**REPORT ON SENTIMENT ANALYSIS AND TWITTER SENTIMETS**

**SENTIMENT ANALYSIS:**

Sentiment Analysis is the automated process of analyzing text to determine the sentiment expressed whether it is either Positive, Negative or Neutral. In general we can say that it describes the nature of the given set of text. It can used in Social Media Monitoring, Customer Support, Customer Feedback, Brand Monitoring and Reputation Management, Voice of Customer, Voice of Employee, Product Analysis, Spam Filtering and Market Research and Competitive Research.

Using machine learning algorithms, natural language processing and steps of preprocessing (identify and remove stop words, count vectorizer) we can extract the subjective information of a document(csv, tsv, text files) or sentence and try to classify it according to its polarity such as positive, neutral or negative.

It is observed that the analysis can be done using two methods, namely: Support Vector Machine(using Support Vector Classifier) and Naives Bayes Theorem(Multinomial Naive Bayes). In this methods Naive Bayes Theorem(previous knowledge of condition sis needed) predicts the sentiment of a text efficiently when compared to the Support Vector Machine.

Machine learning pipelines consist of multiple sequential steps that do everything from data extraction and preprocessing to model training and deployment. In Sentiment Analysis the pipeline concept can be used which is the combination of the estimators(fit, transform, predict functions) used in Support Vector Machine or Naive Bayes Theorem. This is accurate and efficient because it takes less time, less code and has more accuracy score compared to Support Vector Machine or Naive Bayes Theorem.

And a Joblib can be used to dump, load and save pipelines. Using Joblib pipelines can be saved in a correct order. Using Joblib we can predict the new text or existing text of a dataset. And it gives the accurate results.

**TWITTER SENTIMENTS:**

Twitter is a popular microblogging website we is used to share the opinion or emotion of a user in a 140 characters of length. Sentiment Analysis is highly useful to analyze the text tweeted in Twitter.

Twitter sentiment analysis is important because twitter sentiments allows everyone to keep track of what's being said about your product, service or behaviour and character of a person on social media, and can help you detect angry customers or negative mentions before they escalate.

The data in the twitter can not only be classified using words used but also using the emoticons expressed. Like, Happy emoticons, such as “:)”, “:P”, “:­)” etc and Sad emoticons, such as “:(“, “:’(”, “=(.

The presence of a large dataset is always helpful(for better training of the classifier) and twitter makes it possible to obtain any number of tweets during a desired period. However, various difficulties are faced during processing of raw tweets. Like in identify and removing stop words then punctuations, additional spaces, repeated letters, words that are not needed for analyzing are removed. Here arises a problem, as emoticons are expressed in the form of emoticons(which help to classify the sentiment of the tweet) are now being eliminated which makes a difficulty the judge the tweet sentiment.

Classifying the Twitter sentiments is very much necessary in Today's world because one may post positive tweet and one may not. And someone just keeps abusing or posting negative comments on someone then the analysis is helpful to warn them or suspend their twitter accounts to avoid inconvenience to other users and to avoid any legal actions. Here Naive Bayes Classifier can used for analyzing the Twitter Sentiments.