Enhancing Apple Product Review Sentiment Analysis using Machine Learning and Genetic Algorithm Optimization

Sruthy Sugunan

10626762

Applied Research Project submitted in partial fulfilment of the requirements for the degree of

Master of Science in Data Analytics at

Dublin Business School

Supervisor Name: Mr. Yalemisew Abgaz

August 2023

DECLARATION

I declare that this Applied Research Project that I have submitted to Dublin Business School for the award of Master of Science in Data Analytics is the result of my own investigations, except where otherwise stated, where it is clearly acknowledged by references. Furthermore, this work has not been submitted for any other degree.

Signed: SRUTHY SUGUNAN

Student Number: 10626762

Date:29/08/2023

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my family and friends for their invaluable support in various aspects, including financial, emotional, spiritual, and educational, throughout the course of this research. A special acknowledgment goes to Mr. Yalemisew Abgaz, my project supervisor at Dublin Business School, whose constant and dedicated guidance was a true blessing during the process of completing my dissertation. Working under his mentorship has been an exceptional experience, and I am truly fortunate to have had such an outstanding supervisor. Lastly, I extend my thanks to Dublin Business School for generously providing the necessary resources that enabled me to successfully undertake and complete this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT **1**

CHAPTER 1 .... **2**

INTRODUCTION **2**

1.1 BACKGROUND STUDY 2

1.2 AIM AND OBJECTIVES 2

1.3 SCOPE OF THE PROJECT 3

1.4 MOTIVATION 3

1.5 CHALLENGES 4

1.6 REPORT OUTLINE 6

CHAPTER 2 .... **7**

LITERATURE REVIEW **7**

2.1 INTRODUCTION 7

2.2 BACKGROUND 8

2.3 RELATED WORKS 8

2.4 EXISTING GAP 11

CHAPTER 3 .... **13**

DATASET **13**

3.1 INTRODUCTION 13

3.2 DATA COLLECTION 13

3.4 LEGAL AND ETHICAL CONCERS 17

CHAPTER 4 .... **19**

METHODOLOGY **19**

4.1 INTRODUCTION 19

4.2 PROPOSED METHOD 19

4.3 DATA PRE-PROCESSING 21

4.4 MODELLING 29

4.5 OPTIMIZATION 31

4.5 EVALUATION METRIC 32

CHAPTER 5 ....**35**

RESULT ANALYSIS...**35**

5.1 INTRODUCTION 35

5.2 IMPLEMENTATION DETAILS 35

5.3 R RESULT ANALYSIS OF ML MODELS ON POLARITY BASED SENTIMENT ANALYSIS 36

5.4 RESULT ANALYSIS OF ML MODELS ON STARS BASED SENTIMENT ANALYSIS 42

5.5 CUSTOM REVIEW PREDICTION 48

CHAPTER 6 ....**50**

SUMMARY, CONCLUSION AND RECOMMENDATIONS **50**

6.1 LIMITATIONS 50

6.2 FUTURE SCOPE 51

6.3 CONCLUSION 52

REFERENCES...**53**

Abbreviations

AI: Artificial Intelligence

CNN: Convolutional Neural Network

DNN: Deep Neural Network

FN: False Negative

FP: False Positive

GA: Genetic Algorithm

HTML: Hyper Text Markup Language

KNN: K-Nearest Neighbour

LSTM: Long Short-Term Memory

LR: Logistic Regression

ML: Machine Learning

NB: Naive Bayes

NLP: Natural Language Processing

ROC: Receiver Operating Characteristics

SVC: Support Vector Classifier

TN: True Negative

TP: True Positive

TPOT: Tree-based Pipeline Optimization Tool

URL: Uniform Resource Locator

LIST OF FIGURES

Figure 3. 1 Web scrapping from Trustpilot ………………………………………………15

Figure 3. 2 Data scrapped from Trustpilot …..…………………………………………… 15

Figure 3. 3 Web scrapping from Amazon…..……………………………………………. 16

Figure 3. 4 Data scrapped from Amazon. ………………………………………………...17

Figure 4. 1 Proposed Model ………………………………………………………………21

Figure 4. 2 Trustpilot Dataset ……………………………………………………………21

Figure 4. 3 Trustpilot Dataset information……………………………………………… 22

Figure 4. 4 Amazon Dataset ……………………………………………………………. 22

Figure 4. 5Amazon Dataset information……………………………………… …………22

Figure 4. 6 Renaming the attributes……………………………………………………… 23

Figure 4. 7 Pre-processed Amazon scrapped dataset……………………………………...23

Figure 4. 8 Combined datasets…………………………………………………………… 24

Figure 4. 9 Combined dataset information………………………………………….….24

Figure 4. 10 Removing null values from the final dataset ………………………………… 24

Figure 4. 11 Final dataset star ratings ……….…………………………………………25

Figure 4. 12 Positive review ………………………………………………………………25

Figure 4. 13 Negative review …………………………………………………………… 25

Figure 4.14 Combined reviews……………………………………………………….... 26

Figure 4. 15Data splitting………………………………………………………………… 29

Figure 4.16 Confusion Matrix …………………………………………………………..…34

Figure 5.1 Performance metric of SVM……………………………………………...……36

Figure 5.2 ROC of SVM……………………………...…………………………...37

Figure 5.3 Confusion Matrix of SVM…………………………………………….… 37

Figure 5.4 Performance metric logistic regression ………………………………………… 38

Figure 5.5 ROC of logistic regression …………………....……………………...… 38

Figure 5.6 Confusion Matrix of logistic regression……………………………………….38

Figure 5.7 Performance metric naïve bayes ……..………………………………………… 39

Figure 5.8 ROC of naïve bayes ……..…………………....……………………...… 39

Figure 5.9 Confusion Matrix of naïve bayes ……..………………………………………….40

Figure 5.10 Performance metric SVM with optimization ……..…………………………... 41

Figure 5.11 ROC of SVM with optimization ……..………………....…………………. 41

Figure 5.12 Confusion Matrix of SVM with optimization ……..…………………………….42

Figure 5.13 Performance metric logistic regression …………………………………….. 43

Figure 5.14 ROC of logistic regression …………………....……………………...….. 43

Figure 5.15 Confusion Matrix of logistic regression……………………………………...43

Figure 5.16 Performance metric of SVM……………………………………………...… 44

Figure 5.17 ROC of SVM……………………………...…………………………... 44

Figure 5.18 Confusion Matrix of SVM…………………………………………….… 45

Figure 5.19 Performance metric naïve bayes ……..……………………………………… 45

Figure 5.20 ROC of naïve bayes ……..…………………....…………………………... 46

Figure 5.21 Confusion Matrix of naïve bayes ……..……………………………………… 46

Figure 5.22 Performance metric SVM with optimization ……..…………………………... 47

Figure 5.23 ROC of SVM with optimization ……..…………………....………………….. 47

Figure 5.24 Confusion Matrix of SVM with optimization ……..……………………………48

LIST OF TABLES

Table 1 Performance Comparison of ML Models using polarity score …………………..…42

Table 2 Performance Comparison of ML Models using star-based sentiment analysis ….…48

ABSTRACT

Using cutting-edge machine learning techniques along with Genetic Algorithm (GA) for optimization, this research delivers a thorough study of sentiment analysis of customer evaluations for Apple items. Understanding consumer attitudes can be essential in guiding product improvements and marketing tactics given the enormous amount of user-generated data. The goal of this research is to create a precise and effective model that can automatically categorize customer evaluations as either favourable, negative, or neutral. The study initially goes through a lengthy preprocessing stage to purge and appropriately prepare the data. The next step is to train three well-known classification algorithms to identify if sentiment is positive, negative, or neutral. The sentiment classification was done based on the polarity score and star-based method These methods are Logistic Regression, Support Vector Classification (SVC), and Naive Bayes. The best-performing model is determined through a thorough assessment procedure, and it is then further optimized by Genetic Algorithm utilizing the Tree-based Pipeline Optimization Tool (TPOT). AUC, F1-score, accuracy, precision, recall, and other performance indicators are used to carefully examine the optimization's result. The comparison of the classifiers in polarity score-based sentiment analysis and the star-based analysis was done and it was found that the star-based analysis gave maximum accuracy 0.93 for SVM classifier with genetic programming optimization using TPOT. The project's findings offer insightful information about how consumers feel about Apple products and show how successful a coordinated strategy combining conventional ML algorithms with GA optimization can be. The outcomes not only aid in understanding consumer preferences and impressions but also present a novel technique that is transferable to various fields and goods. Throughout the project, legal standards for site scraping are observed, maintaining ethical integrity. Therefore, the study has significant ramifications for Apple's relentless pursuit of product development and client pleasure.

Keywords – Genetic Algorithm, Machine Learning, Sentiment Analysis, Customer reviews; TPOT

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND STUDY

User-generated material, particularly product evaluations, has grown to be a priceless resource in the current digital era for both businesses and customers. These evaluations are direct client feedback that businesses may use to learn about the advantages and disadvantages of their goods, therefore informing future product development and marketing plans and enhancing customer happiness. Reviews provide buyers with information about other people's experiences with the goods, affecting their decisions to buy.

As a market leader in consumer electronics, Apple Inc. reads a ton of user feedback on numerous platforms. Utilizing the value of these evaluations requires an accurate and effective system that can assess and categorize consumer sentiment. However, it is impractical to manually analyse such big datasets.

This study utilized machine learning techniques for sentiment analysis of Apple product evaluations to suggest a robust solution to this issue. The research seeks to create algorithms that can classify reviews into "positive," "negative," or "neutral" feelings with accuracy. To further improve the accuracy of these models, also use a Genetic Algorithm, an optimization method inspired by biological processes.

This comprehensive strategy not only makes it easier to comprehend how buyers view Apple's goods, but it also provides a scalable method that can be used by other companies or situations. As a result, our project aims to contribute to the larger conversation regarding data-driven decision-making in today's corporate environment, rather than merely developing a tool for sentiment analysis.

1.2 OBJECTIVES

- Data collection: To gather online testimonials for Apple goods from diverse sources while adhering to moral and legal standards.
- Data preparation and cleaning using methods including tokenization, stemming, vectorization, and dividing the data into training and test sets.

- Model Evaluation and Selection: To analyse three classification algorithms (Logistic Regression, SVM, and Naive Bayes), train them, and then choose the best one based on performance indicators.
- Application of the Genetic Algorithm with TPOT for optimization aims to optimize the model of choice.
- Performance Analysis: To thoroughly evaluate the optimized model utilizing a variety of measures to comprehend its efficacy and efficiency in foretelling feelings.
- Interpretation and Reporting: Consolidate data into useful insights to illustrate the approach's robustness and practical application.

1.3 SCOPE OF THE PROJECT

The main goal of this project is to use machine learning algorithms and genetic algorithm optimization to assess and comprehend consumer sentiment regarding Apple products.

1.4 MOTIVATION

The motivation for this study to spur innovation in the fields of sentiment analysis as well as more general marketing, consumer behaviour, machine learning, and ethical data gathering serves as the inspiration for this effort. It appeals to intellectual curiosity, real-world corporate applications, and a dedication to ethical research techniques.

- Understanding Consumer Behaviour: Researching consumer attitudes regarding Apple products offers insightful information about consumer tastes, viewpoints, and attitudes. Businesses and researchers may use this knowledge to pinpoint their strengths and areas for development.
- Utilizing Advanced Techniques: A prime example of the usage of cutting-edge data science approaches is the combination of genetic algorithm optimization and machine learning algorithms. This study attempts to demonstrate the usefulness of such cutting-edge approaches in sentiment analysis.
- Brand analysis: Since Apple is a well-known brand throughout the world, it is crucial for both the business and rivals to comprehend consumer attitudes regarding its goods. This study adds to the body of knowledge on brand management and marketing tactics.
- Enhancing Personalization: This project's insights may be used to improve systems for personalized marketing and recommendation. Businesses may adapt goods and

communication to individual needs and preferences by learning what resonates with customers.

- Filling a Research Gap: Previous studies on sentiment analysis may not have focused precisely on the mix of algorithms and optimization strategies employed in this study. As a result, this research may close a knowledge gap by offering fresh perspectives and methods.
- Ethical online Scraping: The project gains a social responsibility component thanks to the focus on legal compliance and ethical online scraping issues. It serves as an example of ethical data collection in the era of big data.
- Understanding public opinion of Apple's goods might have an impact on stock prices, public relations tactics, and other things because Apple is a key participant in the technology sector. Therefore, a variety of stakeholders, such as investors, legislators, and industry experts, could find this study interesting.
- Contribution to Machine Learning Community: This project can aid in the continuing development and improvement of algorithms within the ML community by utilizing and perhaps improving a variety of machine learning methodologies.

1.5 CHALLENGES

The difficulties encountered in this task are varied and intricate, and their resolution necessitates careful thinking and preparation. Each part of the project has its own specific set of hurdles, ranging from the technological challenges involved in online scraping and modelling to the subtle professional and ethical issues that underpin the entire process.

Technical difficulties

- Web Scraping Challenges:
    - The scraping code usually breaks because websites regularly modify their architecture.
    - Accessibility limitations include user-agent restrictions, IP bans, and captchas.
    - It might be challenging to guarantee the constant quality of scraped data.
- Data preparation:

    - Handling data that are inconsistent or missing.
    - If there are unequal numbers of good, negative, and neutral reviews, balance the dataset.
- Complexity modelling:
    - Hyperparameter tuning for various methods.
    - Ensuring that the data is not under or overfit by the models.
    - A thorough grasp of the algorithms and careful implementation may be necessary for the integration of genetic algorithms.

Professional Difficulties

- Observance of Legal Requirements:
    - Respecting the terms of service of many websites.
    - Observing legal requirements across jurisdictions, particularly those relating to data privacy.
- Time Restrictions:
    - Preprocessing, model training, and web scraping can be time-consuming procedures.
- Resource constraints:
    - Large datasets and intricate model training may need a lot of computer power.

Ethical Issues

- Privacy and Consent:
    - Personal information might be present in the web data gathered. It is essential to make sure that such data is anonymised or managed with the highest secrecy.
- Transparency:

    - Keeping the procedures employed open-book and ensuring that the results are accurately and impartially reported.

- Suitable AI Techniques:
    - Ensuring that the models don't unintentionally create or reinforce biases.
    - Carefully using the insights learned, particularly when used in a commercial or marketing environment.

1.6 REPORT OUTLINE

This thesis comprises primarily six chapters.

Chapter 1 explains the introduction section which methodically outlines the project's setting. This chapter describes the project's goals and purposes. Explaining the study's aims and aspirations, the project's purpose is detailed.

The emphasis of **Chapter 2** is the literature review to extract insight from prior research, finding congruencies and gaps that support and guide the current study's methodology. The review uses a unique combination of methodology and ethical standards to explore sentiment analysis of Apple product evaluations.

Chapter 3 explains the main data collection process used in this project.

Chapter 4 provides an overview of the proposed method involves sentiment analysis of Apple product reviews and comprises several stages, integrating data collection, preprocessing, machine learning, optimization, and performance analysis.

Chapter 5 examines a thorough examination of the findings from the many computer experiments carried out during this investigation. The three main areas of concentration are the performance assessment of machine learning models, the investigation of genetic programming methods, and the novel custom review prediction method.

Chapter 6 is the conclusion and future scope of the project.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The literature review chapter extracts insight from prior research, finding congruencies and gaps that support and guide the current study's methodology. The review uses a unique combination of methodology and ethical standards to explore sentiment analysis of Apple product evaluations.

2.2 BACKGROUND

The field of sentiment analysis has been continuously developing, occupying a significant position within the larger field of machine learning and natural language processing. The capacity to effectively comprehend and exploit this information has been of utmost relevance due to the ever-increasing volume of data available online, particularly in the form of user evaluations. This chapter begins a thorough review of the body of literature, concentrating on topics like sentiment analysis, web scraping, machine learning algorithms, genetic algorithm optimization, and ethical issues in data collection and analysis that are closely related to the goals and methods of this study.

- Sentiment Analysis: The examination begins by looking at various sentiment analysis methods, from basic lexicon-based strategies to complex machine learning models. Studies that have focused on product reviews receive significant attention, especially when they relate to technology companies like Apple (Anjaria & Guddeti, 2014).
- Online scraping and data preprocessing: The review includes a section on comprehending the difficulties and best practices in online scraping, highlighting the methods used to guarantee the accuracy and consistency of data collecting. To lay a basis for the methods used in this work, the literature on data pretreatment and transformation is also explored.
- Machine Learning Algorithms for Sentiment Analysis: This section explores the use of Naive Bayes, SVM, and Logistic Regression in sentiment analysis, combining their benefits and drawbacks as stated in various research. These algorithms' comparisons in various settings will provide a sophisticated understanding of how well-suited they are to the current undertaking (Jagdale et al., 2019).

- Genetic Algorithm Optimization: The literature on genetic algorithm optimization is examined, as well as how it relates to machine learning, especially when combined with TPOT-type techniques. This analysis provides insight into how this innovative method has been applied in related situations.

2.3 RELATED WORKS

Customer satisfaction determines a company's long-term success. Brand value, and customer satisfaction, in a company's long-term survival **(Kang et al., 2019)** are prominent factors in the success of a product. Therefore, many organizations spend millions on marketing research to understand client preferences and needs. Effective branding and positioning strategies require understanding how buyers see the things they buy based on these facts. Other parties' interests have complicated business **(Chaki, 2015.** Due to the changing business environment, organizations need an early change detection system to respond (**Kahaner, 1997)**. Companies need time and resources to adapt to a changing environment that may threaten their security and profitability. To achieve common goals, companies should take complementary, synergistic, interactive, and goal-oriented actions **(Tuan, 2016)**. Given this, competitor analysis, also known as competitive research, is a way to compare the firm and its products or services to others in the market. Strategic competitive research must understand customer perceptions of rivals. This will make business strengths and shortcomings easy to analyse. Sentiment analysis may be used to measure interest in certain issues, learn about market circumstances, and investigate rivals. Text analytics sentiment analysis works well for this.

Sentiment research helps marketers comprehend the "voice of the customer." Understanding the target audience's opinions, values, and motives improves targeted advertising. Sentiment analysis classifies people's feelings in text data as negative (1), neutral (0), or positive (+1). It identifies, evaluates, and classifies subjective data from the source material, particularly feelings. Sentiment analysis has been present since the late 1990s, although its use has increased**(Tang et al., 2009)**. This helps marketers pick tactics, platforms, and creative approaches to engage clients. Sentiment research may assist marketers in identifying their target audiences, but it can also show what their rivals' consumers think about them in terms of price, customer service, value, features, goods, mobile apps, etc. Sentiment research compares customers' opinions of a brand's goods and services to those of its competitors. This data may help brands stand out. Artificial intelligence (AI) has enhanced several domains' performance, but sentiment analysis. Sentiment analysis is the method of identifying emotions conveyed in text using artificial intelligence. Instead of only determining if certain words within a piece of text have a favourable or negative meaning, AI understands the tone of a remark.

Lexicon-based and AI/ML algorithms can categorize sentiment analysis processes**(Anjaria & Guddeti, 2014; Yiran & Srivastava, 2019)**. Businesses need AI-powered sentiment analysis. It simplifies client thinking. It helps comprehend employee feelings too. It is usually used to assess customer evaluations, service issues, social media citations, etc. Sentiment analysis may soon reveal people's sentiments. It will help computers recognize emotions in messages, observations, and feedback. Thus, businesses may better customize their replies.

Sentiment analysis detects ambiguity. Sentiment analysis shows representative and user opinions on a topic. Writers' tone and style frequently reflect their feelings. Several algorithms analyse, forecast, and evaluate sentiments from text data like product or customer reviews. Polarity recognition can benefit from sentiment analysis. It also has problems with a huge vocabulary, domain dependency, negation, NLP load, bipolar terminology, spam, and misleading data. Sentiment analysis requires preprocessing to organize text and extract information for text mining and machine learning **(Palomino & Aider, 2022)**. Data preprocessing prepares text for classification **(Haddi et al., 2013)**. Preprocessing categorizes characters from a text stream by letter, moving between stages. Punctuation, emojis, emoticons, and white spaces assist control random text. Since computers can't interpret unstructured material, AI systems must preprocess it. Nhlabano and Lutu's study **(Nhlabano & Lutu, 2018)** found that text preparation improves sentiment analysis algorithms.

Semantic sentence analysis improves technique, consistency, and precision, according to **(Mangold & Faulds, 2009)**. Digital platforms have extended text-based consumer review data. Marketing scholars have examined text reviews in various ways. The augmented NLP method by Kanakaraj et al.**(Kanakaraj & Guddeti, 2015)** improves emotion taxonomy and prediction accuracy by merging word senses and reference sensations into vector features and employing ensemble classification.

NLP evaluates, highlights, and extracts meaning from text and voice using machine learning and algorithms **(Castillo & Taherdoost, 2023)**. An AI framework may search the text for certain words or phrases. In advanced applications, it lets computers recognize the speaker's emotion better than humans. AI-based sentiment analysis may work.

Use web scraping to extract the content from an e-commerce website. It will cycle through the roughly ten pages of comments for each product. Online product reviews were gathered for this project using web scraping. The online reviews that have been gathered are examined using classification models including KNN, SVM, Random Forest, CNN**(Jagdale et al., 2019)**, and the suggested hybrid SVM-CNN. The categorization model experiments have shown encouraging results **(Suganya & Vijayarani, 2020)**.

Sentiment analysis and opinion mining are the most popular methods for evaluating text data from Facebook, Twitter, Amazon, and others. It helps companies improve their strategy and understand user feedback on their goods. Computational analysis of purchasing interests and opinion mining regarding a company's corporate body is involved. An event, person, blog piece, or product experience can represent this item. This article incorporates Amazon statistics on cameras, computers, phones, tablets, TVs, and security cameras. After preprocessing, machine learning classified reviews as positive or unfavourable. This study finds that machine learning classifies product reviews best. Naive Bayes got 98.17% camera review accuracy, while SVM had 93.54% **(Jagdale et al., 2019)**. The articles reviewed **(Taherdoost & Madanchian, 2023)** (2012–2022) examine how competitive market research finds and contrasts key market metrics that aid in differentiating the products and services of the rivals. It is possible to find out what clients of rivals think of them in all facets of their operations by using AI-powered sentiment analysis.

The choice of the best hyperparameters is a major problem for deep learning in methods of sentiment analysis like long-short-term memory. The LSTM-DGWO deep learning model combines LSTM and DGWO in this study. The BERT architecture is used to analyze the app review dataset to create effective word embeddings. The genetic algorithm (GA) extracts the review features, and the firefly algorithm (FA) extracts the best review characteristics. Finally, the DGWO method improves the LSTM model's hyperparameters while the LSTM-DGWO model classifies app reviews**(Barik et al., 2023)**.

(Chinnalagu & Durairaj, 2021), the study developed a high-performance, low-cost model to reliably predict attitudes from large customer rating datasets. This model uses LSVM and fastText from Facebook's AI research (FAIR) Lab to classify text and word embedding. The author's SA-BLSTM model was also compared to this model. The fastest model surpasses the LSVM and SA-BLSTM models by 20% and has 90.71% accuracy, according to the data.

(Naseem et al., 2021) surveys the latest work on applying machine learning and deep learning models, such as support vector machines (SVM), linear regression models, Naive Bayes models, convolutional neural networks (CNN), long short-term memory models, deep neural networks (DNN), and others, to sentiment analysis problems like product and app(s) review analysis, textual and visual analysis, and others.

The development of microblogging sites such as Twitter provides an unprecedented opportunity to develop and apply theories and technology that search for and my sentiments. This paper describes an innovative way to sentiment analysis using Twitter data. When utilizing this strategy, machine learning algorithms (Naive Bayes classifier) can obtain high accuracy in classifying sentiment (Kaur, 2017).

Analysing user feelings toward apps via review comments and ratings can be financially beneficial to app creators.(Sangani & Ananthanarayanan, n.d.) a system that provides a many-to-many mapping from reviews to topics of interest, as well as a list of reviews for each topic that convey user sentiment toward that topic.

Online reviews tell buyers about product quality, performance, and ideas. Websites provide information or reviews from individuals, groups, or organizations about a subject, such as Apple products. (Indrayuni & Nurhadi, 2020) study categorizes 200 Apple product customer evaluations as positive or negative using Support Vector Machine and genetic algorithms (GA) for optimal accuracy. Research stages: First, get data, then preprocess.

2.4 EXISTING GAP

There aren't many studies on sentiment analysis of Apple products, although there are plenty that concentrate on generic product assessments or product categories. Customers of other companies may not be aware of the specific thoughts and opinions made by customers of Apple products.

ML-based sentiment analysis research may not have used genetic algorithm optimization. This discrepancy proposes experimenting with genetic algorithms to improve the efficacy and accuracy of sentiment analysis models. The absence of domain-specific data sets for reviews of Apple products It could be difficult to get high-quality, domain-specific datasets for sentiment analysis of Apple products. A sizable dataset explicitly including a range of Apple devices and user attitudes may be missing from the investigation.

Limited Comparative Studies: There may not be many comparative studies comparing ML algorithms and optimization techniques for Apple product sentiment study. It is uncertain which ML algorithm and optimization method will work best for this task.

Results Interpretability: Sentiment analysis models need to be comprehensible for practical uses like analyzing consumer comments. Interpretable algorithms for sentiment analysis of Apple product reviews might close a knowledge gap.

Handling Multimodal Data: Sentiment analysis of Apple product evaluations may require a textual and visual analysis of multimedia material spreads more widely. There may not be enough studies on multimodal data sentiment analysis of Apple product reviews. The study may have overlooked the use of pre-trained models for sentiment analysis of Apple product evaluations.

Class Unbalance Resolution: Sentiment analysis datasets may contain class imbalance, where specific attitudes predominate. Studies on class inequality may be lacking in Apple product sentiment analyses.

Real-time Analysis: The literature may be lacking research on real-time sentiment analysis of Apple product reviews, which might be helpful for quick responses and customer care.
Sentiment analysis studies on user views and assessments may be lacking, which is an ethical concern. This topic of study may be able to help with potential privacy and bias concerns.

CHAPTER 3

DATA SET

3.1 INTRODUCTION

In this chapter, explains the main data collection process used in this project.

3.2 DATA COLLECTION

Data collection is a critical phase in many projects, and in the context of your study, it involves scraping online reviews related to Apple products. Web scraping solutions like Beautiful Soup offer a potent way to harvest data from the internet to find specific information. Although technically challenging, this procedure makes it possible to extract important insights from user-generated information. The strategy must balance the necessity for knowledge with the duty to collect it legally and ethically. This part of the research process is crucial because the data you gather for your project serves as the foundation for all later analysis and modelling.

- Installing Necessary Libraries: HTML is parsed, and data is extracted using Beautiful Soup. The requests library may be used in conjunction with Beautiful Soup to send HTTP requests to get web pages.

- Determining the URLs of the sites containing the reviews you wish to scrape. This can include Amazon product pages or Trustpilot brand review pages.

- Making HTTP Requests: Use the requests library to retrieve the target pages' HTML content. Make careful to manage any potential issues and abide by the robots.txt file for the website.

- Parsing HTML Content: Add the HTML content to a Beautiful Soup object as its initial value. This enables you to browse the HTML tree and look for components that contain the data you want.

- Extracting Reviews: Determine the HTML elements and attributes that hold the review information (such as the title, rating, date, and review content). To extract the necessary information, iterate over these components.

- Handling Pagination: If reviews span several pages, you'll need to figure out the pattern for the pagination and loop over the associated URLs to scrape all the necessary pages.

- Data Storage: You have the option of saving the scraped information in a CSV file or another organized format. To do this, arrange the data into a panda DataFrame first, and then export it to CSV.

- Compliance and Ethical Considerations: It's critical to make sure that the scraping complies with the Terms of Service of the websites and any relevant legislation. This may entail abiding with rate caps, user-agent guidelines, and any applicable scraping bans.

- Error Handling and Robustness: The scraper may be made more robust and less likely to be blocked by implementing correct error handling and perhaps integrating features like retries, user-agent rotation, and proxies.

Web Scraping from Trustpilot:

- Setting up the environment: Import essential libraries like Beautiful Soup, requests, and pandas. Set the URL of the Trustpilot page or pages that contain the reviews you wish to scrape.
- Scraping Multiple Pages: If the reviews span more than one page, you might need to cycle through many of them. To do this, find the Trustpilot pagination pattern and iterate over the URLs in accordance.
- Data Extraction: Parse the HTML code using Beautiful Soup. The html used in the project - https://www.trustpilot.com/review/www.apple.com
- Determine the HTML elements and attributes (such as title, rating, and review text) that contain the review information. Take this data and store it in an appropriate data structure, such as a list of dictionaries.
- Loading Data into CSV: Create a Pandas DataFrame from the extracted data. To save the DataFrame as a CSV file, use pandas' to_csv function.

The web scrapping from Trustpilot is shown in Figure 3.1 and the screenshot of the extracted csv file is given in Figure 3.2.

```python
# Main function to scrape multiple pages and store data in a CSV file
def scrape_multiple_pages(url, num_pages):
    all_data = []

    for page in range(1, num_pages + 1):
        page_url = f'{url}?page={page}'
        data = scrape_reviews(page_url)
        all_data.extend(data)

    # Create a Pandas DataFrame
    df = pd.DataFrame(all_data, columns=['name', 'Location',␣
 ↪'Headline','review_text', 'date', 'Stars'])

    # Save the DataFrame to a CSV file
    df.to_csv('trustpilot_reviews.csv', index=False)
    print('Data successfully scraped and saved to trustpilot_reviews.csv')

# URL of the Trustpilot page with Apple product reviews
trustpilot_url = 'https://www.trustpilot.com/review/www.apple.com'

scrape_multiple_pages(trustpilot_url, num_pages=309)
```

```
Data successfully scraped and saved to trustpilot_reviews.csv
```

Figure 3. 1 Web scrapping from Trustpilot

name	Location	Headline	review_text	date	Stars
Bryce Pegg	US	My Apple Pay got ha	My Apple Pay got ha	April 25, 2023	1
Ben	US	"Just works" no it doe	This company is goi	April 25, 2023	1
Helen	GB	Apple Care Con	I got a new phone re	April 25, 2023	1
Leahcatty	GB	Customer service at a	Had a wonderful exp	April 25, 2023	5
DC	GB	Incompetence and c	Apple sent my less th	April 25, 2023	1
Valia	GB	No financing in the s	As a fan of Apple pro	April 25, 2023	1
David Sansom	FR	New M2 Pro - is awso	I still don`t get why th	April 25, 2023	5
Adam	GB	Fantastic Service	Fantastic Service - C	April 25, 2023	5
BigBenF	US	Pointless !!!	Pointless !! I hate ha	April 25, 2023	1
Raheese Munir	GB	Rude staff	I went into an apple	April 25, 2023	1
Ashfaq King	BD	Great brand	Great brand. I am big	April 25, 2023	5

Figure 3. 2 Data scrapped from Trustpilot

Web Scraping Reviews from Amazon for Apple Products:

- Define the URLs of the Amazon pages that contain reviews for Apple items to identify the Apple products' sites. The URL used is - https://www.amazon.in/New-Apple-iPhone-Pro-512GB/product-reviews/B08L5WVZJ7/ref=cm_cr_dp_d_show_all_btm?ie=UTF8&reviewerType=all_reviews

- Scraping reviews: Like that of Trustpilot, you must utilize Beautiful Soup to read the HTML code of the Amazon pages.com pages. Extract pertinent review information from the text, rating, and review title.
- Handling Amazon's Anti-Scraping Measures: To get around Amazon's strict anti-scraping policies, you might need to adopt strategies like changing user agents, employing proxies, or slowing down the rate at which you scrape.
- Combining and Saving Data: Depending on your analysis needs, you can add the Amazon data to the Trustpilot Data Frame or keep it separate. As before, save the data to a CSV file.

The web scrapping from Amazon is shown in Figure 3.3 and the scrapped file is given in Figure 3.4.

```
for i in range(1, 500):
    link = "https://www.amazon.in/New-Apple-iPhone-Pro-512GB/product-reviews/
↪B08L5WVZJ7/ref=cm_cr_dp_d_show_all_btm?ie=UTF8&reviewerType=all_reviews" +␣
↪str(i)
    response = requests.get(link)
    soup = bs(response.content, "html.parser")

    # Extracting Review Title
    title = soup.find_all('a', class_='review-title-content')
    review_title = [t.get_text().strip() for t in title]
    wm_title.extend(review_title)

    # Extracting Ratings
    rating = soup.find_all('i', class_='review-rating')
    review_rating = [r.get_text().strip().replace(' out of 5 stars', '') for r␣
↪in rating]
    wm_rating.extend(review_rating)

    # Extracting Content of review
    review = soup.find_all("span", {"data-hook": "review-body"})
```

Figure 3. 3 Web scrapping from Amazon

	Title	Ratings	Comments	Date
0	5.0 out of 5 stars THE BEST Phone T	5	We all heard rumors	03-06-2021
1	5.0 out of 5 stars First time iphone use	3	First impression: Awe	17-08-2021
2	4.0 out of 5 stars Recommended to bu	5	I have been using thi	03-06-2021
3	5.0 out of 5 stars The best iPhone witl	5	I have been into app	28-05-2021
4	5.0 out of 5 stars It's Smooth and fast	4	Out of Five StarsBatt	10-03-2022
5	5.0 out of 5 stars Excellent	5	All Good	18-11-2020
6	5.0 out of 5 stars Currently the best ph	5	This is the 'all we ca	06-09-2021
7	5.0 out of 5 stars Most luxury defining	5	Enough has been sai	30-01-2023
8	5.0 out of 5 stars Happy with the purcl	5	First time using an iP	11-12-2020
9	5.0 out of 5 stars Well it's an iPhone..	5	If you can afford it pl	17-01-2021
10	5.0 out of 5 stars THE BEST Phone T	5	We all heard rumors	15-05-2021
11	5.0 out of 5 stars First time iphone use	5	First impression: Awe	17-06-2021
12	4.0 out of 5 stars Recommended to bu	5	I have been using thi	03-06-2021

Figure 3. 4 Data scrapped from Amazon.

3.3 LEGAL AND ETHICAL CONCERNS

A thorough knowledge of the applicable laws and a dedication to ethical behaviour is necessary for navigating the legal and ethical terrain of online scraping. To guarantee complete compliance, it could be a good idea to speak with a legal professional knowledgeable about the regulations in your country (*Building Trustworthy and Ethical AI Is Everyone's Responsibility*, n.d.). With careful planning and attention to the best standards, it is possible to strike a delicate balance between gaining important insights and maintaining legal and ethical constraints. Web scraping is a useful tool for data collecting, but it also presents several legal and ethical issues that need to be properly explored. The following concerns should be taken care of while scraping data for this project, especially from websites like Trustpilot and Amazon:

- Adherence to Terms of Service (ToS): The ToS of many websites have explicit sections that forbid or restrict scraping. You must read and abide by these conditions to prevent legal issues.
- Privacy Issues: It is essential to handle any scraped material that contains personal information carefully. It may be required to aggregate or anonymize the data to avoid individual identity and to adhere to privacy rules like GDPR.

- Intellectual Property Rights: Some websites' material may be covered by intellectual property regulations. It may be against these rights to use such data without the appropriate authorisation.
- Access Restrictions: Avoiding any technological restrictions set up by a website (such as CAPTCHAs or login requirements) may result in legal problems. Any attempts to get beyond these restrictions can be considered unlawful access.
- Rate Limiting and Fair Use: Scraping a website excessively or fast might overtax its servers, limiting access for other users. This problem can be reduced by observing rate caps and applying fair usage strategies like scraping during off-peak hours.
- Transparency and Purpose of Use: It is morally required to be transparent about the reason for data collection and the intended use of the data. There may be mistrust and even legal issues because of deceptive or covert motives.
- Data Security: It is both ethically and legally required to ensure that the data obtained is stored safely, especially if it involves sensitive data.
- Jurisdictional Considerations: Web scraping is subject to various legal restrictions in various nations and areas. It may be required to abide by the rules of both the region where the scraping takes place and the nation where the website being scraped is situated.
- Getting Permissions When Needed: In some circumstances, requesting express permission from the website owner before scraping may be the most moral and legitimate course of action.

CHAPTER 4

METHODOLOGY

4.1 INTRODUCTION

This chapter provides an overview of the proposed method involves sentiment analysis of Apple product reviews and comprises several stages, integrating data collection, preprocessing, machine learning, optimization, and performance analysis.

4.2 PROPOSED MODEL

The project's proposed methodology combines genetic algorithm optimization with contemporary machine learning techniques to provide an organized and thorough approach to sentiment analysis. It stresses ethical issues in addition to accuracy and efficiency, demonstrating a thoughtful and responsible approach to data science. The method attempts to give actionable insights into consumer attitudes regarding Apple goods through the integration of numerous methodologies and technologies, which can aid in both academic knowledge and real-world commercial decision-making. The proposed method is shown in Figure 4.1.

- Data Collection:
    - Web Scraping: Utilize Beautiful Soup to scrape reviews of Apple products from websites like Trustpilot and Amazon. Multiple pages are accessed to gather a comprehensive dataset.
    - Storing Data: The scraped reviews are loaded into a CSV file for further processing.
- Data Preprocessing:
    - Cleaning: Remove unnecessary characters, HTML tags, and other noise from the text data.
    - Tokenization: Break down the text into individual words or tokens.
    - Normalization: Convert text to a standard form, such as lowercase, and perform stemming or lemmatization.
    - Vectorization: Transform the text data into a numerical format using techniques like TF-IDF or word embeddings.
- Model Building and Training:

    - Algorithms Selection: Choose three algorithms for the initial phase - Logistic Regression, Support Vector Classifier (SVM), and Naive Bayes.
    - Training: Divide the data into training and validation sets and train the selected models on the preprocessed data.
    - Tuning: Fine-tune hyperparameters for optimal performance.
- Model Selection:
    - Evaluation: Compare the performance of the three models using metrics such as accuracy, precision, recall, and F1-score.
    - Best Model Selection: Select the best-performing model based on the evaluation criteria.
- Optimization using Genetic Algorithms (GA):
    - Utilizing TPOT: Apply the TPOT library, which uses Genetic Algorithms, to further optimize the best model.
    - Parameter Optimization: Use GA to find the optimal set of hyperparameters for the chosen model.
- Performance Analysis:
    - Testing: Evaluate the optimized model on a test set to assess its real-world applicability.
    - Interpretation: Analyze the results to understand the underlying sentiments and patterns within the reviews.

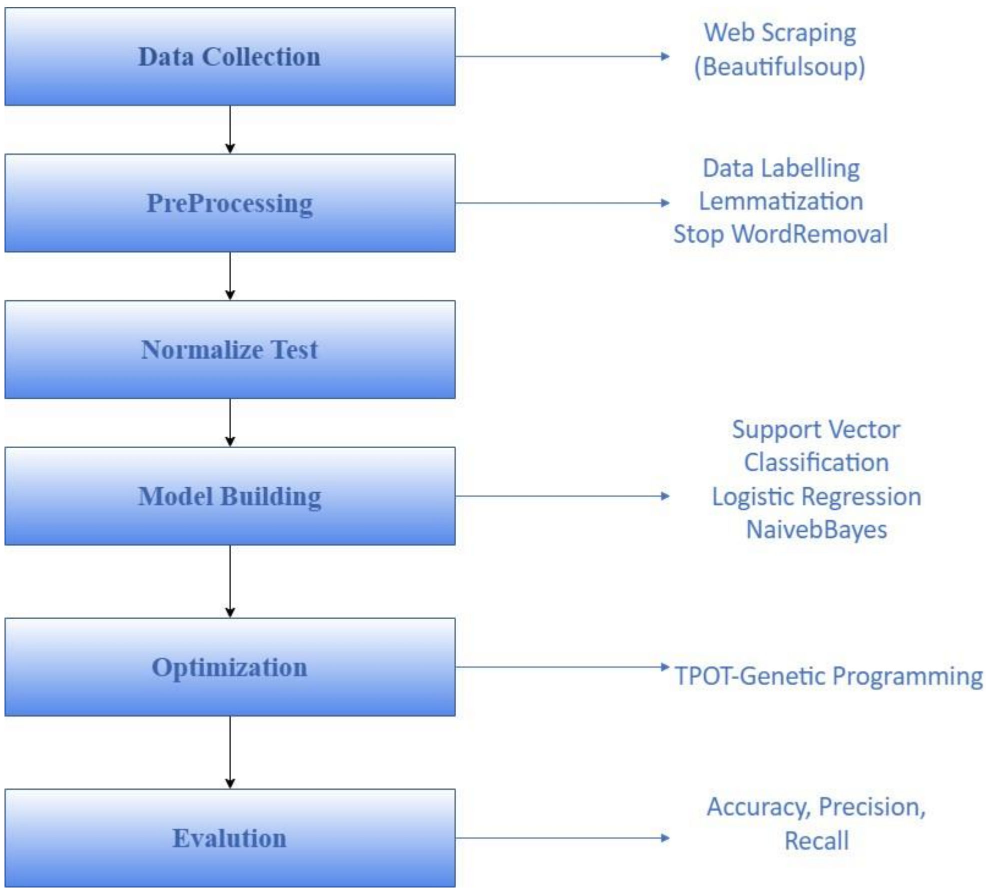

Figure 4. 1 Proposed Model

4.3 DATA PRE-PROCESSING

This data set is collected using the web scrapping technology and the collected data is saved to a csv file and it is used for modelling purposes. Figure 4.2 shows the first 5 entries of the dataset collected from the trust pilot. The dataset information is given in Figure 4.3, and it has 4820 data entries and has attributes name, location, headline, review, date and stars.

```
df.head()
```

	name	Location	Headline	review_text	date	Stars
0	Sky Sanderfer	US	APPLE is not the progressive company...	APPLE is not the progressive company they pret...	August 08, 2023	1
1	Happy Wanderer	CA	Scumbag company	Scumbag company. No decency, no integrity, no ...	August 08, 2023	1
2	pridehappy	GB	Outstanding Experience	From my experience, Apple Support is simply ex...	August 08, 2023	5
3	Melissa Mills	US	Apple sucks in so many ways	Tried to screen cast to a TV and it took forev...	August 08, 2023	1
4	Jacques van den Heever	HK	Worst service experience!	Im in a foreign country, and apple pay decided...	August 08, 2023	1

Figure 4. 2 Trustpilot Dataset

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4820 entries, 0 to 4819
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype 
---  ------       --------------  ----- 
 0   name         4818 non-null   object
 1   Location     4820 non-null   object
 2   Headline     4820 non-null   object
 3   review_text  4690 non-null   object
 4   date         4820 non-null   object
 5   Stars        4820 non-null   int64 
dtypes: int64(1), object(5)
memory usage: 226.1+ KB
```

Figure 4. 3 Trustpilot Dataset information

Figure 4.4 shows the first 5 entries of the dataset collected from Amazon. The dataset information is given in Figure 4.5, and it has 3036 data entries and has attributes unnamed, title, ratings, comments and date.

```
#df_df1 = pd.read_csv("/content/review_ammzon.csv", encoding='latin-1')
```

```
df_df1.head()
```

	Title	Ratings	Comments	Date
0	5.0 out of 5 stars\nTHE BEST Phone That Money ...	5.0	We all heard rumors about several stuffs but a...	2021-06-03
1	5.0 out of 5 stars\nFirst time iphone user rev...	3.0	First impression: Awesome product! I mean the ...	2021-08-17
2	4.0 out of 5 stars\nRecommended to buy	5.0	I have been using this mobile from the past si...	2021-06-03
3	5.0 out of 5 stars\nThe best iPhone with the b...	5.0	I have been into apple ecosystem since iPhone ...	2021-05-28
4	5.0 out of 5 stars\nIt's Smooth and fast	4.0	Out of Five StarsBattery Life ****Screen ****P...	2022-03-10

Figure 4. 4 Amazon Dataset

```
df_df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3036 entries, 0 to 3035
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype         
---  ------    --------------  -----         
 0   Title     2530 non-null   object        
 1   Ratings   3036 non-null   float64       
 2   Comments  2530 non-null   object        
 3   Date      3036 non-null   datetime64[ns]
dtypes: datetime64[ns](1), float64(1), object(2)
memory usage: 95.0+ KB
```

Figure 4. 5 Amazon Dataset information

4.3.1 Data Preparation

Data pre-processing refers to the techniques used to prepare and clean data for analysis or model training in machine learning. This stage is crucial because the quality and quantity of the data that feeds into the model can significantly affect the model's performance. The datasets scrapped from Amazon and the trust pilot are in different forms so need to preprocess the data before combining the datasets. The data scrapped from Amazon is reprocessed by renaming the attributes 'Comments' column to 'Review_text' and the 'Ratings' column to 'Stars' as shown in Figure 4.6. The modified pre-processed dataset is shown in Figure 4.7.

```
# Rename the 'Comments' column to 'Review_text'
df_df1 = df_df1.rename(columns={'Comments': 'review_text'})

# Rename the 'Ratings' column to 'Stars'
df_df1 = df_df1.rename(columns={'Ratings': 'Stars'})
```

```
df_df1.head()
```

	Unnamed: 0	Title	Stars	review_text	Date
0	0	5.0 out of 5 stars\nTHE BEST Phone That Money ...	5	We all heard rumors about several stuffs but a...	03-06-2021
1	1	5.0 out of 5 stars\nFirst time iphone user rev...	3	First impression: Awesome product! I mean the ...	17-08-2021
2	2	4.0 out of 5 stars\nRecommended to buy	5	I have been using this mobile from the past si...	03-06-2021
3	3	5.0 out of 5 stars\nThe best iPhone with the b...	5	I have been into apple ecosystem since iPhone ...	28-05-2021
4	4	5.0 out of 5 stars\nIt□s Smooth and fast	4	Out of Five StarsBattery Life ****Screen ****P...	10-03-2022

Figure 4. 6 Renaming the attributes.

```
df.head()
```

	review_text	Stars
0	My Apple Pay got hacked, they changed my banki...	1
1	This company is going down the toilet.	1
2	I got a new phone recently, and foolishly purc...	1
3	Had a wonderful experience today with a custom...	5
4	Apple sent my less than 6 month old MacBook Pr...	1

Figure 4. 7 Pre-processed Amazon scrapped dataset

After preprocessing the datasets are merged to form the dataset, which is shown in Figure 4.8, the information of the combined dataset is given in Figure 4.9 with attributes review_text and stars with 6676 entries and the different attribute count is shown in Figure 4.11

```
# Combine the DataFrames vertically using concat
demo_df = pd.concat([df, df_df1], ignore_index=True)
```

```
demo_df.head()
```

	review_text	Stars
0	My Apple Pay got hacked, they changed my banki...	1
1	This company is going down the toilet.	1
2	I got a new phone recently, and foolishly purc...	1
3	Had a wonderful experience today with a custom...	5
4	Apple sent my less than 6 month old MacBook Pr...	1

Figure 4. 8 Combined datasets

```
demo_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7856 entries, 0 to 7855
Data columns (total 2 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   review_text  7220 non-null   object 
 1   Stars        7856 non-null   float64
dtypes: float64(1), object(1)
memory usage: 122.9+ KB
```

Figure 4. 9 Combined dataset information

Dropping null values and checking for their presence are common steps in data preprocessing, especially when handling real-world datasets that may contain missing or incomplete information. The null values are checked the count is of null values is zero as shown in figure 4.10.

```
final_df = demo_df.dropna()
```

```
final_df.isnull().sum()
```

```
review_text    0
Stars          0
dtype: int64
```

Figure 4. 10 Removing null values from the final dataset

```
final_df['Stars'].value_counts()
```

```
1.0    3657
5.0    2638
3.0     366
4.0     359
2.0     200
Name: Stars, dtype: int64
```

Figure 4. 11 Final dataset star ratings

4.3.2 Sentiment Intensity Analyzer based on Polarity Score

Sentiment Intensity Analyzer is a tool used to determine the sentiment or emotion expressed in a text. Used in natural language processing (NLP) and sentiment analysis to categorize text as positive, negative, neutral, or a combination thereof. Using the Sentiment Intensity Analyzer (SIA), we can identify the sentiment of a review and categorize it as "Positive," "Negative," or "Neutral." The sentiment ratings for the text in the variable review are determined by the formula sentiment = polarity scores(review). The polarity score's function produces a dictionary with a variety of scores that characterize the text's mood. This often comprises positivity, negativity, neutrality, and a compound score—a summary measure with a range of -1 (most negative) to +1 (most positive)—as well as other ratings.

The compound score is categorized into one of three sentiment categories using the conditional expressions below:

- The sentiment label is assigned to "Positive" if compound score > 0.2.
- Compound score -0.2 results in a "Negative" sentiment label being assigned.
- The sentiment label is changed to "Neutral" if the compound score is in the -0.2 to 0.2 range.

The script loads two CSV files named "Negative_reviews_star.csv" and "Positive_reviews_Stars.csv" containing negative and positive customer reviews, respectively. The negative reviews obtained are shown in Figure 4.13, positive reviews in Figure 4.12 and combined in Figure 4.14.

```
Positive_reviews.head()
```

	Stars	review_text	Sentiment
3	5	Had a wonderful experience today with a custom...	Positive
5	1	As a fan of Apple products, I finally made the...	Positive
6	5	I still don`t get why there are so many haters...	Positive
7	5	Fantastic Service - Oxford Street Store. Thank...	Positive
10	5	Great brand. I am big fan of Apple products.	Positive

Figure 4. 12 Positive review

	Stars	review_text	Sentiment
0	1	My Apple Pay got hacked, they changed my banki...	Negative
1	1	Apple sent my less than 6 month old MacBook Pr...	Negative
2	1	Pointless !! I hate having to deal with Apple....	Negative
3	1	I went into an apple retail store one of the s...	Negative
4	1	Can't help me without phone serial number. Don...	Negative

Figure 4. 13 Negative review

```
combined_df
```

	Stars	review_text	Sentiment
0	1.0	apple progressive company pretend be best comp...	Negative
1	1.0	scumbag company decency integrity ethic refuse...	Negative
2	1.0	im foreign country apple pay decided block tra...	Negative
3	1.0	several charge account minor child game disput...	Negative
4	1.0	dropped watch repair jnder applecare july th j...	Negative
...	...	...	...
6671	5.0	good	Positive
6672	5.0	'all do' offering iphone apple phone run smoot...	Positive
6673	5.0	enough said already phone several review help ...	Positive
6674	5.0	first time using iphone personally really enjo...	Positive
6675	5.0	afford please go it it's device you sometimes ...	Positive

6676 rows × 3 columns

Figure 4.14 Combined reviews

4.3.3 Sentiment Analysis based on Star Based Analysis

Depending on the number of stars provided, classifies the sentiment as "Positive," "Negative," or "Neutral," and then updates a Data Frame with this information. Neutral emotion: If the number of stars is exactly 3.0, the emotion is regarded as 'Neutral'. A 3-star rating appears to be neutral—neither favourable nor negative. Negative emotion: A emotion is termed "Negative" if the number of stars is either 1.0 or 2.0. This illustrates how lower star ratings are typically linked to unfavourable emotions or unhappiness with a good or service. Positive emotion: A star rating of 4.0 or 5.0 indicates that the emotion is positive. This suggests that greater star ratings are linked to contentment or happy emotions.

- 1 or 2 stars corresponds to negative sentiment.
- 3 stars correspond to neutral sentiment.
- 5 stars corresponds to positive sentiment.

4.3.4 Stop Word Removal

Stop Word Recognition NLTK or custom lists created specifically for the context of Apple product evaluations are used to identify a list of stop words during the first preprocessing stage. The recognized stop words are filtered out once the reviews are collected from Trustpilot, Amazon, or other sources and tokenized into individual words. After the tokenization process, this removal ensures that frequently used terms with no significant significance for sentiment analysis are disregarded from further analysis. Eliminating stop words helps to make the data less noisy and dimensional. This simplifies the data and enables the algorithms to concentrate on pertinent words and phrases that express sentiment. Text preprocessing functions are defined to remove numbers, punctuation, URLs, and stop words, and convert the text to lowercase.

4.3.5 Lemmatization

The process of lemmatization involves changing a word to its fundamental or root form. Lemmatization considers the morphological examination of the words and assures that the root word is a part of the language, in contrast to stemming, which frequently removes prefixes or suffixes to identify the base form (sometimes resulting in non-real terms). Lemmatization is frequently used to guarantee that only meaningful words are processed once stop words have been eliminated. Lemmatization functions from libraries like NLTK or spaCy return words in the Apple product reviews to their original forms. It's necessary to understand the word's part of speech (such as a noun, verb, or adjective) for some lemmatization approaches. You may correctly translate the term to its proper lemma using this information. Lemmatization guarantees that diverse versions of a word are handled equally by reducing words to their most basic form, producing data that is more consistently formatted. By combining several inflected versions of a word into a single base form, lemmatization can minimize the dimensionality of the text data. This may improve the effectiveness of further analyses. Lemmatization frequently yields more accurate findings than stemming, which can result in more accurate sentiment analysis since it considers the real morphological structure of each word.

4.3.6 Data Normalization

Machine learning frequently uses the technique of data normalization to scale down features. This has the effect of assisting the learning algorithm in better comprehending the data, which frequently results in improved predicted performance.

Choosing a normalization Technique: Different normalizing techniques exist, including Min-Max Scaling, Z-score Standardization, and other scaling techniques. The decision is based on the project's particular requirements as well as the data's characteristics.

Use with Numerical Features: The numerical depictions of the text data have undergone normalization. These values could be standardized to put them on a consistent scale if features like TF-IDF or word embeddings are employed.

Training and Test Data Transformation: The same scaling parameters should be used for both training and test data transformation. This guarantees that the model reads the characteristics consistently.

Normalization is very crucial for this project:

Improved Convergence: If the features have comparable scales, several machine learning techniques, such as Logistic Regression and SVM, may converge more quickly. As a result, training might go more quickly.

Performance Improvement: Normalized data frequently makes it easier for the algorithm to recognize the underlying patterns in the data, potentially improving performance.

Consistency Across Features: When several numerical features are extracted from the text input, normalization makes sure that they all contribute equally to the learning process of the model and that no one feature dominates because of its size.

4.3.7 Data Splitting

Data splitting is dividing the dataset into different parts (typically into training, validation, and testing sets), to ensure that the model is learning the underlying patterns in the data rather than memorizing the specific data points. The cleaned and pre-processed dataset is split into training and test sets, with 80% for training and 20% for testing as shown in figure 4.15 the training set contains 5340 entries and testing contains 1336.

```
from sklearn.model_selection import train_test_split

df_train, df_test = train_test_split(df_cleaned, test_size=0.2, random_state=42)

print("Shape of df_train:", df_train.shape)
print("Shape of df_test:", df_test.shape)
```

```
Shape of df_train: (5340, 3)
Shape of df_test: (1336, 3)
```

Figure 4. 15 Data splitting.

4.4 MODELLING

The main stage of the sentiment analysis project for Apple product evaluations is model construction, during which machine learning algorithms are taught to recognize and forecast feelings based on the processed text data. A statistical method for binary classification, logistic regression is useful for determining the link between the characteristics of the text and its sentiment (positive or negative). Support Vector Classifier is a non-probabilistic binary classifier that finds the ideal hyperplane to divide classes and performs well with text data. Often employed in text classification applications like sentiment analysis, naive Bayes is a probabilistic classifier based on the Bayes theorem with high independence assumptions.

4.4.1 SVM

Finding the hyperplane that categorizes a dataset into classes most effectively is the goal of the supervised learning method SVM. The Support Vector Machine (SVM) family of algorithms includes SV. It looks for the optimum hyperplane to divide the data into several classes. A hyperplane is a straight line in a two-dimensional setting. A multidimensional space divides the classes into higher dimensions. Finding the best hyperplane to maximize the margin between the two classes is the goal of SV. The distance between the two classes' support vectors, or data points, that are closest to the hyperplane is the margin. A formulation of the issue is an optimization issue. Maximizing the margin while guaranteeing accurate classification of the data points is the objective. The use of quadratic programming can resolve this. SVM enables the conversion of the input space into a higher-dimensional space using various kernel functions. When the data is not linearly separable in the original space, this is especially helpful. Popular kernels include sigmoid, linear, polynomial, and radial basis functions (RBF). The regularization parameter (C) manages the trade-off between margin maximization and classification error minimization.

It would categorize the reviews into several sentiment categories, such as positive or negative, in the context of sentiment analysis. SVM has demonstrated its efficacy in text classification tasks, and its capacity to identify intricate class boundaries may be helpful for sentiment analysis. The model can detect both linear and non-linear correlations in the data since it may use different kernel functions. It works well in high-dimensional spaces and is therefore appropriate for text data, where there may be a significant number of characteristics.

4.4.2 Logistic Regression

The efficiency and readability of the logistic regression model make it appropriate for jobs like sentiment analysis. Although its simplicity may make it difficult to capture complicated interactions, it frequently offers a reliable baseline model. It provides a useful tool for comprehending and forecasting consumer feelings in the project for evaluating Apple product evaluations, with the added benefit of offering insights into the significance of certain words or phrases in deciding sentiment. A statistical technique for examining a dataset in which one or more independent variables affect an outcome is known as logistic regression. A dichotomous variable (one with just two potential outcomes) is used to measure the result. The two conclusions in the sentiment analysis context may be "positive sentiment" and "negative sentiment." The logistic function, an S-shaped curve that can map any real-valued integer between 0 and 1, is the foundation of logistic regression. The model employs the logistic function to identify the best-fitting curve that depicts the connection between the independent variables (features taken from the text) and the dependent binary variable (sentiment). Once trained, the model can analyze the characteristics of an unread review and forecast the likelihood that it would be favourable. If the model predicts a likelihood over a common threshold of 0.5, the review may be categorized as positive; otherwise, it might be labelled as negative.

4.4.3 Naïve Bayes

The multinomial version of the Naive Bayes algorithm, known as multinomial Naive Bayes, is suitable for features that reflect counts or frequency measurements. These characteristics in text categorization might include word counts or frequency in a document. The assumption that the characteristics are conditionally independent given the class is what gives the algorithm its "naive" quality. This indicates that given the sentiment class (such as positive or negative), the existence of one word does not alter the presence of another word.

Multinomial Naive Bayes attempts to determine the likelihood that a given review, given its words, belongs to a certain sentiment class. Using Multinomial Naive Bayes, it is possible to forecast sentiment in the context of the sentiment analysis project for Apple product reviews based on the frequency of terms in the reviews. When dealing with high-dimensional feature vectors and big vocabulary sizes, it can be quite useful. Text categorization using Multinomial Naive Bayes takes a probabilistic approach, making it appropriate for sentiment analysis applications. Its effectiveness and simplicity make it a useful tool for in-depth investigation, such as the assessment of consumer attitudes about Apple goods. Although the simplistic presumption of feature independence may not always hold, in actual use, the algorithm frequently delivers unexpectedly good results, making it a popular option for text categorization jobs.

4.5 OPTIMIZATION

The process of optimization entails selecting the best solution, or set of solutions, from a variety of viable options. This frequently requires figuring out the machine learning parameter values that reduce a loss function or boost accuracy or probability. Optimization, which covers everything from model training to hyperparameter selection, is primarily reliant on the project for sentiment analysis of customer reviews of Apple products. By carefully examining the search space and iteratively improving answers, optimization procedures make sure that the models are both accurate and effective. This results in more insightful research and reliable projections, which aid in understanding consumer attitudes about Apple goods.

4.5.1 Genetic algorithm

A genetic algorithm can find an approximate solution to optimization and search problems. It includes a developing population of individuals over many generations and is based on the ideas of genetics and natural selection.

- Initialization: A population of candidate solutions (individuals) is created. These can represent possible parameter values for a machine-learning model.
- Selection: Individuals are selected to be parents based on a fitness function, which evaluates how well the solution performs the task at hand.
- Crossover (Recombination): Pairs of parents are mated to create offspring. This often involves mixing the genes of the parents in some way.
- Mutation: Some genes in the offspring are altered. This introduces variation and can help escape local minima.
- Replacement: New offspring are added to the population, and less fit individuals may be removed.
- Termination: The algorithm terminates when a stopping criterion is met, such as a maximum number of generations, or a satisfactory fitness level has been reached.

GA provides a flexible and frequently effective approach for resolving challenging problems by replicating the process of natural evolution. Its usage in the project permits the exploration of a range of possibilities, which can improve model performance and show how users view Apple products.

4.6 EVALUATION METRIC

The assessment procedure gives a comprehensive insight into the efficiency of the project-developed methodologies. The right evaluation metrics are an essential part of every machine-learning effort. These metrics provide a detailed picture of how the models are doing in the sentiment analysis of customer reviews of Apple products, allowing for further fine-tuning and advancement. By aligning the evaluation criteria with the unique characteristics and objectives of the inquiry, you can ensure that the models are not only accurate but also relevant and valuable in their results.

4.6.1 Confusion Matrix

A confusion matrix is a tabular representation commonly utilized to evaluate classification problems. As illustrated in Figure 4.6, this matrix encompasses four key values: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

- True Positive (TP): This value denotes the number of instances correctly classified as belonging to the target class.

- True Negative (TN): This count represents the instances that aren't part of the target class and are correctly identified as such.
- False Positive (FP): This refers to the instances wrongly classified as part of the target class even though they are not.
- False Negative (FN): This value indicates the instances that are indeed part of the target class but are mistakenly classified otherwise.

The confusion matrix is instrumental in deriving various evaluation metrics to gauge a model's performance, including:

Specificity: How well the model identifies true negatives.

Sensitivity: The efficiency of the model in recognizing true positives.

Precision: The correctness of positive predictions.

Accuracy: The overall effectiveness of the model in classification.

The F1-score is another essential metric calculated from the confusion matrix. It represents the harmonic mean of precision and recall and serves as a balance between the two. The value of this metric ranges from 0 to 1, with 1 being perfect precision and recall. It's a significant indicator of a test's precision and is especially useful in scenarios where the distribution of classes is uneven.

$$\boldsymbol{Sensitivity\ (Recall) = \frac{TP}{TP + FN}}$$

$$\boldsymbol{Precision\ = \frac{TP}{TP + FP}}$$

$$\boldsymbol{Accuracy = \frac{TN + TP}{TN + FP + FP + TP}}$$

$$\boldsymbol{F1\ Score = \frac{2 * Precsion * Recall}{Precision + Recall}}$$

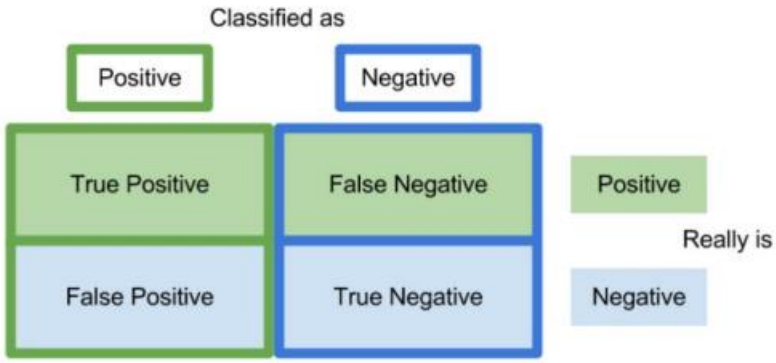

Figure 4.16 Confusion Matrix

CHAPTER 5

RESULT ANALYSIS

5.1 INTRODUCTION

This chapter offers a comprehensive analysis of the conclusions drawn from the numerous computer experiments conducted during this inquiry. The innovative custom review prediction approach, the exploration of genetic programming techniques, and performance evaluation of machine learning models are the three key areas of focus.

5.2 IMPLEMENTATION DETAILS

Python coding is used to implement the model in Colab. The Python emoji package is a helpful resource for managing emoticons in text data. The "review_text" column of the "combined_df" no longer contains emojis. The Python utility VADER (Valence Aware Dictionary and sEntiment Reasoner) from the NLTK package analyzes sentiment intensity. The social media text analysis software VADER is especially useful for brief, casual texts. VADER is an effective tool for sentiment analysis, especially when used on messages that include slang, emoticons, and other colloquial language. It is ready to use straight out of the box because it is trained, which means it doesn't need any more training data.

A free tool for developing and utilizing evolutionary algorithms is the Distributed Evolutionary Algorithms in Python (DEAP) module. DEAP is a very flexible library that uses Genetic Algorithms, Genetic Programming, Evolution Strategies, and other evolutionary techniques. It allows for parallel processing and is applicable to a variety of problem domains. The TPOT installation and genetic programming implementation use DEAP.

To improve machine learning pipelines, TPOT (Tree-based Pipeline Optimization Tool), an automated machine learning tool, uses genetic programming. It effectively automates the steps required in selecting the best machine learning model, performing preprocessing operations, adjusting hyperparameters, and constructing the full ML pipeline for a specific problem. TPOT is a powerful tool that can drastically cut the time and effort needed to develop machine learning models. However, the nature of the genetic programming technique can make it computationally expensive, especially for large datasets and difficult problems. It is possible to reduce the cost of computation by carefully selecting the population size, number of generations, and other factors. The process of transforming text reviews into numerical features uses the TF-IDF vectorizer. The acronym is TF-IDF, or term frequency-inverse document frequency. It is a metric that assesses the significance of a word in a text pertaining to a corpus, or group, of documents. TF-IDF is often used for text analysis and natural language processing applications to translate texts into meaningful numerical representations that may be fed into machine learning models. A phrase's term frequency (TF) indication shows how frequently it appears in a document. It is calculated by dividing the overall word count of a text by the frequency of a particular word in that text. The Inverse Document Frequency (IDF) statistic evaluates the importance of a phrase throughout the whole corpus. You may calculate it by dividing the logarithm of the total number of documents in the corpus by the total number of documents that include the phrase.

5.3 RESULT ANALYSIS OF ML MODELS ON POLARITY BASED SENTIMENT ANALYSIS

The ML models are executed on the data frame formed based on the polarity-based sentiment classification.

SVM - SVM model with a linear kernel. The class_weight='balanced' argument adjusts the weights inversely proportional to class frequencies, which helps in handling class imbalance. The random_state=42 argument ensures the reproducibility of the results. The performance metric is shown in Figure 5.1 and shows average accuracy of 0.85. The ROC curve is shown in Figure 5.2, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.3.

```
SVM Classification Report:
              precision    recall  f1-score   support

         1.0       0.92      0.90      0.91       791
         2.0       0.24      0.25      0.25       118
         5.0       0.90      0.91      0.90       533

    accuracy                           0.85      1442
   macro avg       0.69      0.69      0.69      1442
weighted avg       0.85      0.85      0.85      1442
```

Figure 5.1 Performance metric of SVM

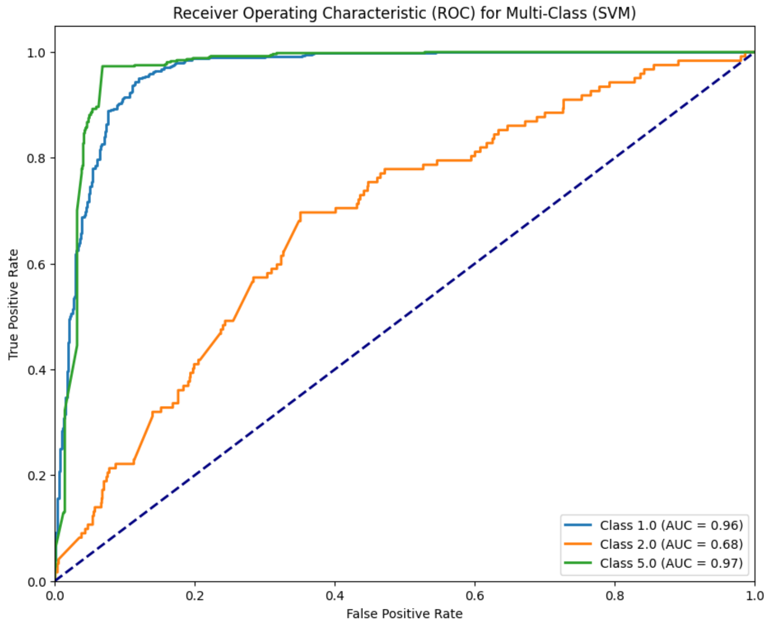

Figure 5.2 ROC of SVM

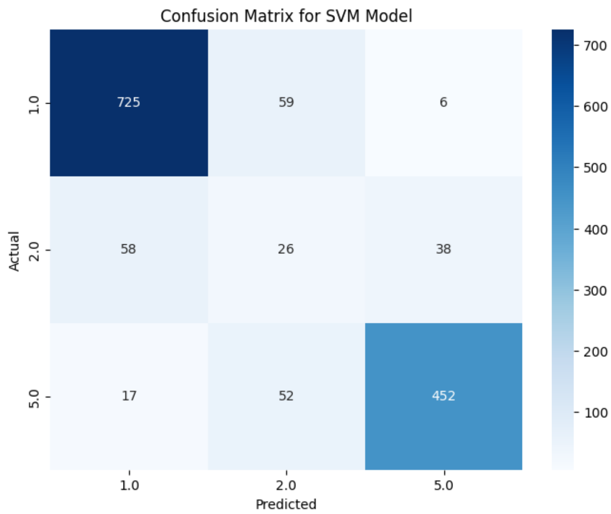

Figure 5.3 Confusion Matrix of SVM

Logistic regression - A standard text classification pipeline using TF-IDF features, and Logistic Regression is implemented. It takes raw review text, transforms it into numerical features using TF-IDF, and then uses Logistic Regression to predict the star ratings. The model is then evaluated on a separate test set to assess its performance. The logistic regression model with class_weight='balanced' argument helps the model pay equal attention to all classes in the dataset, which can be useful if there is an imbalance in the distribution of star ratings. The random_state=42 ensures that the results are reproducible. The performance metric is shown in Figure 5.4 and shows average accuracy of 0.86. The ROC curve is shown in Figure 5.5, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.6.

```
Logistic Regression Classification Report:
              precision    recall  f1-score   support

         1.0       0.92      0.92      0.92       791
         2.0       0.29      0.29      0.29       118
         5.0       0.90      0.89      0.90       533

    accuracy                           0.86      1442
   macro avg       0.70      0.70      0.70      1442
weighted avg       0.86      0.86      0.86      1442
```

Figure 5.4 Performance metric logistic regression

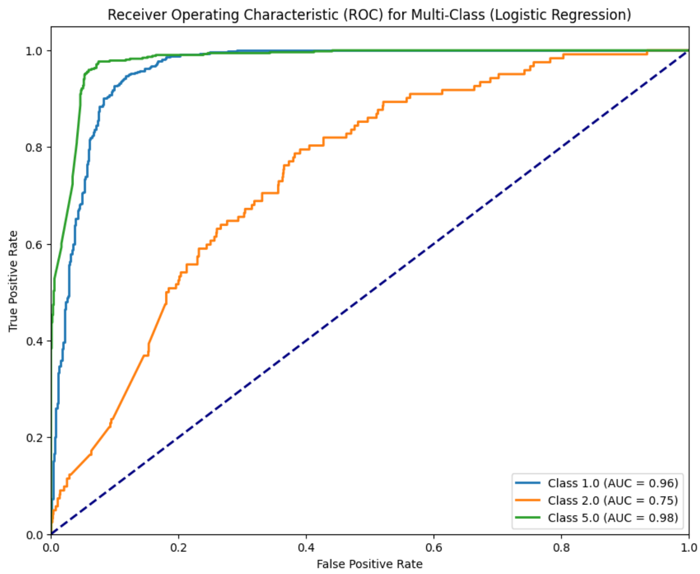

Figure 5.5 ROC of logistic regression

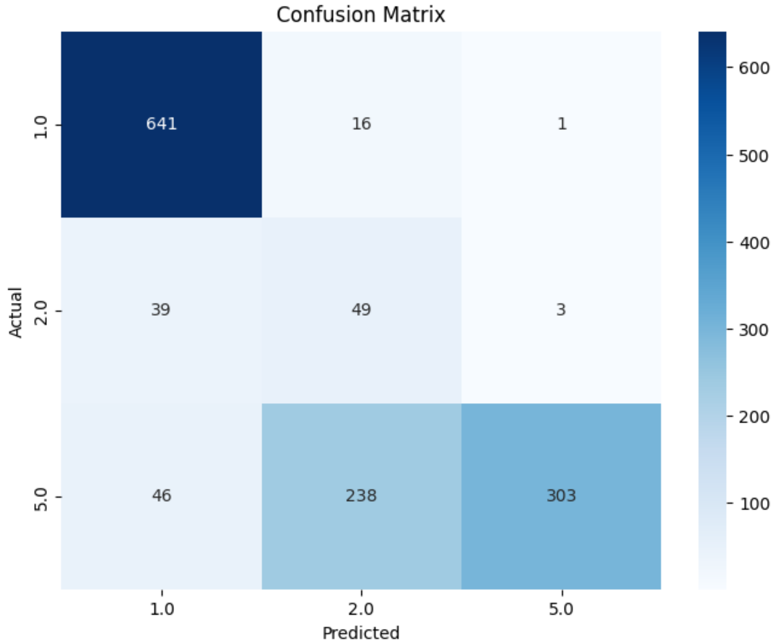

Figure 5.6 Confusion Matrix of logistic regression

Naïve Bayes – The Multinomial Naive Bayes algorithm is particularly suitable for text classification as it handles discrete data and can deal with the type of data produced by the TF-IDF transformation. By specifying class priors, it allows the model to consider any prior knowledge or specific weighting between the classes, providing flexibility in handling imbalanced datasets. The naive Bayes model with class_prior argument is set to [1,10,1], meaning that custom weights are provided for the classes, which can help if there's an imbalance in the class distribution. The alpha parameter is set to 1.0, adding a Laplace smoothing factor to avoid zero probabilities for unseen words. The performance metric is shown in Figure 5.7 and shows average accuracy of 0.60. The ROC curve is shown in Figure 5.8, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.9.

```
Naive Bayes Classification Report:
              precision    recall  f1-score   support

         1.0       0.93      0.60      0.73       791
         2.0       0.12      0.60      0.20       118
         5.0       0.91      0.58      0.71       533

    accuracy                           0.60      1442
   macro avg       0.66      0.60      0.55      1442
weighted avg       0.86      0.60      0.68      1442
```

Figure 5.7 Performance metric naïve bayes

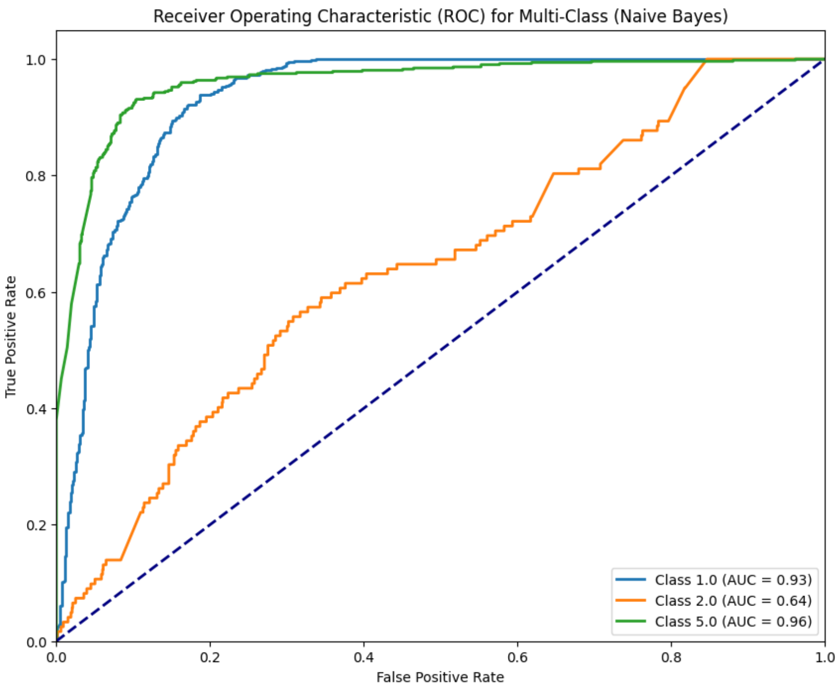

Figure 5.8 ROC of naïve bayes

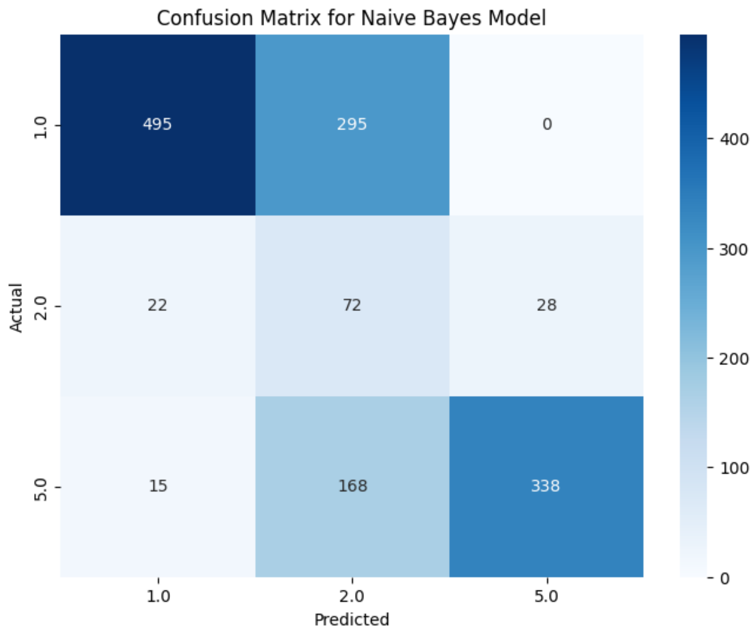

Figure 5.9 Confusion Matrix of naïve bayes

TPOT is used to automate the hyperparameter tuning and model selection process for an SVM classifier. By providing a specific configuration for TPOT to work with, the search space is narrowed down to focus on specific hyperparameters of interest. TPOT then searches this space to find the best model and preprocessing pipeline for the given task. TPOT Classifier with generations=5: Number of iterations to run the optimization process, population_size=20: Number of individuals (pipelines) in the population at each generation, verbosity=2: Controls the verbosity of the output, displaying progress information. random_state=42: Ensures reproducibility by setting seed for random operations. n_jobs=-1: Utilizes all available cores for parallel computation. config_dict: Specifies the model and hyperparameters to optimize, in this case, SVM with alpha values [0.01, 0.1, 1.0] and fit_prior=False. The classification report of the performance, including precision, recall, F1-score, and support for each class of star rating is given in Figure 5.10 with an accuracy of 0.89. The ROC curve is shown in Figure 5.11, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.12. The comparison of the classifiers in polarity score-based sentiment analysis is given in Table 1 and it shows the optimized SVM gives best accuracy of 0.89.

```
Best pipeline: SVC(input_matrix, C=1, gamma=auto, kernel=linear)
TPOT Classification Report:
              precision    recall  f1-score   support

         1.0       0.89      0.99      0.94       791
         2.0       0.67      0.03      0.06       118
         5.0       0.89      0.93      0.91       533

    accuracy                           0.89      1442
   macro avg       0.82      0.65      0.64      1442
weighted avg       0.87      0.89      0.86      1442
```

Figure 5.10 Performance metric SVM with optimization

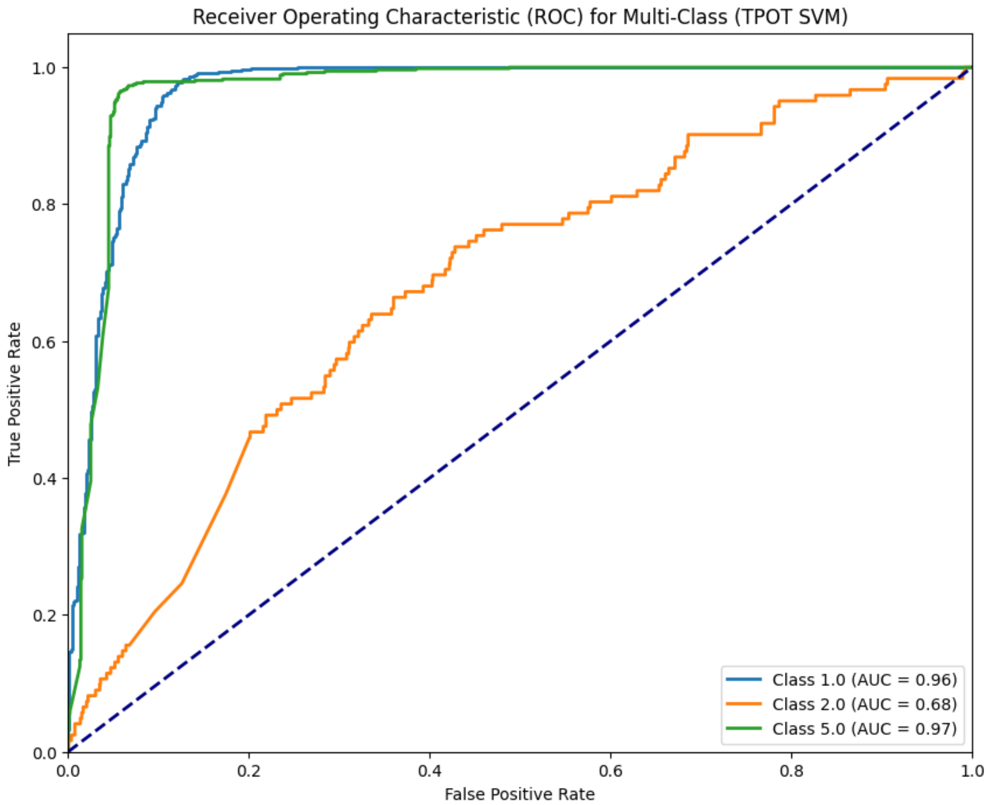

Figure 5.11 ROC SVM with optimization

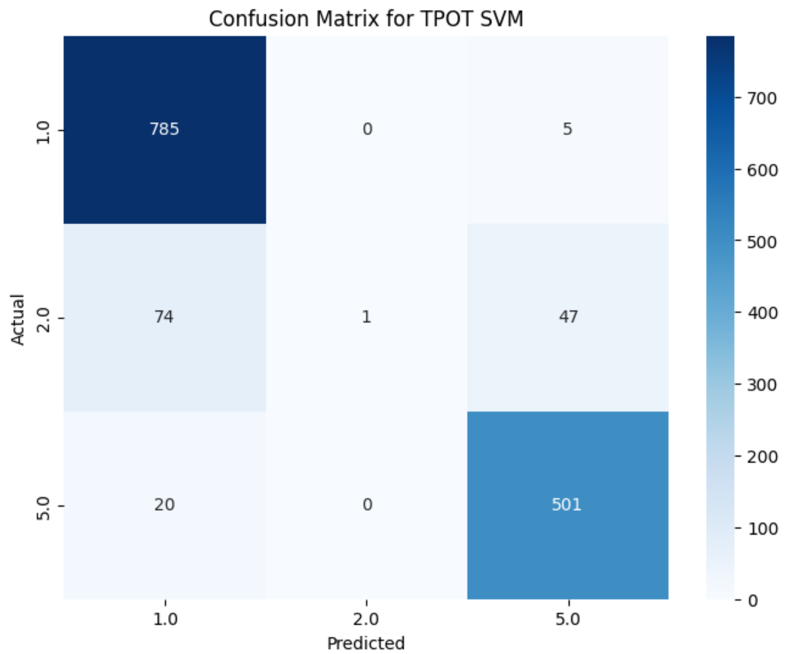

Figure 5.12 Confusion Metric SVM with optimization

Table 1 Performance Comparison of ML Models using polarity score.

Models	**Accuracy**
SVM	0.85
Logistic Regression	0.86
Multinomial Naive Bayes	0.60
SVM + TPOT Optimizer	0.89

5.4 RESULT ANALYSIS OF ML MODELS ON STARS BASED SENTIMENT ANALYSIS

The ML models are executed on the data frame formed based on star-based sentiment classification. The models are built with same configuration as in section5.3 only the data frame is changed.

Logistic regression - The performance metric is shown in Figure 5.13 and shows average accuracy of 0.89. The ROC curve is shown in Figure 5.14, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.15.

```
Logistic Regression Classification Report:
              precision    recall  f1-score   support

         1.0       0.91      0.91      0.91       892
         2.0       0.12      0.21      0.16        57
         5.0       0.98      0.92      0.95       540

    accuracy                           0.89      1489
   macro avg       0.67      0.68      0.67      1489
weighted avg       0.91      0.89      0.90      1489
```

Figure 5.13 Performance of Logistic Regression

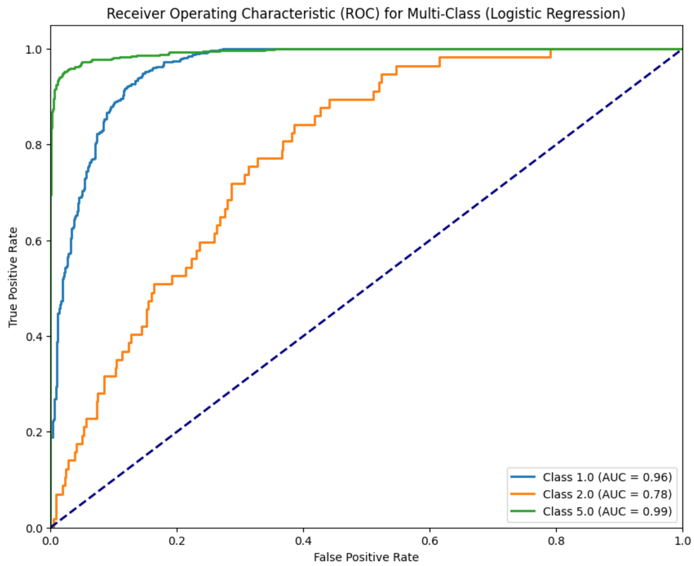

Figure 5.14 ROC of Logistic Regression

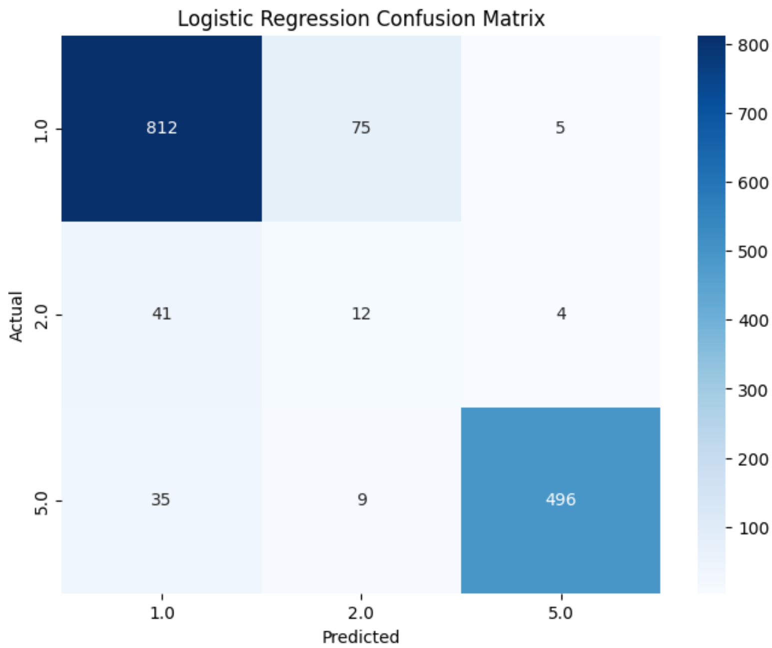

Figure 5.15Confusion Metric Logistic Regression

SVM -The performance metric is shown in Figure 5.16 and shows average accuracy of 0.90. The ROC curve is shown in Figure 5.17, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.18.

```
SVM Classification Report:
              precision    recall  f1-score   support

         1.0       0.92      0.92      0.92       892
         2.0       0.14      0.19      0.16        57
         5.0       0.98      0.94      0.96       540

    accuracy                           0.90      1489
   macro avg       0.68      0.68      0.68      1489
weighted avg       0.91      0.90      0.91      1489
```

Figure 5.16 Performance metric of SVM

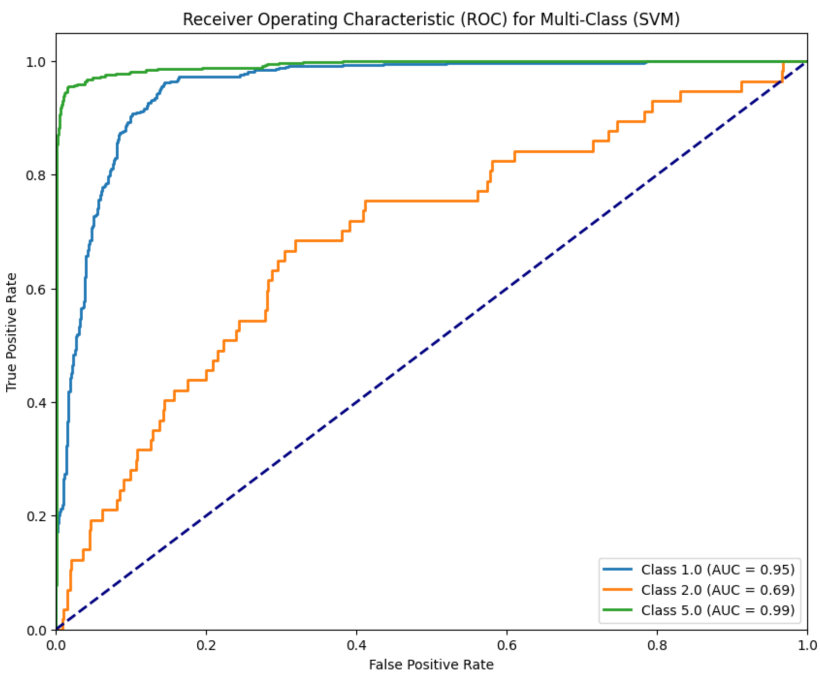

Figure 5.17 ROC of SVM

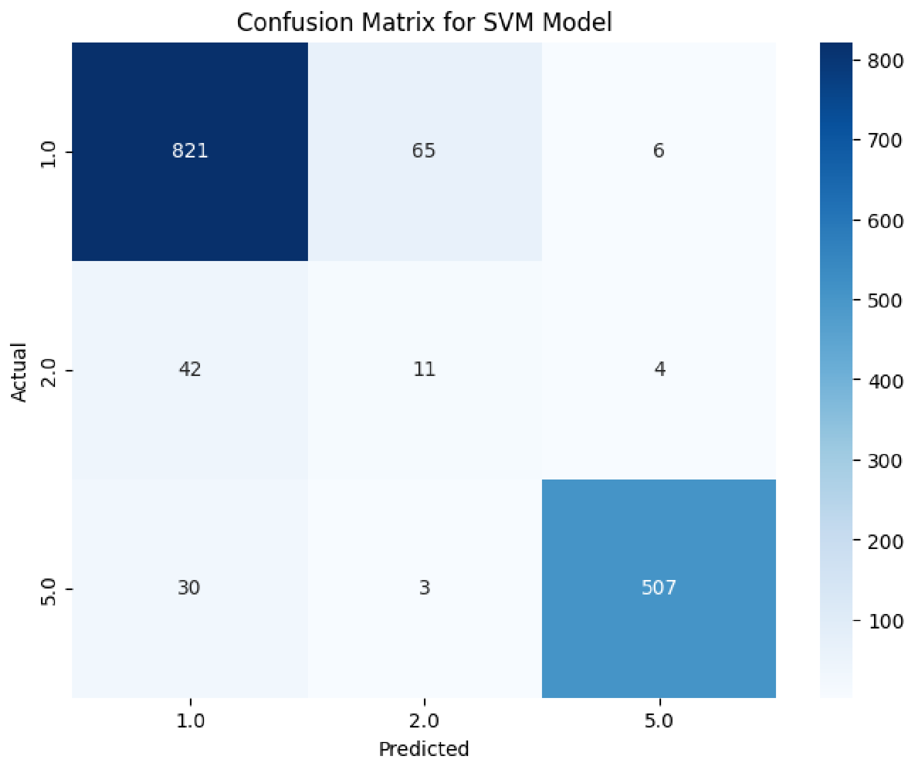

Figure 5.18 Confusion Metric SVM

Naïve Bayes – The performance metric is shown in Figure 5.19 and shows average accuracy of 0.65. The ROC curve is shown in Figure 5.20, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.21.

Naive Bayes Classification Report:

	precision	recall	f1-score	support
1.0	0.91	0.64	0.75	892
2.0	0.08	0.66	0.14	62
5.0	0.99	0.66	0.80	553
accuracy			0.65	1507
macro avg	0.66	0.66	0.56	1507
weighted avg	0.91	0.65	0.74	1507

Figure 5.19 Performance metric of Naïve Bayes

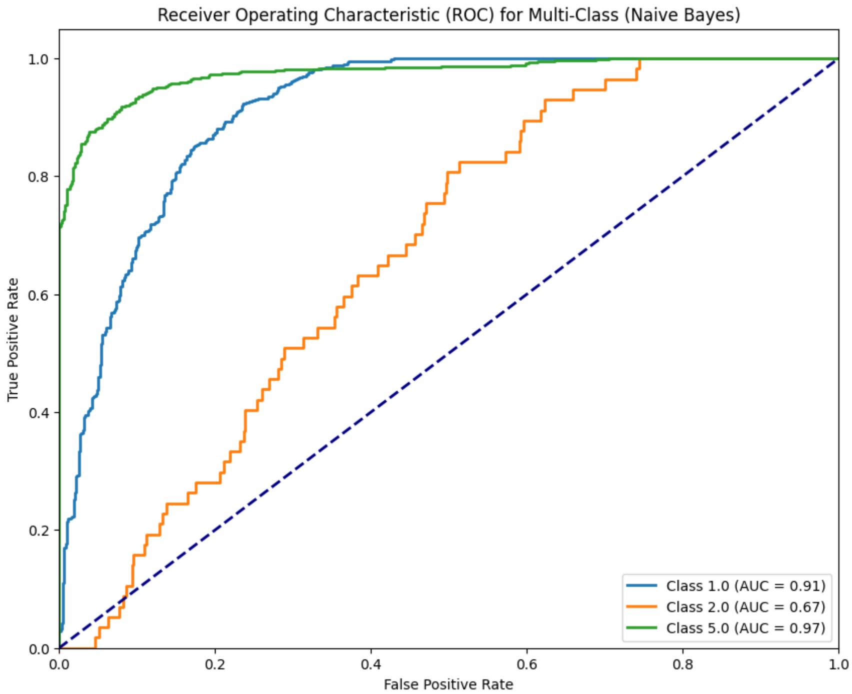

Figure 5. 20 ROC of Naïve Bayes

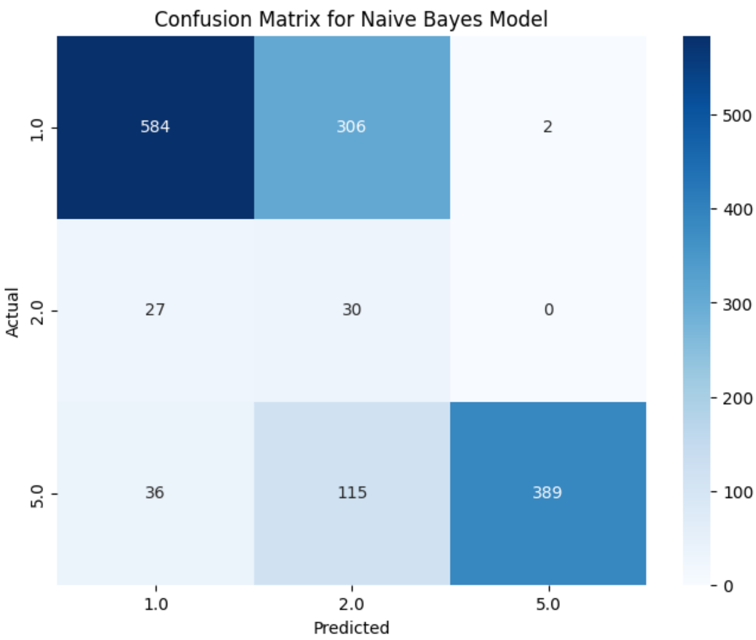

Figure 5. 21 Confusion Metric of Naïve Bayes

The classification report of the performance, including precision, recall, F1-score, and support for each class of star rating is given in Figure 5.22 with an accuracy of 0.92. The ROC curve is shown in Figure 5.23, and it shows the AUC for all three class and confusion matrix is shown in Figure 5.24. The comparison of the classifiers in polarity score-based sentiment analysis is given in Table 2 and it shows the optimized SVM gives best accuracy of 0.92.

```
Best pipeline: SVC(input_matrix, C=1, gamma=auto, kernel=linear)
TPOT Classification Report:
              precision    recall  f1-score   support

         1.0       0.90      0.98      0.94       892
         2.0       0.00      0.00      0.00        62
         5.0       0.96      0.93      0.94       553

    accuracy                           0.92      1507
   macro avg       0.62      0.64      0.63      1507
weighted avg       0.89      0.92      0.90      1507
```

Figure 5.22 Performance of SVM with Optimization

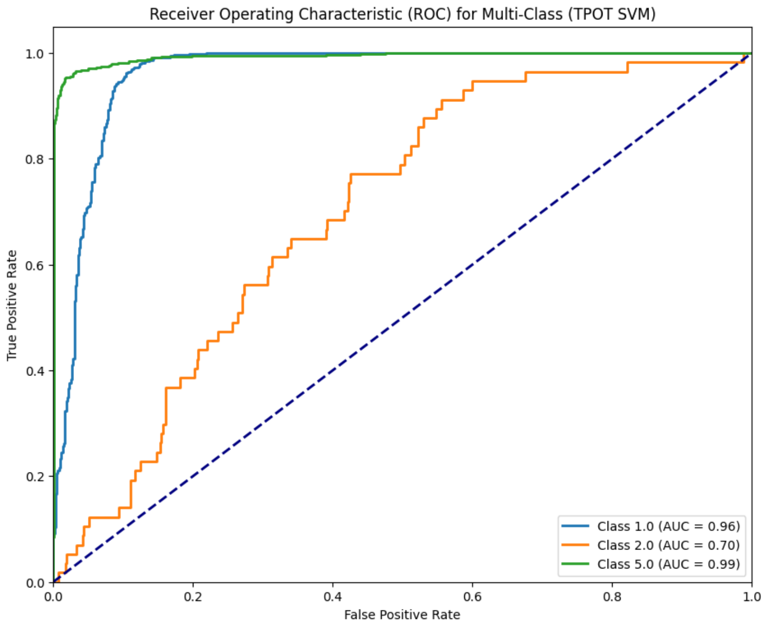

Figure 5.23 ROC of SVM with Optimization

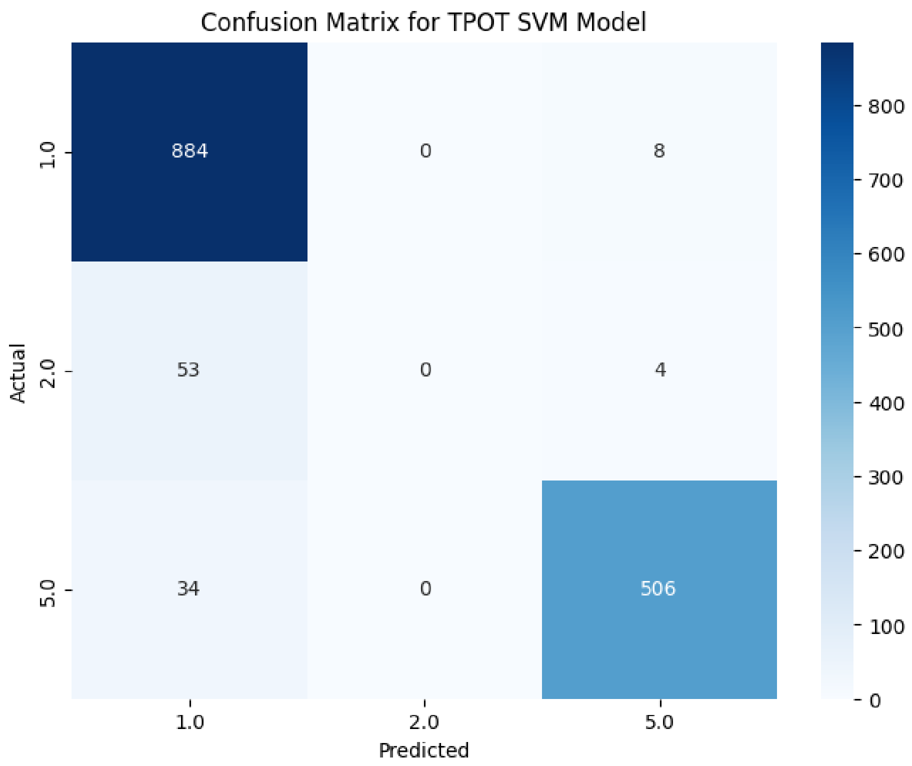

Figure 5.24 Confusion Metric of SVM with Optimization

Table 2 Performance Comparison of ML Models using star-based sentiment analysis

Models	**Accuracy**
SVM	0.90
Logistic Regression	0.89
Multinomial Naive Bayes	0.65
SVM+ TPOT Optimizer	0.92

5.5 CUSTOM REVIEW PREDICTION

A set of custom review texts and applies a previously trained sentiment analysis model to predict their star ratings. It does this by transforming the texts into the same TF-IDF feature space used during training and then applying the best model pipeline found by TPOT to make the predictions. This provides a way to see how the model would rate new, unseen reviews. The custom review prediction using an optimized naïve Bayes model is shown in Figure 5.25

```
Custom Review Predictions:
Review: This product is amazing! I love it.
Predicted Rating: 5.0
Review: Terrible experience, never buying from this company again.
Predicted Rating: 1.0
Review: Good product, but the customer service needs improvement.
Predicted Rating: 5.0
Review: The item was average, nothing special.
Predicted Rating: 1.0
Review: The best purchase I ever made! Highly recommend it.
Predicted Rating: 5.0
Review: Not worth the price, very disappointed.
Predicted Rating: 1.0
Review: I had high expectations, but it did not meet them.
Predicted Rating: 1.0
Review: not a great service,not Happy with my purchase.
Predicted Rating: 5.0
```

Figure 5.25 Custom review prediction by the optimized model

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 LIMITATIONS

Reliance on Web Structure: The scraping approach may malfunction if websites alter their structure or style, necessitating ongoing upgrades and maintenance.

Legal and Ethical Issues: As was already said, scraping must be done following legal and ethical guidelines, which may restrict access to data.

Data Bias and Quality: Depending on the sources, the statistics may not accurately reflect the opinions of the full clientele. Biases in the study might result from missing or uneven data. Unfiltered data may contain spam or fraudulent reviews, which might have an impact on the analysis's accuracy.

Model Restrictions: Under- or Overfitting If not correctly calibrated, algorithms like logistic regression, SVM, and Naive Bayes may result in overfitting or underfitting. Complex models may be difficult to grasp, which makes it more challenging to comprehend how decisions are made.

Limitations specific to genetic algorithms: Genetic algorithms have the potential to be computationally expensive, which makes them potentially unsuitable for real-time applications.

Local Optima: Depending on the implementation, evolutionary algorithms could reach local optimum solutions rather than the ideal ones overall.

Lack of Multilingual Support: If the analysis is confined to one language, it may not account for cultural and linguistic nuances or attitudes from speakers of languages other than English.

Sentiment analysis methods may not completely understand the cultural or contextual subtleties in the reviews due to context sensitivity.

Scalability problems: Handling enormous Amounts of Data: Using enormous amounts of data may be difficult, depending on the infrastructure and techniques employed.

External influences and modifications: Product Line Alterations: Over time, alterations to Apple's product line or the market may have an impact on the analysis's applicability and usefulness.

Considerations for Ethics in Analysis: Potential Bias: If not handled effectively, biases in data collecting and modelling might produce unfair or biased results.

6.2 FUTURE SCOPE

Reflecting the dynamic and complex nature of sentiment analysis, web scraping, and machine learning, the project's methods and conclusions create a foundation that may be expanded upon in a variety of ways. The future scope of this study has the potential for significant contributions to both academia and industry by following developing trends, resolving present limits, and extending into new fields. It represents the fusion of technology, ethics, commerce, and creativity and offers a wide range of opportunities for further research and development. Expansion to Other Brands and Products: The established models are more adaptable and applicable when used to assess feelings for other technology brands or different business sectors. Integration with Real-Time Analysis: By putting real-time data scraping and analysis into practice, companies may use these insights for prompt decision-making, trend forecasting, and dynamic consumer interaction. Including Multilingual study: Including evaluations in other languages as part of the study might give it a more global viewpoint and cater to different markets. Creation of Customized Recommendation Systems: By using the insights from sentiment analysis, customized recommendation systems that improve user experience and boost sales may be made. Improving Ethical Web Scraping Techniques: Upcoming work may focus on creating more ethically sound and transparent web scraping methods to support responsible data science operations. Examining Other Optimization Techniques: The project's use of genetic algorithms for optimization might serve as a springboard for investigating other cutting-edge optimization strategies that would improve model performance. Application in Social Media Monitoring: The approach may be modified to track opinions on products or brands on social media, enabling more flexible public relations plans. Integration with Human-Centered Design: By incorporating user preferences and attitudes, the work might influence human-centred design methodologies, enhancing product design and user pleasure. Addressing

Bias and Fairness: Upcoming studies may concentrate on locating and minimizing biases in sentiment analysis to support impartial and fair modelling. Adherence to Evolving Legal Framework: Future work must be flexible and sensitive to these developments, guaranteeing continuing compliance, as the legal landscapes around data privacy and web scraping advance. Increasing Accessibility for Small Firms: Making these advanced analytical skills more widely available to small and medium-sized firms by adapting the tools and procedures for their usage.

6.3 CONCLUSION

A comprehensive exploration of machine learning techniques, including traditional models like Logistic Regression, Support Vector Machines, and Naive Bayes, and cutting-edge methods like genetic programming and custom review prediction. Through the rigorous application of data preprocessing, feature engineering, model training, and hyperparameter tuning, the study provided valuable insights into the Apple review sentiment analysis. The comparison of the classifiers in polarity score-based sentiment analysis and the star-based analysis was done and it was found that the star-based analysis gave maximum accuracy 0.92 for SVM classifier with genetic programming optimization using TPOT.

REFERENCES

Anjaria, M., & Guddeti, R. M. R. (2014). A novel sentiment analysis of social networks using supervised learning. *Social Network Analysis and Mining*, *4*(1), 181. https://doi.org/10.1007/s13278-014-0181-9

Barik, K., Misra, S., Ray, A. K., & Bokolo, A. (2023). LSTM-DGWO-Based Sentiment Analysis Framework for Analyzing Online Customer Reviews. *Computational Intelligence and Neuroscience*, *2023*, 6348831. https://doi.org/10.1155/2023/6348831

Building Trustworthy and Ethical AI is everyone's responsibility. (n.d.). Retrieved November 24, 2022, from https://www.tavant.com/blog/building-trustworthy-and-ethical-ai-everyones-responsibility

Castillo, M. J., & Taherdoost, H. (2023). The Impact of AI Technologies on E-Business. *Encyclopedia*, *3*(1), Article 1. https://doi.org/10.3390/encyclopedia3010009

Chaki, S. (2015). Components of Enterprise Information Management. In S. Chaki (Ed.), *Enterprise Information Management in Practice: Managing Data and Leveraging Profits in Today's Complex Business Environment* (pp. 15–24). Apress. https://doi.org/10.1007/978-1-4842-1218-9_3

Chinnalagu, A., & Durairaj, A. K. (2021). Context-based sentiment analysis on customer reviews using machine learning linear models. *PeerJ Computer Science*, *7*, e813. https://doi.org/10.7717/peerj-cs.813

Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *First International Conference on Information Technology and Quantitative Management*, *17*, 26–32. https://doi.org/10.1016/j.procs.2013.05.005

Indrayuni, E., & Nurhadi, A. (2020). Optimizing Genetic Algorithms for Sentiment Analysis of Apple Product Reviews Using SVM. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *4*(2), Article 2. https://doi.org/10.33395/sinkron.v4i2.10549

Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment Analysis on Product Reviews Using Machine Learning Techniques. In P. K. Mallick, V. E. Balas, A. K. Bhoi, & A. F. Zobaa (Eds.), *Cognitive Informatics and Soft Computing* (pp. 639–647). Springer. https://doi.org/10.1007/978-981-13-0617-4_61

Kahaner, L. (1997). *Competitive Intelligence: How To Gather Analyze And Use Information To Move Your Business To The Top*. Simon and Schuster.

Kanakaraj, M., & Guddeti, R. M. R. (2015). NLP based sentiment analysis on Twitter data using ensemble classifiers. *2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, 1–5. https://doi.org/10.1109/ICSCN.2015.7219856

Kang, M. Y., Choi, Y., & Choi, J. (2019). The effect of celebrity endorsement on sustainable firm value: Evidence from the Korean telecommunications industry. *International Journal of Advertising*, *38*(4), 563–576. https://doi.org/10.1080/02650487.2019.1601910

Kaur, D. (2017). *Sentimental Analysis on Apple Tweets with Machine Learning Technique*. *7*(9).

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, *52*(4), 357–365. https://doi.org/10.1016/j.bushor.2009.03.002

Naseem, S., Mahmood, T., Asif, M., Rashid, J., Umair, M., & Shah, M. (2021). Survey on Sentiment Analysis of User Reviews. *2021 International Conference on Innovative Computing (ICIC)*, 1–6. https://doi.org/10.1109/ICIC53490.2021.9693029

Nhlabano, V. V., & Lutu, P. E. N. (2018). Impact of Text Pre-Processing on the Performance of Sentiment Analysis Models for Social Media Data. *2018 International Conference* *on Advances in Big Data, Computing and Data Communication Systems (IcABCD)*, 1–6. https://doi.org/10.1109/ICABCD.2018.8465135

Palomino, M. A., & Aider, F. (2022). Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis. *Applied Sciences*, *12*(17), Article 17. https://doi.org/10.3390/app12178765

Sangani, C., & Ananthanarayanan, S. (n.d.). *Sentiment Analysis of App Store Reviews*.

Suganya, E., & Vijayarani, S. (2020). Sentiment Analysis for Scraping of Product Reviews from Multiple Web Pages Using Machine Learning Algorithms. In A. Abraham, A. K. Cherukuri, P. Melin, & N. Gandhi (Eds.), *Intelligent Systems Design and Applications* (pp. 677–685). Springer International Publishing. https://doi.org/10.1007/978-3-030-16660-1_66

Taherdoost, H., & Madanchian, M. (2023). Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research. *Computers*, *12*(2), Article 2. https://doi.org/10.3390/computers12020037

Tang, H., Tan, S., & Cheng, X. (2009). A survey on sentiment detection of reviews. *Expert Systems with Applications*, *36*(7), 10760–10773. https://doi.org/10.1016/j.eswa.2009.02.063

Tuan, L. T. (2016). Organisational ambidexterity and supply chain agility: The mediating role of external knowledge sharing and moderating role of competitive intelligence. *International Journal of Logistics Research and Applications*, *19*(6), 583–603. https://doi.org/10.1080/13675567.2015.1137278

Yiran, Y., & Srivastava, S. (2019). Aspect-based Sentiment Analysis on mobile phone reviews with LDA. *Proceedings of the 2019 4th International Conference on Machine Learning Technologies*, 101–105. https://doi.org/10.1145/3340997.3341012