GROUP 2

Nima Ebadi, Iris Shvartsman, Nam Pham and Victoria Webb-Martin

**A Recommendation System based on Twitter Text Mining and Classification**

Twitter is an increasingly popular micro-blogging site, which enables users to share their thoughts and sentiment expressions via short SMS-like messages (also known as *tweets*). Due to its large number of active users, Twitter is an impetus for broad range of data analytic tasks from business intelligence to national security. In this project, we attempt to exploit text analytics, as well as sentiment classification to recommend the top positively regarded brands on Twitter. This will be achieved by employing robust twitter data mining approaches. The first step, in this regard, would be overcoming data issues such as an unstructured language, misspelled words, emoticons, sarcasm and tweets in different languages to create the recommendations.

In doing this project we believe that this will be useful for consumers who would like to research products before making a purchase. Marketers can use this tool to research public opinions of their company and monitor their public sentiment. They will also be able to analyze customer satisfaction. Companies can also use this to gather critical feedback about problems with newly released products or just as a customer service tool to help any of their consumers with issues.

**Dataset:**

The data is sourced from Sentiment140. The training dataset was created by automatically generated using a Twitter Search API.

<http://help.sentiment140.com/for-students/>

**References:**

A. Go, R. Bhayani, and L. Huang, “Twitter Sentiment Classification using Distant Supervision,” Analysis Semantic Seminar Talk Twitter, Winter 2011. <http://www.stanford.edu/~alecmgo/papers/TwitterDistantSupervision09.pdf>

S. M. Kywe, E. P. Lim, and F. Zhu, “A Survey of Recommender Systems in Twitter,” singapore Management University, Singapore SocInfo, Springer 2012

<https://link.springer.com/book/10.1007/978-3-642-35386-4#page=432>