

FAGON: Fake News Detection Model Using Grammatical Transformation on Deep Neural Network

Youngkyung Seo¹, Seong-Soo Han², You-Boo Jeon³, and Chang-Sung Jeong^{4*}

¹ Department of Electrical Engineering, Korea University
Seoul, Republic of Korea
[e-mail: ygseo@korea.ac.kr]

² Visual Information Processing, Korea University
Seoul, Republic of Korea
[e-mail: postsky0@korea.ac.kr]

³ Department of Computer Software Engineering, Soonchunhyang University, Asan-si, Chungcheongnam-do,
Republic of Korea
[e-mail: jeonyb@sch.ac.kr]

⁴ Department of Electrical Engineering, Korea University
Seoul, Republic of Korea
[e-mail: csjeong@korea.ac.kr]

*Corresponding author: Chang-Sung Jeong

*Received November 27, 2018; revised February 18, 2019; revised March 31, 2019; accepted April 25, 2019;
published October 31, 2019*

Abstract

As technology advances, the amount of fake news is increasing more and more by various reasons such as political issues and advertisement exaggeration. However, there have been very few research works on fake news detection, especially which uses grammatical transformation on deep neural network. In this paper, we shall present a new Fake News Detection Model, called FAGON(Fake news detection model using Grammatical transformation On deep Neural network) which determines efficiently if the proposition is true or not for the given article by learning grammatical transformation on neural network. Especially, our model focuses the Korean language. It consists of two modules: sentence generator and classification. The former generates multiple sentences which have the same meaning as the proposition, but with different grammar by training the grammatical transformation. The latter classifies the proposition as true or false by training with vectors generated from each sentence of the article and the multiple sentences obtained from the former model respectively. We shall show that our model is designed to detect fake news effectively by exploiting various grammatical transformation and proper classification structure.

Keywords: Fake news detection, Grammatical transformation, Deep neural network

1. Introduction

These days, as the amount of fake news is getting large, it becomes essential to figure out how to detect the fake news. Fake news is the ones which deliver wrong information with lack of accuracy. To prevent the growing number of fake news, it is important to develop a fake news detection model based on a grammatical model which can detect the various forms of sentences with the same meaning. Also, it is necessary to train the grammatical model so that it can classify whether the sentence is true or false. Therefore, in this paper, we shall present a new Fake News Detection Model, called FAGON(Fake news detection model using Grammatical transformation On deep Neural network) which determines efficiently if the proposition is true or not for the given article by learning grammatical transformation on neural network with CNN classification. Especially, our model focuses the Korean language. It can verify the truthfulness of the proposition by training the Korean language model with a set of sentences with different grammatical forms with the same meaning.

FAGON consists of two modules: sentence generator and classification. The former generates a proposition sentence group(PSG) which consists of multiple sentences with the same meaning as the proposition, but different grammar by training the grammatical transformation using sentence generator model based on Seq2seq[1]. The latter first trains a classification model based on CNN with vectors generated from each sentence of the article and the multiple sentences obtained from the former model respectively, and then determines the proposition as true or false using the classification model. Unlike other models, FAGON uses fasttext embedding in order to reflect postpositions in Korean language grammar after the words are divided into morphemes[2]. Furthermore, we attach an independent CNN layer for each sentence of PSG and article in order to improve the accuracy of our model by extracting more word features.

We evaluate our model with three experiments: First, we evaluate our model in terms of PSG to figure out whether it is crucial for classification. Second, we compare three different classification methods: cosine similarity, Support Vector Machine(SVM) and CNN to find out an optimal method for classification. Third, we experiment our model to determine the optimal number of layers in CNN.

In Section 2, we review the previous work for grammar transformation and classification. In Section 3, we describe FAGON architecture in detail, and in Section 4, explain about three types of experiments for the performance evaluation of our model. In Section 5, we give a conclusion and future work.

2. Related Work

2.1 Sentence generator using grammar transformation

Developing a sentence generator model is important to detect fake news. There are several ways for sentence generator model which transforms the grammar of sentences including dependency tree and neural network. The dependency tree model is used as a language model for training grammar with unsupervised learning[3]. It builds the tree according to the distribution probability of the words at the boundaries of the sentences. However, when the length of the sentences increases, the accuracy of the dependency tree dramatically decreases and may produce wrong sentences. Recurrent Neural Network Grammars(RNNG) is developed for machine translation on neural network after trying many approaches of different parsing methods and transition methods for machine translation[4]. It divides an input

sentence with the parser, and then generates a target sentence by using parser transition based on top-down parsing. However, the parsing method is focused on English language and difficult to be used for other languages. Also, it uses bidirectional LSTM with parallel corpus sentences, but has difficulty in training grammar when the sequence is longer, and hence predicting the target word. Grammar transformation in sentence generator model is used in grammar error correction[5]. It uses a single hidden layer with mini-batch gradient descent by updating the parameters. Each sentence is divided into word vectors each of which is represented by one hot vector, and after concatenating word vectors for each sentence, they determine if two sentences are grammatically identical or not. However, their model using one hot vector is not adequate for the Korean language, since only the distribution frequency of words cannot reflect the context of Korean language properly. They developed a sentence generator model which make uses of reversing, sorting and replacement of the words for training based on Seq2seq model[6]. But, it uses only four types of grammatical transformation for a input sentence with relatively small length as training dataset. In order to overcome the problem of Korean language with a lot of postpositions and grammar forms with the same meaning regardless of the sentence length, we use fasttext embedding which reflects postpositions in Korean language grammar after the words are divided into morphemes.

2.2 Relation Classification

In order to detect the fake news by comparing the sentences, it is important to figure out the relation between the sentences, called as relation classification. In the past, the patterns of the sentences are compared, and classified without training[7]. However, these days, deep neural networks are widely used for training the features for the classification. In CNN tri-section model, the sentences are embedded into word sequence vector using word-net with the labels[8]. The model generates tri-section with 2 labels and word sequence vector. It extracts the features of word sequence vector through convolutional layers, and then trains by concatenating tri-section at the fully connected layer. Another work exploits modified CNN for relation classification to overcome the problem of CNN which has difficulty in reflecting the sentence context due to the fixed size of hidden vectors[9]. They find out the optimal window size of 2 for training by applying different window size to the input sentence. For another work with the CNN model, they use two kinds of embeddings, position embedding and word embedding[10]. After embedding, the model generates a shortest dependency path from each sentence. The feature of the embedded vectors are convoluted in CNN, and the weights are calculated with the weight matrix from the shortest dependency path. They also try to overcome the disadvantage of CNN for relation classification by using position embedding. Besides, they use attention based bidirectional LSTM for relation classification which exploits 9 different relation types of the dataset for training[11]. They design the greedy search decoder which gets one output with the highest possibility from attention layer, and compare the accuracy of their model with those using bidirectional LSTM, SVM and CNN[12].

In this paper, we develop a modified CNN model based on LSTM. which attaches the independent CNN layer for each sentence of PSG and article, further improving the accuracy of our model by extracting more word features.

3. Model

In this section, we present our model which consists of two modules: Sentence generator and classification. The former generates multiple sentences which have the same meaning as the proposition, but with different grammar by training the grammatical transformation. The latter classifies the proposition as true or false by training with vectors generated from each sentence of the article and the multiple sentences obtained from the former model respectively. Unlike other models, FAGON focuses on the Korean language which uses fasttext embedding to reflect postpositions in grammar after the words are divided into morphemes. Furthermore, we use CNN layers individually for the proposition and the article to extract more word features.

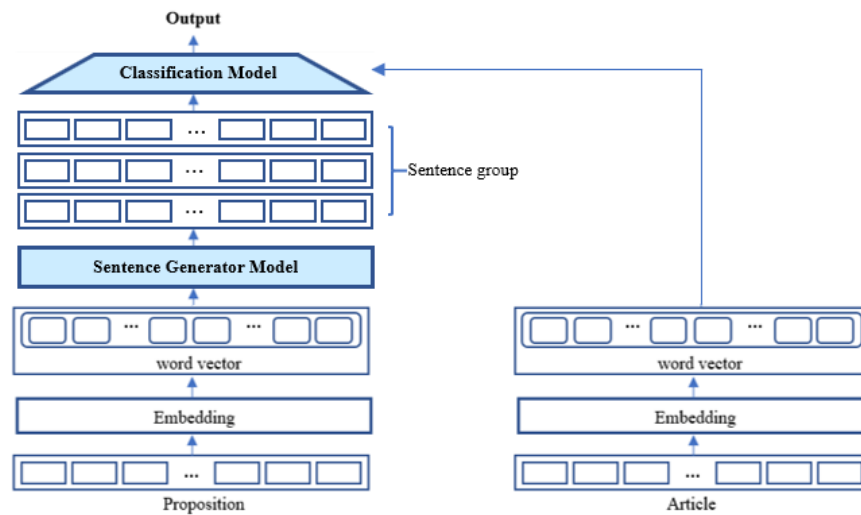


Fig. 1. The overall architecture of our model

Fig. 1 shows the overall process of our model. The sentences in the proposition and the article are tokenized by Mecab library into morphemes. Since the postpositions in the Korean language are important for the meaning of the sequence, we used the external library, Mecab. The split morphemes from the proposition and the article are embedded into fasttext embedded word vector. The word vectors enter sentence generator model and generate grammatically transformed sentences with the same meaning of the proposition. After it generates the sentences in sentence group, the sentences are compared and classified in the classification model. The details about each of the modules are described in the following sections.

3.1 Sentence Generator Model

The sentence generator model has three layers: Context generation layer, Matching layer and Output layer as shown in Fig. 2. The sentences are divided into morphemes, and the model generates a vocab list for fasttext representation for reflecting the word positions. After these are embedded in the vector, they enter the following layers.

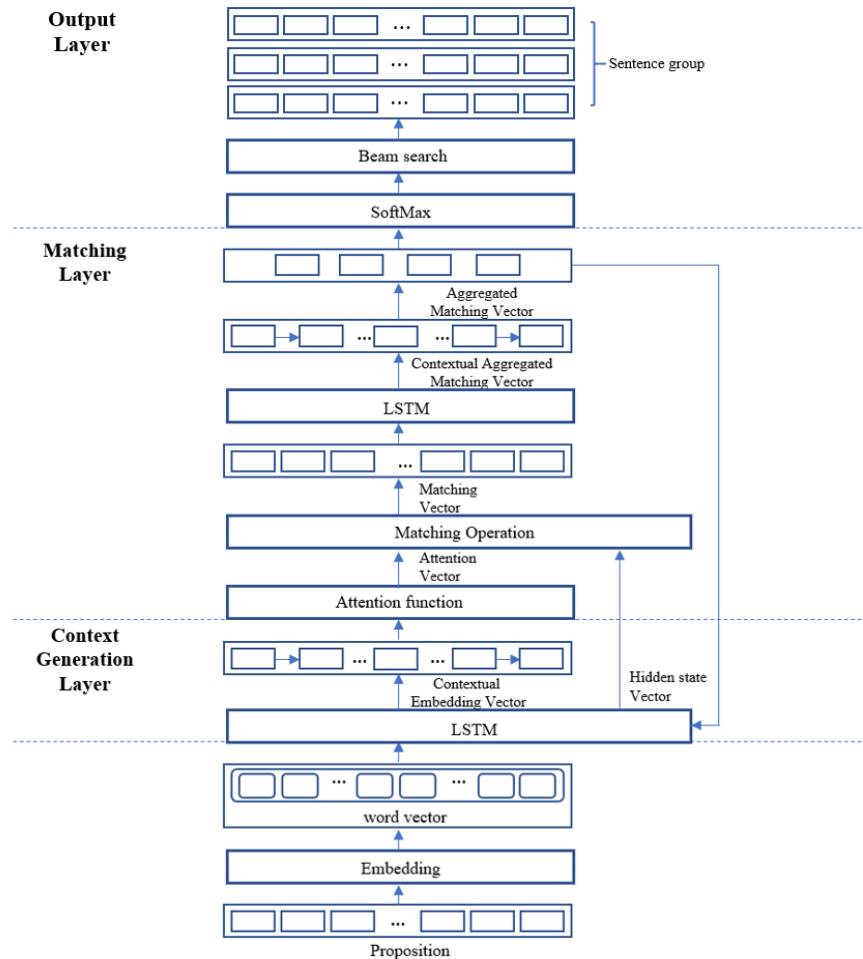


Fig. 2. The structure of sentence generator model

- **Context Generation Layer:** The word vectors enter LSTM layer generating contextual embedding vector. We use the LSTM layer rather than RNN layer to solve the vanishing gradient problem in long sequences of sentences.
- **Matching Layer:** The model computes the weighted sum of the words in the proposition which is called as an attention function which can consider the long sequences. The computed weights are used in matching operation as a matching vector. The vectors enter LSTM layer again for the higher accuracy. After the matching operation with attention vector, the vectors enter one more LSTM layer for higher accuracy. The output vector from LSTM layer is aggregated from the first output vector from LSTM layer and combined as the aggregated matching vector.
- **Output Layer:** The vectors which have the highest possibility are computed and decoded into the word from softmax function. The word combinations are generated as a sentence. When the vectors are decoded, we use beam search decoder so the various number of sentences can be generated. The model uses three size beam decoder and generates three sentences as output.

The sentence generator model can be divided and represented as encoder and decoder as in Fig. 3.

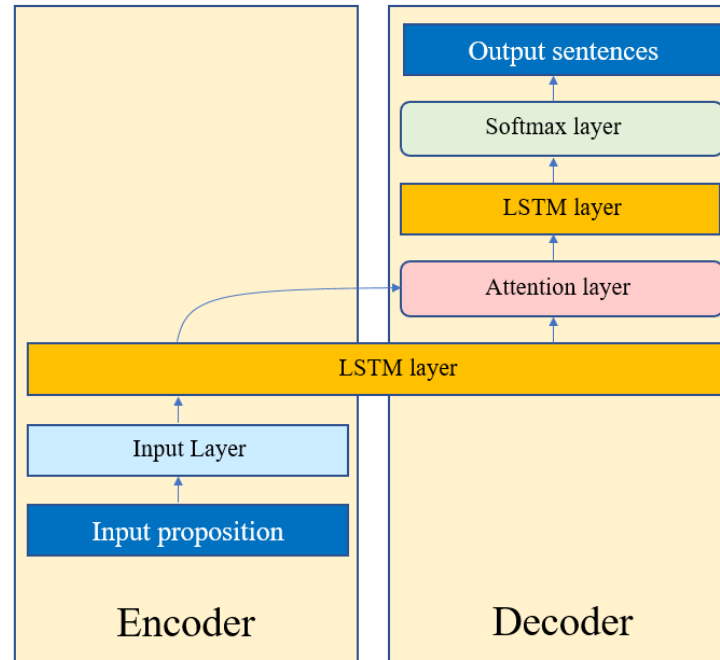


Fig. 3. The encoder and decoder of sentence generator model

From the encoder, the input proposition is encoded and represented as a hidden vector from LSTM layers. When the model generates the sentences, the vectors are encoded through attention layer by computing the weighted sum of the previous words in the proposition. After the softmax layer predicting the target sentences, the model can generate the three sentences. When the sentences are generated, mostly the greedy search decoder is used for the one output with the highest possibility. However, we use the beam search decoder with beam size 3 which can produce three sentences with three possible combinations of the words. We figured out the optimal beam size as 3 on our previous work[13].

3.2 Classification Model

The generated sentences from sentence generator model are needed to be classified for the output. The model for the classification has three convolutional layers, max pooling layer, flatten layer, fully connected layer, dropout layer and output layer for the final output. Unlike other models, we use CNN layers individually for the proposition and the article to extract more word features. Fig. 4 shows the overall model structure.

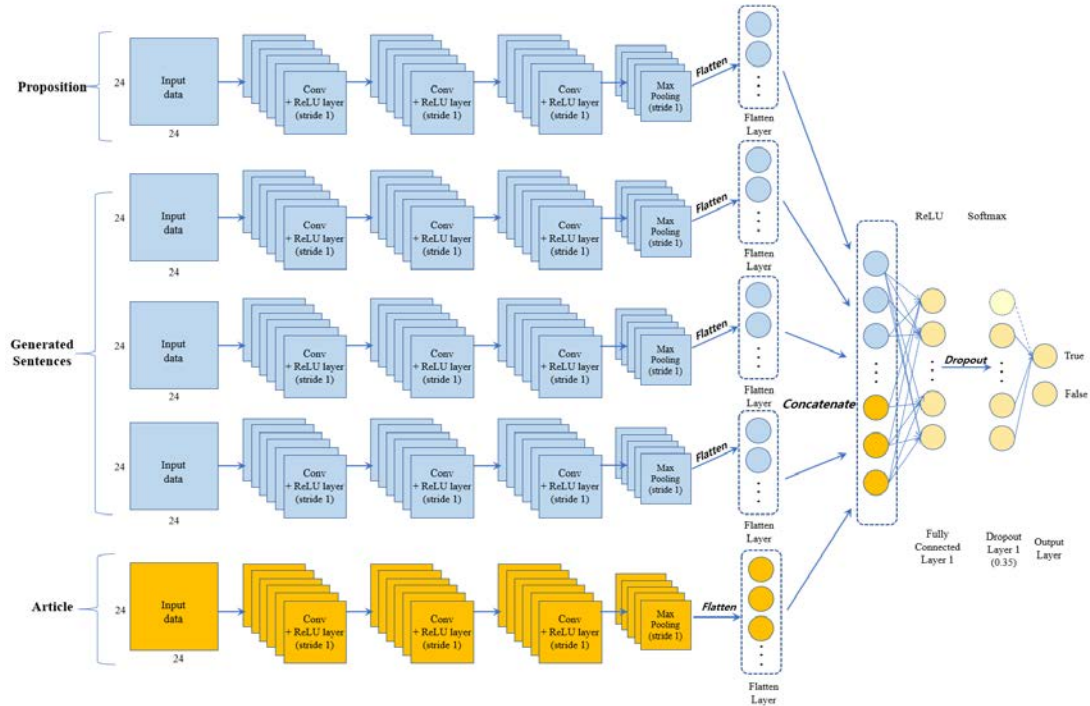


Fig. 4. The structure of classification model

The sentences in the proposition, article and the generated sentences are embedded by the same way in sentence generator model. The sentences are divided into morphemes, embedded into fasttext word vector. The input size of the proposition and the article are set as vector with the shape of (24, 24). We use the same size of every sentence from the article and the generated sentences. For extracting features, we put the generated sentences and article into three convolutional layers respectively. The deeper convolutional layer leads the prediction to the higher accuracy. Then, the vectors from three sentences and each sentence from the article are concatenated into fully connected layer with the activation function, ReLU to resolve vanishing gradient problem. The model computes the possibility of each of the sentences in the article with softmax function to classify the truthfulness of the news. Finally, the model gets the answer with the average value of the calculated values between 0 to 1. When the value is closer to 0, the answer is considered as true and when the value is almost 1, vice versa. Unlike other CNN models that all the sentences are converted into one vector and enter the convolutional layers, the sentences from the proposition and the article enter CNN layers individually for extracting more features in FAGON. Also, the model can classify better by using the generated sentences from the previous model which can consider the grammatical transformation.

4. Experiment

We implement our model in python 3.0.1 on Pycharm environment with 1 CPU(Intel® Core® CPU @ 3.50GHz) and memory of 8GB using two different types of dataset for the experiments with 1 GPU(GeForce GTX 1080) and memory of 8GB. Tensorflow library is used for the setup of the deep learning model.

4.1 Experiment Dataset

We make use of three types of dataset for the experiment. First, we use the parallel corpus dataset with different grammatical forms for training dataset in sentence generator model. Second, we exploit Korean news parallel corpus sentences and labels with true or false in classification model. Third, we test FAGON with 1,000 number of article dataset to evaluate the accuracy. The attribute of the dataset for sentence generator model is shown in **Table 1**.

Table 1. Dataset attributes of sentence generator model

Attribute type	Size
Number of corpus dataset	10,582
Number of epoch	100
Number of vocabulary	42,326
The maximum length of the sentence	200

In sentence generator model, the training dataset is made of the Korean news parallel corpus dataset, and we set the maximum length of the sentence as 200 words for converting the vectors into fixed-size. The details of the dataset for the classification model are presented in **Table 2**.

Table 2. Dataset attributes of classification model

Attribute type	Size
Number of training dataset	35,144
Number of validation dataset	10,040
Number of test dataset	5,021
Number of training epoch	300
Number of classes	2

In classification model, the dataset consists of the Korean news parallel corpus sentences and labels with two classes as true or false. True is labeled as 0, and false is represented as 1. The dataset is separated into training, validation and test dataset. Finally, the details of the test dataset are shown in **Table 3**.

Table 3. Dataset attributes for the test

Attribute type	Size
Number of test article	1,000
Number of vocabulary	3,765
Number of classes	2

We test the model with 1,000 number of the Korean articles to evaluate the accuracy of the model as shown in The correct answers over the total number of answers are calculated as accuracy.

4.2 Experiment Result

We make use of three types of dataset for the experiment. First, the model is tested with and without the sentence generator model which can show that generating grammatical transformed sentences is important for detecting the fake news. Second, we use several classification methods to figure out the optimal text classification method for fake news detection. Third, we perform the experiments for the classification model on the various number of layers to find the optimal depth of layers.

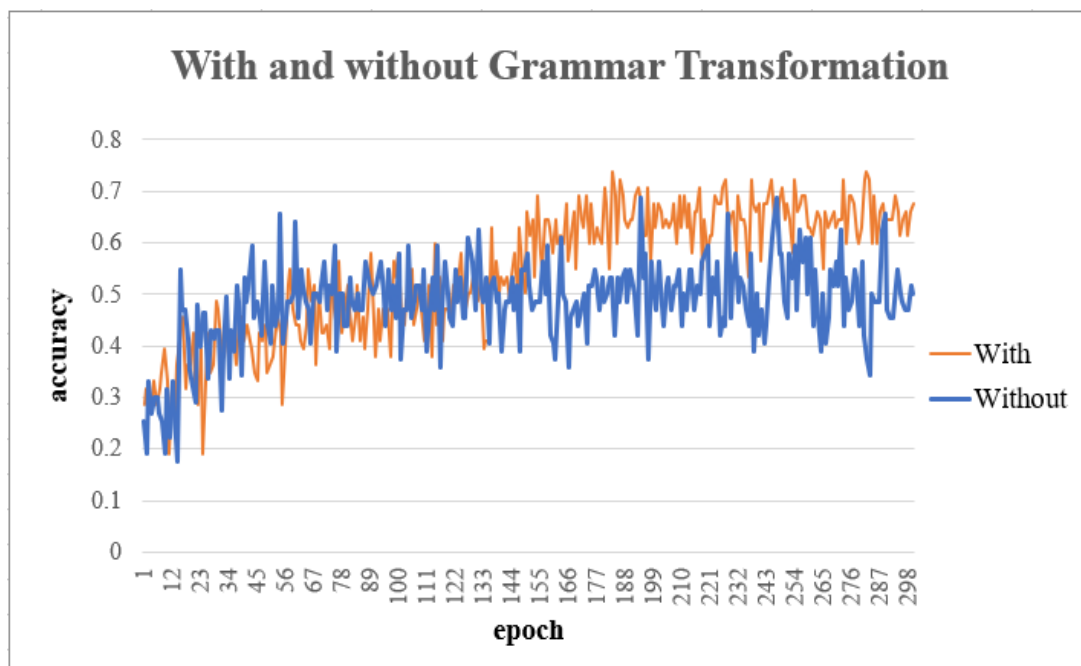


Fig. 5. With and without the grammatical transformation

First, we compare the model with and without the grammatical transformation. When the model classifies the answer without the grammatical transformation, it concatenates the proposition and each sentence in the article in classification model. The model is able to learn more word features when the model detects the fake news with a proposition and the generated sentences. Thus, when the grammatical transformation sentences in the model are generated, the training accuracy of the model shows better than the result without it. The concatenation of the generated sentences and the article also shows better training about the average as 0.72 out of 1. Therefore, we use the grammatical transformation sentences for the classification which can detect the various forms of the sentences.

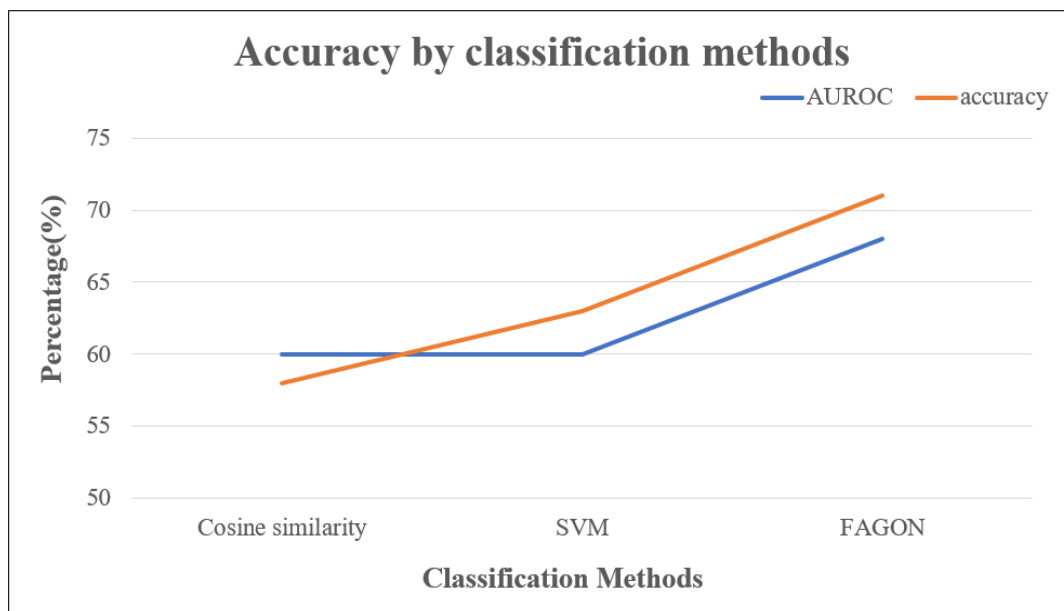


Fig. 6. Accuracy by different classification methods

Second, we make use of various classification methods such as cosine similarity, SVM and CNN classification. With the sentence generator model, the classification model part is switched into the methods above. We calculated the accuracy and AUROC value for better comparison[14]. When the words are represented as a word vector, the relations can be calculated as cosine similarity. Also, clustering the classes with SVM is largely used for the relation methods. For clustering the answers with two classes as true and false, SVM model is used as an unsupervised learning algorithm. When we get the answer from calculating cosine similarity, it is difficult to detect truthfulness when the format of the proposition is changed as passive or active forms. However, SVM model relatively figures out the answer better than cosine similarity by clustering the answers. The SVM model calculates the each vector's distance from the true and false cluster, and finally classifies the answer. But the relation between the sentences is classified better in FAGON since the model trains more sentences. The labels of the concatenated vectors are trained well, so the model figures out the truthfulness and detects the fake news. We compare the model also with AUROC to calculate more accurate accuracy considering the distribution of the answers. Area Under the Receiver Operating Characteristics Curve is called as AUROC considers the distribution of the classes and calculate the accuracy better. AUROC takes into account the number of answer data whether the model answers 0 for 0, 1 for 0, 0 for 1, or 1 for 1. FAGON has higher accuracy and AUROC value over two other methods.

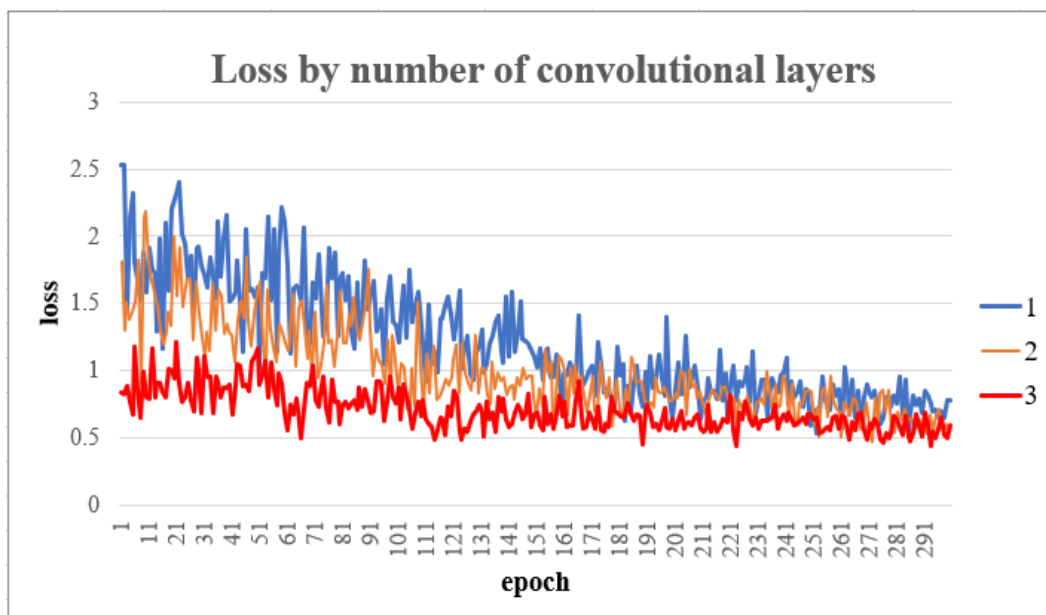


Fig. 7. Loss by number of convolutional layers

Third, we evaluate the model by changing the number of convolutional layers as shown in Fig. 7. As we have chosen using CNN layer with the experiments above, the model is tested with the different number of layers. The depth of the layers is important for training the model, and the graph shows the loss according to the number of the layers. When we set the training epoch as 300, the loss increases from the start of the epoch, and it decreases as the training goes on. When we increment the number of the convolutional layer on the classification model, the loss decreases that means the model trains better. Therefore, we set the three convolution layers.

5. Conclusion

In this paper, we have presented a new Fake News Detection Model, called FAGON(Fake news detection model using Grammatical transformation On deep Neural network) which determines efficiently if the proposition is true or not for the given article by learning grammatical transformation on neural network with CNN classification. Especially, our model focuses the Korean language. It can verify the truthfulness of the proposition by training the Korean language model with a set of sentences with different grammatical forms with the same meaning. In order to overcome the problem of Korean language with a lot of postpositions and grammar forms with the same meaning regardless of the sentence length, we use fasttext embedding which reflects postpositions in Korean language grammar after the words are divided into morphemes. Also, we have developed a modified CNN model based on LSTM, which attaches independent CNN layer for each sentence of PSG and article, further improving the accuracy of our model by extracting more word features.

FAGON consists of two modules: sentence generator and classification. The former generates a proposition sentence group(PSG) which consists of multiple sentences with the same meaning as the proposition, but different grammar by training the grammatical

transformation using sentence generator model based on Seq2seq. The latter first trains a classification model based on CNN with vectors generated from each sentence of the article and the multiple sentences obtained from the former model respectively, and then determines the proposition as true or false using the classification model.

We evaluate our model with three experiments: First, we evaluate our model in terms of PSG to figure out whether it is crucial for classification. Second, we compare three different classification methods: cosine similarity, Support Vector Machine(SVM) and CNN to find out an optimal method for classification. Third, we experiment our model to determine the optimal number of layers in CNN.

Now, we are planning for implementing our fake detection model on distributed parallel environment in order to improve its overall performance. Also, we shall continue to develop a distributed natural language processing platform for processing various type of QA systems very fast in static and dynamic environment such as dealing with streaming data.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03035461), the Brain Korea 21 Plus Project in 2018, and the Institute for Information & communications Technology Promotion(IITP) grant funded by the Korean government (MSIP) (No. 2018-0- 00739, Deep learning-based natural language contents evaluation technology for detecting fake news).

References

- [1] Sutskever, Ilya, Oriol Vinyals and Quoc V, "Sequence to sequence learning with neural networks," *Advances in Neural Information Processing Systems*, pp. 3104-3112, 2014. [Article \(CrossRef Link\)](#)
- [2] Bojanowski, Piotr, "Enriching Word Vectors with Subword Information," *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 135-146, 2017. [Article \(CrossRef Link\)](#)
- [3] Spitkovsky, Valentin I., Hiyun Alshawi and Daniel Jurafsky., "Three dependency-and-boundary models for grammar induction," in *Proc. of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 688-698, 2012. [Article \(CrossRef Link\)](#)
- [4] Dyer, C., Kuncoro, A., Ballesteros, M. and Smith, N. A, "Recurrent neural network grammars," in *Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics, Human Language Technologies*, pp. 199-209, 2016. [Article \(CrossRef Link\)](#)
- [5] Sakaguchi, Keisuke, Matt Post and Benjamin Van Durme, "Grammatical error correction with neural reinforcement learning," *arXiv preprint arXiv:1707.00299*, 2017. [Article \(CrossRef Link\)](#)
- [6] Wang, T., Chen, P., Amaral, K. and Qiang, J, "An experimental study of lstm encoder-decoder model for text simplification," *arXiv preprint arXiv:1609.03663*, 2016. [Article \(CrossRef Link\)](#)
- [7] Riedel, S., Yao, L., McCallum, A. and Marlin, B. M, "Relation extraction with matrix factorization and universal schemas," in *Proc. of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 74-84, 2013. [Article \(CrossRef Link\)](#)
- [8] De Silva, T. N., Zhibo, X., Rui, Z. and Kezhi, M, "Causal relation identification using convolutional neural networks and knowledge based features," *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 11, no. 6, pp. 696-701, 2017. [Article \(CrossRef Link\)](#)

- [9] Luo, Y., Cheng, Y., Uzuner, Ö., Szolovits, P. and Starren, J., “Segment convolutional neural networks (Seg-CNNs) for classifying relations in clinical notes,” *Journal of the American Medical Informatics Association*, vol. 25, no. 1, pp. 93-98, 2017. [Article \(CrossRef Link\)](#)
- [10] Tan, Z., Li, B., Huang, P., Ge, B. and Xiao, W., “Neural relation classification using selective attention and symmetrical directional instances,” *Symmetry*, vol. 10, no. 9, pp. 357, 2018. [Article \(CrossRef Link\)](#)
- [11] Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H. and Xu, B., “Attention-based bidirectional long short-term memory networks for relation classification,” *Proceedings of 54th Annual Meeting of the Association for Computational Linguistics*, Vol. 2, pp. 207-212, 2016. [Article \(CrossRef Link\)](#)
- [12] Boser, B. E., Guyon, I. M. and Vapnik, V. N., “A training algorithm for optimal margin classifiers,” in *Proc. of the fifth annual workshop on Computational learning theory*, pp. 144-152, 1992. [Article \(CrossRef Link\)](#)
- Youngkyung Seo and Chang-sung Jeong, “FAGON: Fake News Detection Model Using Grammatical Transformation on Neural Network,” in *Proc. of The 13th International Conference on Knowledge*, 2018. [Article \(CrossRef Link\)](#)
- [13] Bradley, A. P., “The use of the area under the ROC curve in the evaluation of machine learning algorithms,” *Pattern recognition*, vol. 30, no. 7, pp. 1145-1159, 1997. [Article \(CrossRef Link\)](#)



Youngkyung Seo is currently pursuing a M.S. degree in Department of Electrical Engineering from Korea University. Her research interests include the deep learning and distributed parallel processing.



Seong-Soo Han is a professor in the Division of Liberal Studies from Kangwon National University. He received the Ph.D. degree in Visual Information Processing from Korea University in 2018. He is research ubiquitous computing, AI, blockchain, deep learning and distributed parallel processing.



You-Boo Jeon is a professor in the Division of Computer Science and Engineering at Soonchunhyang University. He received the B.S. degree in Computer Science from Soonchunhyang University, Korea in 1988, the M.S. degree in Electronic and Computer Engineering from Korea University, Korea in 2002, the Ph.D. degree in Visual Information Processing from Korea University in 2013. He was a CTO of the ValonTec Inc. in 2006–2013, and was the CTO of the PowerGrid Inc. in 2013–2014. His research interests include software engineering, ubiquitous computing, m2m, IoT, and mobile communication security.



Chang-Sung Jeong is a professor at the Department of Electrical Engineering at Korea University. Before joining Korea University in 1992, he was a professor at POSTECH during 1982–1992. He was on editorial board for Journal of Parallel Algorithms and Application in 1992–2002. Also, he was a chair of IEEE Seoul Section and has been working as a chairman of Computer Chapter at Seoul Section of IEEE region 10. He was a chairman of EE department in Korea University and a leader of BK21 project. His research interests include distributed parallel computing, cloud computing, networked virtual environment, and distributed parallel deep learning for real-time image processing and visualization.