**Twitter Sentiment Analysis**

**Problem Statement**

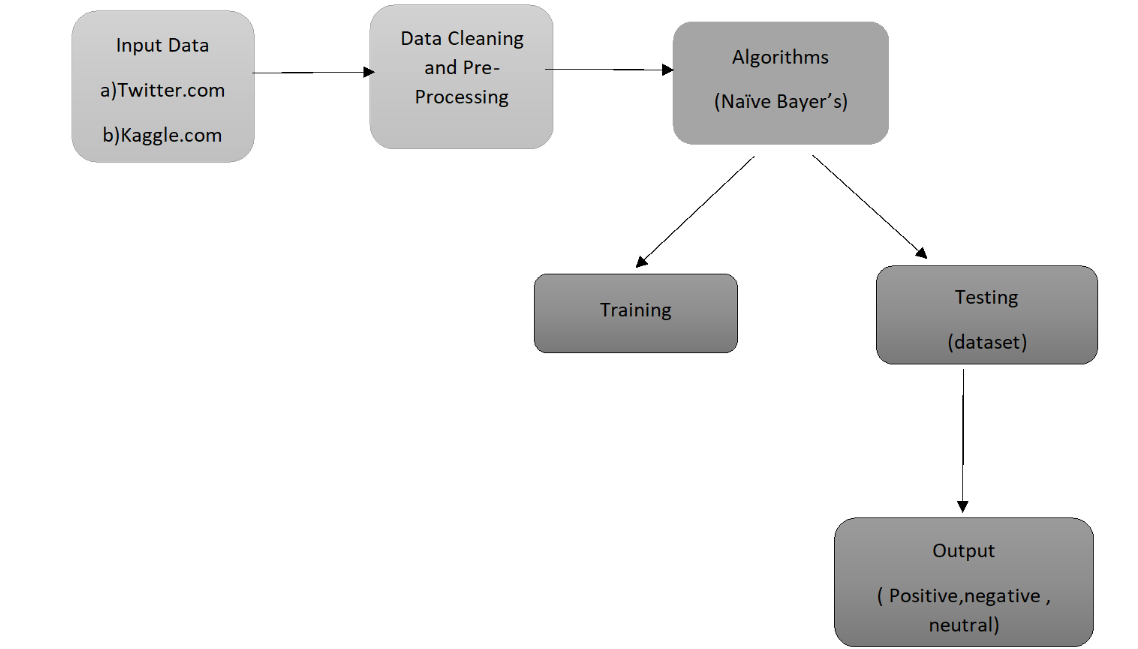
In this project, we try to implement an NLP Twitter sentiment analysis model that helps to overcome the challenges of sentiment classification of tweets. We will be classifying the tweets into positive, negative, or neutral sentiments.

**Introduction: -**

Sentiment analysis, also refers as opinion mining, is a sub-machine learning task where we want to determine which is the general sentiment of a given document. Using machine learning techniques and natural language processing we can extract the subjective information of a document and try to classify it according to its polarity such as positive, neutral, or negative.

Here we choose to try to classify tweets from Twitter into “positive” or “negative” sentiments by building a model based on probabilities. Twitter is a microblogging website where people can share their feelings quickly and spontaneously by sending tweets limited to 140 characters. You can directly address a tweet to someone by adding the target sign “@” or participate in a topic by adding a hashtag “#” to your tweet. Sentiment analysis is a method of data mining to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document.

**Workflow –**



**Pre-requisites: -**

1. Python

2. Machine Learning by using python libraries

**Technologies used**

Python, Machine learning, Power BI

**Libraries used**

1. **Pandas**- Used for handling and managing large amount of data by means of data manipulation (data cleaning, filtering, merging, etc) and data analysis(by data visualization).

 2. **Matplotlib**- Used to create animated, static and interactive visualizations.

3. **Seaborn**- To enhance the use of matplotlib by creating statistical graphs.

4. **NLTK**- It is a natural language processing library and we have used it for lemmatization (combines words into one meaningful word-eg run, ran, running to one word run).

5. **Text blob**- Used to process text data by calculating the polarity of words in terms of positive, negative, and neutral.

8. **Word cloud**- Used for generating word cloud from text data.

9. **RE**- Used regular expression to remove unwanted words, special characters, etc from the tweet.

**CODES**

