

CS 410 Text Information System Final Project

Text Classification Competition

Zhenchen Yu (zy23)

October 24th, 2020

1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.

*I'm planning to complete this competition by myself.
My name is Zhenchen Yu. My NetID is zy23.*

2. Which competition do you plan to join?

Text Classification Competition

3. If you choose the classification competition, are you prepared to learn state-of-the-art neural network classifiers? Name some neural classifiers and deep learning frameworks that you may have heard of. Describe any relevant prior experience with such methods

First, I want to try to use some statistical learning models such as SVM, Random Forest and Logistic Regression. Then I will try to use some deep learning models such as CNN and RNN. Finally, I have heard about some cutting-edge deep learning models such as BERT. I would like to take this opportunity to learn BERT.

Prior to this course, I have completed the CS498 Applied Machine Learning and CS598 Practical Statistical Learning. In these courses, I implemented SVM, Random Forest, Logistic Regression and CNN in the assignments or projects.

4. Which programming language do you plan to use?
Python

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