**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 of all, let me express my gratitude to Allah and to His Beloved Prophet Hazrat Muhammad (PBUH).

I want to dedicate this endeavor to my exemplary parents who have supported and prayed for me during the entire process.

I also want to express my gratitude to my project manager and the teachers for their advice, knowledge, and insightful comments. Your guidance has been of immeasurable value in helping me develop this project and deepen my knowledge of the topic.

Last but not least, this project is dedicated to everyone who has contributed in any way to make it happen, no matter how tiny. Your efforts have had a long-lasting effect, and I sincerely appreciate your help.

**ACKNOWLEDGEMENT**

I'm going to want to take this opportunity to express my gratitude and respect to my mentor **"Miss Amna Bibi"** and **"Virtual University"** for their assistance, oversight, and encouragement during this endeavor. The blessings, kindness, and guidance she has sometimes offered will carry me a long way. She walks you through how to complete the task.

She taught me a lot, and I owe her a big debt of appreciation for that. She gave me important guidance and encouragement, which made the job easier for me to complete.

I'm also appreciative of the Virtual University for providing the useful information. I value their cooperation as I completed my job. I also want to thank everyone who has prayed for me and helped with this project in some way.

**ABSTRACT**

Sentiment analysis, also known as opinion mining or emotion AI, employs text analysis, computational linguistics, biometrics, natural language processing, and natural language understanding to systematically identify, extract, measure, and examine emotional states and subjective data. Businesses regularly use sentiment analysis to find out what customers think of a product.

An online customer has the choice to provide feedback after making a purchase. Feedback comes in the form of comments. It is no longer necessary for a new buyer to read every review before making a purchase. You must develop a tool that analyzes the tone of customer evaluations in order to determine the quality of a product.

It is an online store application where only registered users can browse products (which can be of any type) and their features and leave reviews. The consumer can also see the comments made by other customers. Every comment made regarding a product will be evaluated and rated by the software.

**Table of Contents**

[**Chapter 1:** INTRODUCTION 8](#_Toc140584838)

[1.1. SENTIMENT ANALYSIS: 9](#_Toc140584839)

[1.2. CUSTOMER SENTIMENT ANALYSIS 9](#_Toc140584840)

[1.3. TYPES OF SENTIMENT ANALYSIS: 9](#_Toc140584841)

[1.4. WHY SENTIMENT ANALYSIS IS IMPORTANT: 10](#_Toc140584842)

[1.5. SENTIMENT ANALYSIS OF PRODUCTS: 11](#_Toc140584843)

[1.6. DEFINITIONS, ACRONYMS AND ABBREVIATIONS: 11](#_Toc140584844)

[**Chapter 2:** METHODOLOGY 13](#_Toc140584845)

[2.1. INTRODUCTION: 14](#_Toc140584846)

[2.2. DATA DOMAIN: 14](#_Toc140584847)

[2.3 PREPROCCESSING: 15](#_Toc140584848)

[2.4. SENTIMENT ANALYSIS USING TOOL: 16](#_Toc140584849)

[2.4.1) SENTIMENT SCORES: 17](#_Toc140584850)

[2.4.2) KEYWORDS CLOUD: 18](#_Toc140584851)

[2.5. MANUAL SENTIMENT ANALYSIS: 19](#_Toc140584852)

[2.5.1.) SENTIMENT SCORES: 20](#_Toc140584853)

[2.5.2.) KEYWORD CLOUDS: 20](#_Toc140584854)

[2.6. WORK PLAN: 21](#_Toc140584855)

[2.7. PROJECT STRUCTURE: 22](#_Toc140584856)

[2.7.1.) TEAM STRUCTURE: 24](#_Toc140584857)

[2.7.2.) PROJECT SCHEDULE (SUBMISSION CALENDAR): 24](#_Toc140584858)

[**Chapter 3:** RESULTS AND EVALUATION 25](#_Toc140584859)

[3.1. INTRODUCTION: 26](#_Toc140584860)

[3.2. VISUALIZATION: 26](#_Toc140584861)

[3.3. ACCURACY: 28](#_Toc140584862)

[3.4. ANALYSIS: 31](#_Toc140584863)

[**Chapter 4:** CONCLUSION AND FUTURE WORK 34](#_Toc140584864)

[4.1. CONCLUSION: 35](#_Toc140584865)

[4.1.1. Conclusion: 35](#_Toc140584866)

[4.1.2. Implications and Recommendations: 35](#_Toc140584867)

[4.1.3. Limitations and Challenges: 35](#_Toc140584868)

[4.2. FUTURE WORK: 36](#_Toc140584869)

[**REFERENCES** 36](#_Toc140584870)

[**APPENDIX** 36](#_Toc140584871)

|  |
| --- |
| **Chapter 1:** INTRODUCTION |

## SENTIMENT ANALYSIS:

Sentiment analysis, commonly referred to as opinion mining or emotion AI, is the methodical identification, extraction, quantification, and study of affective states and subjective data using computational linguistics, natural language processing, text analysis, and biometrics. In business, sentiment analysis is frequently used to determine how customers feel about a product.

* 1. CUSTOMER SENTIMENT ANALYSIS**:**

Customer sentiment analysis is the act of examining and comprehending the feelings, thoughts, and behaviors customers exhibit toward a good, name, or service. Customer sentiment analysis in the context of your project refers to the study of user comments and ratings to ascertain the sentiment related to the acquired products.

## 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:

There are various sorts of sentiment analysis approaches that can be applied in the context of customer sentiment analysis. **Here are three typical examples:**

* **Lexicon-Based Sentiment Analysis:** 
  + This method of sentiment analysis makes use of lexicons or dictionaries that have already been created and contain words or phrases along with corresponding sentiment ratings. A polarity score is given to each word or phrase, indicating whether it is positive, negative, or neutral.
  + This method determines the sentiment of a text by computing the overall sentiment score from the sentiment scores of the words that make up the text.
* **Rule-Based Sentiment Analysis:**
  + Sentiment analysis using specific rules or patterns to create sentiment expressions is known as rule-based sentiment analysis. Typically, linguists or domain experts create these rules.
  + Based on the presence of predetermined linguistic patterns or rules connected to positive or negative sentiment, the text is assessed. By assessing the existence and significance of these patterns, the sentiment of the text can be ascertained.
* **Machine Learning-Based Sentiment Analysis:**
  + To identify and categorize sentiment in text, machine learning-based sentiment analysis entails building a model on a labeled dataset. During the training phase, the model discovers patterns and connections between textual features and corresponding sentiment labels. By removing pertinent features and putting the patterns it has learnt to use, the model can predict the sentiment of new, unknown text after training. Naive Bayes, Support Vector Machines (SVM), and Recurrent Neural Networks (RNN) are a few examples of machine learning techniques frequently employed for sentiment analysis.

## 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.

## 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.

## 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 |

## **2.1. 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.

## 2.2. 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

## 2.3 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.

## 2.4. 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:**

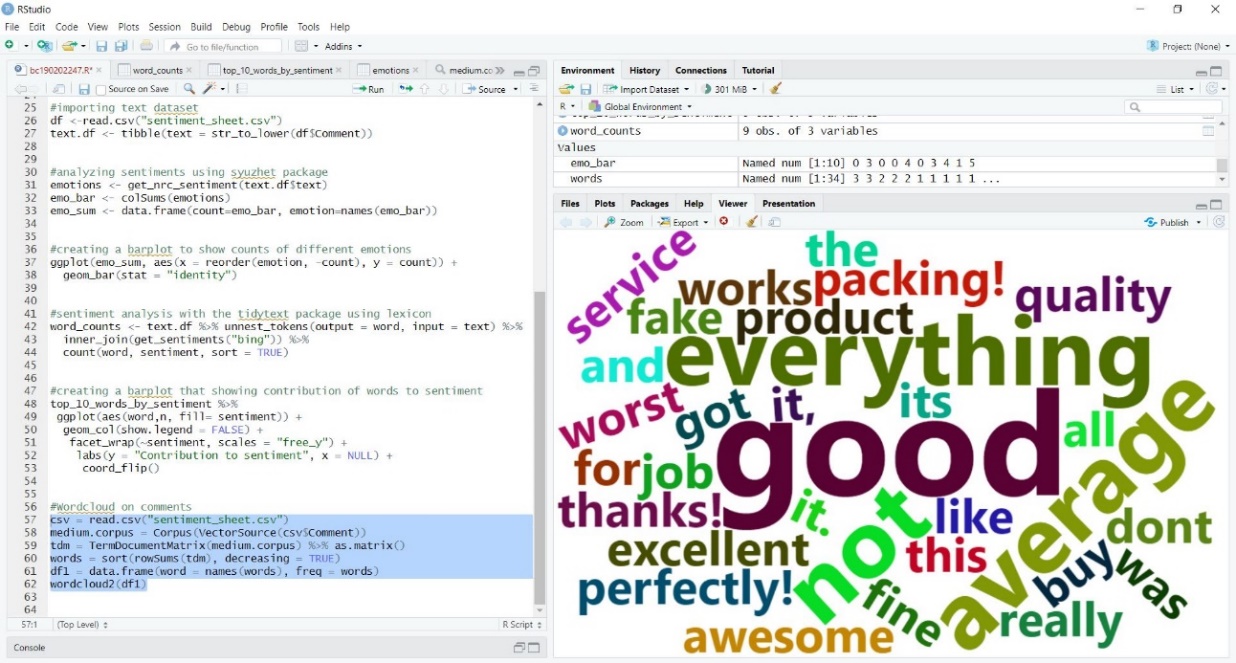
Sentiment scores in the context of sentiment analysis represent numerical values that quantify the sentiment expressed in a given text. These scores indicate the strength or intensity of the sentiment and help in assessing the overall sentiment polarity, whether positive, negative, or neutral. Here's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

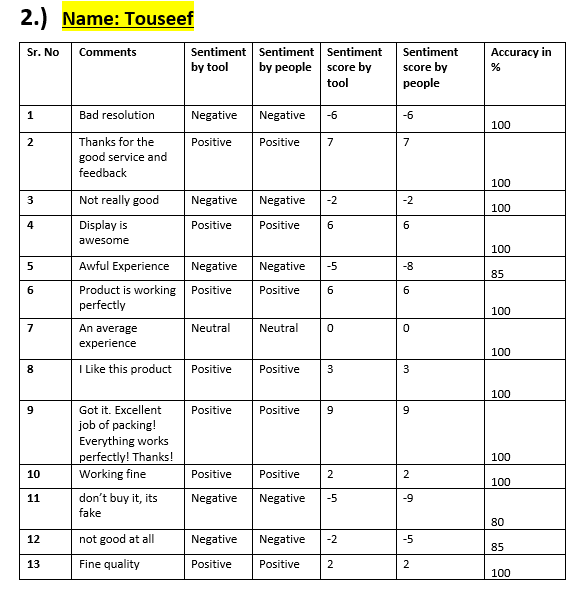


**Fig.01 Keyword Clouds**

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.

## 2.5. 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.

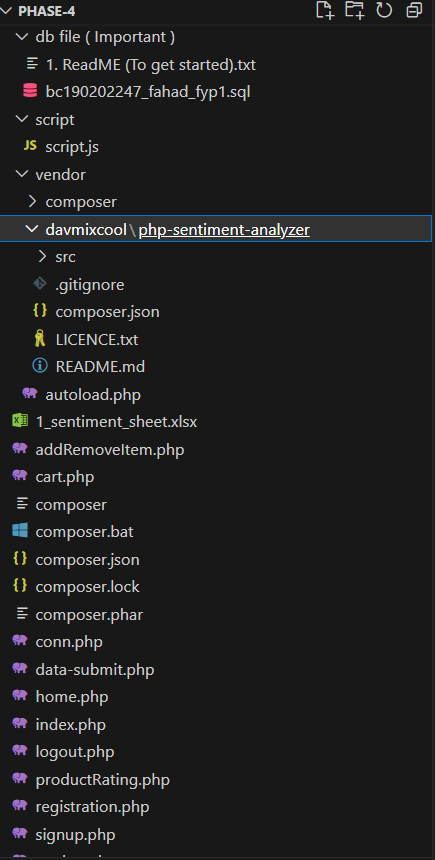
## 2.6. WORK PLAN:

|  |  |
| --- | --- |
| **Phase #** | **Plans** |
| **01** | • Make an online shopping website.  • User can view products listed on the website and can add these into the cart after registering.  • User can purchase one or more products.  • Website must offer the facility of feedback with each product.  • The user will Comment on the products based on their usage experience.  • Store the data in the normalized database.  •Store the comments properly with respect to products. • Register multiple users on the website and add comments against each product.  • Comments should be precise and short; For Example; Product quality is good/ bad/ average etc. |
| **02** | **Preprocessing:**  **a.** I will gathered the data in Phase 01 and perform preprocessing on the collected data. Each product has its own comments and will be handled independently.  **b.** Avoid the irrelevant and meaningless comments.  **c.** Do this activity for each product comments. |
| **03** | **Visualizing:**  **a.** In order to visualize sentiment scores of product comments, use the ggplot2 library of R statistical tool. Using this tool, a scatter plot will be created which shows sentiment categories for product comments. **b.** Create word clouds for each type (positive, negative and neutral) as shown below for the keywords obtained from the comments in phase 02.  **Accuracy of tests:** In this phase conduct a user case study to check the accuracy of analysis outcome. Take a group of 10 users to give feedback as positive (100) and negative (-100) to product comments (according to your own scale used in phase 2). Each user should assess 30 comments and then results of experiments performed on the public tool were compared with score of user feedback. Use the accuracy formula for this purpose. |
| **04** | **Task 1: Product Rating**  So far you have calculated the sentiment scores and the overall tone of the comment that whether it is positive or negative. Now is the time to use this knowledge and rate your product on the website created during Phase I.  This is entirely based upon your own research and coding abilities. You yourself have to search the method to convert these sentiment scores into stars. Based upon this star rating, a new user can decide about the purchase of the product.  **2. Task 2: Final Deliverable Document**  The template of the final deliverable document is included in the folder. Thoroughly read the template and fill the document accordingly. You can take help from internet, books and other resources. |

## 2.7. 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 |

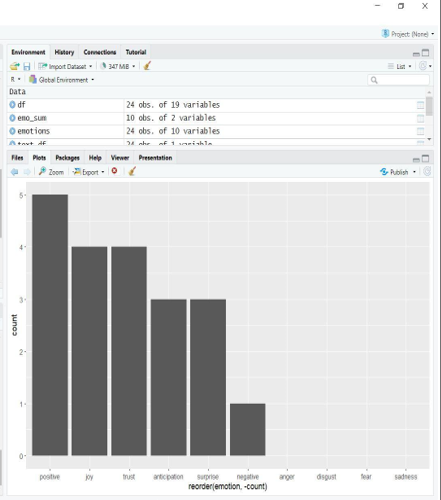
## 3.1. INTRODUCTION:

The "Results and Evaluation" chapter focuses on evaluating the outcomes and effectiveness of the developed online shopping website with sentiment analysis capabilities. This chapter presents an assessment of the project's success in achieving its goals and objectives, along with an analysis of the obtained results. The evaluation process aims to determine the impact and usefulness of the sentiment analysis feature, analyze the sentiments expressed by users, and assess the overall performance of the website. By evaluating the results, we gain valuable insights into the effectiveness of the implemented solution and its potential for enhancing the user experience, decision-making, and customer satisfaction. This section provides an overview of the evaluation process conducted and outlines the key findings and conclusions derived from the analysis of the obtained results.

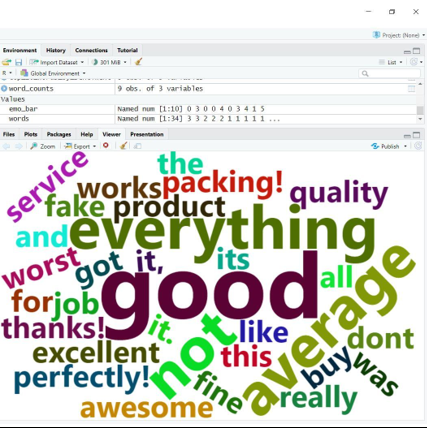
## 3.2. VISUALIZATION:

Visualization plays a crucial role in presenting and interpreting the collected data, analysis, and findings. Visualizations provide a clear and concise way to communicate complex information, patterns, and relationships to the readers. Here are some points to consider regarding visualization:

* **Visualizing Sentiment Scores:** Use visualizations, such as bar charts or line graphs, to present the sentiment scores obtained from the sentiment analysis process. This allows for a quick understanding of the distribution and variations in sentiment across different products or user comments.

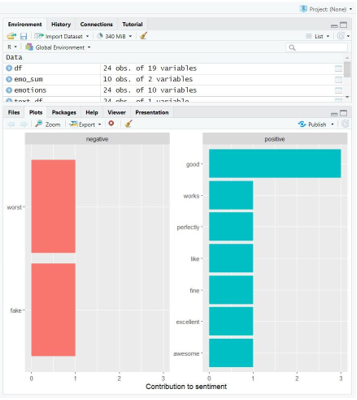


**Fig.04 Bar Plot**

* **Keyword Clouds:** Include the screenshot or generated word clouds that represent the prominent keywords associated with each sentiment category (positive, negative, or neutral). These keyword clouds visually highlight the frequently occurring words, providing insights into the main themes or topics expressed by users.

**Fig.05 Keywords Cloud**

* **Scatter Plots:** Utilize scatter plots to visualize the sentiment categories and sentiment scores of product comments. This type of visualization can help identify patterns, clusters, or outliers in the sentiment analysis results.



**Fig.06 Scatter plots**

* **Comparative Visualizations:** Create visual comparisons between different sentiment categories, such as side-by-side bar charts or stacked area charts. This allows for a visual assessment of the differences in sentiment distribution and helps draw meaningful comparisons between positive, negative, and neutral sentiments.

**Fig.07 Visual image of comments Filled later**

* **Word Frequency Charts:** Present word frequency charts or histograms to display the occurrence frequencies of specific words or phrases within the user comments. This can help identify the most commonly mentioned aspects, features, or issues related to the products.

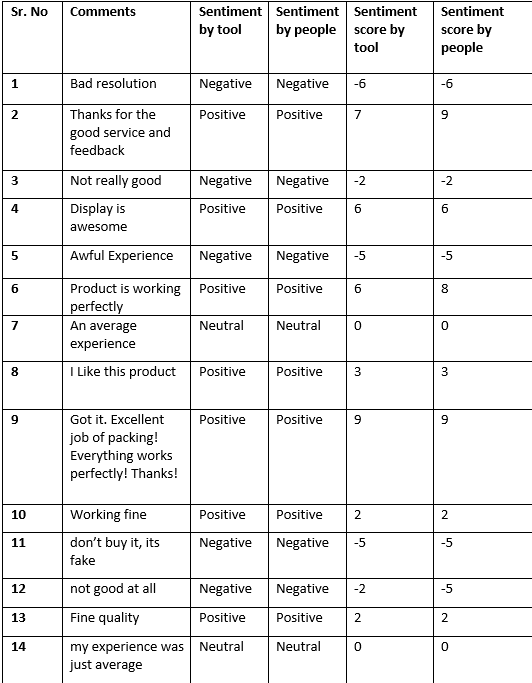


**Fig.08 Frequency Chart**

## 3.3. ACCURACY:

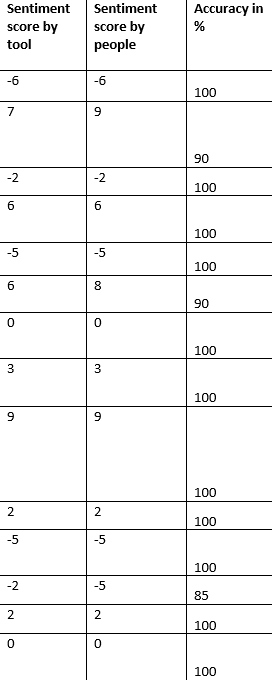
Accuracy refers to the measure of how closely the sentiment scores generated by the sentiment analysis tool align with the sentiment ratings provided by users. It provides an indication of how well the tool's automated sentiment analysis matches the sentiments expressed by users in their comments.

* **Data Comparison:**
  + Compare the sentiment scores generated by the sentiment analysis tool for each comment with the sentiment ratings provided by users
  + This involves matching the sentiment scores and sentiment ratings for each comment to determine if they align in terms of positive, negative, or neutral sentiment issues related to the products.



**Fig.09 Data Comparison**

* **Accuracy Calculation:**
  + Calculate the accuracy by determining the percentage of comments for which the sentiment scores from the tool match the sentiment ratings assigned by users.
  + Accuracy is usually computed by dividing the number of correctly matched sentiment scores by the total number of comments and multiplying by 100.



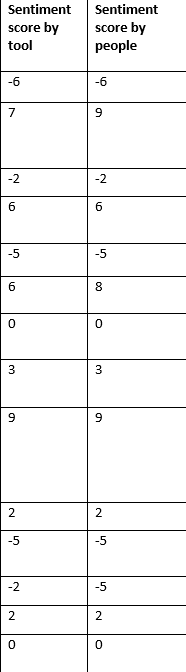
**Fig.10 Accuracy Calculation**

* **Interpretation of Accuracy::**
  + A higher accuracy percentage indicates a stronger alignment between the sentiment analysis tool and user sentiment ratings, suggesting that the tool effectively captures the sentiment expressed in the comments.
  + Conversely, a lower accuracy percentage indicates a weaker alignment, highlighting potential discrepancies between the tool's automated analysis and users' subjective evaluations.
* **Insights and Improvements:**
  + Analyze the accuracy results to gain insights into the performance of the sentiment analysis tool.
  + Identify any patterns or areas where the tool's sentiment scores consistently deviate from the user ratings, and consider potential improvements or adjustments to enhance the accuracy.

## 3.4. ANALYSIS:

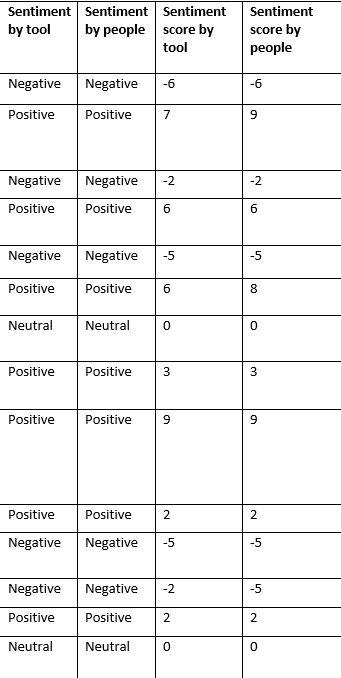
In the context of this project, the "Analysis" section refers to the interpretation and discussion of the collected data, results, and findings from the sentiment analysis and evaluation process. It involves a deeper exploration and understanding of the patterns, trends, and insights derived from the analysis of user comments, sentiment scores, and other relevant data.

* **Interpretation of Sentiment Scores:**
  + Analyze the sentiment scores obtained from the sentiment analysis tool, considering both the automated sentiment analysis results and the sentiment ratings provided by users.
  + Identify any notable trends or patterns in the sentiment scores across different products or user comments.
  + Discuss the distribution of sentiment scores and their implications for overall user sentiment towards the products.



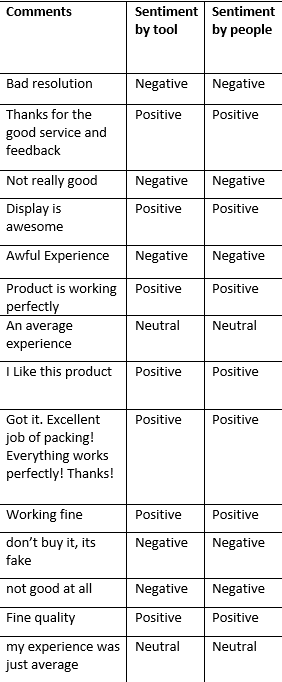
**Fig.11 Analyzing Sentiment Scores**

* **Comparison of Tool Scores and User Ratings:**
  + Discuss the accuracy calculation results, highlighting the percentage of agreement between the sentiment scores generated by the tool and the sentiment ratings assigned by users.
  + Interpret the implications of the accuracy results and identify any discrepancies or alignment between the tool's automated analysis and users' subjective evaluations.



**Fig.12 Tool vs people**

* **Keyword Analysis:**
  + Analyze the keyword clouds generated from the user comments to identify the most frequently occurring words or themes associated with each sentiment category (positive, negative, neutral).
  + Discuss the prominent keywords and their relevance to the sentiment expressed by users, providing insights into the aspects, features, or experiences that influence user sentiment.



**Fig.13 Keyword Analysis**

* **Trends and Insights:**
  + Identify any significant trends, patterns, or shifts in sentiment over time or across different product categories.
  + Discuss any notable findings or insights derived from the analysis, such as common strengths, weaknesses, or areas of improvement based on user feedback.
* **Implications and Recommendations:**
  + Discuss the implications of the analysis results on the overall user experience, customer satisfaction, and decision-making process for potential buyers.
  + Provide recommendations or suggestions for improving the online shopping website based on the analysis findings, such as enhancing product descriptions, addressing common issues, or leveraging positive sentiment to boost sales.

## 

|  |
| --- |
| **Chapter 4:** CONCLUSION AND FUTURE WORK |

## 4.1. CONCLUSION:

### **4.1.1. Conclusion:**

* In the conclusion, I am summarizing the key findings and outcomes of this project. Highlight the successful development of the online shopping website with sentiment analysis capabilities.
* Discussed how the sentiment analysis tool effectively analyzed user comments, providing valuable insights into customer sentiment towards the purchased products.
* Emphasize the positive impact of sentiment analysis on the decision-making process for potential buyers and the overall user experience on the website.

### **4.1.2. Implications and Recommendations:**

* Based on the analysis results, identify the implications for online shopping website. Discuss the strengths, weaknesses, opportunities, and challenges identified during the evaluation process.
* Recommend improvements to enhance the sentiment analysis feature, such as refining the sentiment scoring algorithm or expanding the lexicon to capture a broader range of sentiments.
* Suggested ways to leverage the sentiment analysis results, such as utilizing positive sentiment to promote highly rated products or addressing negative sentiment through proactive customer support.

### **4.1.3. Limitations and Challenges:**

* Acknowledge any limitations or challenges encountered during this project, like limited availability of training data for sentiment analysis or potential biases in user-generated content.
* These limitations may have influenced the accuracy or comprehensiveness of the sentiment analysis results.
* The importance of addressing these limitations in future work to improve the robustness and reliability of the sentiment analysis component.

## 4.2. FUTURE WORK:

* Considering integrating advanced sentiment analysis techniques like aspect-based sentiment analysis or sentiment analysis in multiple languages.
* Propose the integration of user feedback loops to continuously update and enhance the sentiment analysis model.
* Explore the possibility of integrating sentiment analysis with personalized product recommendations to offer tailored suggestions based on individual sentiment preferences.

# **REFERENCES**

1. [**https://www.youtube.com/watch?v=hu-q2zYwEYs&list=PL4cUxeGkcC9ivBf\_eKCPIAYXWzLlPAm6G**](https://www.youtube.com/watch?v=hu-q2zYwEYs&list=PL4cUxeGkcC9ivBf_eKCPIAYXWzLlPAm6G)
2. [**https://www.youtube.com/watch?v=1SnPKhCdlsU**](https://www.youtube.com/watch?v=1SnPKhCdlsU)
3. [**https://www.w3schools.com/php/**](https://www.w3schools.com/php/)
4. [**https://www.w3schools.com/MySQL/default.asp**](https://www.w3schools.com/MySQL/default.asp)
5. [**https://www.youtube.com/watch?v=1WGgtYnmGjU**](https://www.youtube.com/watch?v=1WGgtYnmGjU)
6. [**https://github.com/davmixcool/php-sentiment-analyzer**](https://github.com/davmixcool/php-sentiment-analyzer)

# **APPENDIX**

|  |  |
| --- | --- |
| Phase-1 | Prototype |
| Phase-2 | Gathering & Analyzing Info. |
| Phase-3 | Designing the Project |
| Phase-4 | Final Deliverable |