

CS482/682 Final Project Report Group XX

Fine-tuning BERT for Disaster Response

Zihao Zhao (zzhao71), Yuxuan Li(yli602), Jingkai Guo(jguo90), Tianyu Shi(tshi22)

1 Introduction

Problem Statement In case of emergencies, rapid responses are needed to address the victim's help. The computer can be used as a machine to quickly classify important messages. Previously, SVM(support vector machine) and Naive Bayes were proposed to help text classification. We want to fine-tune BERT(Bidirectional Encoder Representations from Transformers) to increase the classification accuracy.

There are several challenges in classifying the data.

- (i) The text message only contains a few words, so it may need some additional information sometimes.
- (ii) Some texts may belong to several categories. For example, an emergency message may contain the need for food and medicine.
- (iii) The training set has a limited size.
- (iv) The text message may have some mistakes. Also, some of the messages may contain French words.

Related Work There are some works tackling the classification problem of disaster responses. Some techniques include the Naive Bayes algorithm (Aris Ordonez, et al. 2018), and support vector machine (Cornelia Caragea et al. 2011). However, there is limited effort focusing on addressing this problem by using neural networks. In the meantime, there is a bunch of previous work showing the capability of text classification of BERT. (Chi Sun et al. 2019). Therefore, we have a solid foundation on this project and it is promising that we could have better performance than methods other than neural networks.

2 Methods

Dataset The dataset we will be using is disaster response messages from Huggingface. It contains 30,000 messages drawn from 100s of different disasters, 21k entries will be used for training. The data has been encoded with 36 different categories related to disaster response with 40 class labels for intent and content. Each data might have one or more corresponding true labels. The dataset is a multilingual dataset which has the messages in the original language and also its translated English form. Since the messages are already translated into English and the quantity is sufficient, the preprocessing steps including text cleaning are required.

Proposed approach We will first have the preprocessing steps to have the cleaned and normalized messages. Then, the fine-tuned BERT model is applied to multi-label classification tasks. We may build several BERT stacked layers on top of the BERT outputs to fine-tune BERT into the multi-label classification of messages. Then, we add an extra dense layer with a sigmoid activation function. Further, we used two loss functions (binary cross-entropy, and focal loss) to minimize the errors during the training of our models. The multiple labels may be gained by the probability of classes over the threshold.

The model is evaluated by its accuracy in the test set provided by the dataset. This project aims to surpass the performance of all other methods and achieve the state of the arts.

3 Reference

SMS Classification Method for Disaster Response Using Naïve Bayes Algorithm. (n.d.). Ieeexplore.ieee.org. Retrieved November 9, 2023

How to Fine-Tune BERT for Text Classification?. (n.d.). Scholar.google.com. Retrieved November 9, 2023

Caragea, C., Kim, H.-W., Yen, J. (n.d.). Classifying Text Messages for Emergency Response. Retrieved November 9, 2023