

# Assignment A4. 41403



- Title : Twitter data analysis.
- problem statement: ~~we~~ use twitter data for sentiment analysis. The dataset is 3mb in size & has 30k+ tweets. Identify the tweets that are hate tweets.
- Objective: To classify tweets as hate tweets or not.
- outcome: Identifying & removing hate tweets from twitter.
- SW & HW requirements: python 3, jupyter environment.

## Theory:

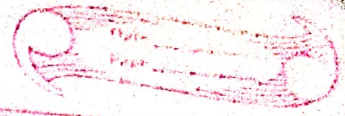
- Natural language processing is a subfield of linguistics computer science & artificial intelligence concerned with interactions b/w computers & human language in particular how to program computers to process & analyze large amounts of natural language data.
- Stop-words are words that are filtered out before or after the natural language data are preprocessed.
- Stemming : for grammatical reasons, text can use different forms of a word. These are also families of derivationally related



words with similar meanings

- stemming reduces inflectional forms & sometimes derivationally linked forms of a word with similar meanings / common base form.
- when applied to a document the result is like ORIGINAL the boy's cars are different colors. STEMMED the boy car be differ color.
- Feature selection is the process of selecting a subset of the terms occurring in the training set.
- This makes the classifier more efficient as well as more accurate because it eliminates noise.
- vectorization is the process of converting the text data into a machine-readable form.
- TF-IDF vectors are related to one-hot encoding but instead of just featuring a count, they feature numerical representation where words aren't just present or not present instead they are represented by their term frequency multiplied by their inverse doc frequency.





- For this particular problem, which is classifying tweets as hate tweets or not.
- The classification methods used were multinomial Naive Bayes, Random Forest & linear support vector classifier.
- Accuracy of  $> 90\%$  was achieved
- The tweets were pre-processed to convert them to lowercase removed @ mentions, removed numbers punctuations
- The tweets were vectorized TFIDF & split into training & test data
- The 3 models were fitted & then used to predict the labels.
- Conclusion: successfully classified tweets as hate tweets or not.