## Project: "Text mining and Twitter Sentiment Analyzer"

Rutgers requires every student to submit a paper for this project by Dec 1st. Your papers should be written in a word document with Time New Roman size 12 with double spaces and should be least 12 pages(no upper limit). The following sections should be included (but not limited to) in your papers:

1. Introduction
2. Dataset
3. Problem statement
4. Methodology
5. Experiment
6. Related work
7. Conclusion

**Problem Statement**

The objective of this task is to build a sentiment analyzer to detect hate speech in tweets. For the sake of simplicity, we say a tweet contains hate speech if it has a racist or sexist sentiment associated with it.

So, the task is to classify racist or sexist tweets from other tweets based on positive or negative polarity.

Data Source: The data source of this project is live tweets from Twitter using API

### What is sentiment analysis?

Sentiment Analysis is the process of ‘computationally’ determining whether a piece of writing is positive, negative or neutral. It’s also known as opinion mining, deriving the opinion or attitude of a speaker.

It is scored using polarity values that range from 1 to -1. Values closer to 1 indicate more positivity, while values closer to -1 indicate more negativity.

There can be two approaches to sentiment analysis.

1. Lexicon-based methods

2. Machine Learning-based methods.

In this project I will be using Lexicon-based methods.

Why sentiment analysis?

•Business: In marketing field companies use it to develop their strategies, to understand customers’ feelings towards products or brand, how people respond to their campaigns or product launches and why consumers don’t buy some

products.

•Politics: In political field, it is used to keep track of political view, to detect consistency and inconsistency between statements and actions at the government level. It can be used to predict election results as well!

•Public Actions: Sentiment analysis also is used to monitor and analyse social phenomena, for the spotting of potentially dangerous situations and determining the general mood of the blogosphere.

I will follow these 3 major steps in our Project:

•Authorize twitter API client (Getting credential to access the tweets).

•Make a GET request to Twitter API to fetch tweets for a query, trends or twitter id.

•Parse the tweets. Classify each tweet as positive, negative or neutral and visual representation.

References:

<https://developer.twitter.com>

<https://twitter.com>

https://www.geeksforgeeks.org/twitter-sentiment-analysis-using-python/

https://www.geeksforgeeks.org/twitter-sentiment-analysis-using-python/

https://towardsdatascience.com/creating-the-twitter-sentiment-analysis-program-in-python-with-naive-bayes-classification-672e5589a7ed

https://www.earthdatascience.org/courses/earth-analytics-python/using-apis-natural-language-processing-twitter/analyze-tweet-sentiments-in-python/

https://www.analyticsvidhya.com/blog/2018/07/hands-on-sentiment-analysis-dataset-python/