

Progress Report

At this time, we have scraped some reviews from Amazon and converted them into formats which are better for processing. We have compared different Natural Languages Processing packages used frequently in Python such as NLTK and SpaCy. After comparing, we decided to use NLTK to do the POS tagging in our case. Also, we learned about sentiment analyzer in NLTK package, and might be going to use it to perform sentiment analysis. Moreover, we had several basic ideas regarding the UI/UX design for the front-end part.

Next week we will work on tokenization and stemming, as well as performing sentiment analysis and topic analysis on the reviews. We will visualize our user interface design and start to implement necessary elements. In the next two weeks, we will build the front-end component and make it as an extension to the browser.

So far our work progresses with hardly any challenge since the scraping technique is taught in one of our previous MPs. Tokenization and analysis may also be simple to implement, but front-end implementation may be a tough task to do since we have not made browser extensions before. Also, we might need to figure out what python packages for sentiment analysis we should use in our case because different packages have different performance.