Twitter Bots and Stock Market Manipulation

Carl Terve
Linköping University
Linköping, Sweden
carte495@student.liu.se

Mattias Erlingsson
Linköpings University
Linköping, Sweden
mater915@student.liu.se

***Abstract*—Stock market manipulation is a growing threat, exacerbated in part by the activities of social bots. In this project, we conduct a study on the activity of these bots on Twitter. Therefore, by focusing on a few of the currently most shorted stocks on the US market, we track their short interest and Twitter activity during the time period 2021-12-31 to 2022-03-31. Then, using an automated sampling methodology, we define and determine the most normal and abnormal period with regard to change in short interest per company. During these periods, we collect all company-related tweets and use the Botometer API to evaluate the twitter bots' share of activity and how it correlates to the change in short interest. The results provide some initial insight into the prevalence of social bots on Twitter for shorted companies. However, no significant difference in bot ratio was identified between the min- and max-periods.**

***Index Terms*—Twitter, Stock market manipulation, Bots, social bot, short-interest**

I. Introduction

A. Social media and the stock market

Stock market manipulation in social media is a growing threat and it has been shown that much of the stock-related content on Twitter is produced by bots [8]. Moreover, Twitter bots have been shown to spread information from low-credibility sources that should not be trusted, especially not for financial investments. Nevertheless, human users are frequently, further spreading the misinformation provided by social bots [20], enabling their ambition and making them more effective. Furthermore, the social bots on Twitter are becoming increasingly sophisticated and difficult to detect [12] which further emphasises how bots can be used to mislead private investors and put distrust in the financial system. Adding to this, it seems to be a rising interest from online communities to track shorted companies. The most prevalent example is perhaps with GameStop (GME) in early 2021, heavily shorted at the time, where investors on the online platform Reddit came together with a strategy to squeeze hedge funds by pumping the stock price [5]. The incident gained large public interest and to some extent proved the potential of social media as a tool for manipulating the stock market.

B. Bot or not

Bots on social media are accounts that are controlled by software, defined and configured to imitate the behavior of ”normal” user accounts. These bots can produce their own content as well as interact with content posted by other users. For certain content on Twitter, such as the sharing of links to popular websites, it has been estimated that 66 % of tweets are derived from automated accounts [25]. While some bots can be considered useful, helping spread news etc. [15], bots used for malicious intentions are a growing concern. Research show how bots have been used to influence political polarization [6, 7] and to manipulate the stock market [11], to mention a few. Evaluating the impact of social bots relies on being able to accurately identify bots with detection algorithms. This has been an evolving field of study during the past decade, with a large computing community engaging in the advancements of more sophisticated algorithms. Ferrara et al. [11] outlines a taxonomy of bot detection algorithms, ultimately dividing them into the three categories briefly mentioned below.

1) Graph based methods: The first category rely on analysing the structure of social graphs, where it is assumed that nets of social bots connect to other bot nets to increase their social ties, making them more trustworthy. Graph based algorithms exploit this trait to detect interconnected groups of bots. However, most of these algorithms rely on the assumption that legitimate users wont interact with bot accounts, which has been disproved in literature [2]. Sophisticated bots that succeed in infiltrating legitimate communities will thus circumvent the graph based detection algorithms.

2) Crowdsourcing: The second category, crowdsourcing social bot detection, assume that bot detection is a simple task for humans due to the ability to assess language usage. Studies testing the efficiency of the strategy show good result, but there are shortcomings to the methodology such as scalability and cost effectiveness, making it unfeasible for large social networks such as Twitter.

3) Feature-based: The third category utilizes behavioral patterns in users to encode features that can be used by machine learning techniques to distinguish between legitimate accounts and bots. One such detection system for Twitter was released and made publicly available in 2014, called *Bot or Not?* [9]. Since then, it has been renamed to *Botometer* and has gone through some changes [27, 19].

In this study the feature based detection system *Botometer API* [3] is used for bot detection. The API extracts over 1000 features that are used to characterize a user based on its temporal behavior, user profile, social network and use of language etc. With machine learning models the system retain bot scores that can be used to determine the likelihood of a user being a bot. Unlike other bot detection tools, Botometer has been around for a long time, is well maintained, and has regularly been updated and upgraded. Moreover, Botometer provides an easy to use interface and has been validated in research [26]. It is therefore a good fit for the use case in this study.

C. *Define the problem*

Evaluating the effect of social bots and to evaluate the significance of Twitter bot activity on the movements of the stock market is an on-going research area within computer science. Previous research show how Twitter mood and social media sentiment predict return on financial instruments [13] [1], but it is not yet clear to what extent the social bots influence the sentiment or the market itself. This paper focuses on gaining a better understanding of the behaviour and activity of social bots surrounding points of interest in short interest. By collecting relevant bot and non-bot created Twitter posts and financial data for a collection of highly shorted companies, statistical tests is performed to calculate the presence of social bots around peaks in short interest change compared to ”normal” times.

D. *Explain the relevance of the problem*

By evaluating the behavior of social bots discussing stocks on Twitter, we gain a better understanding of their influence on the stock market. More specifically, the results can be used to better understand the dynamic of social bot activity for shorted companies and to what extent they may affect the public sentiment of a stock.

E. *Structure of the report*

This paper will first explore related research and their main contributions to the subject area. Following will be the company selection strategy for the set of stocks that will be evaluated in this paper, which leads to the next section covering the collection of financial and Twitter data related to the set of stocks and the bot evaluation of sampled users. The result will then follow which will present the Botometer API evaluation results. Finally a discussion of the results will be presented.

II. Research questions

The main problem that this paper aims to address, is that there is limited knowledge and research regarding social bots, Twitter, and their effects on the stock market. The report addresses the following research questions:

- What is the behaviour of social bots during short-periods?
- Do presence of social bots correlate to change in market sentiment?
- Do social bots manipulate the market?

III. Related work

A lot of research on social media has focused on the spread of fake news. Some contributions include research on how social media affect the public discourse in topics such as the COVID-19 pandemic [18] and the 2016 US election [4]. Other research study the role of social bots in the spread of political propaganda [6], and their effect on the stock market [11], where it has been shown that bots play a central role in the exchange of significant content. Others have focused on social media sentiment as a predictor for the stock market, and it has been shown that the social sentiment on Twitter predicts the return on financial instruments [13] [1]. Furthermore, there has been studies providing some evidence of the effects of social bots on Twitter on the stock market by correlating surges in bot activity to simultaneous changes in a stocks value [10, 21].

Papers [10, 25] employ the Botometer API for bot detection and one of the papers [10] show that their set of detected bots later had $48\,\%$ of the accounts removed by Twitter shortly after. This indicates that these were likely social bots and indeed correctly labeled as bots by the Botometer API. The original paper for the Botometer API [24] claim a high accuracy of $86\,\%$ for the bot detection and estimates that between $9\,\%$ and $15\,\%$ of all Twitter users were bots in 2017 when the paper was published. Since the release of the Botometer API, it has been extensively tested in the field. Understanding the impact of social bots on Twitter seem to be the most common use case in research, with topics including vaccine debates [28], election campaigns [22], climate change [16] and finance [10] among many.

Even if research has been done on correlating social bot activity to the stock market, no one has so far looked into the bot dynamics of heavily shorted stocks specifically. By not analysing correlation to price and volume, our work differentiates itself by instead focusing on the short interest metric.

IV. Company selection

This study focuses on companies that are heavily shorted, which is commonly measured using short interest. This metric, when expressed as a percentage, is calculated as the number of shares shorted for a company, divided by the number of shares outstanding. One reason for investors to engage in short selling is to take profit from a decline in the price of a stock. This makes short interest a great indicator of market sentiment. Short interest data is relatively inaccessible to the public crowd and is only reported every two weeks. This led us to base our selection on an already assembled list of the 200 most shorted stocks, published on Yahoo! Finance [14], as per April 2022. This list is based on US companies with a trading volume of at least 200,000 on average per day during the past three months, and the short data is based on Yahoo! Finance's data and data provided by Morningstar. We decided to focus the analysis on a time period stretching three months back from the latest short interest update on the 31st of Mars 2022. The time period will thus be from 2021-12-31 to 2022-03-31. Given this time period, we gathered the total number of tweets and retweets for each of the 200 companies on the short-list. A threshold for the minimum average Twitter activity per day was set to 500 to further limit the selection. This resulted in 13 companies, all present in the 200 top shorted list with at least 500 tweets or retweets per day during the three month period.

V. Data collection and bot ratio evaluation

This section presents how the data collection for Twitter and financial data related to a company was performed and how a user was evaluated as a bot or not using the Botometer API. This part also discusses rate limits and how it affects the data collection and bot evaluation.

A. Data collection

Due to rate limits for the Twitter API we decided to further investigate two short periods in the time period 2021-12-31 to 2022-03-31, which at the time of writing this paper was the three most recent months. Given the rate limits of the Twitter API the companies with stock symbols AMC and ATOM were not possible to include. From this point on these companies will not be a part of the paper and when referring to the selected companies these are not included. The financial data for the selected companies in Table II was extracted from Morningstars historical data [17] for the time period 2021-12-31 to 2022-03-31. For the purpose of this paper the main financial metric of interest is the short interest. Short interest is updated bi-weekly on Morningstar which gives that one short period spans two weeks and the selected time period mentioned earlier contains seven short interest updates. The two chosen short periods are extracted based on the largest and smallest change in short interest respectively for each company in terms of percentage points. Therefore, the two short period time intervals within the three month time period may vary between companies and the data collection was performed on these somewhat varying short periods for each company.

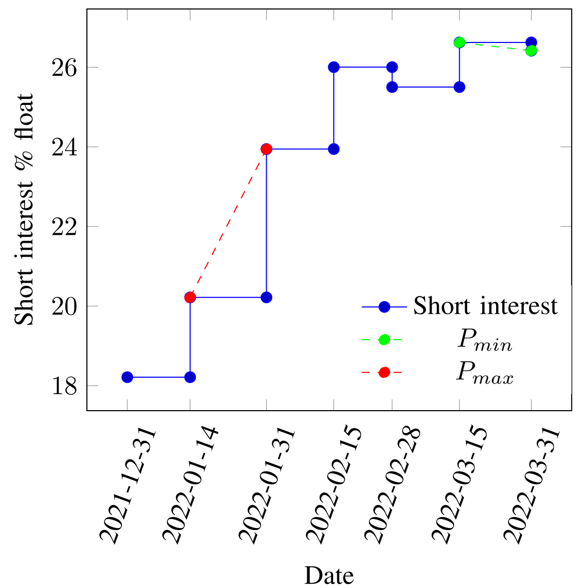
1) Company short period selection: This section aims to further elaborate on how the two minimum and maximum short interest change periods were extracted for each company within the three month time period 2021-12-31 to 2022-03-31. For the time period there are seven short interest updates for all the companies in Table II. The first update being on 2021-12-31 and the last on 2022-03-31. The change in short interest given to a two week period between the bi-weekly updates is calculated as the change from one short interest update to the next. More precisely, it is calculated as the most recent short interest update over the prior short interest update two weeks prior. If there was no change, the ratio would be one. The equation can be formulated as

$$\Delta SI = \frac{SI_t}{SI_{t-1}} \tag{1}$$

where t is a short update date, $t-1$ is the short update date 2 weeks prior to t and SI is the measured short interest value at the given time.

Using equation 1 a total of six short periods spanning two weeks with a calculated short interest change was extracted. The periods with the minimum and maximum short interest

Fig. 1. GME short interest 2021-12-31 to 2022-03-31 and annotated maximum and minimum short interest change.



change were chosen based on change in percentage units provided by equation 1. A visualization of such a calculation can be seen in Figure 1. In the figure the 2-week period that sees the largest short interest change is marked in red, whereas the two-week period with the least short interest change is marked in green.

2) Twitter data: Financial tweets are commonly posted containing one or more cashtags. Cashtags are a way of presenting security tickers in a tweet whether or not they are actually related to the contents of the Tweet or not [8]. Twitter then allows through their developer API [23] to extract tweets and referenced tweets based on cashtags. The structure of a query to the Twitter API for the company GameStop may look as follows, Query: '"GME—"GameStop Corp."—"GameStop"' and the returned header fields for a tweet can be seen in Table I. If a tweet references other tweets another header called *referenced_tweets* is also included. The *referenced_tweets* header contains the *type* (quote or retweet) and the *id* of each referenced tweet. By utilizing the cashtag functionality of Twitter, all tweets and referenced tweets in each companies minimum and maximum short interest change periods were collected. In Table II the collection meta data can be seen.

B. Botometer data

The Botometer API is able to evaluate a Twitter user as either a bot or not. This paper uses the Botometer-V4 endpoint which is a model that allows for bot detection. Botometer gives various metrics in a response but for the purpose of this paper the header *cap:english*, and *cap:universal* will be used to extract a user botscore which is a probability of a user with equal or greater score being a bot. Since the Botometer API comes with rate limits random sampling was applied to produce a valid representation of the entire dataset. The

TABLE I
RESPONSE HEADERS OF A TWEET

id	created_at	text	author_id	lang	conversation_id	retweet_count	reply_count	like_count	quote_count

TABLE II
TWITTER DATA SUMMARY. HERE WE SHOW THE TOTAL NUMBER OF TWEETS FOR THE MINIMUM AND MAXIMUM SHORT INTEREST CHANGE PERIODS FOR EACH COMPANY AND THE SHORT INTEREST CHANGE IN PERCENTAGE POINTS FROM THE PREVIOUS SHORT INTEREST UPDATE.

Symbol	Company	P_{max}	$P_{max}(tweets)$	$P_{max}(SI\ change\ pp)$	P_{min}	$P_{min}(tweets)$	$P_{min}(SI\ change\ pp)$
TLRY	Tilray Brands	21/12/31 - 22/01/14	104770	2.39	22/01/14 - 22/01/31	4285	0.05
RIDE	Lordstown Motors	21/12/31 - 22/01/14	8671	3.10	22/02/15 - 22/02/28	9331	0.32
OCGN	Ocugen	22/02/15 - 22/02/28	18558	3.67	22/01/31 - 22/02/15	8609	0.22
MSTR	MicroStrategy	22/01/14 - 22/01/31	7148	5.31	22/02/15 - 22/02/28	7843	0.16
UPST	Upstart Holdings	22/01/14 - 22/01/31	7109	2.29	22/02/28 - 22/03/15	8126	0.46
GME	GameStop	22/01/14 - 22/01/31	99049	3.73	22/03/15 - 22/03/31	149328	0.21
SFM	Sprouts Farmers Market	21/12/31 - 22/01/14	9071	2.06	22/02/28 - 22/03/15	8604	0.12
BIG	Big Lots	22/01/31 - 22/02/15	39711	12.39	22/02/15 - 22/02/28	5490	1.10
CAKE	Cheesecake Factory	22/03/15 - 22/03/31	60909	2.50	22/01/31 - 22/02/15	40069	0.72
BBIG	Vinco Ventures	22/02/15 - 22/02/28	19595	9.91	22/03/15 - 22/03/31	35631	1.96
SOLO	Electrameccanica Vehicles	21/12/31 - 22/01/14	19903	1.26	22/02/15 - 22/02/28	5722	0.10

data that was sampled was based on the collected tweets for the companies in Table II which shows the metadata of the companies minimum and maximum short interest change periods. In order to remain within the rate limits of the Botometer API, 1400 unique tweeters were sampled from each period of each company, resulting in 2800 total tweeters being evaluated as a bot or not per company.

1) Sampling: The sampling was done in a way to make it replicable. For each companies respective short periods, every unique user was first extracted based on *author_id* into a list for each period of each company. Each list was then sorted based on *author_id* before randomly sampling 1400 users with $seed = 10$. All sets of sampled users were then combined into a combined large set in order to remove any overlap of Twitter users between the samples in order to avoid evaluating a user more than once with the Botometer-V4 endpoint. Before combining the sampled sets into one large set, the number of users summed up to a total of 30800 for the eleven companies respective minimum and maximum short interest change periods. When combined into the large set, 27429 unique users remained when any duplicate entries had been removed. This final set of unique users is the set that was evaluated using Botometer.

In Table III the Botometer API evaluation results can be seen. For each short period of each company the table shows Botometer complete automation probability (CAP) scores for the samples. The table shows the CAP average english botscore where the main language is english, if the majority language is not english the english score is dropped, and the universal score which includes all users. For 118 of the 27429 users the Botometer API failed to evaluate and are therefore excluded.

VI. DATA SET SUMMARY

In Table II a summary of the dataset used for the analysis can be seen. For each company in the list, a period P_{max} and P_{min} has been chosen to narrow down the scope for the Twitter data collection. Each period represent the two-week period with the largest and respectively the least change in short interest between 2021-12-31 and 2022-03-31. Table II also shows the number of tweets collected per company and per period as well as the short interest change in percentage points for the given period. As an example, the largest short interest change belongs to BIG between 2022-01-31 to 2022-02-15 with a change of $12.39pp$ and GME has the most tweets in a single period with $149,328$ tweets in the 2-week period between 2022-03-15 to 2022-03-31. As for the Botometer data, this is summarized in Table III and discussed in more detail in the analysis.

VII. BOT DENSITY ESTIMATION

In Table III the average Botometer score is displayed for each period and company. As is clear in the table, the average score ranges somewhere between 0.6 and 0.85 and there is no significant difference between the average score in the P_{min} period and the P_{max} period for any company. However, there is a difference between the average score when comparing the english and the universal scores. Considering all users, no matter the language (i.e. universal scores) gives on average a higher botscore.

The bot density estimation approach of this report uses a conservative threshold for classifying users as bots. By setting the CAP botscore subvalues *english* and *universal* threshold to 0.95 it allows for a binary classification of a user as a bot or not with a false positive rate of less than $5\,\%$. In Table IV we show the resulting number of english users considered per company, number of english bots, and the english user bot ratio for P_{max} and P_{min}. In total, 69 users are classified as bots in the max-periods and 63 users are classified as bots in the min-periods. The average bot ratio is also slightly higher in the max-periods with $0.5\,\%$ bots compared to the average bot ratio in the min period of $0.45\,\%$.

In Table V we show the resulting number of universal users considered per company, universal bots, and the universal user

TABLE III

BOTOMETER DATASET SUMMARY BROKEN DOWN PER COMPANY OF INTEREST. HERE WE SHOW THE AVERAGE COMPLETE AUTOMATION PROBABILITY (CAP) FOR ENGLISH SCORE WITH MAJORITY LANGUAGE ENGLISH AND UNIVERSAL SCORE FOR ALL SAMPLED USERS. THE DATA IS FOR THE CHOSEN SHORT PERIODS OF EACH COMPANY.

Ticker	P_{max}	$P_{max}(avg\ botscore\ english)$	$P_{max}(avg\ botscore\ universal)$	P_{min}	$P_{min}(avg\ botscore\ english)$	$P_{min}(avg\ botscore\ universal)$
TLRY	22/12/31 - 22/01/14	0.691234	0.727521	22/01/14 - 22/01/31	0.685825	0.730753
RIDE	21/12/31 - 22/01/14	0.765064	0.780305	22/02/15 - 22/02/28	0.767808	0.786381
OCGN	22/02/15 - 22/02/28	0.715926	0.765918	22/01/31 - 22/02/15	0.724327	0.770461
MSTR	22/01/14 - 22/01/31	0.644869	0.687327	22/02/15 - 22/02/28	0.784526	0.796267
UPST	22/01/14 - 22/01/31	0.622465	0.679878	22/02/28 - 22/03/15	0.656387	0.713308
GME	22/01/14 - 22/01/31	0.635222	0.684732	22/03/15 - 22/03/31	0.641408	0.693859
SFM	21/12/31 - 22/01/14	0.745608	0.764109	22/02/28 - 22/03/15	0.782599	0.791590
BIG	22/01/31 - 22/02/15	0.798178	0.808481	22/02/15 - 22/02/28	0.834290	0.835939
CAKE	22/03/15 - 22/03/31	0.786575	0.793118	22/01/31 - 22/02/15	0.789534	0.795785
BBIG	22/02/15 - 22/02/28	0.704707	0.750408	22/03/15 - 22/03/31	0.683254	0.728704
SOLO	21/12/31 - 22/01/14	0.793173	0.802993	22/02/15 - 22/02/28	0.766845	0.788055

bot ratio for P_{max} and P_{min}. Compared to when restricting to english users, the results for the universal scores are slightly higher on average. In total, 132 users are classified as bots in the max-periods and 127 users are classified as bots in the min-periods. Similarly to the average botscores, the average bot ratios for the universal score is higher than for the english botscores. In the max-periods the average bot ratio is $0.86\,\%$ and in the min periods it is $0.83\,\%$ showing again how the ratio during the max-period is higher, but not significantly.

In Figures 2 and 3 the bot score distribution for all companies per period can be seen. For the universal scores the frequency is highest in the botscore range $0.80-0.85$ and for the english botscores in the range $0.75-0.80$.

VIII. DISCUSSION AND CONCLUSION

One of the main research questions of this report was to see what can be learnt about the behaviour of social bots during short periods. In this report we have looked closer at the change in short interest in percentage points during the period 2022-12-31 to 2022-03-31 to extract two periods spanning two weeks with the most and respectively the least change in short interest. From the results presented in Tables IV and V we can see that the choice of a conservative threshold (0.95) for the CAP subvalues *english* and *universal* for what is considered a bot resulted in a low bot ratio and a slim difference between P_{max} and P_{min}. Furthermore, we had that Botometer failed to evaluate 118 users, which likely means that these users have been removed from Twitter, potentially because they were bots. However, with the strict threshold we can almost certainly say that the identified bots in fact are bots (false positive $< 5\,\%$) and given that, we can see a slightly larger average bot ratio over all of the selected companies during P_{max} compared to P_{min}. Which provides an indication that the bot density may be greater during periods of large change in short interest, but it would require further research to establish. However, the bot ratio in our samples is not close to the results presented in [24], where $9\,\%$ to $15\,\%$ of the users on Twitter were classified as bots. There may be several explanations for this, for instance, the sampling may not be representative of the entire set and as previously mentioned some bot accounts have likely been removed. Furthermore, the tweets related to the selected companies may overall have a lower bot density compared to the Twitter average.

Another observation that can be made from Tables III, IV, andV is that the *universal* value on average is greater than the *english* value. Additionally, for the *universal* botscore there were close to the double number of values greater or equal to the 0.95 threshold than for the *english* botscores. This raises the question if the Botometer API is better trained for users with a majority language of english resulting in more universal users easily being labeled as a bot.

REFERENCES

[1] Johan Bollen, Huina Mao, and Xiaojun Zeng. “Twitter mood predicts the stock market”. In: *Journal of Computational Science* 2.1 (2011), pp. 1–8. ISSN: 1877-7503. DOI: https://doi.org/10.1016/j.jocs.2010.12.007. URL: https://www.sciencedirect.com/science/article/pii/S187775031100007X.

[2] Yazan Boshmaf et al. “Design and analysis of a social botnet”. en. In: *Comput. netw.* 57.2 (2013), pp. 556–578.

[3] *Botometer by OSoMe*. en. https://botometer.osome.iu.edu/faq. Accessed: 2022-4-29.

[4] Alexandre Bovet and Hernán A Makse. “Influence of fake news in Twitter during the 2016 US presidential election”. en. In: *Nat. Commun.* 10.1 (2019), p. 7.

[5] Bruce Brumberg and JD. “Reddit and GameStop lessons: Former SEC enforcement chief explains stock manipulation and how to avoid trouble”. In: *Forbes Magazine* (Feb. 2021).

[6] Guido Caldarelli et al. “The role of bot squads in the political propaganda on Twitter”. en. In: *Commun. phys.* 3.1 (2020), pp. 1–15.

[7] M Conover et al. “Political polarization on Twitter”. In: *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*. 2011, pp. 89–96.

[8] Stefano Cresci et al. “Cashtag Piggybacking: Uncovering Spam and Bot Activity in Stock Microblogs on Twitter”. In: *ACM Trans. Web* 13.2 (Apr. 2019). ISSN: 1559-1131. DOI: 10.1145/3313184. URL: https://doi.org/10.1145/3313184.

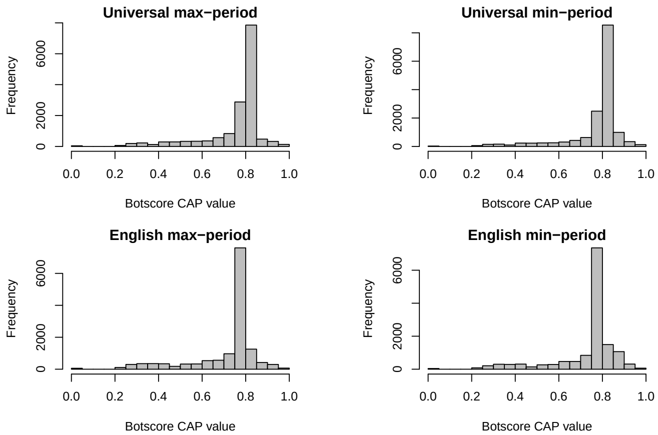


Fig. 2. Botscore distributions for each period and botscore type (english and universal).

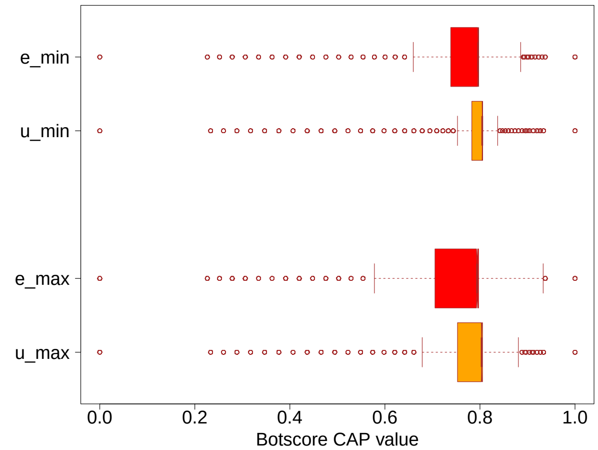


Fig. 3. Botscore distribution comparison for english and universal scores in both the min-periods and the max-periods

TABLE IV
BOTOMETER RESULT ENGLISH LANGUAGE USERS WITH CAP=0.95

Ticker	$P_{max}(\#\ english\ users)$	$P_{max}(\#\ english\ bots)$	$P_{max}(bot\ ratio)$	$P_{min}(\#\ english\ users)$	$P_{min}(\#\ english\ bots)$	$P_{min}(bot\ ratio)$
TLRY	1279	22	0.017	1302	14	0.011
RIDE	1183	2	0.0017	1041	2	0.0019
OCGN	1328	10	0.0075	1350	10	0.0075
MSTR	1245	4	0.0032	1334	0	0
UPST	1196	10	0.0084	1228	12	0.0098
GME	1350	3	0.0022	1324	3	0.0023
SFM	1290	1	0.00078	1195	4	0.0033
BIG	1383	1	0.00072	1390	0	0
CAKE	1236	3	0.0024	1272	1	0.00079
BBIG	1293	9	0.0070	1323	9	0.0068
SOLO	1264	4	0.0032	1238	8	0.0065
Summary	14047	69	0.0050	13997	63	0.0045

TABLE V
BOTOMETER RESULT UNIVERSAL LANGUAGE USERS WITH CAP=0.95

Ticker	$P_{max}(\#\ universal\ users)$	$P_{max}(\#\ bots\ universal)$	$P_{max}(bot\ ratio)$	$P_{min}(\#\ universal\ users$	$P_{min}(\#\ bots)$	$P_{min}(bot\ ratio)$
TLRY	1390	36	0.026	1387	34	0.025
RIDE	1396	6	0.0043	1388	2	0.0014
OCGN	1381	26	0.019	1381	27	0.020
MSTR	1395	7	0.0050	1397	3	0.0021
UPST	1397	12	0.0086	1391	20	0.014
GME	1397	4	0.0029	1390	6	0.0043
SFM	1399	0	0	1400	4	0.0029
BIG	1395	3	0.0022	1399	4	0.0029
CAKE	1391	3	0.0022	1393	6	0.0043
BBIG	1387	21	0.015	1392	13	0.0093
SOLO	1399	14	0.010	1395	8	0.0057
Summary	15327	132	0.0086	15313	127	0.0083

[9] Clayton Allen Davis et al. “BotOrNot: A system to evaluate social bots”. In: *Proceedings of the 25th International Conference Companion on World Wide Web - WWW ’16 Companion*. New York, New York, USA: ACM Press, 2016.

[10] Rui Fan, Oleksandr Talavera, and Vu Tran. “Social media bots and stock markets”. en. In: *Eur. Fin. Manag.* 26.3 (2020), pp. 753–777.

[11] Emilio Ferrara et al. “The rise of social bots”. en. In: *Commun. ACM* 59.7 (2016), pp. 96–104.

[12] Giorgia Guglielmi. “The next-generation bots interfering with the US election”. en. In: *Nature* 587.7832 (2020), p. 21.

[13] Olivier Kraaijeveld and Johannes De Smedt. “The predictive power of public Twitter sentiment for forecasting cryptocurrency prices”. en. In: *J. Int. Financ. Mark. Inst. Money* 65.101188 (2020), p. 101188.

[14] *List of most heavily shorted stocks*. en. Accessed: 2022-4-8. URL: https://stockmarketmba.com/listofmostheavilyshortedstocks.php.

[15] Tetyana Lokot and Nicholas Diakopoulos. "News Bots: Automating news and information dissemination on Twitter". In: *Digit. journal.* 4.6 (2016), pp. 682–699.

[16] Thomas Marlow, Sean Miller, and J Timmons Roberts. "Bots and online climate discourses: Twitter discourse on President Trump's announcement of U.S. withdrawal from the Paris Agreement". en. In: *Clim. Policy* 21.6 (2021), pp. 765–777.

[17] *Morningstar*. https://www.morningstar.com/. Accessed: 2022-04-20.

[18] K S. Jayaraman. "Fake news at the forefront of COVID-19 crisis". en. In: *Nature India* (2020).

[19] Mohsen Sayyadiharikandeh et al. "Detection of novel social bots by ensembles of specialized classifiers". In: (2020). eprint: 2006.06867.

[20] Chengcheng Shao et al. "The spread of low-credibility content by social bots". en. In: *Nat. Commun.* 9.1 (2018), p. 4787.

[21] Serena Tardelli et al. "Characterizing social bots spreading financial disinformation". en. In: *Social Computing and Social Media. Design, Ethics, User Behavior, and Social Network Analysis*. Cham: Springer International Publishing, 2020, pp. 376–392.

[22] R Tobias and Ulrike Keller. "Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications". In: *Political Communication* 36.1 (2019), pp. 171–189.

[23] *Use cases, tutorials, & documentation*. en. https://developer.twitter.com/en. Accessed: 2022-4-8.

[24] Onur Varol et al. "Online human-bot interactions: Detection, estimation, and characterization". In: (2017). eprint: 1703.03107.

[25] Stefan Wojcik et al. "Bots in the Twittersphere". en. In: *undefined* (2018).

[26] Kai-Cheng Yang, Emilio Ferrara, and Filippo Menczer. "Botometer 101: Social bot practicum for computational social scientists". In: (2022). eprint: 2201.01608.

[27] Kai-Cheng Yang et al. "Scalable and generalizable social bot detection through data selection". en. In: *Proc. Conf. AAAI Artif. Intell.* 34.01 (2020), pp. 1096–1103.

[28] Xiaoyi Yuan, Ross J Schuchard, and Andrew T Crooks. "Examining emergent communities and social bots within the polarized online vaccination debate in Twitter". en. In: *Soc. Media Soc.* 5.3 (2019), p. 205630511986546.