UIT2511 - Software Development Project – II

Project Report

Academic Year
2022-2023

BEST PICK

Sentiment Analysis for E-Commerce Products using Natural language Processing

Vijaymurugan N   -   3122 21 5002 124
Vishal Prakash -   3122 21 5002 125
Vishnu Siddharth VR - 3122 21 5002 126

Sri Sivasubramaniya Nadar College of Engineering
(An Autonomous Institution, Affiliated to Anna University)

BONAFIDE CERTIFICATE

Certified that this project titled "The Best Pick" is the bonafide work of "Vijaymurugan N - 3122215002124 , Vishal Prakash - 31222115002125 , Vishnu Siddharth V R - 3122215002126 ", and is submitted for Project Viva-Voce examination held on 29/11/2023

.

Signature of examiner(s)

INDEX

I. Abstract

The project "Best Pick: Sentiment Analysis for E-Commerce Products using Natural Language Processing" aims to develop a robust system that assists users in making informed decisions by analyzing sentiments in product reviews on Amazon. This report provides an overview of the project, its objectives, and the methodologies employed for sentiment analysis. The "Best Pick" application is designed to assist users in finding the top 3 products within a specified budget and product category on Amazon by scraping and analysing user reviews.

II. Introduction

"Best Pick" is a groundbreaking project designed to revolutionize the realm of e-commerce by employing advanced Natural Language Processing (NLP) techniques. The project addresses the challenges faced by online shoppers in navigating through an overwhelming array of product reviews on e-commerce platforms, particularly Amazon. In the age of information overload, making informed purchase decisions becomes increasingly complex.

Our project recognizes the critical role sentiment analysis plays in enhancing the online shopping experience. Sentiment analysis involves extracting valuable insights from user reviews, deciphering sentiments, and using this information to recommend the top 3 products tailored to individual preferences and budgets. Leveraging the power of NLP, our application, "Best Pick," goes beyond conventional product recommendations. It delves into the sentiment expressed by users in their reviews, ensuring that the recommended products not only meet functional criteria but also align with the user's emotional and subjective preferences.

By scraping and analyzing user reviews on Amazon, "Best Pick" employs sophisticated algorithms to rank products based on user-defined criteria. The application's user-friendly interface allows shoppers to input their desired product category and budget, initiating a seamless process that culminates in

personalized and sentiment-driven product recommendations. This innovative approach not only simplifies the decision-making process for users but also adds a human touch to the often impersonal world of online shopping.

In the following sections, we delve into the specific requirements, functionalities, and performance metrics of "Best Pick," outlining its potential for future extensions and providing supporting information. As we navigate through the intricacies of the project, it becomes evident that "Best Pick" is poised to redefine the landscape of e-commerce, offering a more personalized, user-centric, and sentiment-aware shopping experience.

III. Motivation

Motivation for the project stems from the growing importance of sentiment analysis in the e-commerce domain. The need to enhance the online shopping experience by providing users with personalized and sentiment-driven product recommendations serves as a driving force for the project.

IV. Objective

The primary objective of the project is to develop an intelligent system that leverages Natural Language Processing (NLP) techniques to analyze sentiments in user reviews on Amazon. The system aims to assist users in identifying the top 3 products within a specified budget and product category.

V. Problem Statement

To create a fair recommendation engine for e-commerce products. The existing search pages often rank based on promotions, SEO, etc. "Amazon's choice" may not always be the best product to buy when we enter a search term.

"Best Pick" is an AI-based sentiment analysis tool created to provide e-commerce product recommendations to it's users when provided with a search term.

VI. Literature Survey

Understanding Sentiment Analysis in E-commerce:

Literature surrounding sentiment analysis in the e-commerce domain serves as a foundational exploration of the relevance and impact of user sentiments in online product reviews. Researchers have investigated the effectiveness of sentiment analysis in deciphering user preferences, satisfaction levels, and potential areas for improvement. The identification of sentiment-related indicators, such as positive and negative language use, forms the motivation for integrating sentiment analysis into the e-commerce decision-making process.

Exploring Natural Language Processing (NLP) in E-commerce:

The literature survey highlights the pivotal role of Natural Language Processing (NLP) in enhancing the analysis of textual data, particularly in the context of e-commerce product reviews. Various studies delve into the application of NLP techniques, including sentiment analysis, to extract meaningful insights from unstructured text. The integration of NLP in understanding user sentiments becomes crucial for developing intelligent recommendation systems, aligning with the objectives of the "Best Pick" project.

Utilizing Machine Learning Algorithms for Product Recommendations:

Research in the field of machine learning algorithms for product recommendations forms a critical aspect of the literature survey. Studies explore the application of algorithms such as collaborative filtering, content-based filtering, and hybrid approaches in generating personalized recommendations. The effectiveness of these algorithms in capturing user preferences and enhancing the overall shopping experience underscores the importance of leveraging machine learning techniques in the "Best Pick" project.

Addressing Challenges in E-commerce Data Retrieval:

The literature survey delves into the challenges associated with data retrieval from e-commerce platforms, drawing parallels with the "Best Pick" project's reliance on scraping Amazon's website. Studies discuss the ethical considerations, legal implications, and technical challenges involved in web scraping for data extraction. Insights from these studies contribute to the risk management and design constraints considerations within the project.

Analyzing User Interface Design for Recommendation Systems:

User interface design for recommendation systems is explored in the literature to understand the impact of interface elements on user engagement and satisfaction. Research emphasizes the need for intuitive interfaces that facilitate user input, preferences customization, and clear presentation of recommendations. These insights guide the user interface design requirements in the "Best Pick" project, ensuring a seamless and user-friendly experience.

Incorporating Sentiment Analysis in Recommender Systems:

The literature survey discusses the integration of sentiment analysis into recommender systems, aligning with the sentiment-driven approach of the "Best Pick" project. Studies explore how sentiment-aware recommendations can enhance user satisfaction and decision-making. The identification of sentiment-based criteria for ranking products contributes to the project's specific requirements and functionalities

VII. Deliverables

This section outlines the tangible outcomes expected from the project. It includes the "Best Pick" application, which facilitates user-friendly input, data retrieval, sentiment analysis, and the presentation of top 3 product recommendations. The project's deliverables encompass a robust and user-friendly web application, featuring an advanced sentiment analysis system tailored for e-commerce product evaluations.

The primary components include:
1. User Friendly Interface (UI)
2. Data Retrieval System
3. Review Retrieval System
4. Analysis and Ranking Engine
5. Recommendation System
6. User Search History
7. User Feedback Mechanism
8. Data Security Measures
9. Customizable Recommendation Criteria

VIII. Requirements Engineering

Functional Requirements:

1)User Input:
- Users can input product category and budget through a user-friendly interface.
- The interface includes validation checks for accurate and complete input.

2)Data Retrieval:
- The system scrapes Amazon for product information within the specified category.
- It employs robust web scraping techniques to handle various page structures.

3)Review Retrieval:
- The system retrieves user reviews for the top 10 products based on user-defined criteria.
- User reviews are extracted and stored to provide relevant insights.

4)Analysis and Ranking:
- The system analyzes reviews using sentiment analysis algorithms.
- Transparent ranking criteria are applied to identify and rank the top 3 products.

5)Recommendation:
- The system presents the top 3 recommended products.
- Clear and concise display of product information, price, and a link to the product page on Amazon is provided.

6)User Search History:
- Users can see their already searched preferences(ie Search History), including brand preferences, price range, and specific product features.

7)Data Security:
- User data is securely stored, encrypted, and protected from unauthorized access.
- Compliance with data protection regulations and industry standards.

8)Customizable Recommendation Criteria:
- Users can customize the criteria for product recommendations, such as weighting sentiment analysis, price, or brand preferences.
- Fine-tuning recommendations to align with specific user needs.

Non-functional Requirements:

1)Performance Optimization:
- Continuous monitoring and optimization of performance metrics.
- System responds within 30 seconds for user queries.

2)Error Handling and Logging:
- Robust error-handling mechanism to log and report errors for debugging.
- Informative error messages provided to users in case of issues.

3)Redundancy and Backup:
- Redundancy measures implemented for data integrity.
- Regular automated backups conducted for disaster recovery.

4)Internationalization and Localization:
The application designed to support internationalization, adapting to different languages, currencies, and regional preferences.

5)Accessibility:
User interface adheres to accessibility standards (e.g., WCAG) for users with disabilities.

6)Compatibility: The system must be compatible with major web browsers, such as Chrome and Firefox, ensuring accessibility and a seamless user experience.

Epic	Sprint	User Story #	User Story	Essential/ Desirable	Need of requirement	Description of requirement
1	1	1	User Input	Essential	.User enters the product category and budget.	The system presents a user-friendly interface for entering the desired product category and budget.
		2	Data Retrieval	Essential	Collect relevant data for system training.	The system accesses Amazon's product listings and retrieves relevant data.
		3	Review Retrieval	Essential	Reviews data for analysing and retrieves user reviews for top 10 products	The system identifies the top 10 products based on user-defined criteria and retrieves user reviews for the top 10 products.
	2	4	Analysis and Ranking	Essential	Provides users with clear insights into product rankings for informed	Analyze user reviews using sentiment analysis and transparently rank the top 3 products

					decision-making	
		5	Recommendation	Essential	Offers users easily accessible and relevant product recommendations	Present top 3 recommended products with clear details and an Amazon link
		6	User Preferences	Desirable	Enhances user experience by tailoring product recommendations	Allow users to set and save preferences, with the option to update them.
2	3	7	Testing with Various Sample inputs	Essential	Assess system performance with diverse sample inputs	Assess BESTPICK's performance and accuracy using diverse by sample keyword inputs
	4	8	BESTPICK Integration w/ Flask Phase 1	Essential	Integrate BESTPICK with Flask for backend functionality.	Integrate the BESTPICK system with Flask (Phase 1) for backend functionality.
		9	BESTPICK Integration w/ Flask Phase 2	Essential	Ensure seamless frontend-backend communication	Integrate the BESTPICL system with Flask for seamless frontend-backend communication.

		10	Deployment	Desirable	For public access.	Deploy the BESTPICK system online for public access and use.

<div align="center">Table 1</div>

IX. Risk Management

Anticipated Risks:

- **Quality of Sentiment Analysis Data:**
  Poor quality or insufficient sentiment analysis data may impact the accuracy of product recommendations and user satisfaction.

- **Timely Completion of Development:**
  Completing essential development tasks within the specified time frame is critical for the success of the e-commerce sentiment analysis platform.

- **Changes in E-commerce Policies:**
  Changes in e-commerce policies and structures may impact the platform's ability to retrieve and analyze product data effectively.

Mitigation Plan:

1. **Quality of Sentiment Analysis Data:**

- Implement robust sentiment analysis algorithms that can handle various linguistic nuances.
- Regularly update the sentiment analysis model based on user feedback and changing language patterns.
- Incorporate user feedback mechanisms to continuously improve the accuracy of sentiment analysis.

2. Timely Completion of Development:

● Utilize Agile development methodologies, employing tools like JIRA, to break down tasks and enhance productivity.
● Prioritize feature development based on user needs and platform requirements.
● Foster open communication among team members to promptly address any challenges or roadblocks.

3 . Changes in E-commerce Policies:

● Stay informed about changes in e-commerce policies and platforms.
● Design the platform with flexibility to adapt to variations in e-commerce structures.
● Establish a process for regular updates and maintenance to align with evolving e-commerce policies.

Implementation and Risk Management

Name: Vishnu Siddharth VR

Register Number: 3122 21 5002 126

Roles: Scrum Master, Developer

Implementation

Epic	Sprint	User Story #	User Story	Essential/ Desirable	Need of requirement	Description of requirement
1	1	1	User Input	Essential	User enters the product category and budget.	The system presents a user-friendly interface for entering the desired product category and budget.
	2	7	Data Retrieval	Essential	Collect relevant data for system training.	The system accesses Amazon's product listings and retrieves relevant data.
2	3	8	Review Retrieval	Essential	Reviews data for analysing and retrieves user reviews for top 10 products	The system identifies the top 10 products based on user-defined criteria and retrieves user reviews for the top 10 products.
	4	2	Analysis and Ranking	Essential	Provides users with clear insights into product rankings for informed decision-making	Analyze user reviews using sentiment analysis and transparently rank the top 3 products
		6	Recommen dation	Essential	Offers users easily accessible and relevant product suggestions	Present top 3 recommended products with clear details and an Amazon link

Table 2

Risk Management 1:

Risk #	Risk Description	Probability	Impact	Mitigation Plan
1	Scrum Master's lack of availability due to conflicting responsibilities	High	High	Plan and prioritize the Scrum Master's responsibilities to minimize conflicts with backend programming tasks. Allocate additional resources or delegate specific Scrum Master duties to other team members when necessary to ensure the smooth functioning of the Scrum process
2	Insufficient knowledge or expertise in Scrum processes impacting project efficiency	Moderate	Moderate	Provide training and resources to enhance the Scrum Master's knowledge of Scrum methodologies and best practices. Encourage participation in relevant workshops or certification programs. Collaborate with the Scrum Master to create a learning plan to continuously improve their Scrum expertise.

Table 3

Test Log report

TC id	RS #	Test case description/ condition	Test case input	Expected Output	Result (PASS/ FAIL)
1	1	Handling Human Queries	Queries in Colloquial Slang	The output recommendations were as per the input.	PASS
2	2	Error Handling	Triggering errors or exceptions	Appropriate error messages	PASS
3	3	Sprint planning meeting	User stories and backlog	Prioritized backlog and sprint plan created	PASS
4	4	Sprint review meeting	Completed user stories and demos	Stakeholder satisfaction and feedback	PASS

Table 4

Implementation and Risk Management

Name: Vishal Prakash

Register Number: 3122 21 5002 125

Roles: Developer

Implementation

Epic	Sprint	User Story #	User Story	Essential/ Desirable	Need of requirement	Description of requirement
1	1	11	Bestpick Integration w/ Flask Phase 1	Essential	Integrate BESTPICK with Flask for backend functionality.	Integrate the BESTPICK system with Flask (Phase 1) for backend functionality.
	2	12	Bestpick Integration w/ Flask Phase 2	Essential	Ensure seamless frontend-backend communication	Integrate the BESTPICK system with Flask (Phase 2) for seamless frontend-backend communication.
2	3	9	User Preferences	Desirable	Enhances user experience by tailoring product recommendations	Allow users to set and save their search history preferences

Table 8

Risk Management 1:

Risk #	Risk Description	Probability	Impact	Mitigation Plan
1	Privacy and Security Concerns	Moderate	High	Implement strong data encryption methods. Comply with data protection regulations (e.g., GDPR, CCPA). Regularly conduct security audits and updates.
2	Scalability Issues	Moderate to High	High	Employ scalable infrastructure and cloud services. Optimize code and architecture for better scalability.

Test Log report

TC id	RS #	Test case description/ condition	Test case input	Expected Output	Result (PASS/ FAIL)
1	1	Research Accuracy	Conducting research on a topic	Accurate and reliable findings	PASS
2.	2	Vague user input	Text input	Products from different categories	PASS
3	3	Limited Products	Text input	Top three results must be generated	PASS
4	4	Random Keyword	Text Input	Error prompt	FAIL

Implementation and Risk Management

Name: Vijaymurugan N

Register Number: 3122 21 5002 124

Roles: Developer

Implementation

Epic	Sprint	User Story #	User Story	Essential/ Desirable	Need of requirement	Description of requirement
1	4		Testing with Various Sample Inputs	Essential	Assess system performance with sample inputs	Assess BESTPICK's performance and accuracy using diverse sample keyword inputs .

Table 11

Risk Management 1:

Risk #	Risk Description	Probability	Impact	Mitigation Plan
01	Amazon frequently updates its website, and any structural changes may disrupt the web scraping process	Medium	High	Regularly monitor Amazon's website for changes, implement flexibility in the scraping algorithm to adapt to structural changes, and maintain an active communication channel with Amazon for any API updates

02	Storing user preferences and reviews poses a risk to data security and privacy.	Medium	Medium	Implement robust encryption methods, regularly audit and update security measures, anonymize user data whenever possible, and adhere to data protection regulations.
03	Sudden increases in user traffic, especially during peak shopping seasons, may lead to performance issues and decreased system responsiveness	Medium	High	Implement load balancing, employ scalable infrastructure, conduct regular load testing, and optimize system performance for scalability
04	The system relies on external services, such as amazon's website data and sentiment analysis, and their unavailability may impact system functionality.	Medium	High	Implement fallback mechanisms for critical external services, have alternative providers for essential services, and regularly check the availability of dependent services
05	Users may not find the recommendations useful, leading to low user engagement and adoption.	Less	Medium	Conduct user surveys and feedback sessions during development, continuously refine recommendation algorithms based on user feedback, and provide clear explanations for how recommendations are generated.

Test Log report :

TC id	RS #	Test case description/ condition	Test case input	Expected Output	Result (PASS/ FAIL)
1	2	Handling different audio file formats.	Audio file .wav Audio file .acc	Algorithm successfully processes and analyzes audio files in .wav and .acc formats.	PASS
2.	2	Navigation Menu Test	Clicking on menu items	Proper navigation to respective pages	PASS
3	3	Form Validation Test	Entering invalid input in form fields	Error messages displayed appropriately	PASS
4	4	Mobile Responsiveness	Resizing browser to mobile	Content adjusts correctly for mobile view	PASS

Table 13

X. Project Management

During the course of the project, effective process management and team collaboration were essential in ensuring the successful development of the student accommodation finding system. The team followed an iterative and collaborative approach, employing the Agile methodology with the Scrum framework to manage the project effectively.

The project consisted of two main epics.

Epic 1: Foundation and Data Gathering

This epic concentrates on establishing the foundational elements by conducting comprehensive research and gathering a diverse dataset for sentiment analysis in BestPick. Thorough data acquisition and preprocessing techniques are employed to ensure the dataset's quality, cleanliness, and suitability for subsequent model training.

Epic 2: Algorithm Development and Recommendation System Integration

Focused on advancing algorithms, this epic involves the creation and refinement of feature extraction algorithms, sentiment analysis models using Natural Language Processing (NLP) techniques, collaborative filtering, and heuristic evaluation. The integration of these sophisticated algorithms aims to build a cohesive and robust recommendation system in BestPick, enhancing the accuracy and reliability of product recommendations based on sentiment analysis.
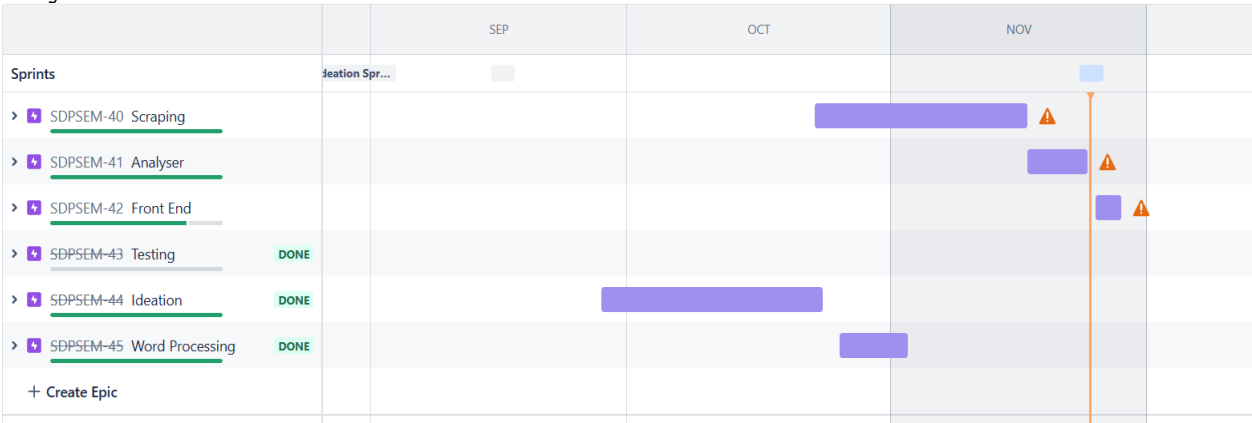
Epic 1: Data Scraping

Sprints	Spr...

SDPSEM-40 Scraping

- SDPSEM-15 scraping fro... **DONE** VS
- SDPSEM-16 extracting a... **DONE** VS
- SDPSEM-29 beautiful so... **DONE** VS
- SDPSEM-34 data structu... **DONE** VS
- SDPSEM-36 flask integra... **DONE** VP

Epic 2: Analyser - For Sentiment Analysis

SDPSEM-41 Analyser

- SDPSEM-14 Analyser m... **DONE** VS
- SDPSEM-27 Naive Bayes... **DONE** VS
- SDPSEM-28 score conve... **DONE** VS
- SDPSEM-18 Integrating ... **DONE** VS
- SDPSEM-31 model build... **DONE** VS
- SDPSEM-30 naive bayes ... **DONE** VS
- SDPSEM-32 prediction a... **DONE** VS
- SDPSEM-33 integrating ... **DONE** VS
- SDPSEM-35 modularizat... **DONE** VS
- SDPSEM-19 front end la... **DONE**
- SDPSEM-25 search page... **DONE** VP
- SDPSEM-24 search page... **DONE** VP
- SDPSEM-26 search page... **DONE** VP
- SDPSEM-21 home page ... **DONE** VP
- SDPSEM-20 front end d... **DONE** VP
- SDPSEM-23 home page ... **DONE** VP
- SDPSEM-22 home page ... **DONE** VP

Miscellaneous Epic:
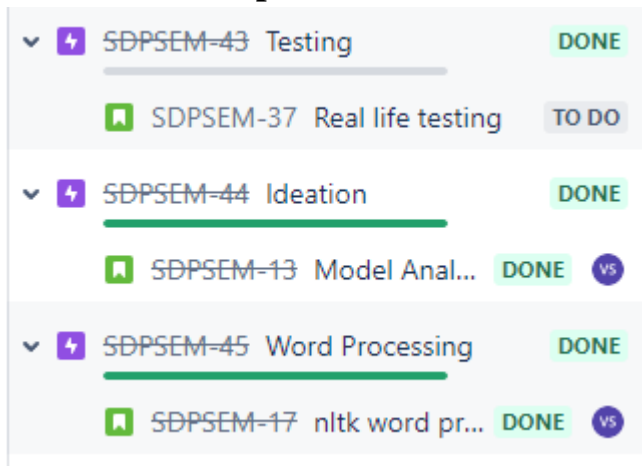
Sprint 1: Setting up Project Infrastructure

1. Decide on Title: Define the project title and scope through collaborative discussions with the team.
2. Setup Jira: Establish the project management framework using Jira, defining workflows, and user roles.
3. Decide on Algorithm: Research and decide on the appropriate algorithm(s) to be used in the system development.
4. Setup Github repo: Create a version-controlled environment on Github, initializing the repository for collaborative coding.
5. Setup Teams: Formulate and organize project teams, assigning roles and responsibilities to team members.
6. Setup Confluence: Establish a documentation platform using Confluence for collaborative content creation and knowledge sharing.
7. Learn about CoCoMo Model and apply: Research and understand the CoCoMo Model, applying its principles to estimate project effort and schedule.
8. Formulate the Effort equation: Develop an equation for estimating project effort, considering factors like team size, project complexity, and the CoCoMo Model.

Sprint 2: Planning System Architecture

9. Block Diagram: Create a detailed block diagram outlining the high-level structure and components of the system architecture.
10. IEEE Reference Paper: Identify and review relevant IEEE reference papers that provide insights into similar projects or methodologies.
11. Find what all libraries to be imported: Research and identify the necessary libraries and dependencies required for the implementation of the chosen algorithm and system.

Sprint 3: Sentiment Analysis Module Development

12. Analyser Module Blueprint: Develop a detailed blueprint for the sentiment analysis module, outlining the key features, functionalities, and the flow of data through the system.
13. Model Analysis: Conduct an in-depth analysis of various sentiment analysis models to determine their strengths and weaknesses, aiding in the selection of the most suitable model for implementation.
14. Naive Bayes Algorithm: Implement the Naive Bayes algorithm as the chosen sentiment analysis model, ensuring a comprehensive understanding of its principles and functionalities.
15. Score Converter from Naive Bayes: Develop a score converter module that transforms the output of the Naive Bayes algorithm into meaningful sentiment scores, facilitating better interpretation and application.

Sprint 4: Review Scraping

16. Scraping from Amazon: Develop a scraping module specifically designed for Amazon, ensuring efficient and ethical extraction of review data.
17. Extracting ASIN from Search Term: Implement functionality to extract Amazon Standard Identification Numbers (ASINs) from search terms, streamlining the process of retrieving reviews for specific products.

18. Data Structure Analysis for Handling Reviews: Analyze and design an effective data structure to handle and organize scraped review data, optimizing storage and retrieval processes.
19. Beautiful Soup Analysis: Conduct a thorough analysis of Beautiful Soup, a Python library for web scraping, and integrate it into the scraping module to enhance parsing capabilities.
20. Integrating Pre-processing with Scraping: Integrate pre-processing steps with the scraping module to ensure that the extracted review data is clean, structured, and ready for further analysis.

Sprint 5: Sentiment Analysis 2

21. NLTK Word Processing: Implement Natural Language Toolkit (NLTK) for advanced word processing tasks, enhancing the efficiency of text analysis for sentiment determination.
22. Prediction Analysis: Conduct an in-depth analysis of prediction methodologies, exploring different approaches to improve the accuracy and reliability of sentiment predictions.
23. Naive Bayes Analysis: Perform a comprehensive analysis of the Naive Bayes algorithm, identifying its strengths and weaknesses in the context of sentiment analysis.
24. Model Building: Initiate the process of building a sentiment analysis model, incorporating insights from previous analyses to ensure the model's effectiveness and adaptability.

Sprint 6: Integration and cleanup

25. Integrating Naive Bayes with Amazon Scraped Reviews: Implement the integration of the Naive Bayes sentiment analysis module with the dataset obtained from Amazon scraping, ensuring seamless compatibility and effective sentiment labeling.
26. Modularization of Different Functions: Modularize various functions within the system, promoting code organization and reusability. This includes breaking down different aspects like scraping, preprocessing, and sentiment analysis into distinct, manageable modules.

Sprint 7: Front End Development

27. Flask Integration with Backend: Integrate Flask with the backend system, establishing communication and ensuring seamless data flow between the user interface and the underlying logic.

28. Front End Design: Ideate and design the front end of the application, focusing on creating an engaging and user-friendly interface for a positive user experience.

29. Front End Ideation: Brainstorm and conceptualize ideas for the front-end design, considering aesthetics, usability, and overall user satisfaction.

30. Home Page Design Improvement: Enhance the design of the home page, incorporating improvements based on feedback and usability considerations.

31. Results Page Design Improvement: Improve the design of the results page to ensure a visually appealing and coherent presentation of search outcomes.

32. Home Page HTML: Develop the HTML structure for the home page, laying the foundation for the front-end design.

33. Home Page CSS: Implement the CSS styles for the home page, defining the visual presentation and layout of the elements.

34. Home Page Flask: Integrate Flask with the home page, establishing dynamic content and ensuring a responsive user interface.

35. Search Page HTML: Develop the HTML structure for the search page, providing the framework for displaying search-related content.

36. Search Page CSS: Implement the CSS styles for the search page, defining the visual presentation and layout to enhance the user experience.

Sprint Burnup Report

Date - September 25th, 2023 - October 2nd, 2023
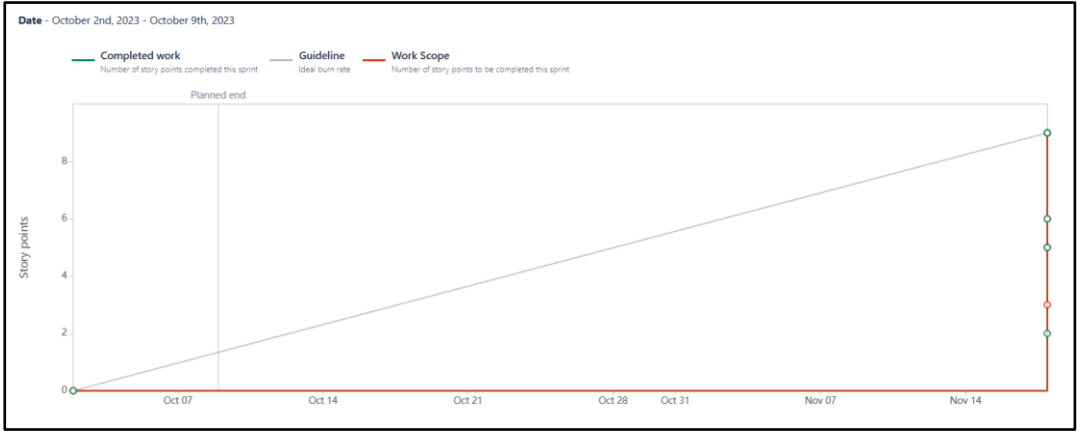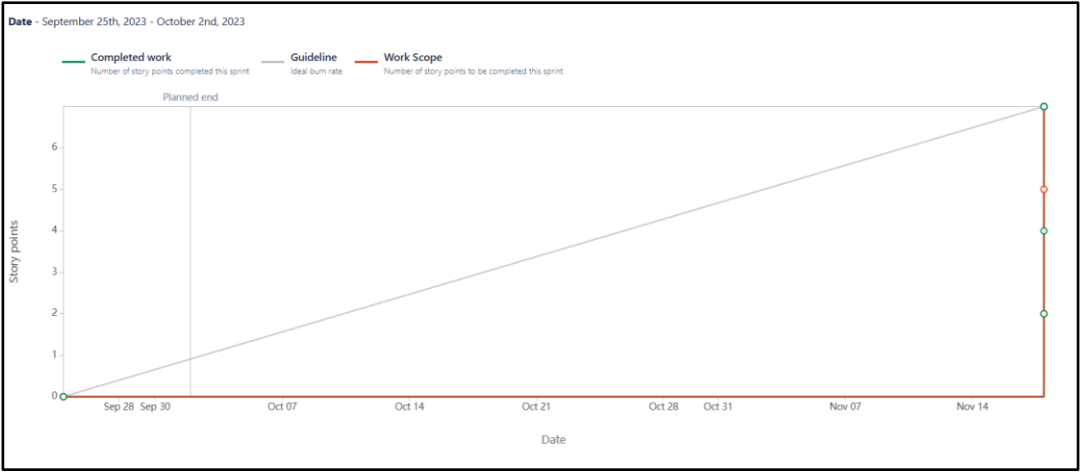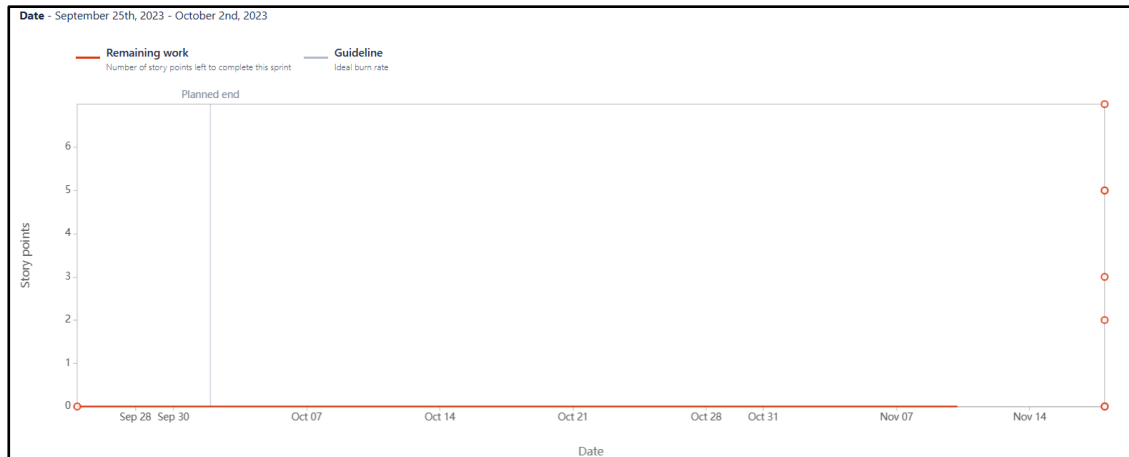
	SDPSEM-36	flask integration with backend	Vishal Prakash	Vishnu Siddharth	=	DONE
	SDPSEM-35	modularization of different functions	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-34	data structure analysis for handling the reviews	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-33	integrating naive bayes with amazon scraped reviews	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-32	prediction analysis	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-31	model building	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-30	naive bayes analysis	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-29	beautiful soup analysis	Vishnu Siddharth	Vishnu Siddharth	=	DONE
	SDPSEM-28	score converter from naive bayes	Vishnu Siddharth	Vishnu Siddharth	=	DONE

▣	SDPSEM-27	Naive Bayes Algorithm	VS Vishnu Siddharth	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-26	search page flask	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-25	search page css	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-24	search page html	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-23	home page flask	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-22	home page css	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-21	home page html	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-20	front end design	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-19	front end ideation	VP Vishal Prakash	VS Vishnu Siddharth	=	DONE ⌄	Doı
▣	SDPSEM-15	scraping from amazon	VS Vishnu Siddharth	VS Vishnu Siddharth	=	DONE ⌄	Do
▣	SDPSEM-14	Analyser module blueprint	VS Vishnu Siddharth	VS Vishnu Siddharth	=	DONE ⌄	Do
▣	SDPSEM-13	Model Analysis	VS Vishnu Siddharth	VS Vishnu Siddharth	=	DONE ⌄	Do
☑	SDPSEM-12	Find what all libraries to be imported	Unassigned	VP Vishal Prakash	=	DONE ⌄	Do
☑	SDPSEM-11	IEEE Reference Paper	Unassigned	VP Vishal Prakash	=	DONE ⌄	Do
☑	SDPSEM-10	Block Diagram	Unassigned	VP Vishal Prakash	=	DONE ⌄	Do
☑	SDPSEM-9	Gather Information about CoCoMo Model	VP Vishal Prakash	VP Vishal Prakash	=	DONE ⌄	Do
☑	SDPSEM-8	Formulate the Effort equation	VP Vishal Prakash	VP Vishal Prakash	=	DONE ⌄	Do
☑	SDPSEM-7	Learn about CoCoMo Model and apply	Unassigned	VP Vishal Prakash	=	DONE ⌄	Do

XI. Project Outcomes

Development of a Robust Sentiment Analysis Model:
The project has successfully developed a powerful sentiment analysis model specifically tailored for e-commerce product reviews. This model demonstrates high accuracy and reliability in discerning sentiments, providing a valuable tool for businesses to understand customer feedback.

Fine-grained Sentiment Classification:
The implemented model excels in fine-grained sentiment classification, distinguishing not only between positive and negative sentiments but also capturing nuances within these categories. This granularity enhances the depth of insights into customer opinions, allowing businesses to identify subtle shifts in sentiment.

Multimodal Sentiment Analysis Capability:

A notable achievement of the project is the integration of multimodal sentiment analysis. Beyond textual data, the model can process and analyze other modalities such as images or videos associated with product reviews. This comprehensive approach provides a more holistic understanding of customer sentiments.

Real-time Sentiment Monitoring System:

The project delivers a real-time sentiment monitoring system that allows e-commerce businesses to track and respond to customer sentiments as they evolve. This dynamic monitoring capability is crucial for timely decision-making and adapting marketing strategies based on emerging trends.

Adaptability to Diverse Product Categories:

The developed sentiment analysis model is designed to be adaptable to a wide range of product categories. Whether dealing with electronics, fashion, or household items, the model can effectively handle the specific language and sentiment nuances associated with different types of products.

User-friendly Integration with E-Commerce Platforms:

The outcomes include a user-friendly integration process, enabling seamless incorporation of the sentiment analysis model into existing e-commerce platforms. This ease of integration ensures that businesses can quickly leverage the benefits of sentiment analysis without significant technical hurdles.

Insights for Product Improvement and Marketing Strategies:
The project contributes actionable insights for product improvement and marketing strategies. By understanding the sentiments expressed by customers, businesses can identify areas for product enhancement, tailor marketing campaigns to address specific concerns, and ultimately enhance the overall customer experience.

Scalable Architecture for Growing E-Commerce Operations:
The architecture of the sentiment analysis model is designed to be scalable, accommodating the growing data volumes of expanding e-commerce operations. This scalability ensures that the model's performance remains robust even as the volume of customer reviews increases.

Contributions to Research and Industry Knowledge:
Through the project, valuable contributions are made to both research and industry knowledge in the domain of sentiment analysis for e-commerce. The integration of advanced natural language processing techniques and the focus on e-commerce-specific challenges enrich the broader understanding of sentiment analysis applications.

Enhanced Customer Relationship Management:
The project outcomes enhance customer relationship management by providing e-commerce businesses with tools to proactively address customer concerns and preferences. Improved understanding of customer sentiments fosters stronger relationships and loyalty.

Screenshots:
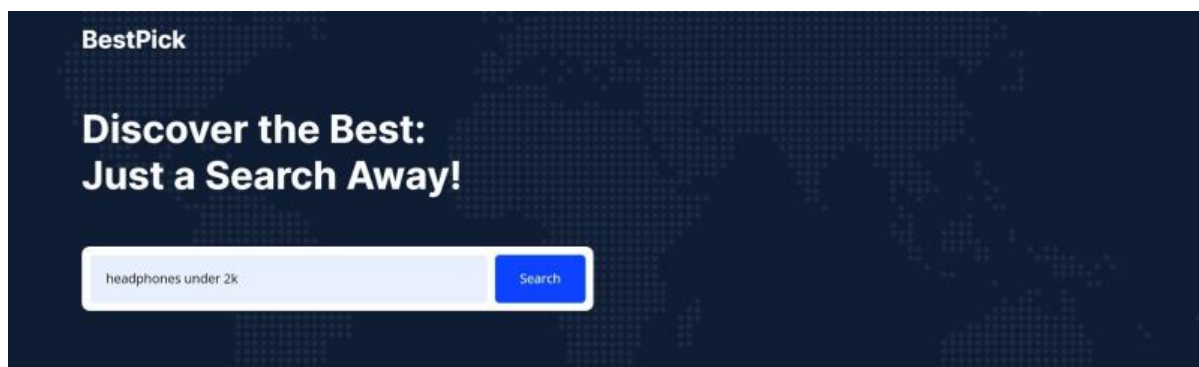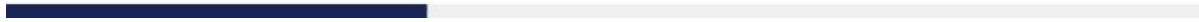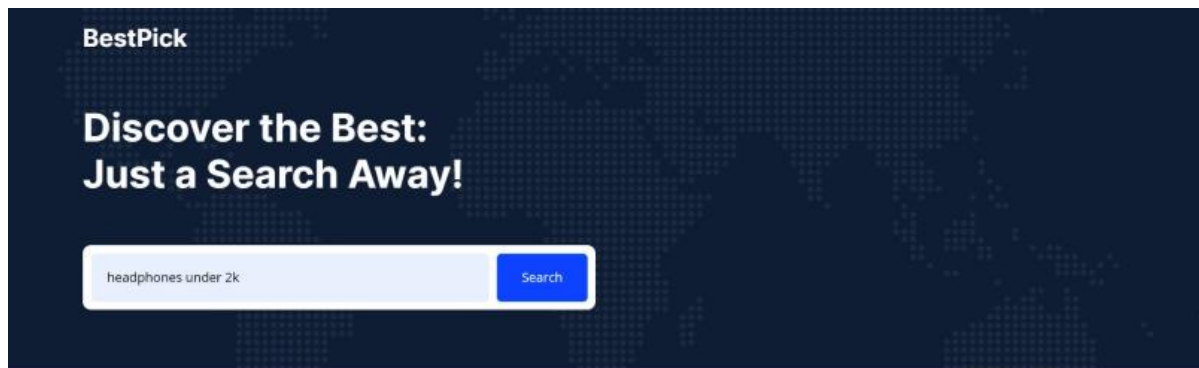
Search Results

pTron Newly Launched Zenbuds Evo TWS Earbuds,Ai-Trutalk Enc Calls,45Ms Movie/Music Modes,Deep Bass,32Hrs Playtime,Bluetooth 5.3 Headphones,Ipx5 Water Resistant&Type-C Fast Charging (Black),in-Ear

Go

Infinity by Harman Tranz 700 On Ear Wireless Headphone with Mic, 20 Hrs Playtime with Quick Charge, Deep Bass, Dual Equalizer, Bluetooth 5.0 and Voice Assistant Support (Black)

Go

boAt Rockerz 551ANC Hybrid Active Noise Cancellation Headphones with Up to 100H Playtime, ASAP™ Charge, Ambient Sound Mode &Dual EQ Modes, ENx™ Technology(Stellar Black)

Go

Results

Code Snippets

index.html

```html
1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4    <meta charset="utf-8">
5    <meta content="width=device-width, initial-scale=1.0" name="viewport">
6    <title>BestPick: A review analysis software</title>
7    <meta content="" name="description">
8    <meta content="" name="keywords">
9
10   <!-- Favicons -->
11   <link href="static/favicon.png" rel="icon">
12   <link href="static/apple-touch-icon.png" rel="apple-touch-icon">
13
14   <!-- Google Fonts -->
15   <link rel="preconnect" href="https://fonts.googleapis.com">
16   <link rel="preconnect" href="https://fonts.gstatic.com" crossorigin>
17   <link href="https://fonts.googleapis.com/css2?family=Open+Sans:ital,wght@0,300;0,400;0,500;0,600;0,700;1,300;1,400;1,600;1,700&f
18
19   <!-- Vendor CSS Files -->
20   <link href="static/vendor/bootstrap/css/bootstrap.min.css" rel="stylesheet">
21   <link href="static/vendor/bootstrap-icons/bootstrap-icons.css" rel="stylesheet">
22   <link href="static/vendor/fontawesome-free/css/all.min.css" rel="stylesheet">
23   <link href="static/vendor/glightbox/css/glightbox.min.css" rel="stylesheet">
24   <link href="static/vendor/swiper/swiper-bundle.min.css" rel="stylesheet">
25   <link href="static/vendor/aos/aos.css" rel="stylesheet">
26
27   <!-- Template Main CSS File -->
28   <link href="static/css/main.css" rel="stylesheet">
29
30
31   <style>
32     .loading-bar-container {
33       display: none;
34       position: fixed;
35       bottom: 0;
```

```html
51  </head>
52  <body>
53    <!-- ======= Header ======= -->
54    <header id="header" class="header d-flex align-items-center fixed-top">
55      <div class="container-fluid container-xl d-flex align-items-center justify-content-between">
56        <a href="index.html" class="logo d-flex align-items-center">
57          <H1> BestPick</H1>
58        </a>
59        <i class="mobile-nav-toggle mobile-nav-show bi bi-list"></i>
60        <i class="mobile-nav-toggle mobile-nav-hide d-none bi bi-x"></i>
61        <nav id="navbar" class="navbar">
62        </nav><!-- .navbar -->
63      </div>
64    </header><!-- End Header -->
65    <!-- End Header -->
66
67    <!-- ======= Hero Section ======= -->
68    <section id="hero" class="hero d-flex align-items-center">
69      <div class="container">
70        <div class="row gy-4 d-flex justify-content-between">
71          <div class="col-lg-6 order-2 order-lg-1 d-flex flex-column justify-content-center">
72            <h3 data-aos="fade-up">Discover the Best:         Just a Search Away!</h3>
73            <p data-aos="fade-up" data-aos-delay="100"></p>
74            <form action="/process_url" method="post" class="form-search d-flex align-items-stretch mb-3" data-aos="fade-up" data-ao
75              <input type="text" class="form-control" name="product_url" placeholder="Enter the keyword">
76              <button type="submit" class="btn btn-primary">Search</button>
77            </form>
78
79          </div>
80        </div>
81      </div>
82      <div class="hero-img" style="background-image: url('static/hero-img.jpg');"></div>
83    </section>
84    <!-- <div class="loading-bar-container">
85      <div class="progress-bar"></div>
```

search_results.html

```html
49    </style>
50  </head>
51  <body>
52    <!-- ======= Header ======= -->
53    <header id="header" class="header d-flex align-items-center fixed-top">
54      <div class="container-fluid container-xl d-flex align-items-center justify-content-between">
55        <a href="index.html" class="logo d-flex align-items-center">
56          <H1>BestPick</H1>
57        </a>
58        <i class="mobile-nav-toggle mobile-nav-show bi bi-list"></i>
59        <i class="mobile-nav-toggle mobile-nav-hide d-none bi bi-x"></i>
60        <nav id="navbar" class="navbar">
61        </nav><!-- .navbar -->
62      </div>
63    </header>
64    <!-- End Header -->
65    <!-- End Header -->
66
67    <!-- ======= Hero Section ======= -->
68    <section id="heroo" class="heroo d-flex align-items-center" style="background: url('static/hero-bg.png') center/cover no-repeat;
69      <div class="container">
70        <div class="row gy-4 d-flex justify-content-between">
71          <div class="col-lg-6 order-2 order-lg-1 d-flex flex-column justify-content-center">
72            <br>
73            <h3 data-aos="fade-up">Discover the Best:                       &
74            <p data-aos="fade-up" data-aos-delay="100"></p>
75            <form action="/process_url" method="post" class="form-search d-flex align-items-stretch mb-3" data-aos="fade-up" data-ao
76              <input type="text" class="form-control" name="product_url" placeholder="Enter the keyword">
77              <button type="submit" class="btn btn-primary">Search</button>
78            </form>
79          </div>
80        </div>
81      </div>
82    </section>
83
84    <!-- ======= Search Results Section ======= -->
85    <style>
86      /* Custom CSS for Search Results */
87      .search-results h2 {
88        font-family: 'Open Sans', sans-serif; /* Change to your preferred font */
89      }
90
91      .result-box {
92        padding: 20px;
93        text-align: center;
94        border: 1px solid #ccc;
95        margin: 10px;
96      }
97
98      .result-box img {
99        max-width: 100%;
100       height: auto;
101     }
102
103     .result-title {
104       font-size: 20px; /* Adjust the font size as needed */
105       margin-top: 10px; /* Adjust the top margin for spacing */
106     }
107
108     .go-button {
109       background-color: var(--color-secondary);
110       color: #fff;
111       text-decoration: none;
112       padding: 5px 20px;
113       border-radius: 5px;
114     }
115   </style>
116   <div class="container">
117     <h2 class="result-title">Search Results</h2>
```

```
115    </style>
116    <div class="container">
117      <h2 class="result-title">Search Results</h2>
118      <div class="row">
119        <div class="col-md-4">
120          <div class="result-box">
121            {% if best %}
122            <img src="{{ best[3] }}" alt="Best Result Image">
123            <p class="result-title">{{ best[1] }}</p>
124            <br>
125            <a href="{{ best[2] }}" class="go-button">Go</a>
126            {% endif %}
127          </div>
128        </div>
129        <div class="col-md-4">
130          <div class="result-box">
131            {% if second %}
132            <img src="{{ second[3] }}" alt="Second Result Image">
133            <p class="result-title">{{ second[1] }}</p>
134            <br>
135            <a href="{{ second[2] }}" class="go-button">Go</a>
136            {% endif %}
137          </div>
138        </div>
139        <div class="col-md-4">
140          <div class="result-box">
141            {% if third %}
142            <img src="{{ third[3] }}" alt="Third Result Image">
143            <p class="result-title">{{ third[1] }}</p>
144            <br>
145            <a href="{{ third[2] }}" class="go-button">Go</a>
146            {% endif %}
147          </div>
148        </div>
149      </div>
```

puremain.py (main backend)

```
55   #     return render_template('search_results.html', best=best, second=second, third=third)
56   def get_product_and_image_urls(asin):
57       contents = {
58       "user-agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/119.0.0.0 Safari/537.36
59       "accept-language": "en-IN,en;q=0.9"
60   }
61
62
63       # Construct the URL for the product page using the ASIN
64       product_url = f"https://www.amazon.in/dp/{asin}"
65
66       # Send a request to the product page
67       response = requests.get(product_url, headers=contents)
68       soup = BeautifulSoup(response.text, 'lxml')
69
70       # Extract the product image URL from the page (you may need to adjust the selector)
71       image_element = soup.select_one('#landingImage')  # Adjust the selector based on the HTML structure of the page
72
73       # If the image element is found, retrieve the image URL
74       image_url = image_element.get('src') if image_element else None
75
76       # Return both the product and image URLs
77       print(product_url,image_url)
78       return product_url, image_url
79
```

37

```python
@app.route('/process_url', methods=['POST'])
def search_endpoint():
    query = request.form.get('product_url')
    #Without threads: 48 seconds
    answers = scrapingandco.get_product_info(query)
    #With threading : risk of getting blocked by amazon
    #answers = scrapingandco.get_product_info_threads(query)
    best = answers[0] if answers else None
    second = answers[1] if len(answers) > 1 else None
    third = answers[2] if len(answers) > 2 else None
    print('ANSWERS')
    print(answers)
    print(best, second, third)
    #best[1] = title
    #best[2] = url of product
    #best[3] = result image

    print(best[0], best[1], best[2], best[3])

    return render_template('search_results.html', best=best, second=second, third=third)

@app.route('/get_image_url_for_product', methods=['POST'])
def get_image_url_for_product(asin):
    # Assuming the ASIN is valid
    product_url = f"https://www.amazon.in/dp/{asin}"
    return product_url

if __name__ == '__main__':
    app.run(debug=True)
```

scrapingandco.py (for scraping reviews)

```python
def search(searchterm):
    print("Searching up stuff")
    words = searchterm.split()
    query = "+".join(words)
    request = f"https://www.amazon.in/s?k={query}"
    contents = {
        "user-agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/119.0.0.0 Safari/53
        "accept-language": "en-IN,en;q=0.9",
    }
    response = requests.get(request, headers=contents)
    soup = BeautifulSoup(response.text, "lxml")
    asin_elements = soup.find_all("div", {"data-asin": True})
    #asin_list = [i["data-asin"] for i in asin_elements if i["data-asin"] != ""]
    final_list = []
    for i in asin_elements:
        if i["data-asin"] != "":
            with open("testing.txt","w",encoding='utf-8') as file:
                file.write(str(i.contents))
            image = i.find("img")
            asin = i["data-asin"]
            url = image["src"]
            final_list.append([asin,url])
    asin_list = final_list
    with open("asindata.txt","w") as file:
        for i in asin_list:
            file.write(str(i)+"\n")
    return asin_list[:10]
```

```python
def extract_from_asin(asin):
    image = asin[1]
    asin = asin[0]
    print("extracting from asin")
    contents = {
        "user-agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/119.0.0.0 Safari/53
        "accept-language": "en-IN,en;q=0.9",
    }


    #time.sleep(random.uniform(1.0, 2.0))
    url = f"https://www.amazon.in/dp/{asin}"

    response = requests.get(url, headers=contents)
    soup = BeautifulSoup(response.text, "lxml")
    title_element = soup.select_one("#productTitle")
    # print(title_element)
    review_element = soup.select("div.review")
    scrapedreviews = []
    for review in review_element:
        content = review.select_one("span.review-text")
        scrapedreviews.append(content.text)
    final = []
    for review in scrapedreviews:
        rev = "".join(c for c in review if c not in emoji.EMOJI_DATA)
        final.append(rev)
        print(rev)
    print("Writing into file")
    with open("scrapereviews.txt","a",encoding="utf-8") as file1:
        d = {'title': title_element.contents[0] if title_element else None,
        'reviews': final}
        print(d)
        file1.writelines(str(d))


def process_query(query):
    print("Query processing")
    asins = search(query)
    #time.sleep(random.uniform(0.2, 1.0))
    master_reviews = {}
    for asin in asins:
        name, reviews = extract_from_asin(asin)
        master_reviews[(asin[0], name, asin[1])] = reviews
    return master_reviews

def thread_work(asin,master_reviews):
    print("Searching Asin")
    #time.sleep(random.uniform(0.2, 1.0))
    name, reviews = extract_from_asin(asin)
    master_reviews[(asin[0], name, asin[1])] = reviews


def process_query_thread(query):
    print("Query processing")
    asins = search(query)
    #time.sleep(random.uniform(0.2, 1.0))
    master_reviews = {}
    threads = []
    for asin in asins:
        threads.append(threading.Thread(target=thread_work,args=[asin,master_reviews]))
    for i in threads:
        i.start()
    for i in threads:
        i.join()
    return master_reviews
```

```python
129    def thread_work2(product,answers,reviews):
130        asin = product[0]
131        title = product[1]
132        image = product[2]
133        print("Analyzing",product)
134        #score = analyzer.save_training_model(reviews, r"dataset2.csv")
135        score = analyzer.train_with_data(reviews, "dataset2.csv")
136        url = f"https://www.amazon.in/dp/{asin}"
137        answers.append([score, title, url, image])
138    def get_product_info_threads(query):
139        d = process_query(query)
140        answers = []
141        threads = []
142        for product in d:
143            threads.append(threading.Thread(target=thread_work2,args=[product,answers,d[product]]))
144        for i in threads:
145            i.start()
146        for i in threads:
147            i.join()
148        answers.sort(reverse = True)
149        return answers
150    if "__name__" == "__main__":
151        query = input("Query: ")
152        d = process_query(query)
153        answers = []
154        for product in d:
155            asin = product[0]
156            title = product[1]
157            image = product[2]
158            reviews = d[product]
159            print("Analyzing",product)
160            #score = analyzer.save_training_model(reviews, r"dataset2.csv")
161            score = analyzer.train_with_data(reviews, "dataset2.csv")
162            answers.append([score, title, image])
163        print(answers)
```

train.tsv

```
1    PhraseId    SentenceId    Phrase    Sentiment
2    1    1    A series of escapades demonstrating the adage that what is good for the goose is also good for the gander , some of whi
3    2    1    A series of escapades demonstrating the adage that what is good for the goose    2
4    3    1    A series    2
5    4    1    A    2
6    5    1    series    2
7    6    1    of escapades demonstrating the adage that what is good for the goose    2
8    7    1    of    2
9    8    1    escapades demonstrating the adage that what is good for the goose    2
10   9    1    escapades    2
11   10   1    demonstrating the adage that what is good for the goose 2
12   11   1    demonstrating the adage 2
13   12   1    demonstrating    2
14   13   1    the adage    2
15   14   1    the 2
16   15   1    adage    2
17   16   1    that what is good for the goose 2
18   17   1    that    2
19   18   1    what is good for the goose    2
20   19   1    what    2
21   20   1    is good for the goose    2
22   21   1    is    2
23   22   1    good for the goose    3
24   23   1    good    3
25   24   1    for the goose    2
26   25   1    for 2
27   26   1    the goose    2
28   27   1    goose    2
29   28   1    is also good for the gander , some of which occasionally amuses but none of which amounts to much of a story .    2
30   29   1    is also good for the gander , some of which occasionally amuses but none of which amounts to much of a story    2
31   30   1    is also 2
32   31   1    also    2
33   32   1    good for the gander , some of which occasionally amuses but none of which amounts to much of a story    2
34   33   1    for the gander , some of which occasionally amuses but none of which amounts to much of a story 2
35   34   1    the gander , some of which occasionally amuses but none of which amounts to much of a story 1
```

scrapedreviews.txt

1 {'title': '   Boult Audio FXCharge Wireless in Ear Bluetooth Neckband with ENC Mic, 32H Playtime, Type-C Fast Charging
(5Mins=7.5Hrs Playtime), Dual Pairing, Made in India, Biggest 14.2mm Drivers Ear Phones (Black)   ', 'reviews': ["\nBeing an
Occupational Therapist, Wellness & Events Consultant and a DJ, I value quality sound in comfort.I need comfort, pairing, sound
quality & playback duration in the same order.In the matter of using Head/Earphones, Comfort is utmost.The Fit of these
earphones within the ear canal is OK but soon turns painful. The outer edges of Magnetic Rear Side starts to feel heavy and
causes pain along the ridges of the Concha Cavum & Cymba. Couldn't keep on for more than 30 minutes max.Forget dual pairing.
Completely Failed the purpose I bought it for. 3 Stars are for the Audio Quality & Average Noise Cancellation.My Overall Rating
is reduced because of Discomfort and Improper Pairing functionality.My Ear Size Top of Ear to Bottom of Lobe - 7.2cm / Earphone
Cavity - 3cm (For reference check Image)People with larger ear cavity size would find it a better fit and enjoy good sound with
bluetooth connectivity.====== FEATURE'S SUMMARY ======COMFORT :- Irrespective of silicon inner ear plugs I tried, I was Very
Unhappy with the discomfort & pain caused by the outer rear edge (the circumference of the magnetic side) along the periphey of
ear notch near lobes where the buds are designed to rest.HEAVY:- The Earbuds felt weighted / heavy in ear, constantly kept
adjusting position for comfort.DUAL PAIRING :- I was not simultaneously connected to two devices. It connected only to one
device until disconnected or switch off, only then it would detect the other device and connect. So, Dual pairing feature
inefficient.By Definition, I am to understand Dual pairing allows 2 Bluetooth devices to be connected to One Pair of head/
earphones at the same time to answer calls and listen to music without having to fumble with device controls to pair, connect,
disconnect and reconnect to & fro. This one does Pair to two devices, but not simultaneously.TAP FUNCTION :- had no effect and
frankly irritating. The inline control worked to receive & end calls. Did not bother much one ears started hurting.ENC :- Not
highy effective. Noise Cancellation was enough for clear call but the environmental noise / sounds don't cut off completely.
PAIRING :- Bluetooth 5.2, Single device, fastRANGE INDOORS :- Good Through walls. My House carpet area approx. 1700sq ft. Placed
device in central area, clear connectivity within the house.SOUND :- Clear, Crisp with Good Balanced Bass (Not Very Punchy until
custom EQ) Not the Best but Much better than Average. I expected more as it's publicized with biggest drivers in ear phones.
PLAYBACK TIME :- Came with a 80% Charge and according to usage of 1-2 Hrs came down to 70%----- SAFETY TIP: Please Avoid Noise
Cancelling while Driving / Riding / Walking on Road ---------- SAFETY TIP: Avoid using Noise Cancelling Earphones/Earpods/
Headphones in Public Areas ---------- SAFETY TIP: Be Aware of your surroundings when in public places -----My main criteria:---
Bluetooth Handsfree Neckband --- Long duration comfort --- Dual pairing with two bluetooth devices --- Sound quality ---
min. 10 Hrs Music Playback time \nRead more", '\ngoood\nRead more", '\nGood sound quality with high bass, but sometimes loose
connectivity automatically... Need to do improve Bluetooth connectivity issues\nRead more', '\nWorth to buy at the price of
somewhere under ₹100.\nRead more', "\nEar Peace (Speaker Buds) are very big in size not comfortable to wear in ear and also not
suitable for bike Riders while wearing helmet it's not comfortable to wear. And Sound Quality is good but Valume is very low.
And it's not returnable if you don't like product only option to return is if product is defect. So don't buy this item if
you're using while bike riding with helmet. Please don't Buy this Boult companies this model headphones other models I don't
know.\nRead more", '\nGood \nRead more', '\nSound quality is ok ok\nRead more', '\n\n\n\n\n   The media could
not be loaded.\n   \n\n\n\n\xa0As you think 14.2 mm drivers are very best it is not there very cheap then getting a
earphone in market and I find it Rupees as the company Bolt audio I buy the earphone but there nothing in the earphones very

XII. Conclusion and Future Directions

The "Best Pick" project has successfully addressed the challenge of enhancing the e-commerce experience through sentiment analysis, natural language processing (NLP), and machine learning algorithms. The developed web-based application provides valuable insights into user sentiments expressed in product reviews on Amazon. By leveraging NLP techniques, sentiment analysis, and machine learning algorithms, the application effectively categorizes reviews, identifies sentiments, and generates personalized product recommendations.

Key Achievements:

Sentiment Analysis Implementation:

The project achieved the successful implementation of sentiment analysis techniques, allowing for the extraction of sentiments from unstructured textual data. Natural Language Processing libraries, including NLTK and spaCy, were instrumental in this process.

Machine Learning for Product Recommendations:
Leveraging machine learning algorithms, the application generated accurate and personalized product recommendations. The incorporation of collaborative filtering and content-based filtering mechanisms enhanced the precision of recommendations.

User-Friendly Interface:
The user interface design prioritized simplicity and intuitiveness, ensuring an optimal user experience. Users can easily navigate the application, input preferences, and receive tailored product suggestions.

Data Scraping and Retrieval:

The project effectively addressed the challenges associated with web scraping, retrieving relevant data from Amazon. Ethical considerations, legal implications, and technical challenges were carefully managed to ensure responsible data extraction.

Risk Management:

Throughout the project, careful attention was paid to ethical considerations, legal compliance, and potential challenges associated with web scraping. Proactive risk management strategies were implemented to mitigate any unforeseen issues.

Future Scope:

While the "Best Pick" project has achieved significant milestones, there are several avenues for future improvement and expansion:

Multilingual Support:

Enhancing the application to support multiple languages can broaden its user base and cater to a more diverse audience.

Real-Time Sentiment Analysis:

Introducing real-time sentiment analysis capabilities can provide users with immediate insights into changing trends and sentiments.

Integration with E-commerce Platforms:

Exploring opportunities for integration with various e-commerce platforms beyond Amazon can extend the application's utility and relevance.

Enhanced Recommendation Algorithms:

Continuously refining and expanding the machine learning algorithms for more accurate and personalized product recommendations.

User Feedback Mechanism:
Implementing a feedback mechanism to gather user insights and preferences, facilitating ongoing improvements based on user experiences.

Mobile Application Development:
Creating a mobile version of the application to increase accessibility and cater to users who prefer shopping on mobile devices.

Advanced Analytics Dashboard:
Developing an analytics dashboard for users to track and visualize trends, providing a comprehensive overview of product sentiments and recommendations.

In conclusion, the "Best Pick" project has successfully contributed to the field of e-commerce by harnessing the power of sentiment analysis and machine learning. The future scope encompasses enhancements in language support, real-time analysis, expanded platform integration, and continuous refinement of recommendation algorithms. This project serves as a testament to the potential of data-driven decision-making in improving the online shopping experience

Challenges Faced

- **Data Handling**: Managing and processing large datasets with varying vocal characteristics posed significant challenges.
- **Algorithm Optimization**: Fine-tuning the heuristic model and optimizing it using a bayseian algorithm required extensive testing and validation.
- **Integration of Technologies**: Seamlessly integrating Flask, Parselmouth, and other technologies to ensure smooth functioning was complex.
- **User Interface Design**: Creating an intuitive and responsive user interface that caters to a diverse user base was challenging.

What Went Right (As Per Plan)

- **Effective Sentiment Analysis**: Successfully performed sentiment analysis of the users' reviews using the Naive Bayes model.
- **Model Accuracy:** Achieved a reliable level of accuracy in predicting whether a product is good or not.
- **Score Generation:** Successfully converted the Naive Bayes probability into a normalized scorable number which is used to rank different products and different reviews to get average score per product.
- **User-Friendly Interface**: Developed a simple and easy-to-use web interface for uploading voice recordings and viewing results.

What Went Wrong (Unplanned)

- **Data Quality Issues**: Encountered inconsistencies in the datasets which impacted the initial model accuracy.
- **Technical Hurdles**: Faced unexpected technical challenges in integrating different software components.
- **Performance Optimization**: Struggled with optimizing the application's performance, particularly in real-time data processing.

Lessons Learnt
- **Importance of Data Quality**: Learned that high-quality, relevant data is crucial for the success of any AI/ML project.
- **Flexibility in Problem-Solving**: Gained insights on the need for adaptability and innovative thinking when faced with technical roadblocks.
- **User-Centric Design**: Understood the significance of designing with the end-user in mind, ensuring accessibility and ease of use.

In conclusion, the **Best Pick** project has successfully contributed to the field of e-commerce by harnessing the power of sentiment analysis and machine learning. The future scope encompasses enhancements in language support, real-time analysis, expanded platform integration, and continuous refinement of recommendation algorithms. This project serves as a testament to the potential of data-driven decision-making in improving the online shopping experience .

This is our contribution towards e-commerce

XIII. References

[1] Anju A, Barath M, Maheshwaran R, Vignesh K, Kishor Kumar SR, "Sentiment Analysis for E-commerce Website", 2022 10th IEEE International Conference on Emerging Trends in Engineering & Technology-Signal and Information Processing

[2] K Sudheer, Valarmathi B, " REAL TIME SENTIMENT ANALYSIS OF ECOMMERCE WEBSITES USING MACHINE LEARNING ALGORITHMS ", International Journal of Mechanical Engineering and Technology (IJMET) Volume 9, Issue 2, February 2018, pp. 180–193, Article ID: IJMET_09_02_018

[3] Kamp"s, J., Marx, M.J, Mokken, R.J. and Rijke, M.,"Using word net to measure semantic orientations of adjectives,"LREC 2004 May 26 (Vol. 4, pp. 1115-1118). 2004

[4] Kim, S.M. and Hovy, E," Determining the sentiment of opinions," In Proceedings of the 20th international conference on Computational Linguistics (p. 1367). Association for Computational Linguistics, August 2004 Prof. K. Sudheer and Dr. B Valarmathi http://iaeme.com/Home/journal/IJMET 192 editor@iaeme.com

[5] Hatzivassiloglou, V.and McKeown, K.R.," Predicting the semantic orientation of adjectives," In Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics (pp. 174-181). Association for Computational Linguistics, July 1997.

[6] Ganapathibhotla, M. and Liu, B., "Mining opinions in comparative sentences," Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1 (pp. 241-248). Association for Computational Linguistics, August 2008.

[7] Jindal, N. and Liu, B," Mining comparative sentences and relations," In AAAI (Vol. 22, pp..1331-1336), July 2006.

[8] Davidov, D., Tsur, O. and Rappoport, A." Semi-supervised recognition of sarcastic sentences in twitter and amazon," Proceedings of the fourteenth conference on computational natural language learning (pp. 107-116). Association for Computational Linguistics, July 2010

[9] Turney, P.D., "Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews," Proceedings of the 40th annual meeting on association for computational linguistics (pp. 417-424). Association for Computational Linguistics, July 2002

[10] Kwon, N. and Hovy, E," Integrating semantic frames from multiple sources," International Conference on Intelligent Text Processing and Computational Linguistics (pp. 1-12). Springer Berlin Heidelberg,Feb 2006

[11] Pang, B. and Lee, L, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," Proceedings of the 43rd annual meeting on association for computational linguistics (pp. 115-124), Association for Computational Linguistics, June 2005.

[12] Anastasia, Sonia, and Indra Budi, "Twitter sentiment analysis of online transportation service providers," Advanced Computer Science and Information Systems (ICACSIS), 2016 International Conference on. IEEE, 2016.

[13] Ramteke, Jyoti, et al. "Election result prediction using Twitter sentiment analysis," Inventive Computation Technologies (ICICT), International Conference on. Vol. 1. IEEE, 2016.

[14] Yang, Nan, Sanxing Cao, and Shufan Zhang,"Data analysis system for online short video comments," Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on. IEEE, 2016.

[15] Agarwal, Apoorv, et al,"Opinion mining of news headlines using SentiWordNet," Colossal Data Analysis and Networking (CDAN), Symposium on. IEEE, 2016.

[16] Tai, Chih-Hua, Zheng-Han Tan, and Yue-Shan Chang," Systematical approach for detecting the intention and intensity of feelings on social network," IEEE journal of biomedical and health informatics 20.4 (2016): 987-995.

[17] Chen, Huan, et al,"Cloud-Based Core Text Processing Services for Sentiment Analysis," Big Data (BigData Congress), 2016 IEEE International Congress on. IEEE, 2016.

[18] Ashok, Meghana, et al, "A personalized recommender system using Machine Learning based Sentiment Analysis over social data," Electrical, Electronics and Computer Science (SCEECS), 2016 IEEE Students' Conference on. IEEE, 2016.

[19] Seki, Yohei, "Use of Twitter for Analysis of Public Sentiment for Improvement of Local Government Service," Smart Computing (SMARTCOMP),2016 IEEE International Conference on. IEEE, 2016.

Client Evaluation Report

Name of the project: Best Pick

Team Members: Vijaymurugan N , Vishal Prakash , Vishnu Siddharth V R

Client details:

Rating System - 1: Strongly disagree 2: Disagree 3: Neutral 4: Agree 5: Strongly Agree

Questions	1	2	3	4	5
The problem was well discussed and the requirements and goals were clear.					
The project plan was well defined and communicated from the start.					
The resources were adequate for achieving the goals.					
The original timeline was realistic and was followed.					
The teamwork was well demonstrated.					
The client was communicated on regular intervals and given updates on the progress of the project.					
The expected project requirements have been satisfied.					