

# CS410 – Fall 2021 – Project Proposal

## Progress Report

**Team Name:** Team21

**Net Ids:**

Subudhi, Pradosh Kumar - [pradosh2](#)

Veera Sikku, Siva Prakasini - spv4

Dunor, Aaron - adunor2

**Captain:**

Pradosh

This report provides an overview of the progress on the Project that we are working on and addresses the review comments.

**Overall Project Plan:**

In addressing our email sentiment analysis project, we decomposed the project into the following smaller manageable tasks:

Emails → Sentiment Annotation → Text Cleaning → Word Embedding → Model Training → Model Evaluation → Model Deployment

**Challenges:**

Finding a good email dataset set has been most challenging. We finally found [customer-service-sentiment-analysis/data](#) dataset which contains some five hundred fifty thousand emails. The dataset needs a lot of cleaning to make it usable for before performing the analysis.

**Tasks Completed so Far:**

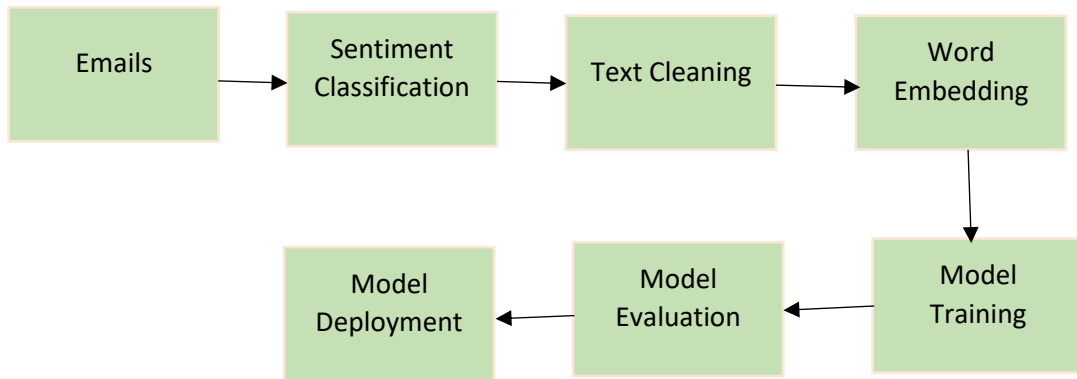
Currently we have sourced the emails and processed them into a new data frame containing the text and the sentiment score. We used the “SentimentIntensityAnalyser” from NLTK library to assign sentiment categories to each email. The categories include Negative, Neutral and Positive.

**Next Stage:**

We are working on word embedding and developing the final model.

**Review Feedback addressed:**

- High level component interaction plan through a diagram



- Can you clarify the novelty of this proposal?

Though this is a project that is like many out there, the machine learning model created from this project could be deployed in many use cases. Unlike a tutorial on the internet that ends when the model is produced, we can release our model for reuse in applications like email clients, customer service ticketing systems and even web sites with some enhancements to the model.

- Interesting project idea! Are you planning on selecting key words from each customer's complaints to study the priority?

We are not building the model on select keywords. We will utilize NLTK SentimentIntensityAnalyser functionality to assign sentiments.

- You are expected to have a preliminary model in the progress report:

The Preliminary model can be found at the following location –

<https://github.com/Pradoshks/CourseProject/tree/main/Model>