Received 5 June 2023, accepted 24 June 2023, date of publication 12 July 2023, date of current version 19 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294613

APPLIED RESEARCH

Constructing a User-Centered Fake News Detection Model by Using Classification Algorithms in Machine Learning Techniques

MINJUNG PARK AND SANGMI CHAI

Ewha School of Business, Ewha Womans University, Seodaemun-gu, Seoul 03760, South Korea

Corresponding author: Sangmi Chai (smchai@ewha.ac.kr)

This work was supported in part by the Ministry of Education of the Republic of Korea and in part by the National Research Foundation of Korea under Grant NRF-2020S1A5A2A01046634.

ABSTRACT As fake news spreads rapidly in social media, attempts to develop detection technology to automatically identify fake news are actively being developed, recently. However, most of them focus only on the linguistic and compositional characteristics of fake news (e.g., source or authors indication, length of a message, frequency of negative words). Compared to them, this study proposes a fake news detection model based on machine learning that reflects the characteristics of users, news content, and social networks based on social capital. To comprehensively reflect the characteristics related to the spread of fake news, this study applied the XGBoost model to estimate the feature importance of each variable to derive the priority factors that preferentially affect fake news detection. Based on the derived variables, we established SVM, RF, LR, CART, and NNET, which are representative classification models of machine learning, and compared the performance rate of fake news detection. To generalize the established models (i.e., to avoid overfitting or underfitting), this study performed a cross-validation step, and to compare the predictive accuracy of the established models. As a result, the RF model indicated the highest prediction rate at about 94%, while the NNET had the lowest performance rate at about 92.1%. The results of this study are expected to contribute to improve the fake news detection system in preparation for the more sophisticated generation and spread of fake news.

INDEX TERMS Classification algorithms, fake news, fake news detection, feature selection, prediction algorithms, predictive models, XGBoost.

I. INTRODUCTION

The content-based recommender system developed to provide customized content to users improved their satisfaction; however, it recently became a decisive opportunity for spreading fake news [1]. These are designed to continuously show content in the feed similar to what the user has previously seen or has shown engagement with in the past by "Likes" or comments, regardless of whether the news is true [2]. In other words, once a user encounters fake news, the system has no choice but to recommend similar content to the user continuously. They even intentionally adjust the appearance of unwanted content in their social media feeds [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Geng-Ming Jiang.

Almost prior studies have focused on detecting or identifying fake news depending on its linguistic or compositional characteristics. They have identified fake news based on whether the article has a clear author and source or whether the article has enough length [4], [5], [6], [7]. This approach assumed differences in linguistic or compositional features between fake and factual news. It is hard to reflect the characteristics of users who accept or spread fake news and the features of the social media networks where fake news spreads.

With the advent of ChatGPT(Generative Pre-trained Transformer), which can describe stylistic features that do not look awkward, such as those written by low-level AI, it is no longer possible to guarantee the accuracy of detecting fake news in the previous way. Recently, it has become so easy to create fake news that looks like real news that users mistake news written by ChatGPT in a few seconds for news written by a professional journalist [48]. Therefore, it is necessary to identify fake news differently than before. In this study, we finally propose a detection model that comprehensively considers not only the visual features of the content, but also the characteristics of the users who generate and share fake news and the networks that spread fake news.

Moreover, fake news can be generated more sophisticatedly due to the technology that can automatically cause it to be identical to real news easily in a short time by using AI (Artificial Intelligence) is spreading today. AI-powered bots in Twitter (i.e., AI Twitterbot) can populate thousands of user accounts that can support and oppose any content, which looks the same as the real news, even if it is fake, the bot controllers target [8]. Therefore, it has become increasingly difficult to detect precisely manipulated fake news based only on its superficial features. It is necessary to approach the user characteristics and networks of social media from a more diverse viewpoint, to overcome the existing fake news detection method that focuses on linguistic characteristics.

This study aims to improve the prediction performance of fake news detection by overcoming the limitation that previous studies did not consider the characteristics of information recipients. Therefore, we establish a fake news detection model by considering various content features and users in social media and the network where fake news is generated and propagated. Among the different explanatory variables to detect fake news, 'feature selection,' which is the priority of the explanatory variable, is first derived through the XGBoost (Extreme Gradient Boosting). By constructing an optimal fake news detection model through the selected explanatory variables by XGBoost, we aim to increase the predictive performance rate. It is constructed by applying five machine learning techniques which are Logistic Regression (LR), Neural Network (NNET), Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Trees (CART). A model with the highest prediction performance rate for detecting fake news is finally derived by comparing their performance rate.

II. RELATED WORKS

A. FAKE NEWS DETECTION

Recent studies demonstrate the diverse range of approaches that researchers are taking to develop fake news detection models using machine learning, and the potential of these models to improve the accuracy of news verification. However, like any technology, fake news detection systems are hard to be perfect and can sometimes make errors in identifying fake news. If a fake news is not detected appropriately by the system,] it can be shared widely on social media in a short time, leading to a significant impact on public opinion and behavior. For example, hundreds of people died in Iran, after drinking methanol for curing COVID-19, due to the fake news which had been accepted fact in the first [47]. Furthermore, errors in fake news detection systems can lead to false accusations or misidentifications of individuals or groups. If a system mistakenly identifies a legitimate news story as fake news, it could lead to accusations of bias or censorship against the news outlet that published it [5].

To summarize the current state-of-the-art in fake news detection systems, most of the previous research assumes that linguistic and compositional features of content are the main criteria for distinguishing fake news and real news [9]. Fake news detection systems typically rely on linguistic and structural features of news articles, but they often fail to capture the context of the news, such as the history of the news source or the socio-political environment in which the news is circulated. These methods which could not detect words semantic meaning and context of the word picked up from a fake news have been identified with low accuracy value [10].

The content-oriented fake news detection, which is the most common approach, focuses on natural language processing (NLP) to identify fake news by concentrating on the characteristics of the text. NLP techniques process news content based on language pattern detection, word occurrences common to satire, irony, sentiment, and topicality [11]. To find deemphasizing the source or design highlights the article headline of the news is also a way of identifying fake news by paying attention to its textual characteristics [12]. This content-oriented approach assumes that fake and real news have different linguistic and composition structures. It proposes a hybrid fake news detection algorithm that combines a linguistic approach and network cues and provides operational guidelines for a feasible fake news detecting system [13]. In addition, based on grammatical characteristics through syntax parsing through Probabilistic Context Free Grammar (PCFG) and the difference between keywords used in fake news and real news, semantic characteristics, rhetorical structure, and discourse analysis results were selected as explanatory variables to determine whether fake news or not [14]. On fake news detection targeting Facebook posts and various articles, Term Frequency - Inverse Document Frequency (TF-IDF) is frequently used to represent text characteristics in text analysis and was used as a criterion for classifying fake news [5].

It is constructed as an automated fake news detection model by extracting linguistic features from the text of online newspaper articles [15]. Using the LIWC (Linguistic Inquiry and Word Count), the ratio using punctuation marks (e.g., periods, commas, question marks, exclamation points) was calculated. In addition, it also figured how many words related to positive and negative were mentioned in each document based on the LIWC lexicon to extract the proportions of observations that fall into psycholinguistic categories and, finally, constructed as readability metrics to apply in the fake news detection model. A naive Bayesian classifier to news attributes [16], the combination of the text of news and clickbait [17], a hybrid model combining spread patterns of fake news, and semantic analysis to develop an automated fake news detection model is being actively conducted [18], [19], [20], [21].

There are several datasets available in literature, such as the LIAR dataset or Fakenewsnet, that can be used to train and evaluate machine learning models for the task of detecting fake news. The LIAR dataset, for example, is a widely used dataset for fake news detection, consisting of statements labeled as either true, mostly true, half true, barely true, false [51]. The Fakenewsnet dataset, on the other hand, contains news articles labeled as either fake, bias, or conspiracy [52]. Both datasets have been labeled by human annotators. The LIAR dataset has been labeled by human fact-checkers, while the Fakenewsnet dataset has been labeled by crowd-sourced workers. They determined fake news depending on its surface-level linguistic patterns. It is difficult to apply their patterns when fake news has the appearance as same as real news and uses a professional expression look like written by professional journalist, these days. Therefore, it has been addressed, with the number of sophisticated fake news has been increased, the needs for reflecting not only its linguistic characteristics but also various approaches including the users' network relationships and the context of fake news.

As fake news detection systems have become more popular, recently, models for spreading fake news have emerged that bypass the algorithmic identification methods of well-known fake news systems to avoid being identified as fake news. Fake news detection systems can be targeted by individuals who deliberately create fake news that is designed to bypass these systems [22]. These adversarial attacks can make it difficult to develop accurate and reliable fake news detection systems. While the technology used to spread fake news continues to evolve, current fake news detection systems are still focused on the linguistic and compositional features used in fake news, similar to the past. Therefore, we aim to improve the existing system that are difficult to distinguish fake news with the increased the more sophisticated fake news in the recent seems like fact news. Recent fake news detection systems have some limitations, which have made it easier for fake news to spread. Some limitations that have recently been identified include as follows. First, the existing fake news detection system mainly rely on keywords. Some systems depend on keywords to identify fake news, which can be easily manipulated by those spreading false information [49]. In other words, fake news can be designed to reach a wider audience by exploiting weaknesses in these systems. Second, it has been identified that fake news system, in the recent, has been hard to detect new forms of fake news. As the methods used to spread fake news evolve, the existing detection systems may not be able to keep up. For example, deepfake videos or manipulated images may be difficult to detect using current systems [50]. Third, many fake news systems rely on analyzing individual pieces of content in isolation, without considering the broader context in which it was shared [22]. This can make it difficult to determine whether a piece of information is intentionally false or simply a mistake or misunderstanding. In other words, it is hard to determine a content is fact or not only to depend on the existing fake news systems which has a lack of contents. Finally, we include the context of fake news and the latest fake news generation styles by reflecting word sentiment, similarity, and users' network relationships to overcome the limitations of the existing fake news system presented above.

B. SOCIAL CAPITAL THEORY

Social influence can be explained that is a structure of social relationships by exchanging interactions between users based on the social network. The main background enabling this mutual exchange of social influence is derived from each individual's social capital in the social network [23]. Therefore, it was found that the type of social network formed by the users differs according to the social capital possessed by them. At the same time, the will to create a social relationship and the degree of persistence of the relationship are different [24]. Social capital arises from different attributes within three dimensions –structural, relational, and cognitive [24]. The structural dimension is the key to whether or not to establish a network connection between actors in the network and includes the overall strength of the connection [25]. The relational dimension of social capital refers to personal relationships between actors formed through interactions between individuals [26]. Finally, the cognitive dimension refers to sharing shared representations, interpretations, semantic content, and systems among agents [27]. Based on the three dimensions of social capital presented above, determinants influencing was selected to detect fake news spread on Twitter. It is to comprehensively consider the characteristics of the network of Twitter and users affecting the individual acceptance and spread of fake news.

1) NETWORK FEATURES OF STRUCTURAL DIMENSION

The network features of Twitter can be defined as factors that can affect users mutually through Twitter. In this study, three factors were adopted to estimate the network features of social network structures: the number of followers and followings and the degree of centrality in the network. First, the number of followers and followings is representative as an indirect proxy variable that shows the user's willingness to interact with others on Twitter [28], [29]. It can be inferred that users have many followers means that the user has high expectations for establishing relationships with others based on the structural network features of Twitter. It can be explained that the user has a higher willingness to engage in networking activities with others when the user has a lot of followings than those with few followings. As the degree of centrality in the network can be measured differently depending on the way users communicate and interact with others on Twitter, considering all these factors (i.e., in-degree, out-degree, betweenness centrality) were regarded as features of the network. In-degree centrality indicates how many users are followed by other users in Twitter's limited network [30]. In other words, it is an index of influence in which the direction of exchange within the Twitter network has a direction from others to oneself. In contrast, out-degree centrality refers to how much information or attention you provide to other users on the Twitter network. Therefore, it is estimated as an indicator of outward connection centrality within the Twitter network, and it means the number of arrows directed away from each user to the others [30]. The in-degree centrality and out-degree centrality of node *i,* denoted respectively with $C_{I,i}$ and $C_{O,i}$, can be defined as:

$$C_{I,i} = \sum\nolimits_{j=1,j\neq i}^{N} \frac{l_{ij}}{N-1} \tag{1}$$

$$C_{O,i} = \sum\nolimits_{j=1,j\neq i}^{N} \frac{l_{ij}}{N-1} \tag{2}$$

Betweenness centrality refers to how information is mediated between different users in Twitter's network. It is often used to find nodes that serve as a bridge from one part of a network to another. It is an indicator showing how much a user can transmit information among users [32]. The betweenness Centrality of node i, denoted as BC_i, can be defined as:

$$BC_i = \frac{2\sum_{1}^{N}\sum_{k}^{N}\frac{9_{jk(i)}}{g_{jk}}}{N^2-3N+2}, \quad j\neq k\neq i \tag{3}$$

Therefore, in this study, the degree centrality index on the Twitter network is divided into three network influence indicators as above and measured. Then a fake news detection model is constructed based on the degree of centrality influence level of each.

2) USER FEATURES OF RELATIONAL DIMENSION

The relational dimension in Twitter is to examine how much and what kind of relational actions the users have taken on Twitter. It is estimated based on the total number of tweets posted by each user and the total number of 'Likes' and 'Retweets' received from others. In addition to tweet posts, 'Retweets' and 'Likes' are essential for inferring user characteristics in the network. They are the most representative and only communication methods on Twitter [23].

A user's network activity in social media means how each user influences the other users based on the number of Twitter users' mentions and retweets [30]. There have been identified that empirically evaluate the effects of influencers based on the probability that a tweet is retweeted on Twitter [33], [34]. It has been confirmed that when a tweet from a user with less than 1,000 followers is retweeted, that tweet is delivered to thousands of additional users. A retweet can be explained as having an additive impact on many Twitter users, regardless of their number of followers [35]. Combining the previous studies presented above, this study also measures the characteristics of Twitter users based on the total number of tweets, retweets, and likes created by users for evaluating the relational dimension of social capital on Twitter.

3) CONTENT FEATURES OF COGNITIVE DIMENSION

Twitter, the subject of this research, is a representative social media where information is shared mainly through short text messages of 140 characters or less. Therefore, since the content of the tweet message is implied by the limited 140 characters, it is difficult for the information recipient to sufficiently acquire the information through only the tweet message. Accordingly, it is not easy for a user to determine the authenticity of the information on Twitter. In this study, the text characteristics of tweet messages were judged as influential factors for identifying fake news, and the characteristics of tweet text messages were examined in terms of 'word similarity' and 'word sentiment.'

Word similarity evaluates how two words/phrases/ documents are similar. Words used with similar meanings to specific words in a sentence into numerical values to check the similarity between words composing the entire sentence. The similarity is calculated through word embedding, which is a method of quantifying a single word constituting a sentence. A word is expressed as a vector, and the distance between a specific term and a similar word is calculated to derive the similarity between the words. Therefore, word similarity analysis estimates the semantics of word meaning based on context according to the word embedding method in calculating the distance between words. When each tweet message consists of frequent use of vocabulary with similar purposes, the value of word similarity increases. Accordingly, in the case of general information messages that convey only the facts without arbitrarily forming the tone of the message, it is common to record a relatively low value of word similarity by using mainly non-emotional words or neutral expressions without duplication or repetition. However, this study predicted that fake news would constitute a tweet message by repeatedly and intentionally using the same or similar words to incite users who encountered the information to accept and spread the information. Therefore, word similarity was also considered a vital antecedent factor for detecting fake news in this study and included as an explanatory variable to identify fake news among general information.

SentiWordNet (SWN) is a dictionary that adds sentiment scores to each word. It gives a positivity, negativity, and objectivity score, which measures how much sentiment a word has. SWN automatically sets the values of positive, negative, and objective sentiment for each synonym set of WordNet. Various methods for sentiment analysis have been developed. However, SWN has a difference in that it is possible to calculate an adequate emotional intensity level because the emotion value is applied differently for each part of speech of the word used.

Fake news is expected to show a high emotional intensity because it uses more provocative vocabulary and negative phrases that appeal to emotions. Analyzing about 320,000 news articles published in the New York Times from 2012 to 2014 by applying SWN revealed that out of 754 days, positive and negative news was recorded 322 days and 432 days, respectively. Analysis of positive and negative news based on SWN showed an improved stock price prediction performance of about 4% or more than when the only technical analysis was performed [36]. Parts of speech of online product reviews were classified into adjectives, adverbs, and verbs, and found that SWN improved performance compared to previous sentiment analysis methods [37]. Furthermore, SWN demonstrated that the performance of identifying the tone of each document was improved even for articles such as relatively crude reviews [38]. Combining previous studies, this study judged that deduction of linguistic characteristics of fake news based on SWN would have a significant effect on identification of fake news.

III. RESEARCH METHOD

This study conducts the following research analysis procedures to identify fake news. First, XGBoost is applied to derive the priority of variables that have a significant effect on fake news detection. Second, we establish a model to distinguish fake news with five representative classification algorithms (i.e., LR, NNET, RF, SVM, CART) among machine learning models based on the derived factors from performing XGBoost. Third, we adopt k-fold cross-validation steps to improve the performance rate and generalized of the established each model, and also performed ablation studies to increase the robustness of the model.

Machine learning techniques have recently been widely used in various research fields for prediction. This study uses supervised learning classification algorithms to identify fake news from given data. Recently, ensemble learning uses multiple models together to improve the performance of an algorithm rather than using a single model [39]. XGBoost is a representative ensemble model, which usually shows superior performance in classification than other single classification algorithms [40]. The existing Gradient Boost Machine has a limitation in that the speed of analysis is significantly slow as the learning weights are sequentially increased. However, unlike this, XGBoost is faster than the existing gradient boosting technique because it can learn in a parallel CPU. In addition, it has the advantage of solid durability against overfitting in that it provides regularization to prevent overfitting [40], [41].

XGBoost model is formed to identify feature importance, the gain for the accuracy of each variable, and the frequency of the appearance of the variable in the entire tree can be reflected together [41]. As the split criterion used and the accuracy contribution point value due to the pruning is derived through each pruning, the direction of the variable can be grasped through this. It is possible to identify which variables and how much they worked in making "important decisions" in the decision-making process among the many variables input to construct the model when XGBoost is used. Therefore, the importance of the variables input for model building is calculated for each variable, and each variable can be sorted by rank.

In general, when constructing a model by inputting many variables, there is an advantage in that the model can be built by considering the various conditions. In contrast, the model's noise is increased. It also has a limitation that the model fitness inevitably deteriorates due to the increase in noise. It becomes an obstacle in constructing the fake news detection model by lowering the accuracy of the prediction model. Feature selection through XGBoost is expected to enable model construction with high accuracy while preventing model overfitting. Accordingly, in this study, the main factors affecting identifying tweet messages-related fake news are first derived and based on them, an optimal model is built.

IV. DATA COLLECTION

We collected a total of 23,592 tweets over a period of 595 days (from Mar. 5, 2019 to Oct. 19, 2020). We first extracted the topics of popular and representative fake news cases that spread globally. Only news that were clearly determined to be fake by authoritative media outlets (e.g., The Wall Street Journal, CNN) could be included in the final data analysis as a fake. The selected fake news was collected considering various fields: medicine, politics, economy, IT, entertainment, and international areas. For example, *'Drinking alcohol-including beverages with high percentages of alcohol-offers protection from COVID 19…'* was collected, which is popular fake news in medicine. Data preprocessing was also performed as follows. First, considering the characteristics of fake news, tweets which include same topics or similar contents were regarded as the same news and excluded from the final analysis. In other words, tweets and retweets that talk about the same content were considered duplicates and removed except for the first tweet posted. Second, tweets in which users unilaterally expressed emotions such as anger, joy, and sadness, including agreeing or disapproving with a specific tweet, were considered tweets that delivered the only personal emotions, were removed from final data set. We also excluded tweets in which users evaluated by the users whether the tweet was fake or fact. 402 tweets, finally, including 202 fake news and 200 true news tweets, were used to establish a fake news detection model.

In this study, each tweet's 'word sentiment' and 'word similarity' were calculated as content features in a cognitive dimension of social capital. And the degree of centrality of users (i.e., in-degree, out-degree, betweenness centrality) in the Twitter network and the user's account information of the numbers of followers and followings were also collected for network features in the structural dimension of the social capital. Additionally, data on the total number of tweet messages posted by each ID and the total number of likes and retweets received for each account were also obtained to evaluate users' relational features on Twitter.

V. RESEARCH ANALYSIS

A. EXTRACTING PRIORITY FACTORS USING XGBoost

This study aims to determine which factors have a significant influence in discriminating between fake news and true news by identifying the feature importance of each variable through XGBoost. The importance of a variable is the result of adding up the gains that each variable contributed to the model's accuracy in constructing the XGBoost model. As a result of performing XGBoost based on a total of 10 explanatory variables input, the importance of the derived variables is as follows FIGURE 1. and TABLE 1.

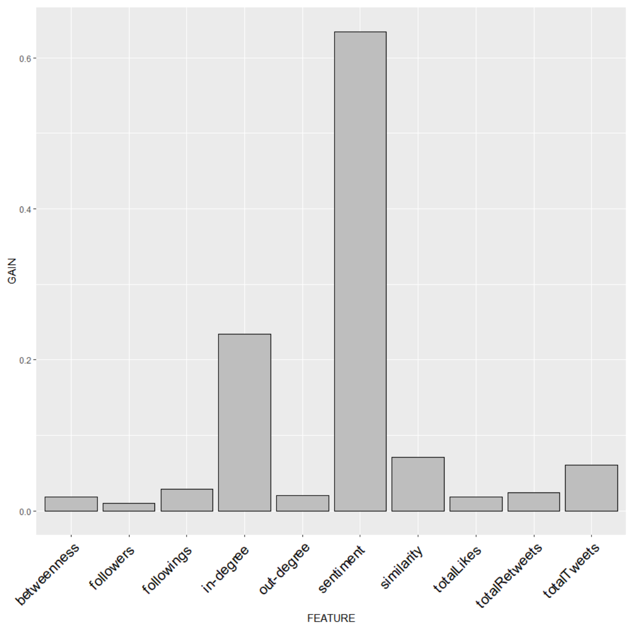

FIGURE 1. Feature importance of each variable.

TABLE 1. Feature importance of each variable.

Feature	Gain	Cover	Frequency
Word Sentiment	0.6346001	0.2993744	0.1063830
In- degree Centrality	0.2339340	0.3708668	0.2127660
Word Similarity	0.0705958	0.2067322	0.4042553
Number of Total Tweets	0.0608700	0.1230265	0.2765957

As shown in FIGURE 1. the most influential factor in identifying fake news among tweets is 'word sentiment.' The more positive or negative the tweet's tone is too one-sided, the more it needs to be suspected as fake news. It was found that Twitter users' in-degree centrality, word similarity, and a total number of tweets sequentially influenced the construction of a fake news detection model. As the in-degree centrality formed by Twitter users appears as an explanatory variable with high importance, it can be confirmed that how actively individuals form relationships and interact with other users in the Twitter network are essential factors in determining fake news. Word similarity is a significant criterion for detecting fake news. A high word similarity value is estimated if each tweet message repeatedly uses the same vocabulary and sentences with similar meanings. Therefore, in the case of general information messages that deliver only the facts without arbitrarily forming the tone of the message, it is common to record relatively low word similarity by mainly using neutral expressions without overlapping or repeating non-emotional words. However, when the word similarity is high, there is a high probability that the writer of the message wrote it without considering the context or repeatedly used a word that appealed to the user's emotions to mislead the reader. Furthermore, with the advent of AI Bots that automatically write tweets, word similarity is high even when similar words are repeated without paraphrasing. We can also suspect that an individual or a few intentionally spread the information if the number of tweets on a specific topic is too many on Twitter for a certain period. Therefore, it can be inferred that the information is intentionally for the profit of a few individuals, not to provide factual information. On the contrary, Twitter user's number of followers and followings was relatively low importance. Therefore, this study intends to establish a fake news detection model based on a machine learning algorithm based on the four explanatory factors (i.e., word sentiment, in-degree centrality, word similarity, the number of total tweets) extracted as significant important variables XGBoost.

B. ESTABLISHING FAKE NEWS DETECTION MODELS USING MACHINE LEARNING

A detection model based on a machine learning algorithm is constructed using categorical binary variables as a dependent variable to determine whether the tweet message is fake news or not. Additionally, the highest performance rate of the fake news detection model is selected by evaluating the performance of classification models based on various machine learning algorithms. A fake news detection model is established based on five machine learning algorithms.

First, the Logistic Regression (LR) performed a stepwise method analysis to overcome the limitation that the complexity of the model increases as all variables are input. As a result, when the AIC value was 540.1684, the best model with the highest performance rate was constructed. Classification and Regression Tree (CART), to increase the performance of the prediction rate, the deviance value was adjusted 100 times repeatedly. As a result, the optimal model was established when the number of nodes was 6. To find the optimal value of deviance, tree pruning was repeatedly performed. This iterative pruning process prevents the node from being split and stopped due to generating more nodes than necessary in the model or preventing performance degradation due to a minimal number of nodes. Finally, 6 (deviance =7.75), the best size value with the smallest deviation, was fixed as the terminal node. In Neural Network (NNET), the parameters tested were size and decay. The optimal model was searched by controlling the number of such hidden layers, and as a result, the three layers (size =1, decay =0.1) showed the best performance. Support Vector Machine (SVM) was analyzed using the nonlinear Radial Basis Function (RBF) of Gaussian Kernel. The parameters tested were Sigma and C. C is used for the soft margin cost function, which involves trading error penalty for stability, while Sigma is the standard deviation. Through this, Sigma =0.1, C =10, which has the smallest difference between the training data and the result of the evaluation data, was selected as the final model. To establish Random Forest (RF), the model constructed through the change of parameter values was implemented in the evaluation data to evaluate the model's performance. As a result, the number of generated trees in the random forest was set to 100 (ntree =100). The optimization of RF's tuning parameter 'mtry' is considered in conjunction with variable selection in each node as 3. Accordingly, the prediction accuracy according to the misclassification rate, which is the final classification performance of each model, is presented in the following TABLE 2.

TABLE 2. Comparisons of prediction performance rates.

LR	FAKE	TRUE
FAKE	201	35
TRUE	0	166
Performance Rate / Misclassification Error Rate	91.3% / 0.087	
CART	FAKE	TRUE
FAKE	194	15
TRUE	7	186
Performance Rate / Misclassification Error Rate	94.6% / 0.054	
NNET	FAKE	TRUE
FAKE	201	22
TRUE	0	179
Performance Rate / Misclassification Error Rate	94.6% / 0.054	
SVM	FAKE	TRUE
FAKE	201	35
TRUE	0	166
Performance Rate / Misclassification Error Rate	91.3% / 0.087	
RF	FAKE	TRUE
FAKE	201	23
TRUE	0	178
Performance Rate / Misclassification Error Rate	94.3% / 0.057	

C. MODEL EVALUATION

As presented above, a fake news detection model was finally established using five machine learning techniques. We applied the k-fold cross-validation method as a method for constructing model optimization. Recently, data resampling methods such as k-fold cross validation and bootstrapping were applied to reduce the uncertainty of input dataset partition [42]. As previously performed, high performance can be achieved in the corresponding data when a model is built using all the given data set. Still, the prediction accuracy may be lowered when new data is added. A representative way to verify whether such overfitting occurs and to solve the overfitting problem is to use some of the given data as training data to build a model and use the rest as a test dataset for the model [43]. In this study, 402 data sets were divided into seven sets by setting the k value as 7 for each model of fake news detection using five machine learning techniques.

We compare the accuracy, precision, recall, and F1-score of the proposed each model to evaluate each model. Multiple evaluation metrics, including the accuracy, precision, F1-score, recall, and specificity, were adopted to evaluate the performances of the established model. The accuracy is the ratio of the number of samples correctly classified to the total number of samples in a given test dataset. The precision is the ratio of the true positive samples to the sum of the true positive and false positive samples. The F1-score is the weighted average of precision and recall. The recall indicates to the ratio of the true positive samples to the sum of the true positive and false negative samples. The F1-score value is used to evaluate the success of machine learning algorithms [44]. The specificity is the true negative rate. The accuracy, precision, F1-score, recall, and specificity are explained as follows, where TP, TN, FP, and FN represent the numbers of true positive, true negative, false positive, and false negative samples in the confusion matrix, respectively.

$$Accuracy = TP + TN/TP + TN + FP + FN \tag{4}$$
$$Precision = TP/TP + FP \tag{5}$$
$$Recall = TP/TP + FN \tag{6}$$
$$F1{-}score = 2 \times Precision \times Recall \, /Precision + Recall \tag{7}$$
$$Specificity = TN/FP + TN \tag{8}$$

TABLE 3. Results of evaluation metrics of five machine learning classifiers.

Model	LR	CART	NNET	SVM	RF
Accuracy	0.918	0.967	0.918	0.917	0.951
Precision	0.857	0.966	0.857	0.852	0.910
Recall	1.000	0.967	0.900	1.000	1.000
F1-score	0.923	0.952	0.923	0.931	0.952
Specificity	0.839	0.967	0.936	0.838	0.903

TABLE 3 shows that the accuracy, precision, recall, F1-score and specificity of the proposed each model.

D. COMPARE THE PERFORMANCE RATE

We simulated 1000 iterations of each model to analyze the effect of data and model learning. Therefore, it measures how frequently the classification model misclassifies in 1000 different simulations. In other words, "misclassification rates in

1000 simulations'' refer to a metric that shows how accurately a specific classification model performs. This type of simulation can help evaluate the accuracy of the classification model and can be used to improve the model's performance.

As a result of the iterations of the suggested model, the accuracy of the RF was about 94.1%, showing the best performance rate among the five models, while the accuracy of the NNET was the lowest at about 92.1%. In addition, the LR and CART showed a performance rate of approximately 93.1% and 92.8%, respectively. As described above, the average value of the misclassification rate according to 1000 iterations of each model and the performance according to the final prediction accuracy is presented in TABLE 4 and FIGURE 2.

TABLE 4. Misclassification rates in 1000 simulations.

Model	LR	CART	NNET	SVM	RF
Misclassification Error Rate	0.0689	0.0721	0.0786	0.0754	0.0590
Performance Rate	93.11	92.79	92.14	92.46	94.10

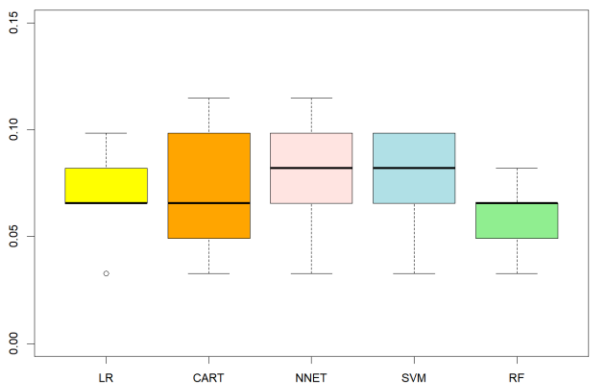

FIGURE 2. Misclassification rates in 1000 simulations.

E. ABLATION STUDY

As an additional experiment to identify the impact of the different features we performed an ablation study for the RF which was the best-performing model for each feature group. As with the prior model analysis, understanding the effect of each feature and which features are redundant is important for future model development [45]. An ablation study can show the effect of removing specific features. Furthermore, the study is needed to confirm results from the established XGBoost model. In particular, we have organized each feature group, based on the feature importance, called the *"Feature Group A, B, C"*, respectively. For example, *"Feature Group A"* include the number of the top four features which were ranked in the XGBoost model. Therefore, *"Feature Group A"* include the feature as follows: word sentiment, in-degree centrality, word similarity, the number of total tweets. *"Feature Group B"*, include additional feature as the number of followings based on *"Feature Group A" and, "Feature Group C"*, include the number of total retweets, based on *"Feature Group B"*. Then the performances of the optimal feature group which was derived by XGBoost model and the other feature groups are compared with one another and the results shown in TABLE 5. The results of the ablation study confirmed that the combination of features in *"Feature Group A"*, as we mentioned above, has the best performance.

TABLE 5. Results of ablation study on the rf model of each feature group.

Feature group	Test Accuracy (%)	Finding
A	94.10	Highest accuracy
B	93.27	Modest accuracy
C	93.02	Lowest accuracy

VI. CONCLUSION

We constructed five classification machine learning models to identify fake news spread and shared through Twitter and compare their performance rate. The main findings of the study can be summarized as follows. First, we derived the feature importance of various explanatory variables estimated to impact the identification of fake news spreading on Twitter.

Four major explanatory factors affecting fake news detection among various factors were finally extracted, and models for each machine learning algorithm were constructed based on those derived factors. These variables could be explained to significantly contribute to the construction of a fake news detection model in the following order: word sentiment, in-degree centrality, word similarity, and a total number of tweets.

Fake news detection models were established based on five machine learning algorithms: LR, NNET, CART, SVM, and RF with the top four derived variables. Second, the CART model and the NNET model showed the highest performance rate, about 94.6% among five classification machine learning models. The LR model and the SVM model, on the other hand, indicated about 91.3% performance rate, which had the lowest prediction rate. Third, the performance of the constructed models was evaluated with the misclassification rate. As the primary purpose of this study is to identify the optimal model for detecting fake news with the highest prediction rate, additional analysis was performed to compare the performance of each model. We established an evaluation model in the following way to solve the data imbalance problem that occurs in constructing the optimal of each model. Cross-validation, the entire data was divided into the training and test data set and inputted into the model establishment process. The data imbalance problem was alleviated through the oversampling technique. Based on the data set configured, a simulation was performed in which each model was repeated 1000 times. As a result, the RF model showed the highest performance rate of 94.1% in identifying fake news among fake and fact news, whereas the NNET model showed the lowest prediction rate of about 92.1%.

Based on the research results presented above, this study derived the following contributions. First, this study is meaningful because it derived various variables that were not previously considered as factors for detecting fake news, including the word sentiment of fake news content. In particular, according to previous studies, as it is known that fake news uses more words that can provoke negative emotions in readers than news of general information, attention has been needed to the negative tone of fake news [38]. It can be inferred that fake news appeals to the negative emotions of readers and urges their acceptance, whereas news containing fact uses neutral words that are relatively unbiased. This is because fake news makes the reader aware of the user's sense of crisis about negative sentiments such as fear, anger, and anxiety. Similarly, this study found that the word sentiment, which measures whether a message has a strong negative or positive tone, has the most significant influence on establishing a model for classifying fake news. Therefore, it can be inferred that the type of word sentiment of the written tweet is a significant factor in identifying fake news if it is written in a state that is overly biased in one direction. It means that it is necessary to doubt fake news if a message contains too much positive meaning and the use of negative words. Fake news related to healthy food, for example, often uses too many positive words to entice the reader to a variety of bodily positive effects that have not been scientifically verified.

Second, this study identifies the relative feature importance of variables, although various factors for determining fake news have been suggested through previous studies. Today attempts to develop fake news detection systems and automated algorithms for classifying fake news are actively performed. However, only a few studies concentrate on users' attitudes toward fake news for constructing models to detect fake news. Therefore, this study contributes to extending future related studies in that the importance of variables among various factors was derived. This study can be a foundation to develop the fake news detection systems that are currently being introduced by various social media platforms. If all aspects are considered in designing the system, it has the advantage of being able to predict in various situations, but the performance of the prediction rate is inevitably lowered due to the model fitness problem. Accordingly, the results of this study are presented in a priority of their relative importance while at the same time deriving various factors that have not been dealt with before and reflecting them in the model. This is expected to contribute to selecting factors that should be prioritized and considered in designing a system for detecting fake news in companies on social media platforms in the future.

Third, the results of this study address the needs for performing further research on ensemble models to detect fake news and XAI (Explainable AI) to extend the reliability of each model. RF model showed the highest accuracy among the five models in this study, and it was verified that ensemble model generally has the better performance than the single model. Ensemble models outperform single models as they use multiple models together to improve the performance of an algorithm, rather than using a single model. However, ensemble models can be more complex and difficult to interpret than single models. This is because the final prediction is generated by combining the outputs of multiple models, which can make it harder to understand how individual models contribute to the final result. Therefore, it can be considered to use model selection techniques such as grid search or Bayesian optimization to increase the accuracy of fake news detection with the ensemble model including RF proposed in this study and to increase the explainability of a single model with high contribution. It can help to avoid including redundant or poorly-performing models in the ensemble to identify the best combination of models and hyperparameters. Finally, it needs to be investigated how we can establish autonomous fake news detection system with not only the higher accuracy but also with explainable AI model.

This study suggests the following contributions as integrating the main findings of results. This study identified that the number of total tweets of each social media account, which has not been considered in past studies, is a significant factor on fake news detection. It can be inferred that users who upload more tweets than necessary, that are intentionally made by AI-bots to automatically generate a large amount of content in a short time. Therefore, it suggests that in the future, when developing a system to identify fake news, the number of uploaded contents of the user account need to be considered and identified comprehensively to improve the detection accuracy.

This study established a model of fake news spread through Twitter. Twitter is different from other social media because it spreads mainly through short text messages, so it can be expected that there will be differences from other social media in spreading fake news. Therefore, in future research, we plan to collect data related to fake news distributed through other platforms to reduce the difference between social media platforms and generalize the research results. This study also could not consider various cultures by collecting fake news messages written in only English without considering various languages. It has been found that the cultures affect the users' behaviors on social media, such as the frequency of posting and the intentions of sharing messages [46]. It can be inferred that there can be differences in the strength of the relationship (i.e., the degree of centrality of users in social media) and the behaviors of accepting or spreading fake news according to the user's cultural norms. Therefore, it provides the future directions for improving the existed fake news detection system.

While various algorithms for detecting fake news are being actively advanced, this study constructed AI models based on the derived priority factors from the perspective of social capital theory. We proposed an optimized model for detecting fake news by reflecting the feature of the information receiver and social network. It finally suggests the need to develop an algorithm with an excellent prediction rate and fully reflect the social network and the characteristics of participants who maintain the network to develop an automated fake news detection system.

REFERENCES

[1] C. Wu, F. Wu, Y. Huang, and X. Xie, "Personalized news recommendation: Methods and challenges," *ACM Trans. Inf. Syst.*, vol. 41, no. 1, pp. 1–50, Jan. 2023.

[2] X. Su, G. Sperlì, V. Moscato, A. Picariello, C. Esposito, and C. Choi, "An edge intelligence empowered recommender system enabling cultural heritage applications," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 4266–4275, Jul. 2019.

[3] F. Zhou, X. Xu, G. Trajcevski, and K. Zhang, "A survey of information cascade analysis: Models, predictions, and recent advances," 2020, *arXiv:2005.11041*.

[4] S. Gaillard, Z. A. Oláh, S. Venmans, and M. Burke, "Countering the cognitive, linguistic, and psychological underpinnings behind susceptibility to fake news: A review of current literature with special focus on the role of age and digital literacy," *Front. Commun.*, vol. 6, Jul. 2021, Art. no. 661801.

[5] M. L. D. Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro, and L. de Alfaro, "Automatic online fake news detection combining content and social signals," in *Proc. 22nd Conf. Open Innov. Assoc. (FRUCT)*, May 2018, pp. 272–279.

[6] E. C. Tandoc, Z. W. Lim, and R. Ling, "Defining "fake news": A typology of scholarly definitions," *Digit. Journalism*, vol. 6, no. 2, pp. 137–153, Feb. 2018.

[7] H. Kim, "An exploratory study on fake news using topic modeling: Focused on fake news published in the online journalism," M.S. thesis, School Manag. Inf. Syst., Kookmin Univ., Seoul, South Korea, 2017.

[8] A. K. Cybenko and G. Cybenko, "AI and fake news," *IEEE Intell. Syst.*, vol. 33, no. 5, pp. 1–5, Sep. 2018.

[9] N. Seddari, A. Derhab, M. Belaoued, W. Halboob, J. Al-Muhtadi, and A. Bouras, "A hybrid linguistic and knowledge-based analysis approach for fake news detection on social media," *IEEE Access*, vol. 10, pp. 62097–62109, 2022.

[10] A. Galli, E. Masciari, V. Moscato, and G. Sperlí, "A comprehensive benchmark for fake news detection," *J. Intell. Inf. Syst.*, vol. 59, no. 1, pp. 237–261, Aug. 2022.

[11] A. P. Salazar, "AI tools on fake news detection: An overview and comparative study," Graduate School Technol. Univ. Philippines, Manila, Philippines, Tech. Rep., 2020.

[12] A. Kim, I. University, A. R. Dennis, and I. University, "Says who? The effects of presentation format and source rating on fake news in social media," *MIS Quart.*, vol. 43, no. 3, pp. 1025–1039, Jan. 2019.

[13] N. K. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proc. Assoc. for Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, Jan. 2015.

[14] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of varying shades: Analyzing language in fake news and political fact-checking," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2017, pp. 2931–2937.

[15] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic detection of fake news," 2017, *arXiv:1708.07104*.

[16] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *Proc. IEEE 1st Ukraine Conf. Electr. Comput. Eng. (UKRCON)*, May 2017, pp. 900–903.

[17] Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: Recognizing clickbait as 'false news,'" in *Proc. ACM Workshop Multimodal Deception Detection*, Nov. 2015, pp. 15–19.

[18] C. Chen, K. Wu, V. Srinivasan, and X. Zhang, "Battling the Internet water army: Detection of hidden paid posters," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2013, pp. 116–120.

[19] Z. Jin, J. Cao, Y. Jiang, and Y. Zhang, "News credibility evaluation on Microblog with a hierarchical propagation model," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2014, pp. 230–239.

[20] V. L. Rubin, Y. Chen, and N. K. Conroy, "Deception detection for news: Three types of fakes," *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, 2015.

[21] R. J. Sethi, "Spotting fake news: A social argumentation framework for scrutinizing alternative facts," in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Jun. 2017, pp. 866–869.

[22] M. Carter, M. Tsikerdekis, and S. Zeadally, "Approaches for fake content detection: Strengths and weaknesses to adversarial attacks," *IEEE Internet Comput.*, vol. 25, no. 2, pp. 73–83, Mar. 2021.

[23] M. M. Wasko and S. Faraj, "Why should I share? Examining social capital and knowledge contribution in electronic networks of practice," *MIS Quart.*, vol. 29, no. 1, pp. 35–57, Mar. 2005.

[24] J. Nahapiet and S. Ghoshal, "Social capital, intellectual capital, and the organizational advantage," *Acad. Manage. Rev.*, vol. 23, no. 2, pp. 242–266, Apr. 1998.

[25] P. Dawson, J. Scott, J. L. Thompson, and D. Preece, "The dynamics of innovation and social capital in social enterprises: A relational sense-making perspective," in *Proc. Massey Univ. Social Innov. Entrepreneurship Conf.*, 2011, pp. 177–191.

[26] N. Lin, "Building a network theory of social capital," *Connections*, vol. 22, no. 1, pp. 28–51, 1999.

[27] C.-M. Chiu, M.-H. Hsu, and E. T. G. Wang, "Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories," *Decis. Support Syst.*, vol. 42, no. 3, pp. 1872–1888, Dec. 2006.

[28] D. R. Bild, Y. Liu, R. P. Dick, Z. M. Mao, and D. S. Wallach, "Aggregate characterization of user behavior in Twitter and analysis of the retweet graph," *ACM Trans. Internet Technol.*, vol. 15, no. 1, pp. 1–24, Mar. 2015.

[29] B. Gonçalves, N. Perra, and A. Vespignani, "Modeling users' activity on Twitter networks: Validation of Dunbar's number," *PLoS ONE*, vol. 6, no. 8, Aug. 2011, Art. no. e22656.

[30] M. Cha, H. Haddadi, F. Benevenuto, and P. K. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," in *Proc. ICWSM*, vol. 10. nos. 10–17, 2010, p. 30.

[31] X. Chen, Z. Chong, P. Giudici, and B. Huang, "Network centrality effects in peer to peer lending," *Phys. A, Stat. Mech. Appl.*, vol. 600, Aug. 2022, Art. no. 127546.

[32] R. Garcia-Gavilanes, D. Quercia, and A. Jaimes, "Cultural dimensions in Twitter: Time, individualism and power," in *Proc. 7th Int. AAAI Conf. Weblogs Social Media*, 2013, pp. 195–204.

[33] D. Gayo-Avello, "Nepotistic relationships in Twitter and their impact on rank prestige algorithms," *Inf. Process. Manage.*, vol. 49, no. 6, pp. 1250–1280, Nov. 2013.

[34] E. Tsourougianni and N. Ampazis, "Recommending who to follow on Twitter based on tweet contents and social connections," *Social Netw.*, vol. 2, no. 4, pp. 165–173, 2013.

[35] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. 19th Int. Conf. World Wide Web*, Apr. 2010, pp. 591–600.

[36] D. Kim and Y. Lee, "News based stock market sentiment lexicon acquisition using Word2Vec," *Korea J. BigData*, vol. 3, no. 1, pp. 13–20, Aug. 2018.

[37] Y. Dang, Y. Zhang, and H. Chen, "A lexicon-enhanced method for sentiment classification: An experiment on online product reviews," *IEEE Intell. Syst.*, vol. 25, no. 4, pp. 46–53, Jul. 2010.

[38] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.

[39] A. Dey, "Machine learning algorithms: A review," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 3, pp. 1174–1179, 2016.

[40] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.

[41] L. Torlay, M. Perrone-Bertolotti, E. Thomas, and M. Baciu, "Machine learning–XGBoost analysis of language networks to classify patients with epilepsy," *Brain Inform.*, vol. 4, no. 3, pp. 159–169, Sep. 2017.

[42] X. C. Nguyen, T. T. H. Nguyen, D. D. La, G. Kumar, E. R. Rene, D. D. Nguyen, S. W. Chang, W. J. Chung, X. H. Nguyen, and V. K. Nguyen, "Development of machine learning–based models to forecast solid waste generation in residential areas: A case study from Vietnam," *Resour., Conservation Recycling*, vol. 167, Apr. 2021, Art. no. 105381.

[43] Y. Bengio and Y. Grandvalet, ''No unbiased estimator of the variance of K-fold cross-validation,'' *J. Mach. Learn. Res.*, vol. 5, pp. 1089–1105, Dec. 2004.

[44] F. Khan, I. Tarimer, H. S. Alwageed, B. C. Karadağ, M. Fayaz, A. B. Abdusalomov, and Y.-I. Cho, ''Effect of feature selection on the accuracy of music popularity classification using machine learning algorithms,'' *Electronics*, vol. 11, no. 21, p. 3518, Oct. 2022.

[45] D. Kauchak, O. Mouradi, C. Pentoney, and G. Leroy, ''Text simplification tools: Using machine learning to discover features that identify difficult text,'' in *Proc. 47th Hawaii Int. Conf. Syst. Sci.*, Jan. 2014, pp. 2616–2625.

[46] I. Pentina, L. Zhang, and O. Basmanova, ''Antecedents and consequences of trust in a social media brand: A cross-cultural study of Twitter,'' *Comput. Hum. Behav.*, vol. 29, no. 4, pp. 1546–1555, Jul. 2013.

[47] N. Karimi and J. Gambrell. (Mar. 27, 2020). *Hundreds Die of Poisoning in Iran as Fake News Suggests Methanol Cure for Virus*. The Times of Israel, erusalem, Israel. [Online]. Available: https://url.kr/w3c8hd

[48] K. Miller. (Feb. 13, 2023). *Human Writer or AI? Scholars Build a Detection Tool*. Stanford Univ. Hum.-Centered Artif. Intell., Stanford, CA, USA. [Online]. Available: https://url.kr/w3c8hd

[49] A. A. A. Ahmed, A. Aljabouh, P. K. Donepudi, and M. S. Choi, ''Detecting fake news using machine learning : A systematic literature review,'' 2021, *arXiv:2102.04458*.

[50] J. Botha and H. Pieterse, ''Fake news and deepfakes: A dangerous threat for 21st century information security,'' in *Proc. 15th Int. Conf. Cyber Warfare Secur. (ICCWS)*. New York, NY, USA: Academic Conferences and publishing limited, Mar. 2020, pp. 1–10.

[51] W. Y. Wang, '''Liar, liar pants on fire': A new benchmark dataset for fake news detection,'' 2017, *arXiv:1705.00648*.

[52] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, ''FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media,'' *Big Data*, vol. 8, no. 3, pp. 171–188, Jun. 2020.

MINJUNG PARK received the M.S. and Ph.D. degrees in data analytics and business administration from Ewha Womans University, Seoul, South Korea, in 2016 and 2021, respectively. Her current research interests include artificial intelligence (AI), big data analytics, blockchain, information security, and privacy.

SANGMI CHAI received the M.S. degree in business administration from Seoul National University, South Korea, and the Ph.D. degree in management information systems from the School of Management, The State University of New York at Buffalo. She is currently a Professor with the School of Business, Ewha Womans University, South Korea. She has published her articles in *Journal of Management Information Systems*, *International Journal of Production Economics*, *International Journal of Logistics Management*, *Decision Support Systems*, IEEE Transactions on Professional Communication, *International Journal of Information Management*, and *Information Systems Frontiers*. Her current research interests include artificial intelligence (AI), blockchain, privacy, and information security.