ISSN 2288-4866 (Print)
ISSN 2288-4882 (Online)
http://www.jiisonline.org
J Intell Inform Syst 2019 September: 25(3): 201~220 http://dx.doi.org/10.13088/jiis.2019.25.3.201

A Study on the Effect of the Document Summarization Technique on the Fake News Detection Model*

Jae-Seung Shim
Graduate School of Business IT, Kookmin University
(*simong_kwnau@naver.com*)

Ha-Ram Won
Graduate School of Business IT, Kookmin University
(*haramy44@kookmin.ac.kr*)

Hyunchul Ahn
Graduate School of Business IT, Kookmin University
(*hcahn@kookmin.ac.kr*)

Fake news has emerged as a significant issue over the last few years, igniting discussions and research on how to solve this problem. In particular, studies on automated fact-checking and fake news detection using artificial intelligence and text analysis techniques have drawn attention. Fake news detection research entails a form of document classification; thus, document classification techniques have been widely used in this type of research. However, document summarization techniques have been inconspicuous in this field. At the same time, automatic news summarization services have become popular, and a recent study found that the use of news summarized through abstractive summarization has strengthened the predictive performance of fake news detection models. Therefore, the need to study the integration of document summarization technology in the domestic news data environment has become evident. In order to examine the effect of extractive summarization on the fake news detection model, we first summarized news articles through extractive summarization. Second, we created a summarized news-based detection model. Finally, we compared our model with the full-text-based detection model. The study found that BPN(Back Propagation Neural Network) and SVM(Support Vector Machine) did not exhibit a large difference in performance; however, for DT(Decision Tree), the full-text-based model demonstrated a somewhat better performance. In the case of LR(Logistic Regression), our model exhibited the superior performance. Nonetheless, the results did not show a statistically significant difference between our model and the full-text-based model. Therefore, when the summary is applied, at least the core information of the fake news is preserved, and the LR-based model can confirm the possibility of performance improvement. This study features an experimental application of extractive summarization in fake news detection research by employing various machine-learning algorithms. The study's limitations are, essentially, the relatively small amount of data and the lack of comparison between various summarization technologies. Therefore, an in-depth analysis that applies various analytical techniques to a larger data volume would be helpful in the future.

Key Words : Fake News Detection, Document Summarization, Automated Fact Checking, Machine Learning, Domestic News

Received : August 8, 2019 Revised : September 8, 2019 Accepted : September 20, 2019
Publication Type : Concise Paper Corresponding Author : Hyunchul Ahn

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A01039377).

1. Introduction

Fake news has recently become a significant global issue, resulting in various studies that investigate this issue (Bondiellia and Marcelloni, 2019). In particular, research on automated fact-checking and fake news detection models using text analysis techniques has been actively conducted. This field of research is basically belongs to the text classification research, and various techniques for extracting the features of texts—such as topic modeling, sentiment analysis, network analysis—have been widely used to extract the features of documents for classification purposes (Yun and Ahn, 2019; Hyun and Kim, 2018; Castillo, 2011; Park and Lee, 2019). However, document summarization techniques have seldom been considered in this field.

But it might be possible to improve detection model performance through summarization techniques. Document summarization may eliminate noise, making the features of fake news more prominent. A recent study showed that the use of news summarized through abstractive summarization has strengthened the predictive performance of fake news detection models (Esmaeilzadeh et al., 2019). Besides, when detecting fake news in Korea, it is necessary to verify whether the same level of judgment can be achieved with summarized news, rather than having to check the original full text. Currently, given the tens of thousands of articles that are produced daily, news summarization services are already popular (Lee and Kim, 2009). For example, NAVER and DAUM launched a News Summary Bot service that extracts and summarizes news articles. Thus, readers can scan a long article at a glance with just one button. As well as NAVER and DAUM, a variety of companies operate 'News Summary' and 'News Briefing' services. (Jin, 2017; Koo, 2017). As the news summary service is active in Korea, the information distortion caused by the news summary service is also raised (Shin, 2017). Thus, we must reflect these developments in current fake news detection research.

Under this background, the present study employs extractive summarization—which is a form serviced in Korea and not covered in any prior studies—to conduct fake news detection research through domestic news data. The study attempts to confirm whether the key information is condensed, lost, or retained. Specifically, this is verified in terms of the model's performance. Two different kinds of models are created—using summarized news data and the original full text, respectively—and statistical hypothesis testing is conducted by comparing the two models' performance. The study's hypotheses are as follows:

Null hypothesis (H0). There is no difference in performance between the summary-based model and full-text-based models. (The key information is preserved as is.)

Alternative hypothesis-a (Ha). The performance of the summary-based model is superior to that of the full-text-based model. (The key information is condensed.)

Alternative hypothesis-b (Hb). The performance of the full-text-based model is superior to that of the summary-based model. (The key information is lost.)

2. Theoretical Background

2.1 Fake news

'Fake news' is a general term in Korean terminology. Its definition and scope are defined in different ways depending on the context in which it is used and the perspective through which it is interpreted. The Korea Press Foundation defines fake news as "False information intentionally formatted and disseminated for the purpose of political and economic advantage." Wikipedia defines the term as "a kind of yellow journalism that draws attention by stimulating people's interests and instincts." From a legal point of view, several bills defining fake news have been proposed as of 2019. The definitions in each bill are similar. However, the definitions differ in their consideration of concepts such as 'intentionality', 'political/economic benefit', and 'third-party harm' (Yoon, 2019). The Fake News Distribution Prevention Act defined fake news more comprehensively as, "information released by the press, information that the press has admitted is untrue through a correction report" and "information determined by the Press Arbitration Commission to contain untrue content." In this study, the scope of fake news is limited to news articles that violate the above-mentioned Fake News Distribution Prevention Act or that are found to be untrue based on the SNU Fact Check.

2.2 Text Summarization

A summary is defined as a text that is generated from one or more texts. It conveys the original text's key information but is shorter than half the original text (Radev et al., 2002). Text summarization refers to a system or method for automatically summarizing text. There are two types of text summarization, depending on how the summary is generated. The first is extraction, which analyzes the original text's elements and extracts the most important ones into summary statements. The second is abstraction, which creates a new summarized document considering the meaning and logic of the original text's whole sentences. Of these two, extraction is relatively easier to perform and is therefore more widely used (Allahyari et al., 2017; Radev et al., 2002; Yun et al., 2019).

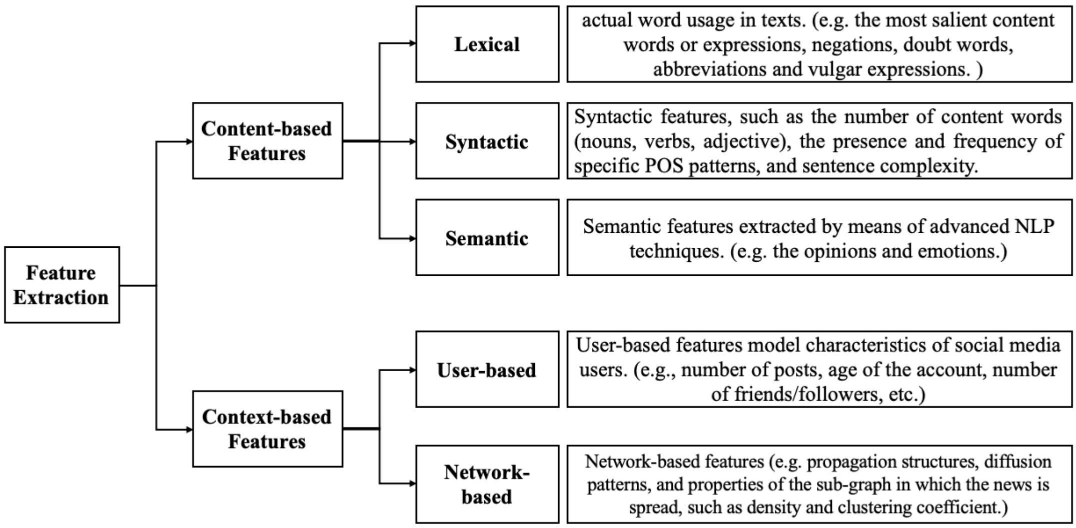
A representative extraction method is the statistical-based approach, which extracts important sentences through statistical methods. Basically, the core sentences are extracted by calculating the importance of each word and each sentence through word frequency. In addition to word frequency, there are various ways to calculate the importance of words and sentences. These additional ways include TF-IDF, sentence position, frequency of positive keywords (keywords that appear frequently in important sentences), frequency of negative keywords (keywords that appear frequently in unimportant sentences), centrality indicators, similarity, sentence length, and the presence or absence of numerical data in each sentence. Extraction can also be conducted through other approaches beyond the statistical-based approach. If the text's subject is identified, a topic-based approach can be used. This is a method of extracting sentences that feature similarities with the subject. A graph-based approach is also commonly used—in this method, a diagram is created between sentences to set each sentence as a node and the similarity for each sentence as an edge. Then, the most important sentences are identified. The rhetorical structure-based approach locates important sentences by applying the theory of rhetorical structure, while the machine-learning-based approach classifies sentences as important and less important by inputting various features into a machine-learning classification model.

On the other hand, abstraction is largely divided into three approaches. First, in the form-based approach, a template is created for an appropriate summary statement based on the pattern of the literature. This template extracts the necessary information from the literature to fill in the gaps in the form. Second, in the graph-based approach, a graph is generated by recognizing words as nodes and by recognizing the adjacency between the words as edges. Through this, the importance of each element is calculated, and the sentence is reconstructed with a grammatically significant path. The third option for abstraction is the semantic-based approach, which uses Natural Language Understanding and Natural Language Generation technologies.

Lexrankr is a system that performs high-quality extractive summarization algorithms for long documents with diverse topics, suitable for use with the Korean language. The present study used Lexrankr to create three-sentence summaries, which is the same as how NAVER's and DAUM's current news summary services operate. Lexrankr is the Korean version of LexRank, which applies Google's PageRank algorithm as a sentence unit. The key principle can be explained through PageRank, which is the foundation of Google's search engine. PageRank calculates the importance of each page through the principle that important web pages are linked to by more sites. LexRank, which applies this approach to sentence importance calculations, is part of the graph-based approach. According to Seol and Lee (2016), Lexrankr conducted tests based on various verb and similarity functions in the process of applying the Lexrank algorithm to the Korean language. Lexrankr was then optimized to perform Korean document summarization by applying the optimal settings identified through the experiment.

2.3 Fake news detection

Thus far, fake news detection research has been conducted under two broad categories of features. These are content-based features—based on vocabulary, sentences, and semantic characteristics—and context-based features, which refer to the



〈Figure 1〉 Feature extraction method of fake news (Alessandro, and Francesco., 2019)

characteristics of users, sources, and/or networks that distribute fake news. Most prior studies have examined a variety of characteristics belonging to both categories.

From a modeling perspective, a variety of machine-learning and deep-learning algorithms were used in the detection model of the SVM (Supporting Vector Machine). Zhang et al.'s (2012) model showed the highest performance in Afroz et al.'s (2012) fake text detection model. The former also exhibited excellent performance in fake news detection models that use only content-based information (Bondielia and Marcelloni, 2019). DT(decision tree) also exhibited effective performance in the fake news study, with a validation accuracy of 86% in the detection model. DT applied both context and content information, as proposed by Castillo et al. (2011), and it has been used in a number of additional studies (Aker et al., 2017; Giasemidis, 2016). In terms of LR (logistic regression), a study predicting the attitudes of media companies dealing with fake news showed a 73% accuracy rate in a validation set (Ferreira and Vlachos, 2016). According to Hardalov et al. (2016), it detected fake online news with 75% accuracy. In addition, CRF (Conditional Random Field), RNN (Recurrent Neural Networks), and CNN (Convolutional Neural Networks) have also been employed in fake news detection research (Lafferty et al., 2001; Zubiga et al., 2016).

Ever since Yun and Ahn (2018) first proposed a fake news detection system based on artificial intelligence that utilizes news data from SNU Fact Check, numerous research studies on fake news detection have been conducted in Korea. Recently, research using Twitter data in combination with news data was conducted to overcome the limitations of fake news detection research in Korea, which include the lack of data and features in detecting fake news. For instance, Hyun and Kim (2018) presented an improved AI fake news classification model by utilizing SNS response data based on the news along with news data in a strict experimental environment. In another study, Park and Lee (2019) applied the network embedding technique DeepWalk to network structure detection of Twitter accounts referring to the news. The researchers then compared the results with the results of traditional frequency-based text analysis, successfully demonstrating the usefulness of their proposed model.

3. Research Model

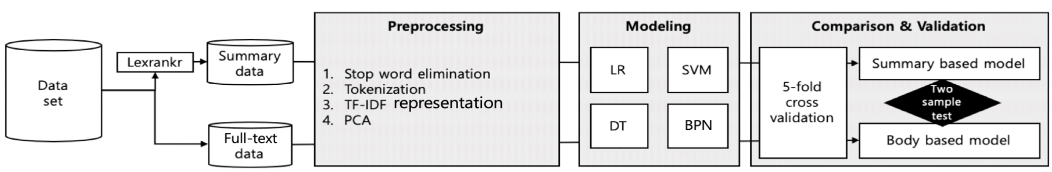
Figure 2 presents a flow chart of this study's research model. First, Lexrankr was applied to the full text of a news article in the collected dataset to create a three-line summary. Next, the summary and full text were preprocessed following the same procedure. Preprocessing consisted of stop word elimination, tokenization, TF-IDF representation, and PCA (Principal Component Analysis). Then, the datasets of the preprocessed "Full-text" and "summary" were divided into training and validation datasets and applied to each of the four machine-learning models to create the full-text-based and summary-based fake news detection models. Finally, the two models were compared and verified.

4. Empirical analysis

4.1 Experimental Data

The data used in this study consist of fake news and general news. A total of 50 fake news data were collected from SNU Fact Check websites－which have been primarily used in existing domestic text-based fake news detection studies－and the Press Arbitration Commission. Another 50 regular news data were collected from Naver News. Table 1 presents examples of fake news data that were collected.

As Figure 3 shows, general news corresponding to fake news was collected using the same keyword-based 1:1 matching method to minimize the possibility of overfitting by the subject, which is one of the vulnerabilities in the text-based news detection model.



〈Figure 2〉 Flow chart of the research model

〈Table 1〉 Fake news examples

Date	Title	Full-text	Source	Evidence	Col.date	Result
2019.3.22	Former President Chun Doo-hwan's penalty expires in October next year.	Former President Chun Doo-hwan's house in Yeonhui-dong has been auctioned off at a public auction, although it is expected to take a considerable amount of time for the actual auctioneer to take over ownership, as the statute of limitations expires in October next year…	SNU Fact Check (http://news.heraldcorp.com/view.php?ud=20190322000182)	http://factcheck.snu.ac.kr/v2/facts/1443	2019.4.9	F
2017.11.1	the revival of the women's military system in 43 years	The Ministry of National Defense has revived the female recruitment system for the first time in 43 years as an alternative to filling the insufficient number of troops due to the reduction of the active duty period…	PCA(2017 coordination arbitration case book1. Pdf)	2017-Seoul-1995,1996	2019.4.9	F

* *Col.date : Date of collection*

〈Figure 3〉 Data collection method (Collected with the same keyword)

After data collection, Lexrankr was applied to the full texts to create a three-line summary of each one. Figure 4 enables a visual comparison of the summaries and the full texts with word cloud.

Genuine news (Full-text)	Fake news (Full-text)
Genuine news (Summary)	Fake news (Summary)

〈Figure 4〉 Top 40 word cloud frequencies

4.2 Data Preprocessing

The data preprocessing stage consisted of the following sequence: stop word elimination, tokenization, TF-IDF vectorizing, and PCA dimensional reduction. First, meaningless special symbols, which are unrelated to an article's context but are likely to be considered as meaningful components of the sentences, were selected as stop words to be removed. Table 2 lists the stop words that were removed during this phase. Second, in the tokenization phase, we used the Korean natural language processing package, KoNLPy - Okt class, for tagging. Morphemes were designated as the extraction units. Through this tagging process, 7,045 de-duplicated morphemes were extracted from the full texts and 2,642 from the summaries. Third, we created the TF-IDF value based on each morpheme. Various embedding techniques were available for use at this stage, but vectorization was ultimately carried out through the TF-IDF, which had already been used for conventional fake news research and user review mining (Jeon and Ahn, 2015; Yun and Ahn, 2019; Hyun and Kim, 2018). Considering the data size, we selected the abovementioned technique rather than embedding techniques such as word2vec. In order to use each morpheme's TF-IDF value as a feature of each news article, all morphemes were replaced by TF-IDF values. The blank values were filled in with zeros. Then, to prevent over-fitting caused by an excessive number of variables compared to the experimental data, a Principal Component Analysis (PCA) was performed. PCA can be used to reduce dimensions and derive several components that are more descriptive as potential variables through linear

〈Table 2〉 Stop word list

Stop word	name	Stop word	name
.	Period	()	Round Brackets
,	Comma	[]	Square Brackets
:	Colon	{}	Brace
;	Semicolon	/	Slash
‘ ’	Quotes	·	Bullet
“”	Double Quotes		

〈Table 3〉 Final independent and dependent variables

Variable Type		Variable Name	Description	
Independent variables	**Full-text**	Full-text_PCA1	Continuous variables	Decimal between 10 and -10
		Full-text_PCA2		
		Full-text_PCA3		
		…		
		Full-text_PCA20		
	Summary	Summary_PCA1		
		Summary_PCA2		
		Summary_PCA3		
		…		
		Summary_PCA20		
Dependent variable		Label	Discrete variables	0: Normal / 1: Fake

combinations that preserve variables' variability in multi-dimensional data (Kim, 2006). PCA was performed through the SPSS, and the criteria for setting up the analysis options were as follows. Since all variables were standardized, a correlation matrix was selected for the analysis, and the top 20 fixed factors were extracted. Varimax was selected as the rotation method. The factor score was the default setting, 'Regression'. Table 3 lists the final independent and dependent variables consisting of 20 fixed variables, each generated from full text and summary data.

4.3 Modeling

For the modeling stage, a full-text-based machine-learning model and a summary-based machine-learning model were created using the final variables from Table 3. Each set of data had a genuine news and fake news ratio of 5:5. In order to test the model using only the training data, 80% of the data were selected as training datasets while the remaining 20% were designated as validation datasets. Experiments were conducted on Python using logistic regression (LR), decision tree (DT), back processing network (BPN), and SVM (Support Vector Machine) algorithms as classification models. A 5-fold cross-validation was performed considering the small size of the data.

4.4 Comparison and Validation

At this stage, each validation dataset was applied to the pre-trained models to calculate the results. Hyper-parameter tuning was also performed to extract the model's best predictive performance in this process. To establish an equally controlled environment for the full-text and summary-based models, the hyper-parameter setting range by algorithm was set and adjusted within.

For LR, a stochastic gradient descent was used as an optimizer, with the alpha adjusted between 0.0001 and 1000. Three penalties were used: 'L1', 'L2', and 'elasticnet'. For DT, the Gini coefficient and entropy were selected as impure indicators, while the maximum number of features, the minimum sample leaves, and the minimum branch were used as hyper-parameters within the range of features. For BPN, the activation, solver, and hidden layers, the learning rate, and the maximum number of repetitions were adjusted. The scopes were set by referring to the setting of GridSearchCv in the Sklearn package. For SVM, the RBF kernel was used and the cost and beta were respectively adjusted between 0.001 to 1 and 1 to 5000. In addition, the non-referred hyper-parameters were used as default values. Moreover, the random state was set to 42 and applied to all classification algorithms. After the tuning and 5-fold validation were completed, the results of the two models were statistically compared through the two-sample test for proportions. Table 4 presents the optimal combination of parameters that can produce the results.

〈Table 4〉 Optimal parameter settings for each model

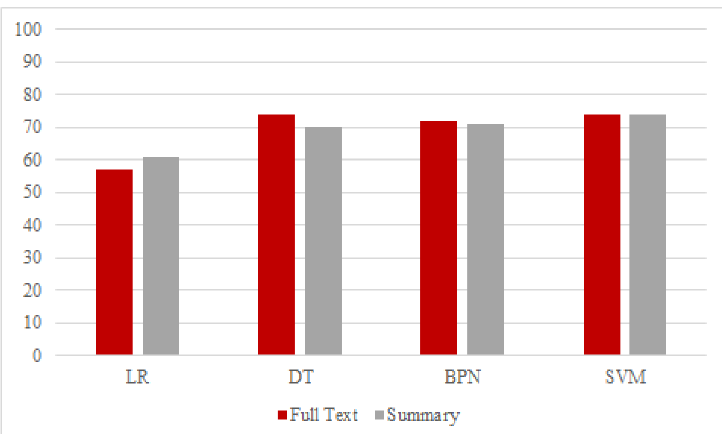
Method	Fold	Parameters	
		Full-text	Summary
LR (Stochastic Gradian Descent)	Fold1	alpha = 0.005, penalty = l1	alpha = 0.1, penalty = elasticnet
	Fold2	alpha = 0.01, penalty = l1	alpha = 0.01, penalty = elasticnet
	Fold3	alpha = 0.05, penalty = l1	alpha = 0.0001, penalty =l1
	Fold4	alpha = 0.01, penalty = elasticnet	alpha = 0.0001, penalty = elasticnet
	Fold5	alpha = 0.0001, penalty = l1	alpha = 0.01, penalty = l2
DT (CART)	Fold1	Criterion = gini, Max features = 9, min_samples_leaf = 4, min_samples_split = 2	Criterion = gini, Max features = 8, min_samples_leaf = 3, min_samples_split = 3
	Fold2	Criterion = entropy, Max features = 4, min_samples_leaf = 1, min_samples_split = 2,	Criterion = entropy, Max features = 4, min_samples_leaf = 1, min_samples_split = 2,
	Fold3	Criterion = entropy, Max features = 9, min_samples_leaf = 2, min_samples_split = 4	Criterion = entropy, Max features = 2, min_samples_leaf = 4, min_samples_split = 2
	Fold4	Criterion = gini, Max features = 3, min_samples_leaf = 3, min_samples_split = 2	Criterion = gini, Max features = 2, min_samples_leaf = 1, min_samples_split = 3
	Fold5	Criterion = gini, Max features = 6, min_samples_leaf = 2, min_samples_split = 2	Criterion = entropy, Max features = 4, min_samples_leaf = 1, min_samples_split = 4
BPN	Fold1	activation = logistic, solver = lbfgs, learning rate = constant, hidden layer sizes = 70, max iterations = 300	activation = tanh, solver = adam, learning rate = constant, hidden layer sizes = 30, max iterations = 800
	Fold2	activation = tanh, solver = sgd, learning rate = constant, hidden layer sizes = 70, max iterations = 300	activation = logistic, solver = adam, learning rate = constant, hidden layer sizes = 20, max iterations = 500
	Fold3	activation = tanh, solver = adam, learning rate = invscaling, hidden layer sizes = 80, max iterations = 800	activation = relu, solver = adam, learning rate = invscaling, hidden layer sizes = 50, max iterations = 200
	Fold4	activation = logistic, solver = lbfgs, learning rate = constant, hidden layer sizes = 30, max iterations = 100	activation = tanh, solver = adam, learning rate = constant, hidden layer sizes = 70, max iterations = 1000
	Fold5	activation = tanh, solver = lbfgs, learning rate = constant, hidden layer sizes = 30, max iterations = 400	activation = tanh, solver = adam, learning rate = adaptive, hidden layer sizes = 100, max iterations = 500
SVM (RBF Kernel)	Fold1	C = 1736, gamma = 0.066	C = 86, gamma = 0.061
	Fold2	C = 2116, gamma = 0.011	C = 1356, gamma = 0.026
	Fold3	C = 421, gamma = 0.036	C = 91, gamma = 0.056
	Fold4	C = 231, gamma = 0.036	C = 21, gamma = 0.036
	Fold5	C = 1906, gamma = 0.076	C = 111, gamma = 0.021

5. Results

Table 5 and Figure 5 feature the results according to each algorithm and folder. The performances of the training dataset and validation dataset were recorded based on the 5-fold mean. The results are as follows. In the logistic regression model, the t performance of the

〈Table 5〉 Experimental results for each classification method and fold

5-fold		LR		DT		BPN		SVM	
		training	validation	training	validation	training	validation	training	validation
1 Fold	Full	62.5	55	91.25	80	100	70	76.25	75
	Summary	62.5	60	90	70	87.5	75	83.75	80
2 Fold	Full	58.75	55	100	70	68.75	60	68.75	65
	Summary	70	65	100	65	73.75	65	73.75	70
3 Fold	Full	62.5	60	97.5	70	95	75	78.75	75
	Summary	66.25	65	85	70	78.75	70	87.5	75
4 Fold	Full	62.5	60	93.75	80	100	75	82.5	80
	Summary	61.25	60	98.75	70	97.5	70	76.25	75
5 Fold	Full	56.25	55	97.5	70	100	80	76.25	75
	Summary	68.75	55	95	75	85	75	82.5	70
Average	Full	60.5	57	**96**	**74**	**92.75**	**72**	76.5	74
	Summary	**65.75**	**61**	93.75	70	84.5	71	**80.75**	**74**



〈Figure 5〉 Experimental result graphs for each classification method full-text-based model exhibited 60.5% training and 57% validation, while it exhibited 65.75% and 61% validation for the summary-based model. In DT, the full-text-based model exhibited 96% training and 74% validation, while the summary-based model exhibited 93.75% training and 70% validation. In BPN, the full-text-based model exhibited 92.75% training and 72% validation, while the summary-based model exhibited 84.5% training and 71% validation 71%. Finally, in SVM, the full-text-based model exhibited 76.5% training and 74% validation, while the summary-based model exhibited 80.75% training and 74% validation. Of the tested algorithms, SVM performed the best with a value of 74% with the summary-based model. For the full-text-based model, SVM and DT performed the best with a value of 74%.

Table 6 presents the results of verifying the performance differences of the summary-based model and the full-text-based model through the two-sample test for proportions. In LR, the Z-value was -0.575 and the summary-based model was higher than the full-text-based model. In SVM, no difference was found. Meanwhile, in DT and BPN, the Z-values were 0.630 and 0.157, respectively, and the performance of the full-text-based model was measured to be slightly higher.

However, the statistical significance between the summary-based model and the full-text-based model was not verified in all algorithms, so the alternative hypothesis-a (Ha) and the alternative hypothesis-b (Hb) were rejected while the null hypothesis (H0) was not rejected. We were able to draw two conclusions through the experiment.

First, the core information was not lost in the fake news detection model using summaries generated by the automatic summary technique. This confirmed the possibility of successful information reduction through the fake news detection model. While 7,045 morphemes were used with the full-text-based model, only 2,642 morphemes were used with the summary-based model－in other words, about 1/3 of the amount of the full-text-based model. This means that performance degradation was possible because all procedures for both models were identical except for this single difference. However, no statistical difference was found between the performance of the summary-based model and the full-text-based model across all algorithms. In addition, the summary-based model using the LR algorithm showed a 4% improvement in prediction. Therefore, when the full news texts were summarized and used in the detection model, small yet positive effects were identified.

〈Table 6〉 Two-sample test for proportions (The full-text-based model/Summary based model)

Full-text/Summary	LR	DT	BPN	SVM
Z-values	-0.575	0.630	0.157	0.000
P-values	0.2826	0.2644	0.4378	0.5000

Second, we were able to check the differences based on algorithm when applying the summary. In the case of LR, which is a statistics-based algorithm, predictive performance increased by 4% with the summary-based model. Conversely, in the case of DT (the tree-based algorithm), the predictive performance of the abstracted model fell by 4% based on the 5-fold average. In this experiment, we found that the two algorithms were adversely affected. On the other hand, in BPN and SVM, no noticeable difference was found when applying summaries. In BPN, when applying summaries, the probability of prediction was reduced by 1% on average, but only marginally. In the case of SVM, the difference in performance could not be determined due to the same prediction probability. Therefore, DT and LR were found to have a small yet opposite effect when summarized, while no direction of impact was found in SVM and BPN.

6. Conclusion

This study empirically analyzed the effect of document summarization techniques on the detection of fake news. We were able to uncover a difference in the effects and direction of the machine-learning algorithms, as well as the possibility of improving performance by applying extractive summarization to fake news detection. Although the news article-based fake news detection model using the document summarization technique was not found to have a statistically significant difference from the full-text-based fake news detection model, this study offers a number of important implications.

First, the study's goal was to apply extractive summarization to fake news detection research and analyze the resulting effects. Currently, the auto-summarizing of news is a common occurrence in daily life, and there is some concern about the infringement of editorial rights and/or information distortion through this auto-summarizing process. This issue has not been studied empirically until now. The present study is the first of its kind to analyze the impact of using summarized news produced by extraction from the point of view of fact-checking.

Second, this study expands the depth of fake news detection research by analyzing full-text domestic news data not previously studied in any other research. Thus far, most conventional domestic fake news detection studies have focused on single-sentenced news, which is different from news we encounter in real life. Differences in news type also exist. The types of fake news investigated in conventional studies include rumors, political celebrity remarks, and fake news from the press—regardless of the different characteristics of each type. According to a survey conducted by the Korea Press Foundation, "fake news from the press caused by a lack of fact-checking" was the most problematic type among the various types of fake news (Yang, 2019). This study focused its investigation on only one type: fake news from the press. Thus, we have expanded fake news detection research in terms of form and type.

Third, we implemented a relatively predictive text-based detection model in a controlled environment. The main problem of the existing text-based fake news detection model has been the issue of overfitting by subjects. In order to overcome this problem, we not only unified the form and type of the entire articles used as data, but we also unified subjects into politics and society. In addition, the genuine news dataset was made up of articles that shared the same keywords as the fake news dataset. The predictive model implemented in this study exhibited a minimum 57% and a maximum 74% predictive performance based on the 5-fold average. Compared to the public's fake news recognition rate of 58.5%, as reported by Sung-Soo Kim's laboratory, we may conclude that this study's model has the predictive power to detect fake news (Kim, 2018).

Nonetheless, this study has some limitations. First, the study's data size was small, so the statistical significance of the experimental results was not obtained. Second, this study does not provide comparisons with other document summarization techniques. Applying different extractive or abstractive summarization methods might produce different results. Third, in the PCA process, the number of fixed variable extractions was set to 20. However, it might be possible to further refine the model by deriving an optimal number of input variables using PCA. Fourth, it might also be possible to improve the model's performance through fine word embedding techniques, which were not used in this study. Therefore, further research should be conducted to secure more data, employ a broader range of summary and word embedding techniques, and finely adjust the number of input variables.

References

Afroz, S., M. Brennan, and R. Greenstadt, "Detecting Hoaxes, Frauds, and Deception in Writing Style Online," *2012 IEEE Symposium on Security and Privacy*, (2012), 461~475.

Aker, A., L. Derczynski, and K. Bontcheva, "Simple Open Stance Classification for Rumour Analysis," *Proceedings of Recent Advances in Natural Language Processing*, (2017), 31~39.

Allahyari, M., S. Pouriyeh, M. Assefi, S. Safaei, E. D. Trippe, J. B. Gutierrez, and K. Kochut, "Text Summarization Techniques: A Brief Survey," *International Journal of Advanced Computer Science and Applications*, Vol.8, No.10, (2017), 397~405.

Bondiellia, A., and F., Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, Vol.497, (2019), 38~55.

Castillo, C., M. Mendoza, and B. Poblete, "Information Credibility on Twitter," *Proceedings of the 20th International Conference on World Wide Web*, (2011), 675~684.

Diab, S., "Optimizing Stochastic Gradient Descent in Text Classification Based on Fine-Tuning Hyper-Parameters Approach," International

Journal of Computer Science and Information Security, Vol.16, No.12, (2018), 155~160.

Erkan, G., and D. R. Radev, "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization," *Journal of Artificial Intelligence Research*, Vol.22, (2004), 457~479.

Esmaeilzadeh, S., G. X. Peh, and A. Xu, "Neural Abstractive Text Summarization and Fake News Detection," arXiv preprint arXiv:1904. 00788, (2019).

Ferreira, W., and A. Vlachos, "Emergent: a novel data-set for stance classification," *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* (2016), 1163~1168.

Gambhir, M., and V. Gupta, "Recent automatic text summarization techniques: a survey," *Artificial Intelligence Review,* Vol.47, No.1, (2017), 1~66.

Giasemidis, G., C. Singleton, I. Agrafiotis, J. R. C. Nurse, A. Pilgrim, C. Willis, and D. V. Greetham, "Determining the Veracity of Rumours on Twitter," *Social Informatics Part Ⅰ, (*2016), 185~205.

Hardalov, M., I. Koychev, and P. Nakov, "In Search of Credible News," *Artificial Intelligence: Methodology, Systems, and Applications,* (2016), 172~180.

Hyun, Y., and N. Kim, "Text Mining-based Fake News Detection Using News And Social Media Data," *The Journal of Society for e-Business Studies*, Vol.23, No.4, (2018), 19~39.

Jeon, B., and H. Ahn, "A Collaborative Filtering System Combined with Users Review Mining : Application to the Recommendation of Smartphone Apps," *Journal of Intelligence and Information System*, Vol.21, No.2, (2015), 1~18.

Jin, H., "Compressed three news articles ... NAVER AI summary bot appeared," Digital Times, 2017, Available at http://www.dt.co.kr/contents.html?article_no=2017112902101131043001 (Accessed 30 August 2019).

Kim, K.-B., "A Passport Recognition and Face Verification Using Enhanced Fuzzy ART Based RBF Network and PCA Algorithm," *Journal of Intelligence and Information System*, Vol.12, No.1, (2006), 17~31.

Kim, S. S., and Green Consumer Network in Korea, Consumer Awareness Survey on Mobile Video Service, Kim Sung Soo Representative Office, 2018. Available at http://theminjoo.kr/inspectionDetail.do?nt_id=16&bd_seq=126493 (Accessed 30 August 2019).

Koo, B.-K., "News Summary Bots and Human Power in the Infinite Information Age," The Hankyoreh, 2017, Available at http://www.hani.co.kr/arti/economy/it/822844.html (Accessed 30 August 2019).

Kwon, S., M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent Features of Rumor Propagation in Online Social Media," *IEEE 13th International Conference on Data Mining,* (2013), 1103~1108.

Lafferty, J., A. McCallum, and F. C. N. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," *Proceedings of the 18th International Conference on Machine Learning 2001*, (2001), 282~289.

Ma, J., W. Gao, P. Mitra, S. Kwon, B. J. Jansen,

K. F. Wong, and M. Cha, “Detecting Rumors from Microblogs with Recurrent Neural Networks,” *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence,* (2016), 3818~3824.

Park, S. S., and K. C., Lee, “A Comparative Study of Text analysis and Network embedding Methods for Effective Fake News Detection,” *Journal of Digital Convergence,* Vol.17, No.5(2019), 137~143.

Radev, D. R., E. Hovy, and K. McKeown, “Introduction to the Special Issue on Summarization,” *Journal of Computational linguistics*, Vol.28, (2002), 399~408.

Seol, J., and S. Lee, " lexrankr: LexRank based Korean multi-document summarization," *Journal of the Korean Institute of Information Scientist and Engineers 2016 Winter Academic Conference*, (2016), 458~460.

Shin, H., “Naver ‘Summary Bot’, is it an evolution or editorial infringement?,” Sisain, 2017, Available at https://www.sisain.co.kr/news/articleView.html?idxno=30828 (Accessed 30 August 2019).

Lee S., and H.-J. Kim, “Keyword Extraction from News Corpus using Modified TF-IDF,” *The Journal of Society for e-Business Studies,* Vol.14, No.4(2009), 59~73.

Wang, W. Y., ““Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection,” *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Vol.2, (2017), 422~426.

Yang, J, “'News' and 'Fake News' that ordinary citizens think,” *Media issue*, Vol.5, No.1, (2019), 1~15.

Yoon, S. O., “A Study on the Characteristics and Problems of Fake News Regulations proposed by the National Assembly,” *Journal of Media Law, Ethics and Policy Research*, Vol.18, No.1, (2019), 103~138.

Yun, T. U., and H. Ahn, “Fake News Detection for Korean News Using Text Mining and Machine Learning Techniques,” *Journal of Information Technology Applications & Management,* Vol.25, No.1, (2018), 19~32.

Yun, Y., Ko, E., and Kim, N., “Subject-Balanced Intelligent Text Summarization Scheme,” *Journal of Intelligence and Information System*, Vol.25, No.2, (2019), 141~166.

Zhang, H., Z. Fan, J. Zheng, and Q. Liu, “An Improving Deception Detection Method in Computer-Mediated Communication,” *Journal of Networks*, Vol.7, No.11, (2012), 1811~1816.

Zhao, Z., P. Resnick, and Q. Mei, “Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts,” *Proceedings of the 24th International Conference on World Wide Web*, (2015), 1395~1405.

Zubiaga, A., M. Liakata, and R. Procter, “Learning Reporting Dynamics during Breaking News for Rumour Detection in Social Media,” arXiv preprint arXiv:1610.07363, (2016).

국문요약

문서 요약 기법이 가짜 뉴스 탐지 모형에 미치는 영향에 관한 연구

심재승* · 원하람** · 안현철***

가짜뉴스가 전세계적 이슈로 부상한 최근 수년간 가짜뉴스 문제 해결을 위한 논의와 연구가 지속되고 있다. 특히 인공지능과 텍스트 분석을 이용한 자동화 가짜 뉴스 탐지에 대한 연구가 주목을 받고 있는데, 대부분 문서 분류 기법을 이용한 연구들이 주를 이루고 있는 가운데 문서 요약 기법은 지금까지 거의 활용되지 않았다. 그러나 최근 가짜뉴스 탐지 연구에 생성 요약 기법을 적용하여 성능 개선을 이끌어낸 사례가 해외에서 보고된 바 있으며, 추출 요약 기법 기반의 뉴스 자동 요약 서비스가 대중화된 현재, 요약된 뉴스 정보가 국내 가짜뉴스 탐지 모형의 성능 제고에 긍정적인 영향을 미치는지 확인해 볼 필요가 있다. 이에 본 연구에서는 국내 가짜뉴스에 요약 기법을 적용했을 때 정보 손실이 일어나는지, 혹은 정보가 그대로 보전되거나 혹은 잡음 제거를 통한 정보 획득 효과가 발생하는지 알아보기 위해 국내 뉴스 데이터에 추출 요약 기법을 적용하여 '본문 기반 가짜뉴스 탐지 모형'과 '요약문 기반 가짜뉴스 탐지 모형'을 구축하고, 다수의 기계학습 알고리즘을 적용하여 두 모형의 성능을 비교하는 실험을 수행하였다. 그 결과 BPN(Back Propagation Neural Network)과 SVM(Support Vector Machine)의 경우 큰 성능 차이가 발생하지 않았지만 DT(Decision Tree)의 경우 본문 기반 모델이, LR(Logistic Regression)의 경우 요약문 기반 모델이 다소 우세한 성능을 보였음을 확인하였다. 결과를 검증하는 과정에서 통계적으로 유의미한 수준으로는 요약문 기반 모델과 본문 기반 모델간의 차이가 확인되지는 않았지만, 요약을 적용하였을 경우 가짜뉴스 판별에 도움이 되는 핵심 정보는 최소한 보전되며 LR의 경우 성능 향상의 가능성이 있음을 확인하였다. 본 연구는 추출 요약 기법을 국내 가짜뉴스 탐지 연구에 처음으로 적용해 본 도전적인 연구라는 점에서 의의가 있다. 하지만 한계점으로는 비교적 적은 데이터로 실험이 수행되었다는 점과 한 가지 문서요약기법만 사용되었다는 점을 제시할 수 있다. 향후 대규모의 데이터에서도 같은 맥락의 실험결과가 도출되는

* 국민대학교 비즈니스IT전문대학원
** 국민대학교 비즈니스IT전문대학원
*** Corresponding Author: Hyunchul Ahn
Graduate School of Business IT, Kookmin University
77 Jeongneung-ro, Seongbuk-gu, Seoul, 02707, Republic of Korea
Tel: +82-2-910-4577, Fax: +82-2-910-4017, E-mail: hcahn@kookmin.ac.kr

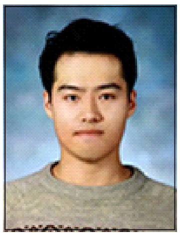
지 검증하고, 보다 다양한 문서요약기법을 적용해 봄으로써 요약 기법 간 차이를 규명하는 확장된 연구가 추후 수행되어야 할 것이다.

주제어 : 가짜뉴스, 문서요약, 자동화 팩트체킹, 기계학습, 국내뉴스

논문접수일 : 2019년 8월 8일 논문수정일 : 2019년 9월 8일 게재확정일 : 2019년 9월 20일
원고유형 : 단편논문 교신저자 : 안현철

저 자 소 개

Jae-Seung Shim
is currently a master’s program at Graduate School of Business IT, Kookmin University, where he earned his bachelor degree in Management Information. His primary research interests include data mining and machine learning for the social sciences and business.



Ha-Ram Won
He holds a bachelor's degree in business administration from Halla University, Korea, and is currently a master's degree in Business Analysis Track at Business IT Graduate School, Kookmin University in Korea. His interests include business analytics and CRM and data-driven marketing analytics.

Hyunchul Ahn
is a professor of management information systems at the Graduate School of Business IT, Kookmin University, Republic of Korea. He obtained his Ph.D. from Korea Advanced Institute of Science and Technology. His major research interests are intelligent IS and IS adoption. His research has been published in *Annals of OR*, *Applied Intelligence*, *Computers & OR*, *Expert Systems with Applications*, *International Journal of Electronic Commerce*, *International Journal of Production Research*, *Information & Management*, etc.