Disaster Tweet Detection using Bidirectional Encoder Representations from transformers (BERT)

Samyak Jain (202IS022)

Department of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal, Mangalore, India
samyakjain.202is022@nitk.edu.in

1 Motivation

In the past several years, social media like Twitter have played a prominent role as an important communication channel in times of emergency. The ubiquitous of smartphones enables people to announce an emergency they are observing in real-time. Therefore, there are growing appeals for agencies or governments to monitor and understand people's Twitter programmatically in real-time to make a prompt reaction like disaster relief and medical help.

However, this seems to be a common problem in clarifying people's words for a program. It's not always clear whether a person's words are actually announcing a disaster or not. One approach to solve this problem is the use of the state-of-art model Bidirectional Encoder Representation from Transformers.

There are many difficulties in disaster detection. The first one is how to evaluate the effectiveness of information. It's challenging to judge whether the event is a disaster since every word in a tweet may come from the arbitrary daily upload. Therefore, the main difficulty in disaster detection is how to segregate the meaningless part from the meaningful part. Many machine learning methods can recognize the word but not judge its context as humans do. On the other hand, there is a phenomenon in life, misunderstanding happened during information transmission. The information transmission is divided into two parts. One is sender's intention, the other is receiver's comprehending. During the processing, the machine plays a receiver's role. It may make a mistake as humans do, simply recognizing a statement without context.

2 Problem definition

In times of crisis, Twitter has become a big communication platform. Deep learning, as one of the most currently remarkable machine learning techniques, has achieved great success in many applications such as image analysis, speech recognition, and text understanding. This work aims to make use of the bidirectional encoder

representation from transformers (BERT) model to a particular problem in our real life, detecting disaster on Twitter. Specifically, we first validate the practicability of BERT in disaster detection, and then improve the model practicability by extracting keyword information to classify the text. The result of this research will be helpful in monitoring and tracking social media content and in discovering how people's descriptions of disasters are like.
As a result, more organizations are interested in tracking Twitter programmatically. The aim of this project is to detect the tweets which contains relevant information regarding the real-time emergencies, sentiments of the people affected during and post catastrophe. However, the detection of such tweets are often difficult due to the tweets' structure's uncertainty.

- ➢ Sentiment analysis of people’s reaction through tweets
- ➢ Training on disaster dataset
- ➢ Correct classification of tweets with ambiguous content

3 Literature survey

To provide a summary of the type of literature, we present a study that focuses on papers based on in-depth reading . Analysis of social media data in transforming disasters has been carried out in several studies [3] [4]. Current studies have used geotagged tweets to measure the degree of social media involvement in areas impacted by disaster to determine whether vulnerable groups remain silent on social media to explore relevant factors.

The ref [5] suggests that social media consumers rely more on rescue and donation-based knowledge using a machine-learning tweet model. However, differences in subjects are consistent across various fields and differing levels of focus. The study of sentiment in online review sites has been thoroughly explored to provide consumers with synthetic views on various product aspects. However, little work has been done to define the polarity of users' emotions during catastrophic incidents. Identification of such feelings on the websites may allow emergency respondents to understand the complexities of their network, such as users' key concerns, panics, and the emotional effects of interactions among participants [1].

On research conducted by [6], they trained sentiment classes in categorizing messages on a concentrated basis and then discussed the demographic variations in these classifiers. They conclude that social media research complements conventional approaches for understanding public awareness of an imminent catastrophe in real-time. In [7], they propose that Twitter feeds data analytics to enhance emergency planning and rescue efforts and services rendered in unusual circumstances. They follow a different approach, which focuses on the thoughts, fears, and opinions of users shared in emergency tweets and evaluates these feelings and expectations within an incident group to ensure that emergency response staff and local authorities have sufficient input. They use analysis and detection methods to store, identify and infer users' spatiotemporal feelings. BERT has become state of the art in NLP. Since its inception, BERT has been widely used and experienced [2]. BERT's main technological advancement is applying transformer bidirectional training, a standard language model, to language modeling. Compared to previous attempts, a text sequence was evaluated from left to right or combined from left to right and from right to left. To solve the fine-grained challenge of classification uses a promising profound learning model called BERT [8].

Experiments demonstrate that their model succeeds other standard models without a sophisticated architecture for this mission. It also explains how efficient transfer learning is in manipulating natural languages. Whereas in [9], BERT was applied to a series of tweets linked to the Jakarta Flood in 2020. In the Jakarta flood tragedy of early 2020, which became a trending issue on Twitter, they tore up the tweet details. From there, the goal is to find specific tweets that can provide valuable information on emergency response to disaster management. Experimental studies have shown positive results. However, according to them, the data collection's consistency substantially affects the efficiency of the system

4 Work Done

4.1. Data Collection

In this study, we used the dataset from Kaggle. The dataset could be downloaded from https://www.kaggle.com/c/nlp-getting-started/data. We have 10873 tweets, and we estimate whether or not a particular tweet is about a specific catastrophe.In fig 1 From 10,873 data, 3054 were not real disasters.

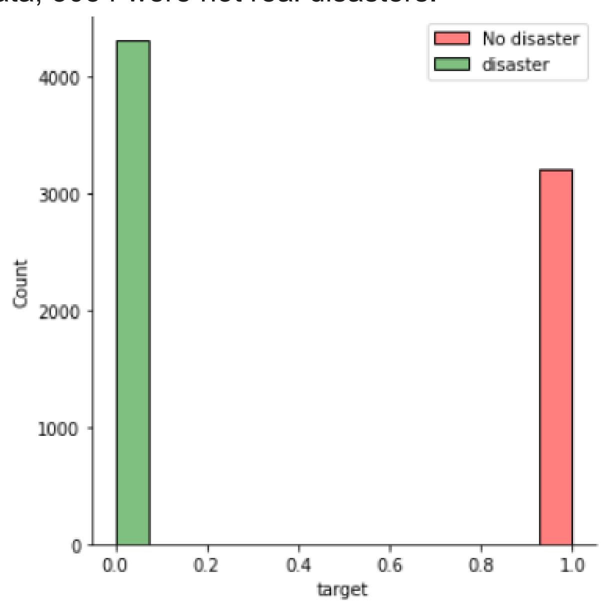

Fig 1 : Dataset Count

Fig 2 described the features and specifications.

- Id - tweet identifier text - text of the tweet
- keyword - A relevant keyword in the tweet (maybe NaN)

- location - location the tweet was sent from (maybe NaN)
- Text - Tweet data
- target - Output that tells if a tweet is a real disaster (1) or not (0)

4.2. Implementation

We have use Pandas to read data, ekphrasis for text pre-processing. Seaborn for data visualization, all are running on Kaggle Notebook. By using pandas in Figure 2, we can see an example of data from the dataset

	id	keyword	location	text	target
0	1	NaN	NaN	Our Deeds are the Reason of this #earthquake M...	1
1	4	NaN	NaN	Forest fire near La Ronge Sask. Canada	1
2	5	NaN	NaN	All residents asked to 'shelter in place' are ...	1
3	6	NaN	NaN	13,000 people receive #wildfires evacuation or...	1
4	7	NaN	NaN	Just got sent this photo from Ruby #Alaska as ...	1

Fig 2: Sample Dataset

First, we do text cleaning of the raw data to make our dataset more applicable. Then we extract the keyword position of each sentence, which has the 'keyword' attribution. Finally, with the proposed data as input, we implement BERT with our proposed Keyword Position Information to classify the text.

A. Data Preprocessing

i. ***Text Cleaning***. When the original data was scrapped, it contained a lot of useless information, such as punctuations. To solve the problem, first, we do the word segmentation, then we implement nltk to get rid of the stopwords. After that, we will remove all the letters and numbers of the word to check if there are illegal characters like '\x89' in the word, if so, we will remove the corresponded word.The cleaning stage is done with library string and further processed the tweets by making them removing punctuation, email, URLs, HTML tags, special chars, unknown words, and duplicates text

ii. ***Keyword Position Extraction***. For each sentence that has a 'keyword' attribute, we locate the position of the keyword and store the position into the 'position' column. As for our crawled data, in order to analyze the coronavirus, for those sentences which contains 'COVID' or 'corona', we label the keyword as well as the position information. The data type of 'position' is float.

B. Implementation of BERT with Keyword Position Information

BERT refers to Bidirectional Encoder Representation from Transformers, which uses the encoder of a bidirectional Transformer due to the normal encoder cannot get the information to be predicted. The main contribution of

BERT is the pre-train method, with the usage of Masked LM and Next Sentence Prediction to capture word and sentence-level representations.

i. Transformer

The transformer is the basic structure BERT has implemented to encoding. It is first proposed by Ashish Vaswani et al. in 2017 [10]. Different from the RNN based on the Seq2Seq model framework, attention mechanism is used instead of RNN to build the whole model framework in there work. Meanwhile, they proposed a Multi-headed attention mechanism, which is widely applied in encoder and decoder.

ii. Masked Language Model and Sentence-level Representation

BERT uses two unsupervised tasks to pre-train. The first step of pre-training is to construct the language model. The masked language model is a deep bidirectional model rather than a left-to-right model or the shallow concatenation of a left-to-right and a right-to-left model. 15% of all WordPiece tokens in each sequence are randomly masked, and the final loss function (cross entropy) only uses the tokens which are masked.

Then is the Next Sentence Prediction (NSP) task. In many tasks, based only on encoding, which is just to learn a bunch of token-level features, is not enough. Some sentence- level patterns need to be captured to achieve tasks like Question Answering (QA) and Natural Language Inference (NLI), which require sentence-level representation. Therefore, BERT introduces the other significant but lightweight task to learn this pattern.

BERT is a sentence-level language model, unlike ELMO [11] whose layers need to be weighted for global pooling when concatenating with downstream specific NLP tasks, BERT can directly obtain the unique vector representation of a whole sentence. As shown in Figure 2, it adds a special mark [CLS] to each input, then uses a transformer to conduct deep encoding on [CLS]. The transformer encodes global information into each position regardless of space and distance, and the highest hidden layer of [CLS] is directly concatenated with the output layer of softmax as the representation of sentence/sentence pair. Therefore it can learn the upper features of the entire input as a 'checkpoint' on the gradient backpropagation path.

iii. Fine-tuning

Fine-tuning is straightforward since the self-attention mechanism in the transformer allows BERT to model many downstream tasks. As for sequence-level classification tasks, BERT directly takes the first [CLS] token of the final hidden state then adds a weight layer, then applied with softmax to predict the label probability:

iv. Keyword Position Information

Instead of the text of tweets, we also use the Keyword Position Information in our input. The keyword position means the position of the keyword in the corresponding tweet. For instance, we set the keyword 'corona', then the keyword position of the tweet 'Company Creates App to Track People Infected with coronavirus' is 9 (index begins with 0). We concatenate the keyword position with the word embeddings

5 References

[1] Neppalli V K, Caragea C, Squicciarini A, Tapia A and Stehle S 2017 Sentiment analysis during Hurricane Sandy in emergency response International journal of disaster risk reduction 21 213-222

[2] Devlin J, Chang M. W, Lee K, and Toutanova K BERT: Pre-training of deep bidirectional transformers for language understanding 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference

[3] Zahera H M, Elgendy I A, Jalota R and Sherif M A 2019 Fine-tuned BERT Model for MultiLabel Tweets Classification TREC 1-7

[4] Frank M R, Mitchell L, Dodds P S and Danforth C M 2013 Happiness and the patterns of life: A study of geolocated tweets Scientific reports 3 (1) 1-9

[5] Fan C, Esparza M, Dargin J, Wu F, Oztekin B and Mostafavi A 2020 Spatial biases in crowdsourced data: Social media content attention concentrates on populous areas in disasters Computers, Environment and Urban Systems 83 101514

[6] Mandel B, Culotta A, Boulahanis J, Stark D, Lewis B and Rodrigue J 2012 A demographic analysis of online sentiment during hurricane irene Proceedings of the second workshop on language in social media 27-36

[7] Singh N, Roy N and Gangopadhyay A 2018 Analyzing the sentiment of the crowd for improving the emergency response services 2018 IEEE International Conference on Smart Computing 1-8

[8] Munikar M, Shakya and S and Shrestha A 2019 Fine-grained sentiment classification using BERT 2019 Artificial Intelligence for Transforming Business and Society 1 1-5

[9] Maharani W 2020 Sentiment Analysis during Jakarta Flood for Emergency Responses and Situational Awareness in Disaster Management using BERT 8th International Conference on Information and Communication Technology 1-5

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin (2017). Attention Is All You Need. arXiv:1706.03762.

[11] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018a. Deep contextualized word representations. In NAACL