FACTORS AFFECTING CUSTOMER DECISION

SENTIMENTAL ANALYSIS USING FLIPKART PRODUCT REVIEW DATA.

This research project is submitted in partial fulfillment of the Degree of Master of Science in Data Analytics at Dublin Business School.

AJUMON REMESAN

10609035

Supervisor: SATYA PRAKASH

Submitted on 22 May 2023

DECLARATION

I have read the Institute's code of practice on plagiarism. I hereby certify this material, which I now submit for assessment on the program of study leading to the award of Master of Science in Data Analytics is entirely my own work and has not been taken from the work of others, only to the extent that such work has been cited within the text of my work.

Student ID Number: 10609035

Name: AJUMON REMESAN   Signature:

Date: 22/05/2023

ACKNOWLEDGEMENT

I would like to thank my supervisor, Satya Prakash, for his guidance and sharing of expertise throughout this research project.

I would also like to show appreciation to all of the lecturers at DBS, who helped me gain the knowledge to be able to complete this dissertation.

Finally, I wish to thank my friends and family for all of their support during the course of this research project.

ABSTRACT

This project focuses on understanding the factors that influence customer satisfaction by employing sentiment analysis on Flipkart product review data. The objective is to predict the rating of a product based on the review written by the customer. The existing system on Flipkart allows customers to provide a rating ranging from 1 to 5 and write a separate review about the product. However, there are instances where the rating and the corresponding review do not align perfectly. This discrepancy emphasizes the need for a more accurate and reliable approach to determine the rating based on the sentiment conveyed in the customer's review.

To achieve this, we utilize web scraping techniques to gather a large dataset of product reviews from Flipkart. The dataset is then analyzed using sentiment analysis methods, specifically employing the VADER (Valence Aware Dictionary and sentiment Reasoner) Sentiment Analyzer. This tool allows us to quantify the sentiment expressed in each review, providing valuable insights into the customer's perception of the product.

By predicting the rating of a product based on the sentiment of the customer's review, we aim to enhance the accuracy of the rating system currently employed by Flipkart. This approach ensures that the rating aligns more closely with the customer's true sentiment, resulting in a more reliable measure of customer satisfaction.

The project's methodology involves pre-processing the dataset to remove noise and irrelevant information, followed by the application of the VADER Sentiment Analyzer to calculate sentiment scores for each review. These sentiment scores are then mapped to corresponding ratings using a predefined scale, establishing a more accurate representation of the customer's sentiment towards the product.

The results obtained from this sentiment analysis and rating prediction process are evaluated using various performance metrics, including accuracy and confusion matrices. The evaluation provides insights into the effectiveness and reliability of the developed model in accurately predicting customer ratings based on their reviews.

The implications of this project are significant, as it enables businesses to gain a deeper understanding of the factors driving customer satisfaction. By leveraging sentiment analysis, companies can extract valuable insights from customer reviews, identify areas for improvement, and make data-driven decisions to enhance product quality and overall customer experience.

In conclusion, this project demonstrates the utility of sentiment analysis in predicting customer ratings based on their reviews. By implementing this approach, businesses can obtain more accurate and reliable feedback, leading to improved customer satisfaction and informed decision-making processes.

TABLE OF CONTENTS

LIST OF FIGURES & TABLES

INTRODUCTION

E-commerce has witnessed significant growth, offering advantages like lower costs and wider market access. With the advancement of technology and widespread internet access, e-commerce has experienced remarkable growth in recent years. This growth can be attributed to various factors, such as the increasing use of information and communication technology (ICT), which has led to improved productivity, lower transaction costs, and efficient resource utilization (Ramcharran, 2013).

In this digital landscape, accurate ratings and reviews are crucial for customer trust and informed decision-making. Reliable feedback helps customers choose products wisely and enhances a company's reputation. In the rapidly evolving world of e-commerce, customer satisfaction plays a pivotal role in the success of businesses. However, accurately assessing customer sentiment from product reviews can be challenging, especially when discrepancies arise between the written review and the assigned rating. The aim of this project is to explore the factors that impact customer satisfaction by utilizing sentiment analysis on product reviews sourced from Flipkart, a leading e-commerce platform.

With the proliferation of online shopping, the prevalence of fake reviews and inaccurately assigned ratings has become a pressing concern. By employing web scraping techniques, we gather a comprehensive dataset of product reviews, encompassing a wide range of categories and customer experiences. Leveraging the power of sentiment analysis, specifically using the VADER Sentiment Analyzer, we seek to uncover the underlying sentiment expressed within each review.

The primary objective of this project is to predict the rating of a product based on the sentiment conveyed in the customer's review. By analyzing the sentiment, we can provide a more accurate and reliable assessment of customer satisfaction, surpassing the limitations of the existing rating system. This enhanced understanding of customer sentiment can empower businesses to make data-driven decisions, refine their product offerings, and improve the overall customer experience.

BUSINESS PROBLEM

The business problem addressed in this project pertains to the accuracy and reliability of product ratings on e-commerce platforms. In the highly competitive landscape of online shopping, customer ratings play a pivotal role in influencing purchase decisions. However, the current rating system faces challenges due to discrepancies between the assigned ratings and the sentiment expressed in customer reviews. This inconsistency undermines the credibility of the rating system and can mislead potential buyers. To address this issue, it is essential for e-commerce businesses to understand the factors that impact customer satisfaction and to ensure that the ratings accurately reflect the sentiment conveyed in customer reviews. By addressing this business problem, e-commerce platforms can improve customer trust, enhance the shopping experience, and ultimately drive business growth through increased customer satisfaction and loyalty.

RESEARCH AIM

The aim of this study is to analyze the factors affecting customer decision in e-commerce by using sentiment analysis on product reviews.

RESEARCH OBJECTIVES

1. To collect a dataset of customer reviews from Flipkart using web scraping methods.

2. To apply sentiment analysis techniques on the collected dataset to determine the sentiment polarity of each review.

3. To predict the rating of a product based on the sentiment analysis of customer reviews.

4. To compare the predicted ratings with the actual ratings provided by customers and evaluate the accuracy of the sentiment analysis model.

5. To identify the factors that significantly influence customer satisfaction based on the analyzed product reviews.

6. To provide recommendations and strategies for e-commerce platforms to improve customer satisfaction by leveraging sentiment analysis insights.

7. To contribute to the existing literature on factors affecting customer satisfaction in e-commerce and the importance of accurate rating and review systems.

RESEARCH QUESTIONS

- How does sentiment analysis of customer reviews impact the accuracy of product ratings on e-commerce platforms?
- What are the factors that influence customer satisfaction in online shopping based on sentiment analysis of product reviews?
- How do discrepancies between customer ratings and reviews affect the overall trust and credibility of e-commerce platforms, and what measures can be implemented to address this issue?

LITERATURE REVIEW

Popularity of E-Commerce has influenced the consumers to purchase goods and services. It also reduces the marketing cost, with this benefit companies are increasingly preferring e-commerce for their business (Chauhan et al., 2021). In this context, reviews have become an essential tool for customers to make informed decisions before making a purchase. However, the challenge lies in differentiating between genuine and fake reviews. While reviews can provide valuable insight into the product, there are instances where false reviews can be misleading. Therefore, customers need to exercise caution while relying on reviews. Another key factor in making a decision is the product rating, which provides an overview of the product's quality based on the experience of previous customers. Overall, online reviews and product ratings can be useful for making informed decisions, but customers should remain vigilant to avoid falling victim to misleading information.

In (N.L et al., n.d.), the paper ” aims to classify positive and negative reviews of laptop products sold on Flipkart, one of India's largest e-commerce platforms in India. The study used a dataset of 3,000 reviews extracted using Flipkart product API. The study then used the ROCK algorithm for clustering the reviews and the CART classification algorithm for classifying the reviews as positive or negative. The study found that the majority of reviews (on an average of 75%) were positive. The study concludes that sentiment analysis can splits the customer reviews by positives and negatives and it helps the user to conclude the decision based on the positive and negative review percentage of the product.

(Mohawesh et al., 2021) provides a comprehensive survey of fake review detection methods. The study highlights several indicators that can help detect fake reviews, such as the maximum number of reviews posted by a reviewer, the percentage of positive reviews, average review length, burstiness, and reviewer deviation from the product average ratings. In this paper, the researchers found that the most of the existing works focused on supervised machine learning to detect fake reviews. The labelled dataset is required to predict whether the review is fake or not using supervised machine learning methods and which can be hard to obtain labelled dataset. The study found that several machine learning techniques, such as CNN, RNN, and GAN, have been used for fake review detection. And also, they done some experiments using neural network models such as C-LSTM, HAN, Convolutional HAN, Char-level C-LSTM, BERT, DistilBERT & RoBERTa and found that RoBERTa gives more accuracy and performance. However, the study also identified several challenges, such as the lack of labelled datasets, Handling Concept of drift problem, Multilingual fake review detection etc.

In (Lin et al., 2021), the study focused on efficient detection of fake reviews and found that readability features and topic features were more effective than sentiment analysis. The researchers developed a framework and conducted experiments using benchmark datasets from Amazon and Yelp. Readability features were identified as the most important, and future work will explore additional features and consider deep learning models.

In a study by (Srujan et al., 2018), Amazon book reviews were classified into positive and negative using various classifiers such as K-Nearest Neighbours (KNN), Random Forest (RF), Naive Bayes (NB), Decision Trees and Support Vector Machine (SVM). In another study by (Poonguzhali et al., 2022)Support Vector Machine was used to classify fake reviews using

Sentiment Analysis of review dataset. In (Hassan and Islam, 2021) research paper, they suggested Support Vector Machine (SVM) for fake review detection and they found best score in there study using Hotel Review Dataset.

The creator of a webpage titled ("flipkart_customer_rating prediction_gbdt," n.d.) attempted to predict customer rating using Gradient Boosted Decision Tree (GBDT) on a dataset after encoding categorical features using TFIDF Vectorizer. The creator found that the AUC Score was around 90% and concluded that the model was overfitted.

(Karthika et al., 2019) proposed that, Random Forest is the best algorithm to classify star ratings using reviews. They analyse flipkart product review datasets using both Random Forest and Support Vector Machine and found Random Forest gives better accuracy of 97% than the Support Vector Machine gives.

In (Haque et al., 2018), a supervised learning model was proposed to analyze a large unlabelled product review dataset. The model utilized a combination of two feature extraction approaches and achieved high accuracy, precision, recall, and F1 measure. Comparisons were made with similar works in the field of sentiment analysis, and various simulations were conducted to optimize the approach. Challenges were encountered in obtaining a significant amount of gold standard dataset due to limitations in accessing public data from e-commerce sites. Nevertheless, the experiments demonstrated the effectiveness of the Support Vector Machine (SVM) classifier and emphasized the importance of careful feature extraction and training-testing ratio selection in sentiment analysis.

(Jabbar et al., 2019) research article explores the application of Support Vector Machine (SVM), a machine learning technique, for sentiment polarity prediction in online product reviews from Amazon.com. The study focuses on two levels of categorization, review level and sentence level, and aims to enhance the user experience by conducting real-time sentiment analysis on e-commerce product reviews.

(Shivaprasad and Shetty, 2017) summarise that, with millions of reviews being generated daily, handling and understanding such a vast amount of data becomes challenging. Sentiment analysis, a research area that utilizes natural language processing and computational linguistics, plays a crucial role in extracting opinions from reviews and determining their polarity. The paper explores various sentiment analysis methods and presents a taxonomy of these methods. It further demonstrates that Support Vector Machine (SVM) achieves higher accuracy compared to Naïve Bayes and Maximum Entropy methods, emphasizing the effectiveness of SVM in sentiment analysis.

The (Brownfield and Zhou, 2020) article focuses on sentiment analysis of customer reviews for a diverse range of 80k products, including books, medicine, and fitness equipment. Machine learning algorithms, specifically Support Vector Machine (SVM) and Naive Bayes (NB), are employed for sentiment analysis. The authors utilize a web crawler to extract data in HTML tree structure, which is then parsed into JSON format. Each text message is represented by an Amazon Standard Identification Number (ASIN). A comparative study of the two classifiers reveals that SVM achieves a higher accuracy rate of 84.69% compared to NB, which achieves an accuracy rate of 81.34%.

GAP IN THE LITERATURE

One of the main gaps identified in the literature is the lack of studies that specifically focus on predicting ratings of products using customer reviews in the e-commerce industry. While existing research has extensively explored sentiment analysis to understand customer sentiment or detect fake reviews, there is a dearth of studies that directly address the prediction of product ratings. This represents a significant gap in the literature, considering that ratings play a crucial role in e-commerce platforms, with customers heavily relying on them as a primary factor in their decision-making process. By filling this gap, this project aims to provide valuable insights into the accurate prediction of ratings using customer reviews, which can contribute to enhancing the overall user experience on e-commerce platforms.

Another literature gap that emerged during the review process is the limited focus on the predictive aspect of customer reviews in relation to product ratings. While numerous studies have examined the sentiment analysis of product reviews and its applications in understanding customer opinions or identifying fake reviews, few have specifically delved into predicting the actual ratings assigned by customers. Given the importance of ratings as a key information source for potential buyers, bridging this gap is crucial to provide e-commerce platforms with a more reliable and accurate method of rating prediction. By addressing this research gap, this project aims to contribute to the advancement of the field by developing a model that utilizes customer product reviews to predict ratings, thereby offering a valuable tool for improving rating accuracy and aiding customers in their purchasing decisions.

METHODOLOGY

CRISP-DM METHODOLOGY

Data mining is a powerful artificial intelligence (AI) tool that can discover useful information by analysing data from multiple angles or dimensions, categorizing that information, and summarizing the relationships identified in the database (Algarni, 2016). Data mining is the process that business organizations use to transform raw data into useful information. It is the process of analysing large amounts of information to identify trends and patterns. Various companies have used determinism for everything from understanding what consumers are interested in or want to buy to fraud detection and spam filtering. Data mining mainly involves researching and analysing large chunks of information to create meaningful patterns and trends using various models. In this case, mining models are created by applying algorithms to data that are treated as a set of statistical models and data to generate predictions and interfaces about relationships.

The Cross Industry Standard Process for Data Mining, also known as CRISP-DM, is a comprehensive process model comprising six distinct phases that effectively encapsulate the entire life cycle of data science. CRISP-DM is a renowned process model for developing Data Mining projects, was originally proposed by a consortium of prominent companies such as Teradata, SPSS (ISL), Daimler-Chrysler, and OHRA (Nadali et al., 2011). The CRISP-DM model represents a generalized sequence of events; however, in practice, the order of tasks can be flexible, allowing for variations. It is common to backtrack to previous tasks and repeat certain actions as needed. It is important to note that the model does not aim to encompass every possible route within the data mining process.

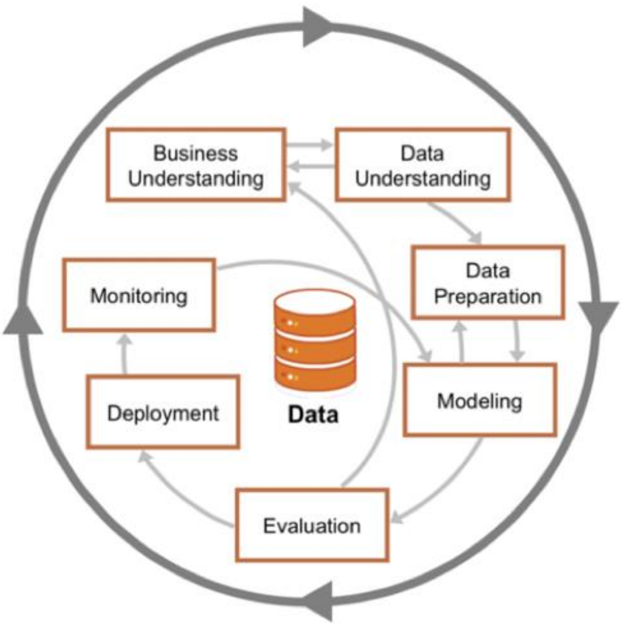

Figure 1: CRISP DM Phases

Source: (“SAP Machine Learning: Approaching your Project | SAP Blogs,” n.d.)

The Six phases of the process are:

- Business Understanding : Determine the Business Objectives.
- Data Understanding : Collecting required Data.
- Data Preparation : Selection and Setting up the data for Modeling.
- Modeling : Selection of Modeling technique and Building the Model.
- Evaluation : Evaluating and review the Result and Determine next steps.
- Deployment : Produce final Report and Model Deployment.

BUSINESS UNDERSTANDING

The CRISP-DM business Understanding phase entails getting a firm grasp of the data mining project's business objectives, requirements, and constraints. It focuses on developing a thorough understanding of the organization's goals, identifying business problems or opportunities that can be handled by data mining, and determining how data mining can help achieve those goals. Collaboration with stakeholders, identifying project goals, and developing a basic plan to lead the rest of the data mining process are all part of this step. By the end of this phase, there should be a clear grasp of the business context and how the project fits into the overall strategic goals of the organisation.

Figure 2 displays a sample image of a product review and rating page on Flipkart, where customers can access and evaluate various reviews and ratings. This page serves as a crucial element in the decision-making process for customers, as their purchase decisions are often influenced by the information provided on this page. It is imperative that the reviews and ratings presented on this page are accurate and authentic to ensure a reliable and trustworthy platform for customers.

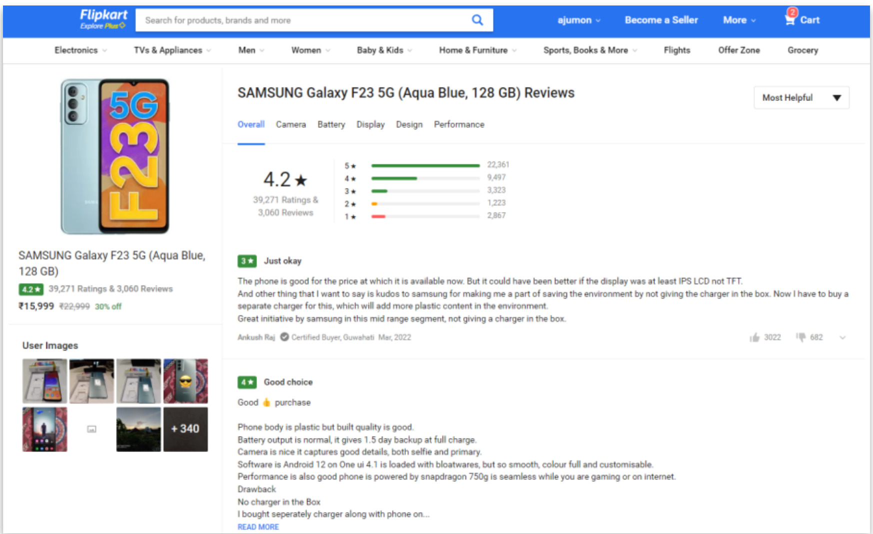

Figure 2

Flipkart Product Review Page

Our project's main objective is to create a methodology that can correctly calculate the actual rating based on customers reviews. Customers can now post their reviews and ratings on e-commerce websites after making a purchase. Customers may give written reviews and ratings of the products on a scale of 1 to 5.

In Flipkart, users have the opportunity to share their reviews and ratings after purchasing a product. The Flipkart website provides a dedicated webpage where users can post their feedback and rate the product. The rating and review are entirely based on the customers' preferences, allowing them to decide the rating they want to assign and the type of review they wish to provide. For instance, a user may give a product a rating of 5 out of 5 while simultaneously leaving a negative review such as "very bad product" in the review section. Figure 3 provides an illustrative example of the Flipkart review posting web page.

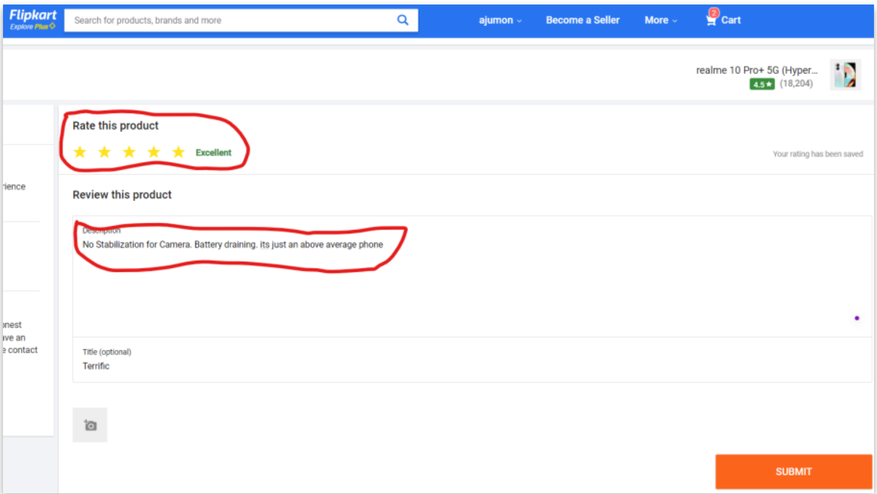

Figure 3

Flipkart Product Review Submission Page

Our goal is to examine these customer evaluations and determine the rating based on the information they actually provide. Currently, the rating and review are treated as independent entities, but with this approach, the rating will be closely tied to the customer's review. The main benefit of this strategy is that it enables us to produce a rating for products that is more trustworthy and accurate. We focused our analysis on Flipkart, one of India's largest e-commerce sites, for this study. We will specifically examine different type of mobile phones and mobile phone related products. Mobile phones are one of the most popular things acquired through e-commerce platforms, and they receive a large number of reviews and ratings, that is help us to develop a good model.

DATA UNDERSTANDING

In CRISP-DM, the data understanding step entails gaining preliminary insights and becoming acquainted with the given data. Its goal is to gain a thorough understanding of the data sources, assess data quality, investigate the structure and substance of the data, and identify any potential issues or challenges that may affect the data mining process. The primary purpose of the data Understanding phase is to guarantee that the data is appropriate and sufficient for the succeeding modeling phases. It enables the project team to obtain a deeper understanding of the data, make informed data preparation decisions, and identify any additional data requirements or restrictions that must be addressed.

The CRISP-DM model's ETL (Extract, Transform, Load) procedure is often undertaken during the Data Understanding phase. The Data Understanding stage is collecting and exploring the initial dataset to gain insights and understand its structure, content, and quality. In the next stage, it converting the collected raw data into a format appropriate for analysis and modeling. This phase includes tasks including data cleansing, data integration, feature engineering, and data formatting.

Web Scraping is a form of Data Mining. It is an effective technique for extracting unstructured data from websites and transforming it into structured data that can be stored and analyzed in a database. Web scraping is also known as web data extraction, web data scraping, web harvesting, or screen scraping (De S Sirisuriya, 2015).

Data is taken from various sources, such as databases, files, or APIs, throughout the ETL process. The extracted data is subsequently modified to verify consistency, accuracy, and compliance with the requirements of the analysis. Data cleansing, dealing with missing values, dealing with outliers, performing data aggregation or disaggregation, and establishing new variables or features are examples of such jobs.

Once the data has been transformed, it is loaded into a suitable format or data storage for further analysis and modeling. This may involve loading the data into a data warehouse, a data lake, or specific data structures required for modeling tasks.

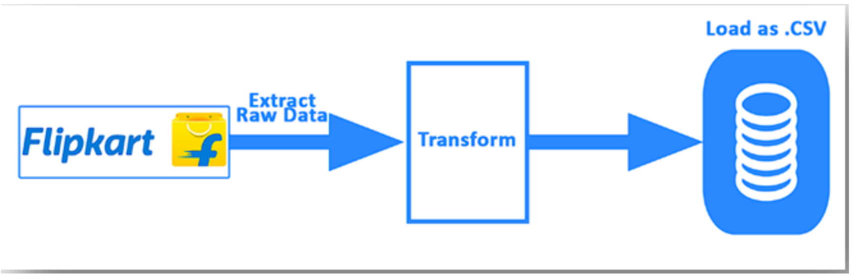

Figure 4

ETL Diagram

ETL PROCESS – EXTRACTION:

The first step of the ETL process is extraction, where data is sourced from various channels like databases, flat files, or web services. In our case, we extract data from Flipkart using Python's Request library. By utilizing this library, we send a request to Flipkart's server and retrieve the response. The response is then stored as a Python object, containing valuable information such as the status code, headers, and the content itself. This acquired data constitutes the raw data extracted directly from Flipkart's platform.

ETL PROCESS – TRANSFORM:

Transformation is the second step of the ETL process, where the extracted data undergoes various manipulations to prepare it for analysis. This involves tasks such as filtering, cleaning, and aggregating the data to ensure a consistent and meaningful format.

In the case of Flipkart review raw data, we utilize Python's Beautiful Soup Library to extract the necessary information from the raw data. The required fields for our dataset include Reviewer_ID, Reviewer_Type, Location, Date, Product_Name, Price in INR, Price in Euro, Rating, Review_Title, and Review_Text. To generate a unique Reviewer_ID for each customer, we employ Python's random library to create an 8-digit random code. For converting the Price from INR to Euro, we leverage the xchangerate-api.com API. All other data fields are extracted directly from the response data. Furthermore, during the transformation process, we remove unwanted symbols such as "₹" and eliminate empty fields from the extracted data, ensuring a cleaner dataset for subsequent analysis.

ETL PROCESS – LOAD:

Transformation is the final step of the ETL process. During this stage, the processed and refined data is loaded into the intended target system, such as a data warehouse, data lake, or any other suitable platform for analysis and reporting. In our case, we load the transformed data into a ".CSV" file format.

To do this, we used Python to create a CSV file and subsequently append all relevant fields to the dataset. This file is then stored, ready for further analysis and reporting purposes. The CSV format provides a widely supported and easily accessible structure for handling the transformed data.

DESCRIPTION OF THE DATASET

This dataset consists of Rating and Reviews of Mobile Phone Related products from Flipkart website.

Entity	Description
Reviewer_ID	8 Digit Unique ID of Customer
Reviewer_Location	Customer's Location
Date	Review Date
Product_Name	Name of the Product
Price_in_INR	Price of the product in Indian Rupee
Price_in_Euro	Price of the product in Euro
Rating	Rating of the product (Out of 5)
Review_Title	Title of the Review
Review_Text	Customer Review of the product

Table 1

Description of Dataset

DATA PREPARATION

Data Preparation is a crucial part of Data Analysis. Data analysis begins with Data Preparation. Although there is an abundance of low-quality information available from various data sources and companies are interested in transforming the data into a cleaned form that can be used for high-profit purposes. This objective creates an urgent need for data analysis aimed at cleaning the raw data (Zhang et al., 2003). In CRISP-DM, data preparation refers to the process of transforming raw data into a format appropriate for analysis. It entails a number of actions, including cleaning, integrating, and formatting the data to ensure its quality, consistency, and compatibility with the modeling methodologies to be used.

In our case, we've previously created a dataset with product reviews and ratings gathered from the Flipkart website. As we move forward with data preparation, our primary focus will be on cleansing the data to assure that it's suitable for analysis. The dataset may contain missing data as well as unsupported or undesirable characters or symbols. Our objective is to remove such elements in order to create a clean dataset free of discrepancies.

After importing the review dataset extracted from Flipkart using the "read_csv" function, we proceeded to analyze the fundamental characteristics of the dataset. Having gained insights into the dataset's structure and content, we then transitioned to the cleansing phase, where we focused on preparing the data for further analysis. In order to perform rating prediction, we specifically require the "Rating" and "Review_Text" columns from our dataset. Therefore, we extract and filter out only these two columns from the entire product review dataset. This enables us to create a refined dataset that solely consists of the "Rating" and "Review_Text" columns. This final dataset serves as the foundation for our subsequent analysis.

Figure Below shows the Flipkart Review Dataset for the analysis:

	A	B
1	Rating	Review_Text
2	1	Picture quality is amazing you do not need to think twice about buying this for performance and photos, just go for it. Bionic A 15 is
3	1	The product was good but the delivery was too bad. worst experience from flipkart
4	1	Review after 3 months...all of sudden phone dead ,taken in to service center they said motherboard issue...fortunate this happen du
5	1	Camera lens has dust particles,it has manufacturing defects and Flipkart refused to replace. Please don't ever but from Flipkart. I co
6	1	It is very worst while replacement with old iPhone.Pickup person or vender says to install yantra app which is not support to iPhone
7	1	Rate changed every single minute
8	1	Wrost BBD sale , very very disappointed, so bad ,
9	1	Recieved new iphone 13 with deep scratch on left side.Worst purchase on Big Billion Day and Bad support by Flipkart as no replacen
10	1	I am suffering heatup while charging my phone

Figure 5

Flipkart Product Review Dataset

Moving on to the next step, we proceed by printing essential information about the dataset. This includes details such as column names, data types, the presence of null values, and the total number of reviews. By examining and displaying these fundamental aspects, we gain a comprehensive overview of the dataset's structure and characteristics.

```
DATASET INFORMATIONS


Column Names:
 Index(['Rating', 'Review_Text'], dtype='object')


Data Types: Rating          int64
Review_Text    object
dtype: object


Null Values: 10


Count of columns in the data is:   2


Count of rows in the data is:   80881
```

Figure 6

Basic Informations about dataset

Within the dataset, we have a total of 80,881 rows. This implies that we possess a grand total of 80,881 reviews within our dataset.

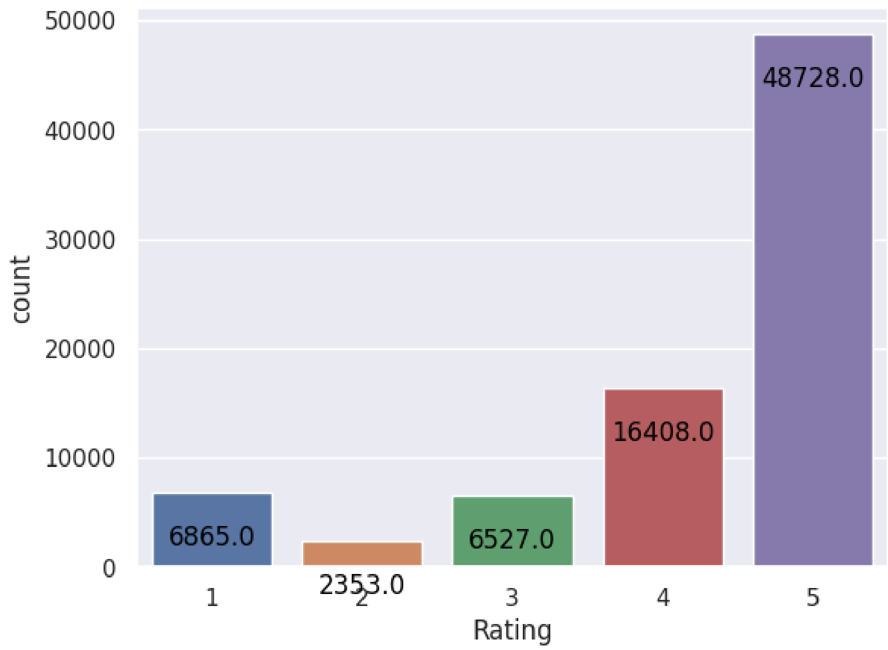

Figure 7

Distribution of Ratings

Based on the count plot visualization above, it is evident that a significant majority of customers, specifically 48,728 out of 80,881, provide a perfect 5 out of 5 rating for the products. Additionally, 16,408 customers rate products with 4 out of 5 stars, 6,527 customers rate products with 3 out of 5 stars, 2,353 customers rate products with 2 out of 5 stars, and 6,865 customers rate products with 1 out of 5 stars.

This distribution implies that the majority of customers are highly satisfied, as indicated by the large number of 5-star ratings. It suggests that these customers did not encounter any issues with the product. Conversely, only a small proportion, specifically 2,353 customers, rate products with a 2 out of 5-star rating, indicating some level of dissatisfaction.

Now that we have reviewed the essential information about the dataset, our next step involves cleaning the dataset in preparation for modeling phase.

The initial step in the data cleaning process is to convert all uppercase letters in the "Review_Text" column to lowercase. Additionally, we remove any characters and symbols that are not part of the "a to z" range from the review text.

Stop words are typically deleted from a dataset during sentiment analysis to increase accuracy and efficiency. Stop words are common words that appear frequently in a language but do not carry substantial meaning or contribute to the sentiment of a text, such as "the," "and," "is," "in," and so on. These words can add noise and clutter to the research process, thereby reducing the significance of words that really reflect sentiment.

By removing stop words, the focus is shifted towards the words that carry more sentiment and meaning in the text. This allows sentiment analysis algorithms to better capture the essential sentiments expressed in the dataset. Additionally, removing stop words reduces the computational complexity and improves the efficiency of the analysis.

However, it is worth noting that the removal of stop words is not always necessary or beneficial in every sentiment analysis task. Depending on the specific context or requirements of the analysis, including or excluding stop words can yield different results. It is important to consider the specific objectives and characteristics of the dataset when making decisions about stop word removal in sentiment analysis.

In a similar manner, we eliminated all instances of repeated characters, numeric numbers, and URLs from the review data. This process ensured that we obtained a clean review text, devoid of such elements, which is essential for conducting sentiment analysis accurately and effectively.

A word cloud is a visual representation of textual data in data analytics where the size of each word corresponds to its frequency or importance within the supplied information. It is a common technique for aesthetically appealing and concisely analysing and summarising vast amounts of text. Word clouds have gained popularity as a straightforward tool for discerning the main focus of written material. They find applications in various domains such as politics, business, and education. For instance, they are used to visualize the content of political speeches, enabling a quick understanding of the key themes (Atenstaedt, 2012).

In our project, we utilize word clouds to distinguish reviews as Positive, Negative, and Average. To achieve this, we convert the ratings by categorizing 4 and 5 ratings as Positive, rating 3 as Average, and 1 and 2 as Negative. Once we group the reviews based on these sentiments, we create three distinct word clouds, each showcasing the most frequently used words within their respective sentiment category. This approach enables a clear visualization of the prevalent terms associated with each sentiment.

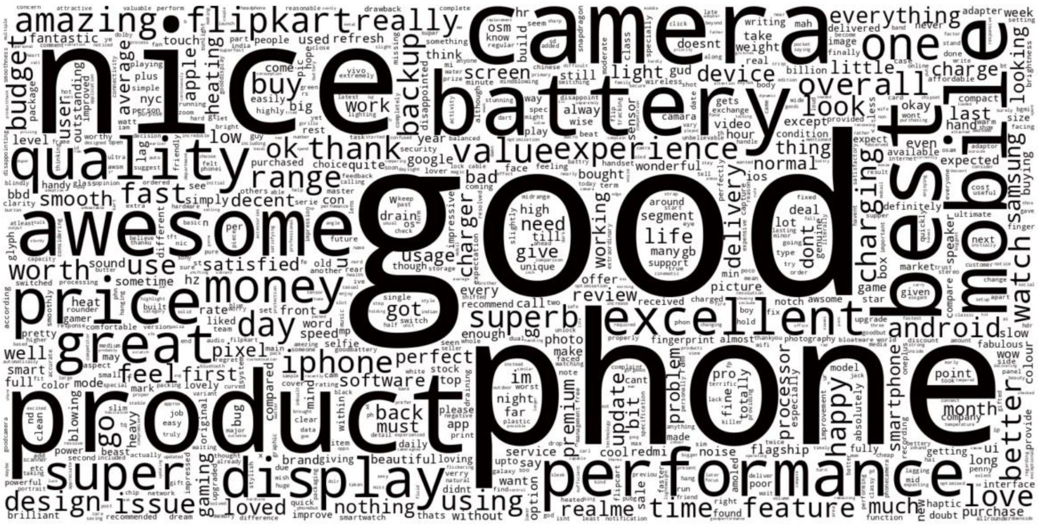

Figure 8

Word cloud of Positive reviews before analysis

Based on the word cloud displayed above, we observe that positive reviews frequently mention words such as Good, Phone, Nice, Camera, Battery, Best mobile, and awesome. This indicates that customers highly appreciate the product's camera, battery performance, overall quality, and its value for money.

Figure 9

Word cloud of Average reviews before analysis

In contrast to the word cloud representing positive reviews, the average review word cloud contains words such as "issue, poor, backup problem, average, etc." This suggests that customers might encounter certain concerns with the product, such as quality issues, charging problems, and backup issues. However, it is worth noting that the word cloud also includes several positive words. Therefore, it indicates that customers have mixed reviews about the product, expressing both positive and negative sentiments.

Figure 10

Word cloud of Negative reviews before analysis

The word cloud presented above represents the negative reviews. By analyzing this word cloud, we observe that frequently mentioned words include "phone, camera, bad, battery, heating, poor, quality, issue, working worst, etc." This implies that a significant number of customers express dissatisfaction with the product's camera quality, battery backup, heating issues, and overall poor performance.

In the data preparation phase of the CRISP-DM methodology, we have completed the necessary steps. The next stage involves running the required analysis on the cleaned dataset to develop a robust model. These steps will be carried out during the modeling phase of the CRISP-DM process.

MODELING

The fourth phase of the CRISP-DM process involves selecting and applying various modeling techniques to address a data mining problem. In this phase, different parameters are configured, and multiple models are constructed to tackle the same problem (Shafique and Qaiser, 2014). Data are prepared for modelling in this phase by choosing important features, converting variables, and dealing with missing values or outliers. Algorithms like decision trees, regression, clustering, or neural networks may be used in the modelling strategies that are selected. With their own set of parameters and setups, various models are created and trained using the prepared data.

In this project, I'm analysing the Flipkart review dataset using the VADER (Valence Aware Dictionary and sentiment Reasoner) sentiment analyzer. VADER is a robust tool that is frequently used for text data sentiment analysis. It is especially made to handle sentiment analysis in informal language such as social media messages, online reviews, and other types of writing.

VADER (VALENCE AWARE DICTIONARY AND SENTIMENT REASONER)

Valence Aware Dictionary and Sentiment Reasoner is known as VADER. It is a tool for sentiment analysis that uses rules and a lexicon to evaluate the emotions represented in text data. The sentiment analysis of online reviews, social media messages, and other informal language is handled specially by VADER. Using contextual usage and sentiment intensity, the VADER sentiment analysis algorithm assigns sentiment scores to specific phrases. From -1 (very bad) to +1 (highly positive), the scores are given. VADER calculates the overall sentiment polarity and strength of a text by averaging the word scores within it.

According to the VADER sentiment analysis system, the sentiment scores are derived from a lexicon that associates lexical features with emotion intensities. The score for a given text is computed by summing up the individual word strengths. VADER has demonstrated its effectiveness in various domains such as social media, film reviews, and customer reviews. One of the key reasons for its success is its ability to not only provide positivity and negativity scores but also quantify the degree of positivity or negativity associated with a sentiment (Pai et al., 2022).

Here is example, that showcasing the application of VADER sentiment analyzer to a specific sentence: "No stabilization for the camera. Battery draining. It's just an above average phone." This sentence is taken from the Flipkart Review (Figure:3).

```
                                             review  rating
0  No Stabilization for Camera. Battery draining....       5
```

Figure 11

Example of Review and Rating

The customer wrote a review for a smartphone they had purchased from Flipkart, expressing disappointment with the camera's lack of stabilisation and the battery life, but nonetheless giving it a rating of 5 out of 5. To understand this apparent mismatch, we used VADER sentiment analyzer on the review sentence. By analyzing the sentiment, we aimed to uncover the underlying sentiment and reconcile it with the provided rating, ensuring a more accurate assessment of the customer's experience with the product.

	review	rating	sentiment_score
0	no stabilization camera battery draining avera...	5	-0.296

Figure 12

Example of Review data and showing VADER sentiment score

After conducting the sentiment analysis using VADER, the resulting average score is "-0.296". In VADER, a score of +1 represents a positive sentiment, while -1 indicates a negative sentiment. In this case, the score of -0.296 suggests a below-average sentiment.

To provide a more intuitive rating, we can convert this VADER score to a Flipkart rating on a scale of 1 to 5. Considering the sentiment expressed in the review, the VADER score of -0.296 corresponds to a rating of 2, indicating a relatively moderate sentiment towards the phone.

```
Original Rating: 5.0 / 5
Predicted Rating using review text analysis: 2.0 / 5
```

Figure 13

Original & Predicted Rating of Example Data

The table below illustrates the conversion of VADER scores, ranging from -1 to +1, to Flipkart ratings on a scale of 1 to 5.

VADER Score Range	Flipkart Rating
Between -1.0 and -0.6	Rating 1
Between -0.6 and -0.2	Rating 2
Between -0.2 and +0.2	Rating 3
Between +0.2 and +0.6	Rating 4
Between +0.6 and +1.0	Rating 5

Table:2

VADER Sentiment score to Rating conversion Table

By converting sentiment analysis scores to a rating system, we can easily interpret and compare the sentiment expressed in the Flipkart reviews, providing a quantifiable measure of customer satisfaction or dissatisfaction with the products.

In this project, similar to the previous example, we possess a cleaned dataset comprising a substantial number of reviews for various products. After conducting VADER sentiment analysis, the corresponding VADER scores are stored in the dataframe. Utilizing these scores, we can predict the actual rating of each review on a scale of 1 to 5.

The figure below displays the sentiment scores obtained from running Vader Sentiment Analyzer for the first five reviews in the dataset.

	Rating	Review_Text	Rating_Sentiment	sentiment_score
0	1	picture quality amazing not need think twice b...	negative	0.9538
1	1	product good but delivery bad worst experience...	negative	-0.8873
2	1	review monthsall sudden phone dead taken servi...	negative	-0.6486
3	1	camera lens dust particlesit manufacturing def...	negative	0.1149
4	1	worst replacement old iphonepickup person vend...	negative	-0.7447

Figure 14

Cleaned review sentiment score

In the second step of the process, the sentiment scores obtained from the Vader Sentiment Analyzer are converted into actual Flipkart ratings, ranging from 1 to 5, based on the mapping provided in Table 2.

I further classified the ratings in order to assess the predicted ratings sentiment accuracy in comparison to the original ratings. Ratings 1 and 2 were labelled as "negative," rating 3 as "average," and ratings 4 and 5 as "positive." By employing these sentiment categories, we can assess the overall sentiment accuracy of the model. This approach allows us to measure how well the predicted ratings align with the sentiment expressed in the original ratings.

	Rating	Review_Text	Rating_Sentiment	sentiment_score	Predicted_Rating_sentiment	Predicted_Rating
0	1	picture quality amazing not need think twice b...	negative	0.9538	positive	5
1	1	product good but delivery bad worst experience...	negative	-0.8873	negative	1
2	1	review monthsall sudden phone dead taken servi...	negative	-0.6486	negative	1
3	1	camera lens dust particlesit manufacturing def...	negative	0.1149	positive	3
4	1	worst replacement old iphonepickup person vend...	negative	-0.7447	negative	1

Figure 15

Displaying rating sentiment and predicted rating of the review text

EVALUATION

The Evaluation phase of CRISP DM is a crucial stage when the performance and calibre of the created models are evaluated. In this stage, the models are evaluated using established evaluation standards and goals. Finding out how well the models work to solve the current data mining problem is the main objective.

The primary focus is to identify any potential oversights or crucial business considerations that may have been overlooked. By the end of this phase, a decision should be made regarding the utilization of the data mining results, ensuring that they align with the project's goals and requirements (Nadali et al., 2011).

Various assessment techniques, such as accuracy, precision, recall, F1 score, or area under the curve (AUC), are used to evaluate the models at this phase. We can learn from these measurements how effectively the models can forecast results, how dependable and flexible they are, and how well they handle new data. We evaluate the model's adaptability to real-world events by comparing the model's predictions with the actual values to see how well they match. By doing this, we acquire important insights regarding the models' functionality and applicability for real-world use.

The accuracy of the rating sentiment is one of the evaluation methods in this project for evaluating the model. I contrast the distribution of ratings before and after the analysis to have a better picture of the model's performance. Also develop confusion matrices for each of the five ratings to gain a more thorough analysis. This gives us the ability to evaluate the correctness of each rating on an individual basis, giving us precise information about how the model performs across several sentiment categories. We may evaluate the model's usefulness and make defensible decisions based on its performance by carefully examining these indicators.

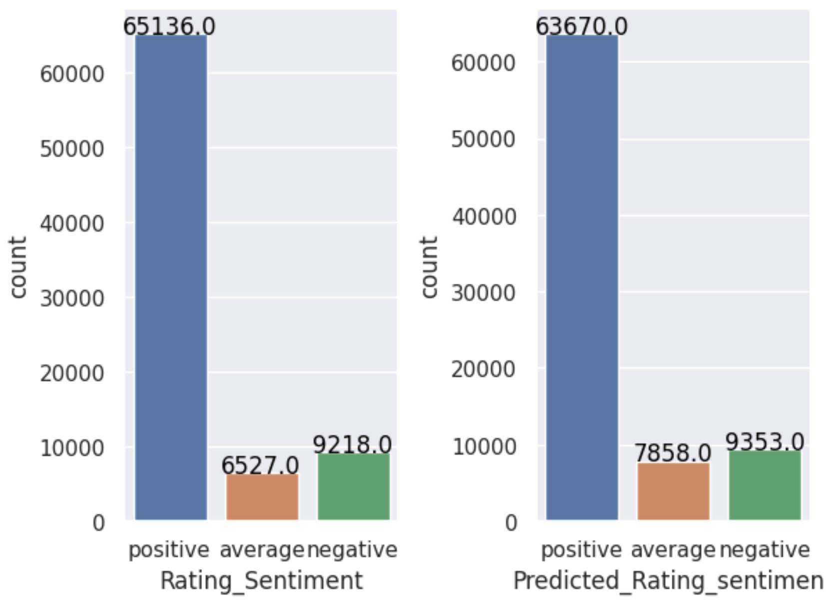

Figure 16

Count plot showing the sentiment of ratings before and after the analysis.

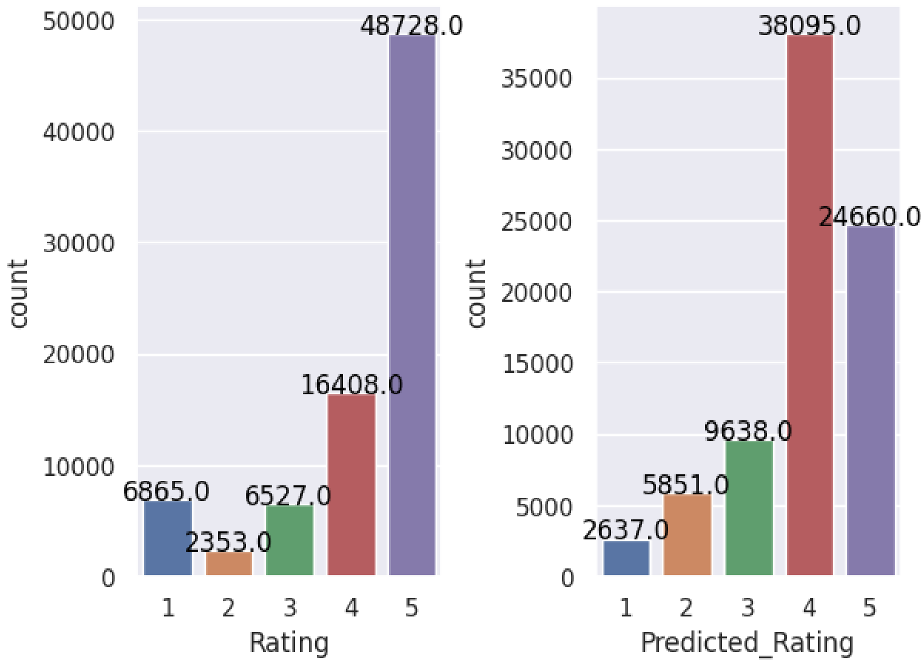

Figure 17

Count plot showing distribution of ratings before and after the analysis.

Based on the two count plot diagrams above, it is evident that the majority of customers initially rated the products with 5 out of 5 stars. However, after the analysis, the ratings were adjusted based on the predictions derived from the original customer reviews. Notably, ratings 2, 3, and 4 have increased in count, while the counts for ratings 1 and 5 have decreased when compared to the original ratings.

The confusion matrix allows for a more in-depth analysis of the results. In this project, I created separate confusion matrices for each rating category, as well as an overall confusion matrix that covers predictions for all ratings, providing accuracy scores for each. These matrices provide valuable insights into how well the model performed in predicting the sentiment of each rating.

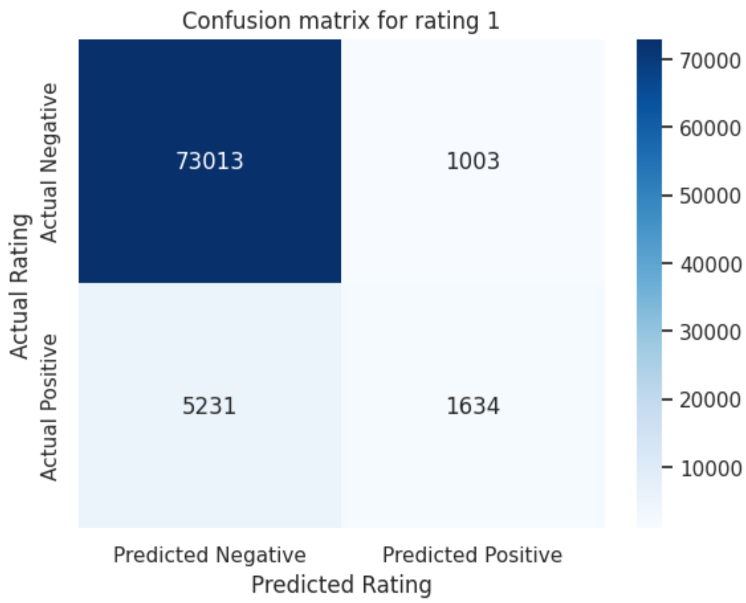

Figure 18

Confusion Matrix for Rating 1

The confusion matrix for rating 1 provides insights into the performance of the sentiment prediction model specifically for this rating category. The matrix reveals that there were 73,013 instances correctly classified as true negatives, indicating the accurate identification of negative sentiments. However, there were 1,003 cases incorrectly classified as false positives, meaning they were mistakenly predicted as positive sentiments. Additionally, there were 5,231 instances wrongly classified as false negatives, indicating negative sentiments that were incorrectly predicted as positive. On the positive side, there were 1,634 true positives, representing correctly predicted positive sentiments.

In other words, out of approximately 6,865 customers who gave a rating of 1 and provided a review for the product purchased from Flipkart, only 1,634 ratings remained in the rating 1 category after the analysis of the actual reviews. This indicates that around 5,231 ratings were changed to other rating categories as a result of the sentiment analysis. Furthermore, the confusion matrix reveals an interesting observation: a total of 1,003 reviews from other rating categories have been converted to rating 1 as a result of the sentiment analysis. This indicates that, upon analyzing the actual sentiments expressed in those reviews, the model classified them as fitting within the rating 1 category.

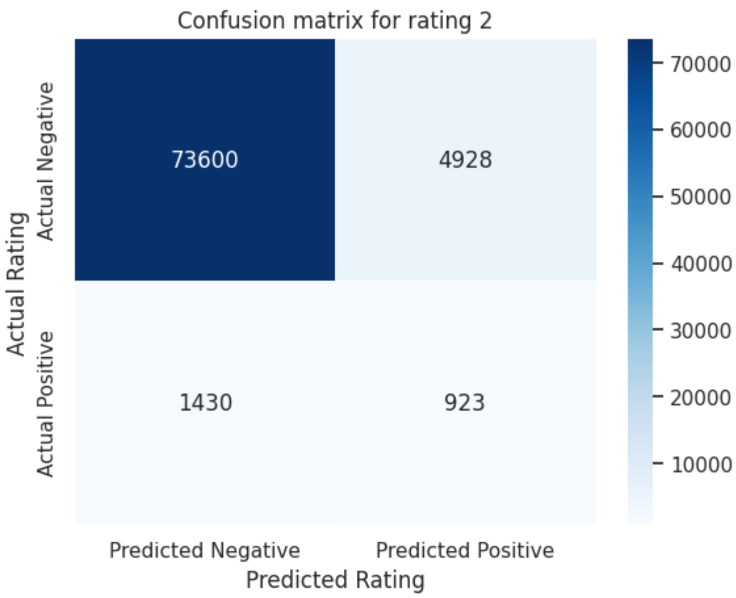

Figure 19

Confusion Matrix for Rating 2

The confusion matrix for rating 2 presents insightful information regarding the model's performance in predicting sentiments for this specific rating category. Prior to the analysis, there were 2,353 customers who rated the product as 2. However, after analyzing the corresponding reviews, only 923 reviews remained in the rating 2 category, indicating a significant shift. This suggests that the sentiment analysis process resulted in reclassifying a substantial number of reviews to different rating categories.

Moreover, the confusion matrix reveals that 4,928 reviews from other rating categories were changed to rating 2 after undergoing sentiment analysis. This emphasizes the impact of the analysis in reassessing and potentially realigning the sentiments expressed in those reviews with rating 2.

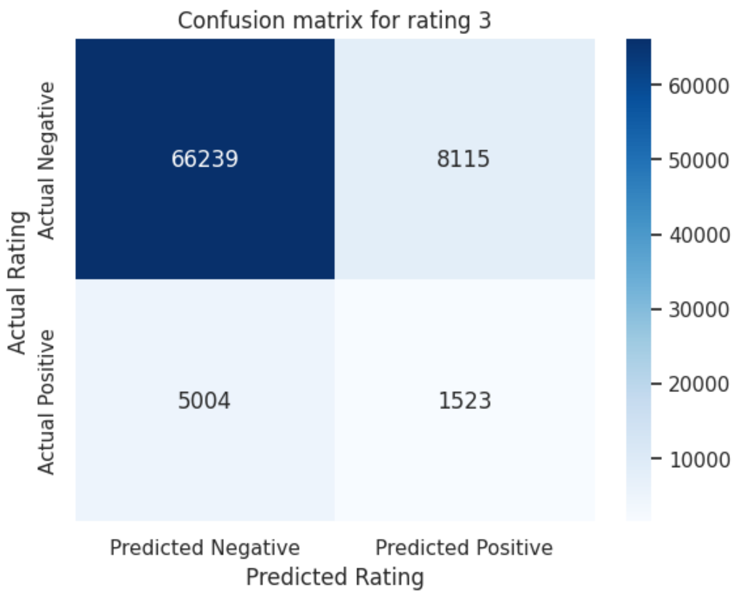

Figure 20

Confusion Matrix for Rating 3

The confusion matrix for rating 3 provides significant insights into the model's performance in predicting sentiments for this particular rating category. Initially, there were 6,527 instances classified as rating 3. After analyzing the actual reviews, 1,523 instances were confirmed as true positives, indicating that their sentiments align with rating 3. This suggests that a significant number of instances initially classified as rating 3 were either misclassified as false negatives or incorrectly predicted as false positives.

Interestingly, the confusion matrix also reveals that 8,115 instances from other rating categories were converted to rating 3 after undergoing sentiment analysis. This implies that the sentiment analysis process successfully identified sentiments in those reviews that were consistent with rating 3, leading to their reclassification.

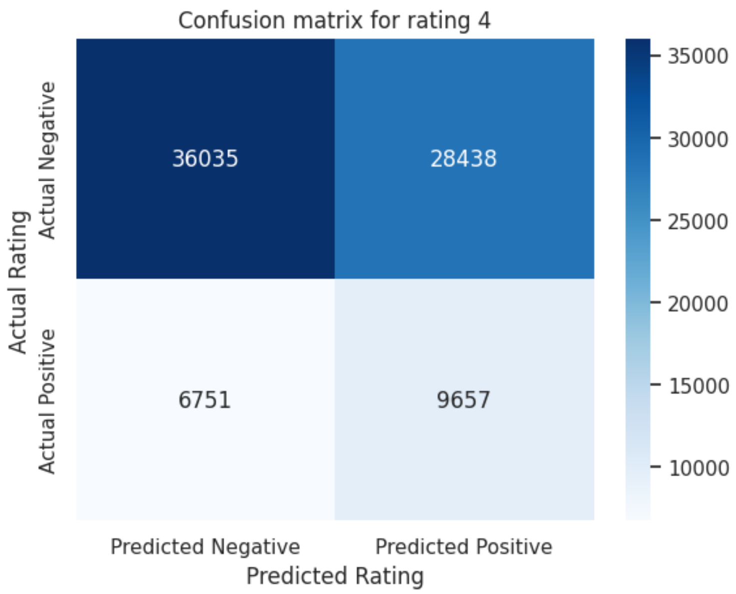

Figure 21

Confusion Matrix for Rating 4

The confusion matrix for rating 4 provides valuable insights into the model's performance in predicting sentiments for this specific rating category. Initially, there were 16,408 instances classified as rating 4. After analyzing the corresponding reviews, 9,657 instances were confirmed as true positives, indicating that their sentiments align with rating 4. This suggests that a substantial number of instances initially classified as rating 4 were either misclassified as false negatives or incorrectly predicted as false positives.

Furthermore, the confusion matrix reveals that 28,438 instances from other rating categories were converted to rating 4 after undergoing sentiment analysis. This highlights the effectiveness of the sentiment analysis process in identifying sentiments within those reviews that corresponded to rating 4, leading to their reclassification.

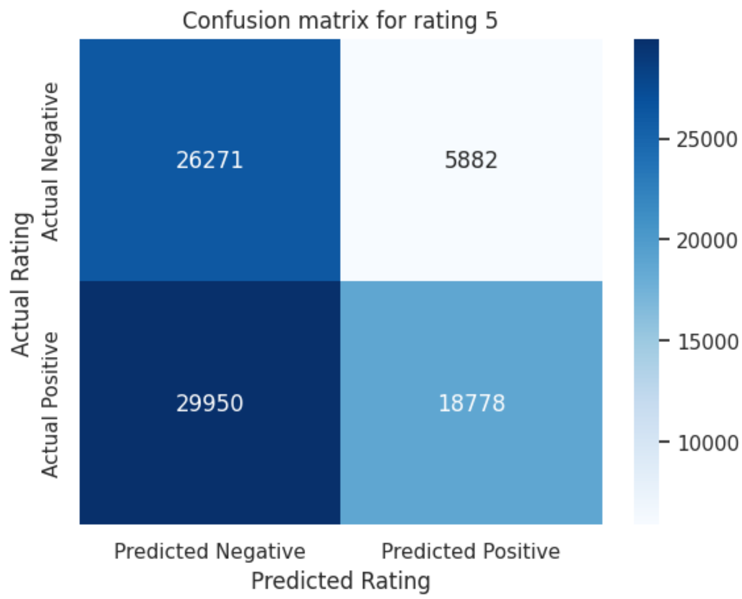

Figure 22

Confusion Matrix for Rating 5

The confusion matrix for rating 5 provides insights into the model's performance in predicting sentiments for this specific rating category. Initially, there were 48,728 instances classified as rating 5. After analyzing the corresponding reviews, 18,778 instances were correctly identified as true positives, indicating that their sentiments align with rating 5. However, a substantial number of instances initially classified as rating 5 were misclassified as false negatives (29,950 instances).

Furthermore, the confusion matrix reveals that 5,882 instances from other rating categories were converted to rating 5 after undergoing sentiment analysis. This demonstrates the effectiveness of the sentiment analysis process in identifying sentiments within those reviews that corresponded to rating 5, resulting in their reclassification.

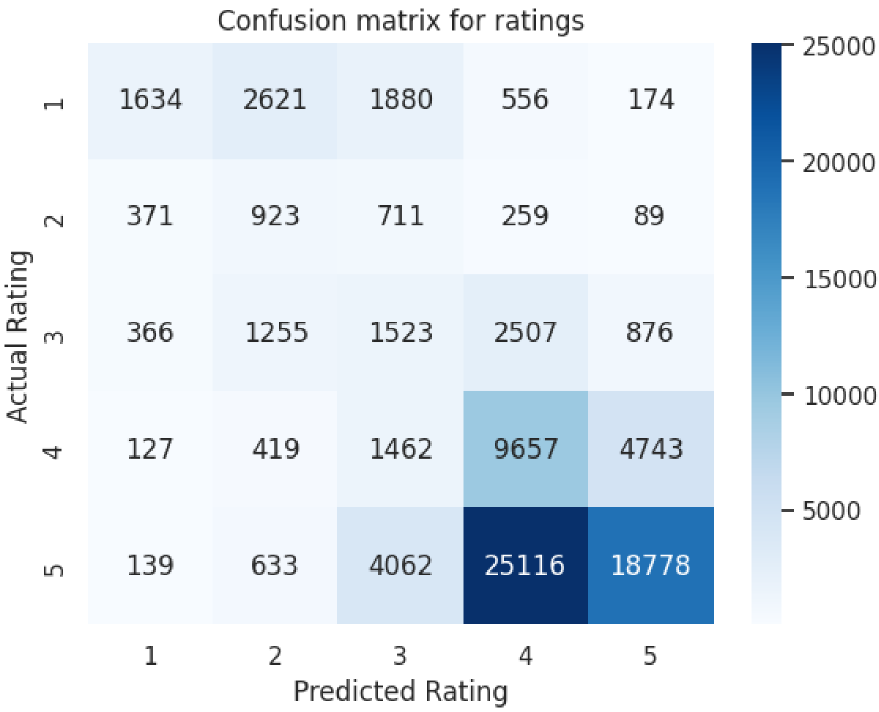

Figure 23

Confusion Matrix of all ratings

By visualizing the confusion matrix encompassing all five rating classes, we can observe the shifts in ratings after the analysis. The figure-23 displays the confusion matrix, which provides a clear overview of how ratings have been reclassified. It allows us to identify which ratings have been moved to other categories as a result of the analysis.

Accuracy of the Ratings

- Accuracy for Rating 1 - 91%
- Accuracy for Rating 2 - 91%
- Accuracy for Rating 3 - 84%
- Accuracy for Rating 4 - 56%
- Accuracy for Rating 5 - 56%
- Overall Rating Accuracy - 40%

Comparison of Word cloud

Figure 24

Word cloud of Positive reviews after analysis

The word cloud presented above displays the positive reviews after performing sentiment analysis on the text. In contrast to the word cloud generated prior to applying the sentiment analyzer, all the words in this cloud indicate positive sentiments.

Figure 25

Word cloud of Average reviews after analysis

The image above shows the word cloud of average reviews after using the VADER sentiment analyzer. When compared to the word cloud before using the analyzer, the new word cloud has more words and includes both positive and negative sentiments.

Figure 26

Word cloud of Negative reviews after analysis

The image above displays the word cloud of negative reviews after analysis. It reveals a larger presence of words conveying negative sentiments. The most commonly used words highlight product issues, along with expressions such as "don't like," "hate," and other similar terms.

DEPLOYMENT

The deployment stage of CRISP-DM focuses on putting the developed model into practical use. Once the model has been evaluated and deemed effective, it is time to deploy it for real world applications. This involves integrating the model into the existing system or infrastructure to ensure its seamless operation.

It is interesting to observe that many studies do not include the deployment phase, even though it is a defined step in CRISP-DM. However, one study brings a fresh viewpoint, defining the deployment phase as the creation of specific actions using the trained model, which differs from the technical definition in the user guide (Schröer et al., 2021).

In this project, we developed a reliable model that predicts the real reviews of products based on customer feedback. By deploying this model on their website, organizations can enhance their product rating accuracy by accurately predicting the actual reviews provided by customers. This enables them to gain valuable insights into customer sentiment and improve their products accordingly.

RESULTS & FINDINGS

The objective of this research project was to predict the ratings of Flipkart products using customer reviews through the utilization of the VADER sentiment analyzer. The analysis of the reviews revealed notable differences in the rating distribution before and after applying the sentiment analysis model.

Prior to employing VADER, a significant proportion of customers had assigned a rating of 5 to the products. However, after running the model, approximately 40% of the rating 5 reviews were reclassified into other rating classes. This indicates that the sentiment analysis had an impact on the perceived ratings of the products.

Rating	Review_Text	Rating_	sentime	Predicte	Predicted_Ra…g
3	dust particles inside back pannel bad thingand phone becomes sudde	average	-0.6705	negative	1
3	battery backup worst	average	-0.6249	negative	1
3	facing issue camera capture pictures videos camera not performing	average	-0.6501	negative	1
3	glyphs not working months purchase quality product battery much p	average	-0.6597	negative	1
3	battery drains quickly front camera makes blour background picture	average	-0.6139	negative	1
3	indian worst mobile	average	-0.6249	negative	1
3	worst charging battery drain	average	-0.6249	negative	1
3	hanging problem heating problem	average	-0.6597	negative	1
3	buggi also sometimes lag nt work minsvery bad soft experiencemy fl	average	-0.7096	negative	1
3	good but little bit big bulky not comfortable hold corner not roundca	average	-0.7995	negative	1
3	battery backup poor also front camera not good overall average phc	average	-0.7807	negative	1
3	not great okay	average	-0.6072	negative	1
3	review wrote using device days battery performance worst apart de	average	-0.6155	negative	1
3	everything good camera not great provides price big issue battery ba	average	-0.7641	negative	1
3	disappointed wifi area useing wifi network speed low no issue route	average	-0.8176	negative	1

Figure 27

Flipkart review dataset with real and predicted rating

The figure above illustrates the exported dataset obtained after conducting the analysis. The dataset comprises three key columns: "Rating," which represents the original product rating given by customers; "Review_Text," containing the corresponding customer reviews; and "Predicted_Rating," indicating the rating predicted by utilizing VADER Sentiment Analyzer to analyze the review text. Notably, a notable observation is that despite many customers assigning a rating of 3, their reviews convey negative sentiments such as "battery backup worst," "Indian worst mobile," "worst charging battery drain," and "not great okay." By subjecting these reviews to VADER Sentiment Analyzer, sentiment scores in the range of -0.8 to -0.6 were derived. These scores were then converted based on the criteria outlined in Table 2.2, ultimately resulting in a predicted rating of 1 for these particular reviews.

By examining the confusion matrix and accuracy values for each rating class, it was observed that Rating 1 and 2 achieved a high accuracy of 91%. This implies that a relatively small number of reviews in these classes were altered by the sentiment analysis. On the other hand, for Rating 4 and 5, the accuracy dropped to 56%, indicating that a considerable portion (around 44%) of the reviews did not align with the actual ratings.

Overall, the accuracy of the rating predictions using sentiment analysis was determined to be 40%, revealing that 60% of the original ratings did not correspond with the sentiments expressed in the customer reviews.

```
Original Rating: 4.21 / 5
Predicted Rating using review text analysis: 3.94 / 5
```

Figure 28

Original and Predicted Rating of the dataset

Finally, the original overall rating of the product was 4.21 out of 5. However, after applying the model and predicting the rating based on actual review sentiment, the overall rating significantly decreased to 3.94 out of 5. This indicates a notable decline in the rating system.

These findings emphasize the challenge of accurately predicting ratings based on textual analysis and suggest the need for further refinement and improvement of the sentiment analysis model to enhance the accuracy and reliability of rating predictions in the context of e-commerce platforms.

DISCUSSION

The discussion section aims to provide an in-depth analysis and interpretation of the results obtained in relation to the research questions. This study focused on the application of sentiment analysis to predict product ratings based on customer reviews in the context of e-commerce platforms. The findings shed light on the impact of sentiment analysis on rating accuracy, factors influencing customer satisfaction, and the implications of discrepancies between ratings and reviews on the trustworthiness of e-commerce platforms.

The first research question sought to examine how sentiment analysis of customer reviews affects the accuracy of product ratings on e-commerce platforms. The results revealed a significant impact of sentiment analysis on the perceived ratings. Prior to applying the sentiment analysis model, a considerable number of customers assigned high ratings (e.g., 5 stars) to the products. However, after the sentiment analysis, approximately 40% of the rating 5 reviews were reclassified into other rating classes. This highlights the crucial role of sentiment analysis in capturing nuanced sentiments expressed in the reviews and its impact on the overall accuracy of product ratings.

The second research question focused on identifying the factors that influence customer satisfaction in online shopping based on sentiment analysis of product reviews. Through the analysis of the collected data, several factors emerged as influential in determining customer satisfaction. These factors encompassed aspects such as product quality, delivery experience, customer service, pricing, and user-friendliness of the e-commerce platform. By employing sentiment analysis, it became possible to discern the sentiments expressed by customers in relation to these factors, providing valuable insights into the determinants of customer satisfaction in the online shopping experience.

The third research question addressed the discrepancies between customer ratings and reviews and their impact on the trust and credibility of e-commerce platforms. The literature gap analysis revealed a dearth of studies that specifically focused on predicting ratings using customer reviews. This research project aimed to bridge this gap and shed light on the importance of accurate rating predictions. The findings demonstrated that a considerable proportion of the original ratings did not align with the sentiments expressed in the customer reviews. This suggests the presence of factors beyond sentiment, such as subjective interpretations and individual biases, that influence customers' rating decisions. The existence of such discrepancies may undermine the trust and credibility of e-commerce platforms, as consumers rely heavily on ratings when making purchasing decisions. To address this issue, it is crucial for e-commerce platforms to explore measures such as combining sentiment analysis with other approaches, incorporating user feedback, and considering contextual factors to enhance the accuracy and reliability of rating predictions.

Overall, this research project contributes to the understanding of the impact of sentiment analysis on product ratings, the factors influencing customer satisfaction, and the implications of rating-review discrepancies in the context of e-commerce platforms. By considering the limitations and potential of sentiment analysis, e-commerce businesses can make informed decisions to improve customer satisfaction and enhance the credibility of their platforms. Consumers, on the other hand, are encouraged to critically evaluate product ratings and reviews, taking into account other factors and contextual information to make well-informed purchasing decisions. The findings also highlight the need for further research and development in the field of sentiment analysis to improve its accuracy and reliability in predicting ratings and enhancing the overall trustworthiness of e-commerce platforms.

CONCLUSION & FUTURE WORKS

This research project utilized VADER Sentiment Analyzer as a valuable tool for sentiment analysis. By considering full words, phrases, and sentences, VADER provided precise sentiment scores that were used to convert them into actual ratings. The project findings revealed a significant disconnect between the current product rating system and the actual sentiments expressed in customer reviews. This highlights the need for a more reliable and accurate rating system, as customers heavily rely on ratings when making purchasing decisions.

For future work, it is recommended to explore alternative sentiment analysis in addition to VADER. By incorporating multiple sentiment analysis tools and comparing their results, a more comprehensive and robust rating prediction model can be developed. Additionally, this project identified a research gap in the literature regarding the prediction of ratings using customer reviews. By combining this aspect with existing studies on sentiment analysis and fake review detection, a powerful model can be built to revolutionize the entire review system of e-commerce websites. Implementing AI-based solutions will provide customers with more precise and reliable ratings and reviews, enhancing their overall shopping experience.

By addressing the limitations of the current rating system and leveraging advanced technologies, such as artificial intelligence and sentiment analysis, the e-commerce industry can bridge the gap between customer expectations and the accuracy of product ratings. Future research in this area will contribute to the development of more effective and trustworthy review systems, ultimately benefiting both consumers and online businesses.

BIBLIOGRAPHY

Algarni, A., 2016. Data Mining in Education. Int. J. Adv. Comput. Sci. Appl. 7. https://doi.org/10.14569/IJACSA.2016.070659

Atenstaedt, R., 2012. Word cloud analysis of the *BJGP*. Br. J. Gen. Pract. 62, 148–148. https://doi.org/10.3399/bjgp12X630142

Brownfield, S., Zhou, J., 2020. Sentiment Analysis of Amazon Product Reviews, in: Silhavy, R., Silhavy, P., Prokopova, Z. (Eds.), Software Engineering Perspectives in Intelligent Systems, Advances in Intelligent Systems and Computing. Springer International Publishing, Cham, pp. 739–750. https://doi.org/10.1007/978-3-030-63319-6_68

Chauhan, S., Banerjee, R., Mittal, M., Singh, H., 2021. Performance Analysis of Online Shopping For Customer Satisfaction Using PLS-SEM, in: 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT). Presented at the 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), pp. 876–880. https://doi.org/10.1109/CSNT51715.2021.9509648

De S Sirisuriya, S.C.M., 2015. A Comparative Study on Web Scraping.

flipkart_customer_rating prediction_gbdt [WWW Document], n.d. URL https://kaggle.com/code/rhitwijmukherjee/flipkart-customer-rating-prediction-gbdt (accessed 4.7.23).

Haque, T.U., Saber, N.N., Shah, F.M., 2018. Sentiment analysis on large scale Amazon product reviews, in: 2018 IEEE International Conference on Innovative Research and Development (ICIRD). Presented at the 2018 IEEE International Conference on Innovative Research and Development (ICIRD), pp. 1–6. https://doi.org/10.1109/ICIRD.2018.8376299

Hassan, R., Islam, Md.R., 2021. Impact of Sentiment Analysis in Fake Online Review Detection, in: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). Presented at the 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), pp. 21–24. https://doi.org/10.1109/ICICT4SD50815.2021.9396899

Jabbar, J., Urooj, I., JunSheng, W., Azeem, N., 2019. Real-time Sentiment Analysis On E-Commerce Application, in: 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC). Presented at the 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), pp. 391–396. https://doi.org/10.1109/ICNSC.2019.8743331

Karthika, P., Murugeswari, R., Manoranjithem, R., 2019. Sentiment Analysis of Social Media Network Using Random Forest Algorithm, in: 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS). Presented at the 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), IEEE, Tamilnadu, India, pp. 1–5. https://doi.org/10.1109/INCOS45849.2019.8951367

Lin, T.-Y., Chakraborty, B., Peng, C.-C., 2021. A Study on Identification of Important Features for Efficient Detection of Fake Reviews, in: 2021 International Conference on Data Analytics for Business and Industry (ICDABI). Presented at the 2021 International Conference on Data Analytics for Business and Industry (ICDABI), pp. 429–433. https://doi.org/10.1109/ICDABI53623.2021.9655845

Mohawesh, R., Xu, S., Tran, S.N., Ollington, R., Springer, M., Jararweh, Y., Maqsood, S., 2021. Fake Reviews Detection: A Survey. IEEE Access 9, 65771–65802. https://doi.org/10.1109/ACCESS.2021.3075573
Nadali, A., Kakhky, E.N., Nosratabadi, H.E., 2011. Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy expert system, in: 2011 3rd International Conference on Electronics Computer Technology. Presented at the 2011 3rd International Conference on Electronics Computer Technology, pp. 161–165. https://doi.org/10.1109/ICECTECH.2011.5942073
N.L, J., J, M., S, K., n.d. Sentiment Analysis of Customer Reviews on Laptop Products for Flipkart.
Pai, A.R., Prince, M., Prasannakumar, C.V., 2022. Real-Time Twitter Sentiment Analytics and Visualization Using Vader, in: 2022 2nd International Conference on Intelligent Technologies (CONIT). Presented at the 2022 2nd International Conference on Intelligent Technologies (CONIT), pp. 1–4. https://doi.org/10.1109/CONIT55038.2022.9848043
Poonguzhali, R., Sowmiya, S.F., Surendar, P., Vasikaran, M., 2022. Fake Reviews Detection using Support Vector Machine, in: 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). Presented at the 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 1509–1512. https://doi.org/10.1109/ICSCDS53736.2022.9760747
Ramcharran, H., 2013. E-Commerce Growth and the Changing Structure of the Retail Sales Industry. Int. J. E-Bus. Res. IJEBR 9, 46–60. https://doi.org/10.4018/jebr.2013040104
SAP Machine Learning: Approaching your Project | SAP Blogs [WWW Document], n.d. URL https://blogs.sap.com/2018/08/28/sap-machine-learning-approaching-your-project/ (accessed 5.17.23).
Schröer, C., Kruse, F., Gómez, J.M., 2021. A Systematic Literature Review on Applying CRISP-DM Process Model. Procedia Comput. Sci. 181, 526–534. https://doi.org/10.1016/j.procs.2021.01.199
Shafique, U., Qaiser, H., 2014. A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA) 12.
Shivaprasad, T.K., Shetty, J., 2017. Sentiment analysis of product reviews: A review, in: 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT). Presented at the 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 298–301. https://doi.org/10.1109/ICICCT.2017.7975207
Srujan, K.S., Nikhil, S.S., Raghav Rao, H., Karthik, K., Harish, B.S., Keerthi Kumar, H.M., 2018. Classification of Amazon Book Reviews Based on Sentiment Analysis, in: Bhateja, V., Nguyen, B.L., Nguyen, N.G., Satapathy, S.C., Le, D.-N. (Eds.), Information Systems Design and Intelligent Applications, Advances in Intelligent Systems and Computing. Springer, Singapore, pp. 401–411. https://doi.org/10.1007/978-981-10-7512-4_40
Zhang, S., Zhang, C., Yang, Q., 2003. Data preparation for data mining. Appl. Artif. Intell. 17, 375–381. https://doi.org/10.1080/713827180