International Journal of Data Science and Analytics
https://doi.org/10.1007/s41060-021-00291-z

REGULAR PAPER

Characterizing and predicting fake news spreaders in social networks

Anu Shrestha[1] · Francesca Spezzano[1]

Received: 24 April 2021 / Accepted: 28 October 2021

Abstract
Due to its rapid spread over social media and the societal threat of changing public opinion, fake news has gained massive attention. Users' role in disseminating fake news has become inevitable with the increase in popularity of social media for daily news diet. People in social media actively participate in the creation and propagation of news, favoring the proliferation of fake news intentionally or unintentionally. Thus, it is necessary to identify the users who tend to share fake news to mitigate the rampant dissemination of fake news over social media. In this article, we perform a comprehensive analysis on two different datasets collected from Twitter and investigate the patterns of user characteristics in social media in the presence of misinformation. Specifically, we study the correlation between the user characteristics and their likelihood of being fake news spreaders and demonstrate the potential of the proposed features in identifying fake news spreaders. Our proposed approach achieves an average precision ranging between 0.80 and 0.99 on the considered datasets, consistently outperforming baseline models. Furthermore, we also show that the user personality traits, emotions, and writing style are strong predictors of fake news spreaders.

Keywords Misinformation · Fake news spreaders · User characterization · User classification

1 Introduction

Online social media platforms such as Facebook and Twitter have drastically changed the landscape of news consumption and the pattern of information flow in the past decade. The majority of the population relies on social media for news on important events, breaking news, and emergencies. According to Pew Research Center 71% of American adults ever get news through social media in 2020 [38]. With the increase in its popularity, social media has significantly transformed the way of creating news content, user interactions, and engagement, reshaping the traditional medium to whole new information ecosystems [39]. Individuals in social media actively participate in creating and sharing news items due to its ease of use, lower cost, and convenience of further sharing [4,45]. This shift of the news paradigm has led to an unprecedented transformation in both news quality and quantity that users encounter in social media, reducing the credibility of news articles and eventually fostering the production and dissemination of misinformation.

Indeed, the rapid spread of fake news has become a concerning problem in online social networks in recent years. Research has found that fake news is more likely to go viral than real news, spreading both faster and wider [51] and people engage more with fake news than real news [47]. Moreover, the worrisome amount of fake news widely spreading over social media can negatively influence users' opinions creating threats on public health [50], emergency management and response [16,49], election outcomes [22], and is responsible for a general decline in trust that citizens of democratic societies have for online platforms [1]. Surprisingly, bots are equally responsible for spreading real and fake news, and human activity causes the considerable spread of fake news on Twitter [46,51] as people are generally not able to accurately identify which news item is fake and which is real [48]. Thus, fake news is successful mainly because people are not able to disguise it from truthful information [21,28] and often share news online without even reading its content [11]. Also, even if people recognize news as fake, they are more likely to share it if they have seen it repeatedly than the news that is novel [10].

✉ Francesca Spezzano
francescaspezzano@boisestate.edu

Anu Shrestha
anushrestha@u.boisestate.edu

[1] Computer Science Department, Boise State University, Boise, ID, USA

Published online: 22 November 2021

Thus, identifying fake news spreaders in social networks is one of the key aspects to mitigate misinformation spread effectively. Examples of strategies that could be implemented include assisting fake news spreaders with credibility indicators to lower their fake news sharing intent [55], and mitigation campaign, e.g., target the most influential real news spreader to maximize the spread of real news [44]. However, less is known about the characteristics of fake or real news spreaders.

Therefore, in this article, we seek to understand the characteristics of fake news spreaders focusing on different attributes such as user writing style, emotions, demographics, personality, social media behavior, and network features. In particular, we leveraged these attributes to perform a comprehensive analysis on two different datasets, namely PolitiFact [44] and PAN [34] to investigate the patterns of user characteristics in social media in the presence of misinformation. We hypothesize that users likely to share fake news hold specific patterns based on these attributes which are different from real news spreaders. To the best of our knowledge, some of the features we considered, such as user stress, needs, values, and tweeting behavior, have not been analyzed before. Furthermore, we investigate to what extent these features can be used to identify users who are likely to share fake news by addressing the problem as a binary classification task.

Our analysis unveils some interesting characteristics of fake news spreaders across the two datasets considered. Specifically, our results show that:

- The majority of users under 18 or over 40 may tend to share more fake than real news.
- Female users may tend to be more fake news spreaders than male users.
- The political orientation of a fake news spreader is more likely to coincide with the source's political bias of the majority of circulating fake news items.
- Fake news spreaders (1) have newer accounts, (2) spend, on average, less time between two consecutive tweets, and (3) tend to tweet more at night.
- Fake news spreaders tend to express more negative emotions and stress in their tweets than real news spreaders.
- Fake news spreaders are estimated to be more extroverted and less neurotic than real news spreaders.
- Classification results using our proposed features outperform the results of baseline approaches with n-grams in both datasets. Specifically, we show that our proposed features can identify fake news spreaders with an average precision of 0.99 on the PolitiFact dataset (vs. 0.96 achieved by the best baseline) and 0.79 on the PAN dataset (vs. 0.78 achieved by the best baseline).
- Emotions and personality features are strong predictors of fake news spreaders in all the considered datasets.

The article is organized as follows. Section 2 summarizes related work, Sect. 3 describes the dataset we used in this article, Sect. 4 presents our proposed features to characterize and classify fake news spreaders, Sect. 5 presents the user characteristics patterns that we found by analyzing the considered datasets, Sect. 6 reports on our experimental evaluations, and finally, conclusions are drawn in Sect. 7.

2 Related work

Several studies have been conducted to understand the characteristics of users that are likely to contribute to spreading fake news on social networks. Vosoughi et al. [51] revealed that the fake news spreaders had, on average, significantly fewer followers, followed significantly fewer people, and were significantly less active on Twitter. Moreover, bots tend to spread both real and fake news, and the considerable spread of fake news on Twitter is caused by human activity. Shrestha and Spezzano showed that social network properties help in identifying active fake news spreaders [41]. Shu et al. [46] analyzed user profiles to understand the characteristics of users that are likely to trust/distrust fake news. They found that, on average, users who share fake news tend to be registered for a shorter time than the ones who share real news and that bots are more likely to post a piece of fake news than a real one, even though users who spread fake news are still more likely to be humans than bots. They also show that real news spreaders are more likely to be more popular and that older people and females are more likely to spread fake news. Guess et al. [15] also analyzed user demographics as predictors of fake news sharing on Facebook and found out political-orientation, age, and social media usage to be the most relevant. Specifically, people are more likely to share articles they agree with (e.g., right-leaning people tended to share more fake news because the majority of the fake news considered in the study were from 2016 and pro-Trump); seniors tend to share more fake news probably because they lack digital media literacy skills that are necessary to assess online news truthfulness. The more people post on social media, the less they are likely to share fake news, most likely because they are familiar with the platform and they know what they share.

Shrestha et al. [43] analyzed the linguistic patterns used by a user in their tweets and personality traits as a predictor for identifying users who tend to share fake news on Twitter data [35,43]. Likewise, Giachanou et al. [13] proposed an approach based on a convolutional neural network to process the user Twitter feed in combination with features representing user personality traits and linguistic patterns used in their tweets to address the problem of discriminating between fake news spreaders and fact-checkers.

Ma et al. [25] went beyond the user and news characteristics and analyzed the characteristics of diffusion networks to explain users' news sharing behavior. They found opinion leadership, news preference, and tie strength to be the most important factors at predicting news sharing, while homophily hampered news sharing in users' local networks. Also, people who are driven by gratifications of information seeking, socializing, and status-seeking were more likely to share news on social media platforms [23].

Moreover, creating hashtags has been widely used to organize campaigns, sharing information and opinion about events and news stories on social media. These hashtags have also been used to draw attention and enhance the topic's visibility, eventually causing its wide spread over social media. Both individuals and news organizations have capitalized on this feature of social media via the massive use of political hashtags to increase readership and user engagement [3]. This target turns true and amplifies if a user shares the piece of news with partisan affiliation [37]. Thus, the political orientation of a user can provide additional cues about the user being a fake news spreader or not.

As compared to previous work, which has been mainly done on the PAN 2020 dataset [34], *this article addresses the problem of characterizing and predicting users that are keen to spread fake news on an additional larger dataset with more reliable ground truth extracted from FakeNewsNet* [44]. *We consider several groups of topic-agnostic features, including new features that have not been used in previous work, such as behavioral features, stress, needs, and values, to profile and predict fake news spreaders on two datasets and evaluate the relative importance of the considered groups of features. We also highlight feature patterns that are common to both datasets.*

3 Datasets

This section describes the datasets we used to carry out our experiments, namely the PAN 2020 and PolitiFact (FakeNewsNet) datasets. The size of these datasets is shown in Table 1.

Table 1 Datasets and statistics

Dataset	# Fake news spreader	# Real news spreader
PAN 2020	250	250
FakeNewsNet (PolitiFact)	648	398

3.1 PAN 2020 dataset

The first dataset we consider is the one provided by the PAN CLEF[1] 2020 shared task on profiling fake news spreaders on Twitter [34]. The dataset has been collected in two languages, namely English and Spanish, and consists of a balanced train and test set for each language. For each considered language, the training set includes 300 Twitter users and 100 tweets for each user from their Twitter feed, resulting in 30,000 English tweets and Spanish 30,000 tweets. The test set contains 200 users in each language and 100 tweets from their feed for each user, resulting in 20,000 English tweets and 20,000 Spanish tweets. In this article, we have considered only the English dataset and combined the train and test set together in a unique (balanced) dataset.

In the PAN 2020 dataset, users that shared fake news in the past are labeled as fake news spreaders and real news spreaders, otherwise. However, it is worth noting that, because the dataset is GDPR compliant,[2] users are labeled as "class 0" or "class 1," and the authors of the dataset did not disclose which one of the two labels corresponds to the class of users who are fake news spreaders. In this article, we assumed one of the two labels to identify fake news spreaders according to feature patterns that result similar to the ones of the PolitiFact dataset. These patterns are described in Sect. 5.

3.2 PolitiFact (FakeNewsNet)

The FakeNewsNet dataset consists of two datasets, PolitiFact and GossipCop, from two different domains, i.e., politics and entertainment gossip, respectively [44]. Each of these datasets contains details about news content, publisher, social engagement information, and user social network. In this article, we only used the PolitiFact dataset, which contains news with known ground truth labels collected from the fact-checking website PolitiFact[3] where journalists and domain experts fact-checked the news items as fake or real. We decided not to use the GossipCop dataset because in our previous work [42] we found that gossip news is quite different than political news; hence we focused our attention on the same news domain as the other dataset we considered, i.e., the PAN 2020 dataset. Overall, the considered PolitiFact dataset contains 295,469 users (after removing self-claimed bot accounts) sharing 701 news items via tweets and retweets. As this dataset only provides ground truth for news, we computed the labels for the users (fake news spreader or real news spreader) as explained here below. First, we filtered out those

[1] PAN CLEF (https://pan.webis.de/) is a well-known forum that focuses on applying text mining for user profiling.

[2] https://www.privacyshield.gov/article?id=European-Union-Data-Privatization-and-Protection.

[3] https://www.politifact.com/.

users who had shared the same news item multiple times, and then we selected only those users who had shared at least eight unique news items. We manually analyzed the profiles of users who shared the same news item multiple times and found that they were bots; hence we excluded them from our analysis as research has shown that false news spreads more than the truth because of humans, not bots [51]. Next, the resulting group of 1,046 users is labeled as fake news spreaders or real news spreaders as follows: (1) a user is a fake news spreader if at least 60% of the news items they shared are fake, or (2) a user is a real news spreader if at least 60% of the news items they shared are real.[4] We labeled 648 users as fake news spreaders and 398 as real news spreaders. Moreover, we retrieved additional user data as follows. For each user who did not have enough tweets, i.e., more than 100 words among all their tweets combined, we crawled all tweets posted one month prior to his first tweet creation time in our dataset. These additional tweets were utilized to generate personality features and political orientation.

4 Features

This section describes the features we analyzed to characterize and classify fake news spreaders in the two datasets considered. Specifically, we study users according to six user features groups: demographics, Twitter behavior and network, emotions, personality, readability, and writing style. Text-based features such as emotions, personality, and readability are computed on the document resulting from the concatenation of all the user tweets. To have a more accurate estimation of user emotions, personality, readability, and writing style, retweets are excluded when computing these features.

4.1 Demographics

The first group of features we consider deals with user demographics, including age, gender, and political orientation. Previous work has shown how these features influence users' news-sharing behavior. For instance, Reis et al. [36] show that white and male users potentially share more news on Twitter. Differently, Shu et al. [46] analyzed user profiles to understand the characteristics of users that are likely to trust/distrust fake news and propagate them on Twitter. They also show that older people and females are more likely to spread fake news.

Demographic features are often not explicitly available on social media platforms. Therefore, as detailed in the following, we used machine learning -based methods to infer such attributes in the PolitiFact dataset users. However, as the required metadata and hashtags are not available for the PAN 2020 dataset, we were not able to compute demographics for this dataset.

4.1.1 Age and gender

We utilized m3inference [53], a deep-learning-based system trained on Twitter data, to infer user demographic characteristics. Based on the available metadata such as username, screen name, description, and profile image, it predicts the *gender* of the user as male or female, *age* of the user grouped in four categories ($\leq$18, 19$-$29, 30$-$39 and $\geq$40) and whether the given account is handled by an *organization* or not. We utilized only two characteristics (age and gender) for both types of users for our analysis. The m3inference has been shown to have an F1 score of 0.918 for gender prediction and 0.522 for age prediction [53].

4.1.2 Political orientation

As the political ideology can provide additional cues about profiling fake news spreaders, we computed a polarization score to identify their political leaning. We used the method defined by Hemphill et al. [18] where a polarization score (#polar score) for each user is defined by using the hashtags from the user tweets to estimate their political ideology. Each of those hashtags is scored according to how political figures with known party affiliation use them. Specifically, we implemented the #polar score as follows. As a political figure dataset, we used the dataset provided by Chamberlain et al. [6] which contains tweets collected from Jan 04, 2007, to Jan 03, 2019, and authored by in-office U.S. Congress members during that time period [6]. Then, we classified each politician as Republican or Democrat by using a TF-IDF vector representation of their tweet hashtags as input features to a binary classifier. We experimented with different classifiers, including support vector machine, logistic regression, extra tree classifier, and random forest with 5-fold stratified cross-validation using class weighting to deal with class imbalance. Random Forest resulted in being the best classifier with 0.69 AUROC and 0.67 average precision, providing us with good confidence in using those hashtags to estimate user political orientation.

Then, we generated Chi-Squared scores for each hashtag, and we leveraged these scores as a polar dictionary to assign polarization scores to the users in the dataset we considered (i.e., PolitiFact). Each hashtag in the tweet is looked up in the polar dictionary, and Chi-Squared scores of matching hashtags are averaged across the entire hashtags included in user tweets defined as polarization score for that user. A positive polarization score indicates that the user tends to

[4] One limitation of this labeling approach is that we may not catch fake news spreaders who camouflage themselves as real news spreaders through their news sharing behavior.

incline toward right-leaning political orientation, and a negative score indicates left-leaning political orientation.

4.2 Behavioral-based features

This group of features measures the tweeting/sharing behavior and engagement of the users and consists of the following features:

Insomnia index We analyzed the user tweeting behavior within the day (24 h). We divided the time into day and night and considered the 'night' window as '9PM-6AM' and the 'day' window as '6:01AM-8:59PM' (we used the local time of the user), and analyzed the normalized difference between the number of tweets shared during these time windows for each user as in [9,40].
Weekend index Similarly to the insomnia index, we computed the normalized difference in the number of tweets on weekdays and weekends.
Time elapsed Average time elapsed between two consecutive tweets of the user.
Account duration The duration (in the number of days) of the account since it is registered.

4.3 Network-based features

Vosoughi et al. [51] have shown that fake news spreaders had fewer followers and followed fewer people than real news spreaders. Thus, in this article, we computed the Twitter follower to following (TFF) ratio as in [46] to measure user connectivity in the Twitter social network. TFF is computed by using the following formula

$$\text{TFF} = \frac{\#\text{Follower} + 1}{\#\text{Following} + 1}$$

which indicates the ratio of the number of followers to the number of followings of the user. The greater the ratio, the higher the popularity of the user.

4.4 Emotions

Fake news is deliberately induced with emotionally charged words to influence public opinion and affects the vulnerabilities of people by triggering their sentiments such as anger, fear, and distrust toward the event, person, and organization. Moreover, Ghanem et al. [12] recently showed emotions play a key role in detecting fake news. Therefore, we computed emotion features such as anger, joy, sadness, fear, disgust, anticipation, surprise, and trust by using the Emotion Intensity Lexicon (NRC-EIL) [29] and happy, sad, angry, don't care, inspired, afraid, amused, and annoyed using Emolex.[5] We started by cleaning tweets by expanding contraction words, correcting misspellings and grammatical mistakes using LanguageTool,[6] replacing negated words with their WordNet antonym, removing stop words, and lemmatizing the words. Next, we computed feature vectors using the approaches proposed by Milton et al. [26,27]. Specifically, each word is looked up in both emotion dictionaries, and the associated affect values of matching words are extracted. Next, we normalized the scores of each emotion category by the total number of emotions retrieved from a tweet to generate an emotion vector. In case the same emotion was present in both lexicons, e.g., sad in NRC-EIL and sadness in Emolex, we considered the average of the two computed values.

4.4.1 Stress

Along with these emotions (i.e., positive and negative emotions), frustrations, worries, and irritations, which are the characteristics of stress expressed through the language used in the user feed, can also progressively accelerate the spread of fake news. Thus, we incorporate a stress feature computed using the lexical dictionary, a Stress Word Count Dictionary created by Wang et al. [52] as the LIWC tool lacks this category. To compute this feature, we concatenated all the tweets by each user to form a single document per user. We removed words like 'RT,' 'Via,' and '&amp' for each document.

4.5 Personality

The IBM Watson Personality Insights service uses linguistic analytics to infer individuals' intrinsic personality characteristics, including Big Five personality traits, Needs, and Values, from digital communications such as social media posts. The tool can work for different languages, including English and Spanish. In our case, we concatenated all the user tweets in a unique document to compute their personality characteristics.

The features computed by this service are detailed in the following (we considered the raw scores provided by the service):

Big five The Big Five personality traits, also known as the five-factor model (FFM) and the OCEAN model, are a widely used taxonomy to describe people's personality traits [30]. This taxonomy's five basic personality dimensions are openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. For each

[5] https://sites.google.com/site/emolexdata/.

[6] https://pypi.org/project/language-tool-python/.

personality dimension, IBM Watson Personality Insights also provides a set of additional six facet features. For instance, agreeableness' facets include altruism, cooperation, modesty, morality, sympathy, and trust.

Needs These features describe a user's needs as inferred by the text they wrote and include excitement, harmony, curiosity, ideal, closeness, self-expression, liberty, love, practicality, stability, challenge, and structure.

Values These features describe the motivating factors that influence a person's decision-making. They include self-transcendence, conservation, hedonism, self-enhancement, and openness to change.

These features ranges from 0 to 1. In terms of how precise is the IBM Watson Personality Insights service, the official documentation[7] reports an average Mean Absolute Error (MAE) for the English language of 0.12 for the Big Five dimensions, 0.12 for the Big Five facets, 0.11 for Needs, and 0.11 for Values. The reported average MAE scores are based on a dataset containing user Twitter feeds from between 1500 and 2000 participants for all characteristics and languages.

4.6 Readability

Readability measures the complexity of the text, and when computed from text written by the user (tweets in our case), it also represents which level of text complexity a user can understand. To determine that, we used popular readability measures in our analysis:

Flesh Reading Ease
Flesh Kincaid Grade Level
Coleman Liau Index
Gunning Fog Index
Simple Measure of Gobbledygook Index (SMOG)
Automatic Readability Index (ARI)
Lycee International Xavier Index (LIX)
Dale-Chall Score

Flesch scores range from 0 to 100. Higher scores of Flesch reading-ease indicate that the text is easier to read, and lower scores indicate difficulty to read. Coleman Liau Index depends on the characters of the word to measure the understandability of the text. The Gunning Fog Index (that generates grade level between 0 and 20), Automatic Readability Index, SMOG Index, Flesh Kincaid Grade Level are algorithmic heuristics used for estimating readability, that is, how many years of education is needed to understand the text. Finally, Dale-Chall's readability test uses a list of words well-known by the fourth-grade students (easily readable words) to determine the difficulty of the text. We use this group of 8 readability features to measure the complexity of a user's writing style.

4.7 Writing style

This set of features captures the writing style of the tweets authored by the same user. Specifically, we computed the average number of certain words, items, and characters per user tweet, which includes the average number of (1) words, (2) characters, (3) lowercase words, (4) uppercase words, (5) lowercase characters, (6) uppercase characters, (7) stop words, (8) punctuation symbols, (9) hashtags, (10) URLs, (11) mentions, and (12) emojis and smileys. Also, we considered the (13) percentage of user tweets that are a retweet and (14) the percentage of user tweets that are a sharing of breaking news; we considered a tweet sharing breaking news if the keyword 'breaking' or 'breaking news' was appearing in the tweet text. All but features (13) and (14) are computed by removing retweets from the user feed.

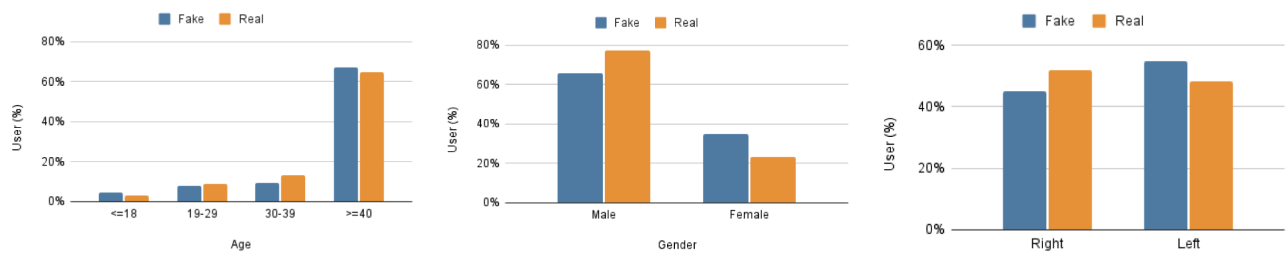
5 User characterization

This section presents the main patterns characterizing users who spread fake news that we found by analyzing the features described in the previous section on the two considered datasets. However, as the PAN 2020 dataset only provides 100 tweets per user, and user profile meta-data and timestamps are not included, and hashtags are blurred, we were not able to compute demographic, behavioral, and network features for this dataset. All the feature differences discussed in this section are statistically significant with a p value < 0.05 (ANOVA or Wilcoxon rank-sum according to the data distribution).

5.1 Demographics

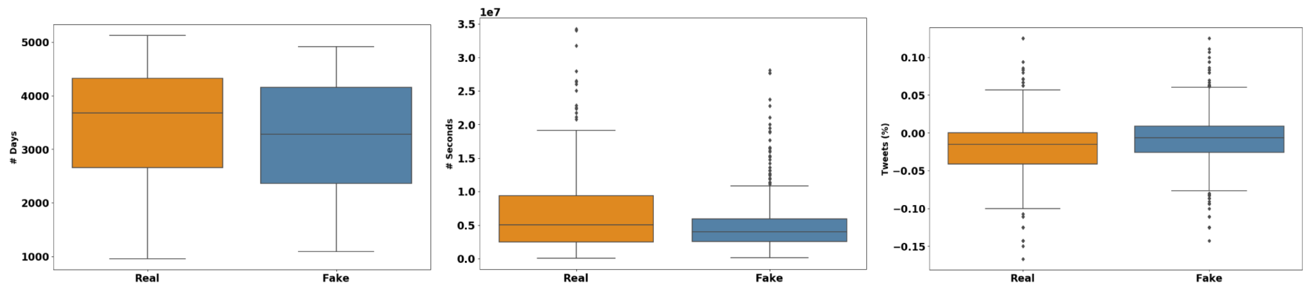
Demographics have been shown to be predictors of fake news spreaders [15]. Figure 1 shows the distribution of age, gender, and political orientation on the PolitiFact dataset. Here, we observe that among users who have been predicted to be under 18 or over 40, the majority of them tend to share more fake news than real one. The trend is the opposite for users whose age is predicted to be in the age range of 19–39. While previous work has shown that people over 65 tend to share more fake news than the younger generations (age range 18–64), the sharing behavior of users under 18 has not been investigated. Here we observe that these users, together with the ones over 40, may be the most vulnerable to fake news, which is somehow aligned with previous findings. The majority of teenagers are, in fact, unable to assess the credibility of the information that floods their devices [7,54], while

[7] https://cloud.ibm.com/docs/personality-insights?topic=personality-insights-science#researchPrecise.



(a) Age ($p \leq 0.003$) **(b)** Gender ($p \leq 0.001$) **(c)** Political Orientation ($p \leq 0.006$)

Fig. 1 Distribution of user demographics on the PolitiFact dataset



(a) Account duration from registered time. ($p \leq 0.001$) **(b)** Average time-lapse between two consecutive tweets. ($p \leq 0.001$) **(c)** Insomnia index (tweeting behavior at night). ($p \leq 0.001$)

Fig. 2 Boxplots of user behavioral features on the PolitiFact dataset

seniors are not as adept as the younger generation in assessing online news veracity [14]. Regarding the role of gender in user sharing behavior, we observe in Fig. 1b that users whose gender is predicted to be female tend to be more fake news spreaders than male users in the considered dataset. One possible explanation could be that female users may be less interested in political news and, consequently, less informed and then more vulnerable on these topics [2,33]. Even if the presented findings for age and gender seem to be somehow aligned with previous research, it is worth noting that these user attributes have been automatically inferred by using a tool whose accuracy is not perfect; hence some errors may have been introduced. Also, the age groups are highly unbalanced, and the last group (≤ 40) is very broad and diverse compared to the other ones. Hence our findings may not be general but just limited to the (not very large) considered dataset.

Figure 1c shows the distribution of fake and real news spreaders according to their political orientation. We see that, in the PolitiFact dataset, left-leaning users are more likely to be fake news spreaders than right-leaning users. Guess et al. [15] have shown that, in 2016, conservatives were more likely to share articles from pro-Trump fake news domains than liberals or moderates because those news items were aligned with their believes, and the majority of fake news items that were circulating were right-leaning. What we observe in the PolitiFact dataset is not in contradiction with this finding. To show that, we gathered the source bias of the news items present in this dataset from the MediaBias/FactCheck website[8] and found that the majority of these news items came from left-leaning sources and were tweeted much more than news coming from right-leaning sources (9,435 tweets about news from left-leaning sources vs. 3,408 tweets about news from right-leaning sources). Thus, we also observe in the PolitiFact dataset that the political orientation of a fake news spreader is more likely to coincide with the one of the sources of the majority of circulating fake news items (left-leaning in this case).

5.2 User behavior

The presence of timestamps in the PolitiFact dataset allows us to investigate fake news spreaders tweeting behavior. Figure 2 shows the box plots of the considered behavioral features on such dataset. Here, we observe that fake news spreaders (1) have newer accounts, (2) spend, on average,

[8] `mediabiasfactcheck.com`. The website's main goal is to educate the public on media bias and deceptive news practices. This website contains a comprehensive list of news sources, their bias, and their credibility of factual reporting scores. Here, the publisher's political bias is defined by using seven degrees of bias: *extreme-right, right, right-centered, neutral, left-centered, left, and extreme-left*.

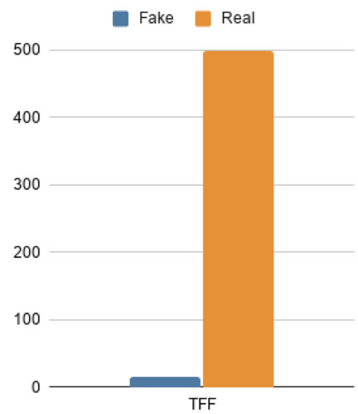


Fig. 3 Average Twitter follower to following (TFF) ratio on the PolitiFact dataset. The difference is statistically significant ($p \leq 0.001$)

less time between two consecutive tweets, and (3) tend to tweet more at night (higher insomnia index) than real news spreaders.

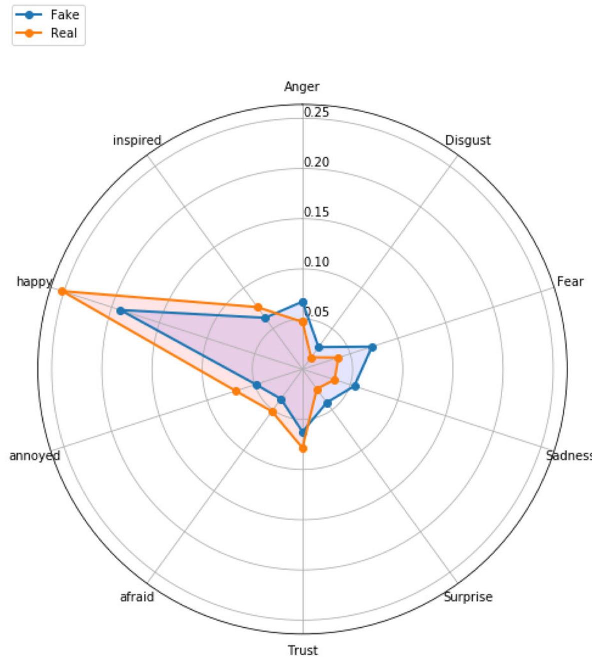
Thus, fake news spreaders are users who are newer to the platform (we are not considering bot accounts) and may be less expert about its functionalities/usage, and who tend to tweet more frequently, perhaps to increase their social capital. Also, their higher nighttime online activity may be connected with the presence of a higher stress condition for fake news spreaders, as shown in Sect. 5.4.

5.3 User network

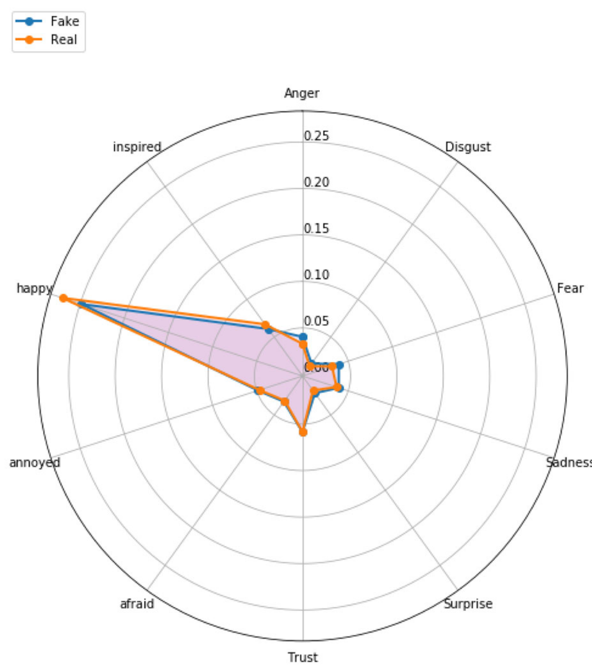
Figure 3 shows the distribution of the average Twitter follower to following (TFF) ratio on the PolitiFact dataset. We observe how non-fake news spreaders are much more popular (they have around 500 times more followers than following, on average) than fake news spreaders. Thus, users with lower TFF may tend to spread fake news more to increase their popularity on Twitter. For instance, users may know a news item is fake and spread it anyway because it is funny or of interest to user's friends and hence generate engagement among Twitter followers. Another motivation could be that a user with a low TFF is new to the platform and is not familiar with its features, hence may mistakenly share fake news.

5.4 User emotions

Figure 4 shows the radar charts of user emotions, while Fig. 5 shows a comparison of user stress levels on both the considered datasets. We notice that, in both cases, fake news spreaders tend to express more negative emotions (fear, sadness, disgust, and angry) and stress in their tweets than real news spreaders (all p values are ≤ 0.001). Conversely, non-fake news spreaders are happier and more inspired, but also



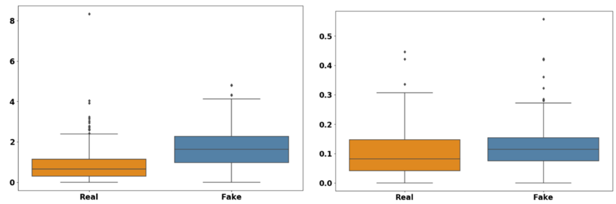
(a) PolitiFact



(b) PAN 2020

Fig. 4 Radar charts of the emotion features: PolitiFact and PAN 2020 datasets

more afraid (all p values are ≤ 0.001). Being induced by negative bias, people generally pay more attention to negative news [19,42]. Hence fake news spreaders tend to frame their tweets with negative emotions targeting to make it catchier and circulate more among people. On the contrary, non-fake



(a) PolitiFact ($p \leq 0.001$) **(b)** PAN 2020 ($p \leq 0.001$)

Fig. 5 Box plots of user stress level on the PolitiFact and PAN 2020 datasets

news spreaders are general individuals whose motive of using social media platforms is to connect with other people or family, share their achievements, advice, and support [31] and are more skeptical about sharing fake news.

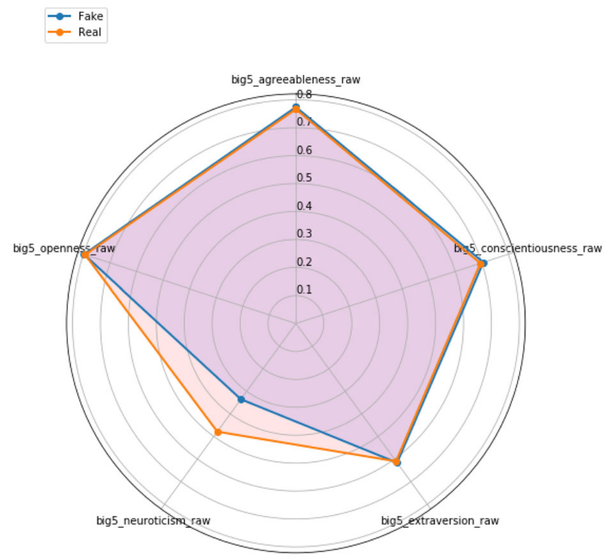
5.5 User personality traits

User Big Five personality traits are shown in Fig. 6 for both types of users. Among the five traits, extroversion and neuroticism are statistically significant features in both datasets (all p values are ≤ 0.001) and show the same trend, namely, fake news spreaders are estimated to be more extroverted and less neurotic than real news spreaders. Extroversion is related to the number of friends a user has, while neuroticism is related to frequency of posting [17]. Thus, fake news spreaders are estimated to be people who may share fake news to capture the interest of and make fun with their friends and/or possibly connect with more people. On the other end, sharing fake news is a rarer phenomenon as compared to real news sharing [15]; hence fake news spreaders are estimated to be less neurotic because they share less than real news spreaders.

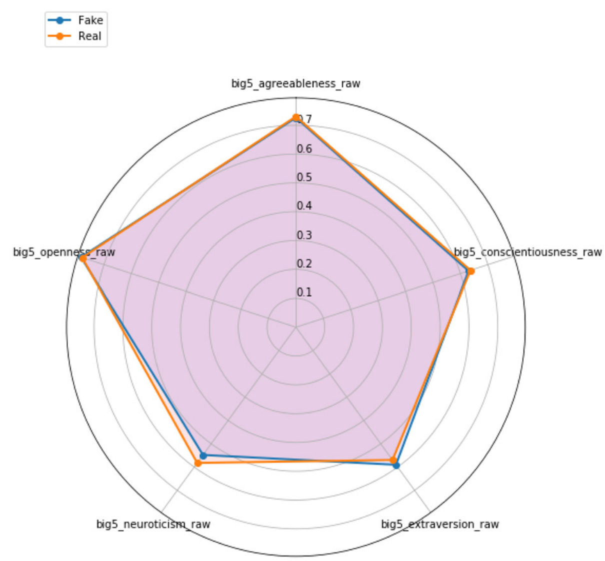
The other three personality traits are statistically significant features only in the PolitiFact dataset (all p values are ≤ 0.001), where we found that fake news spreaders are estimated to be more agreeable, conscious, and open than real news spreaders. Agreeableness is related to the type of feelings (positive or negative) expressed via social media updates, conscientiousness to posting about political news, and openness to the sharing of various forms of media [17]. Thus, fake news spreaders are estimated to be people whose posting behavior is driven by emotions (either positive or negative) and have more interest in political events.

5.6 User readability level and writing style

Different from emotional and personality features, readability features do not generalize across the considered datasets. In general, fake news spreaders in the PolitiFact datasets have a lower readability level than non-fake news spreaders, while the trend is the opposite in the PAN 2020 dataset.



(a) PolitiFact

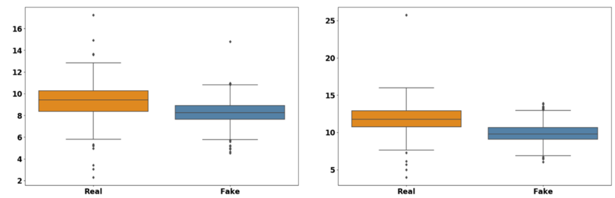


(b) PAN2020

Fig. 6 Radar charts of the Big-Five personality scores: PolitiFact and PAN 2020 datasets

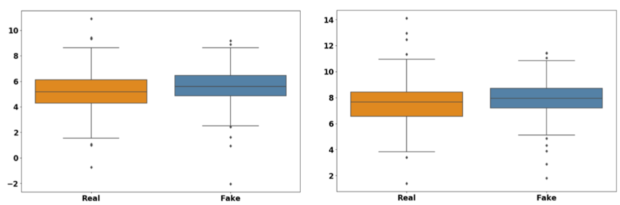
Figures 7 and 8 show the box plots of two of the readability measures we considered on the PolitiFact and PAN 2020 dataset, respectively.

Similarly, Table 2 highlights the pattern of writing style among fake news spreaders and real news spreaders. If the value of a feature was higher (on average) for real news spreaders as compared to fake news spreaders, it is denoted as $R > F$ (and $R < F$ otherwise) in the table. Fake news spreaders tend to use more uppercase characters and fewer hashtags in their tweets but share more breaking news than real news spreaders, and this trend generalizes for both datasets. Moreover, fake news spreaders in PolitiFact incor-



(a) Flesch kincaid ($p \leq 0.001$) **(b)** Gunning fog ($p \leq 0.001$)

Fig. 7 Readability index of tweets written by fake news spreaders versus real news spreaders in PolitiFact



(a) Flesch kincaid ($p = 0.009$) **(b)** Gunning fog ($p = 0.003$)

Fig. 8 Readability index of tweets written by fake news spreaders versus real news spreaders in PAN 2020

Table 2 Writing style features that differ in user feed

Features	PolitiFact	PAN 2020
Hashtags	$\boldsymbol{R > F}$	$\boldsymbol{R > F}$
Retweets		$R > F$
Char	$R > F$	$R < F$
Uppercase char	$\boldsymbol{R < F}$	$\boldsymbol{R < F}$
Lowercase char	$R > F$	$R < F$
Lowercase word	$R > F$	
Uppercase word	$R < F$	
Breaking	$\boldsymbol{R < F}$	$\boldsymbol{R < F}$
Emoji		$R > F$
Trailing period	$R > F$	
Punctuation	$R > F$	
Word count	$R > F$	
Stop words	$R > F$	
URLs	$R < F$	
Mentions	$R > F$	

All differences are statistically significant ($p \leq 0.002$ for PolitiFact and $p \leq 0.04$ for PAN 2020). Bold indicate the same pattern in both datasets

porate more uppercase words and URLs but fewer words, lowercased characters, punctuation, trailing periods ('...'), stop words, and mentions than real news spreaders in their tweets. In the PAN 2020 dataset, fake news spreaders use more lowercase characters, fewer emojis, and retweet less than real news spreaders. This indicates that fake news spreaders aim to gain people's attention by sharing breaking news and using more uppercase words and URLs in their tweets.

6 Experiments

This section reports on our experimental results of using the features described in Sect. 4 to automatically identifying fake news spreaders.

6.1 Experimental setting

We addressed the problem of automatically identifying fake news spreaders as a binary classification task. In particular, we used the combination of all the groups of features for the prediction. Once the features are computed, the classification is performed by using the best classifier among linear support vector machine (SVM), logistic regression, and random forest. We used class weighting to deal with the class imbalance and performed 5-fold cross-validation. Additionally, we also used each group of features as input to the best classifier to examine the contribution of these features in identifying a user likely to spread fake news. As an evaluation metric, we used Average Precision[9] (AvgP) which is a metric commonly used when dealing with unbalanced binary datasets [8], as in the case of the PolitiFact dataset. The average precision is the area under the precision curve, computed by plotting precision against the true positive rate. The average precision score gives the probability that a classifier will correctly identify a randomly selected positive sample (e.g., a fake news spreader in our case) as being positive. In our problem, we are interested in identifying fake news spreaders with high precision. These are the users to target with correction strategies to mitigate the further spread of fake news. In the tables reported in this section, the best average precision values are highlighted in bold.

6.1.1 Baselines for comparison

We compared our proposed approach with the two best performing approaches used by the participants to the PAN CLEF 2020 shared task, namely the approaches proposed by Buda and Bolonyai [5] and Pizarro [32]. These baselines are described here below:

Buda and Bolonyai [5] utilized n-grams based approach and combined them with statistical features from the tweets, such as their average length or their lexical diversity. Specifically, they used an ensemble model of

[9] We used the average precision implementation provided by the Python Scikit-learn library: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.average_precision_score.html.

Table 3 Average precision of our proposed features (in input to a Random Forest classifier) on PolitiFact and PAN 2020 datasets and comparison with baselines

Approach	PolitiFact	PAN 2020
Buda and Bolonyai [5]	0.737	0.783
Pizarro [32]	0.966	0.714
Our features (random forest)	**0.995**	**0.795**
Our features (linear SVM)	0.595	0.687
Our features (logistic regression)	0.672	0.717

Best values are in bold

Table 4 Average precision per feature group on PolitiFact and PAN 2020 datasets

Features	PolitiFact	PAN 2020
Demographics	0.777	–
Emotions	0.976	0.787
Behavior	0.866	–
Network	0.776	–
Readability	0.897	0.635
Personality	0.979	0.786
Writing style	0.990	0.713

Logistic Regression with five sub-models, namely, logistic regression, linear SVM, random forest, and XGBoost with n-grams and XGBoost with statistical features.
Pizarro [32] utilized a character and word n-grams-based approach with a linear support vector machine as the classifier.

6.2 Classification results

Classification results are reported in Table 3 to allow comparison between the performances of baselines and our method on both PolitiFact and PAN 2020 datasets. As we can see, our proposed features consistently outperform both baseline approaches. Specifically, we got an average precision of 0.995 versus the best baseline results of 0.966 achieved by Pizarro [32] on the PolitiFact dataset and an average precision of 0.795 versus the best baseline results of 0.783 achieved by Buda and Bolonyai [5] on the PAN 2020 dataset. Among the considered classifiers, random forest achieved the best performance. Furthermore, the baseline methods are mainly n-grams-based and, consequently, they are not easy to interpret. On the contrary, the features we consider in Sect. 4 achieve better performances and can also be analyzed to provide significant patterns to characterize fake news spreaders as we have shown in Sect. 5.

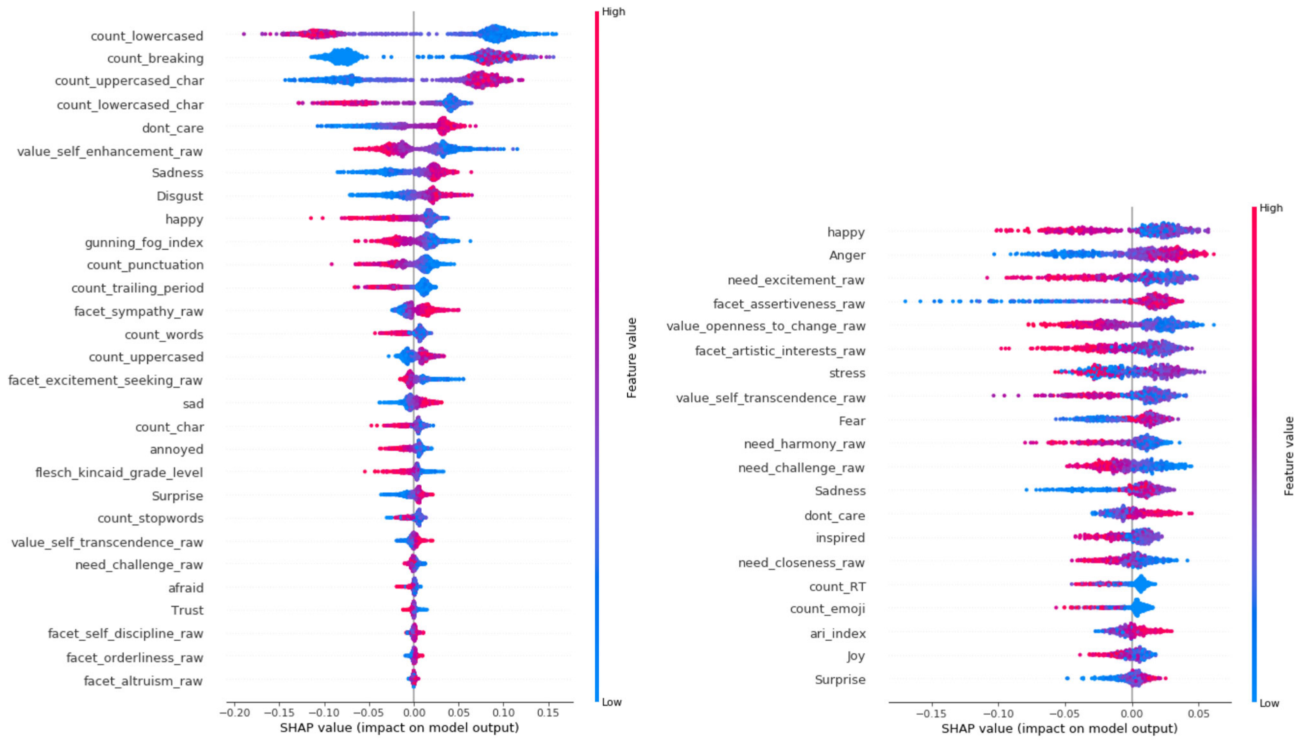
In addition, we investigated the performance of each considered group of features individually (demographics, emotions, behavior, network, readability, personality, and writing style) when the best classifier (i.e., random forest) is used. Results are reported in Table 4. We observe that emotions and personality features are the most important groups of features for the PAN 2020 dataset. In the PolitiFact dataset, the writing style is the most important group of features, while emotions and personality are the second most important groups of features. Hence, our results reveal that emotions and personality are strong predictors of fake news spreaders in both datasets. Since the Twitter IDs of the users in the PAN 2020 dataset are concealed, it was not possible for us to collect the additional user data required to generate some features like demographics, behavior, and network features. However, the features extracted from the text show, in general, better performances than demographics, behavior, and network features in both datasets, as shown in Table 4. Combining all the groups of features together further improves the average precision of the classification task (cf. Table 3).

6.3 Feature importance and Shapley additive explanations

Considering all the features from each group, we have a total of 91 and 99 features for the PAN and PolitiFact datasets, respectively, which can still be too many for the size of the considered datasets (PolitiFact and PAN) to perform real versus fake news spreader classification. Therefore, we used the statistical tests (ANOVA and Wilcoxon rank-sum depending on the data distribution) as in [20,42] to perform feature selection. For each dataset, only features where the two averages (real vs. fake news spreader) were significantly different according to the statistical test (p value < 0.05) were considered. Also, features are sorted by F value in descending order to determine the importance. Among these features, we selected the top-k most important features to feed the classification algorithm, where k is the square root of the training set size (rule of thumb). For each dataset, the selected important features are shown in Fig. 9.

Table 5 shows our classification results with important features using the best classifier, i.e., random forest. We observe that using only important features lowers the performance by a very small margin 0.19% and 0.1% in the PAN 2020 dataset and PolitiFact datasets, respectively. However, it still outperforms the scores of both baselines shown in Table 3.

Further, to explain why users are classified as fake news spreaders or real news spreaders, we used the SHAP values (SHapley Additive exPlanations) of the selected features, a widely used approach inspired by cooperative game theory [24]. We leveraged a tree explainer which is basically used to compute SHAP values for tree-based models. Since we want to learn about how each feature is influencing the decision of the model, we used the global importance, i.e., the



(a) PolitiFact **(b)** PAN

Fig. 9 SHAP summaries of the important features: PolitiFact and PAN datasets. Y-axis represents the features in order of importance. X-axis represents the SHAP values, positive values (greater than zero) represents a higher chance of classifying a user as a fake news spreader and negative values represent a higher chance of classifying a user as a real news spreader

sum of all the absolute Shapley values per feature across the dataset. Figure 9 shows the SHAP summary plot that demonstrates the contribution of each feature in predicting users likely to spread fake news. The higher the SHAP value (i.e., closer to 1.0), the higher the probability of being a fake news spreader. As shown in the figure, writing style features (like frequency of lowercase words, uppercased characters) appear as the most important features in the model for the PolitiFact dataset. We observe that users writing tweets with fewer lowercased words, more uppercased characters, more breaking, less punctuation, shorter text, and fewer stopwords are more likely to be fake news spreaders according to the PolitiFact dataset. On the other end, features indicating emotions like happiness and anger and personality facets such as excitement, assertiveness, openness to change, artistic interests appear as the most important features in the model for the PAN 2020 dataset. We see that the users with less concern about others' welfare and interests (self-transcendence), less concordance (harmony), and having the willingness to change (openness to change) are more likely to be fake news spreaders, according to the PAN 2020 dataset.

Additionally, we further confirm that negative emotions like anger, fear, disgust, stress, and sadness extracted from the tweets of a user are among the most important features and indicate that the users likely to spread fake news seem to embrace a language with more negative valiance than real news spreaders in both datasets.

Table 5 Average precision of important features from Fig. 9 versus all features on PolitiFact and PAN 2020 datasets

Features	PolitiFact	PAN 2020
Important features	0.994	0.776
All features	0.995	0.795

7 Conclusions

In this article, we performed a comprehensive analysis to understand the correlation between user characteristics based on different attributes such as user demographics, personality, emotion, writing style and readability, social media behavior, and the likelihood of a user being a fake news spreader. We considered two datasets to perform our analysis, namely the PolitiFact (FakeNewsNet) and PAN datasets, and investigated new features such as user tweeting behavior and stress level. Furthermore, we addressed the problem of identifying users likely to share fake news using the proposed groups of features in both datasets and compared the performance with baseline approaches from the PAN shared task. Specifically, we obtained an average precision of 0.99 on the PolitiFact dataset (vs. 0.96 achieved by the best baseline) and 0.80 on the PAN dataset (vs. 0.78 achieved by the best baseline).

Our results showed the potential of the proposed features in identifying fake news spreaders by outperforming baseline approaches in both considered datasets. Our findings showed that younger generation under 18 or users over 40 may be more vulnerable in case of fake news sharing, and females may be more likely to be fake news spreaders than male users. Similarly, fake news spreaders tend to express more negative emotion and stress in their tweets, and the political orientation of a fake news spreader is more likely to coincide with the bias of the source of the majority of circulating fake news items. Besides, the behavioral patterns show that fake news spreaders have newer accounts, spend less time but tweet more within a short time interval. Likewise, it shows the inferred user personality, writing styles, and readability of the user's tweets have the potential to identify whether the user is a fake news spreader effectively.

Using an automated tool to infer user demographics based on their screen name, description, and profile image could be a potential limitation of our study. Thus, inferred demographics of some of the users might not be entirely accurate. However, it is impossible to test the tool's efficiency in the considered datasets as such metadata are not explicitly available to be used as ground truth. Labels in the PAN 2020 dataset are another limitation of the work presented in this article as a user is labeled as a fake news spreader if they have shared at least one fake news item in the past. We have proposed a way to compute more reliable labels for the PolitiFact dataset to overcome this limitation. Finally, we have considered only users keen to spread fake political news, and we leave as future work the study of fake news spreaders in other domains, e.g., gossip news.

Acknowledgements This work has been supported by the National Science Foundation under awards No. 1943370 and 1820685 and the Idaho Global Entrepreneurial Mission/Higher Education Research Council under Award No. IGEM22-001. We thank Hongmin (Steven) Kim for collecting additional Twitter data, Ashlee Milton and Maria Soledad Pera for providing us the code used in their article [26] to compute emotional features, and Brian Daley for implementing the code to compute user political orientation.

References

1. Barometer, E.T.: Edelman trust barometer global report. Edelman, available at: https://www.edelman.com/sites/g/files/aatuss191/files/2019-02/2019_Edelman_Trust_Barometer_Global_Report_2.pdf (2019)
2. Benesch, C.: An empirical analysis of the gender gap in news consumption. J. Media Econ. **25**(3), 147–167 (2012)
3. Bonilla, Y., Rosa, J.: # ferguson: digital protest, hashtag ethnography, and the racial politics of social media in the United States. Am. Ethnol. **42**(1), 4–17 (2015)
4. Bruns, A., Highfield, T., Lind, R.A.: Blogs, twitter, and breaking news: the produsage of citizen journalism. Produs. Theory Digit. World Intersect. Audiences Prod. Contemp. Theory **80**(2012), 15–32 (2012)
5. Buda, J., Bolonyai, F.: An ensemble model using n-grams and statistical features to identify fake news spreaders on Twitter. In: CLEF (2020)
6. Chamberlain, J.M., Spezzano, F., Kettler, J.J., Dit, B.: A network analysis of twitter interactions by members of the US congress. ACM Trans. Soc. Comput. **4**(1), 1–22 (2021)
7. Common Sense Media: News and America's kids: How young people perceive and are impacted by the news (2017)
8. Davis, J., Goadrich, M.: The relationship between precision-recall and ROC curves. In: Cohen, W.W., Moore, A.W. (eds.) Machine Learning, Proceedings of the Twenty-Third International Conference (ICML, ACM International Conference Proceeding Series, vol. 148, pp. 233–240. ACM (2006)
9. De Choudhury, M., Gamon, M., Counts, S., Horvitz, E.: Predicting depression via social media. ICWSM **13**, 1–10 (2013)
10. Effron DA, Raj M Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share. Psychological Science. 2020;31(1):75–87. https://doi.org/10.1177/0956797619887896
11. Gabielkov, M., Ramachandran, A., Chaintreau, A., Legout, A.: Social clicks: What and who gets read on twitter? ACM SIGMETRICS Perform. Eval. Rev. **44**(1), 179–192 (2016)
12. Ghanem, B., Rosso, P., Rangel, F.: An emotional analysis of false information in social media and news articles. ACM Trans. Internet Technol. (TOIT) **20**(2), 1–18 (2020)
13. Giachanou, A., Rissola, E.A., Ghanem, B., Crestani, F., Rosso, P.: The role of personality and linguistic patterns in discriminating between fake news spreaders and fact checkers. In: Natural Language Processing and Information Systems: 25th International Conference on Applications of Natural Language to Information Systems, NLDB 2020, p. 181. Springer
14. Gottfriedd, J., Grieco, E.: Younger Americans are Better than Older Americans at Telling Factual News Statements from Opinions. Pew Research Center (2018)
15. Guess, A., Nagler, J., Tucker, J.: Less than you think: prevalence and predictors of fake news dissemination on Facebook. Sci. Adv. **5**(1), eaau4586 (2019)
16. Gupta, A., Lamba, H., Kumaraguru, P., Joshi, A.: Faking sandy: characterizing and identifying fake images on Twitter during hurricane sandy. In: Proceedings of the 22nd international conference on World Wide Web, pp. 729–736 (2013)
17. Hall, J.A., Pennington, N., Lueders, A.: Impression management and formation on Facebook: a lens model approach. New Media Soc. **16**(6), 958–982 (2014)
18. Hemphill, L., Culotta, A., Heston, M.: # polar scores: measuring partisanship using social media content. J. Inf. Technol. Politics **13**(4), 365–377 (2016)
19. Hills, T.T.: The dark side of information proliferation. Perspect. Psychol. Sci. **14**(3), 323–330 (2019)
20. Horne, B.D., Adali, S.: This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In: The 2nd International Workshop on News and Public Opinion at ICWSM (2017)
21. Horne, B.D., Nevo, D., O'Donovan, J., Cho, J., Adali, S.: Rating reliability and bias in news articles: Does AI assistance help everyone? In: Proceedings of the Thirteenth International Conference on Web and Social Media, ICWSM 2019, pp. 247–256 (2019)
22. Isaak, J., Hanna, M.J.: User data privacy: Facebook, Cambridge analytica, and privacy protection. Computer **51**(8), 56–59 (2018)
23. Lee, C.S., Ma, L.: News sharing in social media: the effect of gratifications and prior experience. Comput. Hum. Behav. **28**(2), 331–339 (2012)

24. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 30, pp. 4765–4774. Curran Associates, Inc. (2017). http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf
25. Ma, L., Lee, C.S., Goh, D.H.: Understanding news sharing in social media from the diffusion of innovations perspective. In: 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, pp. 1013–1020. IEEE (2013)
26. Milton, A., Batista, L., Allen, G., Gao, S., Ng, Y., Pera, M.S.: "Don't judge a book by its cover": Exploring book traits children favor. In: Santos, R.L.T., Marinho, L.B., Daly, E.M., Chen, L., Falk, K., Koenigstein, N., de Moura, E.S. (eds.) RecSys 2020: Fourteenth ACM Conference on Recommender Systems, Virtual Event, Brazil, September 22–26, 2020, pp. 669–674. ACM (2020). https://doi.org/10.1145/3383313.3418490
27. Milton, A., Pera, M.S.: What snippets feel: Depression, search, and snippets. In: Cantador, I., Chevalier, M., Melucci, M., Mothe, J. (eds.) Proceedings of the Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020), Samatan, Gers, France, July 6–9, 2020, CEUR Workshop Proceedings, vol. 2621. CEUR-WS.org (2020). http://ceur-ws.org/Vol-2621/CIRCLE20_15.pdf
28. Mitchell, A., Gottfriedd, J., Barthel, M., Sumida, N.: Distinguishing Between Factual and Opinion Statements in the News. Pew Research Center (2018)
29. Mohammad, S.: Word affect intensities. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, Miyazaki, Japan, May 7–12, 2018. European Language Resources Association (ELRA) (2018)
30. Neuman, Y.: Computational Personality Analysis: Introduction, Practical Applications and Novel Directions. Springer (2016)
31. Oh, H.J., Ozkaya, E., LaRose, R.: How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. Comput. Hum. Behav. **30**, 69–78 (2014). https://doi.org/10.1016/j.chb.2013.07.053
32. Pizarro, J.: Using n-grams to detect fake news spreaders on Twitter. In: CLEF (2020)
33. Poindexter, P., Harp, D.: The softer side of news. In: P. Poindexter, S. Meraz, A.S. Weiss (eds.) Women, menand news: Divided and disconnected in the news medialandscape, pp. 85–96. Routledge New York (2008)
34. Rangel, F., Giachanou, A., Ghanem, B., Rosso, P.: Overview of the 8th author profiling task at PAN 2020: profiling fake news spreaders on Twitter. In: Cappellato, L., Eickhoff, C., Ferro, N., Névéol, A. (eds.) CLEF 2020 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (2020). CEUR-WS.org
35. Rangel, F., Giachanou, A., Ghanem, B., Rosso, P.: Overview of the 8th author profiling task at pan 2020: profiling fake news spreaders on Twitter. In: CLEF (2020)
36. Reis, J.C., Kwak, H., An, J., Messias, J., Benevenuto, F.: Demographics of news sharing in the us Twittersphere. In: Proceedings of the 28th ACM Conference on Hypertext and Social Media, pp. 195–204 (2017)
37. Rini, R.: Fake news and partisan epistemology. Kennedy Inst. Ethics J. **27**(2), E-43 (2017)
38. Shearer, E., Mitchell, A.: News Use Across Social Media Platforms in 2020. Pew Research Center (2020)
39. Shoemaker, P.J., Vos, T.P., Reese, S.D.: Journalists as gatekeepers. In: K. Wahl-Jorgensen, T. Hanitzsch (eds.) The handbook of journalism studies, vol. 73, p. 15 (2009)
40. Shrestha, A., Serra, E., Spezzano, F.: Multi-modal social and psycho-linguistic embedding via recurrent neural networks to identify depressed users in online forums. Netw. Model. Anal. Health Inform. Bioinform. **9**(1), 1–11 (2020)
41. Shrestha, A., Spezzano, F.: Online misinformation: from the deceiver to the victim. In: ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining, pp. 847–850. ACM (2019)
42. Shrestha, A., Spezzano, F.: Textual characteristics of news title and body to detect fake news: a reproducibility study. In: Advances in Information Retrieval—43rd European Conference on IR Research, ECIR 2021, Proceedings, Part II, Lecture Notes in Computer Science, vol. 12657, pp. 120–133. Springer (2021)
43. Shrestha, A., Spezzano, F., Joy, A.: Detecting fake news spreaders in social networks via linguistic and personality features. In: CLEF (2020)
44. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: Fakenewsnet: a data repository with news content, social context and dynamic information for studying fake news on social media, vol. 8. arXiv preprint arXiv:1809.01286 (2018)
45. Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: a data mining perspective. ACM SIGKDD Explor. Newslett. **19**(1), 22–36 (2017)
46. Shu, K., Zhou, X., Wang, S., Zafarani, R., Liu, H.: The role of user profiles for fake news detection. In: ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining, pp. 436–439. ACM (2019)
47. This analysis shows how viral fake election news stories outperformed real news on facebook, 2016. https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformedreal-news-on-facebook (2016) facebook, 2016. https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformedreal-news-on-facebook (2016)
48. Spezzano, F., Shrestha, A., Fails, J.A., Stone, B.W.: That's fake news! Investigating how readers identify the reliability of news when provided title, image, source bias, and full articles. In: Proceedings of the ACM on Human Computer Interaction Journal, vol. 5 (CSCW1, Article 109) (2021)
49. Spiro, E.S., Fitzhugh, S., Sutton, J., Pierski, N., Greczek, M., Butts, C.T.: Rumoring during extreme events: a case study of deepwater horizon 2010. In: Proceedings of the 4th Annual ACM Web Science Conference, pp. 275–283 (2012)
50. Vogel, L.: (2017). Viral misinformation threatens public health. CMAJ : Canadian Medical Association journal, 189(50), E1567. https://doi.org/10.1503/cmaj.109-5536
51. Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science **359**(6380), 1146–1151 (2018)
52. Wang, W., Hernandez, I., Newman, D.A., He, J., Bian, J.: Twitter analysis: studying us weekly trends in work stress and emotion. Appl. Psychol. **65**(2), 355–378 (2016)
53. Wang, Z., Hale, S., Adelani, D.I., Grabowicz, P., Hartman, T., Flöck, F., Jurgens, D.: Demographic inference and representative population estimates from multilingual social media data. In: The World Wide Web Conference, pp. 2056–2067 (2019)
54. Wineburg, S., McGrew, S., Breakstone, J., Ortega, T.: Evaluating information: The cornerstone of civic online reasoning. Stanford Digital Repository. Retrieved January 8, 2018 (2016)
55. Yaqub, W., Kakhidze, O., Brockman, M.L., Memon, N., Patil, S.: Effects of credibility indicators on social media news sharing intent. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI'20, pp. 1–14 (2020)