Beatrice Garcia

April 24, 2017

Module 12

Homework

I started off querying data using the Facebook Graph API. I queried for public pages with the word “Apple.” Then, using the top three public pages, I queried each page for about 25 posts, comments, and reactions (Like, Love, etc.) Using this data, I created a graph using the analytic engine, Socrates, to connect the page to the posts and the posts to the comments/likes.

Using the graph, I created an example of a “bad visualization”, in general, the bad visualization looks like a spider web. It has no structure. There aren’t any information that describe the edges, nodes, and it is hard to tell which nodes are connected with what. Because the graph only has one color, one would not be able to tell that it consists of nodes describing different types of data such as posts, comments, and reactions.

Using the “good visualization” with the same information, it is much easier to read information more quickly. I used a connected component clustering to cluster the data in groups. Then, I used degree centrality to size the nodes. The bigger the degree, the bigger the nodes. Then, I styled the nodes by coloring the nodes by their type (post, page, and comment). From the visualization, it is easy and quick to read that there are 4-5 different types of pages because of the 5 clusters. The 2 big clusters show that there are 2 pages that as the most engagement. Then, for each page, you can tell which posts have the most comments and likes. Then, the 2 clusters that are closely positioned together shows that the clusters have users in common.e Overall, using a variety of algorithms and some styling, we can quickly tell which pages and their posts have the more interaction.