determined that correlation does exist between public sentiment in regards to the stock market data in the form of correlation between MOOD and CHANGE; however, the relationship is not yet determined as statistically significant. In relation to Question 2, findings in this investigation suggest that there is evidence of causation between public sentiment and the stock market movements, in terms of the relationship between MOOD and CHANGE, and the time lag findings of MOOD and PRICE, however, they are also yet to be determined statistically significant. We can identify some key limitations of this investigation: the first issue with the sample size is whether it is a good representation of the views of the Twitter population. The final tweet sample in this investigation contained over 1,80,000 tweets over a 21 day period. Considering Twitter is updated hundreds of millions times a day this seems an insignificantly sized sample to be a fair representation of the population mood. Having said that, the sample of this study should aim to be a good representation of the discussions on Twitter by the population, and as the total number of them is unknown, it is difficult to say whether the sample of Tweets is a fair representation of the general mood of the public. The second and more detrimental issue to the study is the size of the specified sample period. However, just because a relationship is found as not statistically significant, it does not mean that the relationship does not exist. Finally, it should be noted that the mood expressed in this study is reflective of the discussions specifically related to the issue of the real-time events occuring during the time of collection of the data. This early study looking into the relationship between political discussions and market performance shows promise for using sentiment analytics on Twitter data for potentially forecasting market movements. The results highlight groups of variables with stronger cause and effect relationships that should be the focal point of similar studies in the future (MOOD and CLOSE). Further studies may wish to explore the relationship between Twitter based political discussions and stock market movements using a broader sample of Tweets or an extended period of observation. Based on the trends as indicated by the results it is believed that in an extended investigation looking at examples, there is a possibility that the mood of the Twitter debates at the time, can be compared to stock market movements in order to measure the influence of political uncertainty in Twitter chatter on businesses, and to predict the performance of businesses via investment indicators based on public perceptions on the economic conditions.

REFERENCES

[1] https://www.omnicoreagency.com/twitter-statistics/, 2020.

[2] https://developer.twitter.com/en/docs/twitter-api/tweets/sampled-stream/introduction, 2020.

[3] https://www.investing.com/equities/trending-stocks, 2020.

[4] https://github.com/cjhutto/vaderSentiment, 2020.

[5] Jonah Berger and Katherine L Milkman. What makes online content viral? *Journal of marketing research*, 49(2):192–205, 2012.

[6] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, P Krishna Gummadi, et al. Measuring user influence in twitter: The million follower fallacy. *Icwsm*, 10(10-17):30, 2010.

[7] Yubo Chen and Jinhong Xie. Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, 54(3):477–491, 2008.

[8] Joel Comm. *Twitter power 2.0: How to dominate your market one tweet at a time.* John Wiley & Sons, 2010.

[9] A Java. Why we twitter: understanding microblogging usage and communities. joint 9th webkdd and 1st sna-kdd workshop'07. san jose, ca, usa. *http://ebiquity. umbc. edu/paper/html/id/367/Why-We-Twitter-Understanding-Microblogging-Usage-and-Communities*, 2007.

[10] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600, 2010.

[11] Yuexin Mao, Wei Wei, Bing Wang, and Benyuan Liu. Correlating s&p 500 stocks with twitter data. In *Proceedings of the first ACM international workshop on hot topics on interdisciplinary social networks research*, pages 69–72, 2012.

[12] Victor Niederhoffer. The analysis of world events and stock prices. *The Journal of Business*, 44(2):193–219, 1971.

[13] Andrew N Smith, Eileen Fischer, and Chen Yongjian. How does brand-related user-generated content differ across youtube, facebook, and twitter? *Journal of interactive marketing*, 26(2):102–113, 2012.