PROJECT PROGRESS REPORT

CS 410 Text Information Systems Fall 2020

Completed Tasks

- Became familiar with BERT and its usage for text classification problems.
- Worked on cleaning of tweets by removing urls , emojis, punctuations, special characters etc. which may not help in classification.
- Implemented an approach using BERT in google collab using Tensor Flow.
- Fine-tuned the model and was able to beat the baseline score.

Pending Tasks

- Source code refactoring, optimization and final submission.
- Documentation.
- Tutorial presentation.

Challenges/Issues

No current issues. Initially, faced issues with setup and computing power with my current laptop to use tensor flow for running BERT. So, started working on google collab for coding and running and was successful in making submissions beating baseline score.