**IS 695 GROUP**

**AMAZON REVIEW ANALYSIS**

Team Members:

Priyanka Malla

Ansuya Patel

**Project Overview**

Is it possible to predict the star rating of an Amazon review given the text of that review? I want to try to answer that question using a collection of 400,000 Amazon reviews of unlocked mobile phones as my dataset. We think it is interesting to see if the sentiment of a review can be decided from the text of the review itself.

The problem of determining the sentiment of text is not new. Research was done to try to extract sentiment from movie reviews. They tried to predict whether the review was positive or negative based on the text. They found that one of the main difficulties with sentiment extraction is that it requires more understanding of the text than traditional topic-based categorization. Topic-based categorization relies on finding keywords in text; however sentiment can be expressed more subtly.

The online store Amazon.com gives its users the ability review products in the form of a written review and a star rating between 1-5 stars. In this report, we try to predict the star rating given the written review.

To build our model we must complete the following tasks:

1. Download and pre-process Amazon.com reviews of unlocked mobile phones.
2. Extract features from the reviews.
3. Train models to predict ratings.
4. Refine the models.
5. Compare/contrast final models.
6. Discuss further possible improvements.