TWITTER SENTIMENT ANALYSIS

1.1 OBJECTIVE:

The use of social media has become an integral part of daily routine in modern society. Social media portals offer powerful public platforms where people can freely share their opinions and feelings about various topics with large crowds.People have always had an interest in what people think, or what their opinion is. Since the inception of the internet, increasing numbers of people are using websites and services to express their opinion.

1.1.1 What is sentiment analysis?

Sentiment Analysis is the process of 'computationally' determining whether a piece of writing is positive, negative or neutral. It's also known as opinion mining, deriving the opinion or attitude of a speaker.Opinions are usually subjective expressions that describe people's sentiments, appraisals, and feelings toward a subject or topic.

In other words,process of computationally identifying and categorizing opinions from piece of text and determine whether the writer's attitude towards a particular topic or the product is positive,negative,or neutral.

Sentiment analysis or Opinion mining, as it is sometimes called, is one of many areas of computational studies that deal with opinion oriented natural language processing. Such opinion oriented studies include among others, genre distinctions, emotion and mood recognition, ranking, relevance computations, perspectives in text, text source identification and opinion oriented summarization.

Sentiment is an attitude, thought, or judgment prompted by feeling. Sentiment analysis is also known as opinion mining. Sentiment analysis or opinion mining is one of the major tasks of NLP (Natural Language Processing). It has gain much attention in recent years. Internet is a resourceful place with respect to sentiment information.

From a user's point of view, people are able to post their own content through various social media, such as forums, micro-blogs, or online social networking sites. From a researcher's perspective, many social media sites release their application programming interfaces (APIs), prompting data collection and analysis by researchers and developers.Data from these online sources can be used in opinion mining and sentiment analysis tasks. For example, we may be interested about "How positive (or negative) are people about it?" or "What do people think about this persons status post(comments) ?".

Sentiment Analysis, an area of Natural Language Processing (NLP), is used to classify the reviews using the sentiment of the words into positive or negative. Using the sentiment expressed in the words, opinions on any entity can be categorized into positive or negative.

For example,

- The sentence, 'I am not excited by this product though it is quite cheap' expresses a negative sentiment about the product.

- The sentence, ‘I love this product’ indicates a more positive sentiment than the sentence ‘I like this product’.

It’s estimated that 80% of the world’s data is unstructured and not organized in a pre-defined manner. Most of this comes from text data, like emails, support tickets, chats, social media, surveys, articles, and documents. These texts are usually difficult, time-consuming and expensive to analyze, understand, and sort through.

With the help of sentiment analysis systems, this unstructured information could be automatically transformed into structured data of public opinions about products, services, brands, politics, or any topic that people can express opinions about. This data can be very useful for commercial applications like marketing analysis, public relations, product reviews, net promoter scoring, product feedback, and customer service.

1.1.2 REAL LIFE EXAMPLES :

1.Sentiment Analysis in Social Media Monitoring

In today’s day and age,interactions with customers and even competition on social networks like Facebook, Twitter, and Instagram. Most marketing departments are already tuned into to online mentions By using sentiment analysis on social media, we can get incredible insights into the quality of conversation that’s happening around

❖ **How Sentiment Analysis Can Be Used**

- Analyze tweets and/or facebook posts over a period of time to see sentiment of a particular audience.

- Run sentiment analysis on all social media mentions to your brand and automatically categorize

2.Sentiment Analysis in Brand Monitoring

Instead of focusing on specific social media platforms such as Facebook and Twitter, we can target mentions in places like news, blogs, and forums –again, looking at not just the volume of mentions, but also the quality of those mentions.

❖ **How Sentiment Analysis Can Be Used**

- Analyze news articles, blog posts, forum discussions, and other texts on the internet over a period of time to see sentiment of a particular audience.

- Automatically categorize urgency of all online mentions to your brand via sentiment analysis.

3.Sentiment Analysis in Market Research and Analysis

Sentiment analysis empowers all kinds of market research and competitive analysis. Whether you’re exploring a new market, anticipating future trends, or keeping an edge on the competition, sentiment analysis can make all the difference.

❖ **How Sentiment Analysis Can Be Used**

- Analyze product reviews of your brand and compare those with the competition

- Generate weekly, monthly, or daily reports –a sort of early-warning system

compare sentiment across international markets

4.Others

- **use of sentiment analysis in the following:**
  - Voice of customer (VoC)
  - Customer service
  - Workforce analytics and voice of employee
  - Product analytics
  - Political reviews
  - Movie reviews

1.1.3 USE OF OUTCOMES OF THE PROJECT:

- The major use of analyzing the variety text is for decision making strategies which can be used in business,marketing,….etc.
- When applied to social media channels, it can be used to identify spikes in sentiment, thereby allowing you to identify potential product advocates or social media influencers.
- It can be used to identify when potential negative threads are emerging online regarding your business, thereby allowing you to be proactive in dealing with it more quickly.
- Sentiment analysis could also be applied to your corporate network, for example, by applying it to your email server, emails could be monitored for their general "tone".

1.2 TECHNICAL PROSPECTIVE :

Sentiment analysis is a type of data mining that measures the inclination of people's opinions through natural language processing (NLP), computational linguistics and text analysis, which are used to extract and analyze subjective information from the Web - mostly social media and similar sources. The analyzed data quantifies the general public's sentiments or reactions toward certain products, people or ideas and reveal the contextual polarity of the information.

Sentiment analysis is also known as opinion mining.Opinions that are mined from such services can be valuable. Datasets that are gathered can be analyzed and presented in such a way that it becomes easy to identify if the online mood is positive, negative or even indifferent.positive sentiment can be identified thereby allowing the identification of product advocates or to see which parts of a business strategy are working.

1.2.1 TECHNICAL TERMS:

We make the project work with help of weepy,textblob ,Naive bayes classifier,matplotlib..etc.

- **Tweepy:** tweepy is the python client for the official Twitter API.

- **TextBlob:** textblob is the python library for processing textual data.
- **Naive bayes classifier:** Naive bayes is used to check the polarity of the sentence.
- **Matplotlib:** matplotlib is to visualization,GUI….etc.

1.2.2 HOW DOES SENTIMENT ANALYSIS WORK:

Step 1: TOKENIZATION

- Tokenization is nothing but dividing the given data into different set of words.

Text: The movie was great!

Result:

- The
- Movie
- was
- Great
- !

Step 2: CLEANING THE DATA

- Removal all the special characters as they do not add any value to the analytics part.

Text: The movie was great!

Result:

- The
- Movie
- was
- Great

!

Step 3: REMOVING STOP WORDS

- The words that do not add any meaning to analytics part are know as stop words like the,was,is,he,she…etc.

Text: The movie was great!

Result:

The

- Movie

was

- Great

!

Step 4: CLASSIFICATION

- The words are classified as either positive,negative or neutral.
- For a positive word we give a sentiment score as +1 .
- For a negative word we give a sentiment score as -1 .
- For a neutral word we give a sentiment score as 0.

Text: The movie was great!

Result:

The

- Movie (neutral word : 0)

was

- Great (positive word :+1)

!

Step 4: APPLY SUPERVISED ALGORITHM

- Train your model with bag of words or lexicons(a dictionary of pre-classified set of words) and test it on analyzing statement.
- More the accuracy score better will be the classification.

Text: The movie was great!

Result:

The

- Movie (neutral word : 0)

was

- Great (positive word :+1)

!

Step 5: CALCULATION

- Add all the polarities obtained in step4 and conclude if it is a positive or negative.
- More the accuracy score better will be the classification.

Text: The movie was great!

Result:

The

- Movie (neutral word : 0)

was

- Great (positive word :+1)

!

0+1=1

(since the resultant polarity is greater than zero ,the given text is positive)

2. POSSIBLE EXTENSIONS/ENHANCEMENT :

- In this project we are use the models derived from trained data sets and algorithms to output the analyzed polarity.we can extended it to a practical level by extaracting data directly from say twitter through API's and output the polarity.
- Determine neutrality.Potential improvement can be made to our data collection and analysis method.Future research can be done with possible improvement such as more refined data and more accurate algorithm.
- Further more we have even enhance the work on more than one social networking sites like collecting the data contents from twitter,facebook,instagram simultaneously and analyzing them in parallel.

3. RESEARCH ARTICLES :

[1] Y. Mejova, "Sentiment analysis: An overview," Comprehensive exam paper, available on http: //www. cs. uiowa. edu/ ymejova/publications/ CompsYelenaMejova.

[2] Boiy, P. Hens, K. Deschacht, and M.-F. Moens, "Automatic sentiment analysis in on-line text," in Proceedings of the 11th International Conference on Electronic Publishing.

[3] A. Balahur, J. M. Hermida, and A. Montoyo, "Building and exploiting emotinet, a knowledge base for emotion detection based on the appraisal theory model," Affective Computing.

[4] Alexander Pak, Patrick Paroubek , "Twitter as a Corpus for Sentiment Analysis and Opinion Mining" , Universit´e de Paris-Sud, Laboratoire LIMSI-CNRS, Bˆatiment 508, F-91405 Orsay Cedex, France alexpak@limsi.fr,pap@limsi.fr

[5] Minqing Hu and Bing Liu, 2004, Mining and summarizing customer reviews, Proceedings of the 10th ACM SIGKDD International conference on knowledge discovery and data mining.

[6] Peter.D.Turney, Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews,Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL),Philadelphia,July 2002, pp.417-424.

[7] C. Lin and Y. He,Joint Sentiment/Topic Model for Sentiment Analysis.,The 18th ACM Conference on Information and Knowledge Management (CIKM),Hong Kong, China,Nov. 2009

[8] Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas,A. & Hołyst, J.A. (2011). Collective emotions online and their influence on community life, PLoS ONE, 6(7): e22207.