# A WEB APPLICATION FOR RESTAURANT REVIEW ANALYSIS

*in partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in*

**INFORMATION TECHNOLOGY**

*Submitted by*

**Ankur Singh**

**Chaitanya Nawghare**

**Janhvi Badwaik**

**Rahel Gonjari**

**Under the guidance of**

**Prof. Vishal M. Tiwari**

**Assistant Professor**

**Academic Year 2023-24**

**Department of Information Engineering**


**ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING AND TECHNOLOGY**

Wardha Road, Gavsi Manapur, Nagpur

# ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING AND TECHNOLOGY -NAGPUR

**DEPARTMENT OF INFORMATION TECHNOLOGY**

## CERTIFICATE

Certified that this project report "**A WEB APPLICATION FOR RESTAURANT REVIEW ANALYSIS**" is the bonafide work of "**ANKUR SINGH, CHAITANYA NAWGHARE, JANHVI BADWAIK, RAHEL GONJARI"** who carried out the project work under my supervision in partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in **INFORMATION TECHNOLOGY** of RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR

| | |
|---|---|
| Dr. Manoj V. Bramhe | Prof. Vishal M. Tiwari |
| Professor | Assistant Professor |
| **HEAD OF THE DEPARTMENT** | **GUIDE** |

**PRINCIPAL**

**ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING AND TECHNOLOGY**

# ACKNOWLEDGEMENT

Our major project is titled, **"A WEB APPLICATION FOR RESTAURANT REVIEW ANALYSIS"**. Any project requires a lot of hard work, sincerity and systematic work methodologies. We express our deepest gratitude to our Project Guide, **Prof. Vishal M. Tiwari**, for giving us an opportunity to be a part of this project and for the guidance, provided throughout the span of our project .

We owe our gratitude to the industry mentor **Mr. Yogesh Joshirao and Mr. Ketan Paithankar, Konverge.AI, Nagpur** for their immense support and guidance throughout the project.

Our special thanks to **Ms. Chhavi Rajput**, an alumni mentor, for her proactive guidance of the student projects.

We would also like to thank **Dr. Manoj V. Bramhe, Head of the Department of Information Technology** and all our faculty members who regularly evaluated our major project and pointed out the shortcomings in the project. They also gave us important feedback for the further improvement of our project. We are highly indebted to them.

We are also grateful to the **Management of the College, Dr. Surendra V. Gole, Principal** for the overwhelming support in providing us the online facilities and other required resources. We would like to thank our Library Department for providing us useful books and the research articles related to our project in online mode.

**PROJECT GROUP MEMBERS :**

**ANKUR SINGH**
**CHAITANYA NAWGHARE**
**JANHVI BADWAIK**
**RAHEL GONJARI**

# ABSTRACT

In the era of the web, a huge amount of information is now flowing over the network. Since the range of web content covers subjective opinion as well as objective information, it is now common for people to gather information about products and services that they want to buy. However, since a considerable amount of information exists as text-fragments without having any kind of numerical scales, it is hard to classify their evaluation efficiently without reading full text. Here we will focus on extracting scored ratings from text fragments on the web and suggests various experiments in order to improve the quality of a classifier. Methodologies like Sentiment Analysis as Text Classification Problem, Sentiment analysis as Feature Classification with mathematical treatment are explored. Of late, the word of mouth opinions expressed online are more valuable as people visit the restaurant by seeing the reviews.

Many platforms enables users to search for a restaurant to dine-out and also let them to share their review and rating about the restaurant. These reviews and ratings help other users who are searching for restaurant to dine out to check what is good and what is bad about any particular restaurant which helps them to have a good meal outside. This application deals with business insights that can be derived from text / opinion mining of restaurant reviews shared on the Zomato by the foodies. The text data available is not structured and so to look for what a majority of the crowd is looking for when they dine can make good business sense for owners. Business owners have to modify their operations based on customer preferences, and there is no better way to understand the customer and what they need, feel and want changed than a review that has no personal agenda. Capturing the emotion of a customer who has written a review, by the choice of words they use in the review is an essential part of improving the overall customer experience at a restaurant. While business owners benefit from constant improvements done to the various aspects of the restaurants, customers who are willing to invest their time are hugely benefitted by online reviews. So, this project looks at how text used in a review can influence a potential costumer and, what is it that works or does not work for a business, why do people go frequently a certain restaurant, what is it that people look for in a restaurant.

# Table Of Content

# LIST OF TABLE

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

# 1. INTRODUCTION

## 1.1 OVERVIEW

A web application (web app) is an application program that is stored on a remote server and delivered over the internet through a browser interface. Web services are web apps by definition and many, although not all, websites contain web apps. A Web application is an advantageous form of software because the use of browsers allows the application to be compatible with most standard computers and operating systems. Moreover, the application does not take up memory on a computer's hard drive and is accessible from nearly any computer or device a person might use. Multiple users can even use the same application at the same time, allowing for simultaneous participation. Although Web applications always require a network connection, this limitation has lessened in importance as the Internet has become more and more ubiquitous. Any service offered over the Internet, by definition, is a form of Web application. Examples of Web applications therefore include online forms, shopping carts, video streaming, social media, games, and e-mail.

The food and hospitality industry has long operated under the assumption that offering good food and service is the key to attracting more customers. However, with the advent of science and technology, particularly the proliferation of online platforms, a wealth of data has become available, pointing towards new findings and opportunities. Nowadays, a significant number of consumers not only rate products online but also take the time to write reviews. Remarkably, around 88% of people trust online reviews, making platforms like Yelp and Google Reviews crucial for both customers and businesses, providing a means for interaction between the two parties. While reviews and ratings offer valuable information, there are substantial challenges in extracting relevant insights and predicting future trends through the analysis and correlation of existing data. Each day, thousands of restaurants and businesses undergo scrutiny through customer reviews. The primary objective of the proposed research is to enhance the user experience by systematically analyzing restaurant reviews and categorizing them into various aspects. This categorization aims to provide users with an easy-to-understand overview of a restaurant, addressing the current issue where businesses struggle to effectively utilize reviews for their improvement.

Restaurant reviews serve as vital sources of information for individuals seeking informed decisions about their dining experiences. These reviews, often accompanied by ratings, are reflective of customers' comfort levels and play a pivotal role in helping others decide whether to patronize a particular restaurant. Simultaneously, reviews offer valuable insights to restaurant owners, aiding in the development and enhancement of their businesses. However, the abundance of textual data on online review platforms like Yelp and Google Reviews poses a challenge in terms of navigation and comprehension. While reviews are inherently textual, many machine learning algorithms

operate on numerical data. Natural Language Processing (NLP) methodologies present a solution by enabling data analysts to apply machine learning and deep learning algorithms to textual datasets. This allows for the classification of reviews and the recommendation of the best restaurants. In the realm of recommended systems, three primary methods are commonly implemented: content-based, collaborative-based, and hybrid-based. Content-based methods involve analyzing the characteristics of the items being recommended and matching them with user preferences. Collaborative-based methods, on the other hand, rely on the preferences and behaviors of similar users. Hybrid-based methods combine elements of both content-based and collaborative-based approaches, offering a more comprehensive and effective recommendation system.

The proposed research aims to bridge the gap between the textual nature of online reviews and the numerical requirements of machine learning algorithms. By employing NLP methodologies and utilizing recommended system approaches, the study seeks to empower both consumers and businesses in the food and hospitality industry. Moreover, the findings of this research are anticipated to have applicability across various industries related to food and hospitality, further expanding the potential impact of the proposed methodologies.

## 1.2 PROBLEM STATEMENT

To classify user reviews of a restaurant into positive and negative reviews using Natural Language Processing

## 1.3 OBJECTIVES

This project focuses on determining the polarity of the emotion a sentence evokes using an input box. to put in place a system that automatically categorizes material as good or negative. Some of the objectives are as follows :

- Sentiment analysis is used to evaluate whether the general public has a good, negative, or neutral view towards the topic at hand. A pie chart is used to depict it.

- Provide users with a platform to easily access and analyze restaurant reviews, aiding in informed decision-making about dining choices.

- Enhance the overall user experience by presenting reviews in a structured and easily understandable format, facilitating quick comprehension of key information.

- Implement a categorization system to organize reviews based on specific aspects such as food quality, service, ambiance, allowing users to focus on areas of particular interest.

- Apply NLP techniques to convert textual reviews into analyzable data, enabling the use of machine learning algorithms for deeper insights.

- Assist restaurant owners in leveraging customer feedback by providing insights into areas of strength and improvement, ultimately supporting business development.

- Integrate a recommendation system that suggests restaurants based on user preferences, utilizing content-based, collaborative-based, or hybrid-based approaches.

- Develop an intuitive and user-friendly interface that simplifies navigation, ensuring accessibility for users of varying technological backgrounds.

- Prioritize the implementation of robust data privacy and security measures to protect user information and maintain trust in the application.

- Enable real-time updates of reviews and ratings, ensuring that users have access to the most current and relevant information about restaurants.

These objectives collectively aim to create a comprehensive and user-centric restaurant review analysis web application that benefits both consumers and businesses in the food and hospitality industry.

## 1.4. ORGANIZATION OF THE REPORT

The organization of the report is how the contents of the project are arranged systematically following certain good practices. A report structure for an organization is a method of arranging business divisions for reporting purposes.

Any project's main body consists of multiple chapters with titles, and each page inside these chapters is numbered following a certain numbering convention. Generally a report organization follows a sequence as :-(i) title and author (ii) acknowledgement (iii) abstract (iv) introduction (v) experimental procedures (vi) results (vii) discussion (viii) references. The usual way of presenting these chapters is:

Chapter 1 gives a brief introduction to the theme of the project. Here, in our project, this chapter gives an overview of what web application and what an Restaurant review Analysis.

Chapter 2 focuses on a review of the literature. This chapter focuses on what conclusions have been derived from the research papers, online forums, the previously done work, and the need for the new proposed system. We have learned and examined and used the different methods for learning the process of web based apps which have been proposed till now.

Chapter 3 comprises the proposed system. It states the features and functionalities of the system. Here we have used certain diagrams to explain the design and analysis of the system. This chapter also describes some important elements such as user reviews and the hardware and software requirements of the system. We have also made use of some flowcharts to explain the modules and their work properly.

Chapter 4 consists of the System release plan. Thus, this chapter mentions the details of the tasks performed by us in this project to date. It also describes, in brief, the implementation methodology used by us.

Chapter 5 gives an overview of test cases and types of testing. It describes the test case used in this project and also the types of testing methodology used to analyze the working of this system.

Chapter 6 is an important chapter of this report as it comprises the conclusions and results derived from this project. This chapter also gives a future scope of this project.

Chapter 7 being the last of the chapters of this report consists of all the references for this project, that is it mentions all the resources we have used as a reference for this project.

# CHAPTER 2

# REVIEW OF LITERATURE

# 2.REVIEW OF LITERATURE

## 2.1 INTRODUCTION

A literature review is an appraising report of knowledge found within the literature relating to your selected area of study. The review ought to describe, summarize, access, and clarify the literature. It should give a theoretical base for the research and help you to determine the nature of your research.

While conducting research, a literature review is a necessary part of the project as a result it covers all previous research done on the subject and sets the platform on which the present research is based. No new analysis will be taken seriously while firstly reviewing the previous research undone on the topic.

The existing work on sentiment analysis can be classified from different points of views: technique used, view of the text, level of detail of text analysis, rating level, etc. From a technical point of view, we identified machine learning, lexicon-based, statistical and rule-based approaches. The machine learning method uses several learning algorithms to determine the sentiment by training on a known dataset. The lexicon-based approach involves calculating sentiment polarity for a review using the semantic orientation of words or sentences in the review. The "semantic orientation" is a measure of subjectivity and opinion in text. The rule-based approach looks for opinion words in a text and then classifies it based on the number of positive and negative words. It considers different rules for classification such as dictionary polarity, negation words, booster words, idioms, emoticons, mixed opinions etc. A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation Statistical models represent each review as a mixture of latent aspects and ratings. It is assumed that aspects and their ratings can be represented by multinomial distributions and try to cluster head terms into aspects and sentiments into ratings. Another classification is oriented more on the structure of the text: document level, sentence level or word/feature level classification. Document-level classification aims to find a sentiment polarity for the whole review, whereas sentence level or word-level classification can express a sentiment polarity for each sentence of a review and even for each word. Our study shows that most of the methods tend to focus on a document-level classification. Most of the solutions on review classification consider only the polarity of the review (positive/negative) and rely on machine learning techniques. Solutions that aim a more detailed classification of reviews (e.g., three or five star ratings) use more linguistic features.

## 2.2 LITERATURE SURVEY

“Restaurant Review Analysis using NLP” [1] various Natural Language Processing (NLP) techniques for analyzing restaurant reviews. This could involve sentiment analysis, topic modeling, or other NLP methods to extract meaningful information from textual data. The authors discuss how they collected restaurant review data and the steps taken to preprocess the data. Data preprocessing is crucial for cleaning and structuring the data before applying NLP techniques. The paper delves into sentiment analysis to determine the sentiments expressed in the restaurant reviews. This involves classifying reviews as positive or negative to understand customer satisfaction levels.

“Sentiment Analysis on Restaurant Reviews” contributes to our understanding of customer sentiments in the restaurant industry. [2] The insights gained inform strategic decision-making for restaurant owners and marketers. It focuses on the major influencers of customer sentiments and those are food taste and service quality. In this paper 31% reviews were negative while rest 69% were positive.

In “Restaurant review Classification and Analysis” the consequences of challenges in the area of sentiment analysis has been discussed. Sentiment review structure is compared with sentiment analysis challenges. The effect of this distinction shows that domain-dependence is an important part of sentiment challenges. [3] Machine learning algorithms like Naïve Bayes and Logistic regression is used for first classifying the reviews in proper aspects then performing sentiment analysis on them. It explains the importance of review analysis over rating system.

## 2.3 FEASIBILITY STUDY

A feasibility study tries to objectively and rationally demonstrate the strengths and weaknesses of a current business or a proposed initiative, as well as the opportunities and risks that exist in the environment, the resources required to stay the course, and the likelihood of success. To establish the project's chances of success, a feasibility study considers all of the project's relevant components, including economic, technical, legal, and scheduling difficulties.

Feasibility studies have several advantages, including assisting project managers in determining the pros and cons of pursuing a project before devoting a considerable amount of time and money to it. Feasibility studies can also provide crucial information to a company's management team, potentially preventing them from embarking on a risky business venture.

Accounting statements, operations and management information, research and policies, monetary data, legal needs, and tax duties should all be included in a well-designed feasibility study. It focuses on the following major issues:

1) What are the user's demonstrable requirements, and how does a potential system meet
2) What are the resources available for a certain candidate system?
3) What are the most likely consequences of the candidate system for the company?
4) Is it worthwhile to address the problem?

### 2.3.1 TECHNICAL FEASIBILITY

A study of resource availability that may affect the capacity to construct an acceptable system is known as technical feasibility. This research assesses whether the technology necessary for the proposed system is currently available on the market. To create the project, both hardware and software, as well as essential technologies, are analyzed/assessed in these current resources. This technical feasibility analysis determines whether the necessary resources and technology are available for the project's development. The formal process of determining whether it is technically possible to manufacture a product or service is known as technical feasibility. It is critical to plan and prepare for each step of the operation before introducing a new offering or taking on a client project.

In this project, the programming languages we have used is python and using python we have implemented the Machine Learning algorithms. The frontend created is user friendly as it is developed using the Python’s Streamlit Library

### 2.3.2 ECONOMICAL FEASIBILITY

Economic feasibility is the cost and logistical outlook for a business project or endeavor. Before embarking on a brand-new venture, most businesses conduct an economic study, which could be a study that analyzes data to determine whether the cost of the prospective new venture will ultimately be profitable to the company. Companies will occasionally hire an outside firm to do economic feasibility studies for them. This feasibility study contains a detailed analysis of the project's development costs, which includes all required costs for final development such as hardware and software resources, design and development costs, and operations costs, among other things. After that, it is determined whether the initiative will be financially beneficial to the organization.

Economic feasibility checks conducted on our system conclude that our project is economically feasible since we have used all the modules as open source and free of cost.

### 2.3.3 OPERATIONAL FEASIBILITY

Operational feasibility is a metric for how well a proposed system solves problems and exploits possibilities discovered during scope definition, as well as how well it meets the criteria identified during the requirements analysis phase of system development. The operational feasibility assessment examines how well the planned development project fits into the existing company environment and objectives in terms of development timetable, delivery date, corporate culture, and existing business practices.

The project which we created is operationally feasible as it is developed in python language. The frontend is user friendly and is easily accessible.

# CHAPTER 3

# PROPOSED SYSTEM

# 3. PROPOSED SYSTEM

Our proposed system is to apply natural language processing techniques to classify a set of restaurant reviews as positive or negative. We implement a set of features that we believe to be relevant to the sentiment expressed in reviews and analyze their effect on performance, providing insights into what works and why sentiment categorization can be so difficult. We analyze how a review's conformance to a particular language model can be affected by the sentiment of the review.

We experiment with different linguistically motivated models of sentiment expression, again using the results to improve the performance of our classifier We examine the effects of part-of-speech tagging on our ability to predict sentiment. We experimented with different methods of preprocessing the data. Because the reviews are unstructured in terms of user input, reviews ca nlook like anything from a paragraph of well-formatted text to a jumble of seemingly unrelated words to a run-on sentence with no apparent regard for grammar or Punctuation. Our initial pass over the data simply tokenized the reviews based on whitespace and treated each token as a unigram, but we were able to improve performance by removing punctuation in addition to the whitespace and converting all letters to lowercase.

In this way, we treat the occurrences of "good", "Good", and "good." all as the same, which gives better predictive power to any test set review containing any of these three forms. Before converting into the unigram stemming was also done which means the various forms (tenses, verbs) of the words were removed and treated as a single word. After the matrix is build the non-frequent words are removed by setting a threshold in order to improve the accuracy. So our matrix includes relevant unigrams as well as bigrams which are occurring more than the threshold times.

## Algorithm used in proposed system (Naive Bayes)

Proposed system uses this Naive Bayes It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as 'Naive'.

Naive Bayes model is easy to build and particularly useful for very large data sets. Along withsimplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

**Advantages of proposed system**

- Good at pattern recognition problems
- Data-driven, and performance is high in many problems
- End-to-End training: little or no domain knowledge is needed in system construction
- Learn of representations: cross-modal processing is possible
- Gradient-based learning: learning algorithm is simple
- Mainly supervised learning methods

## 3.1 DRAWBACKS OF CURRENT SYSTEM AND NEED FOR PROPOSED SYSTEM

The classification of a review is predicted by the average *semantic orientation* of the phrases in there view that contain adjectives or adverbs. A phrase has a positive semantic orientation when it has good associations (e.g., "subtle nuances") and a negative semantic orientation when it has bad associations (e.g., "very cavalier").

The semantic orientation of a phrase is calculated as the mutual information between the given phrase and and the word "poor". A review is classified as recommended if theaverage semantic orientation of its phrases is positive.

### 3.1.1 Maximum Entropy

The Max Entropy classifier is a probabilistic classifier which belongs to the class of exponential models. Unlike the Naive Bayes classifier that we discussed in the previous article, the Max Entropy does not assume that the features are conditionally independent of each other. The Max Ent is based on the Principle of Maximum Entropy and from all the models that fit our training data, selects the one which has the largest entropy. The Max Entropy classifier can be used to solve a large variety of text classification problems such as language detection, topic classification, sentiment analysis and more. “Support Vector Machine” (SVM) is a super vised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n isnumber of features you have) with the value of each feature being the value of a particular coordinate.

- This type of classification is only done when the classifier has to work on the binary data which is not the case with Restaurant Reviews.

- However, from a practical point of view perhaps the most serious problem with SVMs is the high algorithmic complexity and extensive memory requirements of the required quadratic programming in large-scale tasks.

- If categorical variable has a category (in test data set), which was not observed in training dataset, then model will assign a 0(zero) probability and will be unable to make prediction. This is often known as “Zero Frequency”.

## 3.2 SYSTEM DESCRIPTION AND SRS

A software requirements specification (SRS) is a detailed description of software that has to be developed, including both functional and non-functional requirements. The SRS is mostly dependent on the agreement reached between the client and the contractors. It will include all employment scenarios in which the user interacts with the software. The software requirement specification document contains all of the requirements that are required for project development. We must constantly have a comprehensive grasp of the software system when developing it. To do this, we prefer to maintain constant contact with consumers to gather all requirements. This section describes the software module that will be developed. It briefly explains the project's purpose, scope, and work breakdown structure.

In our project we have developed a Software Specification Document (SRS), which will show the types of softwares used and their implementation in the proposed system. The detailed information about Jupyter Notebook, VS Code, MySQL Workbench have been mentioned there. It also includes the intended audience, terminology.

### 3.2.1 PURPOSE

A review/feedback is very significant for the restaurant owner to enhance its services, thus we created a web application to classify the review as positive or negative. The purpose is to provide the feedback to the restaurant from the customer about the ambience, the service they provided, the food they cook etc. These reviews helps the restaurant owner to improve in every way.

1. User Registration and Authentication: Allow users to register and create an account using their email or social media accounts.

2. Review Segmentation : Helps to analyze if the entered customer review is positive or negative.

### 3.2.2 PROJECT SCOPE

The project scope for a web application for Restaurant Review Analysis includes the following:

1. Customer Feedback and Improvement:

- Identifying Pain Points: Analyzing reviews helps restaurants identify specific areas where they excel and areas that may need improvement, such as service, food quality, or ambiance.
- Addressing Issues: Restaurants can use feedback to address specific issues raised by customers, enhancing the overall dining experience.

2. Competitor Analysis:

- Benchmarking: Understanding how a restaurant compares to its competitors in terms of customer satisfaction and overall ratings.
- Market Positioning: Identifying unique selling points and areas where a restaurant can differentiate itself from competitors.

3. Marketing and Reputation Management:

- Highlighting Strengths: Leveraging positive reviews and sentiments in marketing materials to attract new customers.
- Reputation Management: Monitoring and managing online reputation by addressing negative reviews and showcasing positive aspects.

The project scope can vary based on specific requirements, budget, and time constraints. It's important to gather detailed requirements from stakeholders and prioritize features accordingly. Additionally, Agile methodologies can be useful for iterative development and adapting to changing needs during the project.

### 3.2.3 PROJECT PLANNING

Project planning is part of project management that relates to the employment of schedules such as process flow and user flow diagrams. Initially, the project scope is outlined and the acceptable strategies for finishing the project are determined. Following this step, the periods for the various tasks needed to complete the work are listed and sorted into a work breakdown structure. The project schedule has been optimized to achieve the appropriate balance between resources and project duration to comply with the project objectives.

A Gantt chart is a project management tool assisting in the planning and scheduling of projects of all sizes; they are particularly useful for visualizing projects. A Gantt chart is defined as a graphical representation of activity against time; it helps project professionals monitor progress.

In the month of August we started our project, firstly the planning and project flow of the project were discussed. In the month of September we started with our Designing and documentation process. Now in the month of October we are completing the front-end and back-end development of the project.

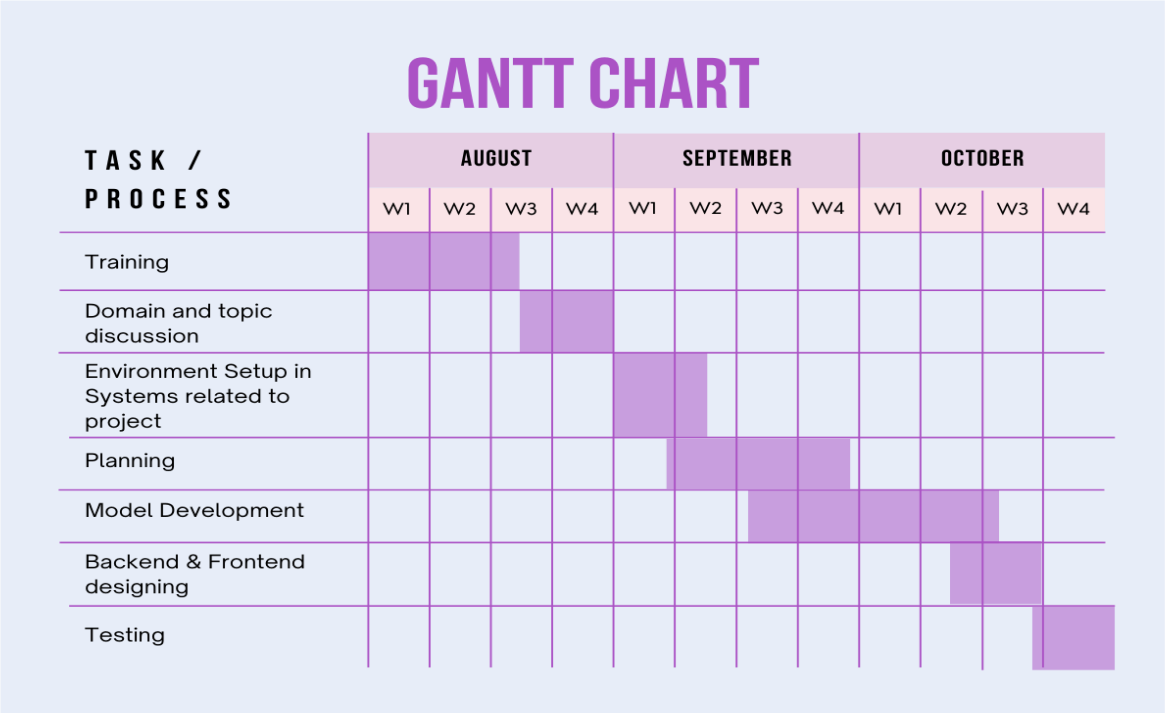


**Table No.3.2.1: Gantt Chart**

### 3.2.4 WORK BREAKDOWN STRUCTURE

A Work Breakdown Structure (WBS) is a hierarchical decomposition of a project into smaller, more manageable work packages or tasks. In the context of a project report, a WBS is a valuable tool for organizing and presenting the project's scope and the tasks involved. A Work Breakdown Structure (WBS) is a hierarchical decomposition of the total scope of work to be carried out by the project team. It is often used in project management to organize and structure the work into manageable sections. In a project report, a WBS provides a visual representation of the project's scope, outlining the tasks, sub-tasks, and deliverables.

| TASK NO. | TASK | DURATION |
|---|---|---|
| 1 | DISCOVERY | 1 Week |
| 2 | PLANNING | 1 Week |
| 3 | LITERATURE SURVEY | 1 Week |
| 4 | SRS | 1 Week |
| 5 | SYSTEM ANALYSIS | 1 Week |
| 6 | SYSTEM DESIGN | 3 Weeks |
| 7 | SPRINT RELEASE | 1 Week |
| 8 | VERIFICATION TESTING | 1 Week |
| 9 | VALIDATION TESTING | 1 Week |
| 10 | DEPLOYMENT | 1 Week |

**Table No.3.2.2: Work Breakdown Table**

## 3.3 SYSTEM ANALYSIS

The major objectives of system analysis are to find answers for what is being done, how it is being done, who is doing it, and when it is doing it. Logical system design occurs as a result of this process, and system analysis is an iterative process that continues until a preferred and accepted solution emerges. This entails acquiring and interpreting data, diagnosing issues, and making recommendations for system improvements. Project goals will be further aided by analysis of end-user information needs and the removal of any inconsistencies and incompleteness in these requirements.

The process of acquiring and interpreting data, diagnosing problems, and recommending system changes is known as analysis. The study of end-user information demands, as well as the eradication of any inconsistencies or incompleteness in these requirements, will help the project achieve its objectives. The study of a problem domain to recommend improvements and identify requirements and priorities for achieving the solution is known as system analysis.

The requirements for the system are gathered during the analysis phase. Company demands are investigated at this stage to improve business operations. The system analysis phase focuses on what the system will do to consider all stakeholders as potential information sources. In the analysis phase, a significant amount of time is spent talking with stakeholders and reviewing the stakeholder's input. Within this analysis phase, the analyst is discovering and fact-finding. Along with meeting with stakeholders, the analyst must meet with end-users to understand what the user's needs are and to learn about problems that affect the current system to assist with designing a new and more efficient system

Requirement analysis also gives software designers a visual representation of data, performance, and behavior that can be converted into data, architectural, interface, and element-level designs. Finally, the requirements specification allows the customer and the developer to evaluate the quality of the product once it has been constructed. There are five areas of endeavor in software requirement analysis: (1) problem recognition, (2) evaluation and synthesis, (3) modeling, (4) specification, and (5) review. The analyst begins by reviewing the system definition (if one exists) as well as the software project plan.

### 3.3.1 SYSTEM ARCHITECTURE

The system architecture is shown in figure. It comprises two main modules, an offline processing module, where the user profiles are being generated and the feature extraction and rating happens, as well as an online module, that generates real time recommendations. The prototype user review data from restaurant. The dataset contains user information, business information and user reviews. These objects are stored on Sqlite3 database. A brief overview of the system is provided as follows:

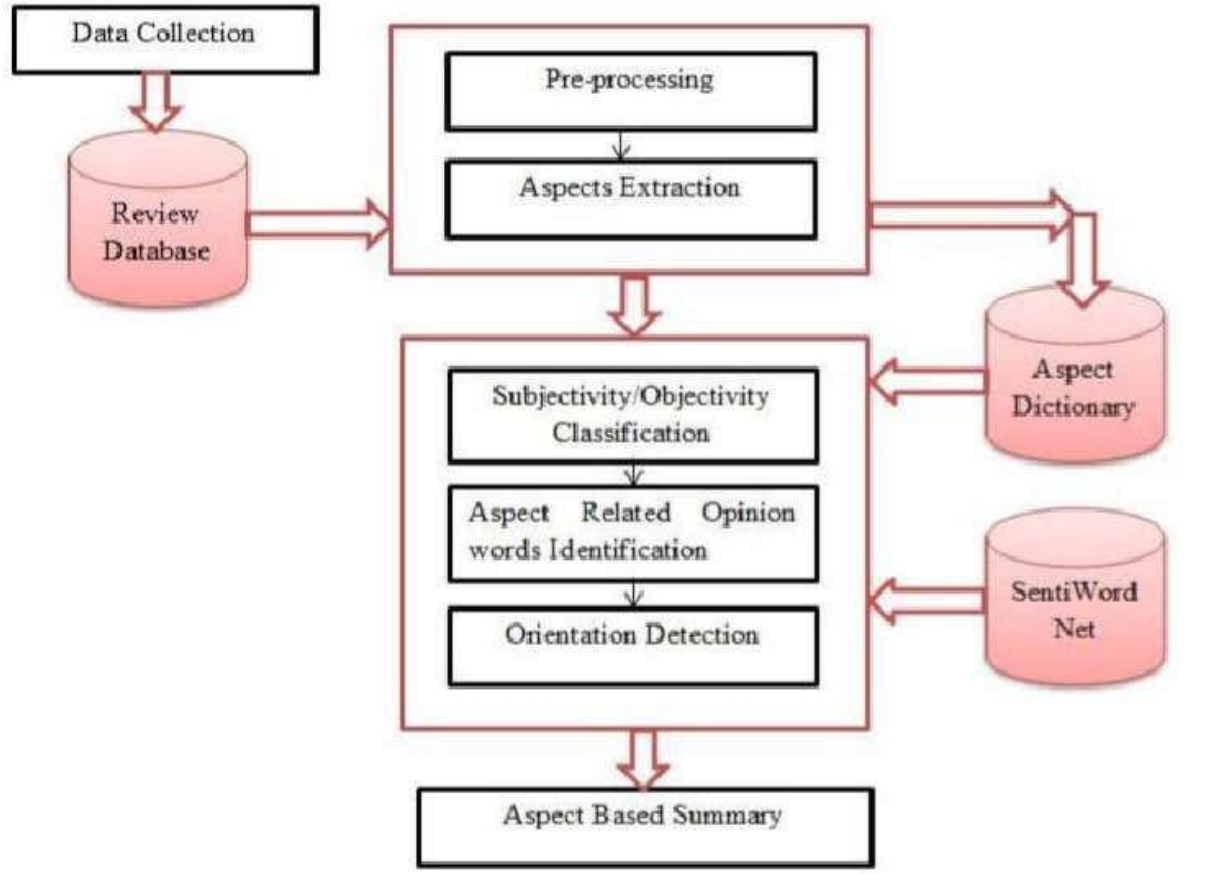


**Fig.No.3.3.1: System Architecture**

## 3.3.2 FLOWCHART OF SYSTEM

A flowchart is a form of a diagram that depicts an algorithm, workflow, or process by depicting the steps as various types of boxes and linking them with arrows. It's a diagram that shows the individual steps of a process in order. It's a general tool that may be used to define a variety of processes, such as manufacturing, administrative, and service processes, and project plans.

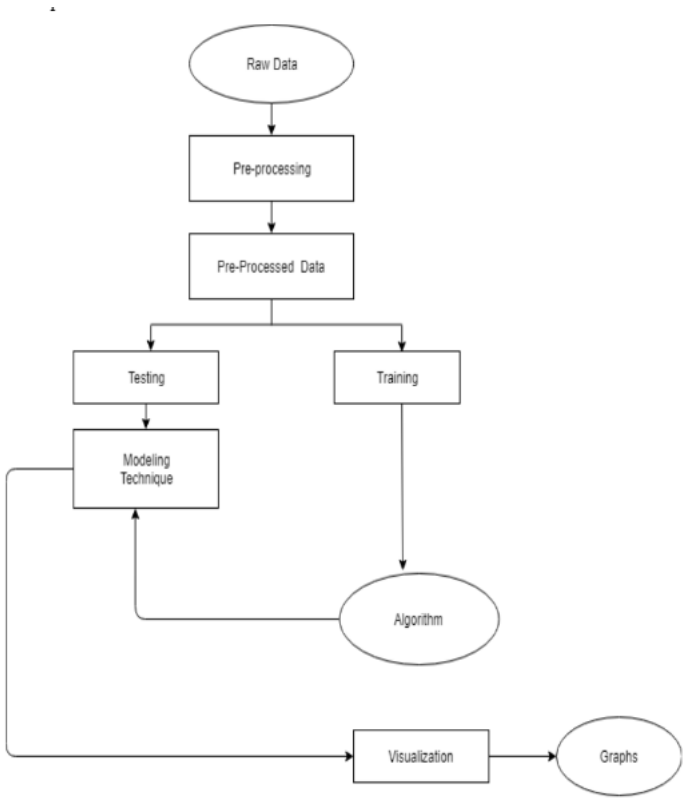


**Fig.No. 3.3.2: Flowchart of System**

## 3.3.3 DATA FLOW DIAGRAM

Data Flow Diagram is abbreviated as DFD. DFD is a diagram that depicts the data flow of a system or process. It also provides information about each entity's inputs and outputs, as well as the process itself. There is no control flow in DFD, and there are no loops or decision rules. A data flow diagram is a graphical representation of data 'flow' across an information system, which models its process features. A data flow diagram (DFD) depicts the types of data that will be input to and output from the system, as well as where the data will come from and go. A data flow diagram can be depicted in a variety of ways. The DFD is a modeling tool for structured analysis.

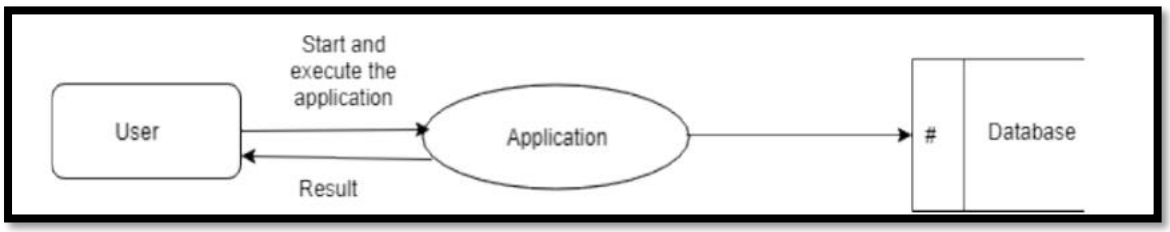


**Fig.No.3.3.3: Level 0 Data Flow Diagram**

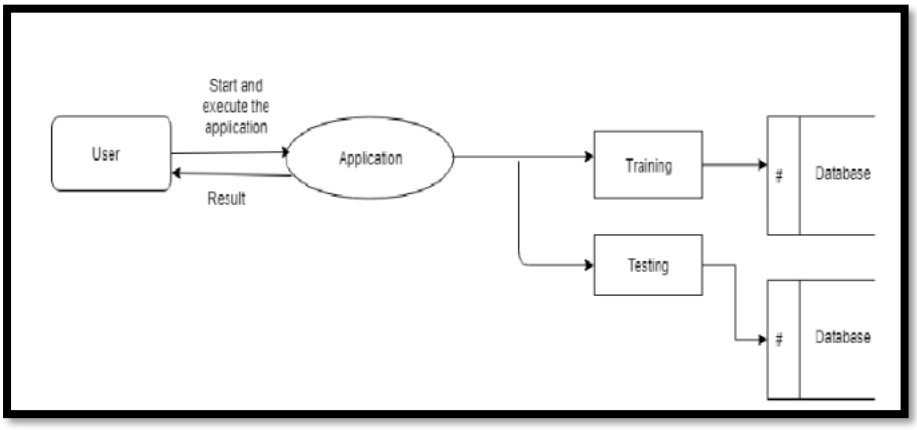


**Fig.No.3.3.3: Top Level Data Flow Diagram**

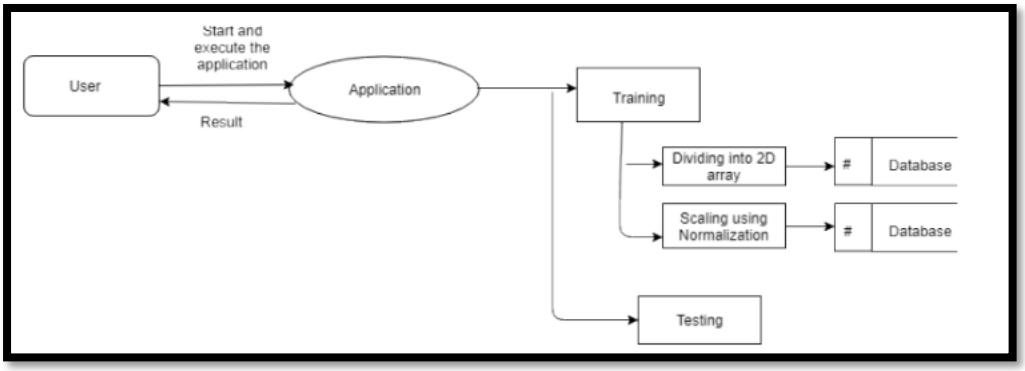


**Fig.No.3.3.3: Detailed Level Data Flow Diagram**

## 3.3.4 USE CASE DIAGRAM

A Use case diagram at its simplest is a representation of a user's interaction with the system and depicts the specifications of a use case. A use case diagram will portray numerous styles of users of a system and the various ways in which they move with the system. This sort of diagram is usually utilized in conjunction with the matter use case and can often be among other types of diagrams as well.

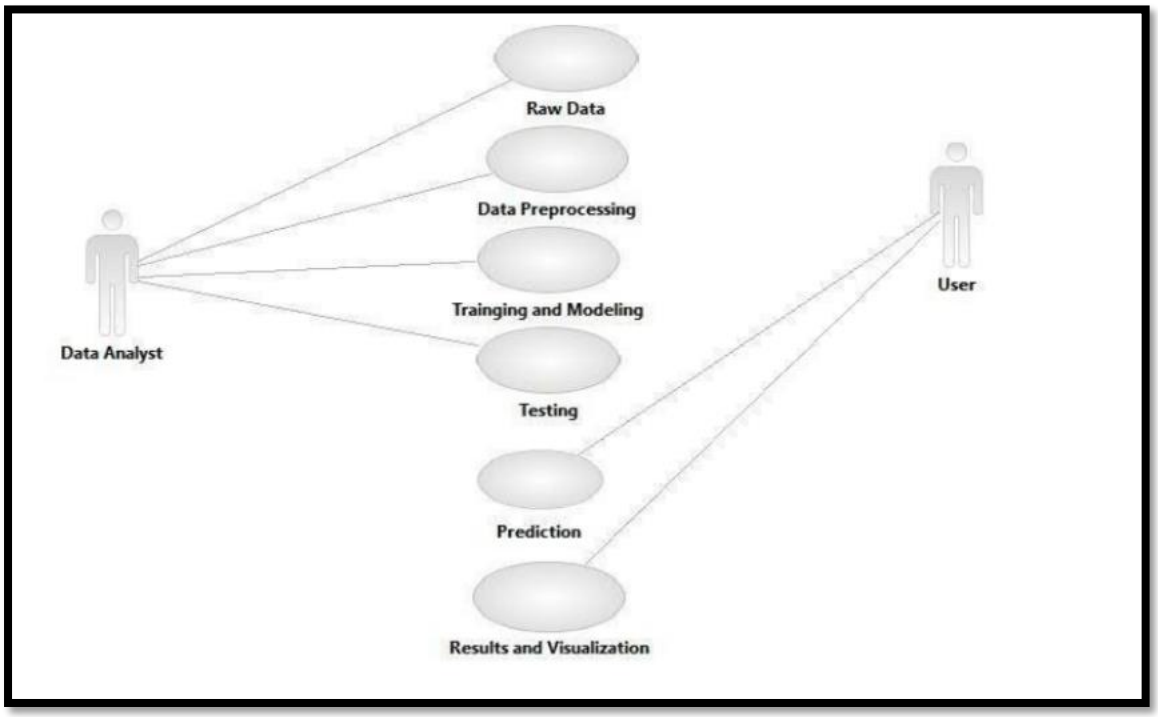


**Fig.No.3.3.4:Use Case Diagram**

## 3.3.5 HARDWARE REQUIREMENT

The foremost common set of necessities outlined by any operating system or software application is the physical computer resources, additionally proverbial as hardware, A hardware requirements list is common among a hardware compatibility list, particularly just in the case of operating systems.

- Minimum Hardware Requirements are as follows:
- Windows XP, Windows 7 (32/64 bit) or higher
- Minimum 8 GB RAM or higher
- Active internet connection minimum speed 512 kbps and above.

### 3.3.6 SOFTWARE REQUIREMENT

The software requirements are a list of the target system's features and functions. Users' expectations from the software product are communicated through necessities. From the client's perspective, the requirements should be evident or concealed, expected or surprising. The following are the software requirements:

- Visual Studio Code (At least Version 1.38 or higher)
- At least one Internet Browser e.g. Chrome, Firefox, Microsoft Edge etc, ( Chrome on windows 112.0.5615.49)
- Jupyter Notebook(v7.0.5 version or higher)
- MYSQL(8.0.19 version or higher)

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft with the Electron Framework, for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add functionality.

Google Chrome browser is a free web browser used for accessing the internet and running web-based applications. The Google Chrome browser is based on the open source Chromium web browser project. Google released Chrome in 2008 and issues several updates a year.

Jupyter Lab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows user to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.
MySQL Workbench is a unified visual tool for database architects, developers, and DBAs. MySQL Workbench provides data modeling, SQL development, and comprehensive administration tools for server configuration, user administration, backup, and much more. MySQL Workbench is available on Windows, Linux and Mac OS X.

## 3.4 SYSTEM DESIGN

System design is the process of defining elements of a system like modules, architecture, elements and their interfaces, and information for a system based on the supported necessities. It's the method of defining, developing, and planning systems that satisfy the particular wants and requirements of a business or organization.

The software modules should be organized in such a way that they are easy to build and alter. Structured design strategies aid programmers in dealing with program size and complexity. Analysts write instructions for developers on how to write code and how to put parts of code together to construct a program.

The Software Development Life Cycle's Design stage is critical. Design decisions are based on the requirements list you created during the definition phase. One or more designs are generated during the design process to accomplish the project's goal. Dioramas, flowcharts, sketches, site trees, HTML screen designs, picture impressions, prototypes, UML schemas, and other design phase products vary depending on the project subject.

Developers and technical architects begin the high-level design of the software and system during the design phase to meet each need. The technical aspects of the design are discussed with stakeholders, and numerous parameters such as risks, technologies to be employed, team capability, project restrictions, time, and budget are evaluated before the optimum design strategy for the product is chosen.

### 3.4.1 CLASS DIAGRAM

The class diagram is a fundamental component of object-oriented modeling. It's utilized for both general conceptual modeling and systematic application modeling. A static diagram is a class diagram. It represents an application's static view. A class diagram is used not only for visualizing, describing, and documenting many parts of a system but also for creating executable code for a software program. Data modeling can also be done with class diagrams.

A class diagram depicts a class's attributes and operations, as well as the system's limitations. Because class diagrams are the only UML diagrams that can be directly mapped with object-oriented languages, they are frequently utilized in the modeling of object-oriented systems. A collection of classes, interfaces, affiliations, collaborations, and constraints are shown in a class diagram. A structural diagram is another name for it. .

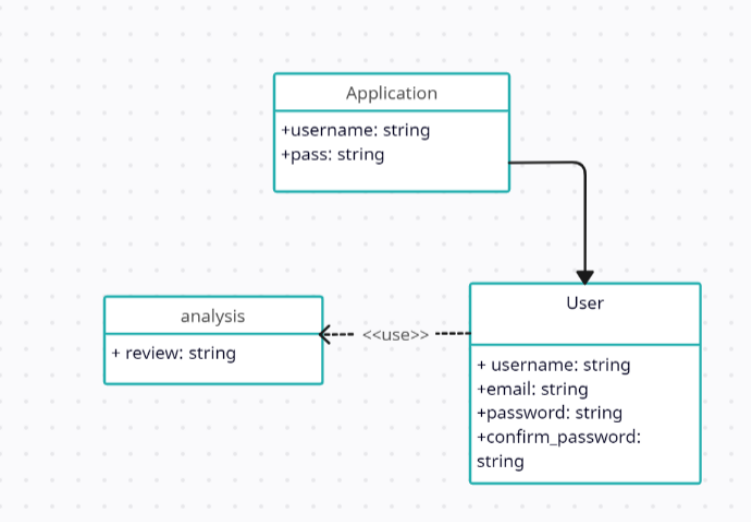


**Fig.No. 3.4.1: Class Diagram**

### 3.4.2 COMPONENT DIAGRAM

A component diagram depicts how software components in a system are organized and wired together. Component diagrams are commonly used to model implementation details and double-check that planned development covers all of the system's required functions. A component is analogous to a black box, with a provided interface and needed interfaces defining its exterior behavior. Component diagrams are used to break down a huge object oriented system into smaller components so that they may be managed more easily. It represents the physical view of a system that resides within the node, such as executables, files, libraries, and so on.

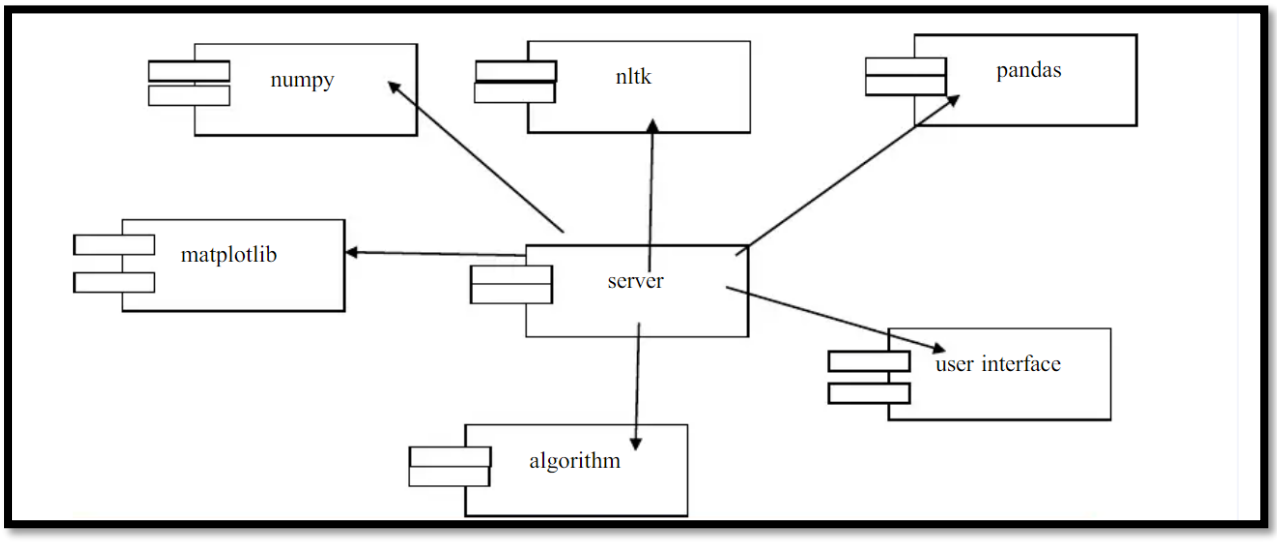


**Fig.No.3.4.2: Component Diagram**

## 3.4.3 DEPLOYMENT DIAGRAM

A deployment diagram depicts the configuration of run-time processing nodes as well as the components that reside on them. Deployment diagrams are a form of structure diagram that depicts the physical characteristics of an object-oriented system. They're typically used to represent the static deployment perspective of a system. The objective of the diagram is described by the phrase Deployment. Deployment diagrams are used to describe the physical components that are utilized to deliver software components. Component and deployment diagrams have a lot in common.

The actual hardware and software of a system are typically depicted using deployment diagrams. It enables you to visualize how the system will be installed on the hardware. Unlike other UML diagram types that primarily outline a system's logical components, deployment diagrams aid in representing a system's hardware topology.

Deployment Diagram represents the usage of applications across different devices and servers.

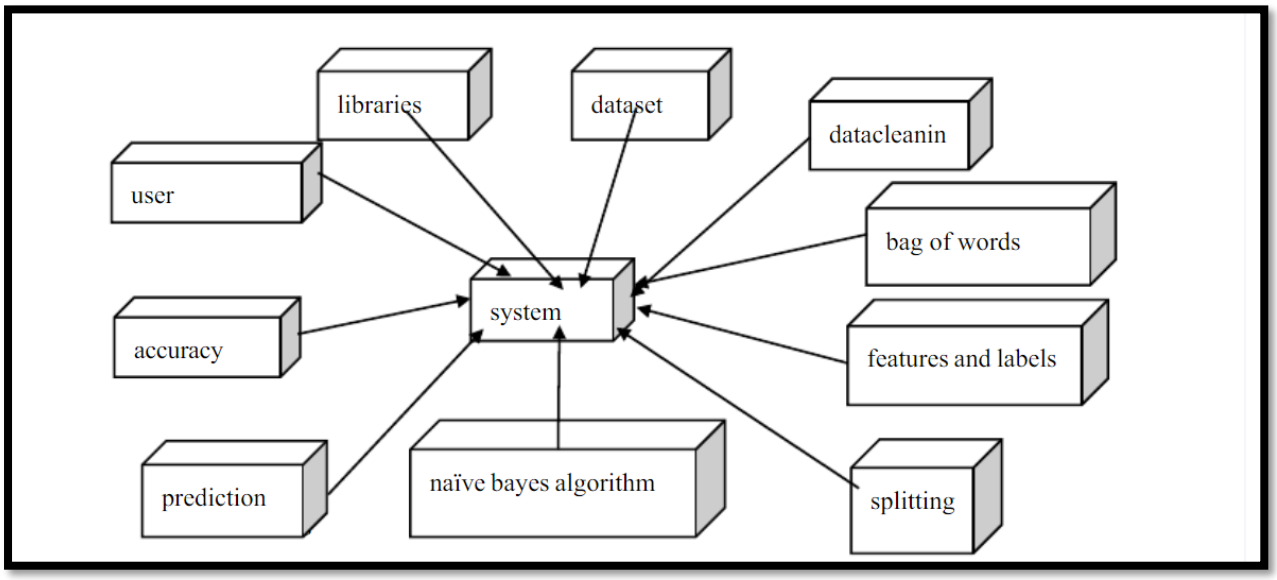


**Fig.No.3.4.3:Deployment Diagram**

## 3.4.4 USER STORIES

A user story is an informal, general explanation of a software feature written from the perspective of the end user. Its purpose is to articulate how a software feature will provide value to the customer. It's tempting to think that user stories are, simply put, software system requirements. But they're not. A key component of agile software development is putting people first, and a user story puts end users at the center of the conversation. These stories use non-technical language to

provide context for the development team and their efforts. After reading a user story, the team knows why they are building, what they're building, and what value it creates. User stories are one of the core components of an agile program. They help provide a user-focused framework for daily work — which drives collaboration, creativity, and a better product overall.

Here are some examples of user stories specific to Restaurant Review Analysis Websites:

- As a visitor, I want to see a list of top-rated restaurants, so I can quickly find popular dining options in my area.
- As a user, I want to be able to search for a specific restaurant and view detailed information about its reviews, ratings, and overall sentiment, so I can make an informed decision before visiting.
- As a user, I want to submit a review for a restaurant, including a text description and a rating, to share my experience with others.
- As a user, I want to see a sentiment analysis of reviews for a specific restaurant, so I can get a quick understanding of the overall feedback.
- As a user, I want to filter restaurant reviews based on criteria such as date, rating, and sentiment, to narrow down the information and find relevant reviews.
- As a user, I want to receive personalized recommendations for restaurants based on my previous reviews and preferences, to discover new dining options.
- As a restaurant owner, I want to be notified when a new review is submitted for my restaurant, so I can stay updated on customer feedback.

# CHAPTER 4

# IMPLEMENTATION AND CODING

# 4. IMPLEMENTATION AND CODING

## 4.1 IMPLEMENTATION METHODOLOGY

Implementation methodology in web-based application development refers to the process of planning, designing, developing, testing, and deploying a web application. There are various implementation methodologies used in web-based application development, each with its own set of advantages and disadvantages.

The implementation is divided into three parts

1. Database : The platform we are using for making databases is MYSQL Workbench.
2. Front-end : For the front-end we used Streamlit
3. Back-end : For the back-end and linking between database and front-end in Mysql.

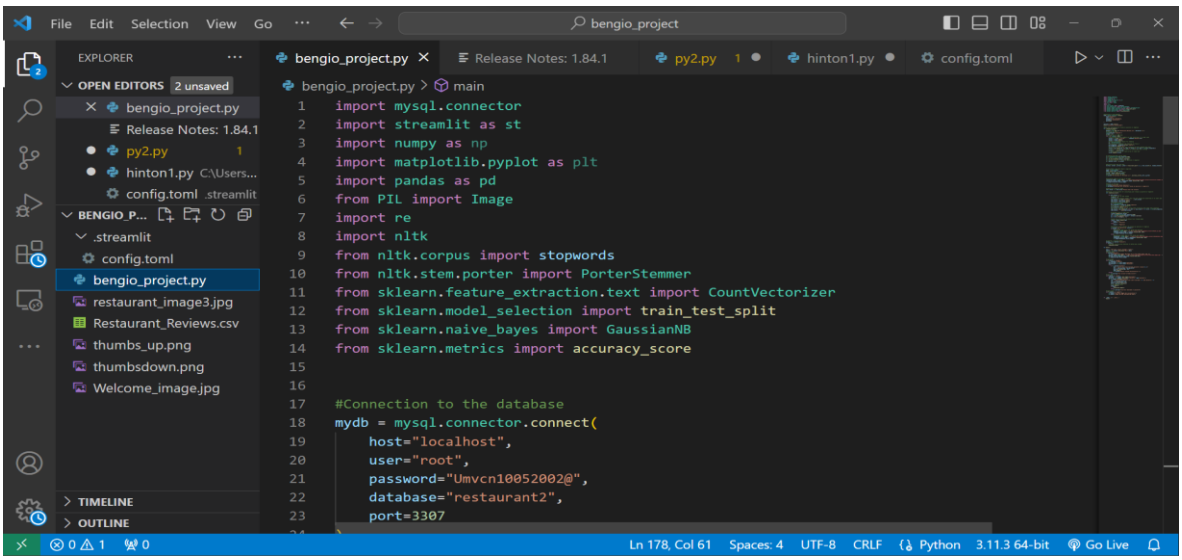


**Fig.No.4.1.1: Project Coding Environment**

## Database :

The tables created for the applications are:

User: created for storing login information of user and authenticating it

| id | username | password |
|---|---|---|
| 1 | chaitanya | 123 |
| 2 | radhika | 456 |
| 3 | bengio | 789 |
| 4 | hinton | 321 |
| 7 | xyz | 567 |
| 9 | ABC | 123 |

**Fig.No.4.1.2: Table & Attributes of the table In Database**

## Coding :

The figures  shows the code for the Welcome/Home Page, Signup, Login, and Logout page.


**Fig.No.4.1.3: Code for welcome and signup page**

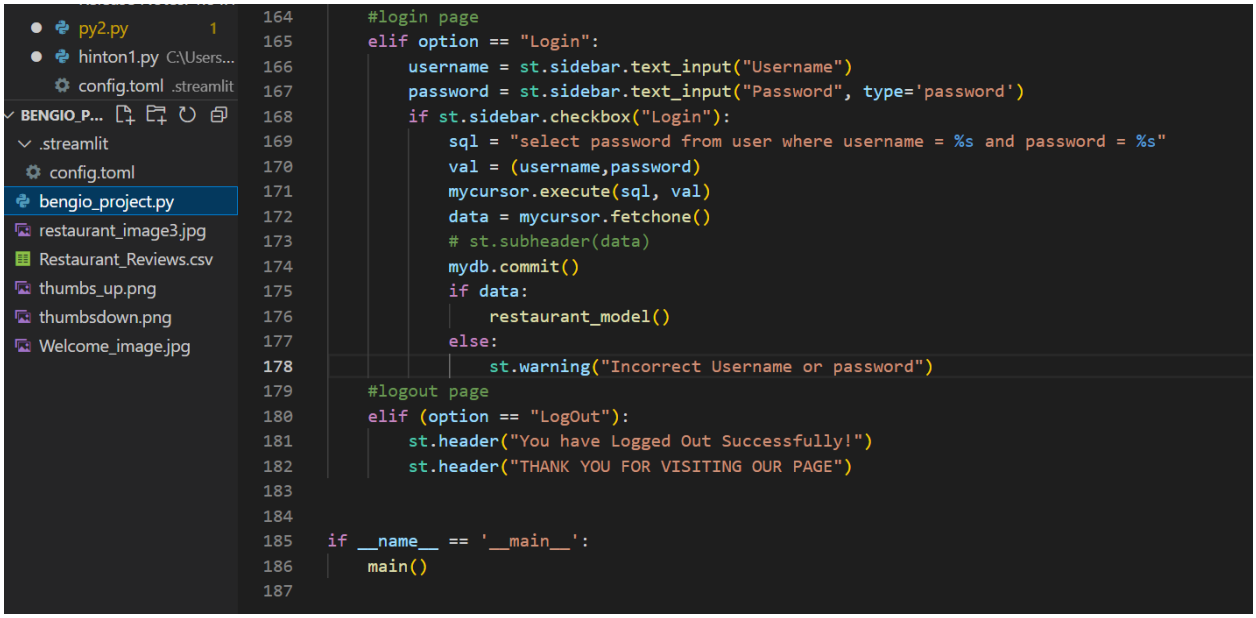


**Fig.No.4.1.4: Code for login and logout page**

The figure shows the code for creating the model using Nave Baye's Algorithm and also for creating a frontend using streamlit as well.

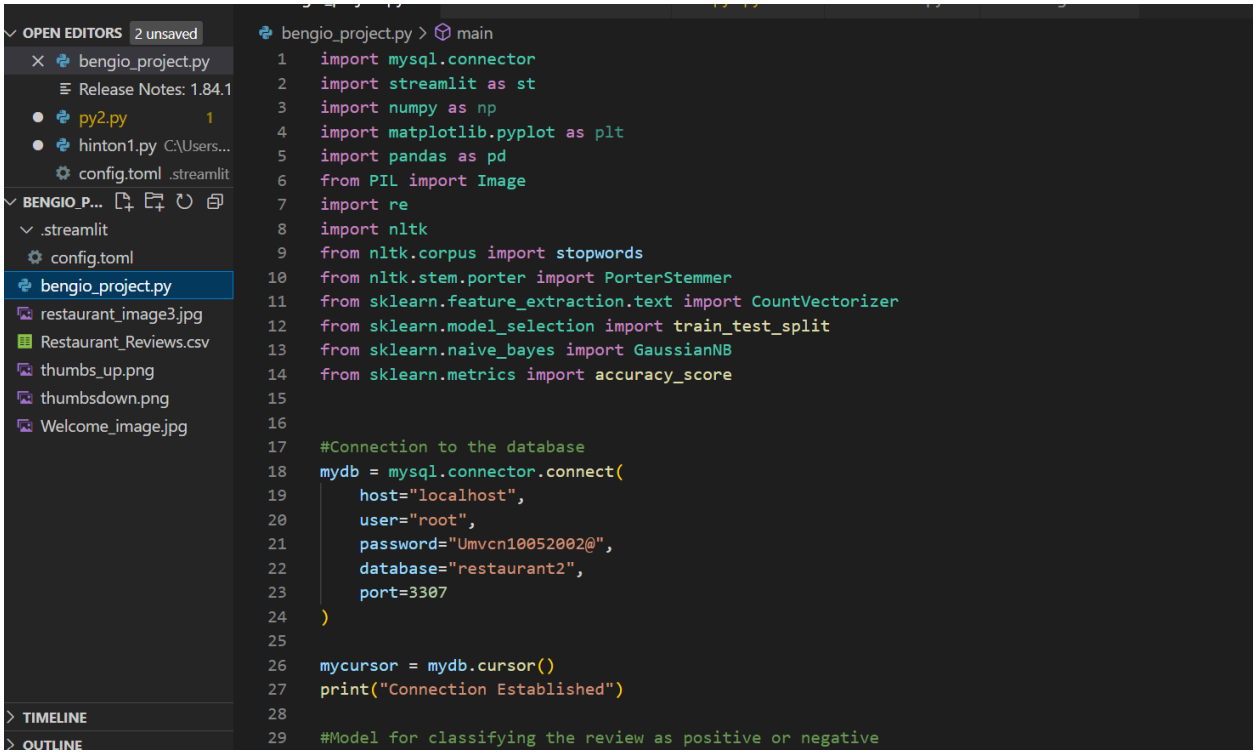


**Fig.No.4.1.5: Code for Machine Learning Model**

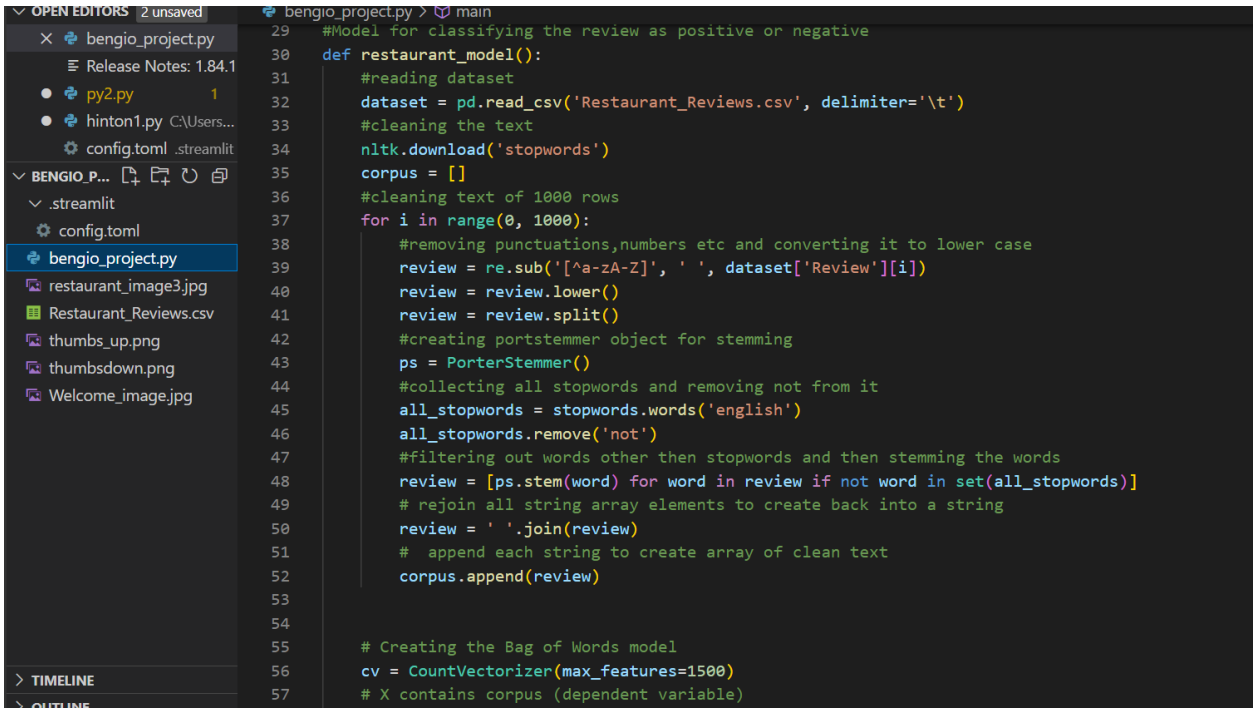


**Fig.No.4.1.6: Code for Machine Learning Model**

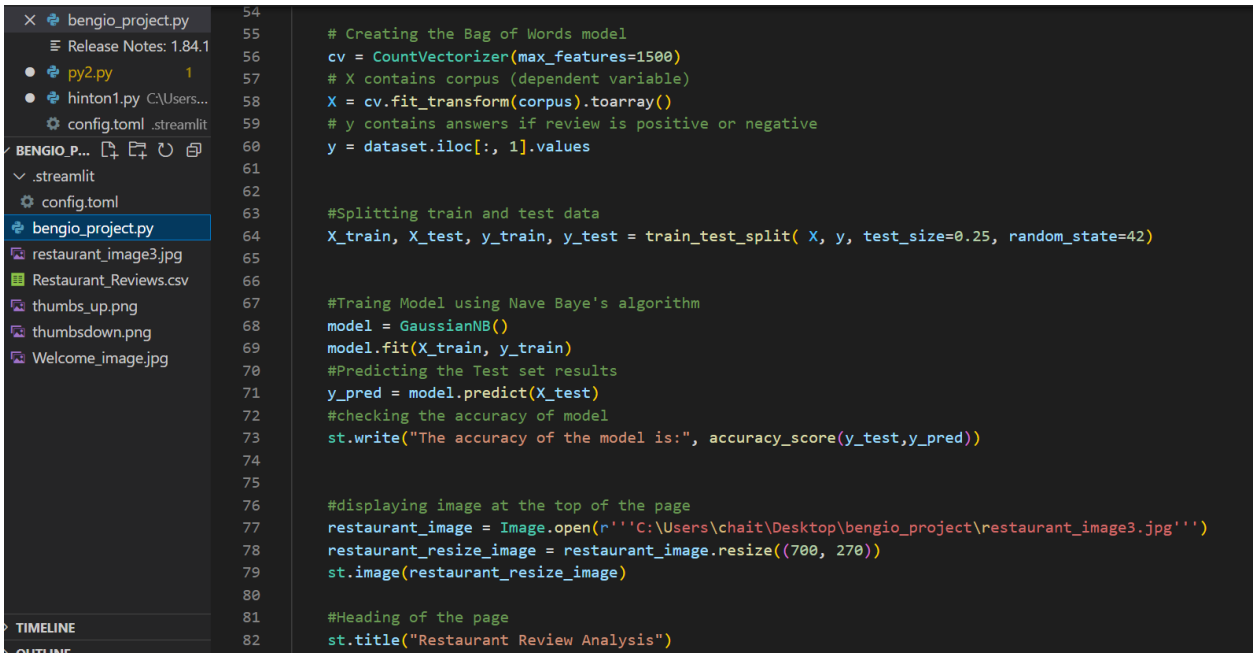


**Fig.No.4.1.7: Code for Machine Learning Model**

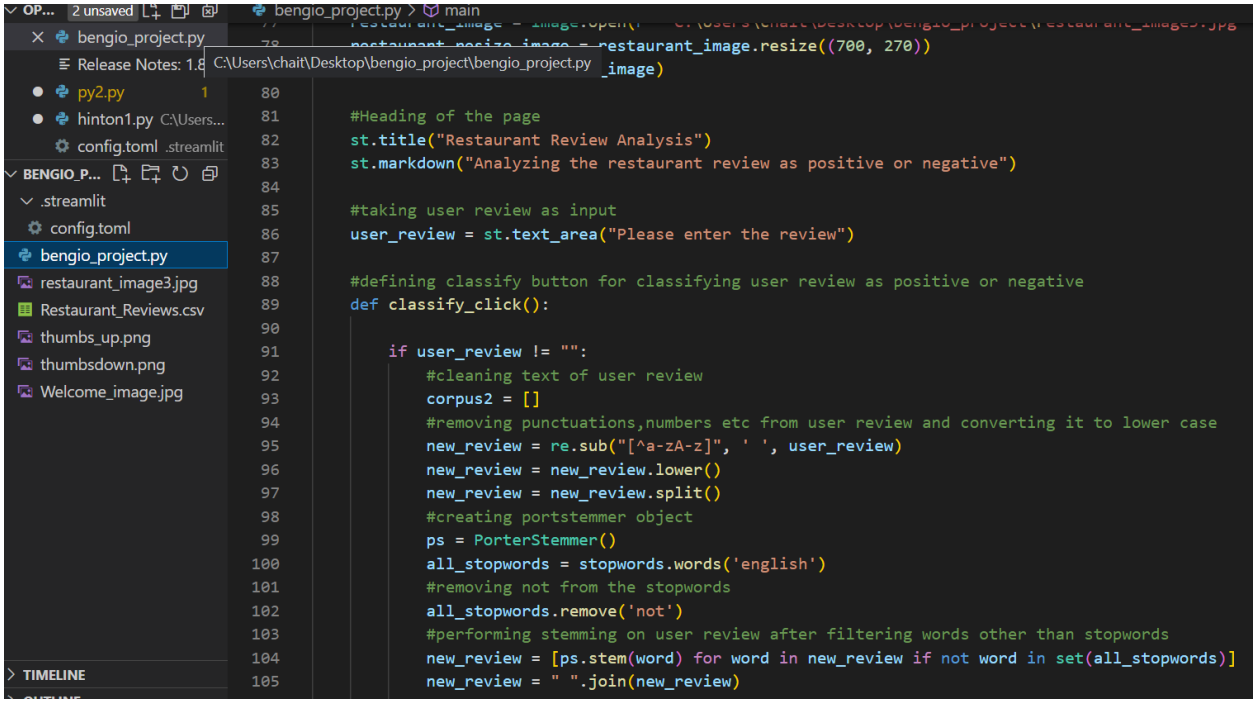


**Fig.No.4.1.8: Code for Machine Learning Model**

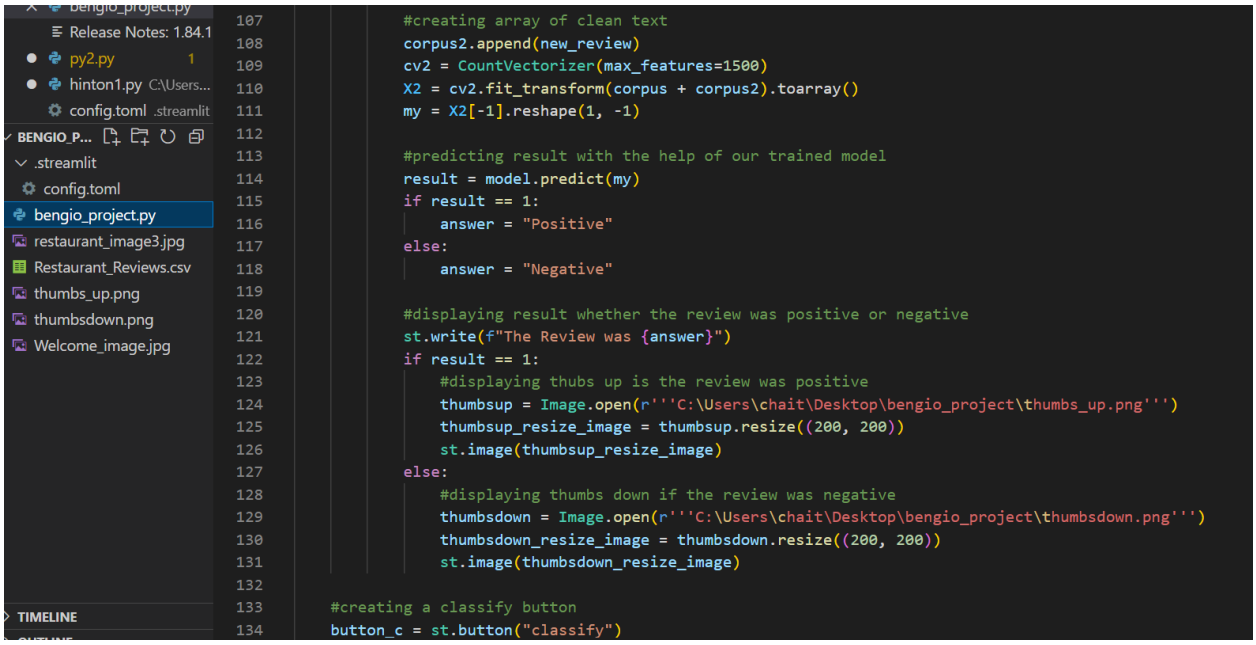


**Fig.No.4.1.9: Code for Machine Learning Model**

## Login Page Output :

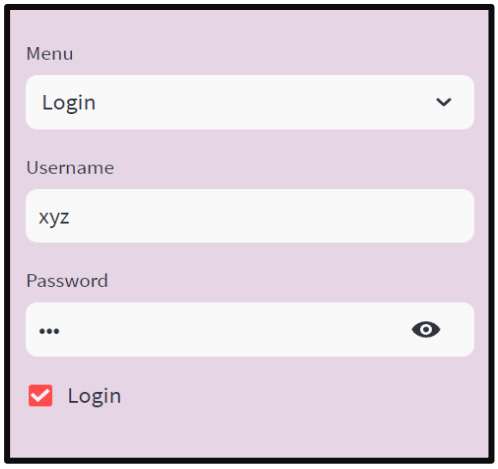


**Fig.No.4.1.10: Login Page**

A login screen is a common component in web applications that allows users to authenticate themselves and access secured areas or personalized features. In our web app user need to enter the username and password for accessing our page.

## SignUp Page Output :

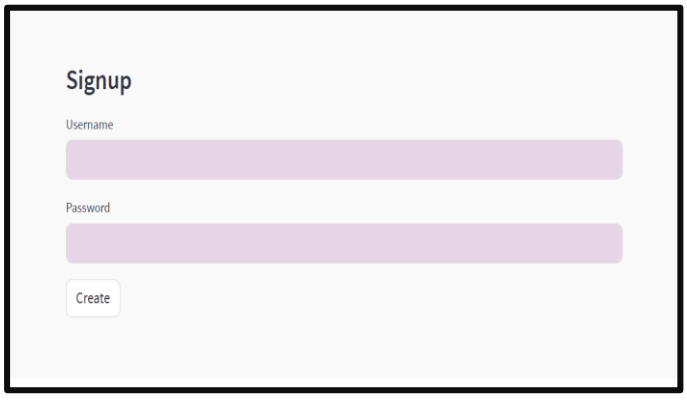


**Fig.No.4.1.11: Sign Up Page**

A sign-up screen is a common component in web applications that allows new users to create an account and access the application's features. For creating the account the user needs to enter the username and password.

## Prediction Page Output:

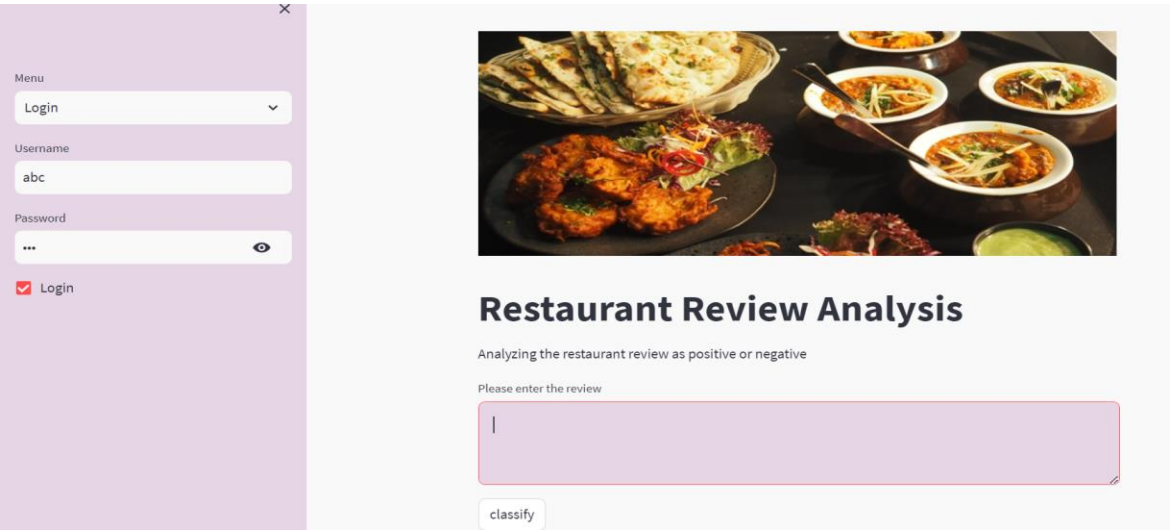


**Fig.No.4.1.12: Prediction page**


**Fig.No.4.1.13: Predicting the positive review**

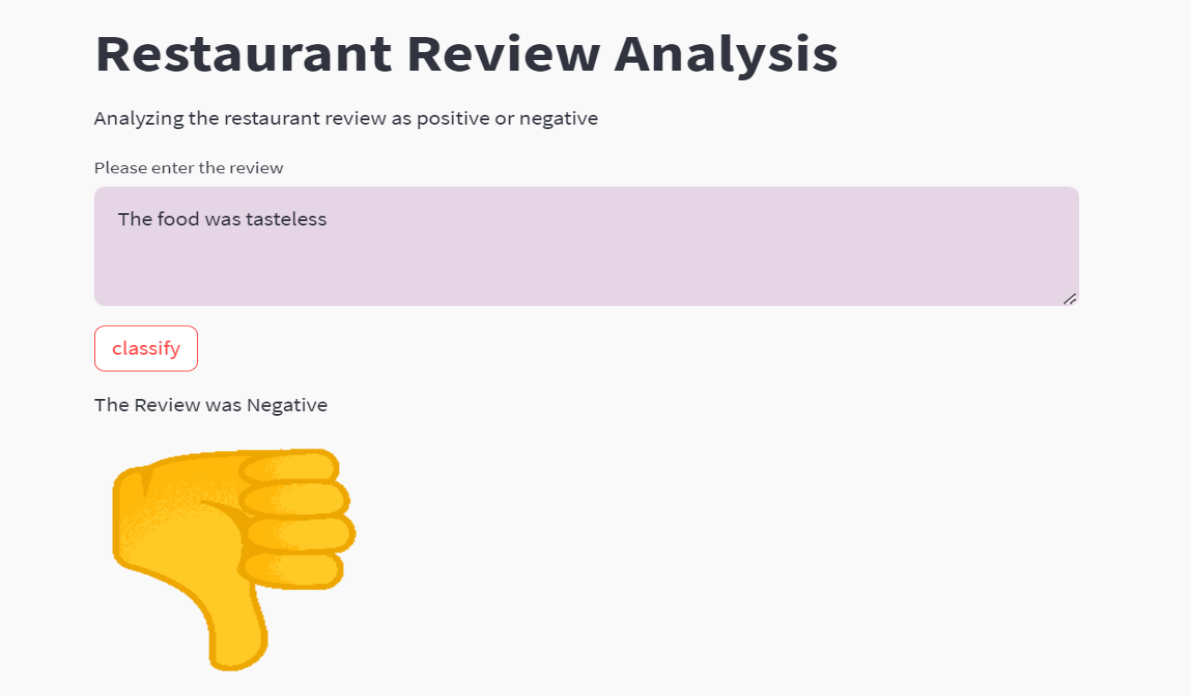


**Fig.No.4.1.14: Predicting the negative review**

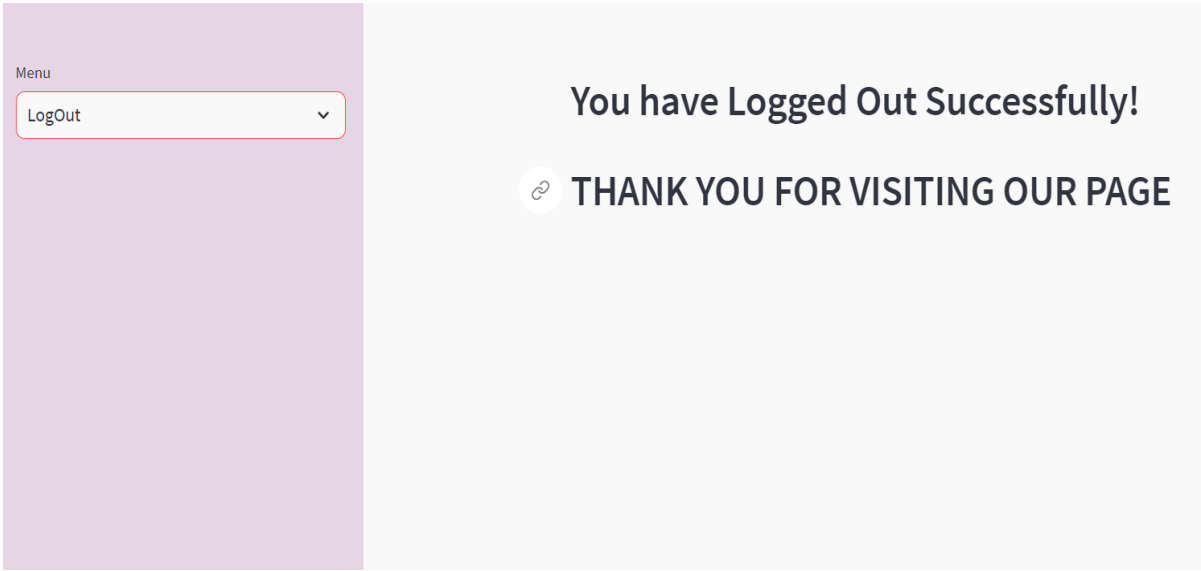


**Fig.No.4.1.15: Logout Page**