**Project Progress Meeting with Project Supervisor, Professor Lo**

Data: 13 February 2020

Time: 0930hrs – 1015hrs

Meeting Agenda: Update project supervisor on project progress for Week 6

Items Discussed:

1. Project Progress Update
2. Suggestions/Feedback
3. Mid Terms

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_­­­­­

1. Project Progress Update

* Data Cleaning
  + Split reviews into smaller sentence
  + Split by conjunction
  + Get the textual form of an emoji
* Visualization and Sketches
  + Gathered feedback from Prof regarding first version
* Model
  + Currently doing model implementation for Topic Analysis and Sentiment Analysis

2. Suggestions/Feedbacks

* Visualization
  + Dashboard Page 1 (Overview Page)
    - Must have a title (what is it and what we are looking for?)
    - Percentage of positive and negative review (please add percentage sign behind make it clearer)
    - Do a comparison as compared to previous month as stakeholders would probably find this important
      * They would want to see the changes as compared to previous month in one go. Suggestion was to use arrows/colors to denote whether something is getting better or worse
    - Consider filters
    - For the funnel at the bottom left hand corner, try to get the opinion of Loreal. What is it they wanted to see from there?
    - Heat map is providing very useful insights – Friday night tend to have high amount of feedbacks coming in
  + Dashboard Page 2 (Correlations between two variables)
    - Quite interesting, but there’s overlapping.
      * Either have no fill to make it better or another standard way to present is using a grouped bar chart
  + Dashboard Page 3 (Positive vs Negative)
    - Reconsiders the placing of the important graphs
    - The word “Delivery” appears on both positive and negative, and takes up a lot of importance
    - Bring out the list of reviews
    - Remove all the stopwords
    - Check with Loreal if they would like the positive/negative word cloud together or separated?
* Model
  + Emoji
    - The emoji is a more correct representation than the text form
    - Fish out for those with emoji first and see if you can classify them with your own eyes first
  + Qn: What is a good accuracy for sentiment analysis
    - > 80%
  + Qn: How do you further improve your accuracy?
    - As we are currently using count vectorization, it might not be able to give very high accuracy
      * Explore word2vec then doc2vec. Word2vec considers the semantics of the reviews. One example is berg. Doc2vec have been doing well with short text as well, can try it out
  + Qn: How to choose baseline?
    - Just choose Naïve Bayes
    - Do a few model and see which one to choose
    - You must improve from baseline and the process of improvement is important

3. Midterms

* Show that you have already showed your stakeholders the prototype (User feedback)
  + What is the feedback from the user?
  + What are the changes you have made?
* Show the overall picture
  + What will you be showing at the end?
  + How does it link to a dashboard?
* What have you already done?
* For the bug fixes/change management, you can show a ghant chart since you already have. Show that whenever there is a discrepancy or delay, why is there a delay? What was the error? What did we do to solve it?
* Technical report (architectural diagram)
* Efforts for model (what have you done so far? Challenges? Solutions?