A Mini Project Synopsis on

CRITIQUE ANALYZER

T.E. - D.S Engineering

Submitted By

Pooja Kumbhar 21107015

Urvi Padelkar 21107054

Pranjal Desai 21107062

Dhanashree Kasar 21107037

Under The Guidance Of

Prof. Poonam Pangarkar

DEPARTMENT OF CSE (DATA SCIENCE)

A.P.SHAH INSTITUTE OF TECHNOLOGY

G.B. Road, Kasarvadavali, Thane (W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic year : 2023-24

CERTIFICATE

This to certify that the Mini Project report on **Critique Analyzer** has been submitted by Pooja Kumbhar(21107015) Urvi Padelkar(21107054) Pranjal Desai(21107062)Dhanashree Kasar(21107037) who are a Bonafede students of A. P. Shah Institute of Technology, Thane, Mumbai, as a partial fulfilment of the requirement for the degree in **CSE(DATA SCIENCE)**, during the academic year **2023-2024** in the satisfactory manner as per the curriculum laid down by University of Mumbai.

Prof. Poonam Pangarkar
Guide

Prof. Anagha Aher
Head Department of CSE(Data Science)

Dr. Uttam D.Kolekar
Principal

External Examiner(s)
1.
2.

Place: A.P.Shah Institute of Technology, Thane
Date:

TABLE OF CONTENTS

1. Introduction

1.1. Purpose.....1

1.2. Objectives 1

1.3. Scope..... 2

2. Problem Definition 3

3. Proposed System..... 4

3.1. Features and Functionality.....4

4. Project Outcomes.....5

5. Software Requirements 6

6. Project Design.....7

7. Project Scheduling.....10

8. Conclusion.....11

References

Acknowledgement

CHAPTER 1

INTRODUCTION

In recent studies it has been found that 70% of users read and consult an online review before making a purchase. 90% of consumers read online reviews before visiting a business. And 88% of consumers trust online reviews as much as personal recommendations. Reviews have become a validator for users to actually consider to buy the product. Online reviews have an incredible power to shape consumer behaviour. Depending upon the size of the business and the volume of reviews it receive, the businesses need to optimize, prioritize, and be very selective in evaluating which reviews to respond to, and how. For large organizations with numerous locations, and a high volume of reviews, the objective may seem impossible. Having a clear and defined strategy is a must for tracking and managing of those reviews and gains a strong sense of trust with the customers.

1.1 Purpose

In the proposed system, the idea is to mine and summarize all the reviews. We have classified all those reviews based on product quality, service related issues and comment. The system provides a statistical report of the product. We have used Support Vector Machine for modelling and prediction of the review dataset. This would enable business to identify frequently reported issues and the trend of the issues. This knowledge would ensure enhanced product and support quality. In a process to derive the target outcome from the unstructured raw text, the first step is to identify suitable data source.

A critique analyzer is a tool or system designed to analyze and process reviews, typically found online, such as product reviews, restaurant reviews, movie reviews, or any other type of user-generated feedback. The purpose of a critique analyzer is to extract valuable insights and information from these reviews to help individuals, businesses, or organizations make informed decisions and improve their products, services, or offerings.

1.2 Objective

Review analysis is like tidying up jumbled-up customer opinions so we can use them to make smarter choices. One big way we do this is by looking at what customers say

about the things they want in a product. It's like listening to them to figure out how to make the product better. We aim to give honest and detailed product reviews that help people make smart buying choices. At the same time, we want to help companies make their products and services better by giving them useful feedback.

- To provide thorough, unbiased, and informative product reviews that empower consumers to make well-informed purchasing decisions, while also assisting companies in improving their products and services through constructive feedback.
- To automatically analyze customer reviews and feedback.
- To determine sentiment expressed in reviews (positive, negative, neutral).
- To identify key topics and issues mentioned in reviews.
- To offer real-time or near-real-time analysis for prompt responses.
- To integrate data from multiple online platforms.
- To provide customization options for businesses' specific needs.
- To ensure high accuracy in sentiment analysis.
- To maintain scalability for handling increasing data volumes.

1.3 Scope

1.Data Collection: Gathering product reviews from various sources, including e-commerce websites, social media platforms, forums, and blogs as a csv file.
Compiling and categorizing reviews based on product type, brand, and other relevant factors.
2.Sentiment Analysis: Analyzing the sentiment of reviews to determine whether they arepositive, negative, or neutral.
Assigning sentiment scores to reviews to quantify the degree of positivity or negativity.
3.Feature Extraction:Identifying and extracting key features, attributes, and characteristics mentioned in the reviews. Grouping reviews based on common product features for comparative analysis. To offer real-time or near-real-time analysis for prompt responses.
4.Customer Feedback Analysis: Analyzing customer feedback to identify common pain points, complaints, and areas of satisfaction. Presenting actionable insights to product manufacturers for improvements.

5. Problem Identification: The phrases are identified manually after reading various review comments for main categories namely Product Related and Support Related and few of them are selected.

CHAPTER 2

PROBLEM DEFINATION

In today's data-driven world, the proliferation of user-generated content and online reviews across a multitude of platforms presents a formidable challenge. We find ourselves in a situation where we are inundated with an unprecedented volume of textual and numerical data in the form of product reviews, customer feedback, and ratings. While this abundance of information holds immense potential for our business, it also poses a daunting problem: how to effectively harness and analyze this wealth of data to derive actionable insights.

The problem we face is multifaceted. First, the sheer volume of reviews, spanning various platforms such as e-commerce websites, social media, forums, and third-party review sites, overwhelms our current capacity to process and make sense of it all. This data deluge leaves us unable to promptly identify emerging trends, customer sentiments, or critical issues affecting our products and services.

The problem we face in review analysis is threefold: the overwhelming volume of reviews, the diversity of data sources and formats, and the inadequacy of our current analysis methods. By implementing a comprehensive, data-driven review analysis solution, we can turn this problem into an opportunity, gaining valuable insights, improving our products and services, and ultimately strengthening our position in the market.

CHAPTER 3

PROPOSED SYSTEM

The proposed review analysis system is designed to tackle the challenges outlined in the problem definition comprehensively. This advanced platform employs state-of-the-art natural language processing (NLP) and machine learning (ML) techniques to systematically collect, preprocess, and analyze customer reviews from diverse online sources. The system's key components include data collection and aggregation from various platforms, robust text preprocessing to standardize data, advanced sentiment analysis models to capture nuanced emotions, and aspect-based analysis to identify specific features driving sentiment. Additionally, the system incorporates topic modeling to uncover recurring themes and custom taxonomy development for structured categorization. To empower decision-makers, the system generates interactive dashboards and reports for data visualization, enabling them to gain actionable insights, track trends, and prioritize improvements effectively. This proposed system offers a comprehensive solution to transform the challenge of review analysis into an opportunity for enhancing customer satisfaction and competitive advantage.

TextBlob is a Pre-trained model: Pre-trained model libraries for sentiment analysis namely Vader, Textblob, and Flair can be used to reduce this burden. This pre-trained model can be used for the process of labeling training data in the supervised learning method approach.

TextBlob uses NLTK (Natural language toolkit) because it is simple, easy to deploy. It can be used for complex analysis and working with textual data. TextBlob returns polarity and subjective of a sentence. It is easy to learn and offers a lot of features like sentiment analysis, postagging, non phrase extraction.

3.1 Features And Functionality :-

- Ability to collect reviews from various platforms, including e-commerce websites, social media, forums, and third-party review sites.
- Clean and preprocess raw text data, removing noise, formatting issues, and irrelevant symbols.
- Determine the sentiment of each review (positive, negative, neutral).
- Identify specific aspects or features mentioned in reviews (e.g., product quality, customer service).
- Keyword Analysis involves finding and listing the most commonly used words and phrases.
- Ensuring the tool can handle a large volume of reviews efficiently.
- Generating charts, graphs, and infographics to visualize review data. Displaying sentiment trends, feature popularity, and topic distribution.

CHAPTER 4

PROJECT OUTCOMES

The project outcome for a critique analyzer project is the development of a software application or system designed to analyze and process text-based reviews from a multitude of sources, including social media, product reviews, and various platforms. The primary objective is to extract valuable insights and information from these reviews, ultimately aiding businesses in making informed decisions and enhancing their products and services. Key features of such a project include sentiment analysis, where the system classifies text as positive, negative, or neutral, providing businesses with a comprehensive understanding of customer sentiment. The system must efficiently collect, preprocess, and clean the vast amounts of review data it gathers, ensuring that it is ready for analysis. To cater to a global audience, language support for multiple languages is crucial. Customization options allow businesses to tailor sentiment analysis models to align with their specific industry or brand terms, enhancing accuracy. Scalability and real-time analysis capabilities are vital for handling varying workloads and providing immediate insights into customer sentiment as reviews are posted. Effective data visualization, including graphical representations, is essential for businesses to grasp customer feedback easily. The specific goals and outcomes of such a project can vary, but here are some common objectives and outcomes for a critique analyzer project:

- Sentiment Analysis Engine: The core of the project would be a sentiment analysis engine that processes user-provided sentences and determines their sentiment. You can use Natural Language Processing (NLP) techniques, machine learning models, or pre-trained models like TextBlob to perform sentiment analysis.
- Sentiment Labels: For each input sentence, your website should provide a sentiment label (positive, negative, or neutral). You can also assign a numerical score to indicate the strength of the sentiment.
- Graphical Visualization: To enhance user experience, you can present the sentiment analysis results in the form of graphs. A bar chart would display the distribution of positive, negative and neutral sentiments for a given set of sentences. This graphical representation can make it easier for users to grasp the overall sentiment trends. As well as from the given set of sentence or csv file we can label them as per the topic

- Topic or Keyword Analysis: Topic or keyword analysis allows users to focus sentiment analysis on specific subjects or terms within a larger body of text. Users can extract valuable insights by specifying keywords or topics of interest. For instance, in a lengthy news article, users might want to sentiment only around certain names, companies, or events mentioned within the text. This feature empowers users to tailor sentiment analysis to their specific needs, providing more targeted and actionable insights. It's useful for market research, content curation, and tracking the perception of particular entities or subjects.
- Trend Analysis: This involves collecting and analyzing reviews or feedback data at different time intervals to identify patterns, shifts, and emerging trends. Businesses can use this information to make data-driven decisions, adapt their strategies, and respond to changing customer sentiment. It's particularly useful for monitoring the impact of marketing campaigns, product launches, or policy changes on public perception.

CHAPTER 5

SOFTWARE REQIUREMENTS

The Software Requirements Specification is produced at the culmination of the analysis task. The function and performance allocated to software as part of system engineering are refined by establishing a complete information description, a detailed functional and behavioral description, an indication of performance requirements and design constraints, appropriate validation criteria, and other data pertinent to requirements.

Operating System	Windows11
Coding Language	HTML, CSS,Python
Tool	Pycharm Community Edition 2022.3.2
Front End	Pycharm 22.3.2,HTML,Python3.11
Backend	Textblob 0.17.1,Flask 2.3.3,Django 4.2.5
Algorithm used	Sentiment Analysis(NLP,LDA)

- Operating System(OS): Windows11: This project can be developed and run on computers using any of these Windows operating systems.
- Front End:
  - Pycharm 22.3.2 : PyCharm is an integrated development environment (IDE) for the Python programming language. It is developed by JetBrains and is one of the most popular and widely used IDEs for Python development. PyCharm provides a comprehensive set of tools and features to assist Python developers in writing, testing, and debugging Python code more efficiently.
  - HTML : (Hypertext Markup Language) is not a programming language but rather a markup language used to structure and format content on the web. HTML is used to create the structure of web pages by defining elements and their relationships, such as headings, paragraphs, lists, links, images, forms, and more. These elements are marked up using HTML tags..
  - Python 3.11

- Backend:
    - Textblob 0.17.1 : TextBlob is a Python library and API for processing textual data. It is built on the NLTK (Natural Language Toolkit) and Pattern libraries and provides a simple and intuitive interface for common natural language processing (NLP) tasks.
    - Flask 2.3.3 : Flask is a lightweight and popular web framework for Python that is used to build web applications, APIs, and web services. It is known for its simplicity, flexibility, and ease of use, making it a great choice for developers, especially those who are new to web development.
    - Django 4.2.5 : Django is a high-level, open-source Python web framework that is designed to make web development faster, easier, and more robust.

Algorithm :

- Sentiment Analysis:

  Sentiment analysis is the process of analyzing digital text to determine if the emotional tone of the message is positive, negative, or neutral.

CHAPTER 6

PROJECT DESIGN

The review analyzer system provides businesses with the tools needed to make data-driven decisions. By performing sentiment analysis, aspect-based analysis, and trend analysis, it empowers organizations to understand customer sentiment, identify areas for improvement, and capitalize on strengths.

The project has streamlined the process of extracting valuable insights from a large volume of reviews. Automation and customizable reporting make it easier for businesses to access the information they need without manual and time-consuming analysis.

- Home Page :

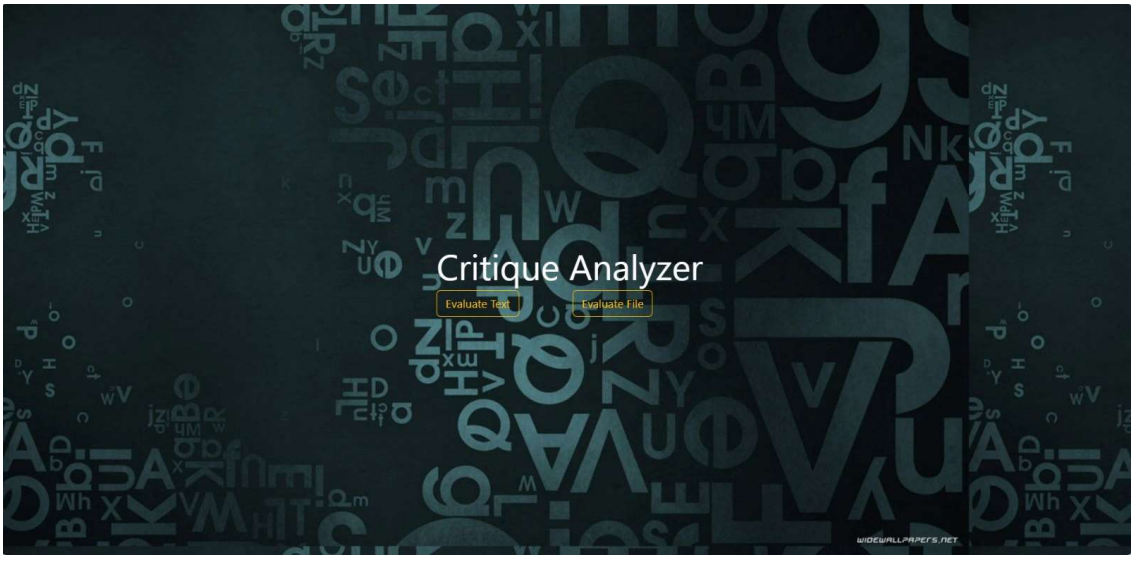


Fig 6.1 Home Page

In the homepage of our Critique Analyzer project, users are presented with two primary options: "Evaluate Text" and "Evaluate File." "Evaluate Text" allows users to input or paste text directly for real-time sentiment analysis, making it a quick and convenient choice for assessing individual comments, social media posts, or short reviews. On the other hand, "Evaluate File" provides the ability to upload entire documents or files containing multiple reviews, enabling users to analyze larger volumes of text efficiently.

- Evalute Text :

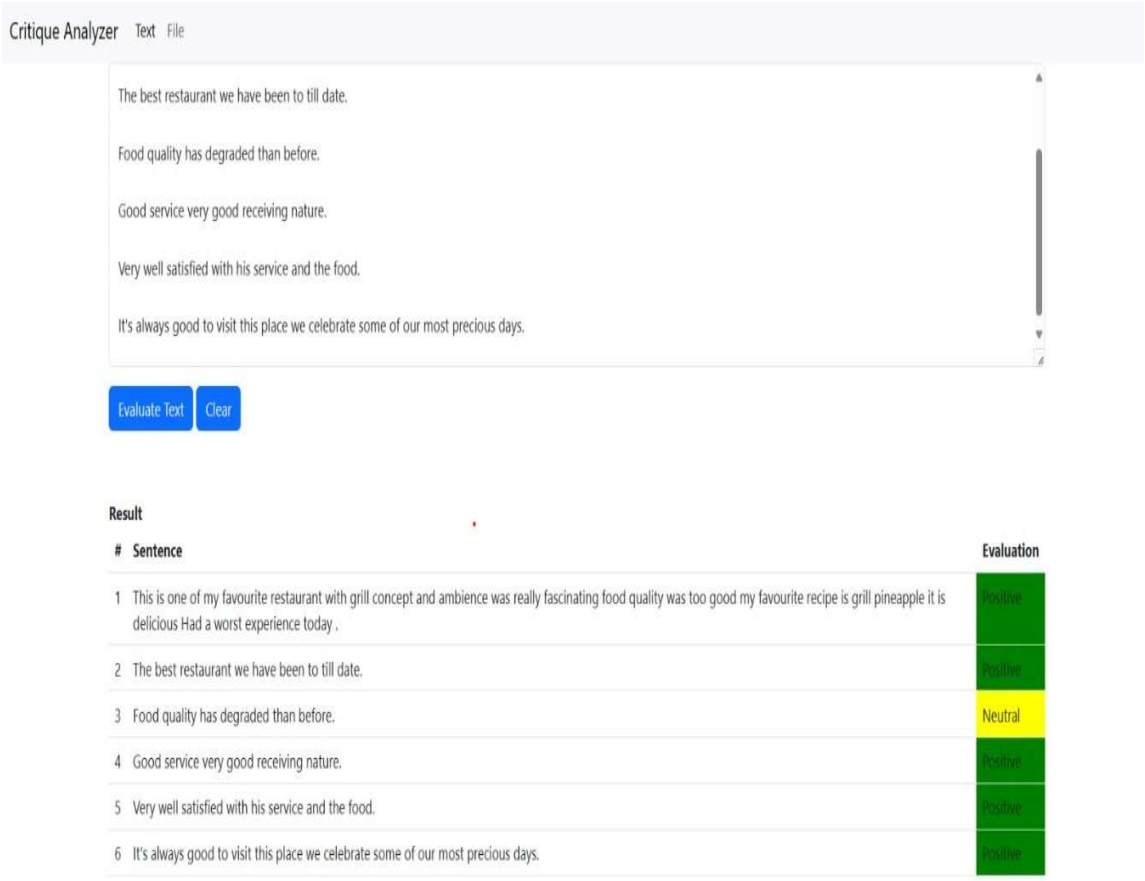


Fig 6.2 Text Page

When users choose the "Evaluate Text" option from the home page, they are directed to a dedicated text input box where they can type or paste their review, comment, or any text they want to analyze. After inputting the text, clicking on the "Evaluate Text" button initiates the sentiment analysis process. The system then swiftly assesses the content and categorizes it into one of three possibilities: Positive, Negative, or Neutral.

- Evalute File :

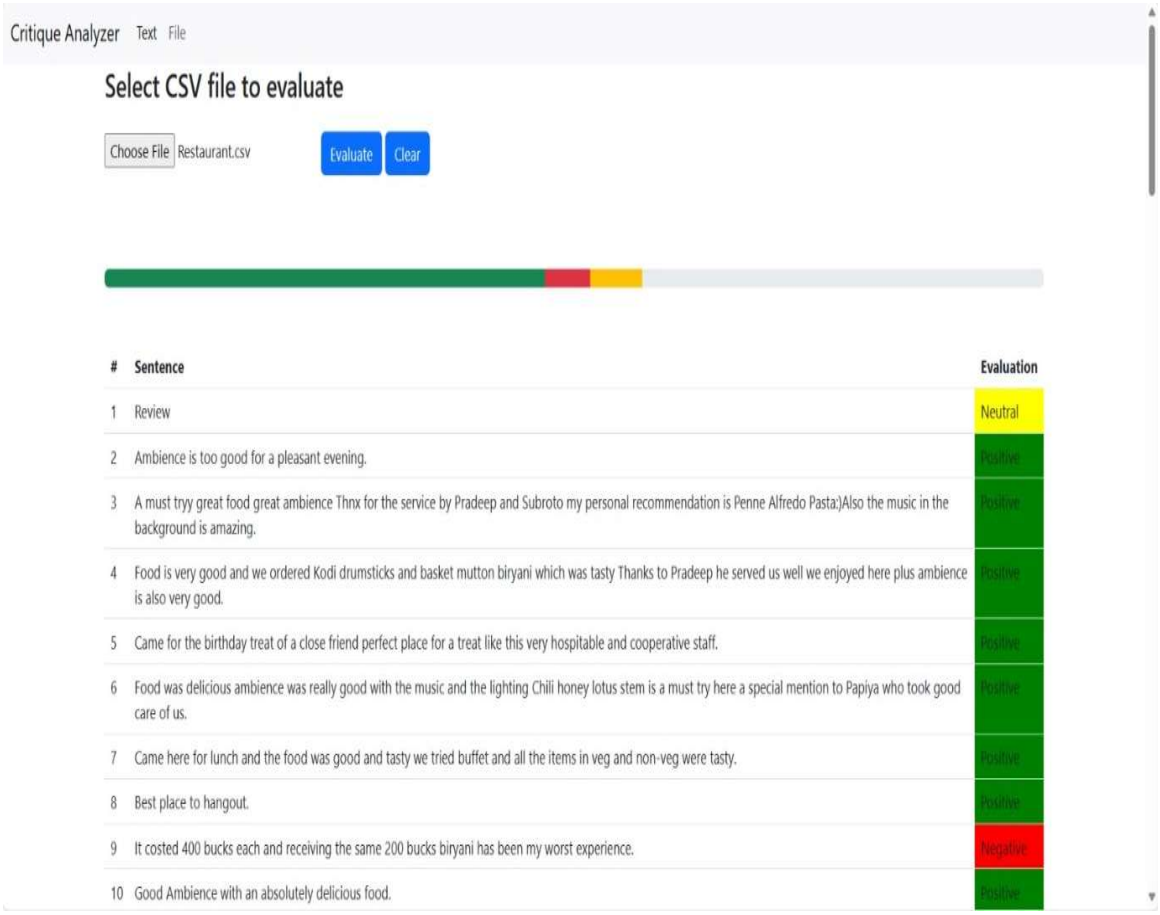


Fig 6.3 File Page

In the "Evaluate File" section accessed from the homepage, users are presented with the option to upload review files in CSV format or with a CSV extension. Once the file is uploaded, the system processes the text data within the file and conducts sentiment analysis on the contents. It then categorizes each review within the file into one of three possibilities: Positive, Negative, or Neutral and also with graph for a quick summary of data. This feature is especially useful for users who need to analyze large volumes of reviews or comments at once, streamlining the process and providing quick insights into overall sentiment trends within the uploaded document.

- Tokenization :

.

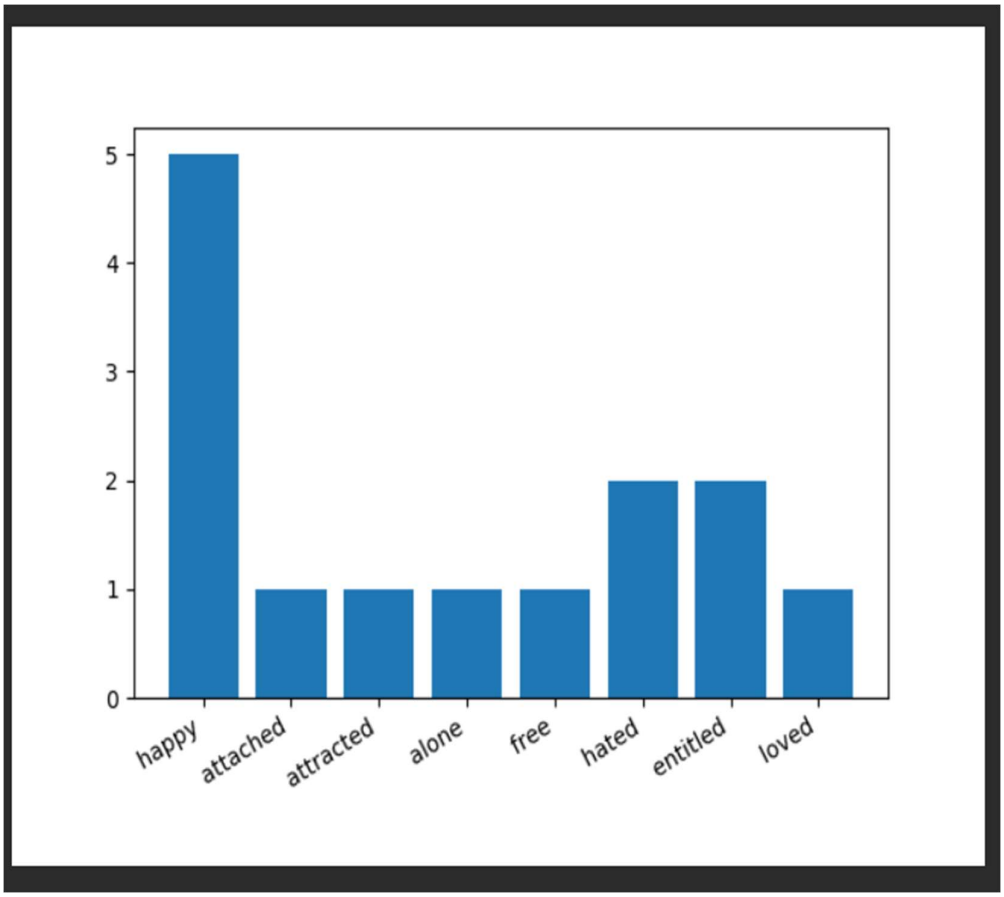


Fig 6.4 Tokenization output

In Tokenization analysis, the system dissects the text, breaking it down into individual words, phrases, or tokens, and then assigns a numerical or categorical value to each token based on factors like sentiment, emotion, or other relevant criteria. These values are subsequently aggregated and visualized in graph format, enabling users to gain a more granular understanding of the sentiment or emotional nuances within the text. Graphical representation makes it easier to identify patterns, trends, and key insights within the review data, facilitating more informed decision-making and a deeper understanding of the text's underlying sentiment.

- Topic Modeling and Labeling :

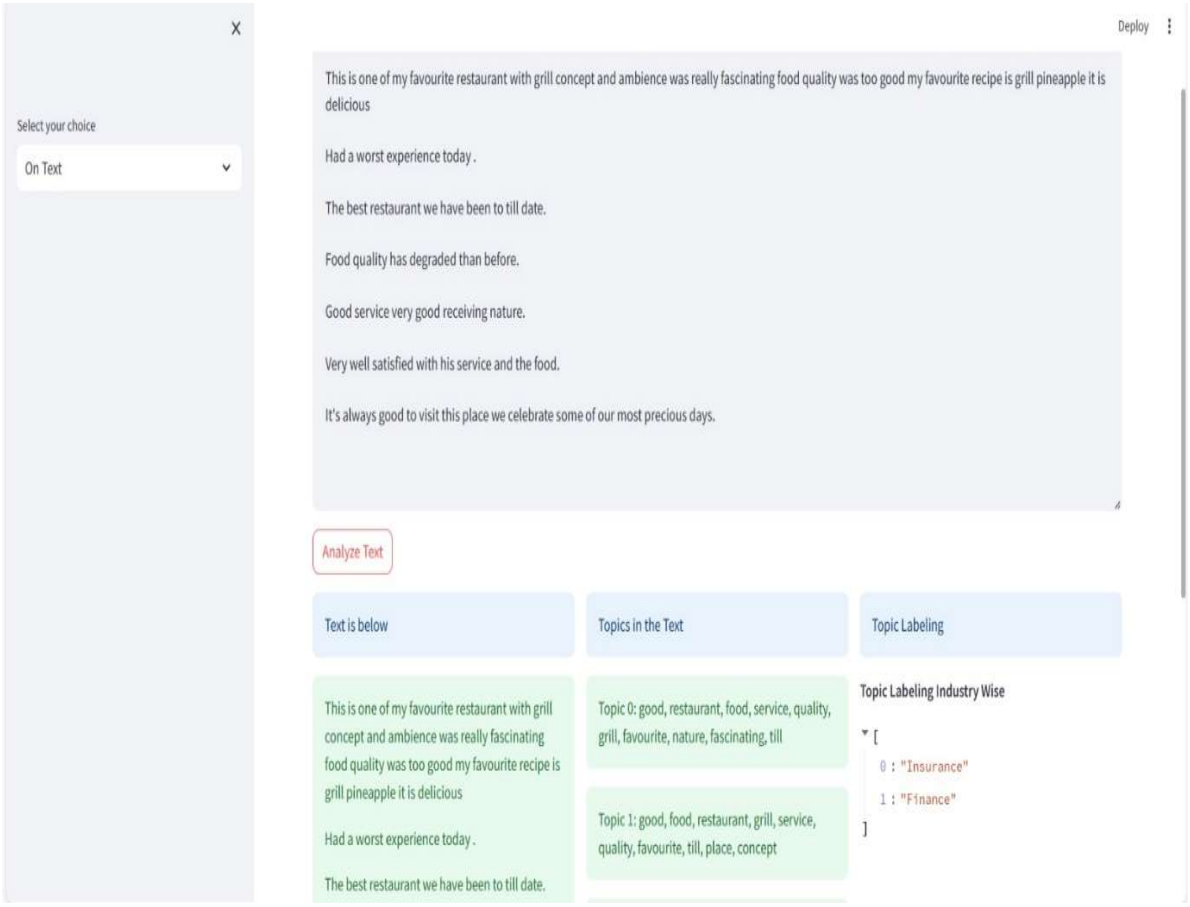


Fig 6.5 Topic Modeling

In the context of Topic Modeling and Labeling, the system conducts a comprehensive analysis of the text to identify and extract the primary themes or topics discussed within the review. It then assigns relevant labels to these topics, effectively categorizing the content based on its subject matter. This feature empowers users to quickly grasp the central themes and subjects in the review, making it a valuable tool for content classification, trend identification, and content organization. The assigned labels enhance the organization and accessibility of the review data, aiding users in their decision-making processes.

CHAPTER 7

PROJECT SCHEDULING

Scheduling in this project management is the listing of activities, deliverables, and milestones within a project. A schedule also usually includes a planned start and finish date, duration, and resources assigned to each activity. Effective project scheduling is a critical component of successful time management, especially for professional service businesses.

Sr.No	Group Member	Time Duration	Work To Be Done
1	Pooja Kumbhar Urvi Padelkar Pranjal Desai Dhanashree Kasar	3 rd week of July	Group formation and Topic finalization. Identifying the scope and objectives of the Mini Project. Discussing the project topic with the help of a paper prototype.
		1 st week of August	Identifying the functionalities of the Mini Project. Designing the Graphical User Interface (GUI).
2	Urvi Padelkar Pranjal Desai	3rd week of August	Model trained

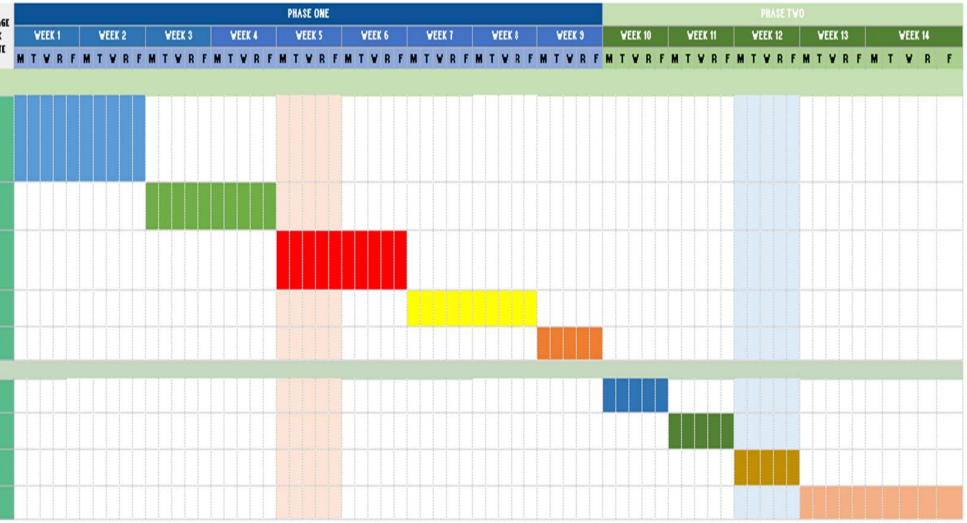
3	Pooja Kumbhar Urvi Padelkar	1 stweek of September	GUIs Connectivity.
4	Pranjal Desai Dhanashree Kasar	1 st week of October	Integration of all modules and Report Writing.

GANTT CHART TEMPLATE

Smartsheet Tip → A Gantt chart's visual timeline allows you to see details about each task as well as project dependencies.

PROJECT TITLE	Critique Analysis	INSTUTUTE & DEPART	AP SHAH INSTITUTE OF TECHNOLOGY(Informatic
PROJECT GUIDE	Prof.Poonam Pangarkar	DATE	10-26-23

WBS NUMBER	TASK TITLE	TASK OWNER	START DATE	DUE DATE	DURATION(Wo.. kr)	PERCENTAGE OF TASK COMPLETE
1	**Project Conception and Initiation**					
1.1	Group formation and Topic finalization. Identifying the scope and objectives of the Mini Project	1.1- Pooja pranjal Dhanashree Urvi	7-26-23	8-3-23	2	100%
1.2	Identifying the functionalities of the Mini Project	1.2- Pooja Dhanshree	8-10-23	8-16-23	1	100%
1.3	Discussing the project topic with the help of paper prototype.	1.3 Pranjal urvi	8-16-23	8-23-23	1	100%
1.4	Designing the Graphical User Interface(GUI)	1.4 Urvi Dhanshree Pranjal Pooja	8-23-23	9-13-23	3	100%
1.5	Presentation I	1.5- Urvi Pooja Pranjal Dhanashree	9-19-23	9-23-23	1	100%
2	**Project Design and Implementation**					
2.1	Database Design	2.1- Pooja Pranjal Urvi Dhanashree	9-26-23	9-27-23	1	100%
2.2	Database Connectivity of all modules	2.2- Pooja Pranjal	9-27-23	10-4-23	1	100%
2.3	Integration of all modules and Report Writing	2.3- Dhanashree Urvi	10-4-23	10-11-23	1	100%
2.4	Presentation II	2.4 - Dhanashree Pooja Pranjal Urvi	10-26-23	10-26-23	2	100%



CHAPTER 8

CONCLUSION

In conclusion, the review analyzer project has successfully addressed the critical need for businesses to gain valuable insights from customer feedback. This project's main objectives were to develop a sophisticated software system capable of analyzing and processing text-based reviews from various sources. Here are the key takeaways and conclusions from this project: The review analyzer system provides businesses with the tools needed to make data-driven decisions. By performing sentiment analysis, aspect-based analysis, and trend analysis, it empowers organizations to understand customer sentiment, identify areas for improvement, and capitalize on strengths.

Project has streamlined the process of extracting valuable insights from a large volume of reviews.Integration and User-Friendly Interface Integration with other systems and platforms enhances the project's utility, while a user-friendly interface simplifies interactions and ensures that users can easily access and interpret reports.

Competitive Advantage: With the capabilities provided by the review analyzer, businesses gain a competitive advantage in today's data-driven market. They can respond to customer feedback more effectively, refine their offerings, and stay ahead of the competition.

REFERENCES

[1] Pankaj Verma, Akib Mohi-Ud-Din Khanday, Syed Tanzeel Rabani, Mahmood Hussain Mir Sanjay Jamwal," Twitter Sentiment Analysis on Indian Government Project", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-3, September 2019.

[2] Francisco Jurado, Pilar Rodriguez,"Sentiment Analysis in monitoring software development processes",Journal of Systems and Software,Volume 104,June 2015.

[3] Sunil B. Mane, Yashwant Sawant, Saif Kazi, Vaibhav Shinde,"Real Time Sentiment Analysis of Twitter Data Using Hadoop", International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5 (3) , 2014.

[4] Marc Herrmann,Jil Klunder,"Applying Sentiment Analysis to a Software Project Meeting",IEEE International Requirements Engineering Conference Workshops (REW).

[5] Jesus Serrano-Guerrero,JoseA. Olivas,FranciscoP. Romero,Enrique HerreraViedma, "Sentiment analysis: A review and comparative analysis of web services",Information sciences,Volume 311, 1 August 2015.

[6] Claudia Iacob, Rachel Harrison," Retrieving and analyzing mobile apps feature requests from online reviews", 2013 10th Working Conference on Mining Software Repositories (MSR).

[7] Anthony R. Wheeler, Kristen K. Shanine, Matthew R. Leon, Marilyn V. Whitman," A review, analysis, and guidelines for future research",Journal of Occupational and Organizational Psychology, Volume 87, Issue 1, March 2014.

[8] Youtube link for Tokenization: https://youtu.be/dyN_WtjdfpA?si=kx6iqrra5RIme2Os

[9] Youtube link for Topic Modeling and Labeling: https://youtu.be/3VfW-UoA3M4?si=nGiN46EBahbQQ0P4

ACKNOWLEDGEMENT

This project would not have come to fruition without the invaluable help of our guide **Prof.Poonam Pangarkar**. Expressing gratitude towards our HoD, **Prof.Anagha Aher**, and **the Department of CSE(Data Science)** for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our teacher Prof. Poonam Pangarkar who gave us her valuable suggestions and ideas when we were in need of them. We would also like to thank our peers for their helpful suggestions.