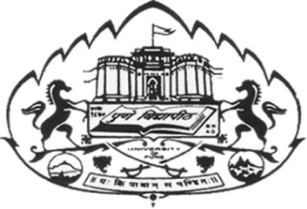
**Use the dataset and classify tweets into positive and negative tweets.**

Mini Project Report submitted to Savitribai Phule Pune University, Pune



In partial Fulfillment for the awards of Degree of Engineering in

Computer Engineering TE(Computer)

**Submitted by**

Ms. Sneha Nimbalkar(Roll No. 52)

Ms. Nikita Aherkar(Roll No. 02)

Ms. Neha Patole(Roll No. 55)

Ms. Priya Gunjal(Roll No. 27)

Ms.Sayali Vadghule(Roll No.74)

**Under the Guidance of**

**Prof. Rupali Wagh**



# Department of Computer Engineering

Parvatibai Genba Moze College Of Engineering, Wagholi, Pune

(2022-23)

# Certificate



This is to certify that, Ms Sneha Nimbalkar (Roll No. 52), Ms Nikita Aherkar (Roll No.02) ,Ms Neha Patole (Roll No.55),Ms Priya Gunjal (Roll No.27),Ms Sayali Vadghule(Roll No.74)have successfully completed the Mini project entitled “ **classify tweet into positive and negative tweets .**under my guidance in partial fulfillment of the requirements for the Third Year of Engineering in Computer Engineering under the Savitribai Phule Pune University during the academic year 2022- 2023

|  |  |
| --- | --- |
| **Prof. Rupali Wagh** | **Prof. Shrikant Dhamdhere** |
| **Project Guide** | **Head Of Department** |

**Date : / /**

**Place :** Pune

# Acknowledgements

With deep sense of gratitude we would like to thank all the people who have lit our path with their kind guidance. We are very grateful to these intellectuals who did their best to help during our project work.

It is our proud privilege to express a deep sense of gratitude to **Prof. Dr. M . G.Jadhav,** Principal of Parvatibai Genba Moze College Of Engineering, Wagholi, Pune for his comments and kind permission to complete this project. We remain indebted to **Prof. Shrikant Dhamdhere,** H.O.D. Computer Engineering Department for his timely suggestion and valuable guidance. The special gratitude goes to (**Prof. Rupali wagh**) excellent and precious guidance in completion of this work .We thanks to all the colleagues for their appreciable help for our working project. With various industry owners or lab technicians to help, it has been our endeavor throughout our work to cover the entire project work.

We are also thankful to our parents who provided their wishful support for our project completion successfully .And lastly we thank our all friends and the people who are directly or indirectly related to our project work.

Ms.Sneha Nimbalkar (Roll No.52)

Ms.Nikita Aherkar (Roll No.02)

Ms.Neha Patole (Roll No55)

Ms.Priya Gunjal (Roll No27)

Ms.Sayali Vadghule (Roll N055)

# 2.Abstract

# 

This project addresses the problem of sentiment analysis in twitter; that is classifying tweets according to the sentiment expressed in them: positive, negative or neutral.

Twitter is an online micro-blogging and social-networking platform which allows users to write short status updates of maximum length 140 characters. It is a rapidly expanding service with over 200 million registered users [24] - out of which 100 million are active users and half of them log on twitter on a daily basis generating nearly 250 million tweets per day [20]. Due to this large amount of usage we hope to achieve a reflection of public sentiment by analysing the sentiments expressed in the tweets. Analysing the public sentiment is important for many applications such as firms trying to find out the response of their products in the market, predicting political elections and predicting cioeconomic phenomena like stock exchange. The aim of this project is to develop a functional classifier for accurate and automatic sentiment classification of an unknown tweet stream.

# 3.Table Of Content

6

* 1. Acknowlegment
  2. Abstract
  3. Table of Content
  4. Introduction
  5. What is Sentiment Analysis
  6. Types Of Sentiment Analysis
  7. Importance of Sentiment Analysis
  8. Application of Sentiment Analysis
  9. Python Library used in Sentiment analysis

# 3.Introduction

With advancements in technology and fields like deep learning, sentiment analysis is becoming more and more common for companies that want to gauge their customers’ sentiments.

Today, businesses use natural language processing, statistical analysis, and text analysis to identify the sentiment and classify words into positive, negative, and neutral categories.

The best companies understand the importance of understanding their customers’ sentiments – what they are saying, what they mean and how they are saying. You can use sentiment analysis to identify customer sentiment in comments, reviews, tweets, or social media platforms where people mention your brand.

As sentiment analysis is the domain of understanding emotions using software, we have prepared a complete guide to understand ‘what is sentiment analysis?’, its tools, and different classifications and use cases.

**What is Sentiment Analysis**

Sentiment analysis can be defined as analyzing the positive or negative sentiment of the customer in text. The contextual analysis of identifying information helps businesses understand their customers’ social sentiment by monitoring online conversations.

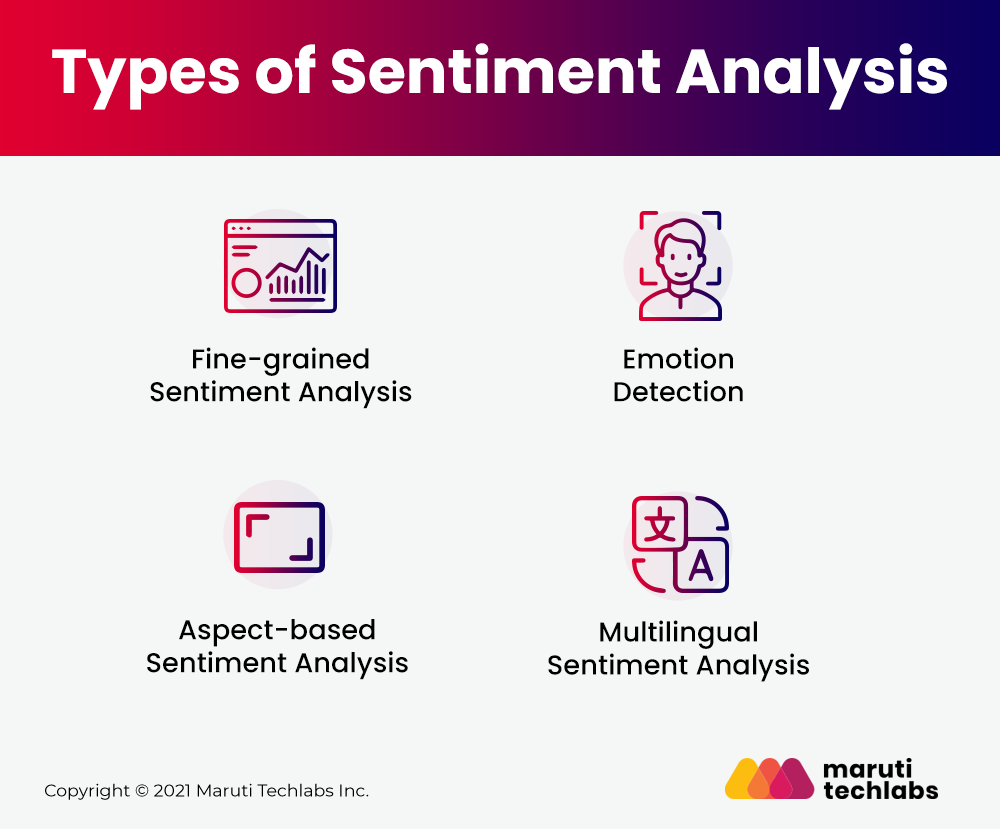
As customers express their reviews and thoughts about the brand more openly than ever before, sentiment analysis has become a powerful tool to monitor and understand online conversations. Analyzing customer feedback and reviews automatically through survey responses or social media discussions allows you to learn what makes your customer happy or disappointed. Further, you can use this analysis to tailor your products and services to meet your customer’s needs and make your brand successful.

Recent advancements in machine learning and deep learning have increased the efficiency of sentiment analysis algorithms. You can creatively use advanced [artificial intelligence and machine learning](https://marutitech.com/artificial-intelligence-and-machine-learning/) tools for doing research and draw out the analysis.

For example, sentiment analysis can help you to automatically analyze 5000+ reviews about your brand by discovering whether your customer is happy or not satisfied by your pricing plans and customer services. Therefore, you can say that the application of sentiment is endless.

**Types of Sentiment Analysis**

The sentiment analysis process mainly focuses on polarity, i.e., positive, negative, or neutral. Apart from polarity, it also considers the feelings and emotions(happy, sad, angry, etc.), intentions(interested or not interested), or urgency(urgent or not urgent) of the text.



# 

The sentiment analysis process mainly focuses on polarity, i.e., positive, negative, or neutral. Apart from polarity, it also considers the feelings and emotions(happy, sad, angry, etc.), intentions(interested or not interested), or urgency(urgent or not urgent) of the text.

Depending on how you interpret customer feedback, you can classify them and meet your sentiment analysis. However, below are some of the popular sentiment analysis classifications:

### ****1. Fine-grained Sentiment Analysis****

If your business requires the polarity precisions, then you can classify your polarity categories into the following parts:

* Very positive
* Positive
* Neutral
* Negative
* Very Negative

For polarity analysis, you can use the 5-star ratings as a customer review where very positive refers to a five-star rating and very negative refers to a one-star rating.

### ****2. Emotion Detection****

This type of sentiment analysis helps to detect customer emotions like happiness, disappointment, anger, sadness, etc. Here, you can use sentiment lexicons or complex machine learning algorithms to identify the customer’s feelings.

One of the disadvantages of using sentiment lexicons is that people tend to express emotions in different ways. So, it may be confusing to understand human emotion clearly while using it.

### ****3. Aspect-based Sentiment Analysis****

Let’s say that you are analyzing customer sentiment using fine-grained analysis. You want to identify the particular aspect or features for which people are mentioning positive or negative reviews. Here, aspect-based sentiment analysis comes into play.

For instance, in the review “The camera quality of this phone is getting worse with time,” an aspect-based classifier will determine that the review expresses a negative opinion from the customer for the phone’s camera feature.

### ****4. Multilingual Sentiment Analysis****

Multilingual sentiment analysis is complex compared to others as it includes many preprocessing and resources available online (i.e., sentiment lexicons). Businesses value the feedback of the customer regardless of their geography or language. Therefore, multilingual sentiment analysis helps you identify customer sentiment irrespective of location or language difference.

## Importance Of Sentiment Analysis

The most crucial advantage of sentiment analysis is that it enables you to understand the sentiment of your customers towards your brand. Your products and services can be improved, and you can make more informed decisions by automatically analyzing the customers’ feelings and opinions through social media conversations, reviews, surveys, and more.

According to the survey, 90% of the world’s data is unstructured. Especially in businesses, emails, tickets, chats, social media conversions, and documents are generated daily. Therefore, it is hard to analyze all this vast data in a timely and efficient manner.

Let us look at the overall benefits of sentiment analysis in detail:

### ****Sort Data at Scale****

There is too much business data to analyze daily. Can you imagine sorting all these documents, tweets, customer support conversations, or surveys manually? Sentiment analysis will help your business to process all this massive data efficiently and cost-effectively.

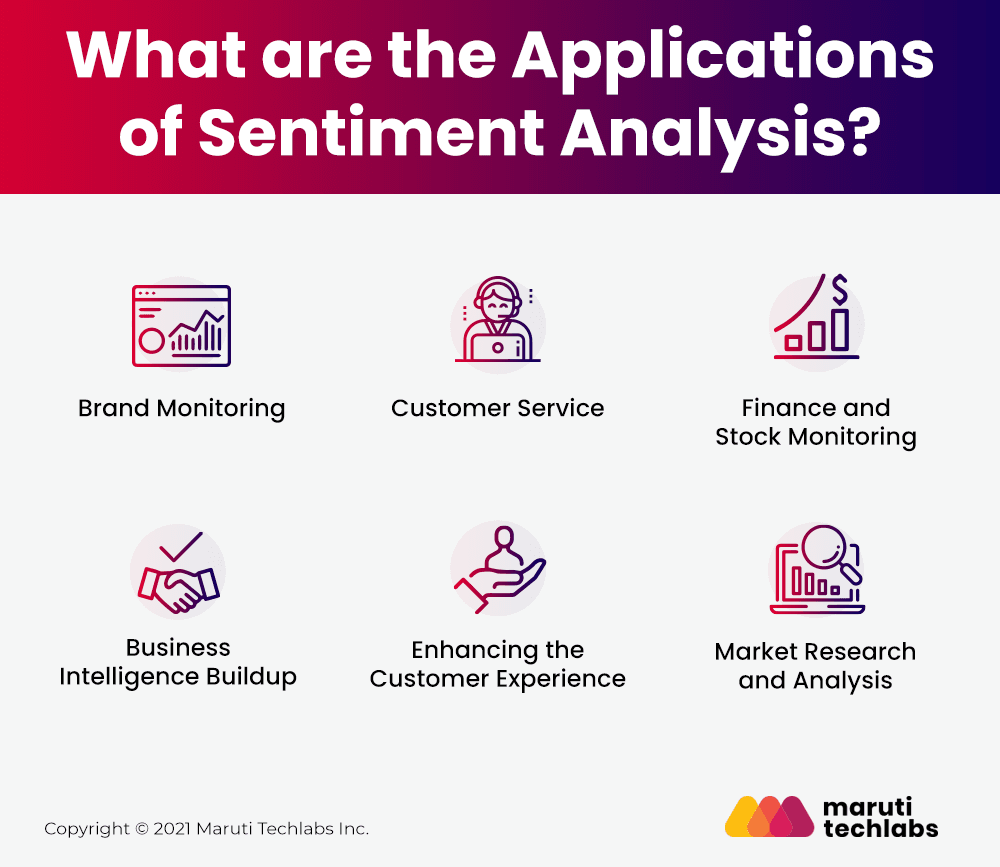
### ****Real-Time Analysis****

Is your angry customer about to churn? Is a PR crisis on social media escalating? Sentiment analysis will help you handle these situations by identifying critical real-time situations and taking necessary action right away.

### ****Consistent Criteria****

According to research, customers only agree for 60-65% while determining the sentiment of the particular text. Tagging text is highly subjective, influenced by thoughts and beliefs, and also includes personal experience. Therefore, you can apply criteria and filters to all your data, improve their accuracy, and gain better insights using sentiment analysis.

## Application Of Sentiment Analysis



## Python Library Used in Sentiment Analysis

## [****1. NLTK****](http://www.nltk.org/)

It stands for **Natural Language Tool Kit**. It is the most popular as well as most useful library for performing Sentiment Analysis.

Released in **2001** and Developed by **University of Pennsylvania.**

NLTK library includes name entity **recognition, tokenizing, part-of-speech (POS) and Topic Segmentation.** NLTK also boasts a good selection of third-party extensions, as well as the most wide-ranging language support of any of the other library.

## ****2.****[****TextBlob****](https://textblob.readthedocs.io/en/dev/)

It is a lightweight Python library which supports both **Python 2**and**Python 3**for sentiment analysis development.

Released in **2013**and Developed by **Steven Loria.**

TextBlob has a rule-based integrated sentiment analysis function with two properties:-

**A. Subjectivity**

**B. Polarity.**

Workflows with TextBlob and VADER (Valence Aware Dictionary and sEntiment Reasoner) are among the most popular approaches to sentiment analysis with TextBlob.

## [****3. SpaCy****](http://www.spacy.io/)

**SpaCy** is a multi-platform environment that runs on **Cython,** a superset of Python that enables the development of fast-executing C-based frameworks for Python. Consequently, SpaCy is the fastest-running solution at the moment according to research by Jinho D. Choi et.al.

Released in **2015** and Developed by **Exlosion AI**

Unlike **NLTK, SpaCy** is focused on industrial usage and maintains a minimal effective toolset, with updates superseding previous versions and tools, in contrast to NLTK. SpaCy’s pre-built models address essential NLP sectors such as named entity recognition, part-of-speech (POS) tagging and classification.

## Conclusion

The era of getting valuable insights from surveys and social media has peaked due to the advancement of technology. Therefore, it is time for your business to be in touch with the pulse of what your customers are feeling. Companies are using intelligent classifiers like contextual semantic search and sentiment analysis to leverage the power of data and get the deepest insights.

Formulate business strategies, exceed customer expectations, generate leads, build marketing campaigns, and open up new avenues for growth through [natural language processing solutions](https://marutitech.com/natural-language-processing-services/).

[Maruti Techlabs’](https://marutitech.com/) developers help you model human language and recognize the underlying meaning behind the words said or the action performed. We take communication beyond words and help to interpret human language and behavior.

Are you looking to interpret customer sentiments for increasing brand value? Drop us a note [here](https://marutitech.com/contact-us/), and we’ll take it from there.