Yuanming Mao – ym18
Jie Zheng – jzheng5
Albert Yeh – ayeh2

CS410 – Progress Report

Our team decided on the Classification competition for the class project. We have beaten the baseline with an F1 Score over 74%. Although, the team each took different approaches to solving the classification problem, each member used some variation of BERT as we stated in our project proposal. The BERT model is significantly more effective in classification problems, using transformers over previous CNN/LTSM type models. Fortunately, there were a lot of resources online, and each of us were able to learn the ins and outs of BERT and transformers, as well as follow examples to solve this classification problem.

The entire group had zero experience in sentiment analysis and text classification going into this project so we were able to get an idea of how to start this competition by reading a few papers that were specifically about Twitter Sarcasm Sentiment Classification. [1][2] Each of these papers were examples that were working with the same exact dataset, so there was a lot of useful discussions on how to pre-process the data, and the test results of which transformer models produced the best F1 scores for classification. After reading the papers, we were able to create a path forward for our team to attempt to work together in solving the classification competition.

After our second meeting we were able to get to 68% for the F1 score, but not enough to pass the competition. However, after our most recent meeting, we were able to solve and pass the baseline. We tried different variations of BERT and Roberta, and ended up just using BERT large uncased. We also tried many different pre-processors both individually and using different libraries (such as ekphrasis), but ultimately settled on tweet-preprocess library. We found most of our inspiration from a multi-classification guide[3]. We played around with the parameters and were able to finally pass the baseline after many different various attempts. We have not only learned how BERT works, but all of us have learned the meaning of parameters and how to solve a problem like text classification using state of the art algorithms.

Lastly, the team will attempt to increase our F1 score by continuing to tweak the parameters, but the majority of the work has been done. There are a few things we can test out, such as single sentence (response w/ context), or combine with LSTM, or other optimizers.

[1] https://www.aclweb.org/anthology/P11-2102.pdf
[2] https://arxiv.org/pdf/1911.10401.pdf
[3] https://analyticsindiamag.com/step-by-step-guide-to-implement-multi-class-classification-with-bert-tensorflow