TWITTER SENTIMENT ANALYSIS

T.Y.B.SC.(COMPUTER SCIENCE) SEM-VI

PROJECT IN DATASCIENCE

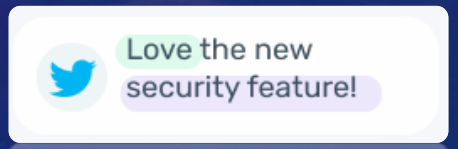
SHUBHAM DHOLE(ROLL NO: 2025007)

CONTENT

- What is sentiment analysis?
- Technologies & libraries used
- Process model
- Dataset file
- Data purification
- Data segregation on basis of polarity and subjectivity of tweet
  - Polarity
  - Subjectivity
- Data segregation as positive, negative and neutral
- Generate word cloud
- Scatterplot in terms of polarity and subjectivity of tweet
- Barplot in terms of positive, negative and neutral

WHAT IS SENTIMENT ANALYSIS?

- Sentiment analysis is the automated process of identifying and classifying subjective information in text data. This might be an opinion, a judgment, or a feeling about a particular topic or product feature.
- It is the process of classifying text as either *positive, negative,* or *neutral.* Machine learning techniques are used to evaluate a piece of text and determine the sentiment behind it.
- Sentiment analysis uses Natural Language Processing (NLP) to make sense of human language, and machine learning to automatically deliver accurate results.
- The most common type of sentiment analysis is 'polarity detection' and involves classifying statements as *positive*, *negative* or *neutral*. A polarity sentiment analysis model, for example, automatically tags this tweet as *positive*:

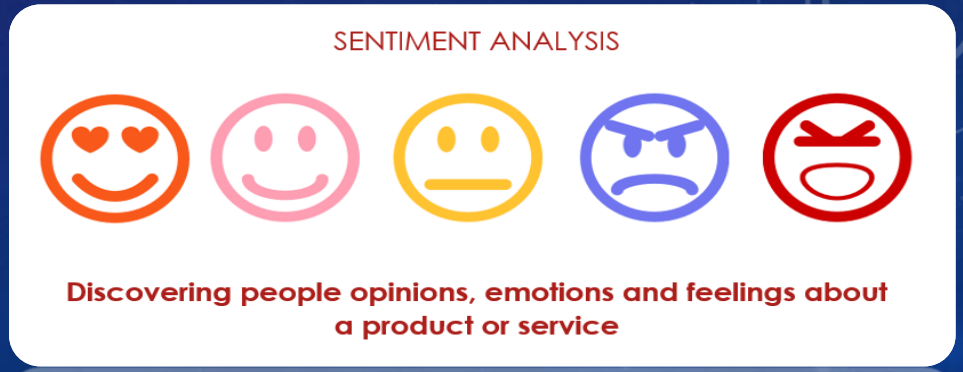


WHY IS SENTIMENT ANALYSIS USEFUL?

- Sentiment analysis is essential for businesses to gauge customer response.

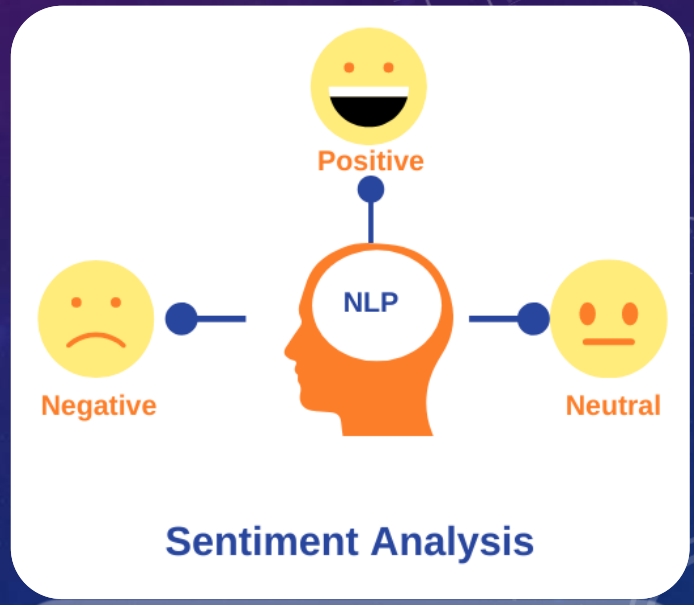
Example: Your company has just released a new product that is being advertised on a number of different channels.

- In order to gauge customer's response to this product, sentiment analysis can be performed.
- Customers usually talk about products on social media and customer feedback forums. This data can be collected and analyzed to gauge overall customer response.
- Taking this a step further, trends in the data can also be examined. For example, customers of a certain age group and demographic may respond more favourably to a certain product than others.
- Based on the information collected, companies can then position the product differently or change their target audience.
- Connect sentiment analysis tools directly to your social platforms , so you can monitor your tweets as and when they come in, 24/7, and get up-to-the-minute insights from your social mentions.



SENTIMENT ANALYSIS USING TWITTER

- Twitter allows businesses to engage personally with consumers. However, there's so much data on Twitter that it can be hard for brands to prioritize mentions that could harm their business.
- That's why sentiment analysis, a tool that automatically monitors emotions in conversations on social media platforms, has become a key instrument in social media marketing strategies.
- Carefully listening to voice of the customer on Twitter using sentiment analysis allows companies to understand their audience, keep on top of what's being said about their brand – and their competitors – and discover new trends in the industry.



HOW TO PERFORM SENTIMENT ANALYSIS ON YOUR TWITTER DATA

Performing sentiment analysis on Twitter data involves five steps:

1. Gather relevant Twitter data
2. Clean your data using pre-processing techniques
3. Create a sentiment analysis machine learning model
4. Analyze your Twitter data using your sentiment analysis model
5. Visualize the results of your Twitter sentiment analysis

TECHNOLOGIES & LIBRARIES USED:

- Jupyter Notebook (Python IDE)

Import libraries:

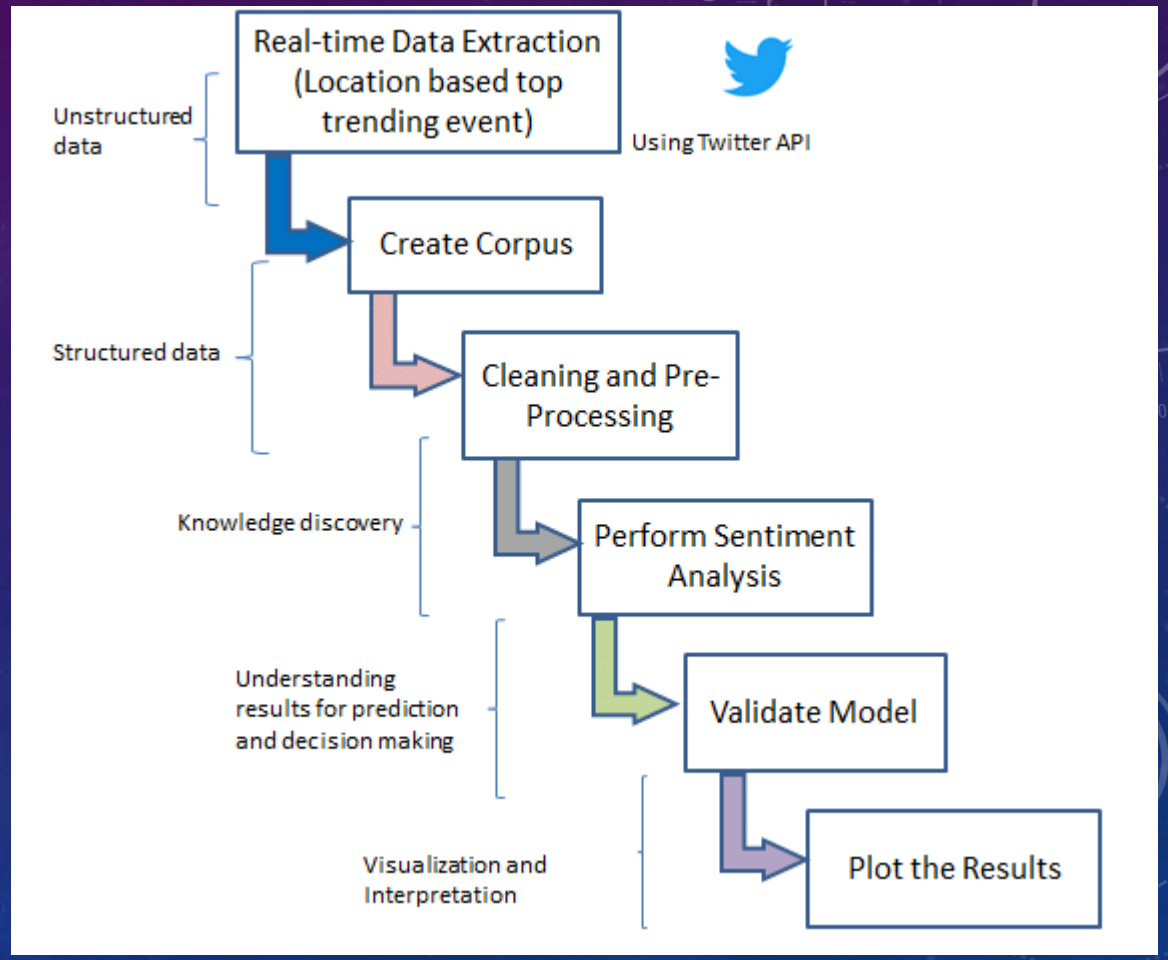
- Textblob : Used to read and identify the human emotion in the text
- Wordcloud : Used to display the clusters of most occurring words in the given amount of data
- Tweepy : Twitter Api module which helps us to connect to twitter via python
- Pandas : Used to create dataframes which are the areas holding the information
- Numpy : Used to perform mathematical operations
- Re : Regular Expression module
- Matplot lib : python library used for data visualization

STEPS

- Creating dataframe
- Reading config.csv file
- Establish tweeter connection
- Loading data from tweeter account

- Data purification
- Data segregation on basis of polarity and subjectivity of tweet
- Polarity
- Subjectivity
- Data segregation as positive, negative and neutral

- Generate word cloud which shows the most frequent or most occurring words in particular tweets.
- Plot scatterplot in terms of polarity and subjectivity of tweet



PROCESS MODEL

Dataset file : config.csv- tweeter authentication token file

To get access to tweeter account e.g wwe(world wide wrestling entertainment)

We have taken record of 1000.

1. Creating dataframe
2. Reading config.csv file
3. Establish tweeter connection
4. Loading data from tweeter account

Data before processing

	Tweets
0	Before @TBARRetribution & @MACEtheWRESTLER...
1	#TheBoss always bounces back. #SmackDown\n\n@S...
2	From swinging @UNBESIEGBAR_ZAR 40 times to cat...
3	Cast your vote now for #WrestleMania 36 in @Th...
4	RT @WWENetwork: So ... which victories made @m...

Data after processing

	Tweets	Subjectivity	Polarity
0	Before & dished out further damage on , ...	0.500000	0.000000
1	TheBoss always bounces back. SmackDown\n\n 💙💚	0.000000	0.000000
2	From swinging _ZAR 40 times to catching oppone...	0.450000	0.300000
3	Cast your vote now for WrestleMania 36 in Peo...	0.000000	0.000000
4	So ... which victories made the first and on...	0.666667	0.125000
...	...	...	...
995	Hell yeah! Congrats champ. _WWE 🤘\n(Can I bor...	0.000000	0.000000
996	WrestleMania NXTTakeOver\n\nWeAreNXT 🖤💛	0.000000	0.000000
997	Yaaaas _WWE !!!! Congrats wifey!!! 💞😘	0.000000	0.000000
998	Your NIGHTMARE is her REALITY. \n\nCongratulat...	0.454545	0.170455
999	Another night of WrestleMania, another sold-o...	0.450000	0.400000

1000 rows × 3 columns

SENTIMENT ANLAYSIS BASED ON SUBJECTIVITY AND POLARITY

Subjective sentences generally refer to personal opinion, emotion or judgment whereas objective refers to factual information.

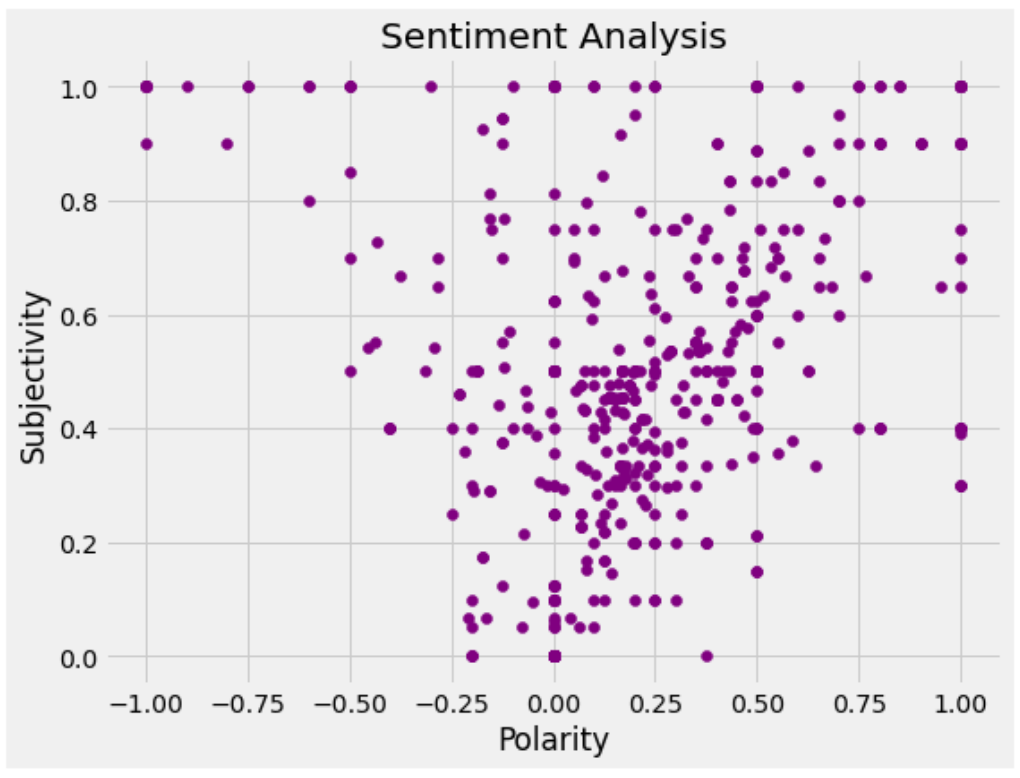
Polarity is float which lies in the range of [-1,1] where 1 means positive statement and -1 means a negative statement.

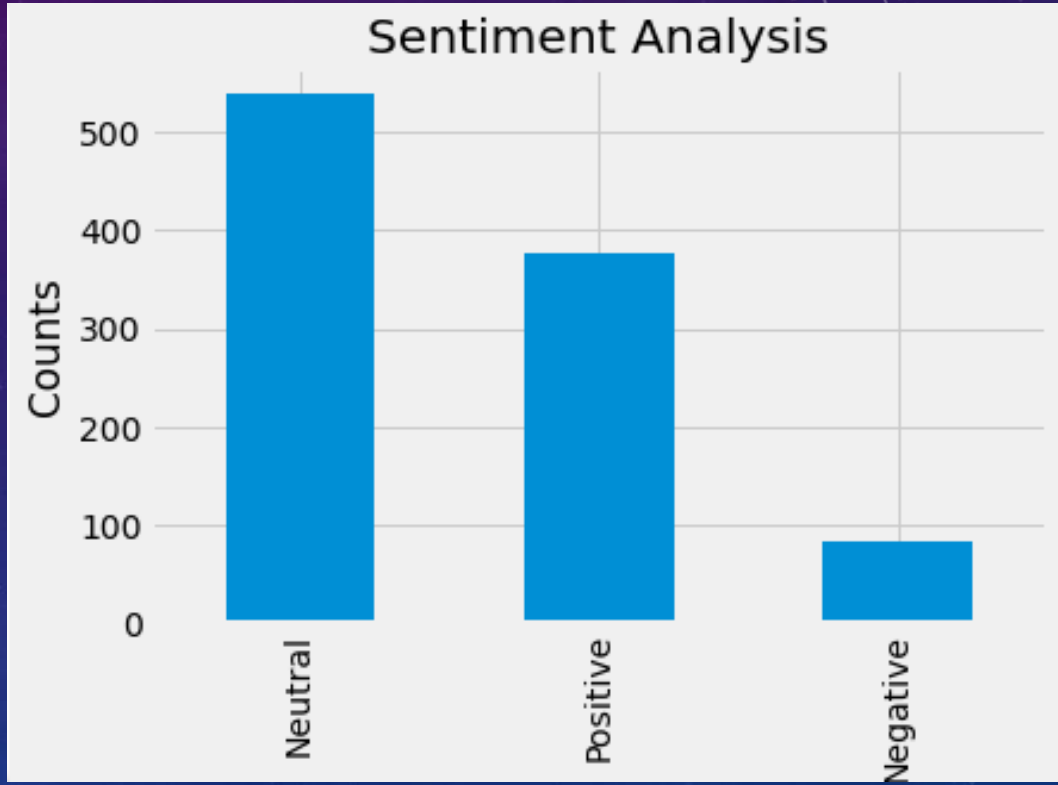
Sentiment polarity for an element defines the orientation of the expressed **sentiment**, i.e., it determines if the **text** expresses the positive, negative or neutral **sentiment** of the user about the entity in consideration.

	Tweets	Subjectivity	Polarity	Analysis
0	Before & dished out further damage on , ...	0.500000	0.000000	Neutral
1	TheBoss always bounces back. SmackDown\n\n 💙💚	0.000000	0.000000	Neutral
2	From swinging _ZAR 40 times to catching oppone...	0.450000	0.300000	Positive
3	Cast your vote now for WrestleMania 36 in Peo...	0.000000	0.000000	Neutral
4	So ... which victories made the first and on...	0.666667	0.125000	Positive
...	...	...	...	...
995	Hell yeah! Congrats champ. _WWE 🤘\n(Can I bor...	0.000000	0.000000	Neutral
996	WrestleMania NXTTakeOver\n\nWeAreNXT 🖤💛	0.000000	0.000000	Neutral
997	Yaaaas _WWE !!!! Congrats wifey!!! 💞😘	0.000000	0.000000	Neutral
998	Your NIGHTMARE is her REALITY. \n\nCongratulat...	0.454545	0.170455	Positive
999	Another night of WrestleMania, another sold-o...	0.450000	0.400000	Positive

1000 rows × 4 columns

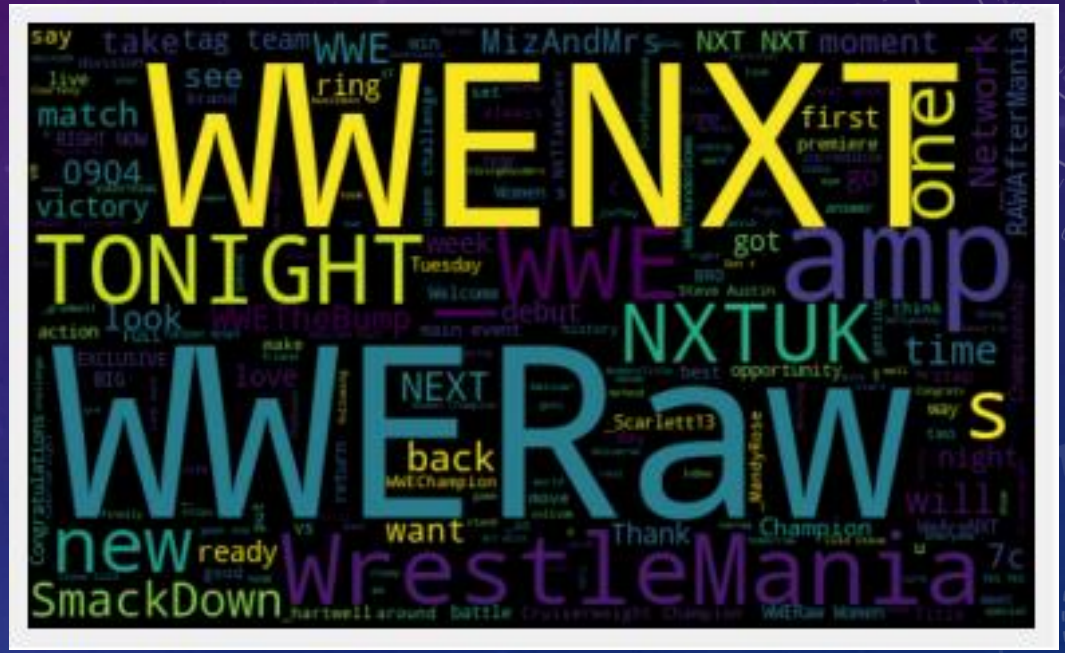
SENTIMENT ANALYSIS





WORDCLOUD:

Word Cloud is a data visualization technique used for representing text data in which the size of each word indicates its frequency or importance. Significant textual data points can be highlighted using a word cloud. Word clouds are widely used for analysing data from social network websites.



Thank You