Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018)

FaNDeR: Fake News Detection Model Using Media Reliability

Youngkyung Seo
Department of Electrical Engineering
Korea University
Seoul, Republic of Korea
ygseo@korea.ac.kr

Deokjin Seo
Nuua
Seoul, Republic of Korea
westevil@nuua.net

Chang-Sung Jeong
Department of Electrical Engineering
Korea University
Seoul, Republic of Korea
csjeong@korea.ac.kr

Abstract— **With the development of media including newspaper written by robots and many unreliable sources, it's getting hard to distinguish whether the news is true or not. In this paper, we shall present a novel fake news detection model, FaNDeR(Fake News Detection model using media Reliability) which can efficiently classify the level of truth for the news in the question answering system based on modified CNN deep learning model. Our model reflects the reliability of various medias by training with the input dataset which contains the truthfulness of each media as well as that of the proposition. Our model is designed for higher accuracy with media dataset in terms of data augmentation, batch size control and model modification. We shall show that our model has higher accuracy over statistical approach by reflecting the tendency of truth level for each media through the training of the dataset collected so far.**

Keywords— Source, Media, Reliability, Fake news, Deep learning, Question Answering System;

I. INTRODUCTION

With the number of medias increasing, it is getting more important to read the right news, and more essential to figure out fake news which delivers wrong information with lack of accuracy[1]. Since the truth of news is verified by people, it takes long to find it. In order to spot fake news, it is not sufficient to consider only one media, but necessary to introspect the truth of the given news in various medias. Moreover, it is more effective to take into account the reliability of each media for more accurate estimation of the fake news.

In this paper, we shall present a novel fake news detection model, FaNDeR(Fake News Detection model using media Reliability) which can efficiently classify the level of truth for the news based on CNN deep learning mode Our model reflects the reliability of various medias by training CNN with the input dataset which contains the truthfulness of each media as well as that of the proposition. Our model consists of two parts: fake news detection for each media and reliability checking on deep learning model. The former detects whether the proposition is true, false or neutral for each media, while the latter determines whether it is true, false or neutral on deep learning model considering the reliability of each media which are implemented on CNN through the training over the dataset obtained from various medias. Our model is designed for higher accuracy with media dataset in terms of data augmentation, batch size control and model modification. Also, we shall show that our model achieves the higher accuracy by reflecting the tendency of truth level for each media through the training of the previous dataset collected so far rather than statistical approach which just measures the frequency of right answers for each media.

The rest of our paper is as follows: In section 2, we describe related works, and in section 3, present our model for fake news detection on CNN which reflects the reliability of various medias. In section 4, we explain the experimental results, and in section 5, give a conclusion.

II. RELATED WORKS

A. Fake News Detection

According to the survey paper explaining fake news detecting system[2], there are various kinds of system. BS detector was developed in Mozila as a plugin to warn fake news website, and it is being used in Facebook[3]. It can identify fake news sites, and show flags to the news measured by people. American government employees evaluate American political news checking whether political news are true in person in the website, PolitiFact[4]. It aims to check the truth by Truth-O-Meter ratings showing six levels of truthfulness. But category of the news that website can check is limited to political news, and also the truth of news is checked by people. There is a system, Flock Fake News Detector to detect fake news with chat service ranking the news URLs statistically figuring out whether they are true by users of the chat service[5]. In the detecting systems mentioned above, the result about deception of news is verified by people, which is a critical limitation of fake news detection due to the lack of dataset as well as the long time taken to get the result. In this paper, we shall present a model which can solve these limitations.

B. Question Answering System

In the past, data was stored in structural form, and the system simply sent queries to the database to get answers to the questions[6,7]. As the development of machine comprehension, various deep learning models are adjusted to the question answering system[8]. However, there have been some problems of language modeling in deep learning models such that it is hard to make meaningful sentences, since hidden unit vector has the information about the word just before it was encoded. When question answering system such as SiamFC that uses template-matching methods has high speed to train and figure out the answer, but it has low accuracy as the target sentences have various forms[9,10,11]. Thus, there could be the loss when long sentences are put into the hidden vector. To avoid the loss of the long input, bi-directional RNN came out to decrease the loss of sentences by encoding the words into hidden vector both forward and backward of the original sentences and combining the vectors together[12,13]. Memory network solves the problem using two modules: saving sentences or words in external memory and finding the sentences that are most related to the question[14]. Thus, the model could overcome the problems when the form of target sentences varies. It produces the answer combining the highest possible answer and the question. Facebook adjusted this model to question answering system, and got 91 percent of accuracy[15]. Since the above two models need to be supervised, it's essential for accepting available values when they are trained. End-to-end memory network is developed to recover those models. It generates the final answer by calculating weighted sum of the similarity of the question and sentences, and updating the answers in memory[16]. To recover small size of memory, dynamic memory network is developed for training set of input sequences and questions by forming episodic memories for using them to generate relevant answers[17].

C. Classification with small amount of data

CNN model is generally used to classify data among deep learning models. Particularly, it is adjusted to image dataset classifying suitable classes extracting features. According to Imagenet, it classifies the image data with various levels of convolutional layers showing high accuracy[18]. Alike Imagenet, CNN model is adjusted to the answer dataset of various media to classify three classes, true, false, or neutral. Our model is based on Imagenet, but the numbers and shape of layers are transformed properly for higher accuracy of our answer dataset. However, the amount of training data is small, leading to overfitting, which happens when a model learns the detail and noise in the training data, and impacts the model ability negatively. There are several ways to solve such problem.

First of all, the size of training batch size can be changed for the dataset. It's important to choose the right batch size depending on the shape and size of the dataset for training. Furthermore, fine-tune deep learning can be used for preventing overfitting. CNN model commonly sets the initial value of weight as random value and updates the weight with training. Moreover, One-shot learning can be adjusted to the small amount of dataset by calculating the similarity of the classes and classifying the dataset into the various number of classes[19]. It is available when the number of classes is large. Next, image augmentation increases the number of training dataset by transforming images, adjusting angles, resizing and flipping horizontal or vertical as given or random setting for the small amount of dataset[20]. Besides the image dataset, image augmentation can be adjusted to the media dataset used in our model by transforming the features of the results from fake news detection model.

III. MODEL

In this section, we shall describe our model in detail. It consists of two parts: fake news detection for each media and reliability checking on deep learning model as shown in Fig. 1. The former detects whether the proposition is true, false or neutral for each media, while the latter determines whether it is true, false or neutral on deep learning model considering the reliability of each media which are implemented on modified CNN through the training over the dataset obtained from various medias.

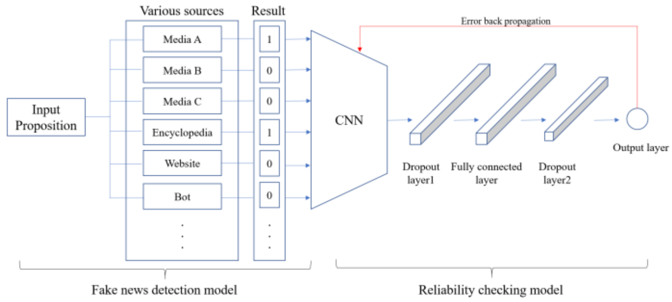

Fig. 1. The overall architecture of FaNDeR

Unlike the traditional fake news detection model, it increases the accuracy by extracting the answers from various medias in the fake news detection part of our model, and then exploiting them as input into modified CNN in the reliability checking part. The reliability of each media is reflected by updating its weights from the input dataset, which results in more accurate fake news detection.

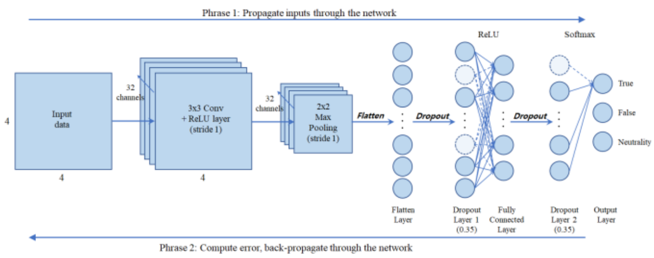

Fig. 2. The structure of reliability checking model based on modified CNN

Our reliability checking model makes use of modified CNN which consists of several layers: input layer, convolutional layer, max pooling layer, two dropout layers, fully connected layer, softmax layer and output layer. Input layer receives a set of input data each with size 4 x 4 x 1, and output layer generates one of classes: true, false, neutral. In convolutional layer, l2 regularization is used in order to spread the weight evenly, and decrease noise values, and the suitable size of filter and stride is adjusted to get the more accurate result. To prevent overshooting, the proper learning rate is adjusted. Our model is designed to be simple rather than complex in order to use small amount of data more effectively. It is achieved by setting dropout value relatively large which eliminates the unnecessary nodes and hence minimizes variables in the deeper layers. Additional dropout layer is added to our model for higher accuracy with the small amount of dataset. In softmax layer, softmax with loss method is adjusted by updating weight appropriately for the efficient classification[21].

IV. 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 three different types of dataset for the experiments. Keras library is used for the set up of the deep learning model.

A. Experiment Dataset

We evaluate the performance of FaNDeR using three types of dataset. The first dataset is MNIST dataset each with 28 x 28 digit image. The original MNIST dataset consists of 10 classes, 0 to 9, and 3 channels, Red, Green and Blue. Our model consists of three classes, True, False and Neutral with 1 channel, that is, 0, 1, 2 centered in fixed-size image, decreasing the number of classes into three.

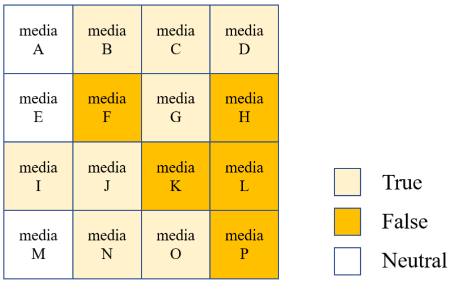

Fig. 3. The format of media dataset

The second dataset is the refined one arranged for accuracy test of our model before using real dataset as shown in Fig. 3. Each data is 4 x 4 matrix from 16 medias, where false, true and neutral are represented by 0, 1 and 2 respectively. Among the various medias, we set only one media as perfect accuracy, and the another one as zero percent of accuracy, and the others as random for the better training of our model.

The final dataset is media dataset similar to the second one, but collected from real news by checking their truth from 16 medias. The following Table 1, 2 and 3 show the specifications for each dataset used in our model.

TABLE I. THE FIRST DATASET ATTRIBUTES

Attribute type	Size
Data size	28 x 28
Number of channels	1
Number of classes	3
Number of training dataset	7,728
Number of validation dataset	1,472
Number of test dataset	6,389

TABLE II. THE SECOND DATASET ATTRIBUTES

Attribute type	Size
Data size	4 x 4
Number of channels	1
Number of classes	3
Number of training dataset	4,149
Number of validation dataset	790
Number of test dataset	244

TABLE III. THE THIRD DATASET ATTRIBUTES

Attribute type	Size
Data size	4 x 4
Number of channels	1
Number of classes	3
Number of training dataset	84
Number of validation dataset	18
Number of test dataset	21

B. Experiment Result

We experiment our model FaNDeR for each of three different dataset respectively. Moreover, we try to improve our model for higher accuracy in terms of data augmentation, batch size control and model modification

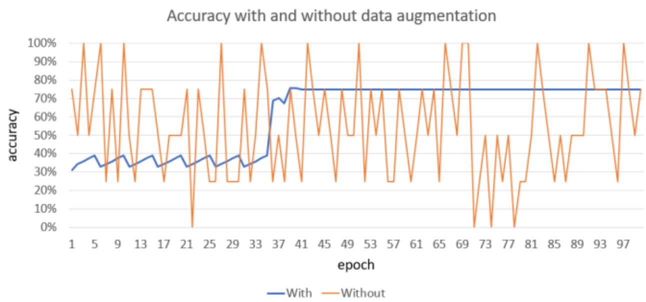

Fig. 4. Accuracy of FaNDeR with and without data augmentation

First, we figure out whether data augmentation is available or not for training. The amount of media data is much smaller than that of MNIST and refined dataset. The experiment for MNIST and refined dataset shows very high accuracy, while that for media dataset lower accuracy and overfitting. To solve the problem, we make use of data augmentation for media dataset for higher accuracy.

Originally, data augmentation is a technique of various transformation for image dataset which is used for the extraction of features in image processing. For higher accuracy with the small amount of media dataset, we exploit data augmentation which enables more efficient extraction of various features in media dataset. In our experiment, augmented media dataset is obtained by rotating in various degrees respectively left and right side, and we get the highest accuracy when rotating left and right side by 5 degrees each, complementing the small amount of media dataset. The accuracy graph with and without image augmentation shows that there arises overfitting with uneven accuracy without the image augmentation in Fig. 4.

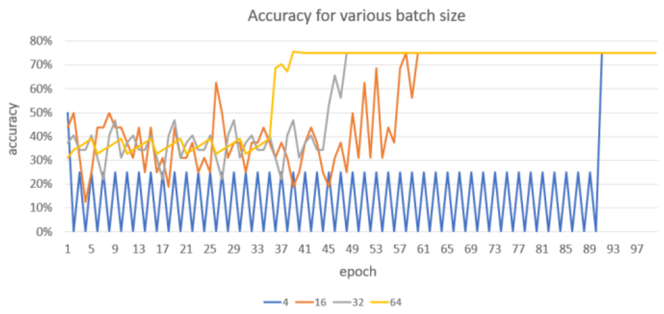

Fig. 5. Accuracy of FaNDeR with media dataset for various batch size

Second, we find out the optimal batch size for the highest accuracy by changing the training batch size such as 4, 16, 32 and 64 as in Fig. 5. Our model shows the highest accuracy when the training size is 64 with MNIST, refined and media dataset. Therefore, the batch size is set as 64 in the experiment.

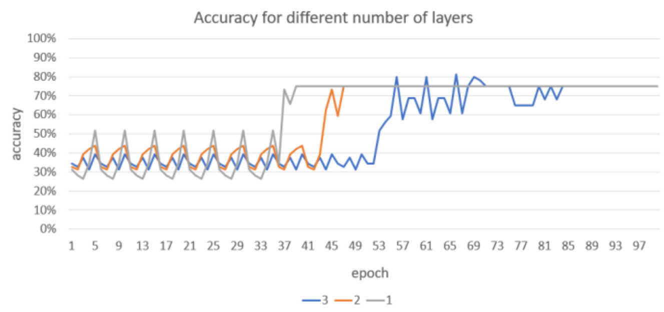

Fig. 6. Accuracy of FaNDeR with media dataset for different number of layers

Third, we evaluate our model with different number of convolutional layers: 1, 2 and 3. Our model has higher accuracy for MNIST and refined dataset with relatively large amount of dataset than media dataset, while lower accuracy in media dataset when the numbers of layers are large. As shown in Fig. 6, our model shows the higher accuracy for media dataset on the simple model with only one convolutional layer rather than the larger numbers of layers. Therefore, our model is designed with one convolutional layer. Besides, our experiment shows that our model has higher accuracy when one dropout layer is added.

TABLE IV. ACCURACY OF EACH MODEL

Model	*MNIST dataset*	*Refined dataset*	*Media dataset*
FaNDeR	100%	97.7%	75.6%
Statistical model	88%	86%	72%

We compare the accuracy of our model for three different kinds of dataset: MNIST dataset, refined dataset and media dataset in Table 4. Also, our model is compared with statistical method which measures the frequency of right answers for each media. Our experiment shows that statistical method does not figure out the right answer very well, and our model gets the right answers about 3% better than the statistical method.

V. CONCLUSION

In this paper, we have presented a novel fake news detection model, FaNDeR which can efficiently classify the level of truth for the news in the question answering system based on modified CNN deep learning model.

We have tested our model with three types of dataset: MNIST, refined and media dataset. The first dataset is MNIST dataset each with 28 x 28 digit image. The second dataset is the refined one arranged for accuracy test of our model. The final dataset is media dataset similar to the second one, but collected from real news by checking their truth from 16 medias.

Our model has been designed for higher accuracy with media dataset in terms of data augmentation, batch size control and model modification. We have used data augmentation for media dataset in order to solve the problem of lower accuracy and overfitting. Furthermore, we find out the optimal batch size for the highest accuracy of MNIST, refined and media dataset. Also, we evaluate our model with different number of convolutional layers in order to figure out the optimal model with the highest accuracy. Eventually, our model achieves higher accuracy than statistical approach by reflecting the reliability of various medias during training with the input dataset which contains the truthfulness of each media as well as that of the proposition.

We believe that our proposed model based on deep learning model can be used efficiently not only for detecting fake news but also for evaluating the reliability of various medias. In the future, we shall continue to improve our fake news detection model with larger dataset, and implement it on distributed parallel environment for fast stream processing of fake news detection coming from various medias in real time.

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] E. Kiely and L. Robertson, "How to Spot Fake News - FactCheck.org", FactCheck.org, 2016. [Online]. Available: https://www.factcheck.org/2016/11/how-to-spot-fake-news/. [Accessed: 10- Jul- 2018].

[2] M. Gahirwal, S. Moghe, T. Kulkarni, D. Khakhar and J. Bhatia, "Fake News Detection", International Journal of Advance Research and Innovations in Technology, vol. 4, no. 1, pp. 817-819, 2018.

[3] "B.S. Detector - Browser extension to identify fake news sites", Bsdetector.tech, 2018. [Online]. Available: http://bsdetector.tech/. [Accessed: 10- Jul- 2018].

[4] "PolitiFact", Choice Reviews Online, vol. 45, no. 12, pp. 45-7061-45-7061, 2008.

[5] N. Shah, "FLOCK: Combating Astroturfing on Livestreaming Platforms", WWW '17 Proceedings of the 26th International Conference on World Wide Web, pp. 1083-1091, 2017.

[6] D. Waltz, "An English language question answering system for a large relational database", Communications of the ACM, vol. 21, no. 7, pp. 526-539, 1978.

[7] E. Sneiders, "Automated Question Answering Using Question Templates That Cover the Conceptual Model of the Database", International Conference on Application of Natural Language to Information Systems, pp. 235-239, 2002.

[8] S. Wang and J. Jiang, "Machine Comprehension Using Match-LSTM and Answer Pointer", arXiv:1608.07905 [cs.CL], 2016.

[9] C. Unger, L. Bühmann, J. Lehmann, A. Ngonga Ngomo, D. Gerber and P. Cimiano, "Template-based question answering over RDF data", WWW '12 Proceedings of the 21st international conference on World Wide Web, pp. 639-648, 2012.

[10] T. Gunawardena, M. Lokuhetti, N. Pathirana, R. Ragel and S. Deegalla, "An automatic answering system with template matching for natural language questions", 2010 Fifth International Conference on Information and Automation for Sustainability, pp. 353-358, 2010.

[11] D. Wang, "A Domain-Specific Question Answering System Based on Ontology and Question Templates", in 2010 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, London, UK, pp. 151-156, 2010.

[12] D. Wang and E. Nyberg, "A long short-term memory model for answer sentence selection in question answering", in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, pp. 707-712, 2015.

[13] M. Seo, A. Kembhavi, A. Farhadi and H. Hajishirzi, "Bidirectional attention flow for machine comprehension", in 5th International Conference on Learning Representations, Palais des Congrès Neptune, Toulon, France, 2016.

[14] A. Bordes, N. Usunier, S. Chopra and J. Weston, "Large-scale simple question answering with memory networks", arXiv:1506.02075, 2015.

[15] J. Weston, A. Bordes, S. Chopra, A. M. Rush, B. van Merriënboer, A. Joulin and T. Mikolov, "Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks", arXiv:1502.05698, 2015.

[16] Sukhbaatar, Sainbayar, End-to-end memory networks. In: Advances in neural information processing systems, p. 2440-2448, 2015.

[17] Xiong, Merity, Socher, “Dynamic memory networks for visual and textual question answering,” International Conference on Machine Learning, pp. 2397-2406, 2016.

[18] Krizhevsky, Alex, Ilya, Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, pp. 1097-1105, 2012.

[19] Vinyals, Oriol, Lillicrap, Wierstra "Matching networks for one shot learning," Advances in Neural Information Processing Systems, pp. 3630-3638, 2016.

[20] Campbell, Lucas-Woodley, Lord, Sykes, Patent No. 8,767,015, U.S. Patent and Trademark Office, 2014.

[21] Liu, Weiyang, "Large-Margin Softmax Loss for Convolutional Neural Networks," ICML, pp. 507-516, 2016.