

# CS 410 Text Information System Final Project

## Text Classification Competition

Zhenchen Yu (zy23)

November 24<sup>th</sup>, 2020

1. Which tasks have been completed?

I have completed the task 1, 2 and 3 in the list below.

In detail, I have implemented SVM, Random Forest, Logistic Regress, CNN, RNN, BERT base and BERT large models to predict the text results.

2. Which tasks are pending?

I still need to complete the documentation and presentation.

3. Are you facing any challenges?

No big challenges. Currently I am trying to fine tune the parameters for BERT large model but it is difficult to further improve the F1 score and it is time consuming.

### Project Plan Timeline

Timeline	Task
10/18/2020 – 10/24/2020	1. Form the team and submit proposal (10%)
10/25/2020 – 11/07/2020	2. Implement SVM / Random Forest / Logistic Regression / CNN / RNN
11/08/2020 – 11/21/2020	3. Implement BERT (or other cutting-edge deep learning models)
11/22/2020 – 11/28/2020	4. Submit the progress report (5%), source code (45%), documentation (20%) and the presentation (20%)
11/29/2020 – 12/14/2020	Buffer time