**Mini Project Report on**



**SENTIMENTAL ANALYSIS USING MACHINE LEARNING MULTIPLE CLASSIFICATION**



**Submitted in partial fulfilment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**SHRISTY CHAUDHARY**  **2019498**

***Under the Mentorship of***

**Dr.A.SURESH KUMAR**

**Designation**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**July-2023**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Twitter Data Sentimental Analysis Using Machine Learning Multiple classification”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. A. Suresh Kumar, Designation**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Shristy Chaudhary 2019498

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | **1-1** |
| Chapter 2 | Literature Survey | **2-5** |
| Chapter 3 | Methodology | **6-9** |
| Chapter 4 | Result and Discussion | **10-12** |
| Chapter 5 | Conclusion and Future Work | **13-15** |
|  | References | **16** |

**Chapter 1**

**Introduction**

Understanding the sentiments of the people is not easy unless they express their feelings, opinions and perspective anything. But if you have such platforms where people are freely speaking up about their thoughts and concerns, you can easily find out their sentiments. Here where Cogito comes in the facility of sentiment analysis.

Sentiment Analysis is the process of determining the conceptions, judgments, feelings, opinions, viewpoints, conclusions, and other notions towards anything. It is a technique to analyze texts, images, emojis and various other actions to know what other people think about a product, service, company, brand name, or a reaction to a specific event, social movement, etc

Sentiment analysis is the process of automatically classifying text data according to their polarity, such as positive, negative, or neutral.[1]

In the context of Twitter data, sentiment analysis is used to analyze the sentiments expressed in tweets and understand how users are talking about products, services, or various topics**[[1]](https://huggingface.co/blog/sentiment-analysis-twitter" \t "_blank)**

One approach to sentiment analysis is using multiple classifications. This involves classifying the extracted data into sentiment polarities, such as positive, neutral, and negative classes**[[3]](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9910766/" \t "_blank)**

By using multiple classifications, more granular insights can be obtained, such as identifying specific themes or topics associated with each sentiment

The goal of sentiment analysis on Twitter data using multiple classification is to gain insights into customer opinions, drive business decisions, and identify potential issues or crises early on

 It can be useful for companies to understand how customers perceive their products or services, as well as to monitor public sentiment on various topics.There are different methods and approaches for sentiment analysis on Twitter data. These include lexicon-based approaches, machine learning-based approaches, and deep learning methods

Each approach has its own advantages and challenges, and researchers continue to explore and improve the performance of sentiment analysis on Twitter data

Overall, sentiment analysis of Twitter data using multiple classification provides a valuable tool for understanding and analyzing public sentiment, customer opinions, and trends on various topics. It allows companies and researchers to make data-driven decisions and gain insights into the sentiments expressed in tweets.

**Chapter 2**

**Literature Survey**

**2.1**

Literature Survey of Sentiment Analysis of Twitter Data:

1. "Twitter Data Sentimental Analysis Using Multiple Classifications"[1]:

- This work analyzed various sentiment analysis techniques using multiple classifications. The evaluation of the experiment concluded that natural language processing (NLP) techniques can effectively classify Twitter data into sentiment polarities.

2. "Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach"[2]:

- This study compared the application of lexicon-based and machine learning-based approaches for sentiment analysis of Twitter data. It emphasized the importance of classifying extracted data into sentiment polarities such as positive, neutral, and negative classes.

3. "A Survey on Sentiment Analysis using Twitter Dataset"[3]:

- This survey paper presents an overview of sentence-level sentiment classification in sentiment analysis. It discusses various techniques and methods used for sentiment analysis on Twitter data, including Naive Bayes and opinion mining.

4. "Sentiment Analysis of Twitter Data: A Survey of Techniques"[4]:

- This survey focuses on sentiment analysis of Twitter data and provides a comparative analysis of existing techniques for opinion mining. It highlights the challenges of analyzing unstructured and heterogeneous opinions expressed in tweets.

5. "A Survey on Sentiment analysis of Twitter using Machine Learning"[5]:

- This survey paper discusses the application of machine learning techniques for sentiment analysis of Twitter data. It explores how sentiment analysis can be applied to extract users' expressed sentiments from tweets.

Based on the literature survey, it is evident that sentiment analysis of Twitter data is a widely studied topic. Researchers have explored various techniques, including lexicon-based, machine learning-based, and NLP-based approaches, to classify tweets into sentiment polarities. The goal is to gain insights into customer opinions, drive business decisions, and identify potential issues or crises early on. The survey papers provide a comprehensive overview of the different methods and approaches used in sentiment analysis of Twitter data.

**2.2**

Here are some major ongoing projects or researches on sentiment analysis:

1. Deep learning for sentiment analysis: Successful approaches and future challenges[1]:

- This research paper explores the use of deep learning for sentiment analysis and discusses successful approaches and future challenges. It highlights the potential of deep learning techniques for improving the accuracy of sentiment analysis.

2. Machine learning and deep learning for sentiment analysis across languages: A survey[2]:

- This survey paper discusses the use of machine learning and deep learning techniques for sentiment analysis across languages. It categorizes the methods and techniques used for sentiment analysis and provides new research directions for future work.

3. A Survey on Sentiment analysis of Twitter using Machine Learning[3]:

- This survey paper provides an overview of sentiment analysis of Twitter data using machine learning techniques. It discusses the challenges and opportunities of sentiment analysis on Twitter data and provides insights into future research directions.

4. Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach[4]:

- This research paper compares the application of lexicon-based and machine learning-based approaches for sentiment analysis of Twitter data. It emphasizes the importance of classifying extracted data into sentiment polarities such as positive, neutral, and negative classes.

5. Sentiment Analysis Guide[6]:

- This guide provides an overview of sentiment analysis and discusses various techniques and methods used for sentiment analysis. It highlights the importance of natural language processing and deep learning for sentiment analysis and provides insights into mastering sentiment analysis.

Overall, ongoing research in sentiment analysis focuses on exploring new techniques and methods for improving the accuracy of sentiment analysis. Researchers are also exploring the use of deep learning and machine learning techniques for sentiment analysis across languages and on social media platforms such as Twitter. The goal is to gain insights into customer opinions, drive business decisions, and identify potential issues or crises early on.

**Chapter 3**

**Methodology**

**3.1 Python**

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together.

Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

**3.2 NLP (Natural Language Processing)**

Natural Language Processing (NLP) is a subfield of [artificial intelligence (AI)](https://monkeylearn.com/blog/nlp-ai/). It helps machines process and understand the human language so that they can automatically perform repetitive tasks. Examples include machine translation, summarization, ticket classification, and spell check.

Take [sentiment analysis](https://monkeylearn.com/sentiment-analysis/), for example, which uses natural language processing to detect emotions in text. This classification task is one of the most popular tasks of NLP, often used by businesses to automatically [detect brand sentiment on social](https://monkeylearn.com/blog/brand-sentiment/) . Analyzing these interactions can help brands detect urgent customer issues that they need to respond to right away, or monitor overall customer satisfaction.

**3.3 Data Set and related Libraries**

A data set was downloaded from Kaggle.com and analysis was performed on it .

Some of the key libraries used are :-

* **Panda , Seaborn and Numpy**

For Data manipulation , analysis and working. Panda offers different types of Data Structure while Numpy helps in working with arrays and seaborn is used for visualization and exploratory data analysis.

* **Metplotlib.pyplot**

For working with 2D graphics. It helps to generate a 2D graph between 2 given coordinates of data itself and helps in plotting of figures , interpolation and so on.

* **Regular Expression and String**

We are familiar with the term Regex or Regular Expressions. It helps to check string validation to pre process data to extract the main content to work upon.

* **NLTK**

It is the main link of NLP in our project. It stands for natural language toolkit and is used to perform on unstructed data and human-readable types.

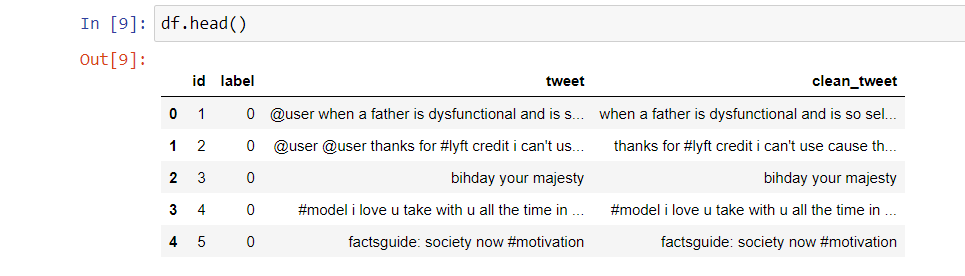
* **Warnings**

It is to handle any exceptions or warn of any situations that aren’t necessary exceptions .This ensures proper functioning of the code .

**WORKING :-**

1. **TOKENIZATION**

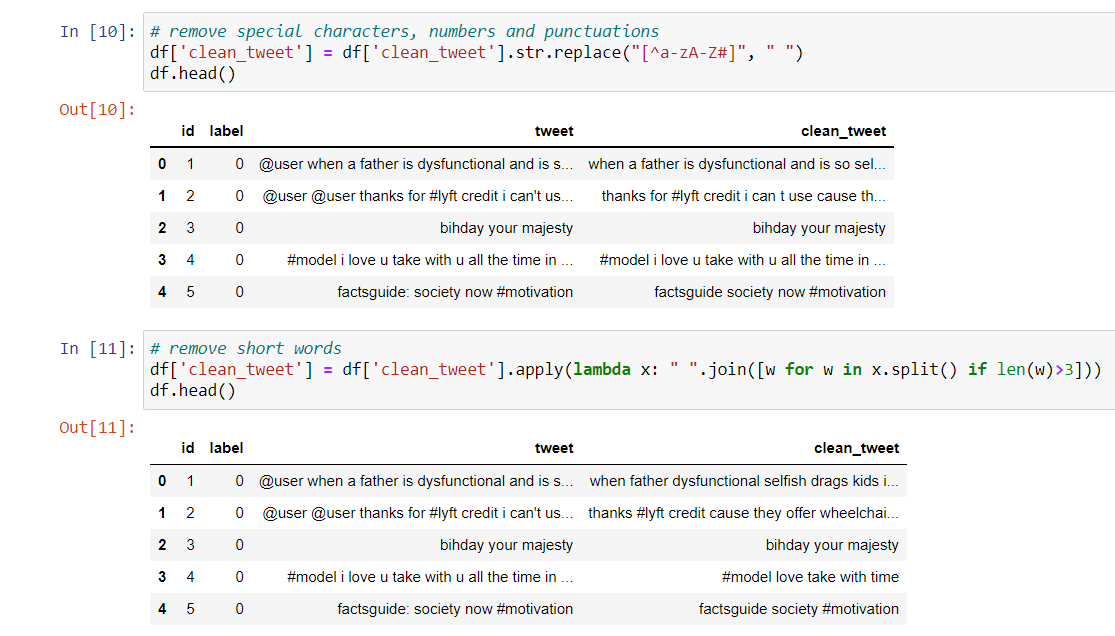
Reduction of words into individual tokens/words to perform operations.

****

**Fig. 3.1**

1. **PROCESSING – (a)**

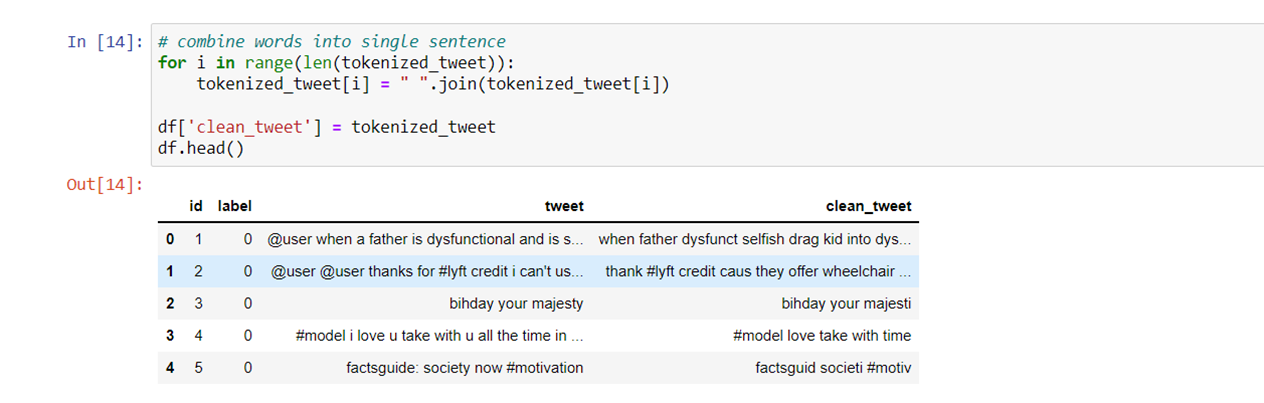
Removing special characters , popular parts of speech and reduce the set to words that contribute to analysis.



**Fig. 3.2**

1. **PROCESSING – (b)**

With the remaining words , we need to classify them as positive , negative or neutral to perform analysis.



**Fig. 3.3**

**Chapter 4**

**Result and Discussion**

We display top results in form of images and graphs.

Pic 4.1 and 4.2:- image of positive sentiments tweets.

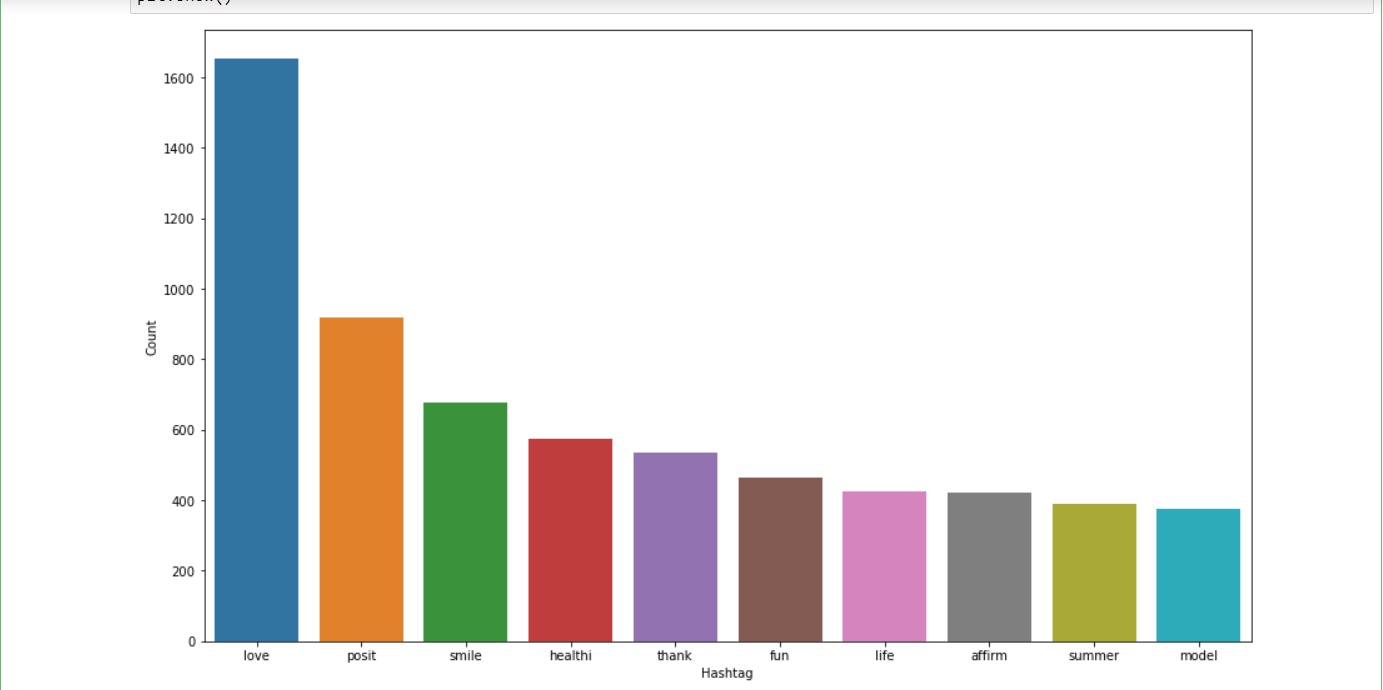
Pic 4.3 and 4.4 :- graph of positive and negative tweets.



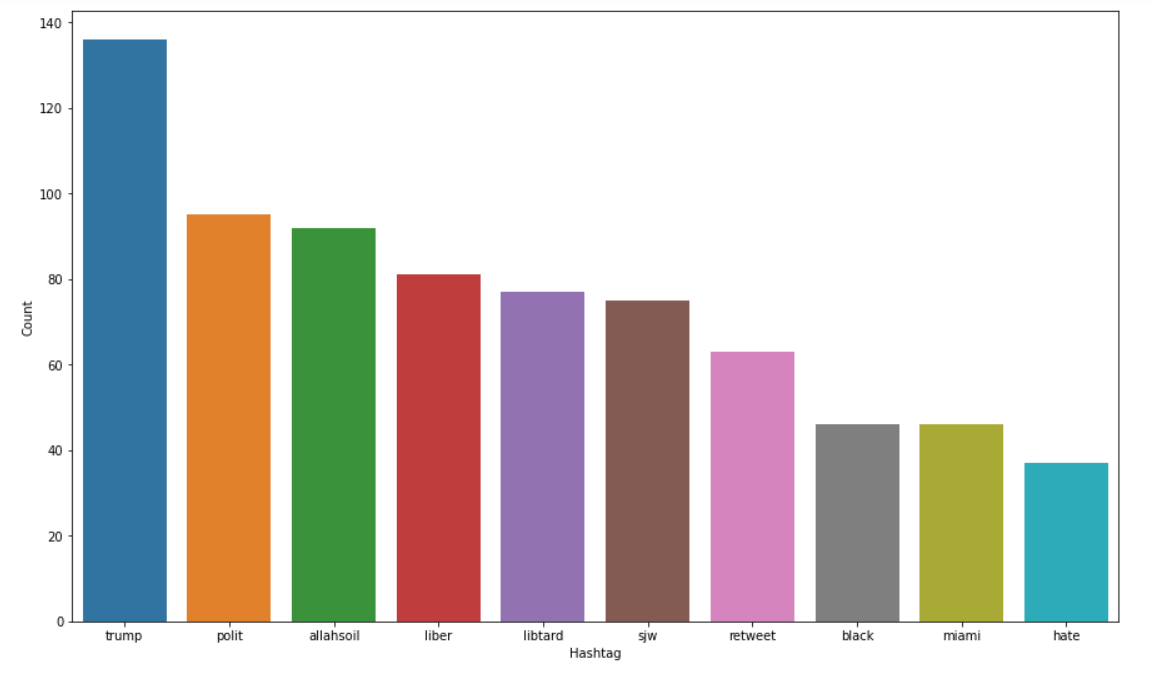
**Fig. 4.1**



**Fig. 4.2**



**Fig. 4.3**

****

**Fig. 4.4**

**Chapter 5**

**Conclusion and Future Work**

**5.1 Conclusion**

Sentiment analysis is a valuable tool for analyzing and understanding public sentiment, customer opinions, and trends on various topics. It involves the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information[2].

The process of sentiment analysis involves classifying extracted data into sentiment polarities, such as positive, neutral, and negative classes[1][5]. Advanced sentiment classification can also look at emotional states such as enjoyment, anger, disgust, sadness, fear, and surprise[2]. There are various types of sentiment analysis, including aspect-based sentiment analysis, grading sentiment analysis, multilingual sentiment analysis, and detection of emotions[2].

Sentiment analysis is important for businesses to improve customer service, increase brand reputation, and make data-driven decisions[4]. It provides objective insights and avoids personal bias associated with human reviewers by using artificial intelligence-based sentiment analysis tools[4]. Sentiment analysis can be applied to various domains, such as product reviews, social media comments, and political debates[5].

Overall, sentiment analysis is a rapidly growing field with ongoing research and development. Researchers are exploring new techniques and methods for improving the accuracy of sentiment analysis, including the use of deep learning and machine learning techniques[1][2][3]. Sentiment analysis provides valuable insights into customer opinions, public sentiment, and trends on various topics, making it a valuable tool for businesses and researchers alike.

**5.2 Future work**

Sentiment analysis of Twitter data is a rapidly growing field with ongoing research and development. Here are some potential future directions for research in this area:

1. Exploring the use of deep learning techniques for sentiment analysis of Twitter data[1][5]: Deep learning has shown promising results in various natural language processing tasks, including sentiment analysis. Researchers can explore the use of deep learning techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for sentiment analysis of Twitter data.

2. Developing multilingual sentiment analysis techniques[2]: Sentiment analysis is often performed on English language tweets. However, Twitter is a global platform with users from different countries and languages. Developing multilingual sentiment analysis techniques can help to analyze tweets in different languages and gain insights into global public sentiment.

3. Incorporating contextual information for sentiment analysis[3]: Contextual information such as user demographics, location, and time can influence the sentiment expressed in tweets. Incorporating such contextual information can improve the accuracy of sentiment analysis and provide more granular insights.

4. Developing real-time sentiment analysis systems[4]: Twitter is a real-time platform, and sentiment analysis can be used to monitor public sentiment in real-time. Developing real-time sentiment analysis systems can help businesses and organizations to respond quickly to emerging issues and crises.

5. Exploring the use of affective computing for sentiment analysis[1][2]: Affective computing is an emerging field that focuses on the development of systems that can recognize, interpret, and respond to human emotions. Researchers can explore the use of affective computing techniques for sentiment analysis of Twitter data to gain insights into emotional states such as enjoyment, anger, disgust, sadness, fear, and surprise.

Overall, future research in sentiment analysis of Twitter data can explore new techniques and methods for improving the accuracy of sentiment analysis, developing multilingual sentiment analysis techniques, incorporating contextual information, developing real-time sentiment analysis systems, and exploring the use of affective computing for sentiment analysis.

**References**

[1] [Getting Started with Sentiment Analysis on Twitter - Hugging Face](https://huggingface.co/blog/sentiment-analysis-twitter" \t "_blank)

Sentiment analysis is the automatic process of classifying text data according to their polarity, such as positive, negative and neutral.

[2]Twitter Data Sentimental Analysis Using Multiple Classifications - ResearchGate

This work analysed various sentimental using multiple classifications. From the evaluation of this experiment, it can be concluded that NLP ...

[[3]. Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach - PMC - NCBI](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9910766/" \t "_blank)

In applying sentiment analysis, the key process is classifying extracted data into sentiment polarities such as positive, neutral, and negative ...

[[4]. Multi-Class Sentiment Analysis in Twitter: What if Classification is Not the Answer](https://ieeexplore.ieee.org/document/8496747" \t "_blank)

For this sake, we propose an approach that automatically attributes different scores to each sentiment in a tweet, and selects the sentiments ...