A Project report on

Analyzing customer sentiments on E-Tail applications using NLP

A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

Bachelor of Technology

in

Computer Science and Engineering

(Artificial intelligence and Machine Learning)

Submitted by

TEEGALA JUSTIN ROY
(20H51A6647)

PALLE SAI TEJA
(20H51A6687)

VENKAIAHGARI UDAY KIRAN REDDY
(20H51A6693)

BODIGE HARSHITH
(20H51A66A1)

Under the esteemed guidance of

Dr.Y.Ambica
(Associate Professor)

Department of Computer Science and Engineering

CMR COLLEGE OF ENGINEERING& TECHNOLOGY (AI &ML)
(UGC Autonomous)
*Approved by AICTE *Affiliated to JNTUH *NAAC Accredited with A^+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

2020- 2024

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (AI & ML)

CERTIFICATE

This is to certify that the Major Project report entitled **"Analyzing customer sentiments on E-Tail applications using NLP"** being submitted by TEEGALA JUSTIN ROY (20H51A6647), PALLE SAI TEJA (20H51A6687) , VENKAIAHGARI UDAY KIRAN REDDY (20H51A6693), BODIGE HARSHITH (20H51A66A1) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering(AI & ML)** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodied in this project report have not been submitted to any other University or Institute for the award of any Degree.

Dr. Y. Ambica	**Dr. P. Sruthi**	**External**
Associate Professor	**Associate Professor and HOD**	
Dept.of CSE(AI&ML)	**Dept. of CSE (AI & ML)**	

ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Dr.Y.Ambica, Associate Professor**, Department of Computer Science and Engineering(AI&ML) for her valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr.P. Sruthi,** Head of the Department of Computer Science and Engineering (AI&ML), CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. G.Devadas** , Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering (AI&ML) for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

Teegala Justin Roy 20H51A6647
Palle Sai Teja 20H51A6687
Venkaiahgari Uday
Kiran Reddy 20H51A6693
Bodige Harshith 20H51A66A1

TABLE OF CONTENTS

List Of Figures

ABSTRACT

With the ever-increasing prevalence of online shopping, understanding customer sentiments is crucial for E-Tail businesses to enhance user experience and optimize their platforms. The study involves the collection of textual data, such as customer reviews and feedback, from various E-Tail applications. NLP algorithms will be employed to process and analyze this textual data, extracting valuable insights into customer sentiments. The project not only focuses on sentiment polarity (positive, negative, or neutral) but also delves deeper into the identification of specific aspects influencing customer opinions, such as product quality, delivery services, and customer support. The findings of this analysis will provide valuable information for E-Tail businesses to make data-driven decisions, improve customer satisfaction, and tailor their services to meet consumer expectations in the dynamic online marketplace.

The project methodology involves the pre-processing of textual data, feature extraction, and the training of machine learning models to accurately classify sentiments. Rigorous validation techniques will be employed to ensure the reliability of the sentiment analysis results. The ultimate goal is to develop a comprehensive understanding of customer sentiments, enabling E-Tail businesses to enhance their marketing strategies, refine product offerings, and address pain points in real-time. The insights gained from this project can potentially revolutionize how E-Tail businesses engage with their customers, fostering a customer-centric approach that goes beyond transactional interactions to build lasting relationships in the competitive online retail landscape.

CHAPTER 1
INTRODUCTION

CHAPTER 1
INTRODUCTION

With the proliferation of online transactions, understanding the sentiments of customers has emerged as a critical determinant of success for E-Tail businesses. This project, "Analyzing Customer Sentiments on E-Tail Applications using Natural Language Processing (NLP)," undertakes a multifaceted exploration of customer sentiments within the dynamic landscape of online retail. Through the lens of NLP, this initiative is structured into three interrelated projects, each tailored to unravel distinct facets of customer sentiments and preferences in the E-Tail ecosystem.

The initial project centers on the development and application of advanced sentiment analysis models, leveraging cutting-edge NLP techniques to categorize customer sentiments gleaned from reviews. By discerning sentiment polarity and identifying influential factors shaping customer opinions, this project provides businesses with granular insights to refine their strategies and enhance user experiences.

Building on this foundation, the second project addresses the intricacies of language by delving into the impact of contextual nuances such as sarcasm and irony on sentiment analysis. Recognizing the need for heightened linguistic sophistication, this project aims to augment sentiment models, ensuring a more nuanced interpretation of customer expressions and a more accurate representation of true sentiments.

The third and final project adopts a holistic approach by correlating specific product features or services with customer sentiments. Utilizing advanced topic modeling techniques, this project seeks to uncover latent patterns and thematic trends within customer feedback, empowering E-Tail businesses with actionable insights to optimize product offerings and promptly address customer concerns. Together, these projects constitute a comprehensive endeavor to decode and respond to customer sentiments, offering E-Tail businesses a strategic advantage in navigating the evolving landscape of online retail.

1.1. Problem Statement

In the rapidly expanding domain of electronic retail (E-Tail), where customer interactions and transactions are predominantly mediated through digital platforms, understanding and effectively responding to customer sentiments have become pivotal for business success. E-Tail applications serve as conduits for consumers to articulate their experiences, preferences, and critiques through reviews and feedback. However, the sheer volume and diversity of textual data generated present a significant challenge for businesses seeking to extract actionable insights. This project, "Analyzing Customer Sentiments on E-Tail Applications using Natural Language Processing (NLP)," aims to address this challenge by employing advanced NLP techniques to systematically analyze and interpret customer sentiments, providing E-Tail businesses with a nuanced understanding of customer experiences and preferences.

The primary objective of this project is to develop and implement web application which can perform sentiment analysis models capable of categorizing customer sentiments expressed in reviews on E-Tail applications. Leveraging state-of-the-art NLP algorithms, the project seeks to discern not only the overall polarity of sentiments (positive, negative, or neutral) but also the nuanced aspects contributing to these sentiments. By doing so, businesses can gain insights into specific factors such as product quality, delivery services, and customer support that significantly influence customer satisfaction or dissatisfaction.

Furthermore, the project aims to enhance the sophistication of sentiment analysis models to navigate the complexities of language, including sarcasm, irony, and other contextual nuances. This refinement is essential for ensuring the accuracy of sentiment interpretation, as customer expressions often embody subtleties that can impact the overall understanding of sentiments. Through these endeavors, the project aspires to equip E-Tail businesses with a robust framework for data-driven decision-making, enabling them to tailor their services, optimize user experiences, and stay responsive to the ever-evolving expectations of their online customer base.

1.2. Research Objective

1. **Identify Key Themes and Topics Impacting Sentiments:** Delve into the underlying themes and topics within customer sentiments to identify the key factors influencing their experiences on E-Tail platforms. Through topic modeling and text analysis, the project seeks to uncover prevalent themes such as product quality, customer service, shipping experiences, and pricing. Understanding these themes is crucial for businesses to pinpoint specific areas of improvement and optimize their operations accordingly.

2. **Address Contextual Challenges in Sentiment Analysis:** Enhance the sentiment analysis models to address contextual challenges inherent in customer reviews. This includes accounting for linguistic nuances, sarcasm, and other forms of expression that may pose challenges to traditional sentiment analysis models. The objective is to refine the models to accurately capture the true sentiments of customers, ensuring a more nuanced and context-aware interpretation of the feedback provided.

3. **Provide Actionable Insights for E-Tail Businesses:** Translate the analyzed sentiments into actionable insights for E-Tail businesses. The project aims to offer practical recommendations based on the sentiments expressed by customers, enabling businesses to make informed decisions. These insights may encompass areas for improvement, strategies for enhancing customer satisfaction, and opportunities for innovation in products or services. The overarching goal is to empower E-Tail businesses to respond proactively to customer sentiments and improve their overall market competitiveness.

4. **Evaluate the Impact of Customer Sentiments on Business Performance:** Investigate the correlation between customer sentiments and key performance indicators for E-Tail applications. By analyzing how sentiments relate to metrics such as conversion rates, customer retention, and overall sales, the project aims to quantify the impact of customer opinions on business success. This objective provides a data-driven foundation for businesses to understand the financial implications of customer sentiments and make strategic decisions accordingly.

CHAPTER 2
BACKGROUND WORK

CHAPTER 2

BACKGROUND WORK AND EXISTING SYSTEMS

2.1. Lexicon-Based Sentiment Analysis:

2.1.1. Introduction:

Lexicon-based sentiment analysis, a foundational approach in the field of natural language processing (NLP), relies on predefined lexicons or dictionaries to assess the sentiment expressed in textual data. In this method, words are assigned sentiment scores, and the overall sentiment of a piece of text is determined by aggregating these scores.

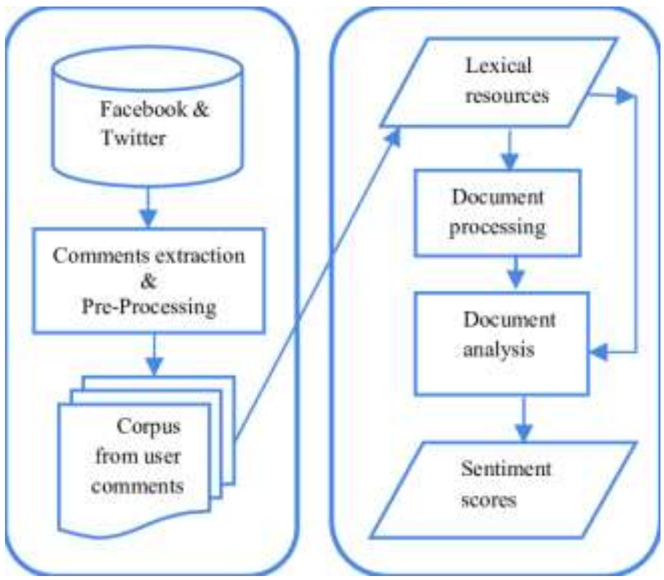


Figure 2.1.1.1 lexicon based sentiment analysis architecture

2.1.2. Implementation Of Lexicon Based sentiment analysis:

The implementation of lexicon-based sentiment analysis involves the creation of a sentiment lexicon, where words are pre-annotated with their associated sentiment scores. Positive, negative, and neutral sentiments are assigned to words, forming the foundation of the analysis. In the implementation process, the textual data to be analyzed is tokenized, breaking it into individual words or phrases. Each token is then compared against the entries in the sentiment lexicon, and sentiment scores are aggregated based on the matched words. The overall sentiment of the text is determined by summing or applying more sophisticated algorithms to the aggregated scores. This implementation is particularly effective in scenarios where understanding the sentiment of individual words contributes to a nuanced interpretation of the overall sentiment expressed in a larger piece of text, such as reviews, social media posts, or customer feedback.

2.1.3. Functions and Limitations:

Lexicon-based sentiment analysis functions by utilizing pre-defined sentiment lexicons to assign sentiment scores to words, enabling the overall sentiment of a given text to be inferred through the aggregation of these scores. This approach is straightforward to implement and computationally efficient, making it suitable for tasks where speed and simplicity are priorities. Lexicon-based methods are particularly effective in capturing explicit sentiment expressions in text and can be useful for domains with specialized vocabularies.

- However, their limitations include challenges in handling context-dependent sentiment, idiomatic expressions, and sarcasm.
- Lexicons may struggle with the evolving nature of language, as they might not encompass newly coined words or shifts in word meanings over time.
- Additionally, lexicon-based sentiment analysis may overlook subtle nuances and struggles to capture sentiments in complex sentences where words interact in intricate ways, limiting its efficacy in handling more sophisticated language constructs.

2.2 **Rule-Based Sentiment Analysis:**

2.2.1 Introduction:

Rule-based sentiment analysis represents a foundational approach in natural language processing (NLP) where sentiment classification is guided by a set of predefined rules. Rule-based systems employ explicit rules to interpret sentiment in text. These rules may range from straightforward keyword matching to more sophisticated grammatical and syntactical analyses. The strength of rule-based sentiment analysis lies in its interpretability and adaptability to specific contextual requirements, allowing analysts to tailor rule sets according to the intricacies of the domain.

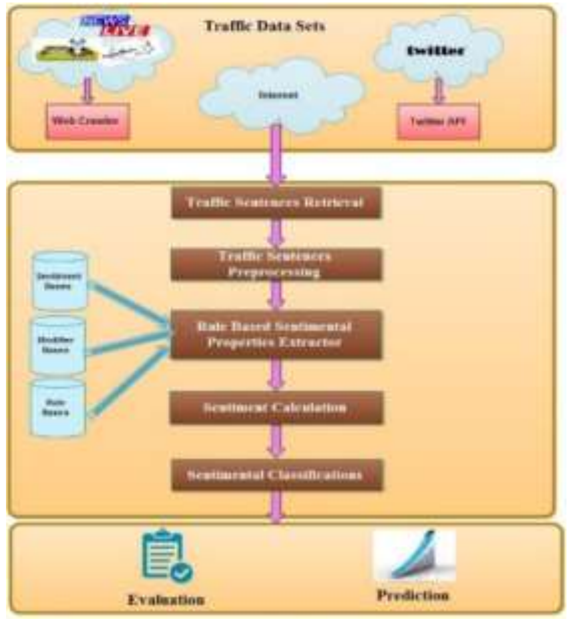


Figure 2.2.1.1: Architecture of Rule based sentiment analysis

2.2.2 Implementation:

The implementation of rule-based sentiment analysis involves creating a set of predefined rules that guide the interpretation of sentiments within textual data. These rules can range from basic keyword matching, where the presence of specific words indicates sentiment, to more sophisticated linguistic rules that consider context, negations, and grammatical structures. In the implementation process, the text is analyzed based on these rules, and sentiment labels are assigned accordingly. Rule-based systems offer transparency and interpretability, allowing for fine-tuning to suit the characteristics of specific domains. This approach is particularly effective when explicit guidelines can be defined for sentiment interpretation, making it adaptable to diverse contexts such as customer reviews, social media content, or financial news analysis. However, it is important to note that rule-based systems may face challenges in handling nuanced language expressions and can require ongoing refinement to address evolving language trends.

2.2.3 Functions and Limitations:

Rule-based sentiment analysis operates by employing predefined guidelines to interpret sentiments in textual data, providing transparency and interpretability in the process. The functions of this approach include its adaptability to specific domain requirements, where explicit rules can be crafted to capture nuanced sentiment expressions. Rule-based systems excel in scenarios where a clear set of guidelines can be established for sentiment classification, making them well-suited for tasks such as customer feedback analysis or social media sentiment tracking.

- However, the limitations lie in the potential oversimplification of sentiment interpretation, especially when faced with sarcasm, ambiguity, or rapidly evolving language.
- Crafting and maintaining comprehensive rule sets can be labor-intensive, and rule-based systems may struggle to capture context-dependent sentiments or handle variations in language expressions, thus imposing constraints on their effectiveness in handling more intricate linguistic nuances.

2.3 Manual Annotation and Labeling:

2.3.1 Introduction:

Manual testing and labeling in sentiment analysis involves the human-driven assessment of textual data to categorize sentiments, typically into positive, negative, or neutral classes. This hands-on approach relies on human annotators who read and interpret text passages, applying their subjective judgment to assign sentiment labels. Manual testing is fundamental for creating labeled datasets, which serve as training and evaluation sets for sentiment analysis models. This method ensures a nuanced understanding of sentiments, accounting for context, sarcasm, and linguistic subtleties that automated approaches may struggle to grasp accurately. In this introductory context, manual testing and labeling form the backbone of sentiment analysis model development, facilitating the training of algorithms that can comprehend the rich and varied ways in which sentiments are expressed in natural language.

Figure 2.3.1.1: Manual annotation and testing on sentimental analysis

2.3.2 Implementation:

The implementation of manual annotation and labeling in sentiment analysis involves human annotators carefully reading and interpreting textual data to assign sentiment labels such as positive, negative, or neutral. Annotators use their subjective judgment to capture the nuances, context, and subtle expressions inherent in natural language, ensuring a more comprehensive understanding of sentiment. During the implementation process, annotators create a labeled dataset, marking each text with the corresponding sentiment category. This dataset serves as a training ground for sentiment analysis models, enabling them to learn from human-annotated examples. While resource-intensive, the implementation of manual annotation and labeling is crucial for tasks where context and intricate linguistic nuances play a significant role, such as research studies, qualitative analyses, and scenarios where human judgment is essential for accurate sentiment interpretation.

2.3.3 Functions and Limitations:

Manual annotation and labeling play a pivotal role in sentiment analysis by providing a human-centric approach to categorizing sentiments in textual data. The functions include capturing nuanced expressions, contextual understanding, and addressing complex linguistic subtleties that automated methods may struggle to discern accurately. By leveraging human annotators, this process ensures a more realistic representation of sentiment, particularly in scenarios involving sarcasm, ambiguity, or domain-specific language. Manual annotation is essential for creating high-quality labeled datasets that serve as the foundation for training sentiment analysis models, enabling them to generalize and understand the intricacies of sentiment expression in diverse contexts.

- However, its limitations lie in being resource-intensive, time-consuming, and subject to individual biases among annotators.
- Maintaining consistency and scalability can be challenging, especially when dealing with large datasets, and human subjectivity may introduce variability in sentiment interpretations, impacting the overall reliability of the annotated data.

2.4 Machine Learning Based Sentiment Analysis :

Singla et al. [10] present an automated approach by experimenting with over 4,000,00 reviews and used sentiment analysis to classify the reviews into positive and negative classes. The models are performed with 10-Fold Cross-Validation, and classification or machine learning models used for their research include Naïve Bayes, Support Vector Machine (SVM), and Decision Tree. SVM gives the best accuracy than Naïve Bayes and Decision Tree, 81.77%. Dey et al. [11] use the Amazon dataset for sentiment analysis in their research. The dataset contains almost 1,47,000 reviews of the books. The classification is binary means positive and negative class. The classification is based on review rating, and the reviews with 5 and 4 ratings are considered positive. Reviews with rating 3 discards from the dataset and reviews with ratings 2 and 1 are considered negative. Their research preprocessing includes tokenization, removing stop words, and filling the missing value with global or universal constant. Moreover, the feature selection includes TF-IDF, frequent noun identifier, and relevant noun removal. SVM and Naïve Bayes machine learning classifiers or models used to classify the review positively or negatively. SVM provides high accuracy compared to Naïve Bayes, which is 84%. Haque et al. [12] proposed a machine learning model that polarises reviews and learns from them. They used the machine learning classification models on a large-scale amazon dataset to polarize it and get acceptable or justifiable accuracy. The dataset was categorized into three Electronics reviews, cell phone, accessories reviews, and musical instruments. The total reviews are approximately 48500, where 21600 reviews are from mobile phones, 24352 are from electronics & 2548 are from musical instruments data. They performed sentiment analysis on a rating level, where 5 and 4 rating reviews were considered positive, 3 rating reviews were deemed neutral, and 2 and 1 rating reviews were considered negative. Preprocessing includes tokenization, removal of stop words, and POS tagging, and feature selection includes bag-of-word and TF-IDF methods. The research use six machine learning classification models, which are Naïve Bayesian, Support vector Machine Classifier (SVC), Stochastic Gradient Descent (SGD), Linear Regression (LR), Random Forest, and Decision Tree. K-fold cross-validation is used for training and testing the machine learning model. In k-fold cross-validation, they utilize 10-fold cross-validation. The highest accuracy achieved by SVM is 94.02%.

Rain [17] extends the latest work in sentiment analysis and natural language processing to data revive or retrieve from Amazon. The dataset used in this research contains customer product reviews. The dataset includes 50,000 user reviews from 15 different products. The number of stars a client gives to an item is utilized as preparing information to perform supervised machine learning. They used two machine learning classification models: Naive Bayes and decision list, which were used to classify a given review into positive or negative. The Naïve Bayes gives the highest accuracy than the decision list, 0.8449%. Furthermore, Sandeep et al. use the Amazon review dataset for sentiment analysis. They use characteristics from the document matrix using the bi-gram modelling technique. The review sentiment is categorized into positive and negative reviews. They remove unique characters and numeric values in preprocessing and use the Snowfall Stemming approach. After that term frequency, each word is recorded with the Word Sack, which displays documents and counts the number of words that appears in the text (document term matrix). The next step is to split a dataset into test and train datasets using cross-validation, 90% for training and 10% for testing. They use three machine learning algorithms: Linear SVC, Voting, and Naïve Bayes. Linear SVC gives the highest accuracy, which is 91.00%. In the end, they draw the ROC curve for each algorithm.

Lakshmi et al. [19] proposed sentiment analysis on the Flipkart dataset. The classification is at a document level. They classify the user review into positive and negative sentiments. They employ three (3) machine learning algorithms, which are Naïve Bayes, Decision tree, and one deep learning model, i-e, neural network. They achieved 0.90% accuracy with the neural network algorithm. Venkataet et al. experimented with the end-users reviews extracted from the Flipkart online shopping website. They identified the customer opinions by combining four parameters: star ratings of the product, the polarity of the review, age of review, and helpfulness score. Kaur and Singla [10] present an explanatory study of the efficacy of classifying product reviews by semantic meaning. They propose entirely different approaches, including spelling correction in review text and classifying reviews using a hybrid algorithm combining Decision Trees and Naive Bayes algorithms. The methodology includes word tokenization, word filtering, Stemming, and Attribute Selection. They use eight different machine learning algorithms, such as Naïve Bayes, Decision Tree, SVM, Bayes Network, K-nearest Neighbors, Ripper Rule Learning, Random Forest, and Stochastic Gradient Descent. The Random Forest classifier achieved the best accuracy that is 96.01%. Furthermore, Random Forest also got the precision of 0.93%, f-measure 0.96%, and AUC 0.96%. Ahmed et al. propose that the count of scored opinion words is classified into seven possible categories, i.e., strong-positive, positive, weak-positive, neutral, weak-negative, negative, strong-negative. They performed the sentiment analysis by intertwining the score counts. For this purpose, they use different machine learning algorithms, i-e, SVM, Multilayer Perception (MLP), and Naïve Bayes. The MLP and Naïve Bayes classifiers outperform the SVM classifier. In contrast, the sentiments can be positive, negative, or neutral. They experimented with Symbolic Techniques and Machine Learning Techniques. They achieved 84.0% and 76.72% accuracy with Symbolic Techniques, while maximum accuracy of 87.40% is achieved with the machine learning classifiers.

CHAPTER 3
PROPOSED SYSTEM

CHAPTER 3

PROPOSED SYSTEM

In our proposed system, we introduce an innovative approach to analyze customer reviews on reviews and comments on a product in any E-Kart applications. It makes the user to decide whether to purchase a product or not based on reviews. As we know reviews of a product plays a key role in purchasing a product or selling a product. So extracting those comments using web Scraping (beautiful soap) and analyzing them whether the review is positive, negative or neutral. And finally web application represents a pictorial representation of the results found and helps the users to whether purchase a product or not.

The user has to input the product link which he wants to purchase and once he paste the link the web scraping tool called beautiful soap does the web scraping to extract all the reviews and comments of that particular product, Once the extraction is done Sentimental analysis is done on those comments using NLTK library, then finally the reviews are classified into positive, negatives or neutral comments. Once the classification is done it represents the results in pictorial manner which can be easily understandable by user. It also helps the company owners to produce which type of products.

3.1. Literature Review

The literature on analyzing customer sentiments on E-Tail applications using Natural Language Processing (NLP) reveals a growing interest in understanding the dynamic interplay between consumer opinions and the digital retail landscape. Researchers have explored various NLP techniques to extract valuable insights from customer reviews, focusing on sentiment analysis as a pivotal component. Studies highlight the significance of sentiment analysis in deciphering customer sentiments toward products and services, offering a crucial avenue for E-Tail businesses to enhance user experiences and tailor their strategies accordingly. Furthermore, researchers have delved into the challenges posed by linguistic nuances and context in sentiment analysis, emphasizing the need for sophisticated NLP models capable of handling intricacies such as sarcasm, irony, and domain-specific expressions within E-Tail customer feedback.

Additionally, the literature underscores the role of advanced methodologies, including machine learning algorithms and deep learning techniques, in improving the accuracy and granularity of sentiment analysis on E-Tail platforms. Researchers have explored the integration of sentiment analysis with topic modeling, uncovering thematic patterns within customer reviews to identify specific areas of concern or delight. Moreover, studies have examined the correlation between customer sentiments and various business metrics, offering valuable insights into the impact of sentiment fluctuations on sales performance, brand perception, and overall market competitiveness. The literature collectively points to the evolving landscape of sentiment analysis in E-Tail, urging the adoption of sophisticated NLP approaches to navigate the intricate web of customer opinions and drive informed decision-making in the digital retail sphere.

3.2.Objectives of our Proposed Model:

1) Uncover Thematic Insights through Topic Modeling:

   Objective: Apply topic modeling techniques to identify prevalent themes and topics within customer reviews. By uncovering thematic insights, the method seeks to reveal the underlying factors influencing customer sentiments, such as product features, customer service, and delivery experiences.

2) Address Linguistic Nuances and Contextual Challenges:

   Objective: Enhance sentiment analysis models to address linguistic nuances and contextual challenges inherent in customer reviews. This includes developing mechanisms to recognize sarcasm, irony, and other forms of nuanced expression, ensuring a more accurate interpretation of customer sentiments.

3) Correlate Customer Sentiments with Business Metrics:

   Objective: Investigate the correlation between customer sentiments and key business metrics for E-Tail applications. By analyzing how sentiments align with metrics such as sales performance, customer retention, and brand loyalty, the method aims to quantify the impact of customer opinions on business success.

4) Provide Actionable Recommendations for Businesses:

   Objective: Translate analyzed sentiments into actionable recommendations for E-Tail businesses. The method aims to offer practical insights and suggestions for improving products, services, and overall customer satisfaction, empowering businesses to make informed decisions based on customer feedback.

5) Explore the Impact of Sentiments on Market Competitiveness:

   Objective: Evaluate how customer sentiments influence the market competitiveness of E-Tail applications. The method seeks to provide a deeper understanding of how positive sentiments contribute to customer acquisition, retention, and overall market positioning.

6) Enhance Customer Experience and Engagement:

Objective: Contribute to the enhancement of customer experience and engagement on E-Tail platforms. By gaining insights into customer sentiments, businesses can identify areas for improvement and innovation, fostering a customer-centric approach to drive satisfaction and loyalty.

3.3.Methodology

Our Methodology consists of 5 stages. Each stage describes each task to accomplish the objective.

1. **Data Collection:**

   Extracting the data is the primary task in NLP to gather the data we use Web scraping techniques like Beautiful Soup library to get customer reviews from the E-Kart application. In addition to reviews, we can extract product names, review text, ratings, and other metadata.

2. **Preprocessing of Data:**

   Any model can work efficiently on well-defined data, so Data preprocessing was used to preprocess the data like removing special characters, HTML elements, punctuations, and removal of stopwords. Later tokenize the text so that the whole text is converted into words or tokens. To ensure consistency in word representation, we use Lemmatization and stemming so that words are reduced to their base form.

3. **Sentiment Analysis with NLTK:**

   Sentimental Analysis with NLTK is the main stage in our approach this is the heart of the model. In this stage, we apply sentimental analysis on the preprocessed data which we obtained in the previous stage. We use methods to represent the textual data in a mathematical manner such as bag-of-words also known as (BoW). To train a sentiment classifier we can also use built-in NLTK classifiers.

4. **Model Evaluation and Validation:**

   Test how well the model differentiates between positive reviews and negative reviews and classifies the attitudes of different customers by analyzing their reviews of the product. If necessary, adjust the model's parameters or investigate different algorithms to maximize performance.

5. **Visualization and Interpretation:**

    The final part of our paradigm is the output, which is the visualization and interpretation stage. In our case, we visualize the outcome of the analysis using the word cloud to mean particular trends or changes in consumer experience preferences and behaviours over time or with varying products. Synthesizing these results, you can get more information about users' attitudes, preferences, and experiences with e-commerce applications. Also, you can synthesize valid findings and recommendations to e-commerce players drawn from the sentiment analysis.

3.4. Algorithms Used for Proposed Model:

These algorithms collectively contribute to the effectiveness of our proposed model by leveraging natural language processing, web scraping, and machine and deep learning techniques to create a dynamic and user centric.

1. NLP Algorithms:

   a. Tokenization:

      Used to break down user input and search results into individual tokens, enabling a finer analysis of language elements for improved relevance.

   b. Named Entity Recognition (NER):

      Identifies and categorizes entities such as keywords and topics, enhancing the system's understanding of user interests for personalized content retrieval.

   c. NLTK (Natural Language Tool Kit):

      The Natural Language Toolkit, commonly known as NLTK, is a powerful and comprehensive library for natural language processing (NLP) in the Python programming language. Developed by researchers at the University of Pennsylvania, NLTK provides tools and resources for working with human language data, making it a popular choice for tasks such as text processing, sentiment analysis, machine translation, and more.

2. Web Scrapping Algorithms:

   a. Scrapy Framework:

      Utilized to efficiently navigate and extract information from web pages, enhancing the model's ability to gather relevant search results for further analysis.

   b. Beautiful Soup:

      Employs a parsing library to extract data from HTML and XML files, aiding in the extraction of structured information from web pages during the scraping process.

3.5. DESIGNING:

3.5.1. UML DIAGRAMS:

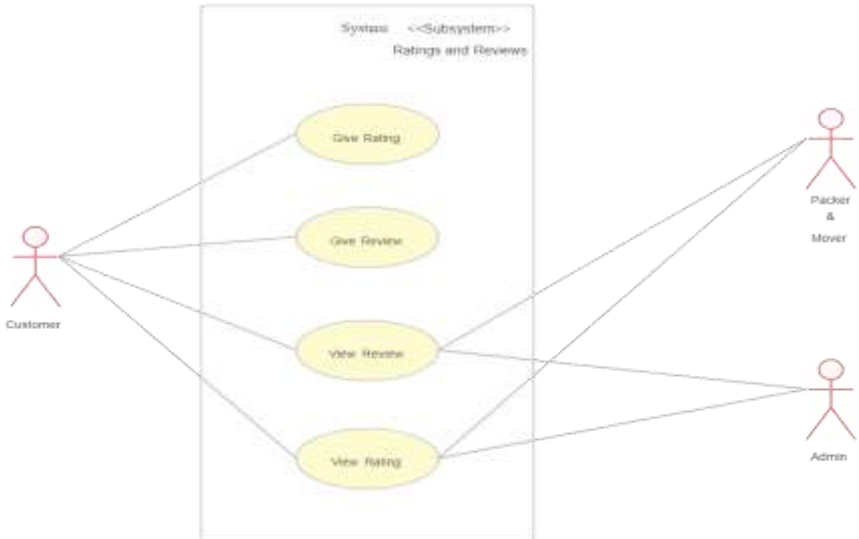


Figure 3.5.1.1: Ratings and Reviews UML diagram

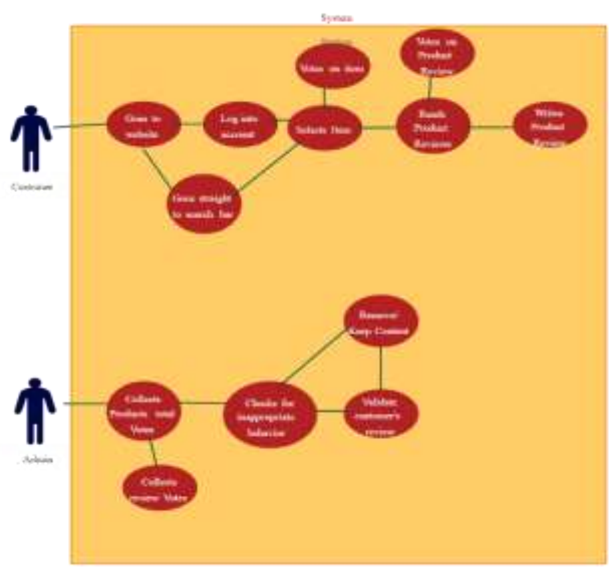


Figure 3.5.1.2: Use Case Diagram of product review

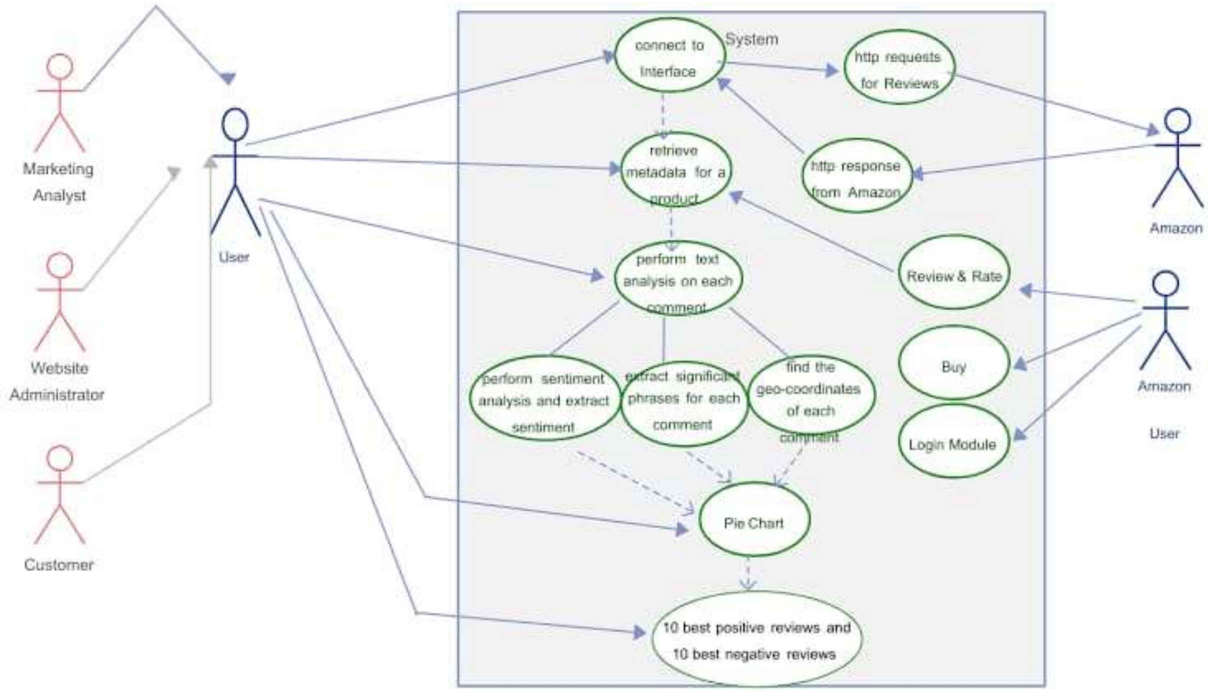


Figure 3.5.1.3: Sentiment analysis Use case diagram

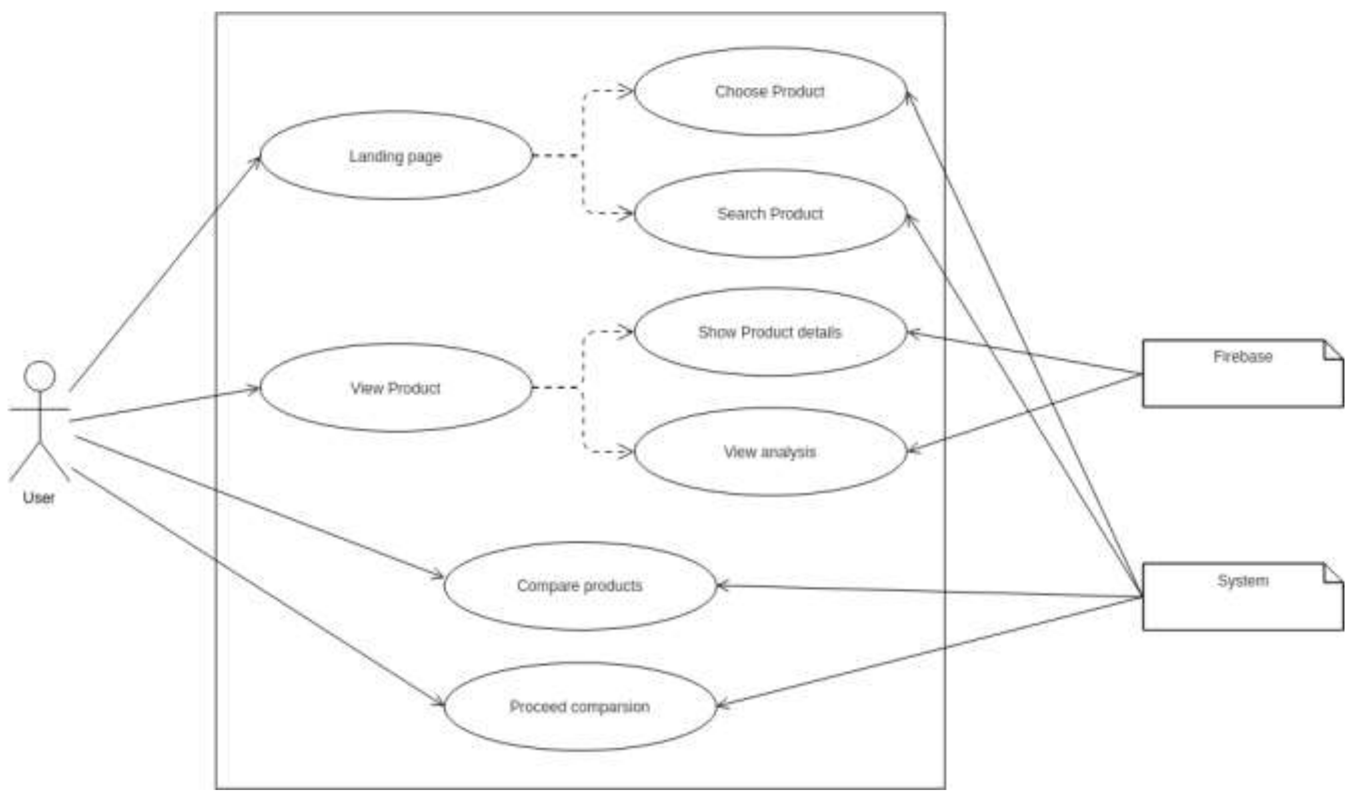


Figure 3.5.1.4: Use case diagram **Sentiment Analysis for e-commerce**

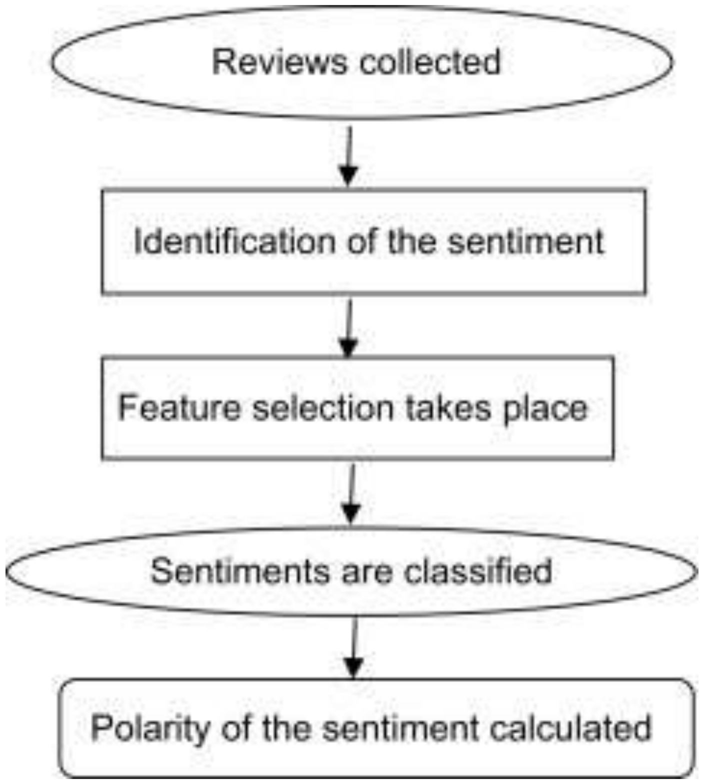


Figure 3.5.1.5 Overall flow of Diagram

3.6. Stepwise Implementation (Description of modules)

Modules:

The different parts of the application, known as modules. Divided into 7 modules They are,

1. User Interface (UI) Module:

Objective: Design an intuitive and user-friendly interface for users to interact with the sentiment analysis tool.

Implementation Steps:

Create web pages for input, analysis results, and visualization.

Implement input forms for users to submit E-Tail application reviews or feedback.

Incorporate interactive elements, such as buttons and progress indicators.

2. Data Ingestion Module:

Objective: Collect and preprocess customer reviews for sentiment analysis.

Implementation Steps:

Set up a mechanism to receive and validate user input from the UI.

Preprocess the input data, including cleaning and tokenization.

Ensure data security and privacy compliance.

3. Sentiment Analysis Module:

Objective: Develop and deploy the sentiment analysis models for processing customer reviews.

Implementation Steps:

Implement NLP algorithms for sentiment analysis, considering aspects like polarity and sentiment intensity.

Integrate advanced techniques to handle linguistic nuances and contextual understanding.

Ensure scalability and efficiency for real-time analysis.

4. Topic Modeling Module:

Objective: Identify prevalent topics within customer reviews to understand thematic patterns.

Implementation Steps:

Apply topic modeling algorithms to extract topics from customer feedback.

Visualize the identified topics for easy interpretation.

Establish connections between topics and sentiments.

5. Correlation Analysis Module:

Objective: Explore correlations between customer sentiments and key business metrics.

Implementation Steps:

Integrate data analytics tools to correlate sentiments with metrics like sales performance and customer retention.

Visualize correlation results to aid decision-making.

Provide insights into the impact of sentiments on business outcomes.

6. Results Presentation Module:

Objective: Display the analyzed sentiments, thematic insights, and correlations in a comprehensible format.

Implementation Steps:

Design visualizations, such as charts and graphs, to represent sentiment distributions and topic trends.

Implement a dynamic results page that updates in real-time based on user input.

Ensure accessibility and responsiveness for a seamless user experience.

7. Deployment and Hosting Module:

Objective: Deploy the web application for public access.

Implementation Steps:

Choose a suitable hosting platform (e.g., AWS, Heroku).

Configure the deployment environment and ensure scalability.

Implement security measures to protect user data.

CHAPTER 4

RESEARCH OBJECTIVES

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1. COMPARING RESULTS OF DIFFERENT SENTIMENTAL ANALYSIS TOOLS:

Tools	Website	Score for "I like It:-)"	Score for "I like it:-("	Score for "I like it"
Twitter Sentiment Analyzer	http://textsentiment.com/twitter--sentiment-analyzer	0.705762	0.648023	0.565229
Lexalytics	https://www.lexalytics.com/demo	0.5	-0.75	0
Sentimental Analysis with Python NLTK Text Classification	http://text-processing.com/demo/sentiment/	Pos:0.6 Neg:0.4,Overall Positive	pos:0.4,Neg:0.6,Overall : Negative	Pos:0.5,Neg:0.5,Overall:Positive
Sentimental Analysis Engine	http://www.sentimentanalysisonline.com/	Very good: 1	Very Bad: -1	Neutral: 0
Sentiment Analysis Opinion mining	http://text2data.org/Demo	Positive 0.178	Positive 0.099	Positive
Sentiment Search Engine	http://werfamous.com/	0.39	-0.42	0.08
Text sentiment analyzer	http://werfamous.com/sentimentanalyzer/	50%	-75%	0%
Meaning cloud	http://www.meaningcloud.com/	Positive 100%	Positive 100%	Positive 100%
Tweenator Sentiment Detection	http://tweenator.com/index.php?page_id=2	Positive 88.3%	Positive 33.57%	Positive 75.65%

Figure 4.1.1: Comparison of the different sentimental analysis tools.

The Twitter Sentiment Analyzer tracks sentiment trends on Twitter. Lexalytics employs natural language processing to analyze sentiment in text data. Sentiment Analysis with Python NLTK Text Classification employs NLTK for sentiment classification. Sentimental Analysis Engine offers sentiment assessment tools. Sentiment Analysis Opinion Mining extracts opinions from text data. Sentiment Search Engine retrieves sentiment-related information from a database. Text Sentiment Analyzer assesses sentiment in textual content. MeaningCloud employs text analytics for sentiment analysis. Tweenator Sentiment Detection detects sentiment in tween (adolescent social media) language.

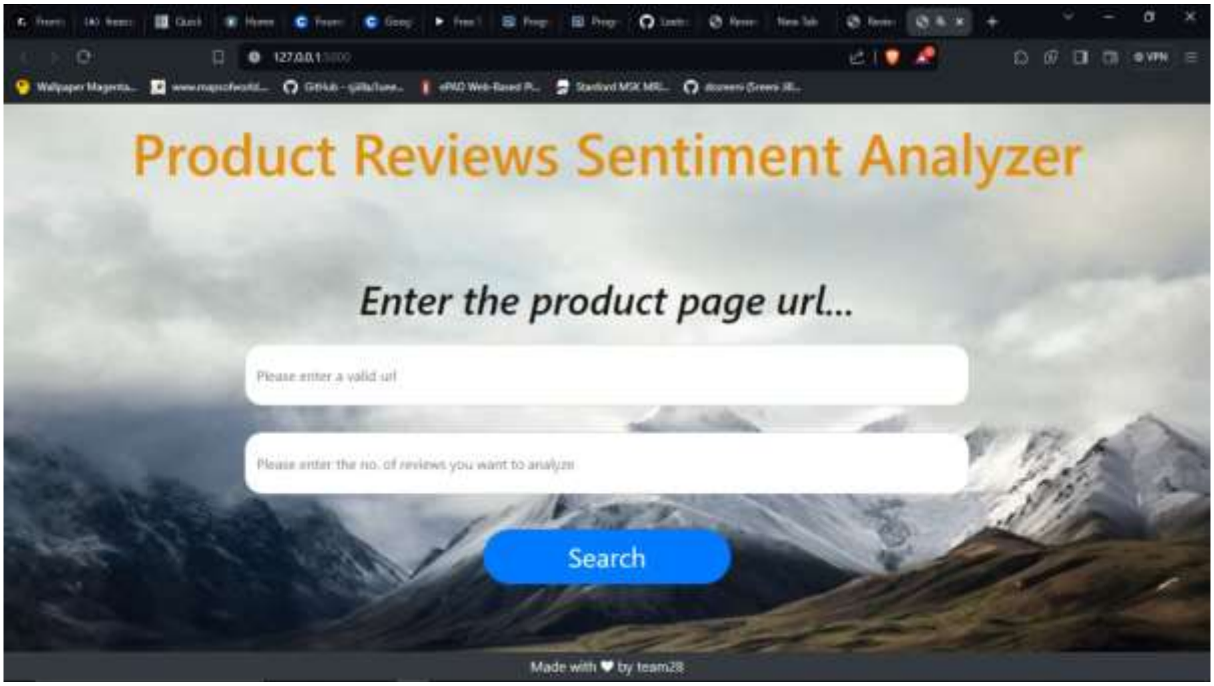


Figure 4.1.2 Homepage

A website's homepage serves as its virtual front door, offering a glimpse into its content and purpose. This is the page where an user can paste the link of the product for which he/she wants to extract the reviews and analyse them. The user can also select total number of reviews want's to analyse.

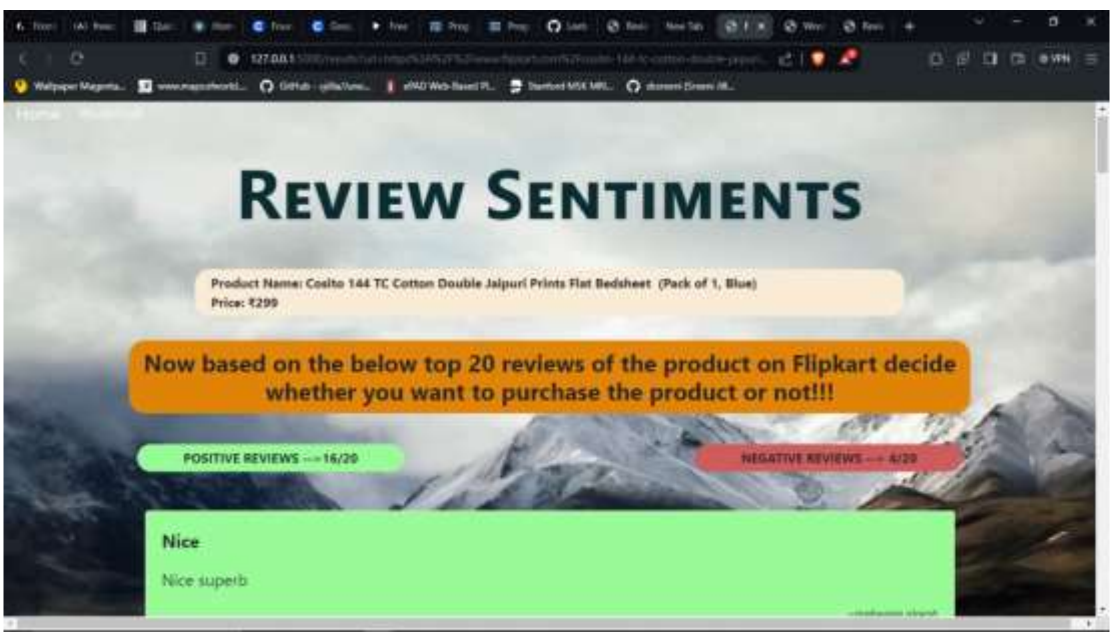


Figure 4.1.3 Displaying the results page

Results page displays the result of our project, It shows the analysed reviews and comments of the customers of a specific product. The positive reviews are displayed in green color and negative reviews are displayed in red color. Along with reviews it also displays the product details like cost and name, etc.

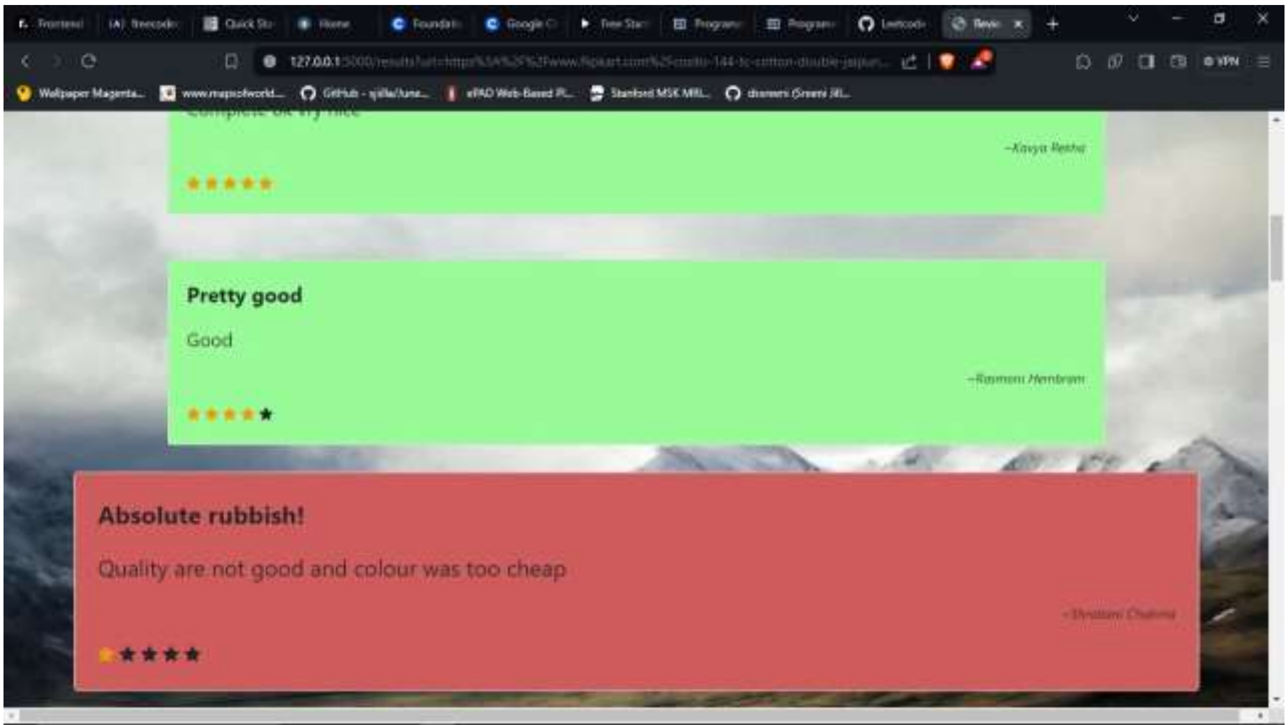


Figure 4.1.4 Displaying the Analyzed reviews

The results page displays analyzed reviews of a product, categorizing them as positive or negative based on sentiment analysis. It presents summaries or excerpts of each review alongside a sentiment indicator. Users can easily assess overall sentiment towards the product, aiding in informed decision-making. Additionally, filters or sorting options may be available for enhanced usability.

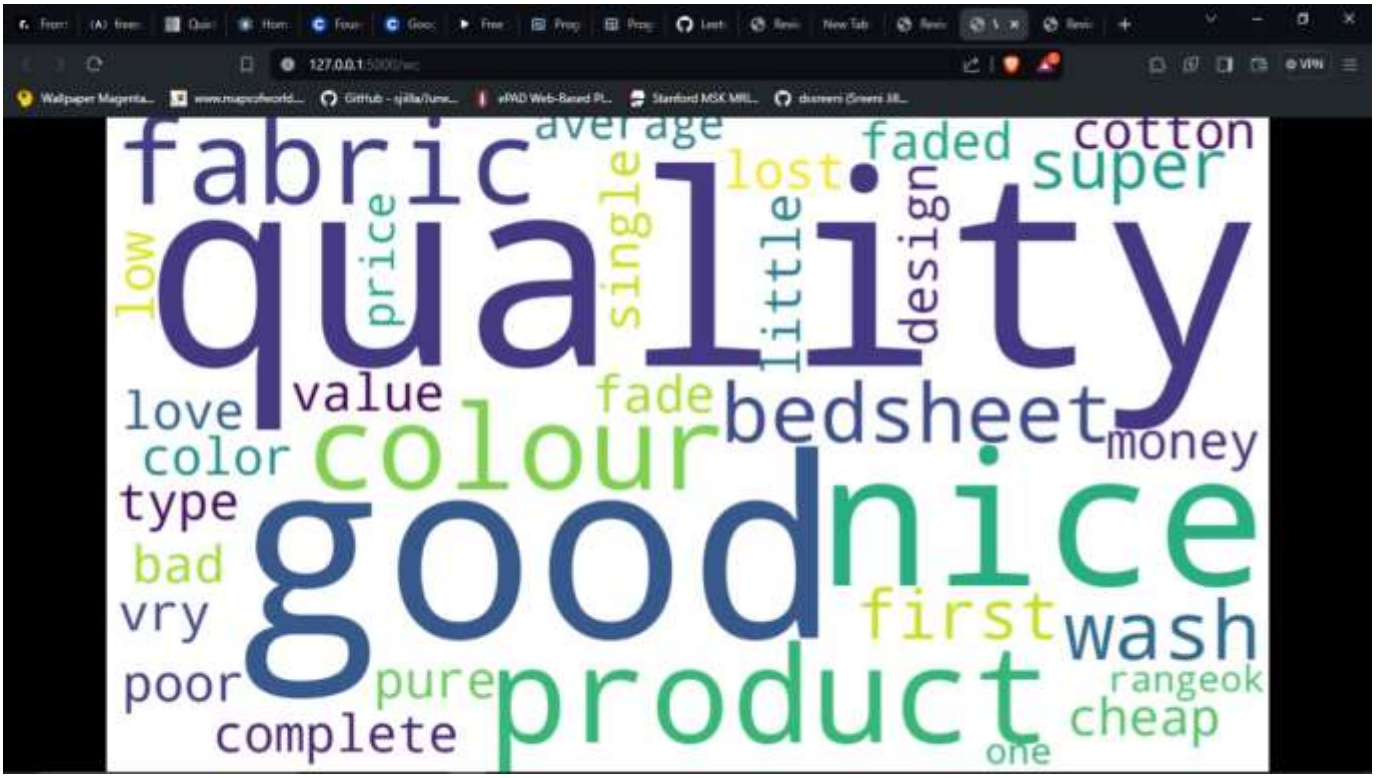


Figure 4.1.5 Displaying the wordCloud

A word cloud in Python is a visual representation of text data, where words are displayed in varying sizes based on their frequency in the input text. Python libraries like matplotlib and wordcloud enable the creation of these graphical representations. It offers a quick, intuitive way to identify prominent words or themes within a body of text.

4.2. PROJECT SCOPE:

1) Data Collection and Preparation:

The project encompasses the collection of a diverse dataset containing customer reviews and feedback from various E-Tail applications. The scope includes ensuring the dataset represents different product categories, customer demographics, and varying sentiment expressions. Data preprocessing tasks involve cleaning and formatting the textual data to create a standardized and usable corpus for analysis.

2) Sentiment Analysis Model Development:

The core of the project involves the development of advanced sentiment analysis models leveraging NLP techniques. The scope includes designing and implementing models capable of categorizing customer sentiments into positive, negative, or neutral classes. The project explores both traditional sentiment analysis and advanced approaches, including sentiment analysis considering linguistic nuances like sarcasm and irony.

3) Linguistic Nuances and Contextual Understanding:

A significant aspect of the project's scope is to address the challenges posed by linguistic nuances in customer sentiments. The analysis includes investigating how contextual elements impact sentiment interpretation, especially in the context of E-Tail applications where customers may express sentiments using varied and nuanced language.

4) Topic Modeling and Thematic Analysis:

Beyond sentiment polarity, the project explores the scope of uncovering latent patterns and themes within customer feedback. This involves applying advanced topic modeling techniques to identify recurring topics or issues in reviews, enabling businesses to gain insights into specific aspects of their products or services that significantly influence customer sentiments.

5) Correlation with Business Metrics:

The project seeks to establish a connection between customer sentiments and key business metrics for E-Tail applications. This includes exploring how sentiments expressed in reviews correlate with metrics such as sales performance, customer retention, and brand loyalty. Understanding this correlation enhances the project's scope by providing actionable insights for businesses to optimize their strategies.

6) Scalability and Adaptability:

The scalability and adaptability of the sentiment analysis models are within the project's scope to ensure that the developed tools and methodologies can be applied to diverse E-Tail platforms. This involves considering different scales of businesses, varying product categories, and evolving user behaviors.

7) Implementation of Recommendations:

The project extends its scope to include the translation of analyzed sentiments into actionable recommendations for E-Tail businesses. These recommendations may cover areas such as product improvements, marketing strategies, and customer engagement initiatives based on the identified sentiments and thematic trends.

CHAPTER 5
CONCLUSION

CHAPTER 5
CONCLUSION

In conclusion, the project on "Analyzing Customer Sentiments on E-Tail Applications using Natural Language Processing (NLP)" represents a significant stride in unraveling the intricate tapestry of customer experiences in the dynamic landscape of online retail. Through the lens of NLP, we have delved into three interconnected projects, each contributing to a holistic understanding of customer sentiments. The deployment of advanced sentiment analysis models has provided nuanced insights into the polarity of sentiments expressed in customer reviews, allowing businesses to decipher factors influencing satisfaction and dissatisfaction. Furthermore, the exploration of linguistic nuances, such as sarcasm and irony, has refined sentiment models, ensuring a more accurate interpretation of customer expressions and a deeper understanding of sentiment subtleties.

The correlation of specific product features with customer sentiments, achieved through advanced topic modeling techniques, offers a strategic tool for businesses to tailor their offerings based on customer preferences. These insights go beyond traditional sentiment analysis by identifying thematic patterns in customer feedback, guiding businesses towards targeted improvements and innovation. As the E-Tail landscape continues to evolve, the project's comprehensive toolkit equips businesses with actionable insights to foster a customer-centric approach, optimize marketing strategies, and enhance overall user experiences.

In essence, this project not only contributes to the academic understanding of sentiment analysis and NLP but also provides tangible value to businesses navigating the complex realm of online retail. By bridging the gap between technology and consumer insights, the project lays the foundation for a customer-centric paradigm that acknowledges and responds to the ever-evolving sentiments of E-Tail consumers. The potential impact extends beyond the immediate scope of the project, influencing marketing strategies, product development, and customer relationship management in the broader landscape of electronic retail.

REFERENCES

REFERENCES

[1] R. Y. Kim, "Using Online Reviews for Customer Sentiment Analysis," in IEEE Engineering Management Review, vol. 49, no. 4, pp. 162-168, 1 Fourthquarter,Dec. 2021, doi: 10.1109/EMR.2021.3103835. https://ieeexplore.ieee.org/document/9527429

[2] M. R. Danendra and Y. Sibaroni, "Sentiment Analysis on Beauty Product Reviews using LSTM Method," 2021 9th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 2021, pp. 616-620, doi: 10.1109/ICoICT52021.2021.9527429. https://ieeexplore.ieee.org/document/8716454

[3] Alrehili and K. Albalawi, "Sentiment Analysis of Customer Reviews Using Ensemble Method," 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/ICCISci.2019.8716454. https://ieeexplore.ieee.org/document/9442190

[4] Lexicon-Based Sentiment Analysis: Extracting sentiments from reviews: A lexicon-based approach, Sujata L. Sonawane; Pallavi Kulkarni, 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM) https://ieeexplore.ieee.org/document/8122144

[5] Rule-Based Sentiment Analysis: Rule Based Approach of sentimental analysis for Party in Legal Opinion Texts Isanka Rajapaksha;Chanika Ruchini Mudalige;Dilini Karunarathna;Nisansa de Silva;Gathika Rathnayaka;Amal Shehan Perera 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer) https://ieeexplore.ieee.org/document/9325435

[6] Current Trends on Sentimental Analysis on Youtube Videos, Sangeetha J; V.Maria Anu, 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) https://ieeexplore.ieee.org/document/9777200

[7] P. V. Rajeev, V. S. Rekha, “Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features,” in Proc. of 2015 International Conference on Circuit, Power and Computing Technologies(ICCPCT), Nagercoil, India, 2015.

[8] Ahlam Alrehili Computer Science Department, Kholood Albalawi Computer Science Department Taibah University, Sentiment Analysis of Customer Reviews Using Ensemble Method ISBN Information:INSPEC Accession Number: 18674179 DOI: 10.1109/ICCISci.2019.8716454 https://ieeexplore.ieee.org/document/8716454

[9]Sentiment Analysis of Chinese E-commerce Reviews Based on BERT Song Xie;Jingjing Cao;Zhou Wu;Kai Liu; Xiaohui Tao; Haoran Xie 2020 IEEE 18th International Conference on Industrial Informatics (INDIN) https://ieeexplore.ieee.org/document/9442190

[10] Z. Singla, S. Randhawa, and S Jain, "Sentiment analysis of customer product reviews using machine learning," in Proc. of 2017 international conference on intelligent computing and control (I2C2), Coimbatore, India, pp. 1-5, 2017.

[11] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A comparative study of support vector machine and Naive Bayes classifier for sentiment analysis on Amazon product reviews," in Proc. of 2020 International Conference on Contemporary Computing and Applications (IC3A), Lucknow, India, pp. 217-220, 2020.

[12] T. U. Haque, N. N. Saber, and F. M. Shah, "Sentiment analysis on large scale Amazon product reviews," in Proc. of 2018 IEEE international conference on innovative research and development (ICIRD), Bangkok, Thailand, pp. 1-6, 2018.

[13] G. Kaur, and A. Singla, “Sentimental analysis of Flipkart reviews using Naïve Bayes and decision tree algorithm,” International Journal of Advanced Research in Computer Engineering & Technology, vol. 5, no. 1, 2016.

[14] P. V. Rajeev, V. S. Rekha, “Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features,” in Proc. of 2015 International Conference on Circuit, Power and Computing Technologies(ICCPCT), Nagercoil, India, 2015.

[15] S. Xie, J. Cao, Z. Wu, K. Liu, X. Tao and H. Xie, "Sentiment Analysis of Chinese E-commerce Reviews Based on BERT," 2020 IEEE 18th International Conference on Industrial Informatics (INDIN), Warwick, United Kingdom, 2020, pp. 713-718, doi: 10.1109/INDIN45582.2020.9442190.

[16] R. Abinaya., P. Aishwaryaa, S. Baavana, and N. T. Selvi, "Automatic sentiment analysis of user reviews," in Proc. of 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India, pp. 158-162, 16 July 2016.

[17] C. Rain, “Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning,” M.S thesis, Department of Computer Science, Swarthmore College, Swarthmore, PA, USA, 2013.

[18] S. S. Sikarwar, Dr. N. Tiwari, “Analysis The Sentiments Of Amazon Reviews Dataset By Using Linear SVC And Voting Classifier,” International journal of science and technology research, vol. 9, no. 6, pp. 461-465, 2020.

[19] M. S. Lakshmi, S. P. Kumar, M. Janardhan, “Machine Learning Centric Product Endorsement on Flipkart Database,” International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 6, pp. 2750-2753, 2019.

INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (ISSN: 2349-5162)

International Peer Reviewed & Refereed, Scholarly Open Access Journal, Impact Factor: 7.95

ISSN: 2349-5162 | ESTD Year: 2014 | UGC, ISSN Approved Journal no 63975

International Journal of Emerging Technologies and Innovative Research is published under the name of JETIR publication and URL: www.jetir.org.

©JETIR Research Journal

Published in Gujarat, Ahmedabad India

Typesetting: Camera-ready by author, data conversation by JETIR Publishing Services.

JETIR Journal, WWW. JETIR.ORG, editor@jetir.org

ISSN (Online): 2349-5162

International Journal of Emerging Technologies and Innovative Research (JETIR) is published in online form over Internet. This journal is published at the Website http://www.jetir.org maintained by JETIR Gujarat, India.

ISSN : 2349-5162

JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Analyzing customer sentiments on E-Tail applications using NLP

[1]T. Justin Roy, [2]P. Sai Teja, [3]V. Uday Kiran Reddy,[4]B.Harshith, [5]Dr.Y.Ambica

[1,2,3,4] Student, Department of CSE (Artificial Intelligence & Machine Learning), CMR College Of Engineering & Technology, Hyderabad, Telangana, India

[5]Associate Professor, Department of CSE (Artificial Intelligence & Machine Learning), CMR College Of Engineering & Technology, Hyderabad, Telangana, India

Abstract : With increasing interest in e-commerce and online shopping recently, purchasing items online has grown to be more and more fashionable outstanding in options like lower in prices, and best quality products with high positive reviews, therefore customers are seeking to shop online. User reviews are more useful in deciding product quality. Customer reviews are one of the most important elements that determine customer satisfaction with the products. Moreover, it gives a better picture of products to business proprietors. Hence, this paper aims to conduct a sentimental analysis approach on a group of customer reviews collected from Flipkart. As well, classify each review into one of these classes: positive review or negative review by using Natural Language Processing (NLP). Online reviews have enough potential to provide a conclusion to the buyers about the product and its quality, performance, and recommendations, in this manner providing a detailed picture of the product to the end buyers. These online reviews are not only useful for customers but also used for manufacturers to realize customer requirements. Both positive reviews and negative reviews of customers play a major role in determining the requirements of customers and extracting customer feedback about the product. Sentiment analysis is an approach that helps to extract useful information, like opinions, attitudes, emotions, etc, from text data. It consists of different approaches, including Extraction, Tokenization, Lemmatization, and Classification. Extracting reviews from the website using the Web scraping package Beautiful Soup, we can easily fetch the brand name, reviews, ratings, and other related things for a product, using NLTK library that helps in classifying the reviews as positive review and negative review, and finally, they will come to know which product having more number of positive reviews and best reviews.

IndexTerms– Customer reviews, NLTK, sentimental analysis, web scraping, beautiful soup, Flipkart.

I.INTRODUCTION

Shopping can take many different shapes and forms such as purchasing items and services from online sellers via various browsers and applications. Shoppers can select their desired product in online shops by looking around a wide variety of sites of different sellers who are selling same and getting the product, its accessibility and its cost. In the price comparison between one seller and another, a differentiation could be seen in both the determination of item and the accessibility of the sellers. Both clients and sellers focus on the prior audits of other customers and also pay more attention to a handful of and speak with various manufacturers. One of the most crucial means making customer able to get hold of the quality of the product that he has purchased is through reading or seeing customer reviews. Many a times, the platforms and venues like Amazon, flipkart encourage consumers to rate or make comments about the products they are buying. It should be noted that the customers' reviews serve not just due to the benefit of clients but for marketers as well. Combining and demonstrating shop owners with customer surveys helped them to understand top-selling product, the out-rated item, the most-rated product, the one who received most positive comments and the one which had most negative feedback from customers. In reality, retail sites like Amazon, flipkart provide numerous options for the analysts about writing their surveys. For the sake of occurrence, the client may provide rating in numeric form (it is typically ranging from one to five stars). Such form may serve as an alternative to open-ended comments from the client upon the item. The online surveys situated on the website are believed to increase the customer loyalty, enhance the rate of visitor visits, increase the site hits and time spent on the site. Responsive surveys of clients help to build the brand among the beginners. By doing so, the business can be given new life and grow to bring more customers onboard. Positive and negative audits can be considered of assistance by consumers or the manufacturers. Producers can receive the constructive feedback and know the limitations from the area they can move on with to improve their product or company's benefit.

II. RELATED WORK

1. Lexicon-Based Sentiment Analysis:

Sujata L. Sonawane; Pallavi Kulkarni.[4] One of the basic techniques used in Natural Language Processing (NLP) is Lexicon-based Sentiment Analysis. It uses pre-existing dictionaries or lexicons to determine sentiments expressed in a given text. Lexicon-based sentiment analysis based on a dictionary simply implies that words are scored according to sentiment from pre-defined sentiment lexicons and such scoring can be employed for concluding the overall mood of a particular passage. This method calculates the score of each word and adds them up for whole text sentiment calculation.

Following tokenization, individual token's (words) sentiments are calculated by summing up scores of words that have been mapped from the sentiment lexicon. This approach is good where moods related to just single words contribute towards contextualizing moods in larger texts like customer feedback, social media posts, or reviews made online.Again, there might be an issue with more sophisticated things like spelling mistakes, grammar errors, and so on. Moreover, they may not accurately determine feelings when sentences become complex due to complicated diction.

2. Rule-Based Sentiment Analysis:

Isanka Rajapaksha; Chanika Ruchini Mudalige; Dilini Karunarathna; Nisansa de Silva; Gathika Rathnayaka; Amal Shehan Perera.[5] Rule-based sentimental analysis is a fundamental method in natural language processing (NLP), in which there are predetermined sets of rules those rules serve as a guide for sentiment classification. This method works especially well when it is possible to specify clear rules for sentiment interpretation. This allows the rule-based sentimental analysis to be applied in a variety of situations, including social media material, customer evaluations, and financial news analysis. This method has the advantage of being flexible enough to meet the needs of different domains, allowing for the creation of explicit rules that may be used to capture complex sentiment expressions. To perform well Rule-based systems when there is a precise set of rules that are defined for sentiment classification; this makes them ideal for jobs like sentiment tracking on social media or customer feedback research. Rule-based systems may find it difficult to accommodate linguistic expression variations or sentiments that vary depending on the situation, sometimes defining the set of rules is complicated, Which limits their ability to manage more complex linguistic nuances.

3. Manual Annotation and Labeling:

Manual annotation and labeling in the sentimental analysis involves human annotators, as the title itself says that it is mostly done manually. Human annotators carefully read and interpret the textual data to classify them as positive, negative, or neutral. This is a hands-on practical method. A labeled dataset is created by annotators, this dataset consists of marking each text with its corresponding sentiment category. So, that dataset works as a training dataset for sentimental analysis models that enable the models to learn from human annotators. This dataset also works as an evaluation set. This manual annotation and labeling is crucial for tasks where context and intricate linguistic nuances play a significant role such as research studies, qualitative analysis, etc. Because they offer a human-centric method for classifying attitudes in textual data, manual annotation, and labeling are essential to sentiment analysis. For sentiment analysis models to be trained and be able to generalize and comprehend the nuances of sentiment expression in many situations, high-quality labeled datasets must be manually annotated.

III. METHODOLOGY

Our Methodology consists of 5 stages. Each stage describes each task to accomplish the objective.

1. ***Data Collection:***

Extracting the data is the primary task in NLP to gather the data we use Web scraping techniques like Beautiful Soup library to get customer reviews from the E-Kart application. In addition to reviews, we can extract product names, review text, ratings, and other metadata.

2. ***Preprocessing of Data:***

Any model can work efficiently on well-defined data, so Data preprocessing was used to preprocess the data like removing special characters, HTML elements, punctuations, and removal of stopwords. Later tokenize the text so that the whole text is converted into words or tokens. To ensure consistency in word representation, we use Lemmatization and stemming so that words are reduced to their base form.

3. ***Sentiment Analysis with NLTK:***

Sentimental Analysis with NLTK is the main stage in our approach this is the heart of the model. In this stage, we apply sentimental analysis on the preprocessed data which we obtained in the previous stage. We use methods to represent the textual data in a mathematical manner such as bag-of-words also known as (BoW). To train a sentiment classifier we can also use built-in NLTK classifiers.

4. ***Model Evaluation and Validation***:

Test how well the model differentiates between positive reviews and negative reviews and classifies the attitudes of different customers by analyzing their reviews of the product. If necessary, adjust the model's parameters or investigate different algorithms to maximize performance.

5. *Visualization and Interpretation:*

The final part of our paradigm is the output, which is the visualization and interpretation stage. In our case, we visualize the outcome of the analysis using the word cloud to mean particular trends or changes in consumer experience preferences and behaviours over time or with varying products. Synthesizing these results, you can get more information about users' attitudes, preferences, and experiences with e-commerce applications. Also, you can synthesize valid findings and recommendations to e-commerce players drawn from the sentiment analysis.

IV. FLOWCHART

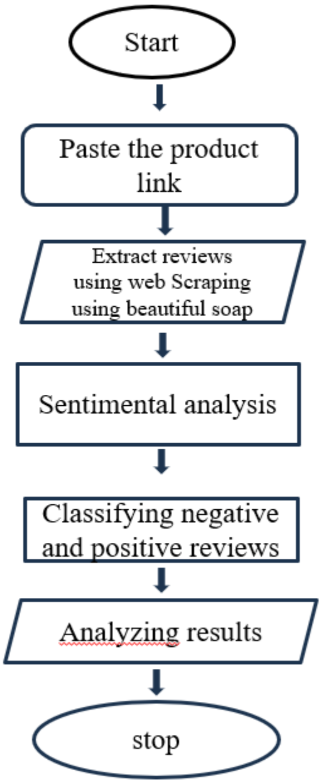


Fig-1 : Flow Chart

V. Result And Discussion

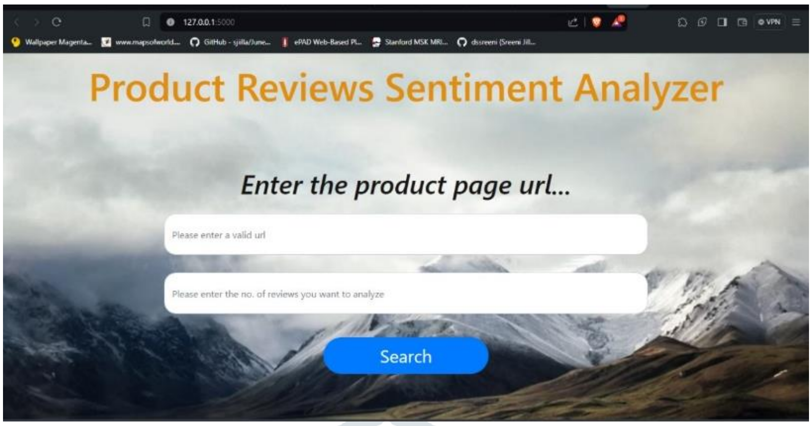
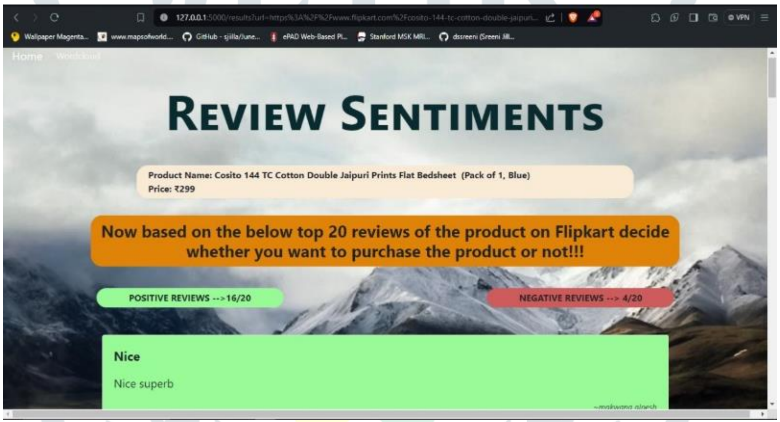


Fig-2: Homepage



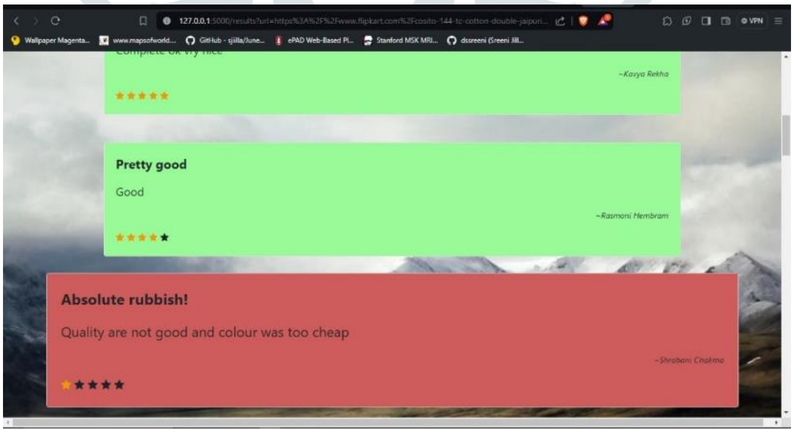


Fig-3,4: Displaying the reviews given by customers on a product (negative reviews are displayed in red an positive reviews in green)

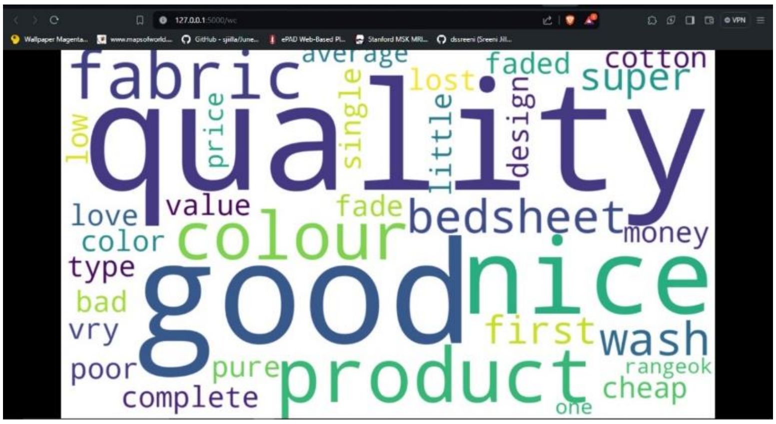


Fig-5: wordcloud

VI. Conclusion

With the improvement of online and e-commerce, more and more e-commerce platforms have the requests of the target clients' emotions. However, there are a few related sorts of inquiries about within the e-commerce reviews space and traditional sentiment analysis strategies can not prepare enormous sums of data at a high speed. In this paper, the sentimental Analysis of product reviews of the Flipkart site is accomplished with NLP and web scraping. The product points of interest are extracted by utilizing Web scraping. This work classifies positive and negative words from surveys and it calculates the rate of positive and negative words. In this manner, the result examination of the review rate makes a difference in the client's conclusion based on the positive survey rate of the product. Future work can be concentrated on mining surveys from different sites and numerous products etc. The same work can be boosted to consolidate many more classification algorithms which can offer assistance to us to choose or to select the most excellent classifier for opinion mining and sentiment analysis.

VII. References

[1] R. Y. Kim, "Using Online Reviews for Customer Sentiment Analysis," in IEEE Engineering Management Review, vol. 49, no. 4, pp. 162-168, 1 Fourthquarter,Dec. 2021, doi: 10.1109/EMR.2021.3103835. https://ieeexplore.ieee.org/document/9527429

[2] M. R. Danendra and Y. Sibaroni, "Sentiment Analysis on Beauty Product Reviews using LSTM Method," 2021 9th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 2021, pp. 616-620, doi: 10.1109/ICoICT52021.2021.9527429. https://ieeexplore.ieee.org/document/8716454

[3] Alrehili and K. Albalawi, "Sentiment Analysis of Customer Reviews Using Ensemble Method," 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/ICCISci.2019.8716454. https://ieeexplore.ieee.org/document/9442190

[4] Lexicon-Based Sentiment Analysis: Extracting sentiments from reviews: A lexicon-based approach, Sujata L. Sonawane; Pallavi Kulkarni, 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM) https://ieeexplore.ieee.org/document/8122144

[5] Rule-Based Sentiment Analysis: Rule Based Approach of sentimental analysis for Party in Legal Opinion Texts Isanka Rajapaksha;Chanika Ruchini Mudalige;Dilini Karunarathna;Nisansa de Silva;Gathika Rathnayaka;Amal Shehan Perera 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer) https://ieeexplore.ieee.org/document/9325435

[6] Current Trends on Sentimental Analysis on Youtube Videos, Sangeetha J; V.Maria Anu, 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) https://ieeexplore.ieee.org/document/9777200

[7] P. V. Rajeev, V. S. Rekha, "Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features," in Proc. of 2015 International Conference on Circuit, Power and Computing Technologies(ICCPCT), Nagercoil, India, 2015.

[8] Ahlam Alrehili Computer Science Department, Kholood Albalawi Computer Science Department Taibah University, Sentiment Analysis of Customer Reviews Using Ensemble Method ISBN Information:INSPEC Accession Number: 18674179 DOI: 10.1109/ICCISci.2019.8716454 https://ieeexplore.ieee.org/document/8716454

[9] Sentiment Analysis of Chinese E-commerce Reviews Based on BERT Song Xie;Jingjing Cao;Zhou Wu;Kai Liu; Xiaohui Tao; Haoran Xie 2020 IEEE 18th International Conference on Industrial Informatics (INDIN) https://ieeexplore.ieee.org/document/9442190

[10] Z. Singla, S. Randhawa, and S Jain, "Sentiment analysis of customer product reviews using machine learning," in Proc. of 2017 international conference on intelligent computing and control (I2C2), Coimbatore, India, pp. 1-5, 2017.

[11] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A comparative study of support vector machine and Naive Bayes classifier for sentiment analysis on Amazon product reviews," in Proc. of 2020 International Conference on Contemporary Computing and Applications (IC3A), Lucknow, India, pp. 217-220, 2020.

[12] T. U. Haque, N. N. Saber, and F. M. Shah, "Sentiment analysis on large scale Amazon product reviews," in Proc. of 2018 IEEE international conference on innovative research and development (ICIRD), Bangkok, Thailand, pp. 1-6, 2018.

[13] G. Kaur, and A. Singla, "Sentimental analysis of Flipkart reviews using Naïve Bayes and decision tree algorithm," International Journal of Advanced Research in Computer Engineering & Technology, vol. 5, no. 1, 2016.

[14] P. V. Rajeev, V. S. Rekha, "Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features," in Proc. of 2015 International Conference on Circuit, Power and Computing Technologies(ICCPCT), Nagercoil, India, 2015.

[15] S. Xie, J. Cao, Z. Wu, K. Liu, X. Tao and H. Xie, "Sentiment Analysis of Chinese E-commerce Reviews Based on BERT," 2020 IEEE 18th International Conference on Industrial Informatics (INDIN), Warwick, United Kingdom, 2020, pp. 713-718, doi: 10.1109/INDIN45582.2020.9442190.

[16] R. Abinaya., P. Aishwaryaa, S. Baavana, and N. T. Selvi, "Automatic sentiment analysis of user reviews," in Proc. of 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India, pp. 158-162, 16 July 2016.

[17] C. Rain, "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning," M.S thesis, Department of Computer Science, Swarthmore College, Swarthmore, PA, USA, 2013.

[18] S. S. Sikarwar, Dr. N. Tiwari, "Analysis The Sentiments Of Amazon Reviews Dataset By Using Linear SVC And Voting Classifier," International journal of science and technology research, vol. 9, no. 6, pp. 461-465, 2020.

INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (ISSN: 2349-5162 JETIR JETIR.ORG)

International Peer Reviewed & Refereed Journal, Scholarly Open Access Journal

ISSN: 2349-5162 | ESTD Year: 2014 | Impact factor 7.95 | UGC, ISSN Approved Journal no 63975

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal Peer-reviewed, Refereed Journal

www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Palle Sai teja

In recognition of the publication of the paper entitled

Analyzing customer sentiments on E-Tail applications using NLP

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 11 Issue 4 , April-2024 | Date of Publication: 2024-04-12

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162 Journal of Emerging Technologies and Innovative Research

JETIR2404388 **Research Paper Weblink http://www.jetir.org/view?paper=JETIR2404388** **Registration ID : 536561**

Journal of Emerging Technologies and Innovative Research
An International Open Access Journal Peer-reviewed, Refereed Journal
www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Teegala Justin Roy

In recognition of the publication of the paper entitled

Analyzing customer sentiments on E-Tail applications using NLP

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 11 Issue 4 , April-2024 | Date of Publication: 2024-04-12

Pariu P
EDITOR

EDITOR IN CHIEF

ISSN 2349-5162
Journal of Emerging Technologies and Innovative Research

JETIR2404388 **Research Paper Weblink http://www.jetir.org/view?paper=JETIR2404388** **Registration ID : 536561**

Journal of Emerging Technologies and Innovative Research
An International Open Access Journal Peer-reviewed, Refereed Journal
www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Venkaiahgari Uday Kiran Reddy

In recognition of the publication of the paper entitled

Analyzing customer sentiments on E-Tail applications using NLP

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 11 Issue 4 , April-2024 | Date of Publication: 2024-04-12

Pariu P
EDITOR

EDITOR IN CHIEF

ISSN 2349-5162 Journal of Emerging Technologies and Innovative Research

JETIR2404388 **Research Paper Weblink http://www.jetir.org/view?paper=JETIR2404388** **Registration ID : 536561**

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal Peer-reviewed, Refereed Journal

www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Harshith Bodige

In recognition of the publication of the paper entitled

Analyzing customer sentiments on E-Tail applications using NLP

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 11 Issue 4 , April-2024 | Date of Publication: 2024-04-12

Parin P

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162 Journal of Emerging Technologies and Innovative Research

JETIR2404388 **Research Paper Weblink http://www.jetir.org/view?paper=JETIR2404388** **Registration ID : 536561**

Journal of Emerging Technologies and Innovative Research

An International Open Access Journal Peer-reviewed, Refereed Journal

www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Y.Ambica

In recognition of the publication of the paper entitled

Analyzing customer sentiments on E-Tail applications using NLP

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 11 Issue 4 , April-2024 | Date of Publication: 2024-04-12

Parin P

EDITOR

EDITOR IN CHIEF

ISSN 2349-5162 Journal of Emerging Technologies and Innovative Research

JETIR2404388 **Research Paper Weblink http://www.jetir.org/view?paper=JETIR2404388** **Registration ID : 536561**