ISABEL DZIFA ATTU

DATA MINNING FINAL REFLECTION PAPER

24TH April, 2015

Admittedly, my sentiments before and during the early part of the semester about this course was nothing short of skepticism about whether or not I would enjoy the course. In fact, even before I began the final project, I didn’t see myself sailing through the task easily. In as much as this had a lot to do the fact that I had not fully grasped the various concepts thought in class yet, what made this project a seemingly herculean task for me was also because I could not find data that I was interested in, not to think of the process to use in mining the obtained data; I was frustrated!!

When I eventually chanced upon UN.ORG and found my data after days of searching, the feeling was heavenly. The task of cleaning and preparing the data for use also another hurdle I eventually overcame by using excel and R. In terms of exploratory data analysis, I opted for R and K-means clustering because my data were unrelated and unsupervised and thus the need to make sense of it using a method like K-means clustering which seemed appropriate.

Coding was not so much of an issue because there were libraries in R that made generating clusters and finding the modularity between them quite easy. If there was any issue at all with coding it was probably finding a pleasant and lucid way to generate and display graphs using the clusters I had obtained. That too was eventually dealt with after I realized there were functions that allowed for k-means clusters to be colorfully plotted in R.

I’m happy, no, elated that I was finally able to get my project done and that I took this course at all, not because I’m exceptionally great at it, but because I cannot over-emphasize its importance enough. While struggling my way through the project I realized data mining not only enhanced my critical thinking abilities in making me find the right algorithms that suited the kind of data I had. I also, in interpreting results from data mining techniques, learnt to make sense of large data and make predictions in a way that can help me in future to do a lot of cool stuff, regardless of whatever field I find myself in.