
A Study of Profanity Effect in Sentiment Analysis on Natural Language Processing Using ANN

Cheong-Ghil Kim¹, Young-Jun Hwang¹ and Chayapol Kamyod^{2,*}

¹*Dept. of Computer Science, Namseoul University, Korea*

²*Computer and Communication Engineering for Capacity Building Research Center, School of Information Technology, Mae Fah Luang University, Chiang Rai 57100, Thailand*

E-mail: cgkim@nsu.ac.kr; hyj765@naver.com; chayapol.kam@mfu.ac.th

**Corresponding Author*

Received 13 December 2021; Accepted 07 January 2022;
Publication 08 March 2022

Abstract

The development of wireless communication technology and mobile devices has brought about the advent of an era of sharing text data that overflows on social media and the web. In particular, social media has become a major source of storing people's sentiments in the form of opinions and views on specific issues in the form of unstructured information. Therefore, the importance of emotion analysis is increasing, especially with machine learning for both personal life and companies' management environments. At this time, data reliability is an essential component for data classification. The accuracy of sentiment classification can be heavily determined according to the reliability of data, in which case noise data may also influence this classification. Although there is stopword that does not have meaning in such noise data, data that does not fit the purpose of analysis can also be referred to as noise data. This paper aims to provide an analysis of the impact of profanity data on deep learning-based sentiment classification. For this purpose, we

Journal of Web Engineering, Vol. 21_3, 751–766.

doi: 10.13052/jwe1540-9589.2139

© 2022 River Publishers

used movie review data on the Web and simulated the changes in performance before and after the removal of the profanity data. The accuracy of the model trained with the data and the model trained with the data before removal were compared to determine whether the profanity is noise data that lowers the accuracy in sentiment analysis. The simulation results show that the accuracy dropped by about 2% when judging profanity as noise data in the sentiment classification for review data.

Keywords: Deep learning, sentiment analysis, opinion mining, natural language processing, stopword.

1 Introduction

The entry into a hyper-connected society in accordance with the start of the recent development of high-speed mobile communication technology is bringing about changes in our Internet environment. Now, web users are actively using services to use the web as a communication space for expressing their opinions, sharing information, and interacting with each other, rather than as a simple search tool in the early days of the Internet. Therefore, the popularity of social media is increasing greatly [1]. Blogs, reviews, tweets, posts, and discussions on social media are scanned to extract the opinions of people. Attitudes, views, feelings, and opinions constitute an essential part in analyzing the behavior of a person, which can be referred to as sentiments. Sentiment analysis, also known as opinion mining, deals with inspecting these sentiments directed towards any entity. The proportion of people's communication through blogs or SNS is continuously increasing. As the communication ratio of such SNS increases, posts from users are being uploaded in large quantities to those social media containing people's emotions and opinions. Therefore, SNS could be undoubtedly utilized as a place to gather users' opinions for corporate marketing strategies [2–5].

Due to these changes, research on not only NLP (Natural Language Processing) [6, 7] but also extracting emotions [8, 9] from texts that analyzes text-type data is being actively conducted. However, such sentiment analysis cannot be utilized only with data. Data preprocessing [10] is necessary for the purpose of use, and the accuracy varies greatly depending on the preprocessing method. In particular, when building a Deep Learning model, data preprocessing is a process that must be completed and is the most time-consuming part. Since the results of deep learning analysis models are also significantly different depending on how the data is preprocessed,

well-refined data must be input to obtain accurate data analysis results. In addition, in sentiment analysis, it is necessary to extract “noise data,” which is data that does not fit the purpose or has no meaning. “Noise data” refers to corrupted, distorted, or duplicated data. It is an important task to remove them because such noise data interferes with the analysis of the data [7]. However, there are words that are difficult to process as noise data, such as profanity. Articles created on SNS contain many profane words. Although such profanity is used as an emphasis, it is difficult to determine whether the profanity has a negative or positive sensibility with only a single sentence.

This paper intends to use a deep learning model to check how profanity with these characteristics affects sentiment analysis using machine learning in written text. The structure of this study is as follows: Chapter 2 introduces the basic understanding of NLP, sentiment analysis, and deep learning. The research method proposed in this paper is described in Chapter 3, the research results from the proposed study are shown in Chapter 4, and the conclusion and future research are explained in Chapter 5.

2 Background

This section is subdivided into several sections which provide background introductions for the implementation and simulation of the proposed method including NLP, semantic analysis, and Deep Learning.

2.1 Natural Language Processing

NLP is a task that enables computers to understand human language through machine learning for smooth interaction between natural language and computers [6, 7]. Furthermore, Deep Learning architectures and algorithms have made significant contributions to recent advances in NLP, image, and speech processing [7]. In general, NLP can be viewed as a process of deriving results through each step of morphological analysis, syntax analysis, semantic analysis, and pragmatic analysis, as shown in Figure 1 [11].

The first stage is morphological analysis that is a step of understanding distinct words by separating the morpheme-based minimum unit of a sentence. Therefore, each word of morphological analysis may contain various parts-of-grammatical elements such as prefixes, suffixes, infixes, or separate roots. The second stage is syntax analysis that is the process of analysing sentence structure by determining sentence components such as subject, verb, and object. For this purpose, a parser is used. After that, it builds a syntax or

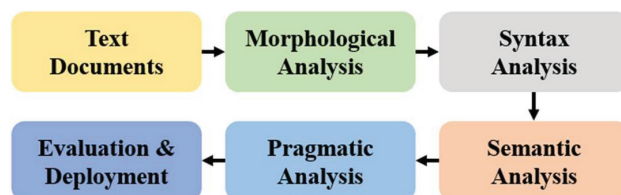


Figure 1 Natural language processing structure.

a parse tree. This tree allows us to determine whether a sentence is structured correctly and grammatically valid and correct. The next step is semantic analysis, which is the process of determining the actual meaning of a morpheme. If a word has multiple meanings and is ambivalent, it is necessary to reveal exactly what meaning it has and to give it a meaning. An example of a task that gives meaning is named entity recognition. Entity name analysis refers to the task of analysing meaning by tagging it with predefined categories such as names, geographical names, organization names, and numbers. Eventually, semantic analysis enables us to understand the speaker's intention. The last step is pragmatic analysis. It may analyse the semantic relationship between sentences through context, in which we can find out what the anaphora indicated by the pronoun is and what the speaker wants. For effective NLP, resources such as meaning and grammatical properties of words and grammar rules, and various knowledge expression techniques such as lexicons and ontology are used [12].

So far, NLP technology has been developed for two purposes: to automate language processing by machines and to improve communication with humans [9]. With the development of language processing automation, it can be said that automatic analysis of literary works has become possible with the development of text digitization, artificial intelligence (AI), and NLP. Through this development, research is being conducted in various fields such as information retrieval, text autocomplete, translation, and sentiment analysis. Chatbots are a representative field for the purpose of improving communication with humans. Chatbot is a compound word of chat and bot, which refers to a robot that communicates with humans based on artificial intelligence and refers to an artificial intelligence program that operates in a request-and-response structure [13].

Recently, research on NLP technology is actively taking place in learning methods through deep learning. It showed excellent performance for natural language processing like the existing RNN (Recurrent Neural Network) [14], LSTM (Long Short-Term Memory) [15], and GRU (Gated

Recurrent Unit) [16], and but there was still a problem that the accuracy decreased as the length increased. The recent trend is to use the transformer model that uses only the encoder and the decoder and combines the attention mechanism. The Transformer is a model presented in [17] and has shown better performance than conventional RNNs. And by applying this model, such as the BERT (Bidirectional Encoder Representations from Transformers) [18] and GPT (Generative Pre-trained Transformer) [19] series are developing, since most of the information on the Internet is in the form of unstructured text data, the application of natural NLP technology is limitless.

2.2 Sentiment Analysis

Sentiment analysis, also called opinion mining, is the area of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. Hence, research in sentiment analysis not only has an important impact on NLP but also has a profound impact on management sciences, political science, economics, and social sciences as they are all affected by people's opinions [20].

In general, sentiment analysis consists of collecting various review data from the web or bulletin board, classifying only text elements to be used for sentiment analysis, and analyzing opinions on the given data. The method of sentiment analysis can be divided into three methodologies: dictionary-based analysis, machine learning-based analysis, and a mixed method using both methods. In the dictionary-based analysis method, if a word that matches the sentiment dictionary exists in the data, sentiment scores are given to classify emotions. Data collection, data preprocessing, sentiment dictionary construction, and finally the sentiment analysis step.

There are two methods for constructing the sentiment dictionary: a method of calculating and analyzing the frequency of positive and negative in an emotional word, and a method of calculating the emotional score by summing up the emotional score. Sentiment analysis based on machine learning is performed through various techniques of supervised, unsupervised, and semi-supervised learning. The machine learning-based method consists of collection, preprocessing, learning, verification, and emotion analysis as shown in Figure 2 [18, 21]. The machine learning-based method also goes through the same collection and preprocessing steps as the dictionary-based method. When the learning is finished, the test data is used to verify the accuracy of the trained model, and when the verification step is completed, sentiment analysis is performed using the trained model.



Figure 2 Sentiment analysis structure.

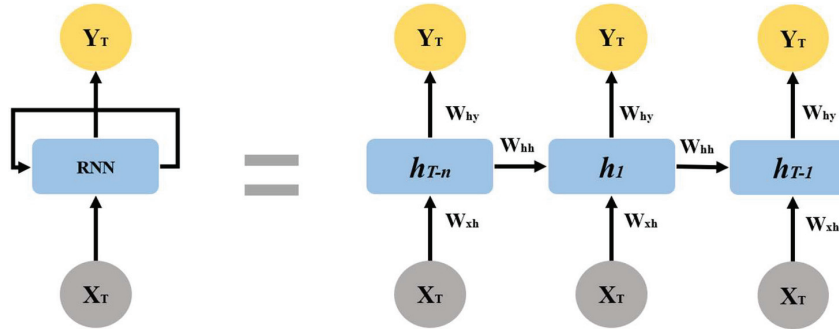


Figure 3 RNN structure.

2.3 Deep Learning

Deep learning, one of the branches of machine learning, generally refers to an algorithm that deepens the layer of a neural network. Traditional machine learning techniques require the separate extraction of feature sets for classification, but deep learning can automatically extract complex features as well as automate the process of feature set extraction [22]. Among deep learning techniques, RNNs are a representative one used for NLP with strength in sequence modelling.

Figure 3 [23] depicts the structure of an RNN with its own recurrent weight (W), which reflects the time series characteristics of predicting future information based on present information through past information through W while recognizing the pattern [24]. For each time step, each weight (W_{hy} , W_{xh} , W_{hh}) is calculated, and the current hidden value is obtained by adding the hidden value of the previous time step and the input value of the current value. The gradient value of the current output part is added to the previous time steps value. If there are many neurons or input values, the gradient value becomes a non-deterministically large value by repeatedly multiplying the weight value, or gradient exploding or 0 value: Gradient vanishing problems that converge to.

To solve this slope problem, an LSTM model was proposed in [15]. LSTM is a recurrent neural network model designed to solve the long-term

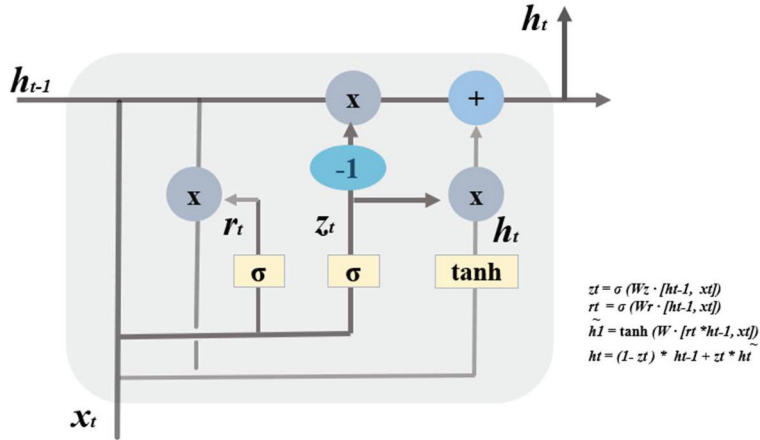


Figure 4 LSTM structure.

dependence problem in which the gradient of the recurrent neural network disappears [25]. It is composed of Figure 4 [23]. structure. It solves the problem of gradient loss by forgetting the cell's internal state value and adjusting whether it accepts a new input value.

3 Proposed Method

This section introduces the proposed method of polarity classification model to analyze how profanity affects sentiment analysis using deep learning. The accuracy of the polarity classification is evaluated using the LSTM model. The first stage of the proposed model is the data collection for training. The next stage is data preprocessing, in which the preprocessing process unfolds in two directions, as shown in Figure 5. One is a profanity removal model, and the other is a profanity-included model in the model training stage. Finally, when the two models are completed, the accuracy is obtained through the same test set. After that, the accuracy of the two models is compared to confirm the influence of profanity.

3.1 Data Collection

In this study, data collection is necessary to build both a general and a profanity removed model through supervised learning. Data collection can be done in two ways. In some cases, data is collected directly by using open data or by crawling. In this study, the open data 'NaverSentimentMoviecorpusv1.0' [26]

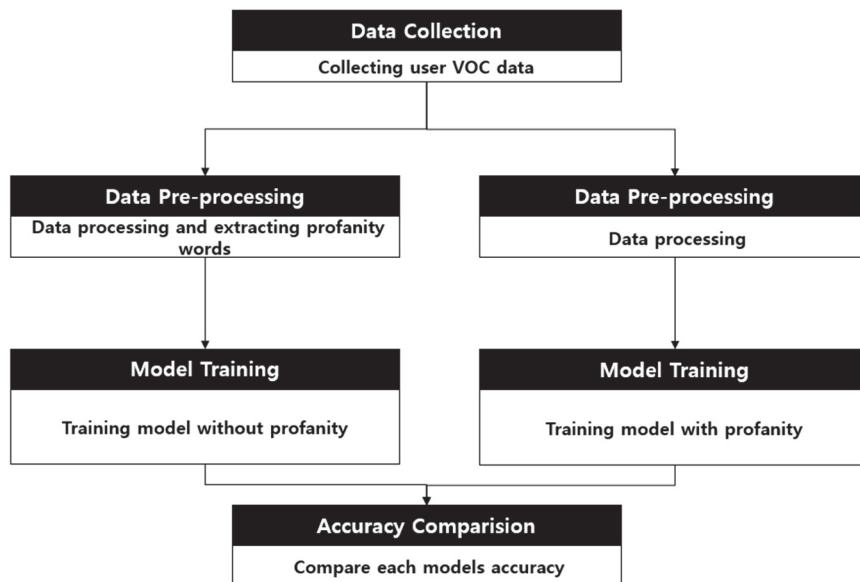


Figure 5 Two-way parallel comparison model.

Table 1 Sentiment data structure

Id	Review	Label
382154	흠 포스터보고 초딩...	0
10248551	교도소이야기는 코믹...	1
154921	이영화는 액션이...	1
648525	재미 액션 둘다 없는...	0
...

was used for the purpose of confirming the effect of profanity in the review data. The open data consists of movie reviews collected from the Naver Movies section operated by Naver, one of the Korean portal sites. The data is labeled as positive or negative, and 1500 reviews were used as the training set and 50,000 reviews as the test set. An example of the data set is tabulated in Table 1.

3.2 Data Pre-processing

Even data that has been labeled must go through a preprocessing process in order to be used in deep learning. Since data with different preprocessing

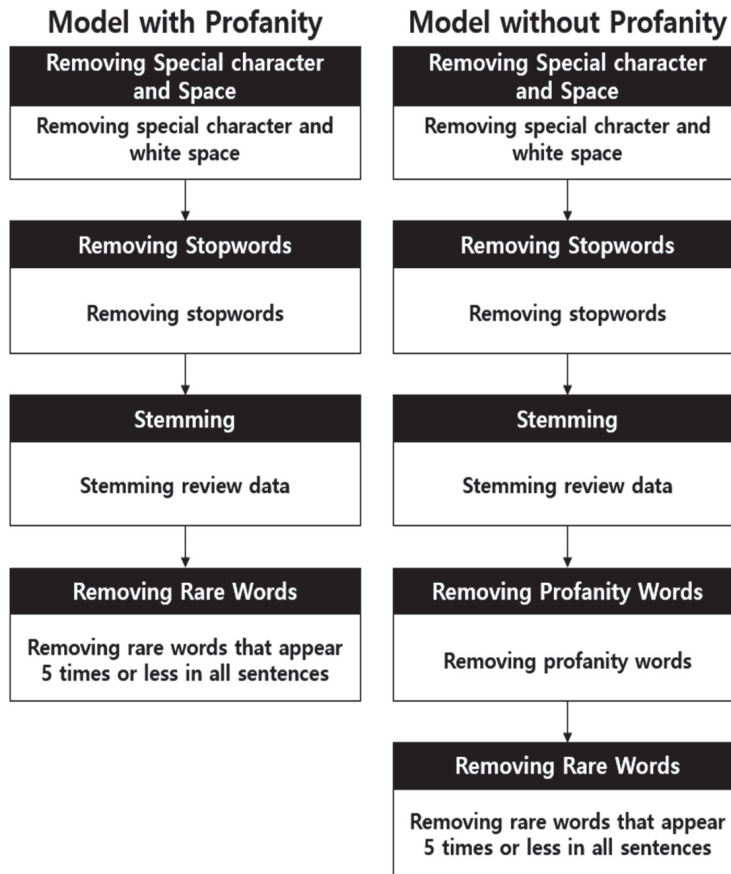


Figure 6 Data pre-processing stages.

needs to be put into the two models, the preprocessing process was performed in the following order, as shown in Figure 6.

First, special characters were removed using regular expressions in order to leave only Korean data before tokenization in the review data in common in both data sets. In the primary preprocessed data, whitespace was additionally removed by regular expression. This is to ensure that the same result is always obtained when morpheme analysis is performed through the morpheme analyzer. After the above-mentioned preprocessing work was completed, tokenization was carried out. Mecab, a Korean morpheme analyzer, was used for the library. After morphological analysis was completed, stopwords that had no meaning were removed.

3.3 Model Training

The data that had been preprocessed was embedded through one-hot-encoding. Here, the single-layer LSTM model was used because the purpose of this study was to check the accuracy according to the data difference, not the accuracy difference according to the model. In order to prevent overfitting from occurring during model training, a dropout layer was added. Dropout ratio was set to 0.5% known as the most optimized ratio [27]. Finally, as an activation function, a sigmoid function is used to output one output value through a fully connected layer called a dense layer.

3.4 Comparison

Finally, after learning the model, an exact comparison target is needed to know the influence of profanity. We used the proposed Two-way Parralel Processing Compare Model to compare the accuracy of models that have undergone different preprocessing on the same review data on the same model, so that we can know the direct influence of the presence or absence of the target data.

4 Experimental Result

This study was conducted in Colab [28], a machine learning environment provided by Google, to investigate the effect of profanity on sentiment analysis of Korean data. The specifications of the Colab environment are shown in Table 2.

Supervised learning was conducted using open data ‘NaverMovieCorpus v0.1’. For evaluation, accuracy was evaluated with 50,000 test sets that had not previously been used for model training. The corresponding data is review data with the same length as shown in Figure 6.

Prior to learning, special characters and stopwords were removed in both models; in addition, rare words with fewer than five occurrences were excluded. For the shape of the AI model, the same single-layer LSTM was used for both models, and the epoch was set to 3, and the batch size

Table 2 Colab specifications

OS	Linux Ubuntu 18.04
Program Language	Python 3.6.9
CPU	Intel(R) Xeon(R) CPU @ 2.30GHz

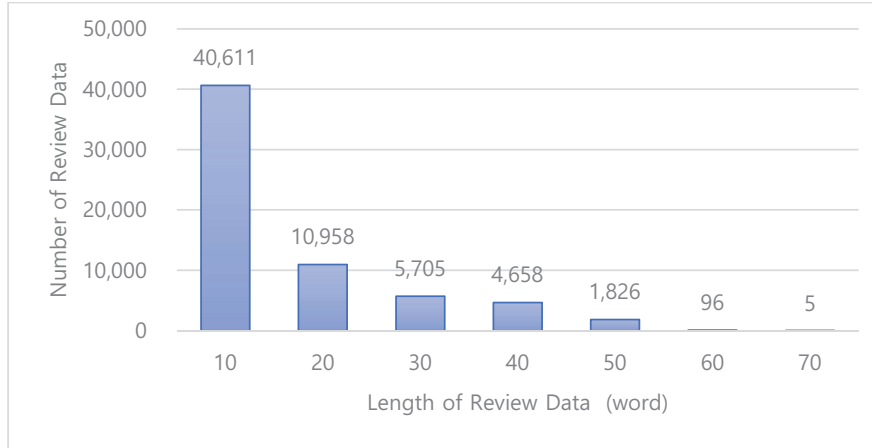


Figure 7 Review data length chart.

Table 3 Comparison results

Model	Accuracy	Data
Profanity Include Model	83.4%	NaverSentimentcorpus v1.0
Profanity Removal Model	81.6%	NaverSentimentcorpus v1.0

was set to 60. After that, the optimizer proceeded with learning with Adam. As a verification method, the accuracy of the model trained on data with profanity removed in the data preprocessing process and the model learned without removing profanity were compared, and the results were impressive.

In the case of performing sentiment analysis after removing profanity, the LSTM model was measured with an accuracy of 81%, and the model without removing the profanity showed an accuracy of 83%. It was confirmed that the model in which the profanity was removed had a decrease in accuracy of about 2% compared to the model that did not remove the profanity.

5 Conclusion

In this paper, a study was conducted to determine whether abusive language is 'noise data' in the sentiment analysis of Korean data. The accuracy of the model was measured using Keras, a high-level deep learning API of TensorFlow. As a result of performing emotion analysis after removing the profanity, the LSTM model showed an accuracy of about 81%, and it was confirmed that the accuracy decreased by about 2% compared to the

83% accuracy of the model trained without removing the profanity. Most of the previous studies dealing with profanity data focused only on filtering profanity data from the chatbot and removing the data. In other words, those studies considered the profanity data only as noise data and concentrated on efficiently removing it. This study confirmed that in the sentiment analysis of user review data, profanity can be used to improve accuracy rather than being considered as noise data that can degrade accuracy. In AI learning, the accuracy of AI varies greatly depending on how the data can be refined and noise data can be reduced. Therefore, it is expected that research to determine the noise data that lowers the accuracy of AI in various fields will continue in the future.

Acknowledgment

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2021R1I1A4A01049755)

References

- [1] Ashima Yadav and Dinesh Kumar Vishwakarma. Sentiment analysis using deep learning architectures: a review, *Artificial Intelligence Review* 53:4335–4385, 2020.
- [2] K. Xu, G. Qi, J. Huang, T. Wu, and X. Fu. Detecting Bursts in Sentiment-Aware Topics from Social Media, *Knowledge-Based Systems*, Vol. 141, pp. 44–54 DOI: 10.1016/j.knosys.2017.11.007, February 2018.
- [3] Umar Ishfaq and Khalid Iqbal. Identifying the Influential bloggers: A modular approach based on Sentiment Analysis, *Journal of Web Engineering*, 16(5 & 6): pp. 505–523, 2017.
- [4] Ruirong Xue, Subin Huang, Xiangfeng Luo, Dandan Jiang, and Yan Peng. Semantic Emotion-Topic Model in Social Media Environment, *Journal of Web Engineering*, 17(1 & 2): pp. 073–092, 2018.
- [5] L. Zhang and B. Liu. Sentiment Analysis and Opinion Mining, In: Sammut C., Webb G.I. (eds) *Encyclopaedia of Machine Learning and Data Mining*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7687-1_907, 2017.
- [6] Julia Hirschberg and Christopher D. Manning. Advances in natural language processing, *Science*, 349(6245), pp. 261–266, DOI: 10.1126/science.aaa8685, 17 July 2015.

- [7] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. Recent Trends in Deep Learning Based Natural Language Processing, *IEEE Computational Intelligence Magazine* 13(3): pp. 55–75 DOI: 10.1109/MCI.2018.2840738, August 2018.
- [8] Geetika Gautam and Divakar Yadav. Sentiment Analysis of Twitter Data Using Machine Learning Approaches and Semantic Analysis, Proc. of 2014 Seventh International Conference on Contemporary Computing (IC3), pp. 7–9 August 2014.
- [9] Erik Cambria, Björn Schuller, Yunqing Xia, and Catherine Havasi. New Avenues in Opinion Mining and Sentiment Analysis, *IEEE Intelligent Systems*, 28(2): 15–21, 2013.
- [10] P. Chandrasekar and K. Qian. The Impact of Data Preprocessing on the Performance of a Naive Bayes Classifier, 2016 *IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, pp. 618–619, doi: 10.1109/COMPSAC.2016.205, 2016.
- [11] P. M. Nadkarni, L. O. Machado, and W. W. Chapman. Natural language processing: an introduction, *Journal of the American Medical Informatics Association*, Vol. 18, pp. 544–551, 2011.
- [12] Sumit Chopra, Michael Auli, and Alexander M. Rush. Abstractive Sentence Summarization with Attentive Recurrent Neural Networks, 2016 *Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, 2016.
- [13] Daniel W. Otter, Julian R. Medina, and Jugal K. Kalita. A Survey of the Usages of Deep Learning for Natural Language Processing, *IEEE Transactions on Ransactions on Neural Networks and Learning Systems*, 32(2) February 2021.
- [14] D. Rumelhart, G. Hinton, and R. Williams, Learning internal representations by error propagation, UCSD, La Jolla, CA, USA, Tech. Rep. ICS-8506, 1985.
- [15] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Comput.*, 9(8): pp. 1735–1780, 1997.
- [16] K. Cho, B. V. Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation, arXiv:1406.1078. [Online]. Available: <http://arxiv.org/abs/1406.1078>, 2014.
- [17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need, in Proc. NIPS, pp. 6000–6010, 2017.

- [18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding, arXiv:1810.04805. [Online]. Available: <http://arxiv.org/abs/1810.04805>, 2018.
- [19] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever. Improving Language Understanding by Generative PreTraining, [Online]. Available: https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf, 2018.
- [20] Bing Liu, Sentiment Analysis and Opinion Mining, Synthesis Lectures on Human Language Technologies, 5(1): pp. 1–167, May 2012.
- [21] D. P. Hong, H. Jeong, S. Park, E. Han, H. Kim, and I. Yun. Study on the Methodology for Extracting Information from SNS Using a Sentiment Analysis, J. Korea Inst. 16(6): pp. 141–155 December 2017.
- [22] H. Y. Park, K. J. Kim, Sentiment Analysis of Movie Review Using Integrated CNN-LSTM Mode, *Journal of Intelligence and Information Systems*, Vol. 25, pp. 141–154, 2019.
- [23] W. Zaremba, I. Sutskever, and O. Vinyals. Recurrent Neural Network Regularization, *International Conference on Learning Representations*, 2015.
- [24] E. J. Lee Basic and applied research of CNN and RNN, *Broadcasting and Media Magazine*, 22(1): pp. 87–95, 2017.
- [25] Y. Bengio, P. Simard, and P. Frasconi. Long-term dependencies with gradient descent is difficult *IEEE Transaction on Neural Networks*, 5(2) March 1994.
- [26] NaverSentimentcorpus <https://github.com/e9t/nsmc>. access 2021-09-21.
- [27] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov Dropout: A Simple Way to Prevent Neural Networks From Overfitting, *Journal of Machine Learning Research*, 1929–1958, 2014.
- [28] Colab <https://colab.research.google.com/>

Biographies



Cheong-Ghil Kim received the B.S. in Computer Science from University of Redlands, CA, U.S.A. in 1987. He received the M.S. and Ph.D. degree in Computer Science from Yonsei University, Korea, in 2003 and 2006, respectively. Currently, he is a professor at the Department of Computer Science, Namseoul University, Korea. His research areas include Multimedia Embedded Systems, Mobile AR, and 3D Contents. He is a member of IEEE.



Young-Jun Hwang is an undergraduate student majoring with Computer Science at the Namseoul University in Korea. He completed the 9th Best of the Best course organized by KITRI. His research interests are cyber security and Machine Learning.



Chayapol Kamyod received his Ph.D. in Wireless Communication from the Center of TeleInFrastruktur (CTIF) at Aalborg University (AAU), Denmark. He received M. Eng. in Electrical Engineering from The City College of New York, New York, USA. In addition, he received B.Eng. in Telecommunication Engineering and M. Sci. in Laser Technology and Photonics from Suranaree University of Technology, Nakhon Ratchasima, Thailand. He is currently a lecturer in Computer Engineering program at School of Information Technology, Mae Fah Luang University, Chiang Rai, Thailand. His research interests are resilience and reliability of computer network and system, wireless sensor networks, embedded technology, and IoT applications.