Name: Chloe Acheampong

Reflection Paper on final project

When we were asked to start the final project for the Data mining class, I was not sure of the topic to work on. However, I knew I wanted to work on sentiment analysis, and started my research for a topic around that area. I finally settled on “Sentiment Analysis on Vodafone tweets” in quest to assess the performance of Vodafone in the market.

Researching on the topic was the easiest task. Obtaining data to explored, on the other hand, was quite cumbersome. For about a week or so, I spent time researching on the internet and hoped to find Twitter’s API to help me download tweet data. But my predicament came to an end when I discovered I could easily do that in R studio. In R studio, I downloaded some libraries like twitterR, wordcloud, plyr, among others, but the sentiment library was difficult to install. I battled with installing this library for about a day until I finally made a break through. Now, I just had get some tokens from Twitter to use for the tweet download.

Initially, I thought using R would be difficult because I was not very familiar with it. But after sometime, the functions and features I discovered in R were mind-blowing. For example, without creating my own functions, I could in one line remove “stopwords”, easily clean my data using the “gsub” function in R. The most amazing and surprising part was when I classified the tweets with just a line of code.

Also, one major lesson from this project was the fact that it is possible to sometimes to manipulate data in a wrong way and get false results, so it imperative that data mining projects should be done with clarity on the techniques and their possible outcomes, as well as being careful on the kind of data being fed into the algorithm or technique.

Overall, working on this project made me feel I would be useful in the real world since every company or business has some data of some sort, which might be useful if mined. And I am really glad I took this course because it gave me the opportunity to learn how to manipulate data, and find meanings and insights using different techniques☺.