ARTICLE

Check for updates

Emotions unveiled: detecting COVID-19 fake news on social media

Bahareh Farhoudinia [1], Selcen Ozturkcan [1,2] & Nihat Kasap [1]

The COVID-19 pandemic has highlighted the pernicious effects of fake news, underscoring the critical need for researchers and practitioners to detect and mitigate its spread. In this paper, we examined the importance of detecting fake news and incorporated sentiment and emotional features to detect this type of news. Specifically, we compared the sentiments and emotions associated with fake and real news using a COVID-19 Twitter dataset with labeled categories. By utilizing different sentiment and emotion lexicons, we extracted sentiments categorized as positive, negative, and neutral and eight basic emotions, anticipation, anger, joy, sadness, surprise, fear, trust, and disgust. Our analysis revealed that fake news tends to elicit more negative emotions than real news. Therefore, we propose that negative emotions could serve as vital features in developing fake news detection models. To test this hypothesis, we compared the performance metrics of three machine learning models: random forest, support vector machine (SVM), and Naïve Bayes. We evaluated the models' effectiveness with and without emotional features. Our results demonstrated that integrating emotional features into these models substantially improved the detection performance, resulting in a more robust and reliable ability to detect fake news on social media. In this paper, we propose the use of novel features and methods that enhance the field of fake news detection. Our findings underscore the crucial role of emotions in detecting fake news and provide valuable insights into how machine-learning models can be trained to recognize these features.

[1] Sabancı Business School, Sabancı University, Istanbul, Turkey. [2] School of Business and Economics, Linnaeus University, Växjö, Sweden. ✉email: selcen.ozturkcan@lnu.se; selcen@sabanciuniv.edu

Introduction

Social media has changed human life in multiple ways. People from all around the world are connected via social media. Seeking information, entertainment, communicatory utility, convenience utility, expressing opinions, and sharing information are some of the gratifications of social media (Whiting and Williams, 2013). Social media is also beneficial for political parties or companies since they can better connect with their audience through social media (Kumar et al., 2016). Despite all the benefits that social media adds to our lives, there are also disadvantages to its use. The emergence of fake news is one of the most important and dangerous consequences of social media (Baccarella et al., 2018, 2020). Zhou et al. (2019) suggested that fake news threatens public trust, democracy, justice, freedom of expression, and the economy. In the 2016 United States (US) presidential election, fake news engagement outperformed mainstream news engagement and significantly impacted the election results (Silverman, 2016). In addition to political issues, fake news can cause irrecoverable damage to companies. For instance, Pepsi stock fell by 4% in 2016 when a fake story about the company's CEO spread on social media (Berthon and Pitt, 2018). During the COVID-19 pandemic, fake news caused serious problems, e.g., people in Europe burned 5G towers because of a rumor claiming that these towers damaged the immune system of humans (Mourad et al., 2020). The World Health Organization (WHO) asserted that misinformation and propaganda propagated more rapidly than the COVID-19 pandemic, leading to psychological panic, the circulation of misleading medical advice, and an economic crisis.

This study, which is a part of a completed PhD thesis (Farhoundinia, 2023), focuses on analyzing the emotions and sentiments elicited by fake news in the context of COVID-19. The purpose of this paper is to investigate how emotions can help detect fake news. This study aims to address the following research questions: 1. How do the sentiments associated with real news and fake news differ? 2. How do the emotions elicited by fake news differ from those elicited by real news? 3. What particular emotions are most prevalent in fake news? 4. How can these feelings be used to recognize fake news on social media?

This paper is arranged into six sections: Section "Related studies" reviews the related studies; Section "Methods" explains the proposed methodology; and Section "Results and analysis" presents the implemented models, analysis, and related results in detail. Section "Discussion and limitations" discusses the research limitations, and the conclusion of the study is presented in Section "Conclusion".

Related studies

Research in the field of fake news began following the 2016 US election (Carlson, 2020; Wang et al., 2019). Fake news has been a popular topic in multiple disciplines, such as journalism, psychology, marketing, management, health care, political science, information science, and computer science (Farhoudinia et al., 2023). Therefore, fake news has not been defined in a single way; according to Berthon and Pitt (2018), misinformation is the term used to describe the unintentional spread of fake news. Disinformation is the term used to describe the intentional spread of fake news to mislead people or attack an idea, a person, or a company (Allcott and Gentzkow, 2017). Digital assets such as images and videos could be used to spread fake news (Rajamma et al., 2019). Advancements in computer graphics, computer vision, and machine learning have made it feasible to create fake images or movies by merging them together (Agarwal et al., 2020). Additionally, deep fake videos pose a risk to public figures,

businesses, and individuals in the media. Detecting deep fakes is challenging, if not impossible, for humans.

The reasons for believing and sharing fake news have attracted the attention of several researchers (e.g., Al-Rawi et al., 2019; Apuke and Omar, 2020; Talwar, Dhir et al., 2019). Studies have shown that people have a tendency to favor news that reinforces their existing beliefs, a cognitive phenomenon known as confirmation bias. This inclination can lead individuals to embrace misinformation that aligns with their preconceived notions (Kim and Dennis, 2019; Meel and Vishwakarma, 2020). Although earlier research focused significantly on the factors that lead people to believe and spread fake news, it is equally important to understand the cognitive mechanisms involved in this process. These cognitive mechanisms, as proposed by Kahneman (2011), center on two distinct systems of thinking. In system-one cognition, conclusions are made without deep or conscious thoughts; however, in system-two cognition, there is a deeper analysis before decisions are made. Based on Moravec et al. (2020), social media users evaluate news using 'system-one' cognition; therefore, they believe and share fake news without deep thinking. It is essential to delve deeper into the structural aspects of social media platforms that enable the rapid spread of fake news. Social media platforms are structured to show that posts and news are aligned with users' ideas and beliefs, which is known as the root cause of the echo chamber effect (Cinelli et al., 2021). The echo chamber effect has been introduced as an aspect that causes people to believe and share fake news on social media (e.g., Allcott and Gentzkow, 2017; Berthon and Pitt, 2018; Chua and Banerjee, 2018; Peterson, 2019).

In the context of our study, we emphasize the existing body of research that specifically addresses the detection of fake news (Al-Rawi et al., 2019; Faustini and Covões, 2020; Ozbay and Alatas, 2020; Raza and Ding, 2022). Numerous studies that are closely aligned with the themes of our present investigation have delved into methodological approaches for identifying fake news (Er and Yılmaz, 2023; Hamed et al., 2023; Iwendi et al., 2022). Fake news detection methods are classified into three categories: (i) content-based, (ii) social context, and (iii) propagation-based methods. (i) Content-based fake news detection models are based on the content and linguistic features of the news rather than user and propagation characteristics (Zhou and Zafarani, 2019, p. 49). (ii) Fake news detection based on social context employs user demographics such as age, gender, education, and follower–followee relationships of the fake news publishers as features to recognize fake news (Jarrahi and Safari, 2023). (iii) Propagation-based approaches are based on the spread of news on social media. The input of the propagation-based fake news detection model is a cascade of news, not text or user profiles. Cascade size, cascade depth, cascade breadth, and node degree are common features of detection models (Giglietto et al., 2019; de Regt et al., 2020; Vosoughi et al., 2018).

Machine learning methods are widely used in the literature because they enable researchers to handle and process large datasets (Ongsulee, 2017). The use of machine learning in fake news research has been extremely beneficial, especially in the domains of content-based, social context-based, and propagation-based fake news identification. These methods leverage the advantages of a range of characteristics, including sentiment-related, propagation, temporal, visual, linguistic, and user/account aspects. Fake news detection frequently makes use of machine learning techniques such as logistic regressions, decision trees, random forests, naïve Bayes, and support vector machine (SVM). Studies on the identification of fake news also include deep learning models, such as convolutional neural networks (CNN) and long short-term memory (LSTM) networks, which can

provide better accuracy in certain situations. Even with a small amount of training data, pretrained language models such as bidirectional encoder representations from transformers (BERT) show potential for identifying fake news (Kaliyar et al., 2021). Amer et al. (2022) investigated the usefulness of these models in benchmark studies covering different topics.

The role of emotions in identifying fake news within academic communities remains an area with considerable potential for additional research. Despite many theoretical and empirical studies, this topic remains inadequately investigated. Ainapure et al. (2023) analyzed the sentiments elicited by tweets in India during the COVID-19 pandemic with deep learning and lexicon-based techniques using the valence-aware dictionary and sentiment reasoner (Vader) and National Research Council (NRC) lexicons to understand the public's concerns. Dey et al. (2018) applied several natural language processing (NLP) methods, such as sentiment analysis, to a dataset of tweets about the 2016 U.S. presidential election. They found that fake news had a strong tendency toward negative sentiment; however, their dataset was too limited (200 tweets) to provide a general understanding. Cui et al. (2019) found that sentiment analysis was the best-performing component in their fake news detection framework. Ajao et al. (2019) studied the hypothesis that a relationship exists between fake news and the sentiments elicited by such news. The authors tested hypotheses with different machine learning classifiers. The best results were obtained by sentiment-aware classifiers. Pennycook and Rand (2020) argued that reasoning and analytical thinking help uncover news credibility; therefore, individuals who engage in reasoning are less likely to believe fake news. Prior psychology research suggests that an increase in the use of reason implies a decrease in the use of emotions (Mercer, 2010).

In this study, we apply sentiment analysis to the more general topic of fake news detection. The focus of this study is on the tweets that were shared during the COVID-19 pandemic. Many scholars focused on the effects of media reports, providing comprehensive information and explanations about the virus. However, there is still a gap in the literature on the characteristics and spread of fake news during the COVID-19 pandemic. A comprehensive study can enhance preparedness efforts for any similar future crisis. The aim of this study is to answer the question of how emotions aid in fake news detection during the COVID-19 pandemic. Our hypothesis is that fake news carries negative emotions and is written with different emotions and sentiments than those of real news. We expect to extract more negative sentiments and emotions from fake news than from real news. Existing works on fake news detection have focused mainly on news content and social context. However, emotional information has been underutilized in previous studies (Ajao et al., 2019). We extract sentiments and eight basic emotions from every tweet in the COVID-19 Twitter dataset and use these features to classify fake and real news. The results indicate how emotions can be used in differentiating and detecting fake and real news.

Methods

With our methodology, we employed a multifaceted approach to analyze tweet text and discern sentiment and emotion. The steps involved were as follows: (a) Lexicons such as Vader, TextBlob, and SentiWordNet were used to identify sentiments embedded in the tweet content. (b) The NRC emotion lexicon was utilized to recognize the range of different emotions expressed in the tweets. (c) Machine learning models, including the random forest, naïve Bayes, and SVM classifiers, as well as a deep learning model, BERT, were integrated. These models were strategically applied to

the data for fake news detection, both with and without considering emotions. This comprehensive approach allowed us to capture nuanced patterns and dependencies within the tweet data, contributing to a more effective and nuanced analysis of the fake news content on social media.

An open, science-based, publicly available dataset was utilized. The dataset comprises 10,700 English tweets with hashtags relevant to COVID-19, categorized with real and fake labels. Previously used by Vasist and Sebastian (2022) and Suter et al. (2022), the manually annotated dataset was compiled by Patwa et al. (2021) in September 2020 and includes tweets posted in August and September 2020. According to their classification, the dataset is balanced, with 5600 real news stories and 5100 fake news stories. The dataset used for the study was generated by sourcing fake news data from public fact-checking websites and social media outlets, with manual verification against the original documents. Web-based resources, including social media posts and fact-checking websites such as PolitiFact and Snopes, played a key role in collecting and adjudicating details on the veracity of claims related to COVID-19. For real news, tweets from official and verified sources were gathered, and each tweet was assessed by human reviewers based on its contribution of relevant information about COVID-19 (Patwa et al., 2021; Table 2 on p. 4 of Suter et al., 2022, which is excerpted from Patwa et al. (2021), also provides an illustrative overview).

Preprocessing is an essential step in any data analysis, especially when dealing with textual data. Appropriate preprocessing steps can significantly enhance the performance of the models. The following preprocessing steps were applied to the dataset: removing any characters other than alphabets, change the letters to lower-case, deleting stop words such as "a," "the," "is," and "are," which carry very little helpful information, and performing lemmatization. The text data were transformed into quantitative data by the scikit-learn ordinal encoder class.

The stages involved in this research are depicted in a high-level schematic that is shown in Fig. 1. First, the sentiments and emotions elicited by the tweets were extracted, and then, after studying the differences between fake and real news in terms of sentiments and emotions, these characteristics were utilized to construct fake news detection models.

Sentiment analysis. Sentiment analysis is the process of deriving the sentiment of a piece of text from its content (Vinodhini and Chandrasekaran, 2012). Sentiment analysis, as a subfield of natural language processing, is widely used in analyzing the reviews of a product or service and social media posts related to different topics, events, products, or companies (Wankhade et al., 2022). One major application of sentiment analysis is in strategic marketing. Păvăloaia et al. (2019), in a comprehensive study on two companies, Coca-Cola and PepsiCo, confirmed that the activity of these two brands on social media has an emotional impact on existing or future customers and the emotional reactions of customers on social media can influence purchasing decisions. There are two methods for sentiment analysis: lexicon-based and machine-learning methods. Lexicon-based sentiment analysis uses a collection of known sentiments that can be divided into dictionary-based lexicons or corpus-based lexicons (Pawar et al., 2015). These lexicons help researchers derive the sentiments generated from a text document. Numerous dictionaries, such as Vader (Hutto and Gilbert, 2014), SentiWordNet (Esuli and Sebastiani, 2006), and TextBlob (Loria, 2018), can be used for scholarly research.

In this research, Vader, TextBlob, and SentiWordNet are the three lexicons used to extract the sentiments generated from tweets. The Vader lexicon is an open-source lexicon attuned
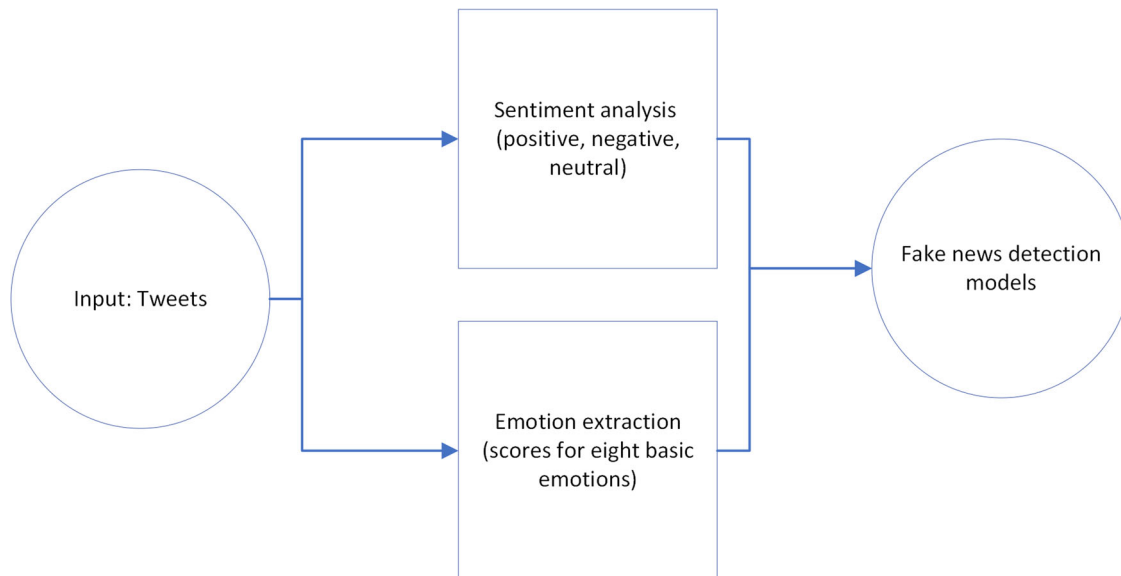
Fig. 1 High-level schematic process. The figure depicts the stages involved in this research in a high-level schematic.
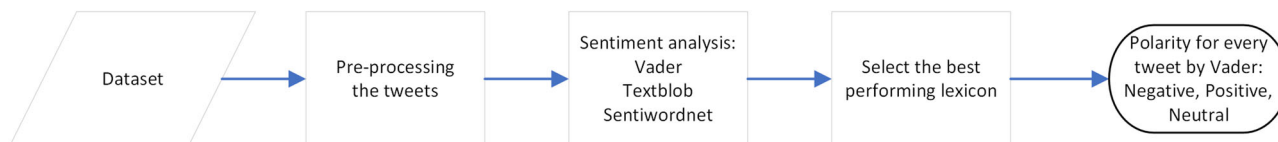

Fig. 2 Sentiment analysis steps. The figure illustrates the steps for the sentiment analysis of tweets.

Table 1 Three-class and two-class classification performance using the three lexicons (Farhoundinia, 2023).

		VADER	TextBlob	SentiWordNet
2-Class classification	Accuracy	0.77	0.55	0.55
	Precision	0.66	0.46	0.65
	Recall	0.84	0.83	0.45
	F1-score	0.73	0.58	0.52
3-Class classification	Accuracy	0.64	0.41	0.40
	Macro precision	0.64	0.46	0.41
	Macro recall	0.64	0.44	0.41
	Macro F1-score	0.63	0.41	0.40
	Weighted precision	0.65	0.49	0.43
	Weighted recall	0.64	0.41	0.40
	Weighted F1-score	0.63	0.41	0.41

specifically to social media (Hutto and Gilbert, 2014). TextBlob is a Python library that processes text specifically designed for natural language analysis (Loria, 2018), and SentiWordNet is an opinion lexicon adapted from the WordNet database (Esuli and Sebastiani, 2006). Figure 2 shows the steps for the sentiment analysis of tweets.

Different methods and steps were used to choose the best lexicon. First, a random partition of the dataset was manually labeled as positive, negative, or neutral. The results of every lexicon were compared with the manually labeled sentiments, and the performance metrics for every lexicon are reported in Table 1. Second, assuming that misclassifying negative and positive tweets

as neutral is not as crucial as misclassifying negative tweets as classifying positive tweets, the neutral tweets were ignored, and a comparison was made on only positive and negative tweets. The three-class and two-class classification metrics are compared in Table 1.

Third, this study's primary goal was to identify the precise distinctions between fake and real tweets to improve the detection algorithm. We addressed how well fake news was detected with the three sentiment lexicons, as different results were obtained. This finding means that a fake news detection model was trained with the dataset using the outputs from three lexicons: Vader, TextBlob, and SentiWordNet. As previously indicated, the dataset includes labels for fake and real news, which allows for the application of supervised machine learning detection models and the evaluation of how well various models performed. The Random Forest algorithm is a supervised machine learning method that has achieved good performance in the classification of text data. The dataset contains many tweets and numerical data reporting the numbers of hospitalized, deceased, and recovered individuals who do not carry any sentiment. During this phase, tweets containing numerical data were excluded; this portion of the tweets constituted 20% of the total. Table 2 provides information on the classification power using the three lexicons with nonnumerical data. The models were more accurate when using sentiments drawn from Vader. This finding means the Vader lexicon may include better classifications of fake and real news. Vader was selected as the superior sentiment lexicon after evaluating all three processes. The steps for choosing the best lexicon are presented in Fig. 3 (also see Appendix A in Supplementary Information for further details on the procedure). Based on the results achieved when using Vader, the tweets that are labeled as fake include more negative sentiments than those of real tweets. Conversely, real tweets include more positive sentiments.

Emotion extraction. Emotions elicited in tweets were extracted using the NRC emotion lexicon. This lexicon measures emotional effects from a body of text, contains ~27,000 words, and is based on the National Research Council Canada's affect lexicon and the natural language toolkit (NLTK) library's WordNet synonym sets (Mohammad and Turney, 2013). The lexicon includes eight scores for eight emotions based on Plutchick's model of emotion (Plutchik, 1980): joy, trust, fear, surprise, sadness, anticipation, anger, and disgust. These emotions can be classified into four opposing pairs: joy–sadness, anger–fear, trust–disgust, and anticipation–surprise. The NRC lexicon assigns each text the emotion with the highest score. Emotion scores from the NRC lexicon for every tweet in the dataset were extracted and used as features for the fake news detection model. The features of the model include the text of the tweet, sentiment, and eight emotions. The model was trained with 80% of the data and tested with 20%. Fake news had a greater prevalence of negative emotions, such as fear, disgust, and anger, than did real news, and real news had a greater prevalence of positive emotions, such as anticipation, joy, and surprise, than did fake news.

Fake news detection. In the present study, the dataset was divided into a training set (80%) and a test set (20%). The dataset was analyzed using three machine learning models: random forest, SVM, and naïve Bayes. Appendices A and B provide information on how the results were obtained and how they correlate with the research corpus.

Random forest: An ensemble learning approach that fits several decision trees to random data subsets. This classifier is popular for text classification, high-dimensional data, and feature importance since it overfits less than decision trees. The Random Forest classifier in scikit-learn was used in this study (Breiman, 2001).

Table 2 Classification scores for the lexicons (Farhoundinia, 2023).

Features	Random forest accuracy
Vader scores + tweets	0.738
TextBlob scores + tweets	0.735
SentiWordNet scores + tweets	0.728

Naïve Bayes: This model uses Bayes' theorem to solve classification problems, such as sorting documents into groups and blocking spam. This approach works well with text data and is easy to use, strong, and good for problems with more than one label. The Naïve Bayes classifier from scikit-learn was used in this study (Zhang, 2004).

Support vector machines (SVMs): Supervised learning methods that are used to find outliers, classify data, and perform regression. These methods work well with data involving many dimensions. SVMs find the best hyperplanes for dividing classes. In this study, the SVM model from scikit-learn was used (Cortes and Vapnik, 1995).

Deep learning models can learn how to automatically describe data in a hierarchical way, making them useful for tasks such as identifying fake news (Salakhutdinov et al., 2012). A language model named bidirectional encoder representations from transformers (BERT) was used in this study to help discover fake news more easily.

BERT: A cutting-edge NLP model that uses deep neural networks and bidirectional learning and can distinguish patterns on both sides of a word in a sentence, which helps it understand the context and meaning of text. BERT has been pretrained with large datasets and can be fine-tuned for specific applications to capture unique data patterns and contexts (Devlin et al., 2018).

In summary, we applied machine learning models (random forest, naïve Bayes, and SVM) and a deep learning model (BERT) to analyze text data for fake news detection. The impact of emotion features on detecting fake news was compared between models that include these features and models that do not include these features. We found that adding emotion scores as features to machine learning and deep learning models for fake news detection can improve the model's accuracy. A more detailed analysis of the results is given in the section "Results and analysis".

Results and analysis

In the sentiment analysis using tweets from the dataset, positive and negative sentiment tweets were categorized into two classes: fake and real. Figure 4 shows a visual representation of the differences, while the percentages of the included categories are presented in Table 3. In fake news, the number of negative sentiments is greater than the number of positive sentiments (39.31% vs. 31.15%), confirming our initial hypothesis that fake news
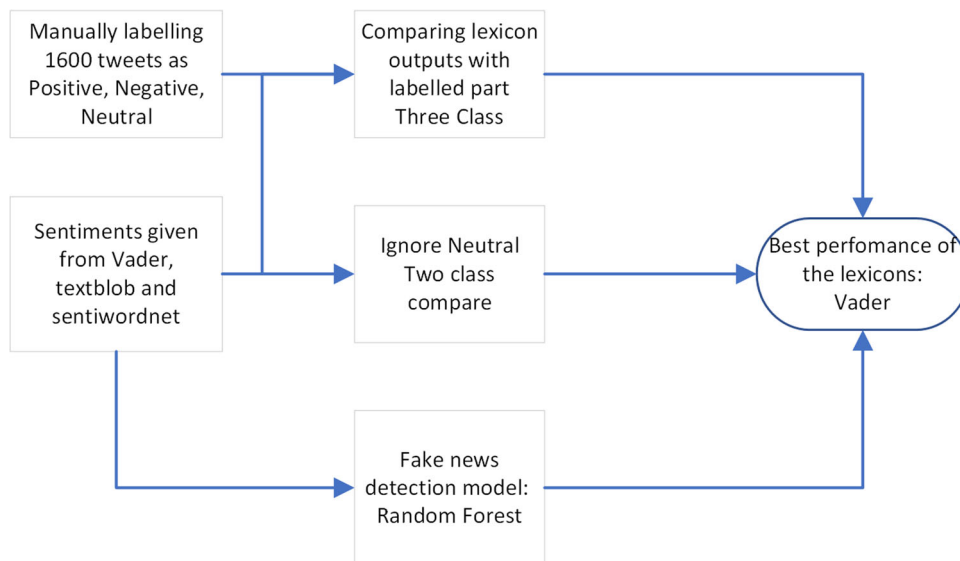

Fig. 3 Process for selecting the best lexicon. The figure exhibits the steps for choosing the best lexicon.
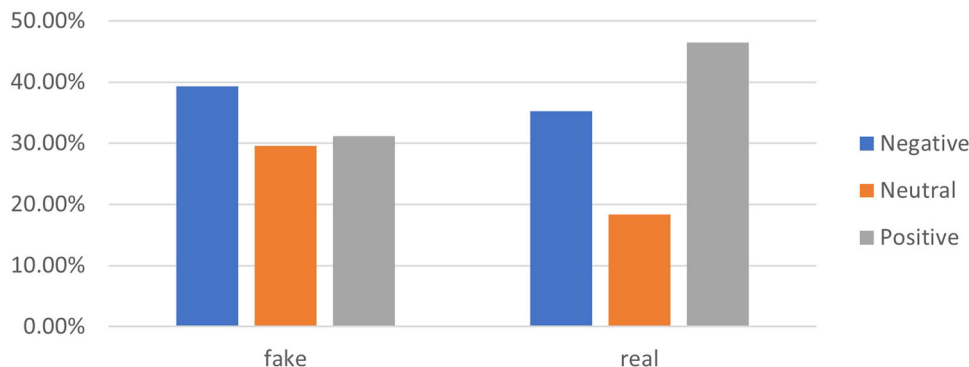
Fig. 4 Frequency of sentiments in the Vader lexicon for each class. The figure displays a visual representation of the differences of sentiments in each class.

Table 3 Vader sentiments categorized by fake and real classes (Farhoundinia, 2023).

Label	Negative (%)	Neutral (%)	Positive (%)
Fake	39.31	29.53	31.15
Real	35.20	18.35	46.45

disseminators use extreme negative emotions to attract readers' attention.

Fake news disseminators aim to attack or satirize an idea, a person, or a brand using negative words and emotions. Baumeister et al. (2001) suggested that negative events are stronger than positive events and that negative events have a more significant impact on individuals than positive events. Accordingly, individuals sharing fake news tend to express more negativity for increased impressiveness. The specific topics of the COVID-19 pandemic, such as the source of the virus, the cure for the illness, the strategy the government is using against the spread of the virus, and the spread of vaccines, are controversial topics. These topics, known for their resilience against strong opposition, have become targets of fake news featuring negative sentiments (Frenkel et al., 2020; Pennycook et al., 2020). In real news, the pattern is reversed, and positive sentiments are much more frequent than negative sentiments (46.45% vs. 35.20%). Considering that real news is spread among reliable news channels, we can conclude that reliable news channels express news with positive sentiments so as not to hurt their audience psychologically and mentally.

The eight scores for the eight emotions of anger, anticipation, disgust, fear, joy, sadness, surprise, and trust were extracted from the NRC emotion lexicon for every tweet. Each text was assigned the emotion with the highest score. Table 4 and Fig. 5 include more detailed information about the emotion distribution.

The NRC lexicon provides scores for each emotion. Therefore, the intensities of emotions can also be compared. Table 5 shows the average score of each emotion for the two classes, fake and real news.

A two-sample *t*-test was performed using the pingouin (PyPI) statistical package in Python (Vallat, 2018) to determine whether the difference between the two groups was significant (Tables 6 and 7).

As shown in Table 6, the *P* values indicate that the differences in fear, anger, trust, surprise, disgust, and anticipation were significant; however, for sadness and joy, the difference between the two groups of fake and real news was not significant. Considering the statistics provided in Tables 4, 5, and Fig. 5, the following conclusions can be drawn:

- Anger, disgust, and fear are more commonly elicited in fake news than in real news.
- Anticipation and surprise are more commonly elicited in real news than in fake news.
- Fear is the most commonly elicited emotion elicited in both fake and real news.
- Trust is the second most commonly elicited emotion in fake and real news.

The most significant differences were observed for trust, fear, and anticipation (5.92%, 5.33%, and 3.05%, respectively). The differences between fake and real news in terms of joy and sadness were not significant.

In terms of intensity, based on Table 5,

- Fear is the mainly elicited emotion in both fake and real news; however, fake news has a higher fear intensity score than does real news.
- Trust is the second most commonly elicited emotion in two categories—real and fake—but is more powerful in real news.
- Positive emotions, such as anticipation, surprise, and trust, are more strongly elicited in real news than in fake news.
- Anger, disgust, and fear are among the stronger emotions elicited by fake news. Joy and sadness are elicited in both classes almost equally.

During the COVID-19 pandemic, fake news disseminators seized the opportunity to create fearful messages aligned with their objectives. The existence of fear in real news is also not surprising because of the extraordinary circumstances of the pandemic. The most crucial point of the analysis is the significant presence of negative emotions elicited by fake news. This observation confirms our hypothesis that fake news elicits extremely negative emotions. Positive emotions such as anticipation, joy, and surprise are elicited more often in real news than in fake news, which also aligns with our hypothesis. The largest differences in elicited emotions are as follows: trust, fear, and anticipation.

We used nine features for every tweet in the dataset: sentiment and eight scores for every emotion and sentiment in every tweet. These features were utilized for supervised machine learning fake news detection models. A schematic explanation of the models is given in Fig. 6. The dataset was divided into training and test sets, with an 80%–20% split. The scikit-learn random forest, SVM, and Naïve Bayes machine learning models with default hyperparameters were implemented using emotion features to detect fake news in nonnumerical data. Then, we compared the prediction power of the models with that of models without these features. The performance metrics of the models, such as accuracy, precision, recall, and F1-score, are given in Table 7.

Table 4 Emotion distribution based on the NRC lexicon (Farhoundinia, 2023).

Label	Anger (%)	Anticipation (%)	Disgust (%)	Fear (%)	Joy (%)	Sadness (%)	Surprise (%)	Trust (%)
Fake	3.70	4.51	1.18	66.81	1.42	3.80	1.20	17.38
Real	2.12	9.93	0.22	50.67	1.50	6.38	1.58	27.14
Total	2.87	7.35	0.68	58.36	1.46	5.39	1.40	22.49
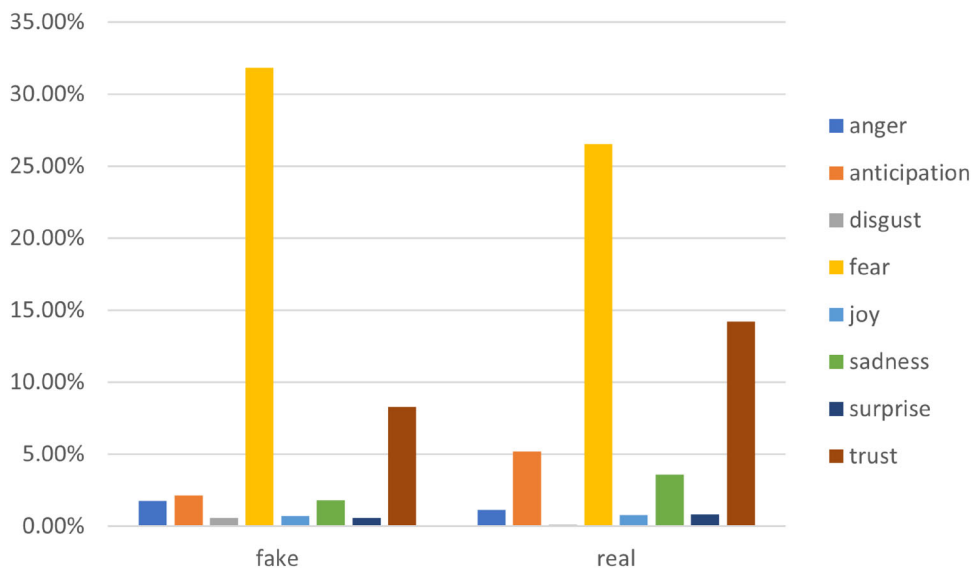
Fig. 5 Visual distribution of emotions. The figure depicts more detailed information about the emotion distribution.

When joy and sadness were removed from the models, the accuracy decreased. Thus, the models performed better when all the features were included (see Table C.1. *Feature correlation scores* in Supplementary Information). The results confirmed that elicited emotions can help identify fake and real news. Adding emotion features to the detection models significantly increased the performance metrics. Figure 7 presents the importance of the emotion features used in the random forest model.

In the random forest classifier, the predominant attributes were anticipation, trust, and fear. The difference in the emotion distribution between the two classes of fake and real news was also more considerable for anticipation, trust, and fear. It can be claimed that fear, trust, and anticipation emotions have good differentiating power between fake and real news.

BERT was the other model that was employed for the task of fake news detection using emotion features. The BERT model includes a number of preprocessing stages. The text input is segmented using the BERT tokenizer, with sequence truncation and padding ensuring that the length does not exceed 128 tokens, a reduction from the usual 512 tokens due to constraints on computing resources. The optimization process utilized the AdamW optimizer with a set learning rate of 0.00001. To ascertain the best number of training cycles, a 5-fold cross-validation method was applied, which established that three epochs were optimal. The training phase consisted of three unique epochs. The model was executed on Google Colab using Python, a popular programming language. The model was evaluated with the test set after training. Table 8 shows the performance of the BERT model with and without using emotions as features.

The results indicate that adding emotion features had a positive impact on the performance of the random forest, SVM, and BERT models; however, the naïve Bayes model achieved better performance without adding emotion features.

Discussion and limitations

This research makes a substantial impact on the domain of detecting fake news. The goal was to explore the range of sentiments and emotional responses linked to both real and fake news in pursuit of fulfilling the research aims and addressing the posed inquiries. By identifying the emotions provoked as key indicators of fake news, this study adds valuable insights to the existing corpus of related scholarly work.

Our research revealed that fake news triggers a higher incidence of negative emotions compared to real news. Sentiment analysis indicated that creators of fake news on social media platforms tend to invoke more negative sentiments than positive ones, whereas real news generally elicits more positive sentiments than negative ones. We extracted eight emotions—anger, anticipation, disgust, fear, joy, sadness, surprise, and trust—from each tweet analyzed. Negative and potent emotions such as fear, disgust, and anger were more frequently found elicited in fake news, in contrast to real news, which was more likely to arouse lighter and positive emotions such as anticipation, joy, and surprise. The difference in emotional response extended beyond the range of emotions to their intensity, with negative feelings like fear, anger, and disgust being more pronounced in fake news. We suggest that the inclusion of emotional analysis in the development of automated fake news detection algorithms could improve the effectiveness of the machine learning and deep learning models designed for fake news detection in this study.

Due to negativity bias (Baumeister et al., 2001), bad news, emotions, and feedback tend to have a more outsized influence than positive experiences. This suggests that humans are more likely to assign greater weight to negative events over positive ones (Lewicka et al., 1992). Our findings indicate that similar effects are included in social media user behavior, such as sharing and retweeting. Furthermore, the addition of emotional features

Table 5 Emotion intensity scores (Farhoundinia, 2023).

Label	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
Fake	0.033	0.023	0.025	0.097	0.028	0.064	0.018	0.097
Real	0.020	0.028	0.015	0.076	0.027	0.064	0.022	0.126

Table 6 *T*-test results (Farhoundinia, 2023).

	P-value	CI 95%	Cohen-*d*	Power
Fear	6.57E−12	[0.01, 0.03]	0.149	1.000
Anger	4.17E−16	[0.01, 0.02]	0.176	1.000
Trust	8.74E−13	[−0.04, −0.02]	0.155	1.000
Surprise	0.007362	[−0.01, −0.0]	0.058	0.764
Sadness	0.984772	[−0.0, 0.0]	0.000	0.050
Disgust	2.16E−14	[0.01, 0.01]	0.166	1.000
Joy	0.318163	[−0.0, 0.0]	0.022	0.170
Anticipation	1.86E−39	[−0.06, −0.04]	0.286	1.000

Table 7 Performance metrics of the FND models (Farhoundinia, 2023).

	Accuracy	Precision	Recall	F1-score
Fake news detection model with emotion features				
Random forest	0.81	0.85	0.94	0.89
Naïve Bayes	0.49	0.69	0.53	0.69
SVM	0.76	0.74	0.95	0.85
Fake news detection model without emotion features				
Random forest	0.79	0.87	0.88	0.87
Naïve Bayes	0.66	0.70	0.91	0.80
SVM	0.71	0.71	0.94	0.83

to the fake news detection models was found to improve their performance, providing an opportunity to investigate their moderating effects on fake news dissemination in future research.

The majority of the current research on identifying fake news involves analyzing the social environment and news content (Amer et al., 2022; Jarrahi and Safari, 2023; Raza and Ding, 2022). Despite its possible importance, the investigation of emotional data has not received sufficient attention in the past (Ajao et al., 2019). Although sentiment in fake news has been studied in the literature, earlier studies mostly neglected a detailed examination of certain emotions. Dey et al. (2018) contributed to this field by revealing a general tendency toward negativity in fake news. Their results support our research and offer evidence for the persistent predominance of negative emotions elicited by fake news. Dey et al. (2018) also found that trustworthy tweets, on the other hand, tended to be neutral or positive in sentiment, highlighting the significance of sentiment polarity in identifying trustworthy information.

Expanding upon this sentiment-focused perspective, Cui et al. (2019) observed a significant disparity in the sentiment polarity of comments on fake news as opposed to real news. Their research emphasized the clear emotional undertones in user reactions to false material, highlighting the importance of elicited emotions in the context of fake news. Similarly, Dai et al. (2020) analyzed false health news and revealed a tendency for social media replies to real news to be marked by a more upbeat tone. These comparative findings highlight how elicited emotions play a complex role in influencing how people engage with real and fake news.

Our analysis revealed that the emotions conveyed in fake tweets during the COVID-19 pandemic are in line with the more general trends found in other studies on fake news. However, our research extends beyond that of current studies by offering detailed insights into the precise distribution and strength of emotions elicited by fake tweets. This detailed research closes a significant gap in the body of literature by adding a fresh perspective on our knowledge of emotional dynamics in the context of disseminating false information. Our research contributes significantly to the current discussion on fake news identification by highlighting these comparative aspects and illuminating both recurring themes and previously undiscovered aspects of emotional data in the age of misleading information.

The present analysis was performed with a COVID-19 Twitter dataset, which does not cover the whole period of the pandemic. A complementary study on a dataset that covers a wider time interval might yield more generalizable findings, while our study represents a new effort in the field. In this research, the elicited emotions of fake and real news were compared, and the emotion with the highest score was assigned to each tweet, while an alternative method could be to compare the emotion score intervals for fake and real news. The performance of detection models could be further improved by using pretrained emotion models and adding additional emotion features to the models. In a future study, our hypothesis that "fake news and real news are different in terms of elicited emotions, and fake news elicits more negative emotions" could be examined in an experimental field study. Additionally, the premises and suppositions underlying this study could be tested in emergency scenarios beyond the COVID-19 context to enhance the breadth of crisis readiness.

The field of fake news research is interdisciplinary, drawing on the expertise of scholars from various domains who can contribute significantly by formulating pertinent research questions. Psychologists and social scientists have the opportunity to delve into the motivations and objectives behind the creators of fake news. Scholars in management can offer strategic insights for organizations to deploy in countering the spread of fake news. Legislators are in a position to draft laws that effectively stem the flow of fake news across social media channels. In addition, the combined efforts of researchers from other academic backgrounds can make substantial additions to the existing literature on fake news.

Conclusion

The aim of this research was to propose novel attributes for current fake news identification techniques and to explore the emotional and sentiment distinctions between fake news and real news. This study was designed to tackle the subsequent research questions: 1. How do the sentiments associated with real news and fake news differ? 2. How do the emotions elicited by fake news differ from those elicited by real news? 3. What particular elicited emotions are most prevalent in fake news? 4. How could these elicited emotions be used to recognize fake news on social media? To answer these research questions, we thoroughly examined tweets related to COVID-19. We employed a comprehensive strategy, integrating lexicons such as Vader, TextBlob, and SentiWordNet together with machine learning models, including random forest, naïve Bayes, and SVM, as well as a deep learning model named BERT. We first performed sentiment analysis using the lexicons. Fake news elicited more negative sentiments, supporting the idea that disseminators use extreme
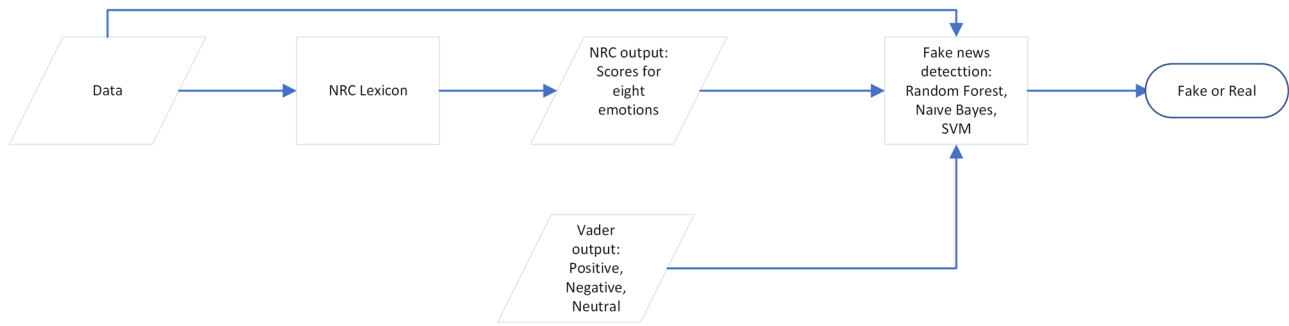
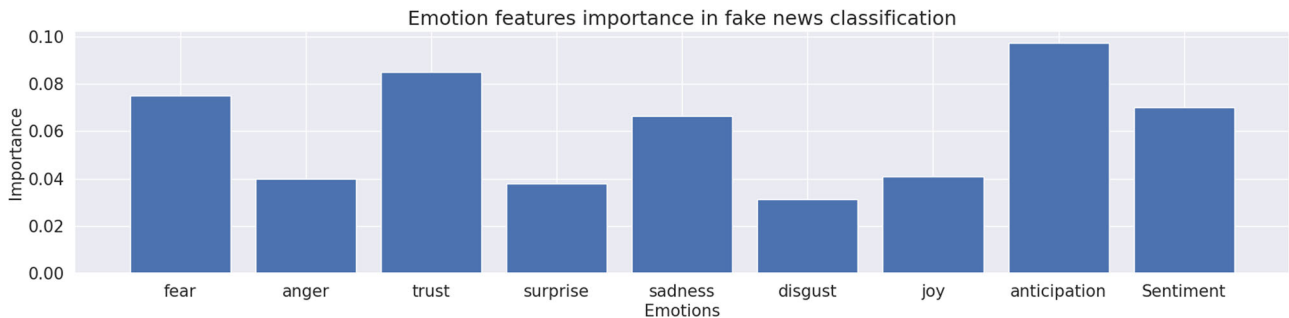**Fig. 6 Fake news detection model.** The figure exhibits a schematic explanation of the model.

Fig. 7 Random Forest feature importance. The figure illustrates the importance of the emotion features used in the Random Forest model.

Table 8 BERT results (Farhoundinia, 2023).

Performance metrics of the BERT model with and without emotion scores

	Accuracy	Precision	Recall	F1-score
With emotion scores	0.972	0.983	0.970	0.976
Without emotion scores	0.961	0.981	0.956	0.967

negativity to attract attention. Real news elicited more positive sentiments, as expected from trustworthy news channels. For fake news, there was a greater prevalence of negative emotions, including fear, disgust, and anger, while for real news, there was a greater frequency of positive emotions, such as anticipation, joy, and surprise. The intensity of these emotions further differentiated fake and real news, with fear being the most dominant emotion in both categories. We applied machine learning models (random forest, naïve Bayes, SVM) and a deep learning model (BERT) to detect fake news using sentiment and emotion features. The models demonstrated improved accuracy when incorporating emotion features. Anticipation, trust, and fear emerged as significant differentiators between fake and real news, according to the random forest feature importance analysis.

The findings of this research could lead to reliable resources for communicators, managers, marketers, psychologists, sociologists, and crisis and social media researchers to further explain social media behavior and contribute to the existing fake news detection approaches. The main contribution of this study is the introduction of emotions as a role-playing feature in fake news detection and the explanation of how specific elicited emotions differ between fake and real news. The elicited emotions extracted from social media during a crisis such as the COVID-19 pandemic could not only be an important variable for detecting fake news but also provide a general overview of the dominant emotions among individuals and the mental health of society during such a crisis. Investigating and

extracting further features of fake news has the potential to improve the identification of fake news and may allow for the implementation of preventive measures. Furthermore, the suggested methodology could be applied to detecting fake news in fields such as politics, sports, and advertising. We expect to observe a similar impact of emotions on other topics as well.

Data availability

The datasets analyzed during the current study are available in the Zenodo repository: https://doi.org/10.5281/zenodo.10951346.

References

Agarwal S, Farid H, El-Gaaly T, Lim S-N (2020) Detecting Deep-Fake Videos from Appearance and Behavior. 2020 IEEE International Workshop on Information Forensics and Security (WIFS), 1–6. https://doi.org/10.1109/WIFS49906.2020.9360904

Ainapure BS, Pise RN, Reddy P, Appasani B, Srinivasulu A, Khan MS, Bizon N (2023) Sentiment analysis of COVID-19 tweets using deep learning and lexicon-based approaches. Sustainability 15(3):2573. https://doi.org/10.3390/su15032573

Ajao O, Bhowmik D, Zargari S (2019) Sentiment Aware Fake News Detection on Online Social Networks. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2507–2511. https://doi.org/10.1109/ICASSP.2019.8683170

Al-Rawi A, Groshek J, Zhang L (2019) What the fake? Assessing the extent of networked political spamming and bots in the propagation of# fakenews on Twitter. Online Inf Rev 43(1):53–71. https://doi.org/10.1108/OIR-02-2018-0065

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–236. https://doi.org/10.1257/jep.31.2.211

Amer E, Kwak K-S, El-Sappagh S (2022) Context-based fake news detection model relying on deep learning models. Electronics (Basel) 11(8):1255. https://doi.org/10.3390/electronics11081255

Apuke OD, Omar B (2020) User motivation in fake news sharing during the COVID-19 pandemic: an application of the uses and gratification theory. Online Inf Rev 45(1):220–239. https://doi.org/10.1108/OIR-03-2020-0116

Baccarella CV, Wagner TF, Kietzmann JH, McCarthy IP (2018) Social media? It's serious! Understanding the dark side of social media. Eur Manag J 36(4):431–438. https://doi.org/10.1016/j.emj.2018.07.002

Baccarella CV, Wagner TF, Kietzmann JH, McCarthy IP (2020) Averting the rise of the dark side of social media: the role of sensitization and regulation. Eur Manag J 38(1):3–6. https://doi.org/10.1016/j.emj.2019.12.011

Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. Rev Gen Psychol 5(4):323–370. https://doi.org/10.1037/1089-2680.5.4.323

Berthon PR, Pitt LF (2018) Brands, truthiness and post-fact: managing brands in a post-rational world. J Macromark 38(2):218–227. https://doi.org/10.1177/0276146718755869

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Carlson M (2020) Fake news as an informational moral panic: the symbolic deviancy of social media during the 2016 US presidential election. Inf Commun Soc 23(3):374–388. https://doi.org/10.1080/1369118X.2018.1505934

Chua AYK, Banerjee S (2018) Intentions to trust and share online health rumors: an experiment with medical professionals. Comput Hum Behav 87:1–9. https://doi.org/10.1016/j.chb.2018.05.021

Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M (2021) The echo chamber effect on social media. Proc Natl Acad Sci USA 118(9). https://doi.org/10.1073/pnas.2023301118

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/BF00994018

Cui L, Wang S, Lee D (2019) SAME: sentiment-aware multi-modal embedding for detecting fake news. 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 41–48. https://doi.org/10.1145/3341161.3342894

Dai E, Sun Y, Wang S (2020) Ginger cannot cure cancer: Battling fake health news with a comprehensive data repository. In Proceedings of the 14th International AAAI Conference on Web and Social Media, ICWSM 2020 (pp. 853–862). (Proceedings of the 14th International AAAI Conference on Web and Social Media, ICWSM 2020). AAAI press

de Regt A, Montecchi M, Lord Ferguson S (2020) A false image of health: how fake news and pseudo-facts spread in the health and beauty industry. J Product Brand Manag 29(2):168–179. https://doi.org/10.1108/JPBM-12-2018-2180

Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint. https://doi.org/10.48550/arXiv.1810.04805

Dey A, Rafi RZ, Parash SH, Arko SK, Chakrabarty A (2018) Fake news pattern recognition using linguistic analysis. Paper presented at the 2018 joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan. pp. 305–309

Er MF, Yılmaz YB (2023) Which emotions of social media users lead to dissemination of fake news: sentiment analysis towards Covid-19 vaccine. J Adv Res Nat Appl Sci 9(1):107–126. https://doi.org/10.28979/jarnas.1087772

Esuli A, Sebastiani F (2006) Sentiwordnet: A publicly available lexical resource for opinion mining. Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)

Farhoundinia B (2023). Analyzing effects of emotions on fake news detection: a COVID-19 case study. PhD Thesis, Sabanci Graduate Business School, Sabanci University

Farhoudinia B, Ozturkcan S, Kasap N (2023) Fake news in business and management literature: a systematic review of definitions, theories, methods and implications. Aslib J Inf Manag https://doi.org/10.1108/AJIM-09-2022-0418

Faustini PHA, Covões TF (2020) Fake news detection in multiple platforms and languages. Expert Syst Appl 158:113503. https://doi.org/10.1016/j.eswa.2020.113503

Frenkel S, Davey A, Zhong R (2020) Surge of virus misinformation stumps Facebook and Twitter. N Y Times (Online) https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html

Giglietto F, Iannelli L, Valeriani A, Rossi L (2019) 'Fake news' is the invention of a liar: how false information circulates within the hybrid news system. Curr Sociol 67(4):625–642. https://doi.org/10.1177/0011392119837536

Hamed SK, Ab Aziz MJ, Yaakub MR (2023) Fake news detection model on social media by leveraging sentiment analysis of news content and emotion analysis of users' comments. Sensors (Basel, Switzerland) 23(4):1748. https://doi.org/10.3390/s23041748

Hutto C, Gilbert E (2014) VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Iwendi C, Mohan S, khan S, Ibeke E, Ahmadian A, Ciano T (2022) Covid-19 fake news sentiment analysis. Comput Electr Eng 101:107967–107967. https://doi.org/10.1016/j.compeleceng.2022.107967

Jarrahi A, Safari L (2023) Evaluating the effectiveness of publishers' features in fake news detection on social media. Multimed Tools Appl 82(2):2913–2939. https://doi.org/10.1007/s11042-022-12668-8

Kahneman D (2011) Thinking, fast and slow, 1st edn. Farrar, Straus and Giroux

Kaliyar RK, Goswami A, Narang P (2021) FakeBERT: fake news detection in social media with a BERT-based deep learning approach. Multimed Tools Appl 80(8):11765–11788. https://doi.org/10.1007/s11042-020-10183-2

Kim A, Dennis AR (2019) Says who? The effects of presentation format and source rating on fake news in social media. MIS Q 43(3):1025–1039. https://doi.org/10.25300/MISQ/2019/15188

Kumar A, Bezawada R, Rishika R, Janakiraman R, Kannan PK (2016) From social to sale: the effects of firm-generated content in social media on customer behavior. J Mark 80(1):7–25. https://doi.org/10.1509/jm.14.0249

Lewicka M, Czapinski J, Peeters G (1992) Positive-negative asymmetry or when the heart needs a reason. Eur J Soc Psychol 22(5):425–434. https://doi.org/10.1002/ejsp.2420220502

Loria S (2018) Textblob documentation. Release 0.15, 2 accessible at https://readthedocs.org/projects/textblob/downloads/pdf/latest/. available at http://citebay.com/how-to-cite/textblob/

Meel P, Vishwakarma DK (2020) Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 153:112986. https://doi.org/10.1016/j.eswa.2019.112986

Mercer J (2010) Emotional beliefs. Int Organ 64(1):1–31. https://www.jstor.org/stable/40607979

Mohammad SM, Turney PD (2013) Crowdsourcing a word–emotion association lexicon. Comput Intell 29(3):436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Moravec PL, Kim A, Dennis AR (2020) Appealing to sense and sensibility: system 1 and system 2 interventions for fake news on social media. Inf Syst Res 31(3):987–1006. https://doi.org/10.1287/isre.2020.0927

Mourad A, Srour A, Harmanai H, Jenainati C, Arafeh M (2020) Critical impact of social networks infodemic on defeating coronavirus COVID-19 pandemic: Twitter-based study and research directions. IEEE Trans Netw Serv Manag 17(4):2145–2155. https://doi.org/10.1109/TNSM.2020.3031034

Ongsulee P (2017) Artificial intelligence, machine learning and deep learning. Paper presented at the 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE)

Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A 540:123174. https://doi.org/10.1016/j.physa.2019.123174

Patwa P, Sharma S, Pykl S, Guptha V, Kumari G, Akhtar MS, Ekbal A, Das A, Chakraborty T (2021) Fighting an Infodemic: COVID-19 fake news dataset. In: Combating online hostile posts in regional languages during emergency situation. Cham, Springer International Publishing

Păvăloaia V-D, Teodor E-M, Fotache D, Danileţ M (2019) Opinion mining on social media data: sentiment analysis of user preferences. Sustainability 11(16):4459. https://doi.org/10.3390/su11164459

Pawar KK, Shrishrimal PP, Deshmukh RR (2015) Twitter sentiment analysis: a review. Int J Sci Eng Res 6(4):957–964

Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG (2020) Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol Sci 31(7):770–780. https://doi.org/10.1177/0956797620939054

Pennycook G, Rand DG (2020) Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. J Personal 88(2):185–200. https://doi.org/10.1111/jopy.12476

Peterson M (2019) A high-speed world with fake news: brand managers take warning. J Product Brand Manag 29(2):234–245. https://doi.org/10.1108/JPBM-12-2018-2163

Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Plutchik R, Kellerman H (eds) Theories of emotion (3–33): Elsevier. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

Rajamma RK, Paswan A, Spears N (2019) User-generated content (UGC) misclassification and its effects. J Consum Mark 37(2):125–138. https://doi.org/10.1108/JCM-08-2018-2819

Raza S, Ding C (2022) Fake news detection based on news content and social contexts: a transformer-based approach. Int J Data Sci Anal 13(4):335–362. https://doi.org/10.1007/s41060-021-00302-z

Salakhutdinov R, Tenenbaum JB, Torralba A (2012) Learning with hierarchical-deep models. IEEE Trans Pattern Anal Mach Intell 35(8):1958–1971. https://doi.org/10.1109/TPAMI.2012.269

Silverman C (2016) This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook. BuzzFeed News 16. https://www.

buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

Suter V, Shahrezaye M, Meckel M (2022) COVID-19 Induced misinformation on YouTube: an analysis of user commentary. Front Political Sci 4:849763. https://doi.org/10.3389/fpos.2022.849763

Talwar S, Dhir A, Kaur P, Zafar N, Alrasheedy M (2019) Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. J Retail Consum Serv 51:72–82. https://doi.org/10.1016/j.jretconser.2019.05.026

Vallat R (2018) Pingouin: statistics in Python. J Open Source Softw 3(31):1026. https://doi.org/10.21105/joss.01026

Vasist PN, Sebastian M (2022) Tackling the infodemic during a pandemic: A comparative study on algorithms to deal with thematically heterogeneous fake news. Int J Inf Manag Data Insights 2(2):100133. https://doi.org/10.1016/j.jjimei.2022.100133

Vinodhini G, Chandrasekaran R (2012) Sentiment analysis and opinion mining: a survey. Int J Adv Res Comput Sci Softw Eng 2(6):282–292

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359(6380):1146–1151. https://doi.org/10.1126/science.aap9559

Wang Y, McKee M, Torbica A, Stuckler D (2019) Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 240:112552. https://doi.org/10.1016/j.socscimed.2019.112552

Wankhade M, Rao ACS, Kulkarni C (2022) A survey on sentiment analysis methods, applications, and challenges. Artif Intell Rev 55(7):5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Whiting A, Williams D (2013) Why people use social media: a uses and gratifications approach. Qual Mark Res 16(4):362–369. https://doi.org/10.1108/QMR-06-2013-0041

Zhang H (2004) The optimality of naive Bayes. Aa 1(2):3

Zhou X, Zafarani R (2019) Network-based fake news detection: A pattern-driven approach. ACM SIGKDD Explor Newsl 21(2):48–60. https://doi.org/10.1145/3373464.3373473

Zhou X, Zafarani R, Shu K, Liu H (2019) Fake news: Fundamental theories, detection strategies and challenges. Paper presented at the Proceedings of the twelfth ACM international conference on web search and data mining. https://doi.org/10.1145/3289600.3291382

Author contributions

Bahareh Farhoudinia (first author) conducted the research, retrieved the open access data collected by other researchers, conducted the analysis, and drafted the manuscript as part of her PhD thesis successfully completed at Sabancı University in the year 2023. Selcen Ozturkcan (second author and PhD co-advisor) provided extensive guidance throughout the research process, co-wrote sections of the manuscript, and offered critical feedback on the manuscript. Nihat Kasap (third author and PhD main advisor) oversaw the overall project and provided valuable feedback on the manuscript.

Funding

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

Informed consent was not required as the study did not involve a design that requires consent.

Additional information

Correspondence and requests for materials should be addressed to Selcen Ozturkcan.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.