## 

Sentiment Analysis of a Tweets

**Lohith DS**

**Harsha N**

**Nitinraj**

### **Problem Statement**

Sentiment Analysis examines the problem of studying text data such as posts and reviews regarding the opinions the users have about a product or service or person. These reviews or posts are uploaded by users on various platforms, forums, social media etc. Sentiment analysis has been popular for identifying sentiments in user reviews and classify those into categories such as happy, sad, angry or generally positive, negative, or neutral. Finding or searching all the user reviews about a particular entity from a large corpus of user reviews of different global entities and retrieving the sentiments from the review takes time.

**Requirements :**

* Data itself. Without proper datasets for training your machine learning algorithms there's no point in machine learning.
* Capabilities for data preparation.
* ML algorithms.
* Automation and iterative processes.

**Design :**

* Library: NumPy, matplotlib, pandas, re, nltk, sklearn, tweepy, flask.
* Data set :We are downloading tweets using tweepy which uses a twitter developer account and we stored tweets as a csv file
* Machine learning model :RandomForestClassifier.

**Implementation :**

* For a dataset we collected tweets using tweepy.
* Then we removed RT,@username,url, but kept emoji's punctuation for vader sentiment analysis and # for plotting graph of top most hashtags
* After vader sentiment analysis plotting a hashtag graphs punctuation and emoji's are no longer required so it is removed.
* Then we converted tweets into lowercase, removed stopwords,applied lemmatization.
* For visualization we have done wordcloudgraphs..etc.
* emotional analysis of tweets are also done.
* we used TfidfVectorizer for feature extraction
* Then we split our dataset into training and test set.
* After that we trained our model using training set .
* Then we tested our model with unseen dataset which is test set and determined the accuracy.
* We also tested the model with new reviews manual giving input.
* Then we evaluated the model using confusion matrix, precision and recall.

**Conclusion :**

We are able to classify the sentiments of tweets using machine learning with good accuracy and also we are able to get the results in web application using flask and HTML. Then we get live visualization of tweets on the dashboard.

**Challenges faced :**

* Figuring out how to use twitter api to download tweets.
* Since there are lots of Machine learning model we were not sure which to choose.
* To create a web page we did lots of research in youtube and references

**References :**

* we searched for research papers,github and Youtube videos for reference