Deep Learning-based Question Answering System for Proactive Disaster Management

Yohan Kim[a], Jiu Sohn[a], Seongdeok Bang[a], Hyoungkwan Kim[a*]

[a]Department of Civil and Environmental Engineering, Yonsei University, Seoul, Korea

E-mail: homez815@yonsei.ac.kr, jiujohn@yonsei.ac.kr, bangdeok@yonsei.ac.kr, hyoungkwan@yonsei.ac.kr

Abstract –

As climate change increases the frequency and intensity of natural disasters, proactive disaster management is needed to reduce the damage caused by the natural disasters. Existing reports that record the scale, damage, and response of natural disasters can be used as references for proactive disaster management. However, it is labor-intensive and time-consuming to manually find the necessary information from a number of reports. Thus, this study proposes a natural language processing (NLP)-based question answering system (QA system) for proactive disaster management using the existing reports. This study is focused on paragraphs retrieval, which retrieves paragraphs that have a high similarity to a given question based on the word embedding. The National Hurricane Center's Tropical Cyclone Reports are used to evaluate the proposed method.

Keywords –

Deep Learning; Disaster Management; Natural Language Processing; Question Answering System

1 Introduction

Natural disasters have negative impacts on infrastructure such as structural failure. In recent years, as the frequency and intensity of natural disasters have increased due to the impact of climate change, the importance of disaster prevention has increased [1].

Analyzing historical data and extracting the necessary information can help disaster prevention. With the advancement of deep learning and natural language processing, many studies for information extraction use unstructured data such as text data. Wang and Taylor proposed a method for detecting urban emergencies using Twitter data and topic modelling techniques [2]. Sit et al. conducted a study to identify disaster-related tweets using deep learning, natural language processing, and spatial analysis [3]. Ragnini et al. proposed a data analysis method for disaster response and recovery using Twitter and sentiment analysis [4].

The Question Answering system (QA system) provides users with answers for questions regarding the data. The QA system is useful in finding the necessary information in a large amount of data. Chan and Tsai proposed a dialogue system that combines a QA module and a knowledge base for emergency operations [5]. Tsai et al. proposed a chatbot system called Ask Diana that provides users with water-related information [6].

Social media data, mainly used in previous studies, can satisfy the need of big data for deep learning models. However, they contain inaccurate or unnecessary data. Previous studies that proposed QA systems were mostly focused on extracting keywords. This keyword-based information is intuitive, but it is difficult to grasp the context of the information.

In this study, we propose a QA system for proactive disaster management. The purpose of the proposed system is to provide users with the necessary information to reduce the damage to the infrastructure caused by natural disasters, especially tropical cyclones. In this study, the Tropical Cyclone reports provided by the National Hurricane Center are used. Since tropical cyclones are one of the major weather phenomena that cause enormous damage to infrastructure, they are chosen as the subject of disaster management. We conduct paragraphs retrieval, one of the steps in the QA system, and represent its results.

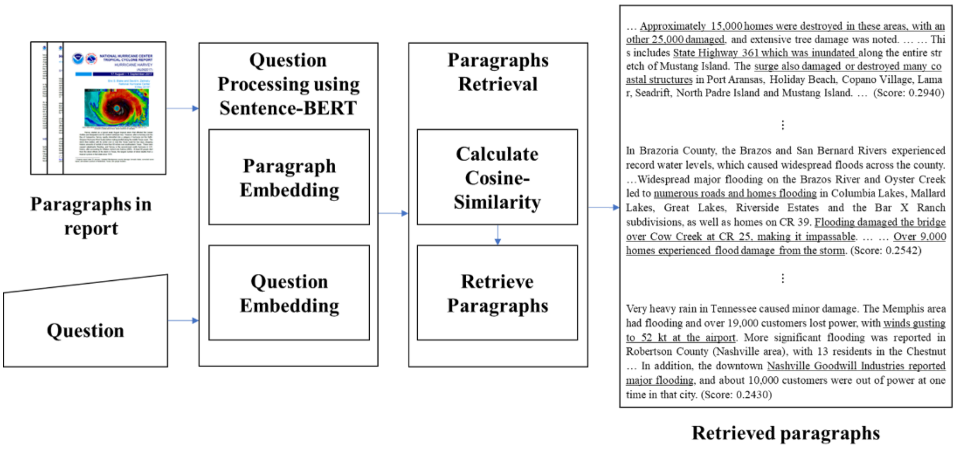

Figure 1. Overview of the proposed paragraphs retrieval in the QA system

2 Paragraphs retrieval using sentence-BERT

2.1 BERT and sentence-BERT

In the past, deep learning models mainly used in natural language processing were Recurrent Neural Network (RNN) models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). The performance of RNN models is excellent, but they have a vanishing gradient problem. When the input sequence is long, the prediction accuracy is decreased by the vanishing gradient problem. To solve the vanishing gradient problem, Sutskever et al. proposed a seq2seq model using the attention mechanism [7]. In the Seq2seq model, the output word is predicted considering all the input words of an encoder. The purpose of the Attention mechanism is to improve the accuracy of prediction by focusing on the part of the input words associated with the predicted word. Vaswani et al. proposed a model called Transformer consisting of an encoder and a decoder using only the attention mechanism [8].

Bidirectional Encoder Representations from Transformers (BERT) [9] is a language representation model developed by Google based on Transformer. BERT is pre-trained with a large amount of Wikipedia and BookCorpus data using unsupervised learning. At the time of the publication, BERT achieved the state-of-the-art from 11 NLP tasks, including Question Answering tests [9].

Sentence-BERT is a model proposed by Reimers and Gurevych [10]. The existing BERT model is a word-based language representations model, and each word that passes the BERT has a dimension of 512. Tasks such as similarity comparison, clustering, and information retrieval take too much computational cost. To solve this problem, Sentence-BERT performs word-embedding in sentence or higher units.

2.2 Paragraphs retrieval

A QA system typically consists of three steps: 1) Question processing, 2) Document and Passage Retrieval, 3) Answer extraction [11]. A paragraph has one subtopic and is a unit that can be clearly classified within the whole text. Therefore, in this study, retrieval is performed on a paragraph basis. In particular, paragraphs retrieval is performed using sentence-BERT and the results are presented. These results show that sentence-BERT can be used in a QA system for proactive disaster management. Figure 1 shows the framework of this study.

2.3 Tropical cyclone reports

As mentioned above, this study uses public data to ensure data reliability, focusing on the tropical cyclones. To this end, Tropical cyclone reports provided by the National Hurricane Center (nhc.noaa.gov) are used. 12 reports, including the Hurricane Harvey report, and 404 paragraphs in them are used for this study.

3 Experiments and Results

To conduct the experiment, the 12 reports provided by the National Hurricane Center were reformatted into paragraphs and saved as Microsoft Excel file. Question embedding and paragraph embedding were performed using the Sentence-BERT. Using the cosine similarity, the similarity score between the question embedding and each paragraph embedding was calculated and the top five paragraphs with high scores were retrieved. Table 1 is the result of paragraphs retrieval on the Hurricane Harvey report. This study uses the same

Table 1. Results of paragraphs retrieval on the Hurricane Harvey report; underlined parts show the information regarding infrastructure damage.

Question	Top 5 most similar paragraphs in reports
What infrastructure has been damaged and What kind of damage has happened to the infrastructure?	Near the initial landfall location in Texas, wind damage was extreme in Aransas County, Nueces County, Refugio County and the eastern part of San Patricio County. Approximately 15,000 homes were destroyed in these areas, with another 25,000 damaged, and extensive tree damage was noted. … … This includes State Highway 361 which was inundated along the entire stretch of Mustang Island. The surge also damaged or destroyed many coastal structures in Port Aransas, Holiday Beach, Copano Village, Lamar, Seadrift, North Padre Island and Mustang Island. … … (Score: 0.2940)
	Major-to-record flooding occurred in Liberty County along the Trinity River with numerous roads inundated including FM 787. Many homes and subdivisions were either cut off or inundated, specifically north of the city of Liberty and in the Grenada Lakes Estates subdivision. … … High flows caused significant scouring of the state 105 (business) road; other roads were washed out as well, with bridge washouts or closures observed in many parts of the county. At least 1,000 homes were damaged in the county. (Score: 0.2636)
	In Brazoria County, the Brazos and San Bernard Rivers experienced record water levels, which caused widespread floods across the county. The hardest hit communities were in Baileys Prairie, Richard and West Columbia. Widespread major flooding on the Brazos River and Oyster Creek led to numerous roads and homes flooding in Columbia Lakes, Mallard Lakes, Great Lakes, Riverside Estates and the Bar X Ranch subdivisions, as well as homes on CR 39. Flooding damaged the bridge over Cow Creek at CR 25, making it impassable. … … Over 9,000 homes experienced flood damage from the storm. (Score: 0.2542)
	Major lowland flooding occurred in Matagorda County along the Tres Palacios River. Many roadways were under water, and homes in the El Dorado Country, Oak Grove, and Tres Palacios Oaks subdivisions flooded. Major flooding also occurred on the Colorado River at Bay City as levees were overtopped by 2 ft of water. High flows from the Colorado and Tres Palacios Rivers impacted river navigation for several weeks. Roughly 2,900 homes were damaged in the county. (Score: 0.2469)
	Very heavy rain in Tennessee caused minor damage. The Memphis area had flooding and over 19,000 customers lost power, with winds gusting to 52 kt at the airport. More significant flooding was reported in Robertson County (Nashville area), with 13 residents in the Chestnut Flats Apartment near the Nashville Fairgrounds evacuated due to the high water. In addition, the downtown Nashville Goodwill Industries reported major flooding, and about 10,000 customers were out of power at one time in that city. (Score: 0.2430)

uses the same question for all of the 12 reports. Examples of infrastructure damage information with in paragraphs are shown in Table 1.

Table 2 shows the results of paragraphs retrieval. all of the paragraphs retrieved from four reports (Harvey, Ingrid, Irma, and Issac) include information about infrastructure damage. In the other eight reports (Alex, Bill, Cindy, Dolly, Hermine, Imelda, Lee, and Michael), only a few paragraphs include infrastructure damage information. This is because the total number of paragraphs with infrastructure damage information in the eight reports was less than 5. In brief, the paragraphs retrieval has been performed well in all the reports.

Table 2. Results of paragraphs retrieval (the number of paragraphs showing infrastructure damage information / the preset number of paragraphs retrieval)

Tropical Cyclone	Result
Alex	1/5
Bill	2/5
Cindy	1/5
Dolly	1/5
Harvey	5/5
Hermine	1/5
Imelda	3/5
Ingrid	5/5
Irma	5/5
Isaac	5/5
Lee	2/5
Michael	2/5

Table 3 shows two paragraphs with different similarity scores. The one with higher score (0.2189) shows no information related to infrastructure damage, whereas the one with lower score shows infrastructure damage information. It seems to be because many words related to damage were used such as *"flood"* and *"death"*, even if there are no words related to infrastructure. This problem could be solved by fine-tuning for text classification using labeled data.

4 Conclusion

In this study, a paragraphs retrieval model, one of the steps in the QA system for proactive disaster management, was proposed and the experimental results were represented. The proposed model well retrieved the paragraphs including infrastructure damage information for all the 12 reports. The retrieved paragraphs can be used as an input in the next module (the answer extraction model) of the QA system. The proposed methods are expected to help disaster prevention and reduce the damage to the infrastructure.

Table 3. Retrieved paragraphs of the Hurricane Lee report; underlined parts show the information regarding infrastructure damage.

Paragraphs
Paragraph not including infrastructure damage information (Score: 0.2189)
Media reports indicate that flooding largely related to the remnants of Lee was responsible for at least 12 additional deaths in the eastern United States; seven people in Pennsylvania, four in Virginia, one in Maryland, and one in Georgia. Nearly all of these deaths occurred when individuals tried to cross flooded roadways in vehicles or were swept away in flood waters.
Paragraph including infrastructure damage information (Score: 0.1565)
Most of the damage from Lee was the result of storm surge or freshwater flooding. Storm surge flooding from Lake Pontchartrain <u>inundated more than 150 houses</u> in Jefferson and St. Tammany Parishes in Louisiana. Minor storm surge flooding was also reported outside the hurricane protection levees in St. Bernard and Orleans Parishes. Freshwater flooding was reported in low-lying areas of southeastern Louisiana and southern and central Mississippi. Several roads were inundated by floodwaters in Hancock, Jackson, and Harrison Counties Mississippi, while in Neshoba County in the central portion of the state, <u>35 roads were damaged with 5 of those completely washed out</u>.

Acknowledge

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (No. 2018R1A2B2008600) and the Ministry of Education (No. 2018R1A6A1A08025348).

References

[1] Field, C. B., V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, P. R. Mastrandrea, and L. L. White. "Intergovernmental panel on climate change. Climate change 2014: Impacts, adaptation, and vulnerability." In *Contribution of Working Group II to the Fifthe Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press Cambridge/New York, 2014.

[2] Wang Y. and Taylor J. E. "DUET: Data-driven approach based on latent Dirichlet allocation topic modeling." *Journal of Computing in Civil Engineering* 33.3: 04019023, 2019.

[3] Sit M. A., Koylu C., and Demir I. Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma. *International Journal of Digital Earth*, 12(11), 1205-1229, 2019.

[4] Ragini J. R., Anand P. R., and Bhaskar V. Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management*, 42, 13-24, 2018.

[5] Chan H. Y., and Tsai M. H. Question-answering dialogue system for emergency operations. *International Journal of Disaster Risk Reduction*, 41, 101313, 2019.

[6] Tsai M. H., Chen J. Y., and Kang S. C. Ask Diana: A Keyword-Based Chatbot System for Water-Related Disaster Management. *Water*, 11(2), 234, 2019.

[7] Sutskever I., Vinyals O., and Le Q. V. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104-3112, Montreal, Canada, 2014.

[8] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L. and Polosukhin, I. Attention is all you need. In *Advances in neural information processing systems*, pages 5998-6008, Long Beach, CA, USA, 2017.

[9] Devlin J., Chang M. W., Lee K., and Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.

[10] Reimers N., and Gurevych I. Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084, 2019.

[11] Jurafsky D. Speech & language processing. Pearson Education India, 2000.