Computer Science and Information Systems 20(4):1439–1457 https://doi.org/10.2298/CSIS230315048T

A Framework for Fake News Detection Based on the Wisdom of Crowds and the Ensemble Learning Model

Hai Bang Truong[1,2] and Van Cuong Tran[3,⋆]

[1] Faculty of Computer Science, University of Information Technology
[2] Vietnam National University
700000 Ho Chi Minh City, Vietnam
bangth@uit.edu.vn
[3] Faculty of Engineering and Information Technology, Quang Binh University
47100 Quang Binh, Vietnam
vancuongqbuni@gmail.com

Abstract. Nowadays, the rapid development of social networks has led to the proliferation of social news. However, the spreading of fake news is a critical issue. Fake news is news written to intentionally misinform or deceive readers. News on social networks is short and lacks context. This makes it difficult for detecting fake news based on shared content. In this paper, we propose an ensemble classification model to detect fake news based on exploiting the wisdom of crowds. The social interactions and the user's credibility are mined to automatically detect fake news on Twitter without considering news content. The proposed method extracts the features from a Twitter dataset and then a voting ensemble classifier comprising three classifiers namely, Support Vector Machine (SVM), Naive Bayes, and Softmax is used to classify news into two categories which are fake and real news. The experiments on real datasets achieved the highest F1 score of 78.8% which was better than the baseline by 6.8%. The proposed method significantly improved the accuracy of fake news detection in comparison to other methods.

Keywords: Fake news detection, Social interaction, User's credibility, User's opinion

1. Introduction

Nowadays, social network services have developed rapidly and become a popular information channel for the community. Social networks have an important role in spreading news and people can rate the news easily. In 2021, about half of U.S. adults often or sometimes got news from social media[4]. Social network services have attracted a lot of people who publish and share news and personal opinions every day. Social networks with multiple useful characteristics, like faster transformation and less expensive, have become an important channel of communication and information sharing. However, due to the increasing popularity of social networks and the ease of posting news on those platforms, it becomes an ideal way of communicating and spreading fake news. Fake news has negative effects on society. It could cause political instability [5][14]. In recent years,

⋆ Corresponding author

[4] https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/

social media and technology companies face criticism for not doing enough to stem the flow of fake news on social network platforms.

Fact-checking is a process that seeks related information to verify factual information and to promote the veracity and correctness of news. In reality, fact-checking action can be conducted before or after the news is posted. Internal fact-checking is such checking done in-house by the publisher or system and the news is analyzed by a third party called external fact-checking. Preventing the spread of misinformation can be done through human intervention by verifying the authenticity of the content. Using the international fact-checking network and manual fact-checking websites such as washingtonpost.com, snopes.com, Politifact.com, etc. are the methods commonly used in practice [10]. The assessment of experts on fact-checking websites can bring forward the reader's realistic viewpoints about related information. Figure 1. is a flowchart of a fact-checking system for the news on social media. Fact-checking websites is an efficient way to verify the authenticity of the information, especially news on social media. However, handling large volumes of data is a great obstacle. It is very difficult for humans to evaluate all up-to-date information on social media. Automatic fact-checking is a promising method to overcome this problem.

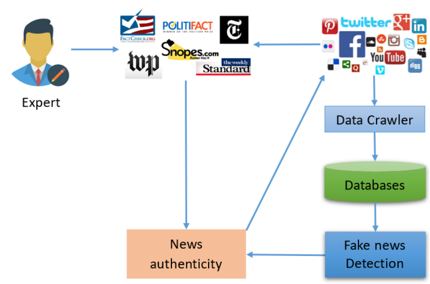

Fig. 1. The flowchart of a fact-checking system

Fake news detection topics have got a lot of attention from researchers. There are different fake news detection approaches have been proposed in the literature. The most common approach is news content analysis [2], [18], [20], [25]. The language features are analyzed to determine the authenticity of the news. Generally, news content analysis is suitable for formal and long news. However, social news is quite different from formal and long news. The short and abbreviation characteristics are the big challenges for the algorithms applying to formal news. Therefore, the approaches combining the content and news context for fake news detection have gotten a lot of attention from researchers [26], [27]. Using the user's profile and news content features to analyze the behavior of the user is an approach receiving much attention currently [21]. Machine learning approaches have been applied quite effectively for detecting fake news [5]. SVM, Random forest, and logistic regression models have been used and achieved high accuracy in the fake news detection experiments [6], [12], [16], [19]. Ensemble methods aim at improving the accuracy of individual models by combining multiple models. Generally, the combined models can increase performance significantly [10]. Social networking service is an online social media platform in which people can share personal viewpoints. The tremendous growth of social networks has attracted research communities to discover collective intelligence by mining social data [3], [4]. Collective intelligence has been enhanced under the emergence of social networks [8]. Exploiting the wisdom of crowds is another research direction that has received a lot of attention to detect fake news on social networks. In this approach, the authenticity of the news is identified by analyzing the interaction data from users [13], [30]. Combining social interactions can significantly improve the performance of fake news detection problems on social networks. The user's interactions such as comments are an important feature. Figure 2. is an example of tweet content and comments representing the user opinion on Twitter. The misinformation can be detected by exploiting the user's comments.

Fig. 2. News and comments representing the user opinion on Twitter

Limited social-related information is a challenge for fake news detection. First, newly emerged events often have not been stored in existing knowledge-based so it is difficult to be inferred. Second, the features of fake news in the past may not be available in the present, especially due to the constant evolution of misleading writing styles. Finally, limited information may reduce the performance of machine learning models. Analyzing short messages to detect fake news is a difficult task. Some characteristics such as up-to-date information, the lack of context, misspellings, and abbreviations are challenges for machine learning algorithms. The user insight extracted from engagement data is valuable information for fake news detection problems. The user's viewpoints on social networks, especially from high-credibility users, could help to identify fake news effectively. In this study, we present a fake news detection method based on exploiting the wisdom of crowds.

We focus only on analyzing the user's viewpoints to detect fake news without considering news content. The proposed method exploits the user data on Twitter such as the user's profile and interaction data of news. Three features are extracted from the dataset to classify news that are the user's opinion, the user's credibility level, and the exhortation level. The users' viewpoints are useful information to determine the veracity of the news. The users' insights and opinions are presented by actions such as comments, likes, and shares. Exploiting interaction data can reveal the user's viewpoint toward the news. However, the accuracy of opinions from users is different. In our work, we assessed the accuracy of viewpoints based on the user's credibility level. The user's profiles including account information and historical data were analyzed to determine their credibility level. The user's news posted in the history was also matched against the fact-checking database to discover fake news. Fake news in history is important in determining a user's credibility level. The user's trustworthiness level will be reduced if they have previously posted fake news. The historical data provides a better assessment of a user's trustworthiness level and enhance the accuracy of classifying a user's opinion.

Besides, spreading news actions and relationships between users were also exploited to determine the user's credibility level. To overcome the weakness of the single model, we used an ensemble learning model that combines the predictions from three models namely, Support Vector Machine (SVM), Naive Bayes, and Softmax. The proposed method classified news into two categories which are fake and real news. The proposed method experimented on datasets collected from Twitter. The experiments revealed that the proposed method significantly improved the performance of fake news detection problems. The highest F1 score was 78.8% and it was better than the baseline by 6.8%. The experiments demonstrated the effectiveness of using features extracted from social interactions and user profiles.

The major contributions of this paper are as follows:

– We propose to exploit the users' interaction data to detect fake news on social networks. We focus only on analyzing the user's profile and viewpoints to detect fake news without considering shared content.
– We applied the voting ensemble method to the extracted features to classify news on Twitter. The achieved results demonstrated the effectiveness of the proposed method in improving the performance of the fake news detection problems.

The remainder of the paper is structured as follows: Section 2 presents an overview of related works; Section 3 presents the problem statement and the proposed methodology; The experiments are presented in Section 4; Section 5 concludes the paper.

2. Related Works

Social media has developed rapidly and become a popular information channel for everyone. The users create news pies spreading them in the community. Misleading information propagating on social networks is a big challenge that has attracted much more attention from researchers in recent years. Automatic fake news detection is very essential on social media. In this section, we discuss some commonly used perspectives in the literature such as content-based, social context-based, knowledge-based, user's credibility, etc.

2.1. Content and Social Context-based Approach

Style-based concerning how fake news is written. Analyzing the content to detect fake news is a popular approach, especially for long and formal news. The vocabulary and writing style are important features to determine fake information [2], [20]. Pérez-Rosas et al. analyzed text to determine linguistic features that express fake information [18]. Their experiments achieved high accuracy in fake news detection. The linguistic characteristics change when people try to hide their lying writing style [2]. Afroz et al. proposed a method to identify deception in the stylistics of documents. The experiments on the real dataset achieved an F1 score of 96.6%. The linguistic characteristics are suitable for fake news detection in formal news. However, the linguistic characteristics of news on social media are usually different from formal news. The writing style of social news is informal. This is a reason why the methods of analyzing content are inefficient and reduce accuracy when applied to social news. Therefore, exploring auxiliary information has the potential to improve the performance of fake news detection problems.

Combining content and social context analysis is an appropriate method for fake news detection on social networks [24], [27]. Kai Shu proposed a tri-relationship of social context during spreading the news on social media [27]. They exploited the relationship among publishers, news, and users. The authors used a tri-relationship framework to model relations between the publishers and news, and interactions between the user and news to identify fake news. Kai Shu et al. also proposed to exploit the combination of content, social and temporal in the news spread ecosystem on social media [24]. They analyzed the network structure of spreading the news to discover fake news. The experiments demonstrated that the social context analysis methods overcome the disadvantages of traditional text analysis.

2.2. Knowledge-based Approach

Fake news detection based on analyzing content and writing style is a hard problem, especially with social media news. Knowledge-based approaches utilize the fact-checking method to compare news content with external sources to verify veracity. The fact-checking methods can be categorized as manual and automatic fact-checking [11]. The manual fact-checking method uses assessments from experts or crowdsourcing. The expert-based method relies on human experts working in specific domains (e.g, fact-checking websites like Snopes, PolitiFact, GossipCop, etc.). The expert-based method is reliable but it is time-consuming and inappropriate for the huge volume of data on social media. Crowdsourcing uses the wisdom of crowds to check the accuracy of news. The user can provide their discussion on news via a platform. Although crowdsourcing fact-checking is difficult to manage and has low reliability, it is more suitable than expert-based fact-checking [33]. On social networks, users can present their viewpoints about the news by leaving comments. Jin et al. exploited collective intelligence to detect fake news in microblogs by exploiting the user's conflicting viewpoints [13]. A topic model was used to discover the conflicting viewpoints of news and then, a propagation network with supporting or opposing relations was built to infer credibility. Their experiments demonstrated the effectiveness of the proposed method on a real dataset. Wei et al. applied a Bayesian aggregation model to relevant human and machine judgment extracted from news content and crowd judgment [30]. The proposed framework demonstrated the effectiveness of exploiting crowd intelligence in fake news detection. Using ontologies in fake news detection is a promising research approach. Seddari et al. proposed a hybrid method for fake news detection that combines linguistic and knowledge-based [22]. They used ontologies to model fake news domain knowledge. Ontologies and linguistic features were used to distinguish fake news. Groza et al. also used ontology reasoning to detect deceptive information about COVID-19 [9]. The major challenge of fake news detection based on ontologies is the lack of scientific knowledge related to a specific category of news. Although knowledge-based approaches can achieve good results in fake news detection, they are not suitable for new topics without corresponding entries in any knowledge base. This approach is often combined with others to get better results.

2.3. Credibility-based Approach

The credibility-based approach investigates the credibility of creators, readers, and spreaders. The credibility characteristic can be extracted from the user's profile such as account information, news in the history, community, following users, etc. The user's credibility plays an important role in fake news detection problems. The users can propagate news in the community by posting, sharing, commenting, etc. The user with a low-credibility level has high probability to post or share fake news than others [1]. News posted on an unreliable website and shared by low-credibility users is more likely to be fake news than news posted by authoritative and credible users [32]. Yang et al. treated the authenticity of the news and the user's credibility as latent random variables [31]. The user's interactions on social networks were mined to determine their opinions towards the news. They used a Bayesian network model to capture the conditional dependencies among the authenticity of the news, the user opinions, and the user's credibility. Shao et al. analyzed the news spreading thousands of articles on Twitter during the 2016 U.S. presidential campaign and election [23]. They found evidence that bots played a big role in amplifying low-credibility news. The accounts actively spread content from low-credibility sources that are more likely to be bots.

In our previous study [29], we used the SVM model to classify news based on a set of four features that are the content, the user's credibility, the user opinion, and the exhortation level. The method is still based on analyzing the news content to determine fake news. In this study, we propose an improved method to detect fake news automatically by exploiting multiple features extracted from social interactions and the user's profile. The methodology and experiment results are presented in the next sections.

3. Proposed Methodology

In this section, we describe the problem of fake news detection from the wisdom of crowds and the features used to classify news. The proposed model is presented at the end of this section.

3.1. Problem Statement

In social networks, once a user posts a piece of news that spreads to the group of users following that user. Other users who read that news can present their viewpoints towards the news by emotional action or reposting or leaving comments about their opinions.

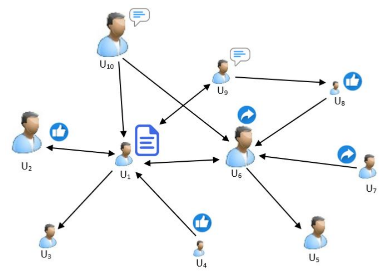

Fig. 3. The illustration of user's relationships on Twitter

Let $U = \{u_1, u_2, ..., u_n\}$ be a set of n users on a social network, $A = \{T_1, T_2, ..., T_m\}$ be a set of news posting by users in U.

For example, Figure 3. describes a simple graph of relationships among users on Twitter. In this graph, user u_1 has five followers $u_2, u_4, u_6, u_9, u_{10}$. Once user u_1 posts news that shows on the timeline of his/her followers. The users can present their viewpoints on that news through comments or emotional actions or retweets. Intuitively, each user has a credibility level to the community/group (e.g., politicians, journalists, and influencers are high credibility; new members and users having few friends are low credibilities). In this example, the high-trust users are drawn to a bigger size and a less size in the opposite case. The credibility level of each user is a parameter to evaluate the validity of their opinion. The comments from high-trust users are more valuable than those of low-trust users.

Given a targeted tweet T posted by user u_j. The news spreads to the community of users following u_j. In general, when followers read the news either they can leave comments expressing their viewpoints or they can present their opinions such as like, retweet, etc. Assuming that each user has a credibility level determined relying on his/her profile. A set of social interactions related to that news is collected to extract the features. Let's determine the authenticity of news if it is fake or real news by analyzing data from the user's social interactions related to the news and the user's profiles.

A tweet is a short text and lacks context to clearly understand its content. Therefore, detecting fake news only relying on its content usually obtains low accuracy. Generally, the reader has to have a wide knowledge of the related subject to be able to understand the inside meaning of the news. A deep understanding helps the reader to leave helpful comments that point out the authenticity of the news. In this study, we exploit the wisdom of crowds which is social interactions related to the news from users to detect fake news on social networks. A framework of the proposed method for fake news detection is described in Figure 4. It contains the following steps:

- First, crawling the interactive data related to the targeted tweet.
- Second, collecting user data on Twitter and measuring the user's credibility level.
- Third, determining the user opinions and exhortation level of the targeted tweet based on comments, likes, and retweets.
- Fourth, building a feature vector.
- Finally, an ensemble model is applied to the features to classify news.

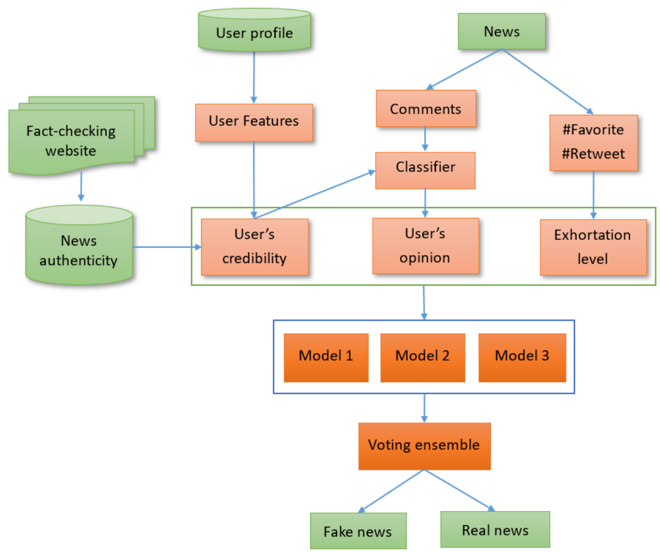

Fig. 4. A framework for detecting fake news based on exploiting the user's interactions

3.2. Feature Extraction

User's Credibility Feature The user's profiles and historical data contain valuable information to determine the user's credibility. Some properties of the user's profile should be considered such as created date, the number of followers, the number of followings, the total of tweets, the total of retweets, etc. These features reflect the seniority of the account, influence level, and the community's level of interest in their posts. Besides, the authenticity of news in history is an important feature to determine the user's reliability. An account that has published fake news is more likely to publish fake news in the future [26]. People who have posted fake news have less credible than people who have not posted fake news yet. The user's news in history needs to be analyzed to determine whether fake news or not and the number of published fake news includes in their profile.

The user's credibility feature is extracted as described in Algorithm 2. To determine the posted fake news in the history of the user, as described in Algorithm 1, the user's historical news is matched with the fact-checking source to find out the fake news. The user's features are a 6-dimensional vector U described as follows:

$$U = \{t, fo, fa, co, re, fn\}, \tag{1}$$

where t is an account age, fo is the ratio of the number of followers to the number of followings, fa is the ratio of the number of likes to the number of tweets, co is the ratio of the number of comments to the number of tweets, re is the ratio of the number of retweets to the number of tweets, fn is the percentage of published fake news to the number of tweets.

User Opinion Feature In social networks, users present viewpoints on news through actions such as emotion, comments, sharing, etc. The user's comments usually express their

Algorithm 1 Fact-checking phase for historical data

Input: Historical data of user, D
Fact-checking source, FS
Output: The percentage of published fake news to the number of news
function CHECK_NEWS_IN_HISTORY(D, FS)
$count \leftarrow 0$;
for $T \in D$ **do**
Pre-processing T;
if T in FS **then** ▷ news exists in the fake news database
$count \leftarrow count + 1$;
end if
end for
return $count/|D|$;
end function

understanding and opinion about news content. On Twitter, users can comment on any tweet even if they are not following the person who posted it. A tweet is a short message and lacks context so the users generally leave comments relying on their understanding of that content. Users have a wide knowledge and can analyze related problems to understand the news content. The comments from users provide useful information for fake news detection. They can leave comments to express their personal opinions. Therefore, these opinions can be a useful source for exploiting the collective intelligence on social networks to detect fake news.

Algorithm 2 Determining user's credibility

Input: User's profile and historical data of user, D
Fact-checking source, FS
Output: A feature vector expressing user's credibility, C
function USER_CREDIBILITY(D)
Determining t, fo, fa, co, re;
$fn \leftarrow Check\_news\_in\_history(D, FS)$; ▷ Algorithm 1
$C \leftarrow t \oplus fo \oplus fa \oplus co \oplus re \oplus fn$; ▷ Integrating features
return C;
end function

To exploit the user opinions, first, the comments related to a targeted tweet are collected. Then, data is cleaned by removing inessential elements such as hyperlinks, hashtags, symbols, mentions, and emoticons. To construct the comment vector, we use the implementation of doc2vec from the gensim package[5]. The comment vector Co is a k-dimensional vector denoted as follows:

$$Co = (w_1, w_2, ..., w_k) \tag{2}$$

A vector of the user opinion OV is constructed by concatenating the vector of the user's comment Co and the user's credibility C.

[5] https://radimrehurek.com/gensim/models/doc2vec.html

$$OV = Co \oplus C \tag{3}$$

Algorithm 3 Classify users' opinions

Input: Users' data interact with a target news, I
Users' profiles, D
Output: users' opinions, O
function USERS_OPINION(D, I)
$O \leftarrow \emptyset$;
for comments $t \in I$ **do**
Pre-processing t;
$Co \leftarrow$ Doc2vec(t);
$C \leftarrow$ User_credibility(D); ▷ Algorithm 2
$OV = Co \oplus C$;
$Op \leftarrow$ multiSVM(OV);
$O \leftarrow O \cup Op$;
end for
return O;
end function

The combination of the user's comment and his/her credibility can increase the efficiency of exploiting the user's opinion. Since the comments from high-credibility users are more valuable than those of low-credibility users. A multiSVM model[6] is utilized to classify the comments into three categories (i.e., $positive$, $neutral$, and $negative$). In which $negative$, $positive$ denote comments that rate the news as fake news, real news, respectively. The comments do not present the veracity of news classified as $neutral$. Since a tweet can receive many comments from different users. Therefore, the users' opinions toward a tweet are determined based on combining the opinions. Pseudocode describes the idea of user opinions extraction described in Algorithm 3. The users' opinions toward a tweet are presented as follows:

$$O = \{pos, neu, neg\}, \tag{4}$$

where pos is the percentage of positive comments, neu is the percentage of neutral comments, and neg is the percentage of negative comments.

Exhortation Feature The comments from users usually present their viewpoints toward news. Analyzing users' comments to know their opinions is helpful for fake news detection problems. Otherwise, the exhortation is another feature to enhance the user opinion feature. In general, users usually encourage the news by like or retweet actions. Once users are interested in the news they will click on the Like button to express their interest. To spread the news, they can also repost the news, called Retweet. Retweet is a way to spread the news to more people. In general, like and retweet actions express the user's agreement with the news. Since fake news is intentionally written to mislead readers to

[6] https://scikit-learn.org/stable/modules/svm.html

believe false information so the readers are easy to believe news content. Fake news has a great ability to attract users. The users usually encourage fake news by like or retweet actions to spread the news to the community. This is one reason that fake news spreads so quickly. Spreading rapidly combined with the user's comments is a good indicator of fake news detection. Since the user's tweet only shows on the timeline of followers. Therefore, in this work, we only consider the ratio of the number of likes to the number of followers ($like$) and the ratio of the number of retweets to the number of followers (ret). The news exhortation E is presented as:

$$E = \{like, ret\}. \tag{5}$$

3.3. Ensemble Model Selection

The ensemble learning approach helps improve machine learning results by combining several models to improve predictive performance compared to a single model. The ensemble learning approach is a widely-used in machine learning. A voting ensemble is an ensemble machine learning model that combines the prediction results from multiple models. The core idea is that the results obtained from a combination of models can be more accurate than any individual machine learning model in the group. In this technique, multiple classifiers are used to make predictions for each data point. The outputs of each classifier are passed into the voting classifier to choose the outputs based on the majority of voting. There are two different types of voting prediction for classification that are hard voting/majority voting and soft voting.

Algorithm 4 Feature vector construction

Input: A target news, T
User's profiles, D
User's data interact with target news, I
Output: A feature vector, F
function FEATURE_VECTOR(T, D, I)
$C \leftarrow$ User_credibility(D); ▷ Algorithm 2
$O \leftarrow$ Users_opinion(T, D, I); ▷ Algorithm 3
$E \leftarrow$ Exhortation_level(T, D, I); ▷ Section 3.2
$F \leftarrow C \oplus O \oplus E$; ▷ Integrating features
return F;
end function

The majority voting approach involves summing the predictions for each class. The majority voting is considered differently when weights associated with the different classifiers are equal or not. The predictions that we get from the majority of the models are used as the final prediction. The prediction of the ensemble classifier can be mathematically represented as the following:

$$\hat{y} = \underset{i}{\operatorname{argmax}} \sum_{j=1}^{m} w_j (C_j(x) = i), \tag{6}$$

where $C_j(x)$ is the predicted label of classifier C_j applied to sample x, x is a sample, w_j represents the weight associated with the prediction of the classifier C_j and $\hat{y}$ is a class label.

Soft voting involves averaging the predicted probabilities for each class. Classifiers can be assigned weights corresponding to the classifier's importance. The class label with the highest averaging of weighted probabilities wins the vote. The prediction of the ensemble classifier can be mathematically represented as the following:

$$\hat{y} = \underset{i}{\operatorname{argmax}} \frac{1}{m_{classifiers}} \sum_{j=1}^{m} w_j(C_j(x) = i), \quad (7)$$

where $C_j(x)$ is the predicted probability of classifier C_j applied to sample x.

Algorithm 5 Ensemble learning model for fake news detection

Input: A target news, T
Output: Veracity of T, V
  function FAKE_NEWS_DETECTION(T)
    Pre-processing T;
    Determining a set of users related to T (i.e., publisher, users have interacted with T);
    Collecting user's profile and historical data of users, D;
    Collecting social interaction data related to T (i.e., Comments, #Like, #Retweet), I;
    $F \leftarrow$ Feature_vector(T, D, I); ▷ Algorithm 4
    V = VotingClassifier(αNaiveBayes(F), γSVM(F), βSoftmax(F));
      ▷ α, γ, β is the weights assigned to classifiers
    return V;
  end function

In this study, we implement the voting ensemble model for fake news detection by using three supervised algorithms that are Naive Bayes, SVM, and Softmax Regression. We use the implementations of the models from the scikit-learn library[7]. The experiments were conducted with the adjusted parameters to test the extracted features and the proposed method, especially the weight representing the importance of classifiers. Algorithm 5 illustrates the proposed idea of the ensemble learning model. The interaction data and the users' profiles interacting with the news are collected to construct the feature vector as described in Section 3.2. The feature vectors are generated by combining the user's credibility, the user's opinion, and the exhortation feature. The steps of feature vector construction are presented in Algorithm 4. The models are trained individually to produce the classifiers and then applied to the feature vectors to achieve the prediction results.

[7] https://scikit-learn.org

4. Experiments

4.1. Evaluation metrics

In this experiment, we used precision and recall measurements to evaluate the results. Precision (P) and Recall (R) are calculated as follows:

$$P = \frac{|TruePositive|}{|TruePositive| + |FalsePositive|} \tag{8}$$

$$R = \frac{|TruePositive|}{|TruePositive| + |FalseNegative|} \tag{9}$$

where True Positive is the correctly predicted fake news, False Positive is the incorrectly predicted fake news, False Negative is the incorrectly predicted real news. The $F1$ score is the harmonic mean of P and R defined as follows:

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{10}$$

Table 1. Data statistics

	Quantity
#News	515
#Fake news	255
#Real news	260
#Publisher's profiles	357
#Comments	2532
#Negative	847
#Positive	840
#Neutral	845

4.2. Datasets

To carry out experiments with the proposed method, we explored data from two sources. First, the data from fact-checking sites are collected to have an expert-validated dataset. In this study, we got data from a fact-checking website that is www.politifact.com. News on this website rated the accuracy of claims by elected officials and others who speak up in American politics. Fact-checking data is utilized to match the historical news of users to determine the user's credibility. Second, we used the data crawled from Twitter that was used in our previous study [29]. The dataset included interaction data related to news and user profile data from all users interacting with the news. Based on a set of rated news, Twitter's search API was used to crawl related tweets. Finally, the experiments were conducted on a dataset including 255 fake news and 260 real news. These tweets belong to 357 different users. The publisher's data and interaction data are also collected to construct the feature vectors as described in Section 3.2. There 2532 comments were selected and annotated into three classes. Comments that reflect the news as fake news are labeled as negative, comments that believe news content are labeled as positive and other comments are labeled as neutral. Statistics of the dataset are provided in Table 1. In this study, we used the available tool Twitter4J[8] to collect data. Data were cleaned to eliminate unnecessary elements such as hyperlinks, hashtags, symbols, mentions, and emoticons.

Table 2. Performance comparison among individual classifiers and voting classifiers

Model	Precision	Recall	F1
Baseline	73.5	70.6	72.0
Naive Bayes	73.1	74.5	73.8
SVM	75.5	72.5	74.0
Softmax	76.0	74.5	75.2
Hard voting	78.0	**76.5**	77.2
Soft voting	**81.3**	**76.5**	**78.8**

4.3. Experiment Results

In this work, we conducted the experiments of the proposed method by using the scikit-learn and gensim libraries. The features as described in Section 3.2 were crawled and built into feature vectors. The data were crawled from Politifact and Twitter as presented in Section. The models were applied to feature vectors to train classifiers. The training and testing were implemented with the scikit-learn library in Python. For the multiSVM classifier, the kernel parameter was set to 'rbf'. The multi_class parameter was set to 'multinomial' for the softmax classifier. For the voting classifiers, the weights parameter was set to 'None' in the case of hard voting, and a sequence of weights in order to weight the occurrences of class probabilities before averaging in the case of soft voting. The other parameters were set to default. To evaluate the effectiveness of the proposal, we experimented with a 5-fold cross-validation strategy. At a time, 1-fold was chosen for a test set and 4-fold remaining was the training set. The process was repeated 5 times for each model to cover all tests and training sets. Five models were used to train on 4-fold treated as a training set and evaluated with a 1-fold test set remaining. The results of the 5-fold cross-validation runs were summarized with the mean of all experiments. We conducted the fake news detection experiments with five classifiers which were three individual classifiers (i.e., Naive Bayes, SVM, and Softmax), hard voting, and soft voting classifiers. Each model was analyzed in terms of Precision, Recall, and F1 Score. The results from [29] were treated as the baselines.

The results of the experiments are shown in Table 2. The highest performance was achieved by the soft voting ensemble model and the Naive Bayes model was the lowest. More specifically, the soft voting ensemble classifier achieved an F1 score of 78.8%, the

[8] http://twitter4j.org

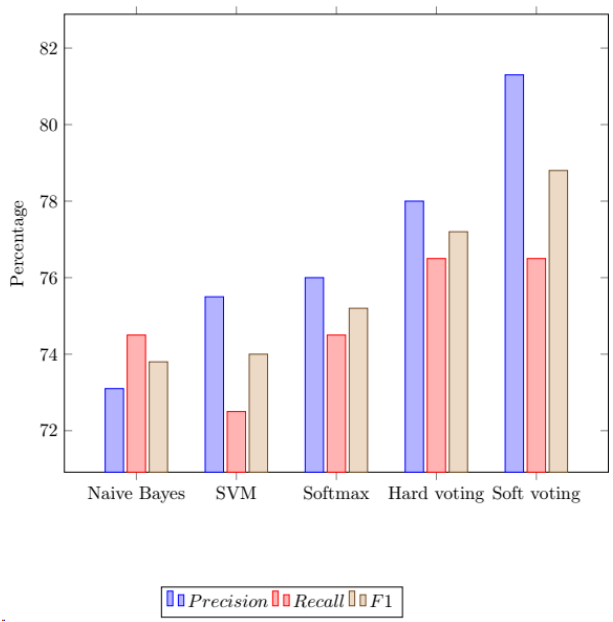

Fig. 5. Comparison of performance in different classification models

best among all the classifiers. The Naive Bayes and the SVM model achieved the same performance. The Softmax classifier and the SVM classifier predicted quite correctly. However, the Softmax classifier surpassed the SVM classifier in the number of predicted samples. Both voting ensemble classifiers outperform the remaining individual classifiers. In terms of precision, the soft voting classifier achieved the best result. It was better than the hard voting, softmax, SVM, and Naive Bayes classifiers by 3.3%, 5.3%, 5.8%, and 8.2%, respectively. The soft voting ensemble classifier only surpassed the hard voting ensemble classifier in precision. However, the predicted results of these two classifiers were the same. As shown in Figure 5, the soft voting ensemble classifier improved the accuracy compared to the hard voting ensemble classifier. The soft voting classifier achieved the F1 score of 78.8%. It was better than the best individual classifier by 3.6% and better than the baseline by 6.8% [29]. In this work, the user's credibility was enhanced based on the fact-checking source. The results demonstrated the effectiveness of the feature set and the proposed method.

The experiments demonstrated the effectiveness of the voting ensemble technique applied to extracted features in the fake news detection problem. The soft voting classifier exceeds the hard voting classifier in detecting fake news. In the above experiments, we set the weights of classifiers to be equal. To evaluate the impact of each classifier, in the soft voting method, we changed the weights associated with classifiers. Table 3 shows the results with different weights of the soft voting technique. In the experiment, one classifier was assigned the weight of 2, remaining classifiers were assigned 1. In Table 3, the first column presents the weights α, γ, and β assigned to the classifiers Naive Bayes, SVM, and Softmax, respectively. Figure 6 is a graphical representation of the performance of

Table 3. The performance of the soft voting classifiers with different weights

Weights assign for classifiers (α, γ, β)	Precision	Recall	F1
2, 1, 1	73.6	76.5	75.0
1, 2, 1	76.5	76.5	76.5
1, 1, 2	**80.0**	**78.4**	**79.2**

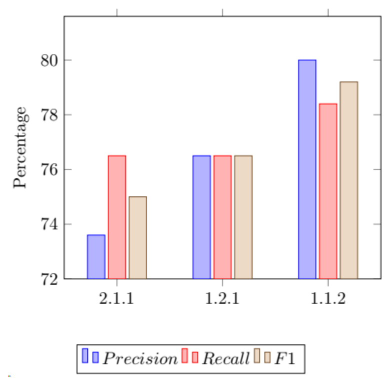

Fig. 6. The performance of the soft voting classifier with different weights

the soft voting classifier with different weights. The experiment with the weight of the softmax classifier of 2 was the most effective. It achieved an F1 score of 79.2%. Meanwhile, the Naive Bayes classifier with a weight of 2 did not perform well, especially the precision was low. In general, the weight associated with the classifiers should be set proportional to the effectiveness of the model. However, the experiments have not revealed the best weight value for each classifier. Obviously, efficient classifiers play an important role in improving the performance of the voting ensemble method.

5. Conclusions

In this paper, we proposed a voting ensemble model to detect fake news on social media based on exploiting the features of the wisdom of crowds. The voting ensemble model comprised three classifiers namely, SVM, Naive Bayes, and Softmax. The feature set was extracted from the user's interaction data and the user's profiles. The method mainly focused on the user's viewpoints to determine the authenticity of the news. We exploited three features that were user's opinions, user's credibility, and exhortation level. The user's credibility was enhanced effectively by matching the user's historical data with fact-checking sources. The individual classifiers and ensemble classifiers were applied to the extracted features to evaluate the effectiveness of the proposed methods. The proposed method experimented on a Twitter dataset. The experimental results showed that the voting ensemble classifiers achieved high performance on all metrics. The soft voting classifier achieved an F1 score of 78.8%. It was better than the best individual classifier by 3.6% and better than the baseline by 6.8%. The experiments demonstrated the effectiveness of using features extracted from the wisdom of crowds on social networks.

For future works, we are planning to build a multi-agent system model for simulating the process of fake new propagation and to use consensus-based methods for making decisions regarding assigning fake status for news [7], [28]. For determining user opinions, consensus methods will be used to unify different user opinions for determining a common opinion that best represents a given set of user opinions [15], [17].

Acknowledgments. This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number DS2021-26-03.

References

1. Abbasi, M.A., Liu, H.: Measuring user credibility in social media. In: International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction. pp. 441–448. Springer (2013)
2. Afroz, S., Brennan, M., Greenstadt, R.: Detecting hoaxes, frauds, and deception in writing style online. In: 2012 IEEE Symposium on Security and Privacy. pp. 461–475. IEEE (2012)
3. Awal, G.K., Bharadwaj, K.K.: Team formation in social networks based on collective intelligence–an evolutionary approach. Applied intelligence 41, 627–648 (2014)
4. Böhringer, M., Helmholz, P.: "what are they thinking?"-accessing collective intelligence in twitter (2011)
5. Bondielli, A., Marcelloni, F.: A survey on fake news and rumour detection techniques. Information Sciences 497, 38–55 (2019)
6. Della Vedova, M.L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M., de Alfaro, L.: Automatic online fake news detection combining content and social signals. In: 2018 22nd Conference of Open Innovations Association (FRUCT). pp. 272–279. IEEE (2018)
7. Duong, T.H., Nguyen, N.T., Jo, G.S.: A hybrid method for integrating multiple ontologies. Cybernetics and Systems: An International Journal 40(2), 123–145 (2009)
8. Gholami, B., Safavi, R.: Harnessing collective intelligence: Wiki and social network from end-user perspective. In: 2010 International Conference on e-Education, e-Business, e-Management and e-Learning. pp. 242–246. IEEE (2010)
9. Groza, A.: Detecting fake news for the new coronavirus by reasoning on the covid-19 ontology. arXiv preprint arXiv:2004.12330 (2020)
10. Hakak, S., Alazab, M., Khan, S., Gadekallu, T.R., Maddikunta, P.K.R., Khan, W.Z.: An ensemble machine learning approach through effective feature extraction to classify fake news. Future Generation Computer Systems 117, 47–58 (2021)
11. Hangloo, S., Arora, B.: Fake news detection tools and methods–a review. arXiv preprint arXiv:2112.11185 (2021)
12. Hardalov, M., Koychev, I., Nakov, P.: In search of credible news. In: International Conference on Artificial Intelligence: Methodology, Systems, and Applications. pp. 172–180. Springer (2016)
13. Jin, Z., Cao, J., Zhang, Y., Luo, J.: News verification by exploiting conflicting social viewpoints in microblogs. In: Proceedings of the AAAI conference on artificial intelligence. vol. 30 (2016)
14. Karnyoto, A.S., Sun, C., Liu, B., Wang, X.: Transfer learning and gru-crf augmentation for covid-19 fake news detection. Computer Science and Information Systems (00), 53–53 (2022)
15. Katarzyniak, R., Nguyen, N.T.: Reconciling inconsistent profiles of agents' knowledge states in distributed multiagent systems using consensus methods. Systems Science 26(4), 93–119 (2000)

16. Kwon, S., Cha, M., Jung, K., Chen, W., Wang, Y.: Prominent features of rumor propagation in online social media. In: 2013 IEEE 13th International Conference on Data Mining. pp. 1103–1108. IEEE (2013)
17. Nguyen, N.T.: Conflicts of ontologies–classification and consensus-based methods for resolving. In: Knowledge-Based Intelligent Information and Engineering Systems: 10th International Conference, KES 2006, Bournemouth, UK, October 9-11, 2006. Proceedings, Part II 10. pp. 267–274. Springer (2006)
18. Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic detection of fake news. arXiv preprint arXiv:1708.07104 (2017)
19. Qin, Y., Wurzer, D., Lavrenko, V., Tang, C.: Spotting rumors via novelty detection. arXiv preprint arXiv:1611.06322 (2016)
20. Rashkin, H., Choi, E., Jang, J.Y., Volkova, S., Choi, Y.: Truth of varying shades: Analyzing language in fake news and political fact-checking. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. pp. 2931–2937 (2017)
21. Sahoo, S.R., Gupta, B.B.: Multiple features based approach for automatic fake news detection on social networks using deep learning. Applied Soft Computing 100, 106983 (2021)
22. Seddari, N., Derhab, A., Belaoued, M., Halboob, W., Al-Muhtadi, J., Bouras, A.: A hybrid linguistic and knowledge-based analysis approach for fake news detection on social media. IEEE Access 10, 62097–62109 (2022)
23. Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F.: The spread of low-credibility content by social bots. Nature communications 9(1), 1–9 (2018)
24. Shu, K., Bernard, H.R., Liu, H.: Studying fake news via network analysis: detection and mitigation. In: Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining, pp. 43–65 (2019)
25. Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter 19(1), 22–36 (2017)
26. Shu, K., Wang, S., Liu, H.: Understanding user profiles on social media for fake news detection. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). pp. 430–435. IEEE (2018)
27. Shu, K., Wang, S., Liu, H.: Beyond news contents: The role of social context for fake news detection. In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. pp. 312–320. ACM (2019)
28. Sliwko, L., Nguyen, N.T.: Using multi-agent systems and consensus methods for information retrieval in internet. International Journal of Intelligent Information and Database Systems 1(2), 181–198 (2007)
29. Tran, V.C., Nguyen, V.D., Nguyen, N.T.: Automatic fake news detection by exploiting user's assessments on social networks: A case study of twitter. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. pp. 373–384. Springer (2020)
30. Wei, X., Zhang, Z., Zhang, M., Chen, W., Zeng, D.D.: Combining crowd and machine intelligence to detect false news on social media. Mis Quarterly (2019)
31. Yang, S., Shu, K., Wang, S., Gu, R., Wu, F., Liu, H.: Unsupervised fake news detection on social media: A generative approach. In: Proceedings of the AAAI conference on artificial intelligence. vol. 33, pp. 5644–5651 (2019)
32. Zhou, X., Zafarani, R.: Fake news: A survey of research, detection methods, and opportunities. arXiv preprint arXiv:1812.00315 2 (2018)
33. Zhou, X., Zafarani, R.: A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys (CSUR) 53(5), 1–40 (2020)

Hai Bang Truong received a Ph.D. degree in computer science from Viet Nam National University Ho Chi Minh City, in 2016. He is currently a professor at the Faculty of Computer Science, University of Technology, Ho Chi Minh City, Vietnam. He has authored 7 journal articles and 14 conference papers. His research interests include collective intelligence, knowledge integration methods, and social network analysis.

Van Cuong Tran received a B.S. degree in computer science from the Hue University of Sciences, Vietnam, in 2012, and a Ph.D. degree in computer engineering from Yeungnam University, South Korea, in 2017. He is currently a professor at Quang Binh University, Vietnam, and the assistant dean of the Faculty of Engineering and Information Technology. He has authored 10 journal articles and 14 conference papers. His research interests include named entity recognition, sentiment analysis, recommendation systems, and social network analysis.

Received: March 02, 2023; Accepted: June 22, 2023.