**CS583 – Research Project**

**Introduction**

Sentiment analysis is the process of analyzing digital text to identify if the message's emotional tone is neutral, positive, or negative. With the prevalence of social media, Sentiment Analysis as a field has been brought to the forefront as it is now used for various applications such as product marketing, political analysis, customer support and feedback, market research and so on.

In this project, I aim to build a model that can classify the sentiment of tweets into one of three classes: positive, negative or neutral. As a part of the set up for this project, we are provided with two datasets: a collection of tweets related to Barack Obama, and the other a collection of tweets related to Mitt Romney. These tweets were downloaded using the Twitter API during the height of the Presidential race in the United States of America in the year 2012.

Using this data, I compare different methods of feature-extraction, models for classification, and evaluate their performance. I also discuss what I’ve learned from undertaking this process, and what could be some improvements to make in the future.

**Methodology**

To start with the training of the models, I load the tweets for Barack Obama from the excel sheet into a dataframe. The data that we are presented with requires some cleaning to get into a format that would work for us, and for that I first drop the first row of the dataframe since that has

**Results**

**Discussion**

**Conclusion**

**Recommendations**

**References**