

CS410 Project Proposal - SoCal Classifiers

1. The names and netIDs of our team members are:
Yuanming Mao - ym18@illinois.edu,
Jie Zheng - jzheng5@illinois.edu,
Albert Yeh- ayeh2@illinois.edu.
Albert Yeh is the captain.
2. We are choosing the text classification competition for our project.
3. The team is committed to learning state of the art classification for the competition. We are researching the most recent state of the art technologies which rely on the idea of transformers which is introduced in 2017 in the paper Attention is all you need. <https://arxiv.org/pdf/1706.03762.pdf>. The idea of transformers supersedes the previous models of RNN, CNN, LSTM, and GRU. Then Google released the paper about BERT, <https://arxiv.org/pdf/1810.04805.pdf>. This was built on transformers which is a much larger improvement on current methods. No one in our group has any relevant prior experience with any of these models or other classifiers. We understand that we will also learn about text preprocessing using technologies such as word2vec and avgword2vec.
4. Because we will primarily be focusing on BERT, we will be programming in python which is the language it is written in.