--Goal

“Opinion Mining”: Classify text into positive/negative/neutral based on the overall sentiment

Step A: Preparing Test Set

1) Create twitter application and request for authentication credentials

<https://developer.twitter.com/en/docs>

2) Authenticating our python script

-- Create an Twitter.API object with the credentials and test if authentication works

3) Create function to get the data

-- A function that takes the keyword we want to search for and returns the tweets that we can iterate through

Step B: Preparing Training Set

--Data that already labeled with positive/negative/neutral

(Where to get?)

Step C: Preprocessing Tweets

--Import:

a. re: RegEx library

b. ntlk: natural language processing toolkit

--Remove information that would not contribute to the sentiment classification: include but not limited to URLs, usernames, punctuations, emojis, hashtag sign, duplicate characters (ex. delllllicious)

--Tokenized for processing

Step D: Training and testing

--Naïve Bayes Classifier

--build feature vector for training

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