The University of Texas at Austin
Center for Identity

Social Media Trustworthy User Identification Across Politics and Finance Domains

Teng-Chieh Huang

Razieh Nokhbeh Zaemm

Suzanne Barber

UTCID Report #22-07

July 2022

Social Media Trustworthy User Identification Across Politics and Finance Domains

TENG-CHIEH HUANG, RAZIEH NOKHBEH ZAEEM, and K. SUZANNE BARBER, The University of Texas at Austin, USA

This research seeks to answer two ever-pressing questions for those who rely on social media for information – What information can I trust? and Who can I trust? Specifically, this research addresses how to recommend trustworthy information from specific information sources using an innovative method of combining interpersonal trust on the Internet and classification algorithms for various target domains. A variety of domains have been studied to see how information from social media can predict or explain the phenomena of the domain, from politics (who wins the next election) to finance (what happens to a company's stock). It is important, however, that as we improve our methods to interpret the information retrieved from social media, we also pay attention to the quality of that information. Social media information is noisy and possibly contains useless, misleading, or even malicious information distributed by untrustworthy users. It is paramount to filter credible and trustworthy information generated by trustworthy users and domain experts from contaminated data, advertisements, or scams.

In this paper, we develop a novel, domain-independent method to distinguish three types of social media (specifically Twitter) users from one another: typical users, domain-related users, and experts. We take a comprehensive list of trust attributes (i.e., measurable properties of a social media user account or posts such as the number of tweets or the number of followers) from previous work and study the value of these trust attributes for the three types of users. The value of the trust attribute for typical users is the average value for all the users. Domain-related users are those who tweet with a set of domain-related handles (e.g., @a_stock_ticker). Experts are real world experts of the filed, who we extract from reputable journalistic sources outside social media. By applying random forest to compare trust attribute performance, we identify which trust attributes can best distinguish expert, domain-related, and typical users from one another. Most importantly, we keep our work independent of the subject domain by performing the same set of experiments in the finance as well as the politics domain. We compare the distributions of trust attributes between finance and politics domains, to test if the classification method can be applied to various target domains and still provide robust results.

Overall, we identify trust attributes capable of distinguishing trustworthy users from malicious ones, or to act as *trust filters*. Our work increases the reliability and utility of social media data for better decision-making. By applying trust filters to filter out untrustworthy and malicious users, the bad impact of social media such as fake news or media manipulation can thus be minimized.

CCS Concepts: • **Information systems** → **Web and social media search**; **Trust**.

Additional Key Words and Phrases: social media, Twitter, trust, misinformation, disinformation, trust filters, stock market, elections

ACM Reference Format:

Teng-Chieh Huang, Razieh Nokhbeh Zaeem, and K. Suzanne Barber. 2018. Social Media Trustworthy User Identification Across Politics and Finance Domains. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The invention of the Internet and the emergence of social media has exponentially increased the *quantity* of sources a person can contact, but some of these sources may not provide high *quality* information. Social media data is noisy, loaded with contaminated data, advertisements, and scams. The presence of misinformation and disinformation on social media is a rising concern. For those who rely on social media for information, be it for personal use or modern-day analytic of social media data, it is important to know who they can trust. There exist users who abuse the capability of social media by spreading spam, fake news, or biased information. These malicious users not only fool their audience but may also distort the information retrieved from social media for analytical or prediction purposes.

This paper promises advances by answering the following questions: (1) What are the indicators of online information trustworthiness? (2) Which indicators are the best predictors of information trustworthiness? (3) Can trustworthy users be reliably distinguished from malicious ones? (4) Can recommendations be developed and implemented by social media users to improve their trustworthiness on the Internet? This research aims to establish a universally applicable trustworthiness recommendation and prediction system for social media platforms.

We aim to find trustworthy information on social media by finding trustworthy users. Specifically, this research addresses how to recommend trustworthy information from specific information sources using an innovative method of combining interpersonal trust on the Internet and classification algorithms for various target domains. We improve the existing and develop new *trust filters*–measurable properties of social media user profile, behavior, connections, posts, etc.–used to identify trustworthy users.

We take a comprehensive list of trust attributes (i.e., properties of a social media user account or posts such as the number of tweets or the number of followers) as potential trust filters from previous work. Further, we consider some established trust filters from previous work.

Using these potential trust filters, we develop a novel, domain-independent method to distinguish three types of social media (specifically Twitter) users from one another: typical users, domain-related users, and experts. We measure trust filters and see how they can rate the three types of users. The value of the trust filter for typical users is the average value for all the users. Domain-related users (e.g., stock-related users) are those that tweet with a set of domain-related handles (e.g., @a_stock_ticker). Experts are real world experts of the filed, who we extract from reputable journalistic sources outside social media.

By (1) observing the distribution of the trust filter values across user types, (2) studying the correlation between trust filters, and (3) applying the random forest classifier to compare their performance in distinguishing the types of users, we identify which trust attributes can best distinguish expert, domain-related, and typical users from one another. Most importantly, we keep our work independent of the subject domain by performing the same set of experiments in the finance as well as the politics domain. Moreover, this model can be applied to other target domains, such as health and entertainment.

The remaining sections of this work are organized as follows. Section 2 introduces related literature that utilize Twitter-retrieved information as a predictor in the politics and finance domains. Section 3 elaborates how the experiment was set up, the definitions of user groups, and the trust attributes. Section 4 shows the distribution of trust attributes and the effectiveness of the attributes as the classification factors. The section 5 finalizes this work and suggests the potential of the trust attributes.

2 RELATED WORK

2.1 Social media predictions in the finance domain

There exists a body of research exploring the power of crowd wisdom as an indicator or a predictor of certain phenomena. Take the finance domain for instance, high social media coverage at the stock level can predict high subsequent return volatility and trading activity [26]. The extracted sentiments of social media users may be used to predict the stock market [27]. Techniques such as text mining [39] and machine learning are frequently used to enhance the predicting power of social media.

For the sake of completeness, we cover some of the most notable related work that seek to predict stock markets, our example target domain, albeit with very different methods. A widely cited work [3] demonstrated how the Twitter mood, in general, predicts the stock market closing values by high accuracy. Nassirtoussi et al. [32] tried to predict foreign exchange markets based on the text of breaking financial news headlines. Dimpfl et al. [12] studied the dynamics of stock market volatility using Internet search queries and found that high stock market volatility can follow high volumes of Internet searches. Nguyen et al. [33] performed stock market prediction based on social media analysis as well. Instead of taking all sentiments into account, they considered only the sentiments of specific topics of the company to predict stock price movement to increase the forecast accuracy. Oliveira et al. [34] sought to predict stock markets through Twitter posts. They found among all the factors, sentiment and posting volume to be particularly important for the forecasting of S&P 500 index.

2.2 Social media predictions in the politics domain

The politics domain, particularly elections, is another hotspot for testing the prediction power of social media. Ever since Twitter started becoming an essential online social platform, researchers have been exploring its potential of predicting the election outcome [41]. The 2016 presidential election of the United States, in particular, brought Twitter under the spotlight of public attention. Compared to the Clinton campaign's strategy, the Trump campaign's more amateurish yet authentic style in social media points towards de-professionalisation and even amateurism as a counter trend in political communication [15]. Moreover, fake news spreading on Twitter [19], social bots distorting public opinion [2], and even Russian interference [31] all played roles on Twitter during the 2016 presidential election. Therefore, it is critical to distinguish real users from fraudulent or malicious sources before utilizing social media as a prediction tool.

The following sections introduce different prediction methods for elections.

2.2.1 Number of Tweets. Many of the earlier work like [38, 41] use the number of tweets which mention the supporting parties or candidates as an indicator. However, this method may result in a higher error rate because not all tweets mentioning the parties or candidates represent the approval to them. One candidate could have a high exposure on social media while most of the comments are negative or totally irrelevant to the campaign.

2.2.2 Sentiments of Tweets. To further improve the accuracy of the forecast, sentiment analysis became popular on top of simply counting the number of tweets among the researchers. Chung et al. [8] categorize each tweet as positive, negative or neutral, then counting the sum of supporting tweets and objecting tweets to another side. Burnap et al. [5] apply sentiment scores (+5 to -1) on tweets and sum the scores up. Different sentiment analysis methods are also applied in [25, 37, 42].

2.2.3 Hashtags as a predicting attribute. Bovet et al. [4] applied hashtags as the opinion finder which are used to train a machine learning classifier. Four clusters has been classified as pro-Trump, anti-Clinton, pro-Clinton and anti-Trump

which show a clear boundary between the usage of hashtags. They first considered only the strongly connected giant components (SCGC), which are formed by the users that are part of interaction loops and are the most involved in discussions. From the distribution of supporters, they pointed out there exists a huge gap between the number of tweets having hashtags exclusively in the Trump supporters and in the Clinton supporters. They then used the same collection of hashtags to calculate the whole Twitter dataset and found the situation reversed - Clinton supporters became the majority of the users. This is due to a huge number of Trump supporters belonging to the SCGC. This paper shows a huge potential of using hashtags as predicting attributes.

2.2.4 Hybrid Methods/Machine Learning. Tsakalidis et al. [40] collect several Twitter-based potential features which originate from the number of posted tweets, positive or negative tweets proportion and the proportion of Twitter users as well. In this research, a poll-based feature is also taken into account. Utilizing the above features as inputs, they have tested several data mining algorithms such as linear regression, Gaussian process and sequential minimal optimization for regression.

2.2.5 Difficulties in Twitter derived election prediction. Even though using Twitter to predict the election seems to be promising and is convenient compared to the traditional polls, there are questions brought up by some researchers which cannot be ignored. In [30], some suggestions on how to correctly predict the election are given. First, you cannot actually "predict" the election retroactively, so anyone who intends to predict the election should choose the methods or words carefully. Second, social media is fundamentally different from real society - there is more likely to exist spammers and propagandists on the Internet than in the real world. Therefore, researchers should consider the credibility of tweets prior to taking all tweets into account. Another literature survey paper [16] also possesses a similar view as [30] that not all tweets are true, so it might be required to filter out the untrustworthy tweets before the main process. As Gayo-Avello [17] pointed out, the researchers cannot consider the Twitter environment as a totally representative and unbiased sample of the voting population. Needless to say, there are a considerable amount of malicious users or spammers spreading misleading information on Twitter.

2.3 Trust filters and social media

The idea of trust filters can be traced back to [28], where the original idea was to filter and rank trustworthy social media users. There were six trust filters: Authority, Experience, Expertise, Identity, Proximity and Reputation. All six filters stem from the traditional way of estimating one's trustworthiness. In succession to the concept, an Internet biosurveillance application [43] was developed to investigate and detect possible outbreak of the disease in the early stage. By comparing the scores generated by trust filters against hospital historical data, this work is able to conclude the significance of trust filters. The following papers [22, 23] applied trust filters on two different topics as the trust indicators. Here trust filters were inserted as an extra layer before the prediction procedure. Both papers show a certain level of improvement on the prediction results when adding trust filters.

In previous work [24], we showed the distinction among different types of user groups: typical users, target domain-related users and experts of specific domains based on the trust attributes. Our work presented the possibility of utilizing social media as a tool to distinguish the more experienced and possibly credible users. That work also suggested some tweet-derived attributes could become promising candidates to differentiate the trustworthy users in specific target domains. That work, however, only shows how trust attributes work in a *single* domain. Not more than one domain was explored.

3 METHODOLOGY

This section introduces data retrieval, the definition for each user category, and trust attributes extracted from tweets used to analyze the user distribution.

3.1 User group categorization

For the finance and politics domain, the users are both categorized based on their expertise, involvement, and reputation to the specific target domain [24].

The definitions of three groups of Twitter users in the finance domain are specified as below.

(1) Typical users: The 1% random sampling of all Twitter users, which stands for the baseline among all groups.
(2) Stock-related users: The users who posted at least one tweet with a reference to the symbols of the stock market companies, such as $AAPL or $TSLA, during the sampling period. This user group represents the users who might have interest in the stock market, while they may or may not be the financial experts.
(3) Financial experts: This group of users are retrieved from well-known online articles which recommended the top-notch financial experts to follow on Twitter [9, 11, 29]. These lists are compiled by their authors or by Wall Street analysts and journalists, who are traditionally considered financial authorities.

To keep the consistency between different target domains, the definitions of three groups of Twitter users were similar to the ones defined in the finance domain, which are shown below.

(1) Typical users: The same user sets as the finance domain.
(2) Political-related users: The users who posted tweets with hashtags of the candidates, during the sampling period.
(3) Political experts: Following same concept as the financial experts, the political experts were recommended by online articles regarding the specific expertise [1, 6, 10, 13, 14, 21, 35].

3.2 Data retrieval and extraction

To retrieve the publicly available collection of tweets, a non-profit digital library–Internet Archive–is selected as the tweet source in this work. The database contains the complete "Spritzer" version of tweets, which stands for 1% sample of overall Twitter public posts in the world.

To keep the consistency of the sampling period, the sampling period all terminated on October 23rd, 2020, the last weekday before the 2020 presidential election in the United States. The number of typical users were limited to 1,000,000 to alleviate the computational burden and hence reduce the computation time.

Originally, there were 257 financial experts' accounts on our list. After removing deactivated users, only 123 of them remain. Likewise, there were 107 out of 188 political experts that were still considered active.

To increase the robustness of the analysis, users who posted less than 3 tweets during the sampling period were excluded from the experiment. In Table 2, there still have around one third of original users after the truncation, which is able to provide sufficient quantities for the following comparisons.

3.3 Tweet-retrieved trust attributes

We derive a set of trust attributes retrieved from tweets for each user and use the trust attributes as indicators of the trustworthiness of users. Here an attribute refers to a user, text, or social connection information of a single social media user. 11 trust attributes (shown in Table 3) are selected based on the analysis of previous research [24] which

Table 1. Sampling period, user counts, and tweet counts for all compared user groups.

User Group	Sampling period	User Counts	Tweet Counts
Typical - Finance	Oct 19th 2020 to Oct 23th 2020	1,000,000	4,132,294
Stock-related	Oct 19th 2020 to Oct 23th 2020	67,159	295,865
Financial experts	June 8th 2020 to Oct 23rd 2020	123	45,801
Typical - Politics	Oct 19th 2020 to Oct 23rd 2020	1,000,000	4,131,777
Political-related	June 8th, 2020 to Oct 23rd, 2020	1,048,575	7,476,859
Political experts	June 8th, 2020 to Oct 23rd, 2020	107	59,507

Table 2. User counts before and after the truncation.

User Group	User Counts	User Counts (Tweets >= 3)
Typical - Finance	1,000,000	336,942
Stock-related	67,159	21,903
Financial experts	123	123
Typical - Politics	1,000,000	336,927
Political-related	1,048,575	246,369
Political experts	107	107

Table 3. The definition of trust attributes.

Trust attributes	Definition	From previous work
Expertise_score	Ratio of tweets which is domain related	[20]
Statuses_count	Total number of posted tweets in user history	[7, 20]
Followers_count	Number of followers of a user	[7, 20, 36]
Friends_count	Number of friends of a user	[7, 36, 36]
Avg_len_tweet	Average tweet length in characters per tweet	[7]
Avg_n_word_tweet	Average number of words per tweet	[7]
Avg_hashtag	Average number of hashtag symbols per tweet	[7, 20, 36]
Avg_tweet_URL	Proportion of tweets containing URLs	[7, 20, 36]
Avg_tweet_question	Proportion of tweets containing "?"	[7]
Avg_tweet_exclamation	Proportion of tweets containing "!"	[7, 20]
Avg_tweet_uppercase	Proportion of tweets composed by upper-case letters	[7]

include tweets content, Twitter user information, and social network structure. The trust attributes we chose are neutral, suitable for universal purpose across all domains and do not require any specialized retrieval technique to calculate the attribute for a user. The *expertise_score* is determined by the "keywords" of the corresponding target domain. In the finance domain the keywords are defined as stock symbols, which is the same as the definition to retrieve stock-related users. In the politics domain the keywords are sets of frequently used election-related hashtags among political-related users. Those tweets containing keywords are counted as domain-related tweets.

4 EXPERIMENT RESULTS

This section presents all the experiment results and the observations from them. To better compare data with different sizes of expert groups, we have normalized the charts. The Y-axis shows the probability density of users instead of

the absolute number throughout this section. Furthermore, for trust attributes which change dramatically with the proportion of users, we will show the y-axis in log scale.

4.1 Inter-domain Comparison

After calculating attribute values for every user, the distribution for three types of users in two target domains are used for the comparison. There are some trust attributes worth mentioning. First, compared to stock-related users, political-related users have higher *expertise_score* on average (Fig. 1). This might be due to the sampling period being close to the election date.

The *statuses_count* distribution in Fig. 2 shows another different trend between two domains. The political-related users in the politics domain have a much lower ratio of high-statuses-count users compared to typical users, which is inconsistent with the finance domain. Therefore, we can infer that there were more political-related users who had fewer history or interaction on Twitter than other users.

We also found the peaks of *avg_len_tweet* and *avg_n_word_tweet* for political-related users are higher than the peaks of the same trust attributes for stock-related users (Fig. 5 and Fig. 6, respectively). The possible explanation could be more frequent usage of hashtags and the more emotional sentences were used in the political tweets such as repeated words. *avg_hashtag* is another trust attribute showing distinct distribution between two target domains (Fig. 7). Political-related users tend to put multiple hashtags in a single tweet. As a comparison, the majority of stock-related users put only one or less hashtag in a single tweet. This could possibly be due to the popularity of using Twitter as a platform to spread the propaganda or certain political slogans especially since the 2016 presidential election [18].

4.2 Intra-domain Comparison

We noticed both the stock-related users and political-related users had higher *expertise_score* (Fig. 1) and *friends_count* (Fig. 3) than typical users. For *friends_count*, we think the reason is that people who are interested in a certain topic would more actively follow the sources regarding the specific topic., hence increasing their average number of friends (i.e., followed sources).

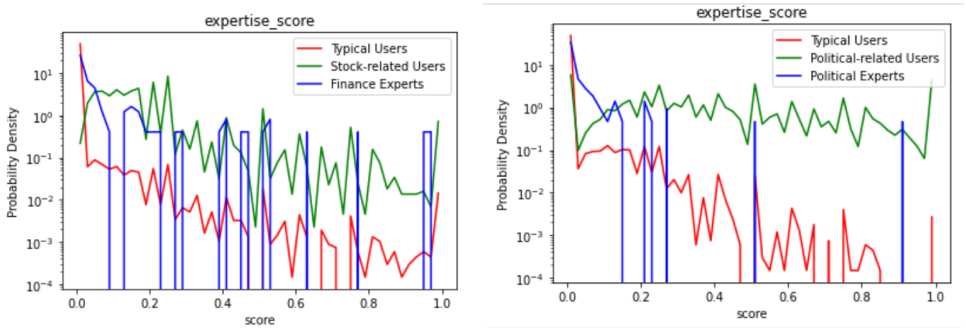

Fig. 1. (Left) Finance domain and (Right) Politics domain. Distribution of *expertise_score* among the three user groups.

To provide a more thorough understanding of the meaning of the trust attributes, we performed the cross-comparison between two attributes which might be highly correlated. First, we compared the correlation between *friends_count* and *followers_count* among all user groups of the two target domains (Fig.8). The cut-off threshold shown in the charts is not a calculation error, it is because of two Twitter follower regulations: (1) The 5,000 friends limit: If a user has less

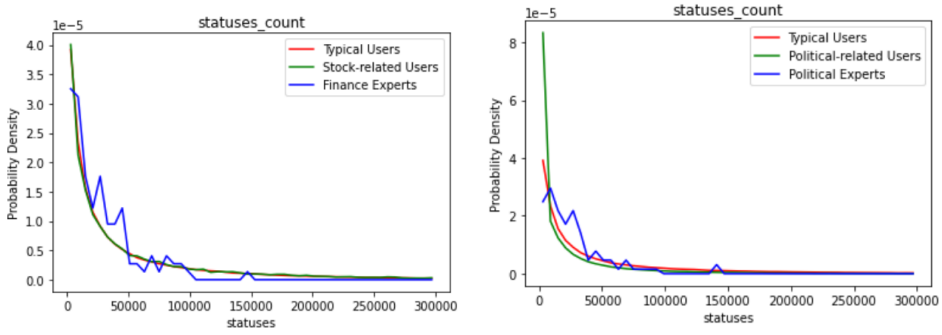

Fig. 2. (Left) Finance domain and (Right) Politics domain. Distribution of *statuses_count* among the three user groups.

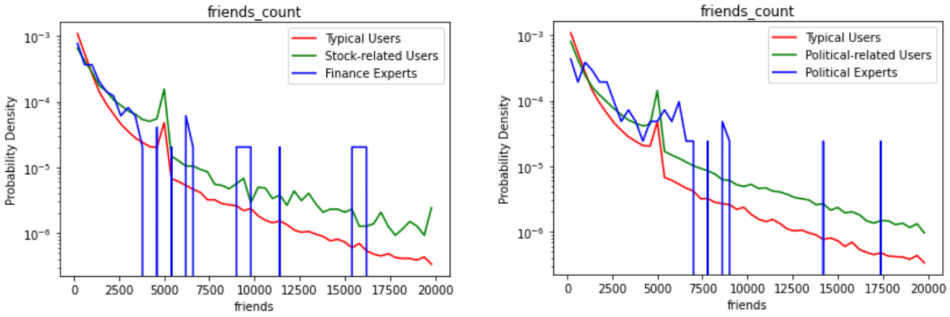

Fig. 3. (Left) Finance domain and (Right) Politics domain. Distribution of *friends_count* among the three user groups.

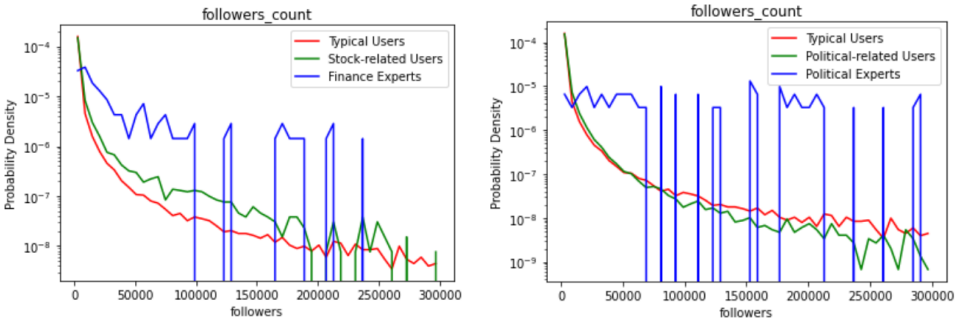

Fig. 4. (Left) Finance domain and (Right) Politics domain. Distribution of *followers_count* among the three user groups.

than 5,000 followers, the maximum he/she can follow is 5,000 accounts. (2) The 10% limit: If a user follows more than 5,000 users, it would be limited to 10% more than the number of people that follow the user. For example, if 5,000 people follow user X, then X can follow a max of 5,500 people; if 10,000 people follow user Y, then Y can follow up to 11,000 people, and so on.

As expected, with similar *friends_count* level, experts usually possess more *followers_count* than the other two user groups no matter in which target domain. It means that with the same amount of friends, the experts can attract more

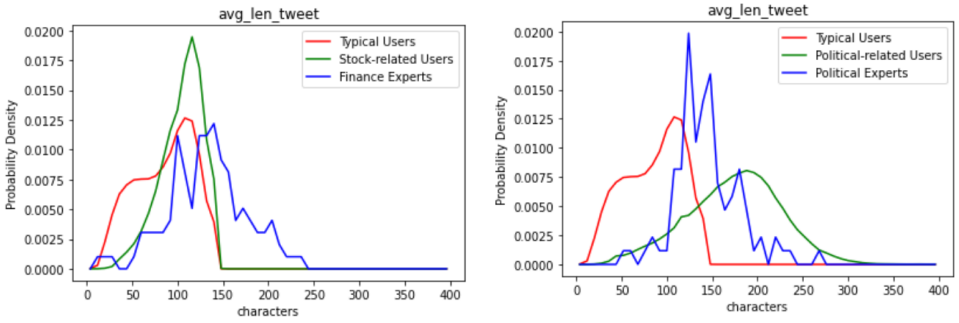

Fig. 5. (Left) Finance domain and (Right) Politics domain. Distribution of *avg_len_tweet* among the three user groups.

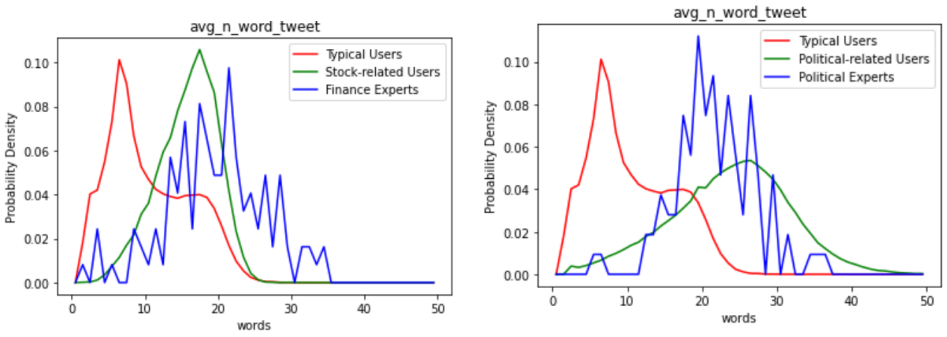

Fig. 6. (Left) Finance domain and (Right) Politics domain. Distribution of *avg_n_word_tweet* among the three user groups.

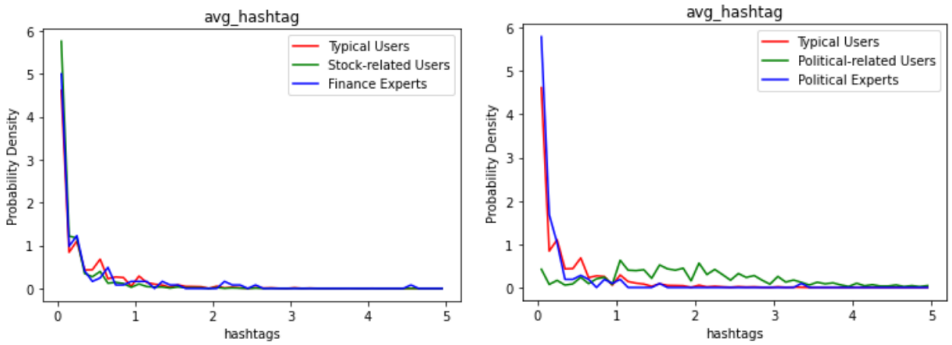

Fig. 7. (Left) Finance domain and (Right) Politics domain. Distribution of *avg_hashtag* among the three user groups.

users as followers then other users. Here we defined the slope as *friends_per_follower*, which will be used in the next section.

The other trust attribute pair is *avg_len_tweet* and *avg_n_word_tweet*. As Fig.9 shows, there is no obvious difference that can be observed. The domain-related users and experts in both domains seem to have steeper trendlines than typical users, while not by large amount. We defined the reciprocal of slope in Fig.9 as *len_per_word*. In the next section,

we will add the two new trust attributes (*friends_per_follower* and *len_per_word*) to the classification to see if the correlation trust attributes can further improve the prediction accuracy.

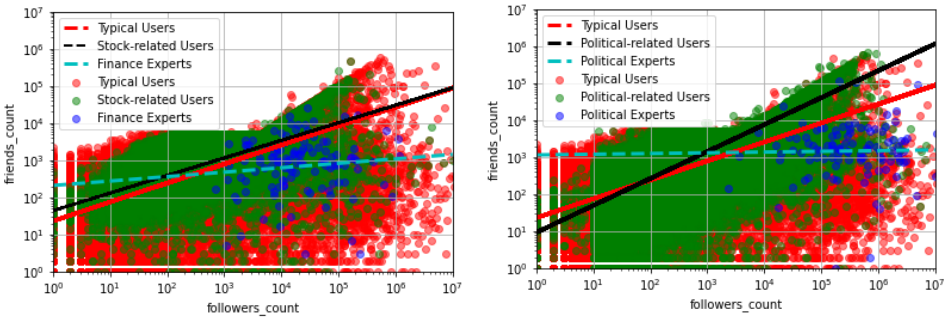

Fig. 8. (Left) Finance domain and (Right) Politics domain. The distribution of users in terms of *friends_count* and *followers_count*. Each dot represents one user. There are three trendlines representing the three user groups, respectively. Slope for each trendline: (left) Typical: 0.5101, Stock-related: 0.4723, Expert: 0.1192. and (Right) Typical: 0.5101, Political-related: 0.7267, Expert: 0.01928.

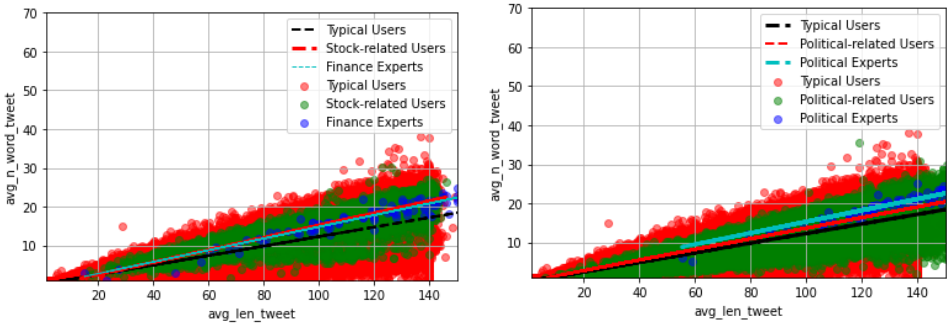

Fig. 9. (Left) Finance domain and (Right) Politics domain. The distribution of users in terms of *avg_len_tweet* and *avg_n_word_tweet*. Each dot represents one user. There are three trendlines representing the three user groups, respectively. Slope for each trendline: (Left) Typical: 0.1226, Stock-related: 0.1579, Expert: 0.1508 and (Right) Typical: 0.1226, Political-related: 0.1357, Expert: 0.1451.

4.3 Random Forests Classification

Random forest is a technique combining decision tree, bagging and random selection of features. One of the benefits of random forest is it can rank each variable according to its importance, which is particularly suitable for this work since it can help us understand the influence of each trust attribute. Random forest also can handle large quantities of data very well since it can be implemented with parallelism, and it can maintain fast prediction and training speed. In this section we compared different combinations of user groups for each target domain, from experts with other users, experts with domain-related users, to domain-related users with typical users. We aimed to discover what distinguishes different types of users.

Fig.10 illustrates how one of the decision trees works in the random forests classification. By combining a set of decision trees, a high quality of classification can be achieved. Fig.11 and Fig.12 show the top level of decision trees for politics and finance domains, respectively.

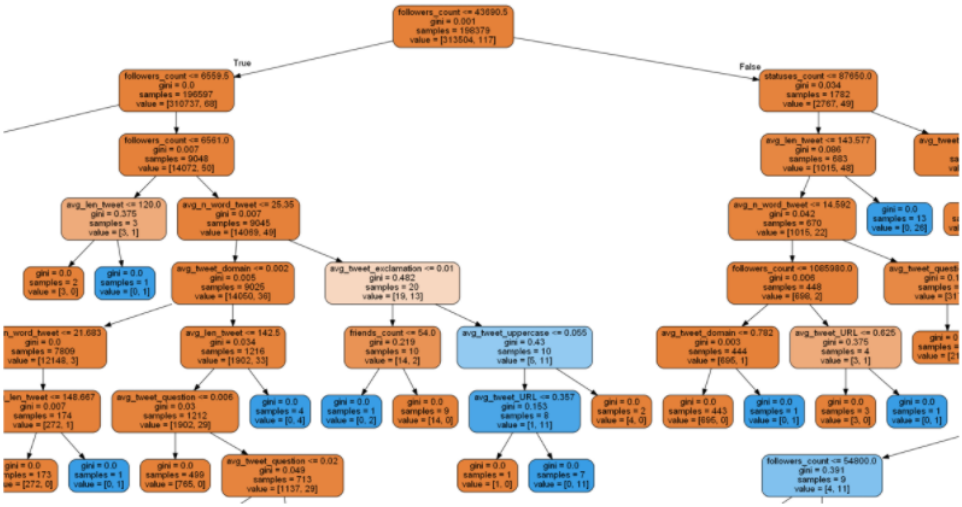

Fig. 10. Partial view of Decision tree0 for the comparing group: financial experts with other users.

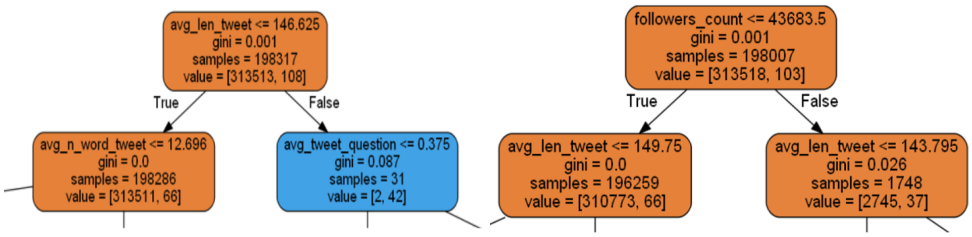

Fig. 11. (Left) Top of decision tree10 (Right) Top of decision tree20 for the comparing group: financial experts with other users.

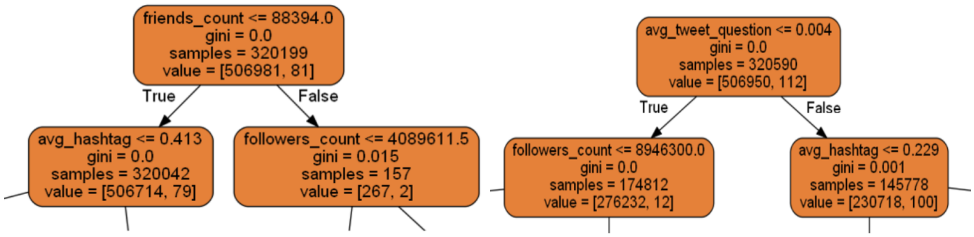

Fig. 12. (Left) Top of decision tree10 (Right) Top of decision tree20 for the comparing group: political experts with other users.

Table4. shows the accuracy and Area Under the Curve of Receiver Operating Characteristic curve (AUC) for the various combinations of classification. Although all the accuracy is high, it is important to note that due to the class imbalance between the experts and other user groups, accuracy alone may become insensitive to the false prediction. AUC, on the other hand, can better represent the actual predicting correctness in this case.

The first observation is the politics domain overall has higher AUC than finance domain with the similar comparing sets. For classification of experts with other users (Group 1 and 2), topic-related users with typical users (Group 3 and 4)

Table 4. The accuracy and AUC derived from random forest classification. The number of estimators is 50. For each target domain, there are three comparison groups: Experts & other users, domain_related users & typical users and domain-related users and experts.

Comparing groups	Accuracy	AUC
1. Political Experts & others	0.9998	0.6923
2. Financial Experts & others	0.9996	0.5833
3. Political-related users & typical users	0.9863	0.9866
4. Stock-related users & typical users	0.9778	0.9227
5. Political-related users & political experts	0.9999	0.8571
6. Stock-related users & financial experts	0.9967	0.7855

or topic-related users with experts (Group 5 and 6), the AUC of politics domain is higher than the AUC of finance domain. The observation shows the selected trust attributes can better classify user groups in the politics domain than those in the finance domain. The second observation is that compared to the classification of experts with other users (Group 1 and 2), the AUC is higher when only comparing the topic-related users and experts (Group 4 and 5). This implies that to better differentiate the experts of a specific target domain from a group of users, we can first filter out the topic-related users, then run the random forests algorithm to gain the predicting accuracy.

To see how each trust attribute contributes to the classification, a set of the feature importance for each attribute is visualized in Fig.9. Table 5 shows the top 3 trust attributes in terms of feature importance for each comparing set. Among all trust attributes, some are more crucial than others when differentiating the user groups. Three peaks (*Followers_count*, *Avg_len_tweet*, and *Expertise*) represent the top three trust attributes that influence the classification the most. The reason for Expertise is obvious, because the definition for topic-related users is the users with a positive Expertise score. Nevertheless, it is still the top 3 trust attributes when identifying the experts.

Followers_count is another important trust attribute, especially when it comes to identifying the experts. This finding is consistent with the distribution differences in Fig.4 where experts possess a higher portion of high *followers_count* users.

Avg_len_tweet is effective for most of the classification, except for the identification of political experts. It is true that the group of stock-related users and typical users has an even lower importance factor, but it is due to the dominance of Expertise in this classification. *Avg_len_tweet* is still the top three attributes in this case.

We also find *statuses_count* is useful when identifying the experts in both domains. *Avg_hashtag*, on the other hand, only works for the politics domain but not as effective for the finance domain.

To test if adding the correlation trust attributes, *friends_per_follower* and *len_per_word*, can improve the prediction accuracy of random forests classification, we re-calculated the classification with the same setting except including the two new trust attributes. The results show in Table 6, Table7, and Fig.14. When comparing the Table 6 with Table 4, we found only the compare groups 1 and 2, i.e. Classification of experts with all other users, achieved noticeable improvement by adding the two trust attributes. Other than that, there is no significant benefit of the new implementation. Likewise, the top 3 trust attributes in Table 7 are mostly unchanged when compared to Table 5. Therefore, we can conclude that there is no significant difference by adding the correlation trust attributes. The only exception is when we directly put the raw data to the random forests classification, which is the case of comparing groups 1 and 2. However, as mentioned previously, we can simply extract the domain-related users first and make it become the case of comparing group 5 and 6, which provide an even better AUC.

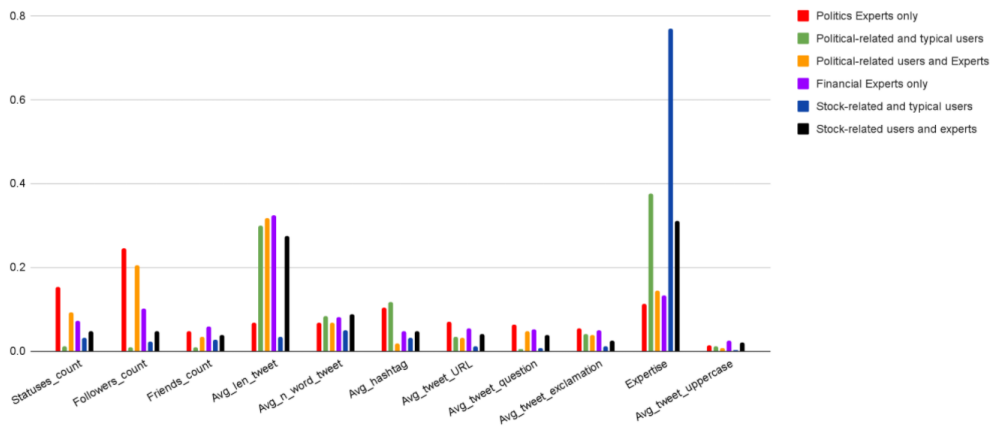

Fig. 13. Feature importance derived from random forests algorithm for all the comparison in Table 4.

Table 5. Top 3 trust attributes and the corresponding importance values derived from random forests algorithm.

	Attribute 1	Importance	Attribute 2	Importance	Attribute 3	Importance
Political Experts & others	Followers _ count	0.245	Statuses _count	0.1543	Expertise	0.1117
Financial Experts & others	Avg_len _tweet	0.3249	Expertise	0.1321	Followers _ count	0.1024
Political-related users & typical users	Expertise	0.3751	Avg_len _tweet	0.2996	Avg _hashtag	0.1183
Stock-related users & typical users	Expertise	0.77	avg_n_word _tweet	0.0493	Avg_len _tweet	0.0333
Political-related users & political experts	Avg_len _tweet	0.3168	Followers _count	0.2049	Expertise	0.1434
Stock-related users & financial experts	Expertise	0.3107	Avg_len _tweet	0.2745	avg_n _word_tweet	0.089

Table 6. The accuracy and AUC derived from random forest classification. The number of estimators is 50. For each target domain, there are three comparing groups: Experts & other users, domain_related users & typical users and domain-related users and experts.

Comparing groups	Accuracy	AUC
1. Political Experts & others	0.9999	0.7222
2. Financial Experts & others	0.9998	0.65
3. Political-related users& typical users	0.9878	0.9865
4. Stock-related users& typical users	0.9772	0.9188
5. Political-related users& political experts	0.9998	0.8462
6. Stock-related users& financial experts	0.9967	0.75

4.4 Recommendations on how to improve users' trustworthiness

Based on all the analysis and observation above, we would like to provide some recommendations for individuals who want to improve their online trustworthiness. First, the easiest way is remodeling the text style. As we discovered,

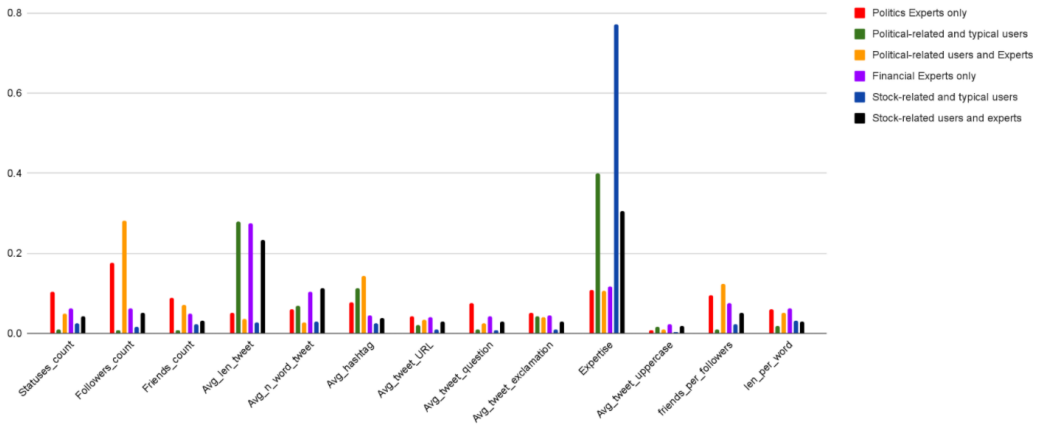

Fig. 14. Feature importance derived from random forests algorithm for all the comparison in Table 6.

Table 7. Top 3 trust attributes and the corresponding importance values derived from random forests algorithm after adding new attributes (friends_per_followers and len_per_word).

	Attribute 1	Importance	Attribute 2	Importance	Attribute 3	Importance
Political Experts & others	Followers _count	0.1767	Expertise	0.1077	Statuses_count	0.1041
Financial Experts & others	Avg_len _tweet	0.2739	Expertise	0.1181	Avg_n _word_tweet	0.1031
Political-related users & typical users	Expertise	0.3994	Avg_len _tweet	0.2794	Avg_hashtag	0.1129
Stock-related users & typical users	Expertise	0.7715	len_per _word	0.0309	Avg_n _word_tweet	0.0301
Political-related users & political experts	Followers _count	0.2813	Avg_hashtag	0.1423	Friends_per _follower	0.1237
Stock-related users & financial experts	Expertise	0.3048	Avg_len _tweet	0.234	Avg_n _word_tweet	0.1134

there exists peaks in the distribution of *avg_len_tweet* and *avg_n_word_tweet*, where most experts have around 150 characters and 20-25 words per tweet. Too many or too few may lose audiences' attraction or not be able to express important opinions. Second, the expertise score is important, but too high may also harm a user's reputation. Many experts' Expertise scores range from slightly above 0 to 0.1, which means they only posted one domain-related content in every 10 or more tweets. With that being said, quantity is not the issue, quality is. The last recommendation is to increase the *followers_count*. This suggestion seems to be meaningless. After all, in the world of social media, more followers inevitably means one obtains more reputation and credibility. Nevertheless, the most robust way is always unpretentious. Only when we can persevere in one specific goal can the accomplishment be built up. We hope by following the recommendations, a user can efficiently improve trustworthiness and at the same time, contribute great quality information to the online community.

5 CONCLUSION

This work concentrated on developing and evaluating trust filters, measurable trust attributes of social media users that may be able to predict their level of trustworthiness. In this paper, we looked into two target domains, politics and finance, on Twitter. We split social media users into three different groups: typical, domain-related and expert users. The list of experts was distilled from reputable online sources outside social media. A list of trust attributes were selected, shown to be effective by previous work, as proposed trust filters. Furthermore, some previously established trust filters were selected. We measured the value of these established and proposed trust filters for Twitter users and generated the distribution of each trust filter value for each of the three user types. We applied the random forest classification algorithm to gauge the effectiveness of trust filters in classifying user types. We found that some trust filters (Expertise, the number of followers, the average length of tweet, and the total number of tweets) are more important than others in the classification of both target domains. Furthermore, we find that other trust filters are effective distinguisher for one target domain but not the other: for example, the average number of hashtags in the finance domain. Our work paves the way for improving the quality of information extracted from social media by focusing on users' trustworthiness and increases the reliability and utility of this data to explain phenomena, help with decision making, and even predict trends.

REFERENCES

[1] Richard Adams. 2020. *Top 50 US politics Twitter accounts to follow.* Retrieved September 3rd, 2020 from https://www.theguardian.com/world/richard-adams-blog/2010/oct/05/top-50-twitter-accounts-us-politics-election

[2] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 US Presidential election online discussion. *First monday* 21, 11-7 (2016).

[3] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of computational science* 2, 1 (2011), 1–8.

[4] Alexandre Bovet, Flaviano Morone, and Hernán A Makse. 2018. Validation of Twitter opinion trends with national polling aggregates: Hillary Clinton vs Donald Trump. *Scientific reports* 8, 1 (2018), 1–16.

[5] Pete Burnap, Rachel Gibson, Luke Sloan, Rosalynd Southern, and Matthew Williams. 2016. 140 characters to victory?: Using Twitter to predict the UK 2015 General Election. *Electoral Studies* 41 (2016), 230–233.

[6] Dylan Byers. 2015. *Twitter's most influential political journalists.* Retrieved September 3rd, 2020 from https://www.politico.com/blogs/media/2015/04/twitters-most-influential-political-journalists-205510

[7] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information Credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (Hyderabad, India) *(WWW '11)*. ACM, New York, NY, USA, 675–684. https://doi.org/10.1145/1963405.1963500

[8] Jessica Elan Chung and Eni Mustafaraj. 2011. Can collective sentiment expressed on twitter predict political elections?. In *Twenty-Fifth AAAI Conference on Artificial Intelligence.*

[9] Simon Constable. 2015. *"Must-Follow" Twitter Feeds On Markets And Economics.* Retrieved November 7, 2019 from https://www.forbes.com/sites/simonconstable/2015/08/14/must-follow-twitter-feeds-on-markets-and-economics/#46f90dc0b09e

[10] Mathew Cruz. 2020. *5 Journalists to Follow on Twitter: Politics & News Edition.* Retrieved September 3rd, 2020 from https://onepitch.co/blog/5-journalists-to-follow-on-twitter-news-and-politics-edition/

[11] Jared Cummans. 2015. *100 Insightful Futures Traders Worth Following on Twitter.* Retrieved November 7, 2019 from https://commodityhq.com/investor-resources/100-insightful-futures-traders-worth-following-on-twitter/

[12] Thomas Dimpfl and Stephan Jank. 2016. Can internet search queries help to predict stock market volatility? *European Financial Management* 22, 2 (2016), 171–192.

[13] Joe Donatelli. 2017. *The 25 Must-Follow Twitter Accounts for The Anti-Trump Resistance.* Retrieved September 3rd, 2020 from https://www.lamag.com/culturefiles/social-media-twitter-resistance/

[14] Stuart Emmrich. 2020. *The 9 Twitter Accounts to Follow on Election Day.* Retrieved September 3rd, 2020 from https://www.vogue.com/article/the-nine-best-twitter-accounts-to-follow-on-election-day

[15] Gunn Enli. 2017. Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. *European journal of communication* 32, 1 (2017), 50–61.

[16] Daniel Gayo-Avello. 2012. " I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper"–A Balanced Survey on Election Prediction using Twitter Data. *arXiv preprint arXiv:1204.6441* (2012).

[17] Daniel Gayo-Avello. 2012. No, you cannot predict elections with Twitter. *IEEE Internet Computing* 16, 6 (2012), 91–94.

[18] Yevgeniy Golovchenko, Cody Buntain, Gregory Eady, Megan A Brown, and Joshua A Tucker. 2020. Cross-platform state propaganda: Russian trolls on Twitter and YouTube during the 2016 US presidential election. *The International Journal of Press/Politics* 25, 3 (2020), 357–389.

[19] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363, 6425 (2019), 374–378.

[20] Manish Gupta, Peixiang Zhao, and Jiawei Han. 2012. Evaluating event credibility on twitter. In *Proceedings of the 2012 SIAM International Conference on Data Mining*. SIAM, 153–164.

[21] Marcus Hawkins. 2020. *The Top 15 Conservatives to Follow on Twitter.* Retrieved September 3rd, 2020 from https://www.thoughtco.com/the-top-conservatives-to-follow-on-twitter-3303615

[22] Teng-Chieh Huang, Razieh Nokhbeh Zaeem, and K. Suzanne Barber. 2019. It Is an Equal Failing to Trust Everybody and to Trust Nobody: Stock Price Prediction Using Trust Filters and Enhanced User Sentiment on Twitter. 19, 4 (2019). https://doi.org/10.1145/3338855

[23] Teng-Chieh Huang, Razieh Nokhbeh Zaeem, and K Suzanne Barber. 2021. Finding Trustworthy Users: Twitter Sentiment Towards US Presidential Candidates in 2016 and 2020. In *Proceedings of SAI Intelligent Systems Conference*. Springer, 804–821.

[24] Teng-Chieh Huang, Razieh Nokhbeh Zaeem, and K Suzanne Barber. 2021. Identifying Real-world Credible Experts in the Financial Domain. *Digital Threats: Research and Practice* 2, 2 (2021), 1–14.

[25] Mochamad Ibrahim, Omar Abdillah, Alfan F Wicaksono, and Mirna Adriani. 2015. Buzzer detection and sentiment analysis for predicting presidential election results in a twitter nation. In *2015 IEEE international conference on data mining workshop (ICDMW)*. IEEE, 1348–1353.

[26] Peiran Jiao, André Veiga, and Ansgar Walther. 2020. Social media, news media and the stock market. *Journal of Economic Behavior & Organization* 176 (2020), 63–90. https://doi.org/10.1016/j.jebo.2020.03.002

[27] Chao Liang, Linchun Tang, Yan Li, and Yu Wei. 2020. Which sentiment index is more informative to forecast stock market volatility? Evidence from China. *International Review of Financial Analysis* 71 (2020), 101552.

[28] Guangyu Lin, Razieh Nokhbeh Zaeem, Haowei Sun, and K Suzanne Barber. 2017. Trust filter for disease surveillance: Identity. In *2017 Intelligent Systems Conference (IntelliSys)*. IEEE, 1059–1066.

[29] Linette Lopez and Lucinda Shen. 2015. *The 129 finance people you have to follow on Twitter.* Retrieved November 7, 2019 from https://www.businessinsider.com/117-finance-people-to-follow-on-twitter-2014-9

[30] Panagiotis T Metaxas, Eni Mustafaraj, and Dani Gayo-Avello. 2011. How (not) to predict elections. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. IEEE, 165–171.

[31] Robert S Mueller and Man With A. Cat. 2019. Report on the investigation into Russian interference in the 2016 presidential election.

[32] Arman Khadjeh Nassirtoussi, Saeed Aghabozorgi, Teh Ying Wah, and David Chek Ling Ngo. 2014. Text mining for market prediction: A systematic review. *Expert Systems with Applications* 41, 16 (2014), 7653–7670.

[33] Thien Hai Nguyen, Kiyoaki Shirai, and Julien Velcin. 2015. Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications* 42, 24 (2015), 9603–9611.

[34] Nuno Oliveira, Paulo Cortez, and Nelson Areal. 2017. The impact of microblogging data for stock market prediction: using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications* 73 (2017), 125–144.

[35] Jeff John Roberts. 2020. *14 Twitter and Instagram accounts you should follow for election news Day.* Retrieved September 3rd, 2020 from https://fortune.com/2020/11/03/election-night-2020-who-to-follow-twitter-instagram-accounts-follow-updates-trump-biden/

[36] Eduardo J. Ruiz, Vagelis Hristidis, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. 2012. Correlating Financial Time Series with Micro-blogging Activity. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (Seattle, Washington, USA) *(WSDM '12)*. ACM, New York, NY, USA, 513–522. https://doi.org/10.1145/2124295.2124358

[37] Parul Sharma and Teng-Sheng Moh. 2016. Prediction of Indian election using sentiment analysis on Hindi Twitter. In *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, 1966–1971.

[38] Juan M Soler, Fernando Cuartero, and Manuel Roblizo. 2012. Twitter as a tool for predicting elections results. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 1194–1200.

[39] Andrew Sun, Michael Lachanski, and Frank J Fabozzi. 2016. Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis* 48 (2016), 272–281.

[40] Adam Tsakalidis, Symeon Papadopoulos, Alexandra I Cristea, and Yiannis Kompatsiaris. 2015. Predicting elections for multiple countries using Twitter and polls. *IEEE Intelligent Systems* 30, 2 (2015), 10–17.

[41] Andranik Tumasjan, Timm Oliver Sprenger, Philipp G Sandner, and Isabell M Welpe. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM* 10, 1 (2010), 178–185.

[42] Lei Wang and John Q Gan. 2017. Prediction of the 2017 French election based on Twitter data analysis. In *2017 9th Computer Science and Electronic Engineering (CEEC)*. IEEE, 89–93.

[43] Razieh Nokhbeh Zaeem, David Liau, and K Suzanne Barber. 2018. Predicting Disease Outbreaks Using Social Media: Finding Trustworthy Users. In *Proceedings of the Future Technologies Conference*. Springer, 369–384.

WWW.IDENTITY.UTEXAS.EDU