**Final Project Report**

**Sentiment Analysis to Rate a Product**



**Project Supervisor**

**Amna Bibi**

**Submitted By**

Project Group ID: **F220227F86**

Group Member Name: **FAHAD**

VU ID: **BC190202247**

**Software Projects & Research Section,**

**Department of Computer Sciences,**

**Virtual University of Pakistan**

|  |
| --- |
|  |



**CERTIFICATE**

This is to certify that FAHAD (BC190202247) have worked on and completed their Software Project at Software & Research Projects Section, Department of Computer Sciences, Virtual University of Pakistan in partial fulfillment of the requirement for the degree of BS in Computer Sciences under my guidance and supervision.

In our opinion, it is satisfactory and up to the mark and therefore fulfills the requirements of BS in Computer Sciences.

**Supervisor / Internal Examiner**

<<Project Supervisor Name>>

Supervisor,

Software Projects & Research Section,

Department of Computer Sciences

Virtual University of Pakistan

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

**External Examiner/Subject Specialist**

<<External Supervisor Name>>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

**Accepted By:**

**\_\_\_\_\_\_\_\_\_\_\_\_\_**

(For office use)

**EXORDIUM**

**In the name of Allah, the Compassionate, the Merciful.**

**Praise be to Allah, Lord of Creation,**

**The Compassionate, the Merciful,**

**King of Judgment-day!**

**You alone we worship, and to You alone we pray for help,**

**Guide us to the straight path**

**The path of those who You have favored,**

**Not of those who have incurred Your wrath,**

**Nor of those who have gone astray.**

**DEDICATION**

First, I would like to thanks to ALLAH and His Beloved Prophet Hazrat Muhammad (PBUH)

I want to dedicate this project to my respected and honorable parents as they guide and pray for me throughout the project.

After that to my project supervisor and faculty members, I am grateful for your guidance, expertise, and valuable insights. Your mentorship has been invaluable in shaping this project and enhancing my understanding of the subject matter.

Lastly, this project is also dedicated to all those who have played a part, big or small, in making it a reality. Your contributions have made a lasting impact, and I am truly grateful for your support.

**ACKNOWLEDGEMENT**

I would like to use the moment to convey my sincere appreciation and respect to my mentor **"Miss Amna Bibi"** and **"Virtual University"** for their support, supervision, and encouragement during this project. The blessings, assistance, and direction she has occasionally provided will help me get far. She demonstrates how to finish the job.

I owe her a huge debt of gratitude for sharing her knowledge and her skills with me. Her invaluable advice and support helped me with the project and made it simple to complete.

I am also grateful to the Virtual University for the helpful information they have offered. Throughout the course of my assignment, I appreciate their cooperation. Additionally, I want to express my gratitude to everyone who has prayed for me and been directly or indirectly engaged in this endeavor.

**ABSTRACT**

Sentiment analysis, commonly referred to as opinion mining or emotion AI, uses computational linguistics, biometrics, natural language processing, and text analysis to systematically detect, extract, quantify, and investigate emotional states and subjective data. To determine how customers feel about a product, businesses frequently utilize sentiment analysis.

When purchasing a product online, a customer has the option of leaving feedback. Comments can be used as types of feedback. A new customer cannot read every remark any longer in order to make a purchasing decision. You must create a program that can assess a product's quality using an analysis of the sentiment of user reviews.

It is an online store application where only customers who have registered can browse products (which can be of any type) and their features and leave comments about them. Other customers' remarks are also visible to the user. The program will evaluate and rate each comment made about a product.

**TABLE OF CONTENTS**

|  |
| --- |
|  |

CHAPTER NO. 1

introduction 10

1.1 Sentiment Analysis

1.2 customer sentiment analysis

1.3 Types of sentiment analysis

1.4 why sentiment analysis is important

1.5 Sentiment analysis of products

1.6 definitions, acronyms and abbreviations

CHAPTER NO. 2

methodology 11

2.1 Introduction

2.2 Data domain

* 1. preprocessing
  2. sentiment analysis using tool
     1. SENTIMENT SCORES
     2. KEYWORDS CLOUD
  3. manual sentiment analysis
     1. SENTIMENT SCORES
     2. KEYWORDS CLOUD
  4. Work Plan
  5. Project Structure
     1. Team Structure
     2. Project Schedule (Submission Calendar)

CHAPTER NO. 3

results and evaluation 12

3.1 Introduction

3.2 Visualization

3.3 Accuracy

* 1. Analysis

Chapter no.4

conclusion and Future work 13

**<<Dear Students, you have to take special care of page numbers, which are written against each Chapter and sub headings. You have to mention your own page numbers and correct one against each document heading. Make proper table of contents. Remember that if there are any new headings or you have included more headings in your documents then you also have to write these in table of contents and modify table of contents according to your documents material>>**

**<<Dear Students, before starting each chapter the following would be the title page for each chapter on a separate page>>**

**CHAPTER 1**

Introduction

* 1. **SENTIMENT ANALYSIS**:

Sentiment Analysis (also known as opinionmining or emotion AI) is the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. Sentiment analysis is widely used in business to detect sentiments of consumers about products.

* 1. **CUSTOMER SENTIMENT ANALYSIS:**

Customer sentiment analysis is the process of analyzing and understanding the emotions, opinions, and attitudes expressed by customers towards a product, brand, or service. In the context of your project, customer sentiment analysis refers to the evaluation of user comments and ratings to determine the sentiment associated with the purchased products.

* 1. **TYPES OF SENTIMENT ANALYSIS:**

In the context of customer sentiment analysis, there are several types of sentiment analysis techniques that can be utilized. Here are three common types:

**Lexicon-Based Sentiment Analysis:**

Lexicon-based sentiment analysis relies on predefined sentiment lexicons or dictionaries that contain words or phrases with associated sentiment scores. Each word or phrase is assigned a polarity score indicating its sentiment (positive, negative, or neutral).

In this approach, the sentiment of a given text is determined by calculating the overall sentiment score based on the sentiment scores of individual words present in the text.

**Rule-Based Sentiment Analysis:**

Rule-based sentiment analysis involves the creation of specific rules or patterns that define sentiment expressions. These rules are typically crafted by domain experts or linguists.

The text is analyzed based on the presence of predefined linguistic patterns or rules associated with positive or negative sentiment. The sentiment of the text is determined by evaluating the presence and strength of these patterns.

**Machine Learning-Based Sentiment Analysis:**

Machine learning-based sentiment analysis involves training a model on a labeled dataset to recognize and classify sentiment in text. The model learns patterns and relationships between textual features and corresponding sentiment labels during the training phase. Once trained, the model can predict the sentiment of new, unseen text by extracting relevant features and applying the learned patterns. Machine learning algorithms commonly used for sentiment analysis include Naive Bayes, Support Vector Machines (SVM), and Recurrent Neural Networks (RNN).

* 1. **WHY SENTIMENT ANALYSIS IS IMPORTANT:**

Sentiment analysis is important in the context of this project for several reasons:

**Customer Insights:** Sentiment analysis provides valuable insights into customer opinions, emotions, and experiences regarding the purchased products. It helps you understand how customers perceive the quality, features, and overall satisfaction with the products offered on your online shopping website.

**Decision Making:** Sentiment analysis results aid potential buyers in making informed decisions. By considering the sentiments expressed by previous purchasers, prospective customers can gain a deeper understanding of the product's strengths and weaknesses, allowing them to make more confident purchasing choices.

**Feedback and Improvement:** Analyzing customer sentiments helps gather feedback on specific products. Positive sentiments indicate areas where the product excels, while negative sentiments highlight aspects that need improvement. This feedback can guide you in making informed decisions about product enhancements, updates, or addressing any issues raised by customers.

**Customer Satisfaction and Retention:** By actively analyzing customer sentiments, you can identify satisfied customers and engage with them for potential testimonials or referrals. Additionally, addressing negative sentiments and resolving customer issues promptly can contribute to improved customer satisfaction and retention.

**Marketing Insights:** Sentiment analysis allows you to identify positive sentiments associated with specific product features, aspects, or overall customer experience. These insights can be leveraged for targeted marketing campaigns, highlighting the positive aspects of the products to potential customers and enhancing their perception of your brand.

**Competitive Advantage:** By understanding customer sentiments, you can gain a competitive advantage in the market. Monitoring and responding to customer feedback promptly can help differentiate your online shopping website from competitors and build a reputation for excellent customer satisfaction.

* 1. **SENTIMENT ANALYSIS OF PRODUCTS:**

Sentiment analysis of products refers to the process of analyzing and evaluating the sentiment expressed by customers towards specific products. It involves assessing customer opinions, emotions, and attitudes associated with the purchased products to determine the overall sentiment.

In the context of this project, sentiment analysis of products entails analyzing the comments and ratings provided by users after they have purchased and used a particular product from your electronic products online shopping website. The sentiment analysis tool PHP Sentiment Analyzer, is used to analyze these comments and assign sentiment scores and labels based on the sentiment expressed in the text.

By performing sentiment analysis of products, we can gain insights into how customers perceive and feel about specific items available in our website. The analysis allows you to understand the overall sentiment associated with each product, whether it is positive, negative, or neutral.

* 1. **DEFINITIONS, ACRONYMS AND ABBREVIATIONS:**

**Definitions:**

* **Sentiment Analysis:** The process of determining and evaluating the sentiment or emotional tone expressed in a piece of text, such as customer reviews or comments.
* **Electronic Products Online Shopping Website:** A website where users can browse and purchase electronic products online.
* **User Interface (UI):** The visual and interactive components of a website or application that users interact with.
* **PHP:** Hypertext Preprocessor, a server-side scripting language used for web development.
* **MySQL:** An open-source relational database management system (RDBMS) used for storing and managing the website's data.
* **Lexicon:** A collection of words or phrases with associated sentiment scores used in sentiment analysis.
* **Rule-Based Analysis:** An approach to sentiment analysis that relies on predefined rules or patterns to determine sentiment.

**Acronyms and Abbreviations:**

* **SRS:** Software Requirements Specification.
* **IDE:** Integrated Development Environment.
* **UI:** User Interface.
* **PHP:** Hypertext Preprocessor.
* **MySQL:** Structured Query Language.
* **RDBMS:** Relational Database Management System.
* **AI:** Artificial Intelligence.
* **API:** Application Programming Interface.
* **UX:** User Experience.

**CHAPTER 2**

Methodology

* 1. **INTRODUCTION:**

* A software development methodology, also known as a system development methodology, is a framework or approach used in software engineering to guide and manage the process of developing an information system or software application. It provides a structured and systematic way to plan, organize, and control the various activities involved in the development lifecycle.
* The purpose of a software development methodology is to establish a set of guidelines, processes, and practices that help ensure the successful completion of a software project. It provides a framework for teams to collaborate effectively, manage resources, and deliver high-quality software products.
  1. **DATA DOMAIN:**

The data domain in the context of this project would encompass the specific types and categories of data that are relevant to your electronic products online shopping website. It includes the scope of data that will be stored, processed, and managed within the system. Here are some examples of the data domain:

* **User Data:**
  + User registration information (e.g., name, email, password)
  + User profile details (e.g., contact information, preferences)
  + Order history and details
  + Cart and shopping preferences
* **Product Data:**
  + Product details (e.g., name, description, price, image)
  + Product reviews and ratings
* **Sentiment Analysis Data:**
  + User comments and reviews
  + Sentiment scores and labels associated with comments
  + Sentiment analysis results (positive, negative, neutral)
* **Administrative Data:**
  + Advertisements and product listings
  + System logs and activity records
* **System Configuration and Settings:**
  + Database configurations
  + Website settings (e.g., theme, language)
  + Security configuration
  1. **PREPROCCESSING:**

In the context of this project, preprocessing refers to the steps and techniques applied to raw data before it is used for sentiment analysis or other tasks. Preprocessing helps to clean, transform, and prepare the data for further analysis. Here are some common preprocessing techniques that can be applied to the data in your project:

**Data Cleaning:**

* This involves removing any irrelevant or redundant information from the data, such as special characters, symbols, or HTML tags.
* It may also involve handling missing data by imputing or removing incomplete or null values.

**Text Normalization:**

* Text normalization techniques standardize and transform text data to make it consistent and easier to analyze.
* This can include converting text to lowercase, removing punctuation marks, and handling abbreviations or contractions.

**Tokenization:**

* Tokenization involves breaking down text into individual words, phrases, or tokens. This step helps in preparing the text for analysis, as each token represents a separate unit of analysis.

**Stop Word Removal:**

* Stop words are common words that do not carry significant meaning or sentiment, such as "the," "is," or "and."
* Removing stop words can reduce noise and improve the efficiency and accuracy of sentiment analysis.

**Stemming or Lemmatization:**

* Stemming and lemmatization are techniques used to reduce words to their base or root form.
* This helps in consolidating variations of words and reducing redundancy. For example, converting "running," "runs," and "ran" to the base form "run."

**Handling Negations and Emphasis:**

* In sentiment analysis, handling negations is crucial to accurately interpret sentiment. For example, converting "not good" to "not\_good" or "not\_good" to "bad."
* Emphasis can also be handled by identifying and preserving intensified words or phrases that may affect the sentiment.
  1. **SENTIMENT ANALYSIS USING TOOL:**

Sentiment analysis using a tool involves utilizing a specific software or library that is designed to perform sentiment analysis tasks. In this project, I mentioned using the PHP Sentiment Analyzer tool for sentiment analysis. Here's an overview of how sentiment analysis can be conducted using this tool:

**Integration:**

Integrate the PHP Sentiment Analyzer tool into your project by incorporating the necessary code and dependencies. Ensure that the tool is properly installed and configured within your development environment.

**Text Input:**

Provide the tool with the text input that needs to be analyzed for sentiment. In your case, the input would typically be the user comments or reviews related to the purchased products.

**Preprocessing:**

Preprocess the text input as discussed earlier in order to clean and normalize the text data before passing it to the sentiment analysis tool. Apply techniques such as text normalization, tokenization, and stop word removal.

**Sentiment Analysis:**

Utilize the sentiment analysis functionality provided by the PHP Sentiment Analyzer tool to evaluate the sentiment expressed in the text input.

The tool employs the VADER (Valence Aware Dictionary and sentiment Reasoner) lexicon and rule-based approach to assign sentiment scores and labels to the text. It calculates an overall sentiment score based on the sentiment scores of individual words or phrases present in the text.

**Sentiment Classification:**

Based on the sentiment scores calculated by the tool, classify the sentiment as positive, negative, or neutral. Determine the sentiment label that best represents the sentiment expressed in the text.

**Output and Analysis:**

* Capture and process the sentiment analysis results provided by the tool. This may include storing the sentiment scores, sentiment labels, and other relevant information in your database or data structure.
* Analyze and interpret the sentiment analysis results to gain insights into customer sentiments, product evaluations, and overall user experiences.
* By utilizing the PHP Sentiment Analyzer tool, you can automate the sentiment analysis process and efficiently analyze customer comments and reviews for sentiment. The tool's lexicon and rule-based approach allow for sentiment evaluation based on predefined sentiment scores, providing a quick and effective way to assess the sentiment expressed in the text data.

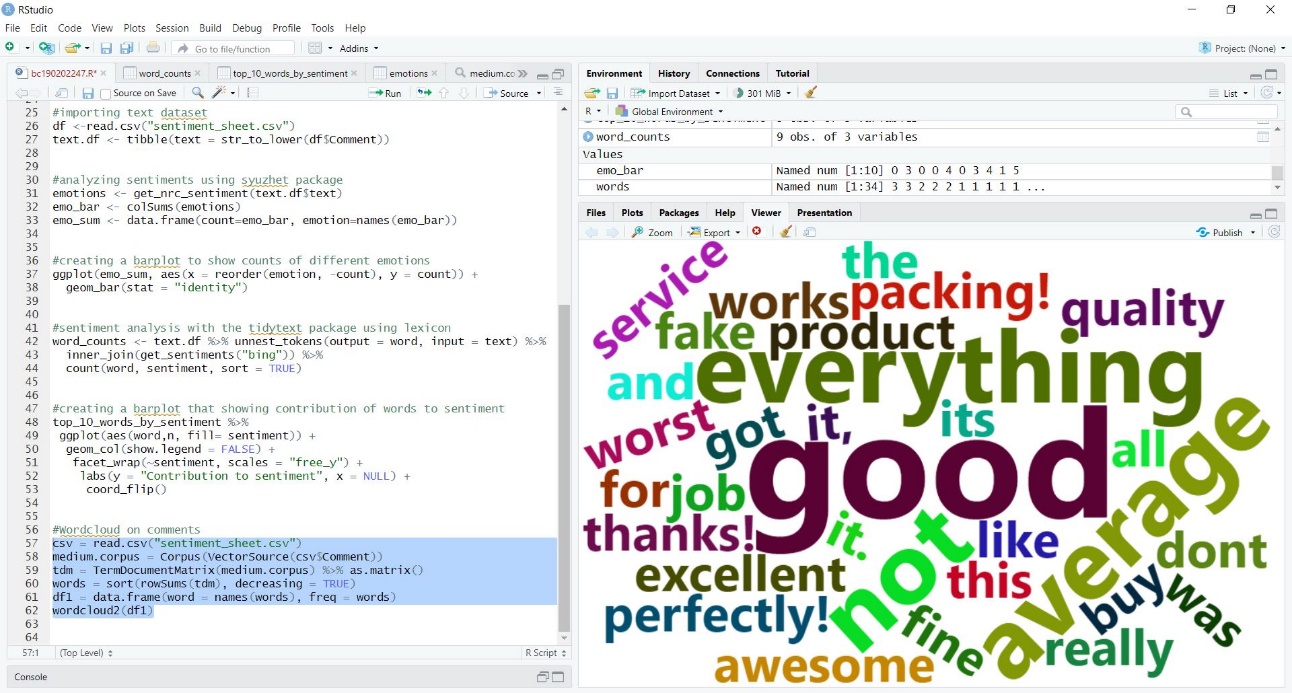
**2.4.1.)** **SENTIMENT SCORES:**

Sentiment scores in the context of sentiment analysis represent numerical values that quantify the sentiment expressed in a given text. These scores indicate the strength or intensity of the sentiment and help in assessing the overall sentiment polarity, whether positive, negative, or neutral. Here's some information about sentiment scores in this project:

* **Calculation:**
  + Sentiment scores are typically calculated using predefined lexicons or dictionaries that assign sentiment scores to words or phrases.
  + Each word or phrase in the text is assigned a sentiment score based on its association with positive or negative sentiment.
  + The sentiment scores of individual words or phrases are then aggregated or combined to calculate an overall sentiment score for the entire text.
* **Valence:**
  + Sentiment scores often have a valence or polarity associated with them, indicating whether the sentiment is positive or negative.
  + Positive sentiment scores indicate positive or favorable sentiment expressed in the text.
  + Negative sentiment scores indicate negative or unfavorable sentiment expressed in the text.
  + Neutral sentiment scores are typically close to zero or fall within a certain range and indicate a lack of strong positive or negative sentiment.
* **Range:**
  + The range of sentiment scores can vary depending on the sentiment analysis tool or library used.
  + In this project PHP sentiment analysis tool uses a range of -10 to +10, where negative scores indicate negative sentiment, positive scores indicate positive sentiment, and zero represents neutral sentiment.
  + Other tools may use a different scoring system or range based on their specific implementation or requirements.
* **Interpretation:**
  + Sentiment scores can be interpreted in various ways depending on the context and application.
  + Higher positive scores indicate stronger positive sentiment, while lower negative scores indicate stronger negative sentiment.
  + The magnitude of the sentiment score can provide insights into the intensity or strength of the sentiment expressed in the text.
* In this project, sentiment scores obtained from the PHP Sentiment Analyzer tool would help assess and quantify the sentiment expressed in user comments or reviews for the purchased products. By analyzing these sentiment scores, we can understand the overall sentiment polarity and use it to make informed decisions, evaluate customer satisfaction, and improve the user experience on electronic products online shopping website.

**2.4.2) KEYWORDS CLOUD:**

Keyword clouds, also known as word clouds or tag clouds, are visual representations of text data where the size of each word corresponds to its frequency or importance within the text. In the context of this project, keyword clouds have been generated to visually display the most commonly occurring words in different sentiment categories (positive, negative, and neutral) of the product comments.

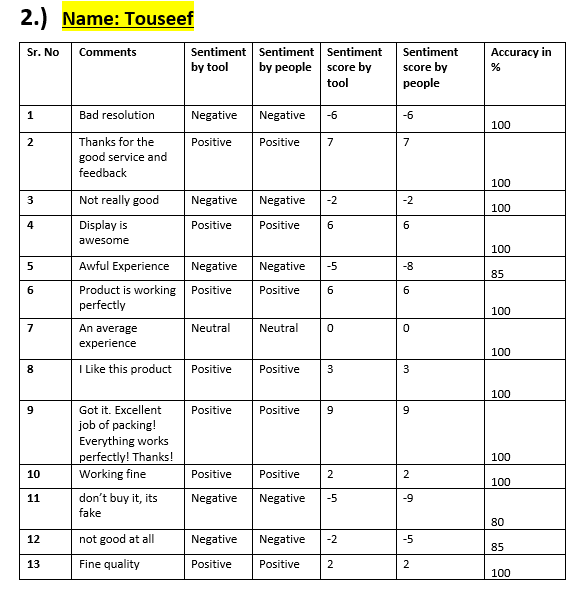


**Fig.01 Keyword Clouds**

The above figure would show the keyword cloud for a specific sentiment category (positive, negative, or neutral). The keyword cloud visually emphasizes words that appear more frequently or have greater importance within that sentiment category. The size of each word in the cloud indicates its frequency or significance relative to other words.

* 1. **MANUAL SENTIMENT ANALYSIS:**

Manual sentiment analysis, in the context of this project, refers to the process of gathering user feedback and opinions on the purchased products through manual ratings and reviews. It involves requesting users to provide their own subjective assessments of the products based on their experiences, and then analyzing and categorizing these ratings and reviews into sentiment categories (positive, negative, or neutral).



**Fig.02 Manual Sentiment Analysis**

In a above figure, I have collected manual sentiment analysis data in the form of a Word document file. This likely contains user ratings and comments, which we can manually analyze and assign sentiment labels to. It is how manual sentiment analysis is conducted.

**2.5.1.)** **SENTIMENT SCORES:**

In manual sentiment analysis, sentiment scores can be assigned to user ratings or reviews to quantify the sentiment expressed in a more numerical or quantitative form. These sentiment scores can provide a measurable representation of the sentiment associated with each user's feedback. Here's how sentiment scores is applied in manual sentiment analysis:

**Numerical Rating Scale:**

I use a numerical rating scale (e.g., on a scale of -10 to +10), you can directly consider the rating as the sentiment score. Visually you can see in the above figure02.

**2.5.2.)** **KEYWORD CLOUDS:**

In the context of manual sentiment analysis, keyword clouds can be generated to visually represent the most frequently occurring words or phrases within the user comments or reviews. These keyword clouds provide a quick and intuitive overview of the prominent keywords associated with specific sentiment categories (positive, negative, or neutral) in the manual sentiment analysis. **Here's how keyword clouds can be useful:**

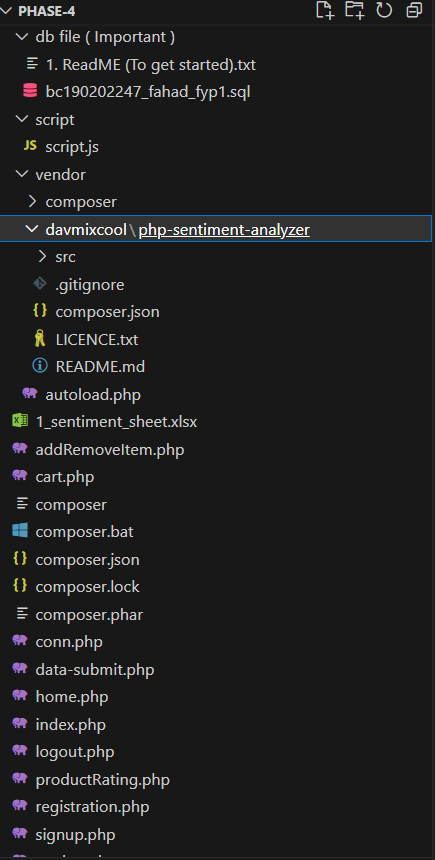
* **Visualizing Frequent Words:**
  + Keyword clouds visually emphasize words or phrases based on their frequency or importance within the user comments or reviews.
  + The size of each word in the cloud represents its occurrence frequency, with larger words indicating higher frequency.
* **Identifying Prominent Themes:**
  + By examining the keyword clouds for different sentiment categories, you can identify the most common themes or topics mentioned by users.
  + Prominent keywords in the clouds provide insights into the aspects, features, or experiences that have the most significant impact on users' sentiment.
* **Comparing Sentiment Categories:**
  + Keyword clouds allow for easy visual comparison between sentiment categories.
  + By comparing the keyword clouds for positive, negative, and neutral sentiment, you can quickly identify differences in the most frequent keywords used in each category.
* **Supplementary to Sentiment Scores:**
  + Keyword clouds can complement sentiment scores by providing a visual representation of the language and vocabulary used by users within different sentiment categories.
  + They help to capture the essence of sentiment beyond just scores, showcasing the specific words that contribute to positive, negative, or neutral sentiment.
  1. **WORK PLAN:**

Filled later ---------------------------------------------------

* 1. **PROJECT STRUCTURE:**

The project structure refers to the organization and arrangement of files, folders, and components within your project. It provides a systematic way to manage and organize the various elements of the project. While the specific project structure can vary based on personal preferences and project requirements, here is a suggested project structure for my sentiment analysis-based online shopping website:

* **Root Directory:**
  + This is the main directory of your project, containing all the project files and folders.
  + It serves as the starting point and typically holds configuration files and high-level project documentation.
* **Source Code:**
  + This directory contains all the source code files for my website's functionality.
  + It can be further organized into subdirectories based on the programming language (e.g., PHP) or the specific modules or components of your project.
* **Database**:
  + This directory includes files related to the database, such as database schema, table definitions, and data migration scripts.
  + It may also contain backup files or documentation related to the database design.
* **Resources:**
  + This directory holds various resources used in this project, such as images, CSS files, JavaScript libraries, and other static assets.
  + It can be further categorized into subdirectories based on the resource type or purpose (e.g., images, styles, scripts).



**Fig.03 Project Structure**

The above figure03 showing the project structure visually.

**2.7.1.)** **TEAM STRUCTURE:**

**2.7.2.)** **PROJECT SCHEDULE (SUBMISSION CALENDAR):**



**CHAPTER 3**

Results and Evaluation

**CHAPTER 4**

Conclusion and Future Work

**REFERENCES**

**APPENDIX**