PROJECT PROPOSAL

For the final term project, I am doing Sentiment Analysis using different kinds of neural networks to get the best accuracy possible. I chose this problem so that I could learn how to use neural networks effectively in Natural Language Processing. Also, in the future, I would like to integrate sentiment analysis with Recurring Neural Networks to predict the stock market more accurately.

The dataset being used for this project is Amazon Reviews for Sentiment Analysis taken from Kaggle website. The dataset is a plain text file of size 1.6GB. This dataset is large enough to be used for deep learning.

From the deep learning section, I will be using Multi-Layer Perceptron network, Convolution network, Recurring Neural Network, and/or pre-trained models. I will be training all of these models individually and/or in combination to find the best network that has the best accuracy possible. I would choose the results of standard networks as baseline and then aim to customize the network. To implement these networks, I will be using PyTorch framework because currently this framework is the most widely used framework. Also, the training process is more comprehensible since it allows breakpoints to be used within the training process. Performance of the networks mentioned above will be measured using cross entropy.

Currently, I am relying on the Neural Network Design book and Google for information. I am also relying on the example codes provided by the professor. I will be mentioning all the referenced sources in the final report.

Finally, the rough schedule for this project would be, a week for data pre-processing, a week for establishing baselines for different kinds of networks. The last week would be for customizing the networks for better results.