Progress Report

1. We have implemented the 5.1 mentioned in the paper, which includes importing data set downloaded from the website, converting word into lower cases, removing punctuations, stop words, and stemming each word to its root.
2. Our remaining task is to implement the rest of the functions from the paper, including 5.2 – 5.4, and get the result from our program by the end of this project.
3. The formulas mentioned in the paper are a bit confusing to us, and we will make sure we can understand the math behind it so that we can finish this project.