**Progress Report – CS410**

1. Which tasks have been completed?

I have completed the extraction of data from Amazon reviews. I have used the following link to scrape the data

https://www.scrapehero.com/how-to-scrape-amazon-product-reviews/

However, due to amazon’s anti-scraping measures, I was unable to do so. Thus, I have instead used the data in the link mentioned below https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products

1. Which tasks are pending?

I have the following pending task:

Split the data into train-test and fit Naïve Baiyes

Check bias trade off and model performance on the test data

Develop additional models and check performance

Check performance using precision, recall

1. Are you facing any challenges?

I tried to scrape from Amazon website but it was preventing me from scraping them. Captcha issues and IP blocks were a major roadblock.