**Twitter Sentimental Analysis**

Project Description:

● Goals and objectives

🡪 The main objectives of this project is to find if the tweets made are racist or sexist. The purpose of this analysis is to find out what kind of emotions(positive, negative and neutral) are being expressed through a piece of information

● Project Requirements

🡪 We will implement this in Python using Python Library’s NTLK, CSV, Time, Json, RE, Twitter, word cloud, , . We will take the data from the Twitter dataset and use Twitter API as our data source. Below is the link for Twitter dataset <https://www.kaggle.com/chirag19/twitter-sentiment-analysis/data>

● Problems to be addressed

🡪 The difficult problem of sentiment composition—how to predict the sentiment of a combination of terms. The determination of sentiment of phrases (that may include negators, degree adverbs, and intensifiers) and sentiment of sentences and tweets. We will use Naïve Bayes Classifier for predicting the accuracy. Using NLP, we will handle the unstructured data and find a way to process string mathematically. Along with the classification algorithms we will also implement and discuss in depth in our writing along with the hyper-parameters tuning(grid search)

● Potential pitfalls & challenges

🡪 The potential challenge will be using the real time tweets on twitter through Twitter API and Tweepy to get the offline tweets made in the past

Background Research:

Background research for the sentimental, comments and aftereffects of the emotions shared.

Data sources:

<https://www.kaggle.com/chirag19/twitter-sentiment-analysis/data>

And Twitter API

References:

<https://towardsdatascience.com/twitter-sentiment-analysis-classification-using-nltk-python-fa912578614c>

<https://www.nltk.org/>

<http://docs.tweepy.org/en/latest/getting_started.html#hello-tweepy>

Project members:

Anand Kumar

Sohail Budhwani