Classify positive and negative emotions with multiple machine learning methods

Yihang Yang, CSE 6520

In past few years, chatbots like ChatGPT and Gemini based on Large Language Model (LLM) have grown in incredible popularity and been widely applied in numerous fields. As chatbots interpret the input from users, it is important for them to identify whether the motion of content they receive is positive and negative, thus deciding to reply with a congratulatory or encouraging tone. Before training and adjusting the parameters of those chatbots for processing emotional tasks, a great amount of input data with accurate emotional labels (positive/negative) should be prepared to ensure the training effect.

A renowned dataset for sentiment analysis is Stanford Sentiment Treebank (SST-2). SST-2 is a corpus based on 11,855 sentences of movie reviews, and each sample belongs to one of 5 classes: very negative, negative, neutral, positive and very positive. Many useful models of Recursive Neural Networks (RNN) and DistilBERT have been trained based on SST-2 to perform emotional classification. Though these models can reach even over 85% accuracy (Socher et. al, 2017), GPUs and hardware of high performance are qualified to do the task. Therefore, I wonder if simpler learning techniques can be applied in the emotional classification task to get good results with less time and money spent.

In this project, I intend to extract feature from text samples including SST-2 and some dialogues that take place in chatbots with NLP methods, such as tokenization, stop word removal and stemming. The techniques will be similar to those applied in scam detection mentioned by Reddy et al. (2023) Afterwards, I will train SST-2 dataset based on several traditional machine learning methods including Naïve Bayes, Support Vector Machine and Random Forest. Then I will also employ the model “distilbert-base-uncased-finetuned-sst-2-english” developed by Sanh et al. (2020) to classify the same dataset as control. Finally, I will compare not only the accuracy of classification results but also the running time between each training algorithm and with the DistilBERT model, drawing a conclusion whether traditional NLP can work as a substitute for more complex models to classify large number of unlabeled texts with acceptable accuracy loss and obvious elapsed time decrement.

References

Socher, R. et al. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. https://www.kaggle.com/datasets/atulanandjha/stanford-sentiment-treebank-v2-sst2, 2017.

Reddy, A. et al. Email Spam Detection Using Machine Learning. http://sifisheriessciences.com/index.php/journal/article/view/1249, 2023.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2020.