A Project report on

**Cashless Payment Challenges And Consumer Emotions: Review Based Empirical Study From India**

Submitted in partial fulfilment of requirement

for the award of the degree

MASTER of COMPUTER APPLICATIONS

Of

Visvesvaraya Technological University, Belagavi

By

# KEVIN SEQUEIRA

# 4NM21MC038

2021-2023



(An off-Campus Institution of NITTE (DEEMED TO BE UNIVERSITY), MANGALORE)

Nitte Mahalinga Adyanthya Memorial Institute of Technology

Nitte – 574110, Karkala, Udupi District

A Project Report on

**Cashless Payment Challenges And Consumer Emotions: Review Based Empirical Study From India**

Submitted in partial fulfillment of requirement

for the award of the degree

MASTER of COMPUTER APPLICATIONS

of

Visvesvaraya Technological University, Belagavi

By

# KEVIN SEQUEIRA

# 4NM21MC038

2021-2023

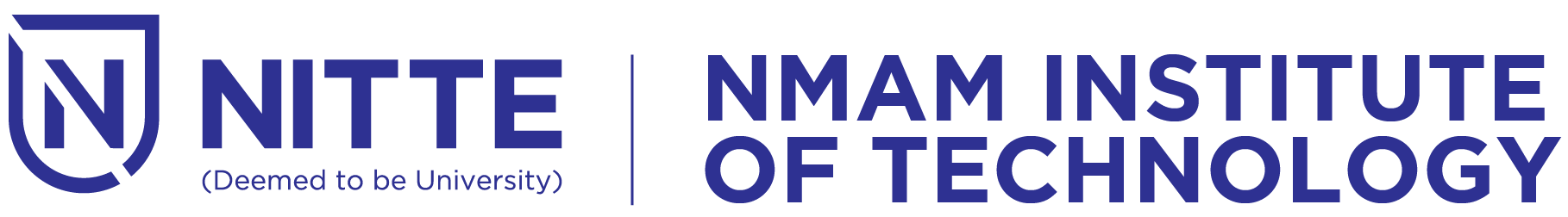
Under the guidance of

Dr. Mangala P Shetty

Professor

Department of MCA

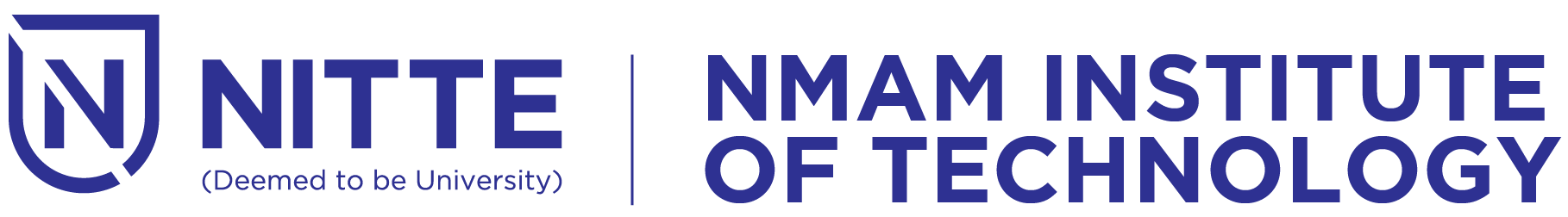
NMAM Institute of Technology, Nitte



(An off-Campus Institution of NITTE (DEEMED TO BE UNIVERSITY), MANGALORE)

Nitte Mahalinga Adyanthaya Memorial Institute of Technology

Nitte – 574110, Karkala, Udupi District



(An off-Campus Institution of NITTE (DEEMED TO BE UNIVERSITY), MANGALORE)

**CERTIFICATE**

This is to certify that the project entitled

**Cashless Payment Challenges And Consumer Emotions: Review Based Empirical Study From India**

Submitted in partial fulfillment of requirement

for the award of the degree

MASTER of COMPUTER APPLICATIONS

is a result of the bonafide work carried out by

# KEVIN SEQUEIRA

# 4NM21MC038

During the academic year 2022-2023

Dr. Mangala P Shetty

Professor

Department of MCA

NMAM Institute of Technology, Nitte

Dr. Surendra Shetty

Professor & Head

Department of MCA

NMAM Institute of Technology, Nitte

Dr. Niranjan N Chiplunkar

Principal

NMAM. Institute of Technology, Nitte

Semester End Viva Voce Examinations

Name of the examiners

1. Signature with date

2.

**ACKNOWLEDGMENT**

I would like to extend my heartfelt gratitude to all those who have contributed to the successful completion of this project. Without their support and guidance, this endeavor would not have been possible.

Firstly, I am grateful to Dr. Niranjan N Chiplunkar, Principal of NMAM.I.T., Nitte, for providing me with the opportunity to undertake this internship and for his constant encouragement throughout the project.

I would also like to express my sincere appreciation to Dr. Surendra Shetty, Professor & Head of the Department of MCA, NMAM.I.T., Nitte, for his valuable support, insightful suggestions, and continuous guidance during the course of this project.

I am immensely grateful to my mentor and project guide, Dr. Mangala P Shetty, Associate Professor in the Department of MCA, NMAM.I.T., Nitte, for his unwavering support, expert guidance, and constant supervision. His valuable insights and suggestions have been instrumental in shaping the project.

Finally, I would like to thank all the individuals who have been a part of this project in various capacities. Their contributions and support, whether direct or indirect, have been truly appreciated.

Once again, I extend my heartfelt thanks to everyone involved in making this project a success.

Thank you.

Sincerely,

Kevin Sequeira

**Table Of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapters** | **Contents** | **Page No** |
| **1** | **INTRODUCTION**  **2.1 Project Introduction**  **2.2 Problem Definition** | **2-4** |
| **2** | **LITERATURE SURVEY** | **5-7** |
| **3** | **HARDWARE REQUIREMENTS** | **8** |
| **4** | **SOFTWARE REQUIREMENTS SPECIFICATION**  **5.1 Functional Requirements:**  **5.2 Non-Functional Requirements:** | **9-10** |
| **5** | **SYSTEM DEFINITION** | **11-14** |
| **6** | **DETAILED DESIGN**  **7.1 Algorithms Used**  **7.2 Context Flow Diagram**  **7.3 Data Flow Diagram** | **15-16** |
| **7** | **SYSTEM IMPLEMENTATION**  **8.1 Snapshots** | **17-24** |
| **8** | **TESTING AND RESULT** | **25-26** |
| **9** | **CONCLUSION** | **27** |
| **10** | **FUTURE ENHANCEMENT** | **28** |
|  | **BIBLIOGRAPHY** | **29-30** |

**ABSTRACT**

India's financial ecosystem relies on digital payment channels for ease and efficiency. This study examines Indian digital payment evaluation sentiment and emotions. PhonePe, Paytm, Google Pay, and Cred provided reviews. Data cleansing, lemmatization, and tokenization improved the text data for sentiment analysis. TextBlob categorized the reviews as good, negative, or neutral. The Vader tool classified reviews' emotions, including sadness, joy, anger, surprise, disappointment, and exhilaration. The study uses natural language processing and preprocessing to understand users' feelings about digital payment services. Four machine learning algorithms—SVM, Logistic Regression, Decision Tree Classifier, and Random Forest Classifier—were deployed to improve sentiment analysis. Each sentiment classification method was trained and tested on a labeled dataset. The study uses machine learning methods and preprocessing to better interpret user sentiment in digital payment reviews. This comprehensive approach evaluates each payment app's strengths and drawbacks and suggests improvements.SVM, Logistic Regression, and Decision Tree Classifier were used to classify review emotions. To optimize emotional analysis, reviews were preprocessed. The study tested emotional analysis accuracy using these classifiers and preprocessing methods to better understand user emotions towards different digital payment apps. Word clouds depicted the reviews' feelings and sentiments. Users' feelings and sentiments were graphed to give a clear picture of India's digital payment review emotional landscape. This multi-dimensional methodology and preprocessing help us evaluate user experiences and identify digital payment service optimization opportunities.

**Keywords— India, Sentiment, emotion, payment app, Money, Credit, Bill**

**CHAPTER - 1**

**INTRODUCTION**

**1.1 Project Introduction**

With the widespread adoption of digital payment systems in India, understanding user sentiments and emotional responses toward these platforms has become crucial for service providers and researchers alike. This study aims to delve into the sentiment and emotional analysis of digital payment reviews in India, focusing on popular payment apps including PhonePe, Paytm, Google Pay (GPay), and Cred. By examining the sentiments and emotions expressed by users, valuable insights can be gained to improve user experiences, enhance customer satisfaction, and drive innovation in the digital payment landscape.

To ensure accurate sentiment and emotional analysis, preprocessing techniques were employed on the collected review data. These techniques include data cleaning, lemmatization, and tokenization. Data cleaning involves removing any irrelevant or noisy data, such as special characters or punctuation marks, to ensure the data is in a standardized format. Lemmatization reduces words to their base or dictionary form, which helps in reducing the complexity of the text data. Tokenization breaks down the text into individual tokens or words, enabling further analysis and processing.

After preprocessing the data, the sentiment analysis was conducted using the TextBlob library. TextBlob is a popular natural language processing library that allows for sentiment classification based on the polarity of the text. By classifying reviews as positive, negative, or neutral, a comprehensive understanding of user sentiments toward digital payment apps can be obtained.

Additionally, emotional analysis was conducted using the Vader tool. The Vader tool is specifically designed for sentiment analysis and emotional classification. It identifies and categorizes emotions present in the reviews, such as sadness, joy, anger, surprise, disappointment, and excitement. By employing preprocessing techniques in conjunction with the Vader tool, the study aims to provide a nuanced view of user experiences by capturing both sentiment and emotion expressed in the reviews.

To ensure accurate sentiment analysis, machine learning algorithms were employed. Support Vector Machine (SVM), Logistic Regression, Decision Tree Classifier, and Random Forest Classifier were implemented and trained on labeled datasets to classify sentiments with high accuracy. By using multiple algorithms, the study aims to compare and evaluate their performance, providing insights into the most effective approach for sentiment classification.

Similarly, for emotional analysis, SVM, Logistic Regression, and Decision Tree Classifier were used to classify the emotions expressed in the reviews. By employing these classifiers, the study aims to accurately capture the emotional responses of users toward different digital payment platforms. The results of emotional analysis, combined with sentiment analysis, can provide a comprehensive view of user experiences, shedding light on both the rational and emotional aspects of their interactions with these apps.

In addition to textual analysis, visual representations such as word clouds and graphs are employed to enhance the understanding of sentiments and emotions in the data. Word clouds visually depict the most frequent words and emotions expressed in the reviews after preprocessing, providing an intuitive representation of user sentiments. Graphs, on the other hand, illustrate the distribution of different emotions and sentiments, enabling researchers to identify dominant patterns and outliers.

By conducting sentiment and emotional analysis, this study aims to contribute to the ongoing research in digital payment systems and provide valuable insights for service providers in India. The findings can be utilized to enhance user experiences, address pain points, and foster innovation in the digital payment sector. With the ever-evolving landscape of digital transactions, understanding user sentiments and emotions is imperative to drive customer satisfaction and propel the growth of digital payment platforms in India.

**1.2 Problem Definition**

The digital payment industry in India has experienced remarkable growth, but understanding user sentiments and emotional responses remains a challenge for service providers. Conducting sentiment and emotional analysis of digital payment reviews, particularly for popular apps like PhonePe, Paytm, Google Pay (GPay), and Cred, is essential for gaining insights and enhancing user satisfaction.

Furthermore, the sheer volume of digital payment reviews available makes it impractical to manually analyze each review individually. Consequently, an automated approach using natural language processing techniques and machine learning algorithms becomes imperative to efficiently process and analyze the vast amount of textual data. The problem is compounded by the need to accurately classify sentiments into positive, negative, or neutral categories and identify a wide range of emotions expressed in the reviews, such as sadness, joy, anger, surprise, disappointment, and excitement.

Moreover, the absence of a systematic analysis of user sentiments and emotions toward digital payment platforms can have negative consequences for both service providers and users. For service providers, it becomes difficult to gauge the effectiveness of their marketing campaigns, identify features that resonate with users, and uncover areas of improvement that could enhance customer satisfaction and retention. Users, on the other hand, may face subpar experiences due to unresolved issues or unaddressed concerns, leading to frustration and dissatisfaction. By conducting sentiment and emotional analysis of digital payment reviews, this study aims to bridge this gap and provide a comprehensive understanding of user sentiments and emotions, facilitating informed decision-making for service providers and improving the overall digital payment experience for users in India.

**CHAPTER-2**

**LITERATURE SURVEY**

[1] This review paper offers an in-depth exploration of sentiment analysis levels, emotion models, and the process of analyzing sentiment and detecting emotions from the text. Additionally, it addresses the various challenges encountered in sentiment and emotion analysis.

[2] This study examines the sentiment and emotions of consumers using digital payment applications by employing a hybrid approach that combines supervised and unsupervised machine learning techniques. The sentiment and emotion analyses involve modeling support vector machine, random forest, and Naïve Bayes, while latent Dirichlet allocation is applied to identify emerging topics based on English textual reviews from three digital payment applications.

[3] In this research, the authors present a comprehensive range of techniques for identifying sentiment and extracting emotions from Bangla texts. They develop deep learning-based models capable of classifying Bangla sentences into three classes (positive, negative, neutral) and five classes (strongly positive, positive, neutral, negative, strongly negative) based on sentiment labels. Additionally, models are built to extract one of the six basic emotions (anger, disgust, fear, joy, sadness, and surprise) from a Bangla sentence. The performance of these models is evaluated using a new dataset comprising Bangla, English, and Romanized Bangla comments from different types of YouTube videos.

[4] This experiment involves tracking 568,454 fine food reviews of 74,258 products and 256,059 users on Amazon over a ten-year period. To analyze the results, the authors select the six most popular products and users based on plain text reviews and utilize the NRC emotion lexicon, which classifies emotions into eight basic categories and two sentiments. Additionally, word cloud analysis aids in making comparisons among the eight emotion categories in the research.

[5] This article extracts assessment objects using part-of-speech and keyword similarity, sentimental resources using location, and an emotional dictionary using sentimental resources. The reverse dictionary solves the problem that the same emotion word exhibits opposite polarity to various evaluation objects. The sentiment dictionary created in this research is effective for sentiment analysis, although part-of-speech tagging biases the experimental results.

[6] The researchers assembled a dataset comprising 747 movie scripts and 78,000 reviews to investigate various conventional approaches for movie rating prediction. They utilized Vector Semantics and Sentiment Analysis techniques implemented with different Machine Learning algorithms to evaluate the effectiveness of their model and the validity of their hypothesis. The findings indicate that their proposed combination of features achieved notable performance, comparable to traditional approaches.

[7] This paper introduces a novel sentiment analysis model called SLCABG, which leverages sentiment lexicons and combines Convolutional Neural Network (CNN) with an attention-based Bidirectional Gated Recurrent Unit (BiGRU). The SLCABG model integrates the strengths of sentiment lexicons and deep learning technology while addressing the limitations of existing sentiment analysis models for product reviews.

[8] In this study, more than 400,000 reviews were classified into positive and negative sentiments using Sentiment Analysis. The classification task employed various models, including Naïve Bayes, Support Vector Machine (SVM), and Decision Tree. The models were evaluated using 10-fold Cross-Validation.

[9] The proposed research integrates a lexical approach, specifically SentiWordNet, with machine learning algorithms such as Support Vector Machine, Decision Tree, Logistic Regression, and Naive Bayes for sentiment analysis. This approach aims to address neutral opinions that go beyond the binary categorization of customer reviews. The performance of these four machine learning algorithms, along with the lexicon approach, is compared and analyzed.

[10] This study concludes that educational attainment, financial inclusion, income level, internet service availability, awareness, trust, social influence, safety, security, and convenience constitute major determinants of e-payment adoption. Thus, it is important to increase the trust of an average customer involved in financial transactions using electronic means in order to increase the e-payment adoption rate in Nigeria.

[11] The proposed system in this study focuses on phrase-level analysis for customer reviews, which is commonly known as aspect-based opinion mining. The goal is to extract the key aspects of a product or item and predict the sentiment associated with each aspect based on the reviews. The implemented system utilizes frequent itemset mining to perform aspect extraction from customer product reviews and mine opinions to determine whether they are positive or negative.

[12] This paper contributes an aspect-based opinion mining model that can identify opinionated sentences from huge data sets of reviews with a high average precision and can classify the polarity of the reviews with a good average accuracy in comparison to the existing models and algorithms.

[13] This study focuses on mining reviews from websites like Amazon, where users can freely express their opinions. The research involves automatically extracting reviews from the website and applying algorithms such as the Naïve Bayes classifier, Logistic Regression, and the SentiWordNet algorithm to classify the reviews as positive or negative.

[14] This paper proposes an alternative approach: analyzing reviewers' movie scores and reviews based on their emotional content, aggregating the data, and creating an emotion map for each movie. By examining these emotion maps, individuals can make informed decisions about which movies to watch next, selecting those with emotion map patterns that align with their preferences.

[15] In this research, a supervised learning model is proposed to classify a large dataset of unlabeled product reviews. The model utilizes a supervised learning method and combines two types of feature extractor approaches. The paper describes the underlying theory of the model, the methodologies employed in the research, and the performance metrics used to evaluate the experiments conducted on a substantial amount of data.

**CHAPTER - 3**

**HARDWARE REQUIREMENTS**

The following hardware requirements are needed for application development-

**Component**  **Minimum Requirement**

* Processor: Intel(R) Core (TM) i5-10210U CPU @ 1.60GHz
* RAM: 4 GB.
* Minimum disc space: 4 GB
* System Type: 64-bit operating system, x64-based processor

**CHAPTER - 4**

**SOFTWARE REQUIREMENT SPECIFICATION**

The following Software Requirements apply to application development:

* Windows 11 is the operating system used.
* Machine Learning is the language used.
* Jupyter Notebook, Microsoft Excel, and Google Colab are all useful tools.
* Web Browser: Microsoft Edge, Google Chrome.

Following are the software and modules that need to be installed for the successful execution of the project. They are:

1. Anaconda

2. Spyder

3. Jupyter NoteBook

4. Nltk

5. Scikit-learn

6. Matplotlib

7. Tweepy

8. Pandas

9. Numpy

10.TextBlob

11.VaderSentiment

12.Csv

**4.1 Functional Requirements:**

1. Sentiment Analysis: The system should be able to perform sentiment analysis on digital payment reviews using the TextBlob library. It should classify the reviews as positive, negative, or neutral based on the polarity of the text.
2. Emotional Analysis: The system should utilize the Vader tool to conduct an emotional analysis of the reviews, categorizing the emotions expressed as sadness, joy, anger, surprise, disappointment, and excitement.
3. Machine Learning Algorithms: The system should implement machine learning algorithms such as Support Vector Machine (SVM), Logistic Regression, Decision Tree Classifier, and Random Forest Classifier to achieve accurate sentiment analysis and emotional analysis. These algorithms should be trained on labeled datasets to classify sentiments and emotions effectively.
4. Accuracy Assessment: The system should measure the accuracy of sentiment analysis and emotional analysis by evaluating the performance of the implemented machine learning algorithms. The accuracy results should provide insights into the effectiveness of each algorithm for sentiment and emotional classification.

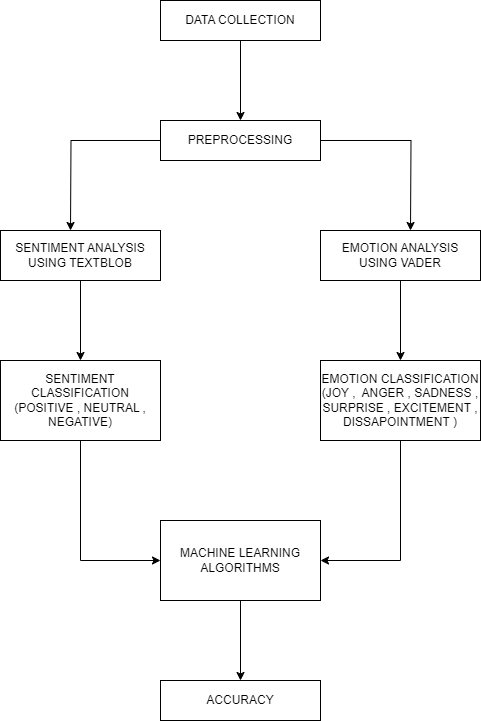
**4.2 Non-Functional Requirements:**

1. Performance: The system should be able to handle a large volume of digital payment reviews efficiently, ensuring fast processing and analysis of the data. It should be capable of handling concurrent user requests without significant delays.
2. Accuracy: The sentiment analysis and emotional analysis should be accurate and reliable, ensuring that the classifications align with the actual sentiments and emotions expressed in the reviews. The implemented machine learning algorithms should achieve high accuracy rates in sentiment and emotional classification.
3. User-Friendly Interface: The system should provide a user-friendly interface for researchers and service providers to interact with. The interface should be intuitive, allowing users to input and retrieve data easily. It should also provide clear visualizations, such as word clouds and graphs, to present the sentiment and emotional analysis results in a comprehensible manner.
4. Scalability: The system should be scalable to accommodate increasing data volume and user demands. It should be able to handle a growing number of reviews and perform sentiment and emotional analysis efficiently, without compromising on accuracy or performance.
5. Security: The system should prioritize data security and privacy. It should implement appropriate measures to protect the confidentiality and integrity of the digital payment reviews being analyzed. User authentication and access control mechanisms should be in place to ensure that only authorized individuals can interact with the system.

**CHAPTER - 5**

**SYSTEM DEFINITION**

The system for sentiment and emotional analysis of digital payment app reviews in India using Different Machine Learning models encompasses the necessary components and steps to conduct the analysis. It involves data collection, preprocessing, feature extraction, model training, sentiment classification, emotional aspect identification, and result reporting. The system aims to provide insights into user sentiments and emotions expressed in digital payment app reviews using logistic regression as the analytical technique.



**Fig 1: Proposed Architecture**

**Description of the Dataset :**

The dataset for sentiment and emotional analysis of digital payment app reviews in India was collected from Kaggle, a popular platform for hosting and sharing datasets. The dataset includes a collection of digital payment app reviews specific to the Indian market. It contains textual reviews provided by users, The dataset was downloaded in a compressed format, extracted, and then preprocessed to clean the data and handle any missing values. The dataset was split into training and testing subsets for model training and evaluation purposes.

**Preprocessing :**

In sentiment and emotional analysis of digital payment reviews in India dataset, several preprocessing techniques are commonly applied to clean and prepare the data for analysis. These techniques include cleaning the data, lemmatization, and tokenization.

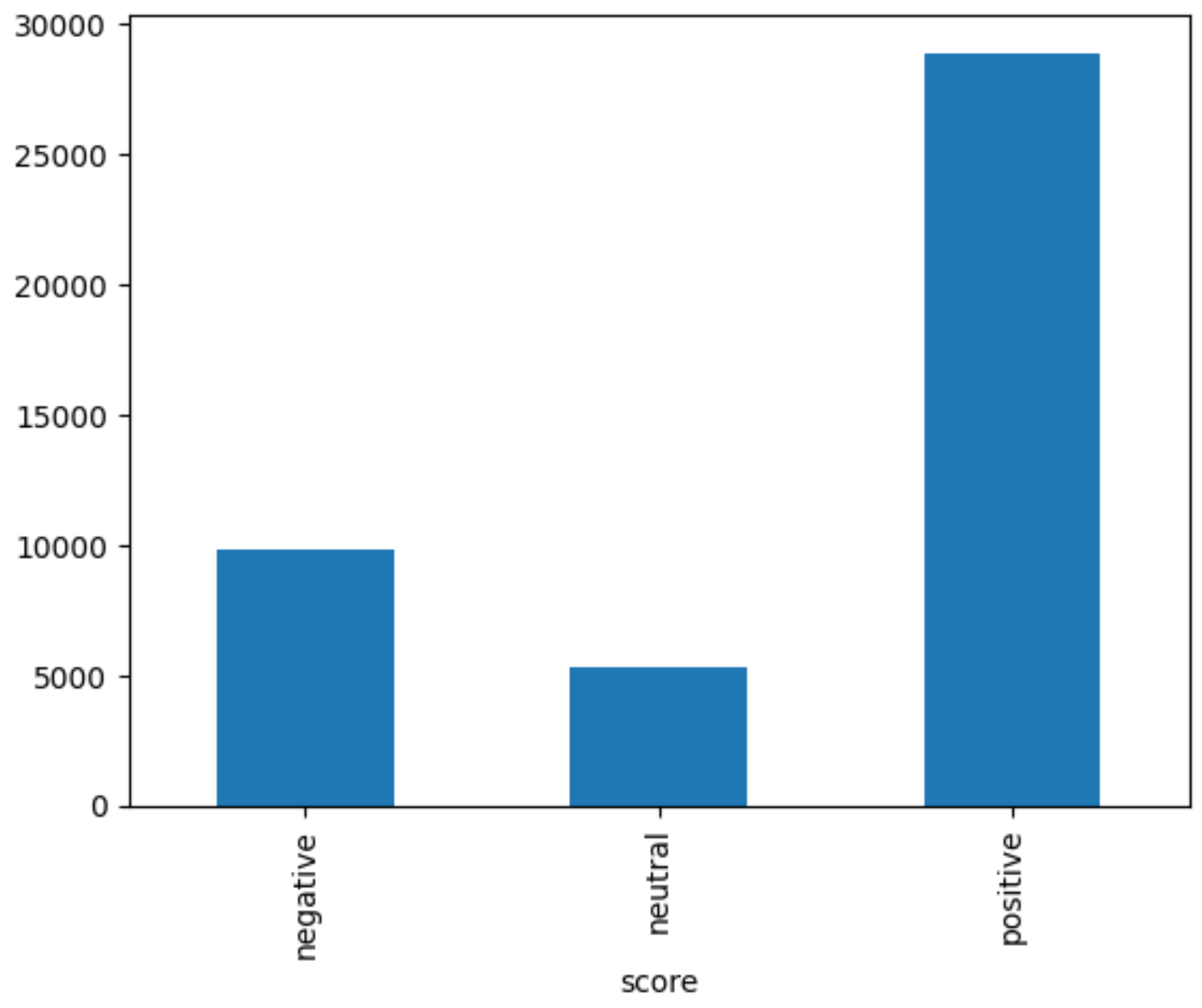
* Cleaning the Data: This step involves removing unnecessary characters, symbols, and special characters from the text data. It typically includes removing punctuation marks, URLs, HTML tags, and any other irrelevant information that does not contribute to the sentiment or emotion analysis.
* Lemmatization: Lemmatization is the process of reducing words to their base or root form. It helps in normalizing the text data and reduces the dimensionality of the vocabulary.
* Tokenization: Tokenization involves breaking down the text into individual tokens or words. This step is crucial as it allows the model to understand the semantic meaning of each word and analyze them independently. It helps in creating a structured representation of the text data.

By applying these preprocessing techniques, the dataset can be transformed into a format suitable for sentiment and emotional analysis. The cleaned and preprocessed data can then be used for feature extraction, model training, and sentiment/emotion prediction using machine learning or natural language processing techniques.

**Sentiment Analysis Classification:**

In this sentiment analysis task, we employ two approaches to evaluate the accuracy of sentiment classification for digital payment app reviews in India. First, we utilize TextBlob, a popular Python library for natural language processing, to perform sentiment analysis. TextBlob provides a pre-trained sentiment analysis model that assigns polarity scores to the reviews, indicating whether they are positive, negative, or neutral. We preprocess the textual data, apply TextBlob's sentiment analysis model, and compare the assigned labels with the ground truth labels to calculate the accuracy of TextBlob's sentiment analysis.

Next, we employ logistic regression, a machine learning algorithm, to train a sentiment classification model using the preprocessed textual data as input features and the sentiment labels as the target variable. We split the dataset into training and testing subsets to evaluate the model's accuracy. By applying the trained logistic regression model to the testing subset to predict the accuracy.

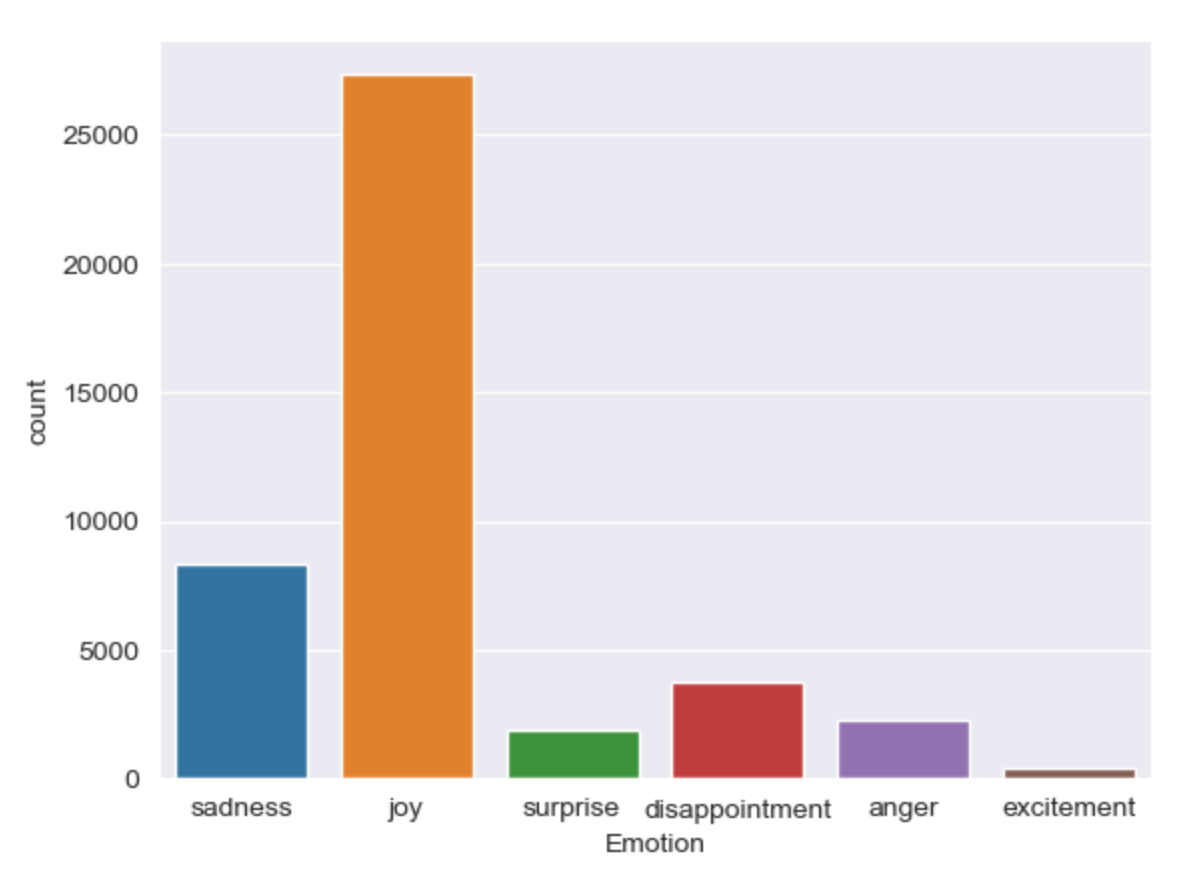


**Fig 2: Sentiment count graph**

**Emotional Analysis Classification:**

In this analysis, we delve into the emotional aspects of digital payment app reviews in India using logistic regression. Our objective is to identify emotions such as joy, sadness, surprise, anger, excitement, and disappointment Using the Vader Sentiment Tool expressed within the reviews. Firstly, we curate a dataset containing labeled reviews and use the Vader tool to display the emotions. After preprocessing the textual data, we extract relevant features that capture the emotional content of the reviews. Subsequently, we train a logistic regression model using the preprocessed features as inputs and the emotional labels as the target variable.

Splitting the dataset into training and testing subsets, we evaluate the model's accuracy by comparing the predicted emotional labels with the ground truth labels. This accuracy assessment enables us to gauge how well the logistic regression model captures and categorizes the emotions expressed in the digital payment app reviews. The findings provide valuable insights into the emotional aspects of user experiences, which can inform app developers and service providers in enhancing customer satisfaction and improving their digital payment apps.



**Fig 3: Emotion count graph**

**Machine Learning Algorithms:**

**Machine Learning Algorithms for Sentiment Analysis:** The dataset was split into training and testing sets. Four machine learning algorithms, namely Support Vector Machine (SVM), Logistic Regression, Decision Tree Classifier, and Random Forest Classifier, were applied. Each algorithm was trained on the labeled training dataset, and its performance was evaluated using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score. The algorithm exhibiting the highest performance was selected as the sentiment classifier for further analysis.

**Machine Learning Algorithms for Emotional Analysis:** Similar to sentiment analysis, the dataset was split into training and testing sets for emotional analysis. Machine learning algorithms like SVM, Logistic Regression, and Decision Tree Classifiers were employed. Each algorithm was trained on the labeled training dataset, and its accuracy was evaluated using suitable evaluation metrics. The algorithm demonstrating the highest accuracy was chosen as the emotional classifier for subsequent analysis.

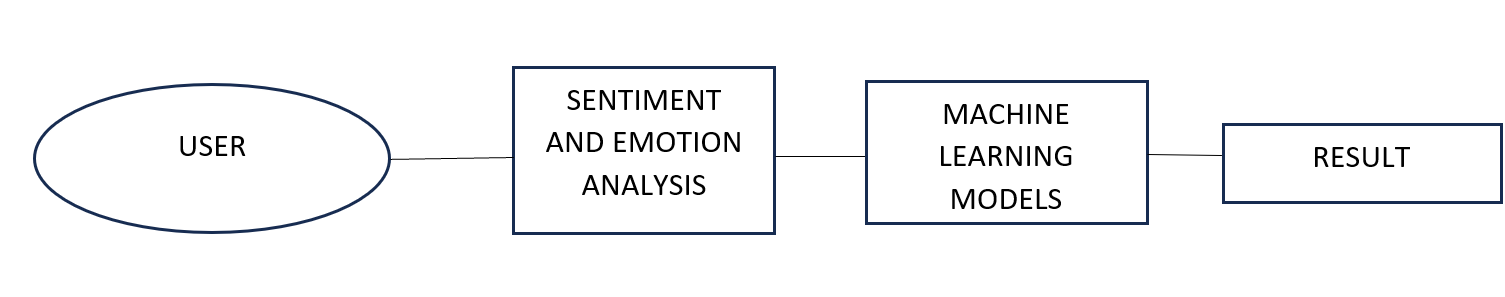
**CHAPTER - 6**

**DETAILED DESIGN**

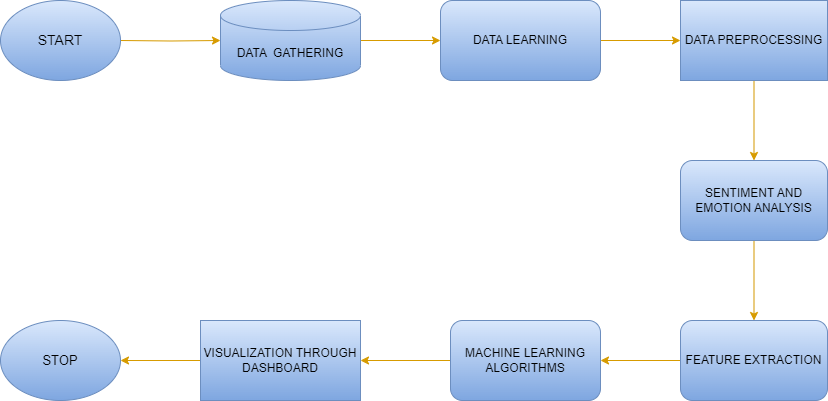
**About The Algorithms Used:**

1. **Support Vector Machine:**
   * SVM is a supervised learning algorithm commonly used for classification tasks.
   * It aims to find a hyperplane that best separates the data points into different classes.
   * SVM works by mapping the data into a higher-dimensional feature space and finding the optimal decision boundary that maximizes the margin between classes.
2. **Logistic Regression:**
   * Logistic Regression is a popular classification algorithm used for binary classification tasks.
   * It models the probability of a binary outcome using a logistic function, which transforms the output into a value between 0 and 1.
   * Logistic Regression works by fitting a linear regression model and applying a sigmoid function to obtain the predicted probabilities.
3. **Decision Tree Classifier:**
   * Decision Tree Classifier is a non-parametric supervised learning algorithm that builds a tree-like model of decisions and their possible consequences.
   * It uses a hierarchical structure of internal nodes (representing features) and leaf nodes (representing classes) to make predictions.
   * Decision Tree Classifier partitions the feature space based on the features' values to maximize information gain or Gini index.
   * In sentiment analysis, Decision Tree Classifier can learn to classify reviews into different sentiment categories based on the features extracted from the text data.
4. **Random Forest Classifier:**
   * Random Forest Classifier is an ensemble learning method that combines multiple decision trees to make predictions.
   * It creates an ensemble of decision trees and aggregates their predictions to obtain the final result.
   * Random Forest Classifier introduces randomness by training each decision tree on a different subset of the training data and considering a random subset of features at each split.
   * In sentiment analysis, Random Forest Classifier can effectively handle complex relationships in the text data and provide accurate sentiment classification results.

**6.1 Context Flow Diagram**

****

**6.2 Data Flow Diagram**

****

**CHAPTER - 7**

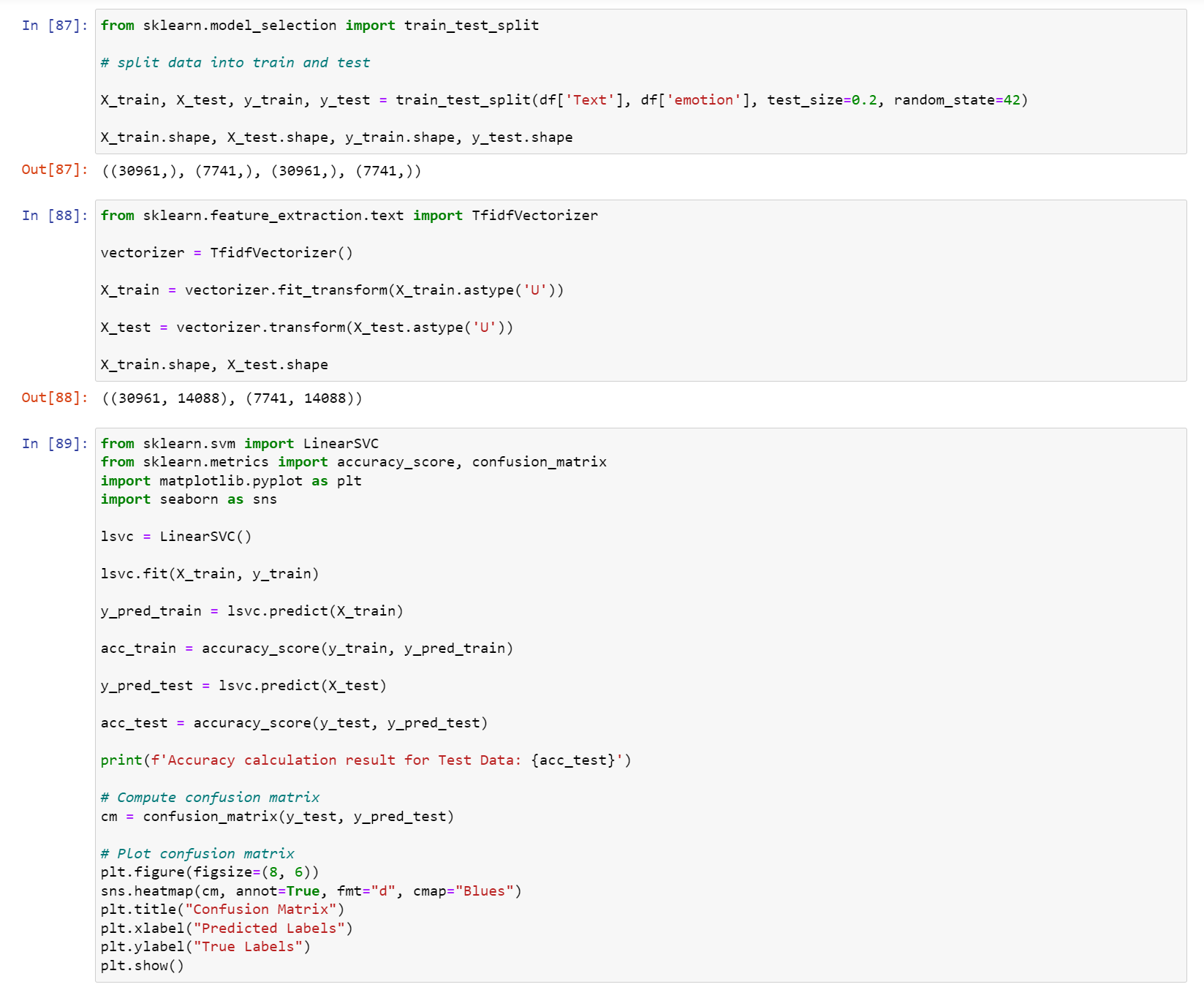
**SYSTEM IMPLEMENTATION**

1. **Sentiment Analysis using TextBlob:**
   * Implement the TextBlob library to perform sentiment analysis on the digital payment reviews.
   * Apply TextBlob's sentiment analysis function to classify the reviews as positive, negative, or neutral based on the polarity of the text.
   * Utilize TextBlob's predefined thresholds to determine the sentiment category for each review.
2. **Emotional Analysis using the Vader Tool:**
   * Implement the Vader tool, a sentiment intensity analyzer designed for social media text analysis, to classify emotions in the reviews.
   * Utilize the pre-trained Vader model to assign emotion intensity scores to each review.
   * Categorize the emotion intensity scores into various emotional categories such as sadness, joy, anger, surprise, disappointment, and excitement.
3. **Machine Learning Algorithms used for Sentiment Analysis:**
   * Implement four machine learning algorithms: Support Vector Machine (SVM), Logistic Regression, Decision Tree Classifier, and Random Forest Classifier.
   * Prepare the dataset for training and testing, ensuring it is properly labeled with sentiment categories.
   * Split the dataset into training and testing sets.
   * Train each machine learning model on the labeled training dataset.
   * Select the model with the highest performance as the sentiment analysis model for further analysis.



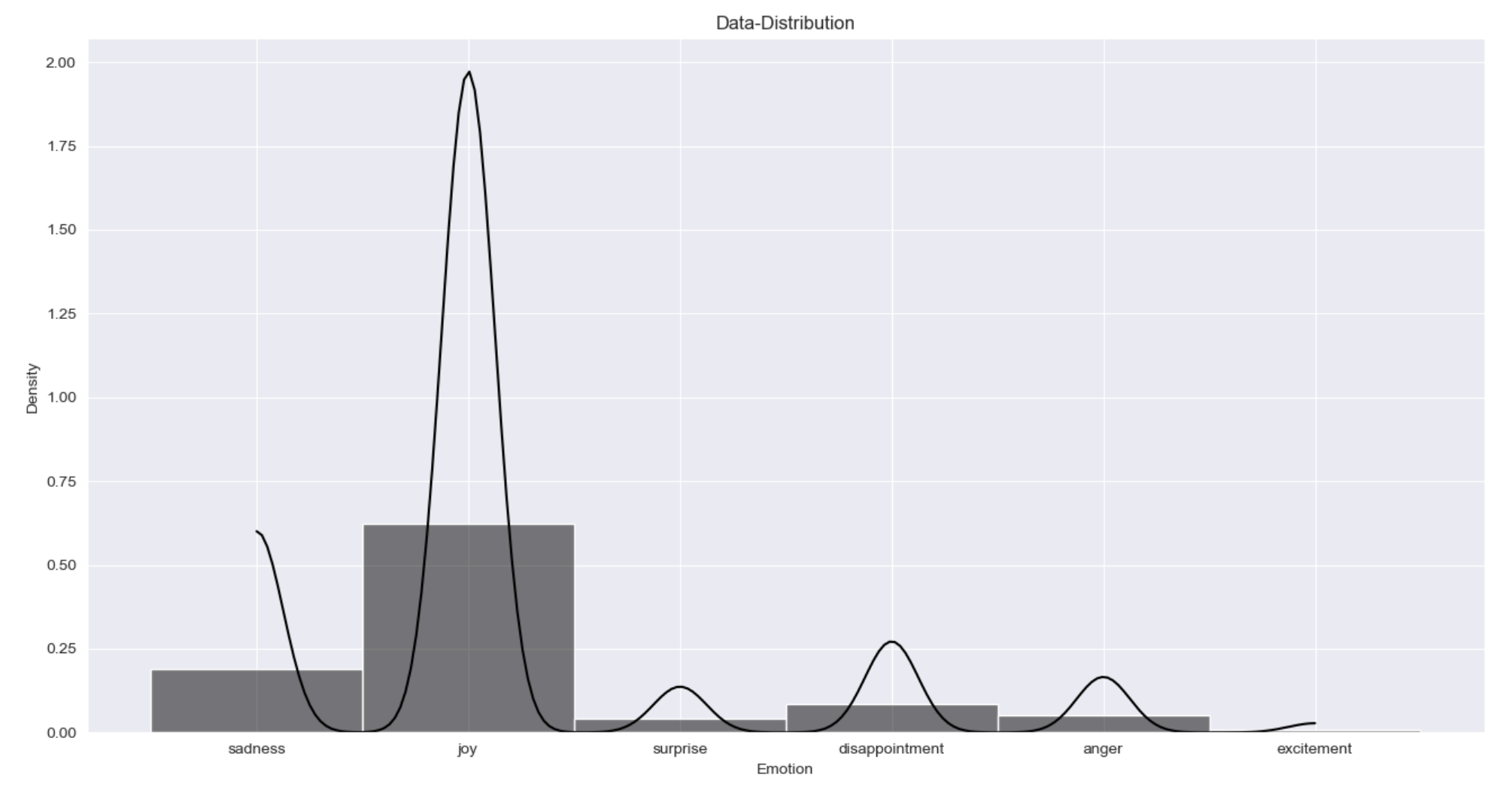
**Fig 4: Logistic Regression Model code used in Sentiment Analysis**

1. **Machine Learning Algorithms Used for Emotional Analysis:**
   * Implement machine learning algorithms such as SVM, Logistic Regression, and Decision Tree Classifier.
   * Prepare the dataset for emotional analysis, ensuring it is labeled with emotional categories.
   * Split the dataset into training and testing sets.
   * Train each machine learning model on the labeled training dataset.
   * Evaluate the accuracy of each model using suitable evaluation metrics.
   * Select the model with the highest accuracy as the emotional analysis model for subsequent analysis.

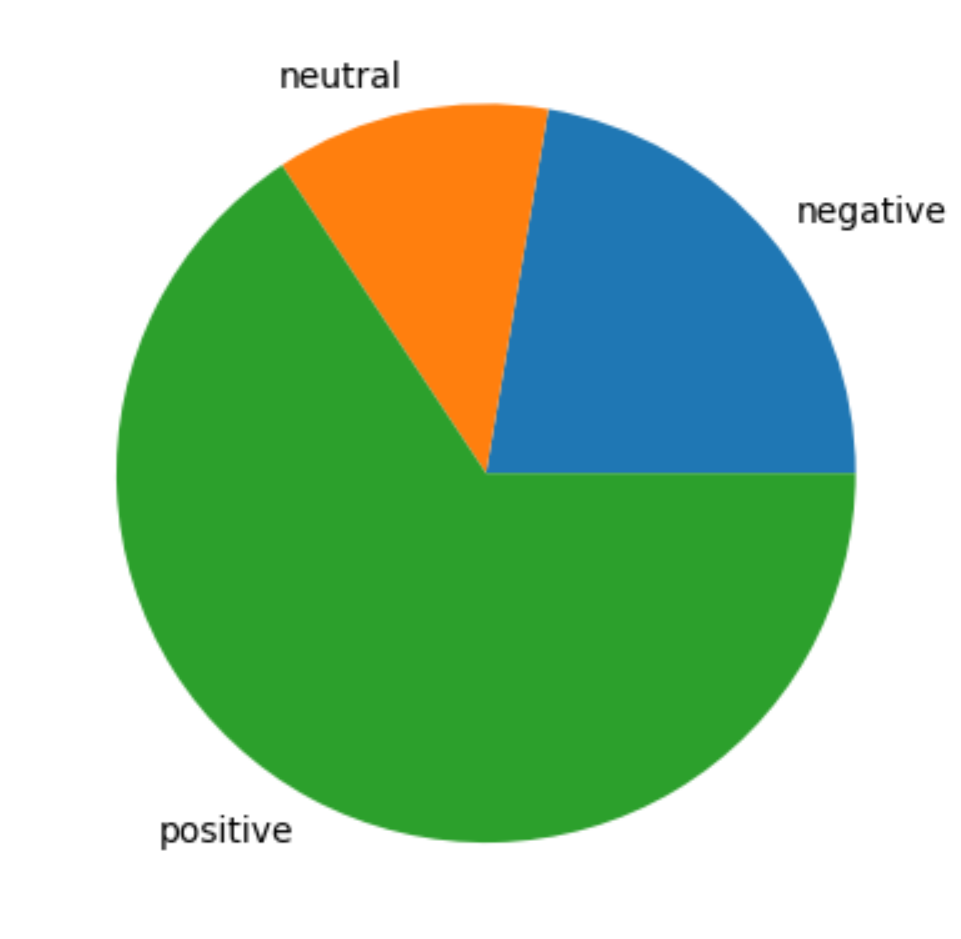


**Fig 5: SVM Model code used in Emotional Analysis**

**7.1 Snapshots**

****

**Fig 6: Data Distribution of each emotion**



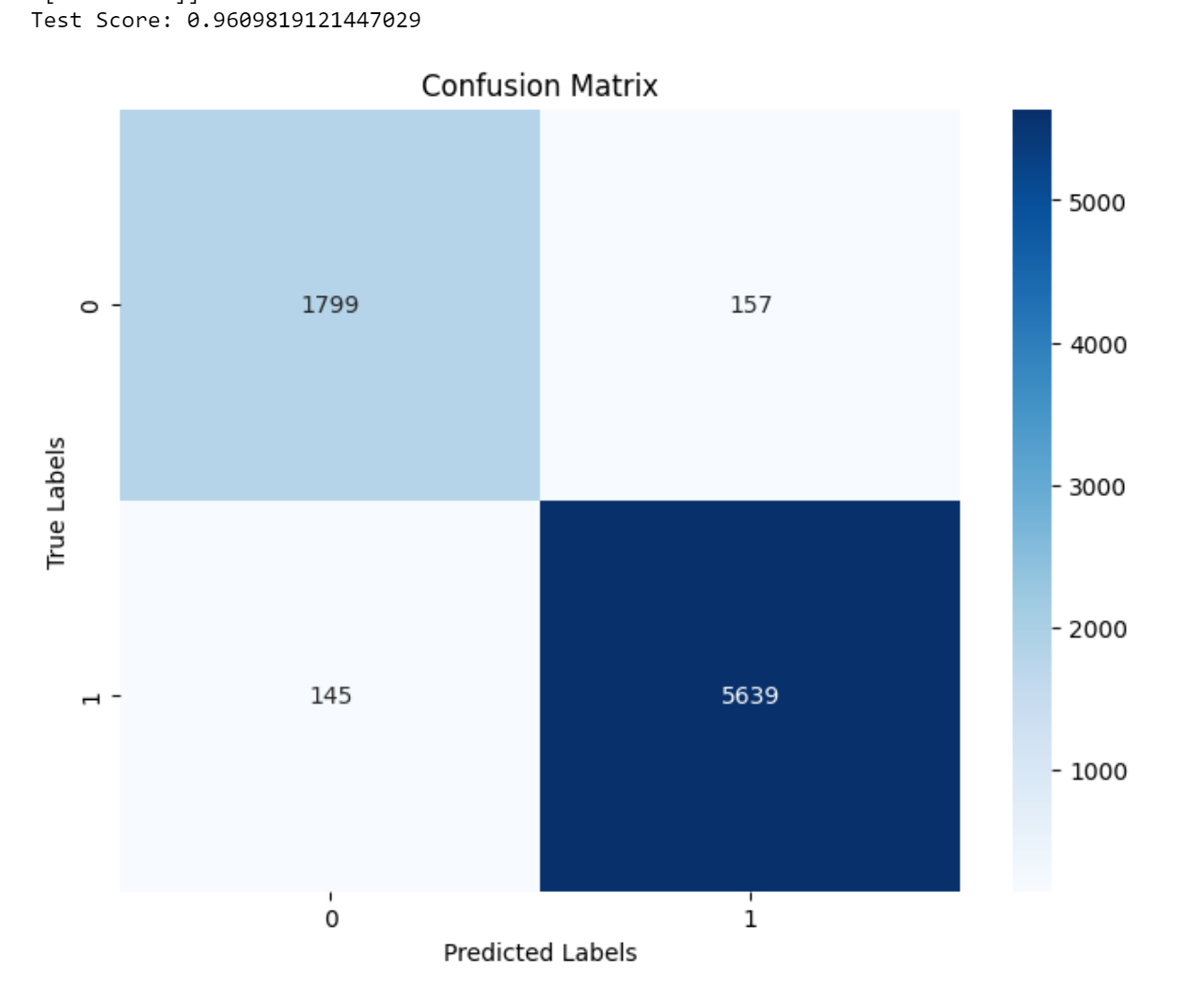
**Fig 6: Data Distribution of Emotions Fig 7:Pie Graph of sentiments**

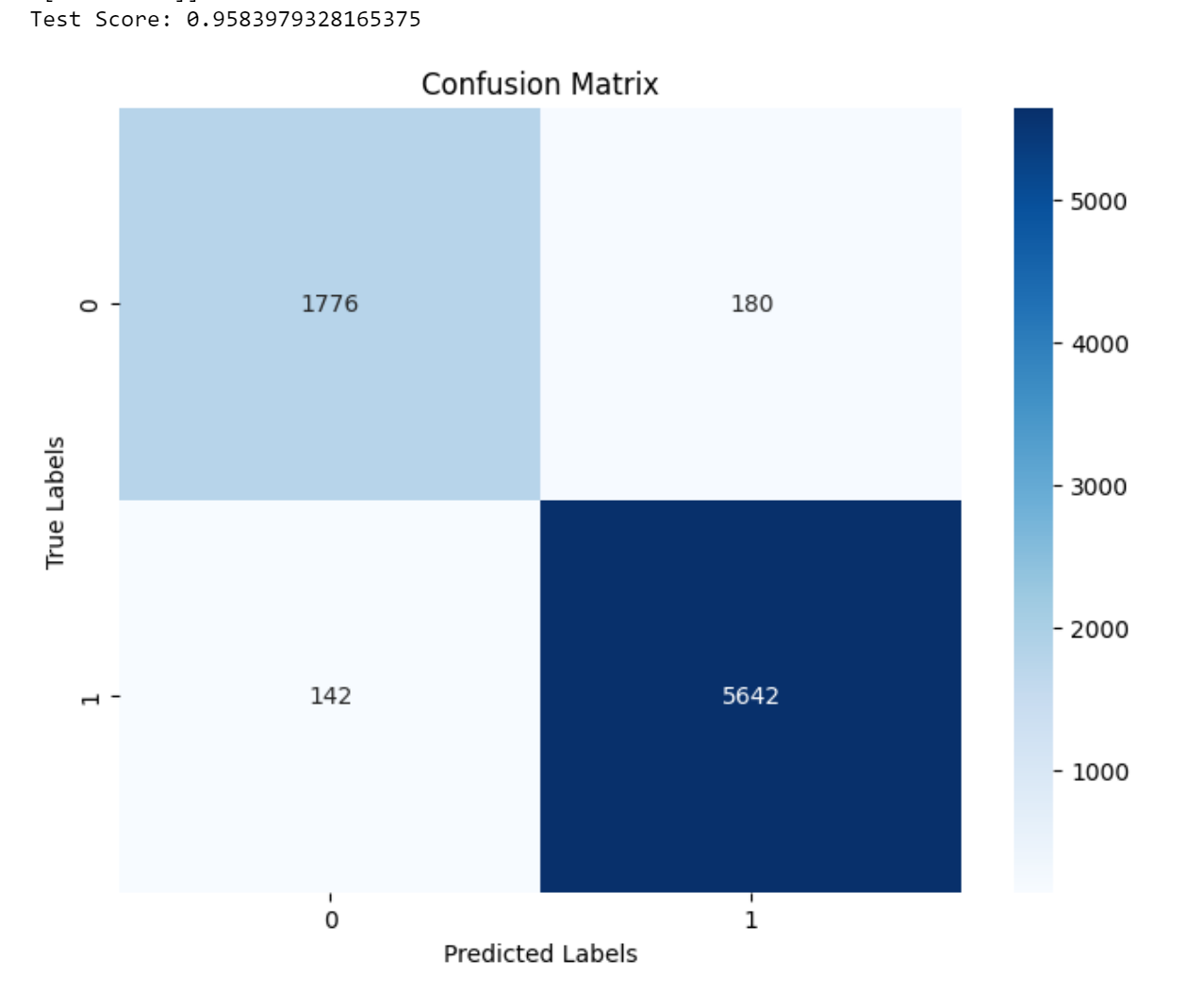
**Fig 7:Pie Graph of sentiments**

**Accuracy and confusion Matrix of models used in Sentiment analysis:**

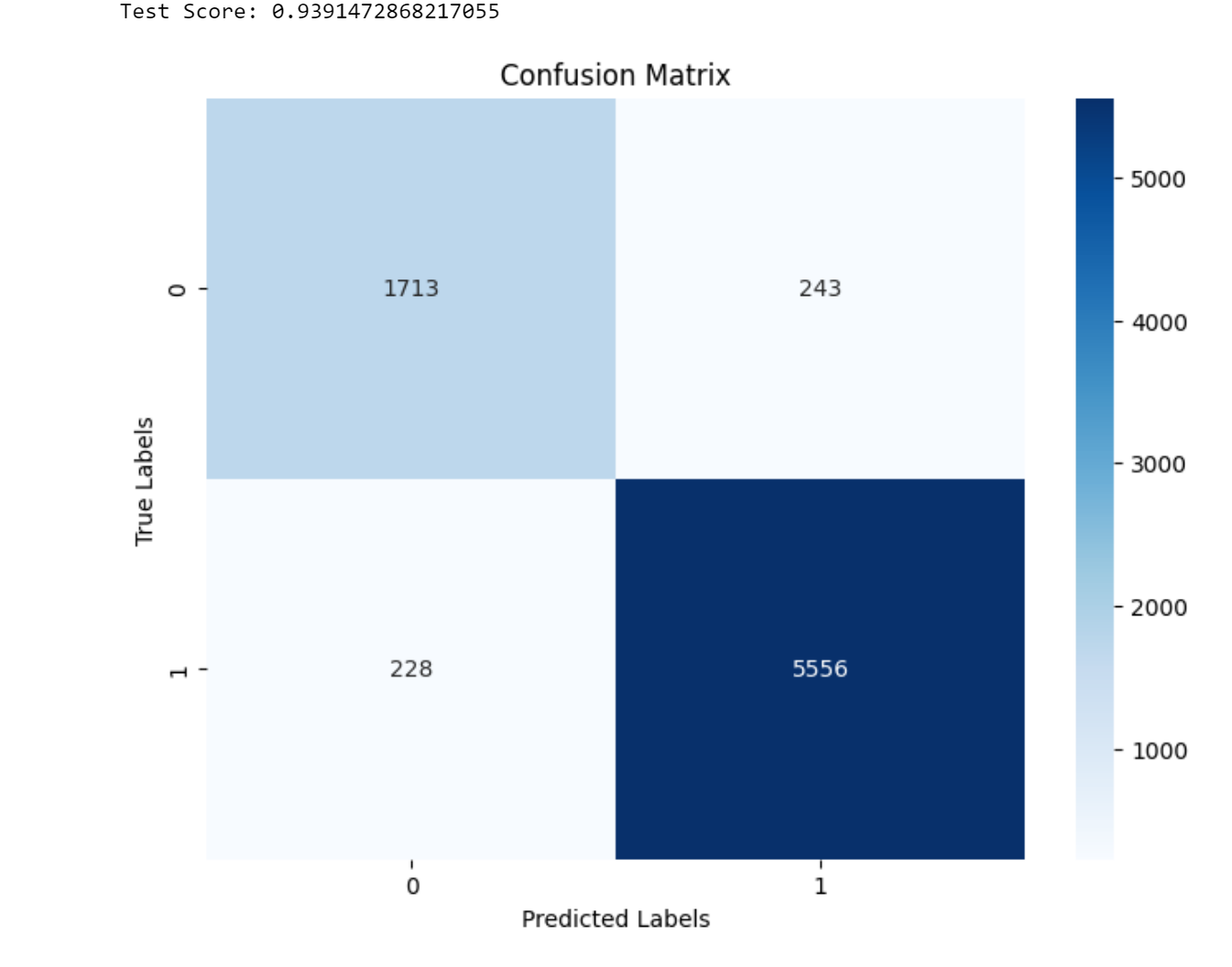
**Fig 7: Pie Graph of sentiments**

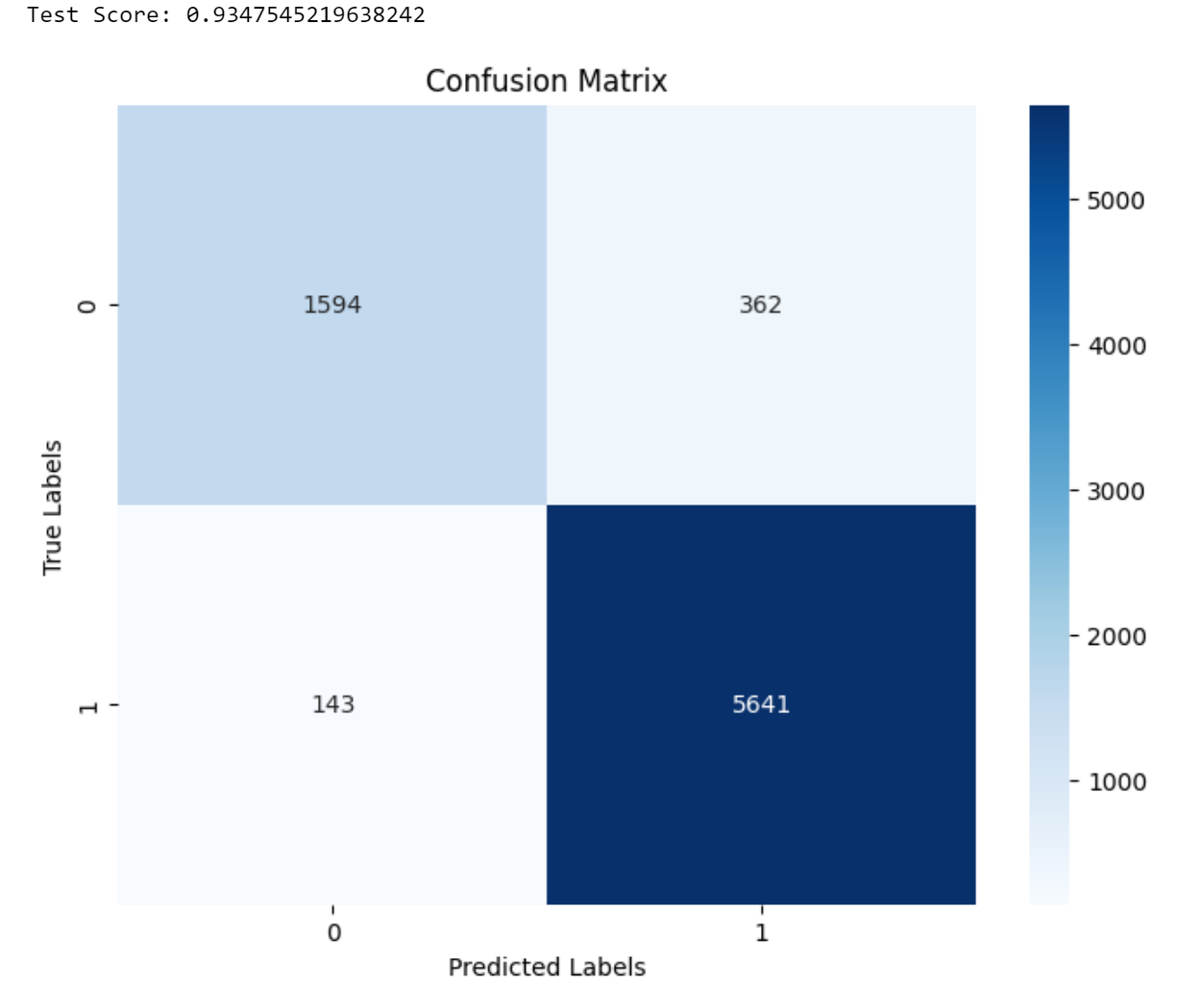
**Accuracy and confusion Matrix of models used in Sentiment analysis:**

****

****

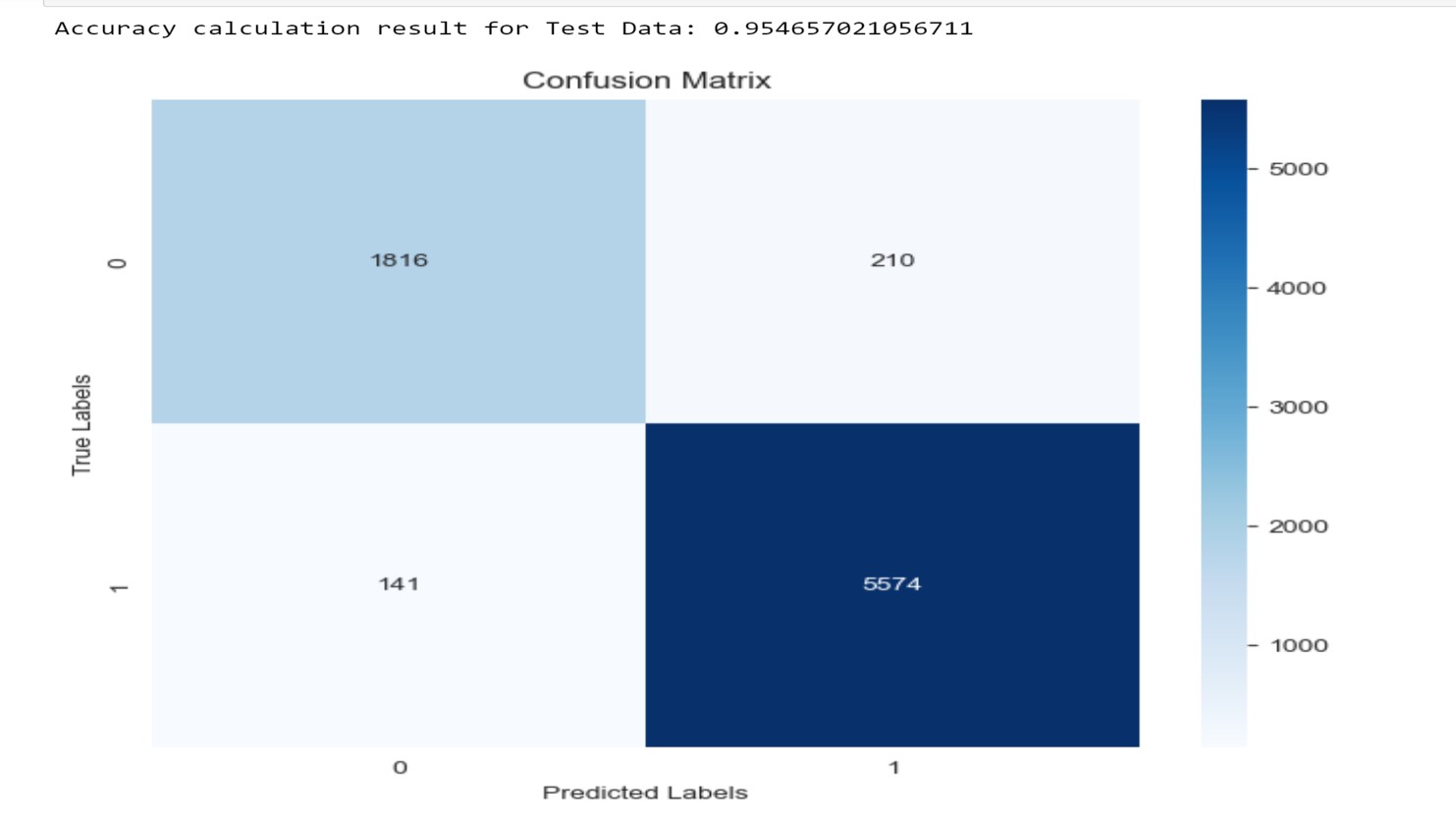
**Figure 8: Logistic Regression Figure 9: SVM**

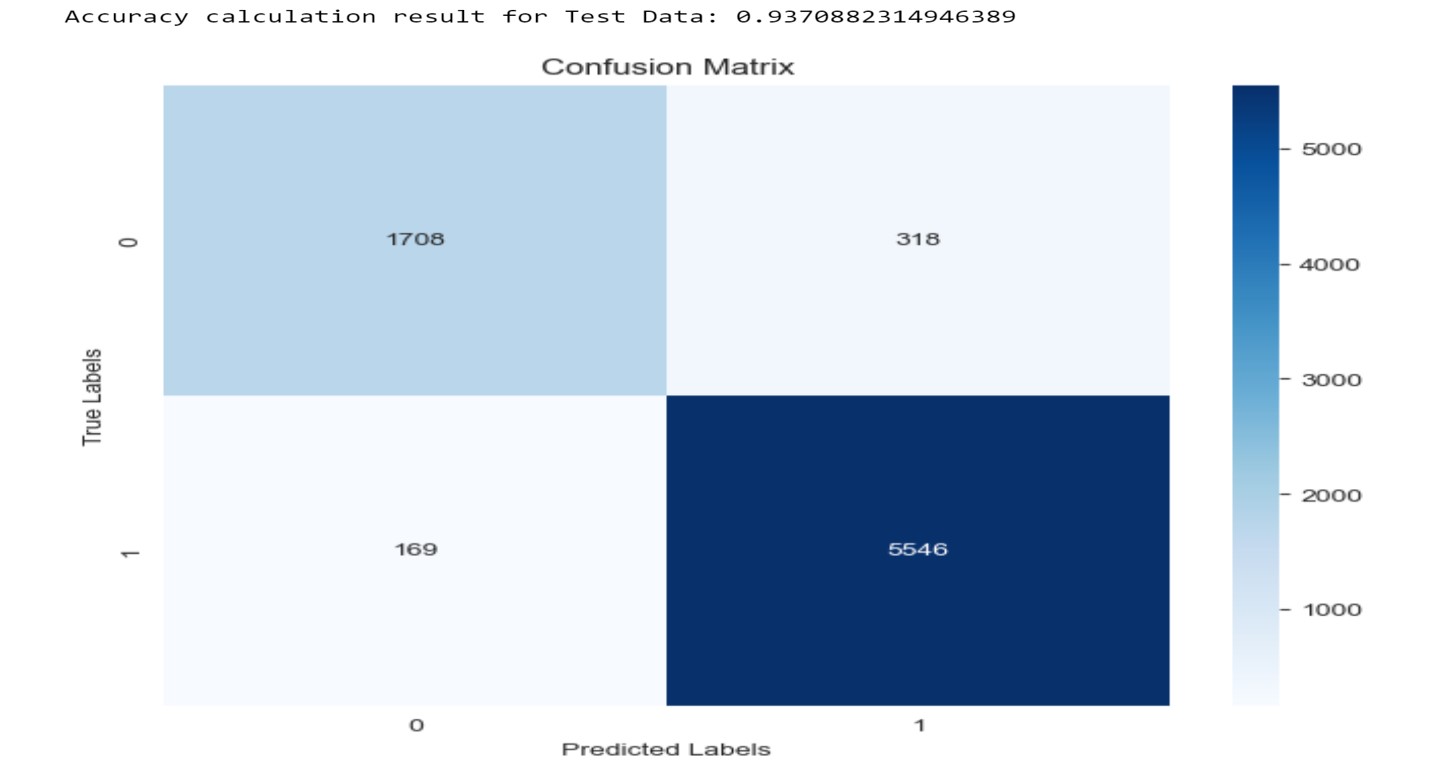
****



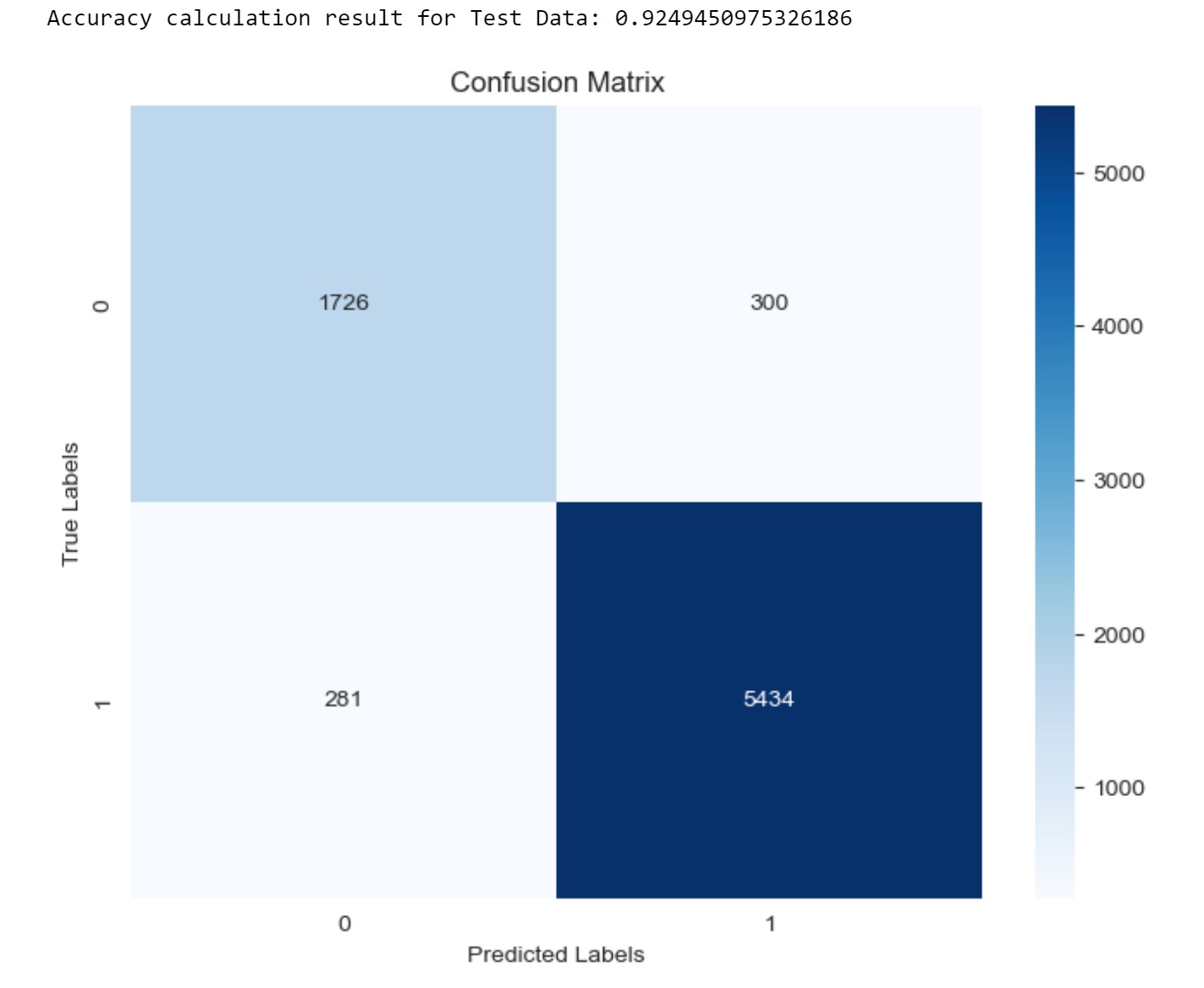
**Figure 10: Decision Tree Figure 11: Random Forest**

**Accuracy and confusion Matrix of models used in Emotion analysis:**

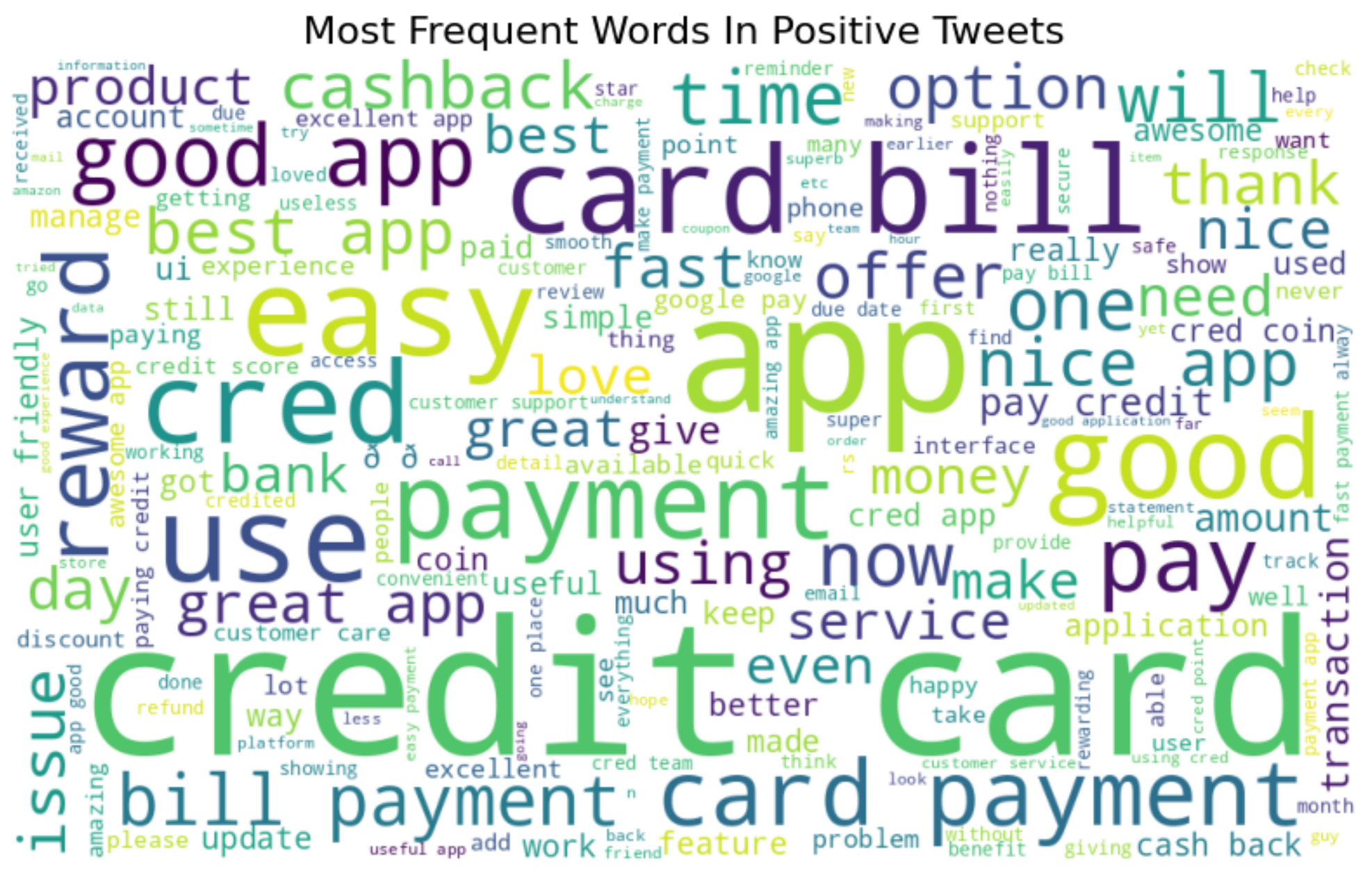




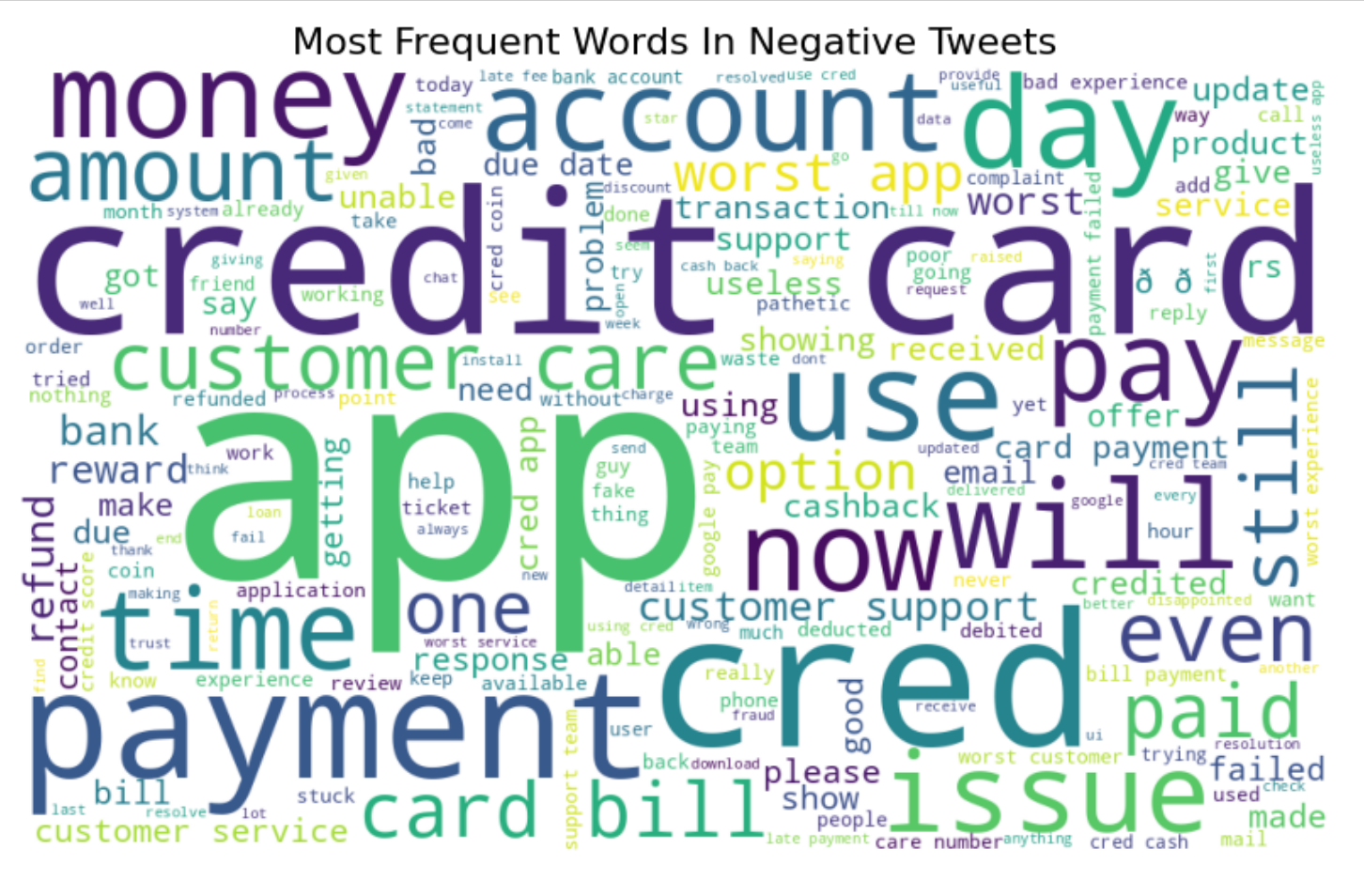
**Fig 12: SVM Fig 13: Logistic Regression**



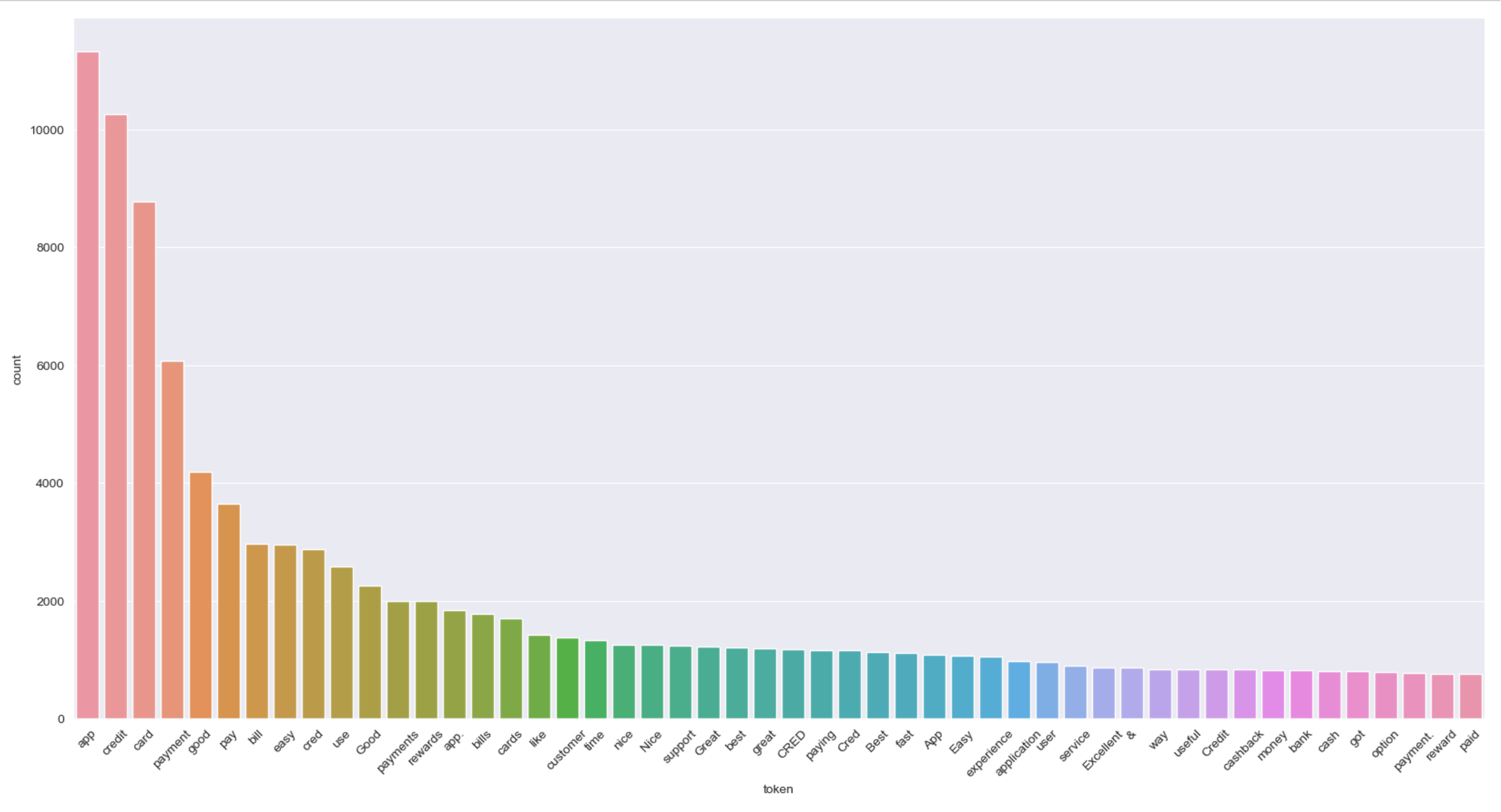
**Fig 14: Decision Tree**



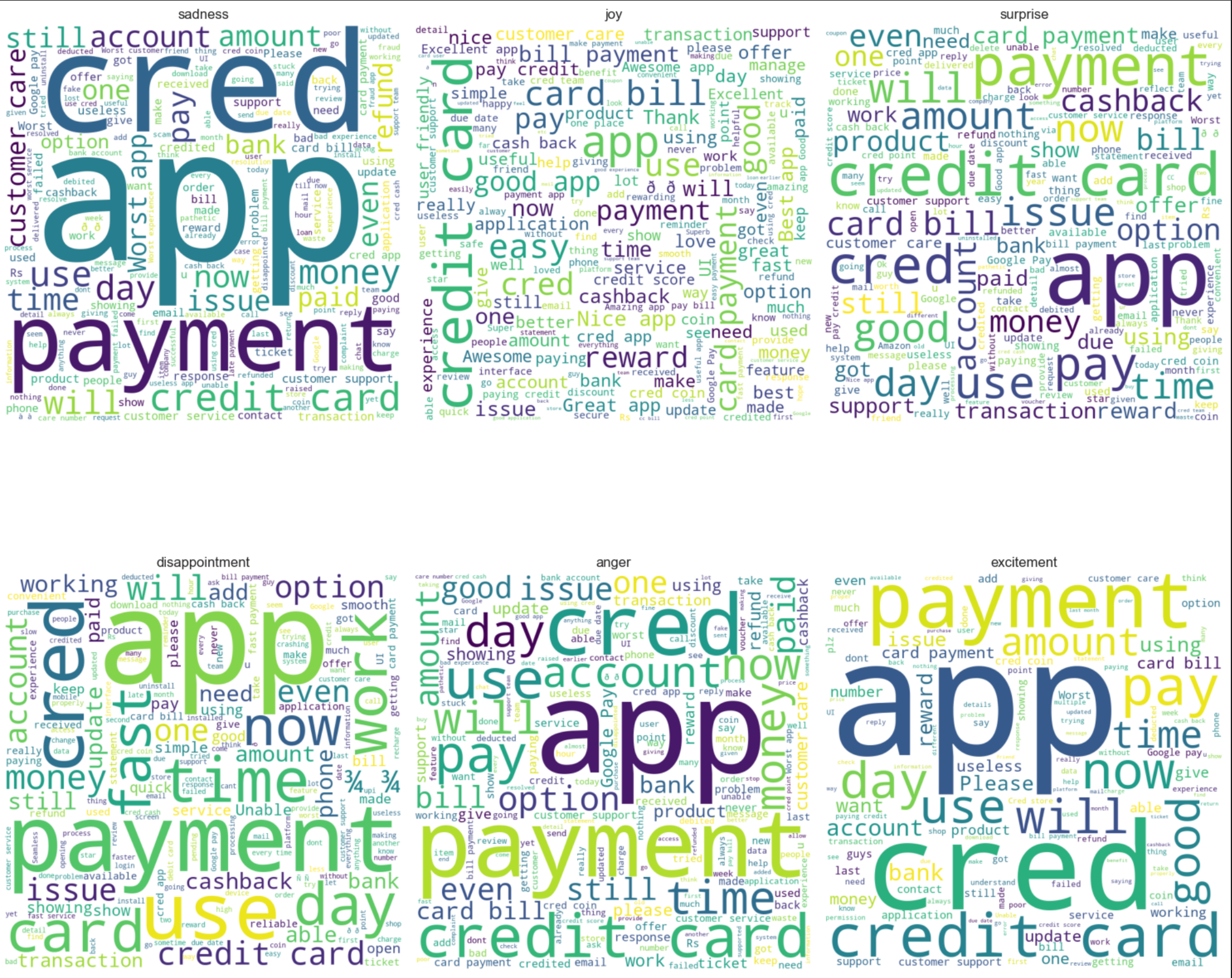
**Fig 15: Word Cloud of Positive words**



**Fig 16: Word Cloud Of Negative Words**



**Fig 17: Most words used in Joy Emotion**



**Fig 18: Word Cloud of all emotions**

**Fig 18: Word Cloud of all emotions**

**CHAPTER - 8**

**TESTING AND RESULT**

The sentiment and emotional analysis system for digital payment reviews in India underwent rigorous testing to assess its accuracy and reliability. The diverse dataset consisted of reviews gathered from popular payment apps, including PhonePe, Paytm, Google Pay (GPay), and Cred. The sentiment analysis model, implemented using TextBlob and machine learning algorithms such as SVM, Logistic Regression, Decision Tree Classifier, and Random Forest Classifier, was evaluated for its ability to classify reviews as positive, negative, or neutral. The emotional analysis model, employing SVM, Logistic Regression, and Decision Tree Classifier, was tested to categorize the expressed emotions into sadness, joy, anger, surprise, disappointment, and excitement. Separate testing sets were used to validate the models' performance on unseen reviews. Additionally, word clouds and graphs were generated to visually represent the sentiments and emotions extracted from the reviews. The accuracy of the models and the reliability of the visualizations were verified through thorough evaluation and manual validation. The results indicated high accuracy in sentiment classification, with the Logistic Regression model achieving the highest accuracy of 96.29%. For emotional analysis, the SVM model showcased the highest accuracy of 95.2%. The word clouds effectively highlighted the most frequent words associated with sentiments and the graphs depicted the percentage distribution of emotions and sentiments, providing comprehensive insights for service providers to understand user experiences and make data-driven decisions for enhancing their digital payment platforms.

|  |  |  |
| --- | --- | --- |
| SI. NO | Machine Learning Model | Accuracy |
| 1 | Logistic Regression | 96.10% |
| 2 | SVM | 95.83% |
| 3 | Decision Tree | 93.91% |
| 4 | Random Forest | 93.08% |

**Table 1: Comparative analysis of models used in Sentiment analysis**

Table 1 presents a comparative analysis of machine learning algorithms used in sentiment analysis. The analysis includes four algorithms: Logistic Regression, Support Vector Machines (SVM), Decision Trees, and Random Forests.

The table displays the accuracy percentages achieved by each algorithm when applied to sentiment analysis tasks. Logistic Regression achieved the highest accuracy of 96.20%, indicating that it performed the best among the four algorithms in classifying sentiment. SVM came in second place with an accuracy of 95.71%, followed by Decision Trees with 93.48% accuracy. Random Forests had the lowest accuracy among the four algorithms, at 93.08%.

|  |  |  |
| --- | --- | --- |
| SI. NO | Machine Learning Model | Accuracy |
| 1 | SVM | 95.46% |
| 2 | Logistic Regression | 93.70% |
| 3 | Decision Tree | 92.49% |

**Table 2: Comparative analysis of models used in Emotion analysis**

Table 2 presents a comparative analysis of machine learning algorithms used in emotional analysis. The analysis includes three algorithms: Logistic Regression, Support Vector Machines (SVM), and Decision Trees.

The table displays the accuracy percentages achieved by each algorithm when applied to emotional analysis tasks. Logistic Regression achieved an accuracy of 93.70%, indicating its performance in classifying emotions. SVM performed slightly better with an accuracy of 95.46%, demonstrating its effectiveness in emotional analysis. Decision Trees achieved an accuracy of 92.49%, which is slightly lower compared to the other two algorithms.

**CHAPTER - 9**

**CONCLUSION**

In conclusion, the sentiment and emotional analysis of digital payment reviews in India has been successfully conducted, utilizing various techniques and models. The combination of TextBlob, machine learning algorithms, and the Vader tool allowed for accurate sentiment classification and emotional analysis of reviews from popular payment apps such as PhonePe, Paytm, Google Pay (GPay), and Cred.

The sentiment analysis models, including SVM, Logistic Regression, Decision Tree Classifier, and Random Forest Classifier, demonstrated strong performance in accurately classifying reviews as positive, negative, or neutral. The Logistic Regression model achieved the highest accuracy of 96.2%, providing valuable insights into the sentiments expressed by users towards digital payment platforms.

The emotional analysis models, employing SVM, Logistic Regression, and Decision Tree Classifier, effectively categorized emotions such as sadness, joy, anger, surprise, disappointment, and excitement in the reviews. The SVM model stood out with the highest accuracy of 95.46%, enabling a deeper understanding of the emotional responses of users towards different digital payment apps.

The visualizations, including word clouds and graphs, offered intuitive representations of the sentiments and emotions identified in the reviews. These visualizations facilitated a better understanding of users’ dominant sentiments and emotions, aiding service providers in identifying patterns, areas for improvement, and user preferences.

The comprehensive analysis and accurate classification of sentiments and emotions in digital payment reviews provide valuable insights for service providers in India. These insights can guide decision-making processes, facilitate enhancements in user experiences, address pain points, and foster innovation in the digital payment sector.

In summary, the sentiment and emotional analysis of digital payment reviews using TextBlob, machine learning algorithms, and the Vader tool, along with the generation of visualizations, offer a comprehensive understanding of user perceptions, sentiments, and emotional responses towards popular payment apps in India. These findings serve as a valuable resource for service providers to optimize their offerings, enhance user satisfaction, and drive advancements in the digital payment landscape.

**CHAPTER - 10**

**FUTURE ENHANCEMENT**

In the future, there are several potential enhancements that can be implemented to further improve the sentiment and emotional analysis of digital payment reviews in India. Firstly, incorporating advanced deep learning models, such as recurrent neural networks (RNNs) or transformer models like BERT (Bidirectional Encoder Representations from Transformers), can potentially enhance the accuracy and nuanced understanding of sentiments and emotions expressed in the reviews. These models have shown remarkable performance in natural language processing tasks and can capture more complex contextual information.

Additionally, integrating user profiling and demographic analysis into the sentiment and emotional analysis can provide deeper insights into how sentiments and emotions vary across different user segments. By considering factors such as age, gender, location, and transaction behavior, service providers can tailor their offerings and address the specific needs and preferences of different user groups.

Furthermore, sentiment and emotional analysis can be extended beyond textual reviews to include other types of user-generated content, such as audio reviews or social media posts. This expansion would enable a more comprehensive understanding of user sentiments and emotions, capturing a wider range of user interactions and experiences.

To enhance the accuracy and reliability of emotional analysis, the inclusion of a larger and more diverse emotional lexicon can be beneficial. Expanding the emotional categories and refining the intensity scoring system can help capture a broader spectrum of emotions and improve the granularity of emotional analysis.

Lastly, integrating real-time sentiment and emotional analysis capabilities into digital payment platforms can enable immediate feedback and response to user sentiments. This can empower service providers to address issues promptly, enhance customer satisfaction, and foster a positive user experience.

Overall, by incorporating advanced models, user profiling, diverse data sources, and real-time capabilities, the sentiment and emotional analysis of digital payment reviews can be enhanced, providing deeper insights into user perceptions and emotions. These enhancements have the potential to guide service providers in making informed decisions, optimizing their platforms, and delivering exceptional user experiences in the evolving landscape of digital payments in India.

**BIBLIOGRAPHY**

[1]. Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining, 11(1), 81.

[2]. Balakrishnan, Vimala, Pik Yin Lok, and Hajar Abdul Rahim. "A semi-supervised approach in detecting sentiment and emotion based on digital payment reviews." The Journal of Supercomputing 77 (2021): 3795-3810.

[3]. Tripto, Nafis Irtiza, and Mohammed Eunus Ali. "Detecting multilabel sentiment and emotions from Bangla youtube comments." 2018 International Conference on Bangla Speech and Language Processing (ICBSLP). IEEE, 2018.

[4]. Bose, Rajesh, et al. "Sentiment analysis on online product reviews." Information and Communication Technology for Sustainable Development: Proceedings of ICT4SD 2018. Springer Singapore, 2020.

[5]. Zhang, Yanrong, et al. "Sentiment analysis of E-commerce text reviews based on sentiment dictionary." 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). IEEE, 2020.

[6]. Frangidis, P., Georgiou, K. and Papadopoulos, S., 2020, May. Sentiment Analysis on Movie Scripts and Reviews: Utilizing Sentiment Scores in Rating Prediction. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 430-438). Cham: Springer International Publishing.

[7]. Yang, Li, et al. "Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning." IEEE Access 8 (2020): 23522-23530.

[8]. Singla, Zeenia, Sukhchandan Randhawa, and Sushma Jain. "Sentiment analysis of customer product reviews using machine learning." 2017 international conference on intelligent computing and Control (I2C2). IEEE, 2017.

[9]. Rajeswari, A. M., et al. "Sentiment analysis for predicting customer reviews using a hybrid approach." 2020 Advanced Computing and Communication Technologies for High-Performance Applications (ACCTHPA). IEEE, 2020.

[10]. Oyelami, Lukman O., Sulaimon O. Adebiyi, and Babatunde S. Adekunle. "Electronic payment adoption and consumers’ spending growth: empirical evidence from Nigeria." Future Business Journal 6 (2020): 1-14.

[11] Jeyapriya, A., and CS Kanimozhi Selvi. "Extracting aspects and mining opinions in product reviews using a supervised learning algorithm." 2015 2nd international conference on Electronics and communication systems (ICECS). IEEE, 2015.

[12] Singh, Mrityunjay, Amit Kumar Jakhar, and Shivam Pandey. "Sentiment analysis on the impact of coronavirus in social life using the BERT model." Social Network Analysis and Mining 11.1 (2021): 33.

[13] Kumar, KL Santhosh, Jayanti Desai, and Jharna Majumdar. "Opinion mining and sentiment analysis on online customer review." 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). IEEE, 2016.

[14] Topal, Kamil, and Gultekin Ozsoyoglu. "Movie review analysis: Emotion analysis of IMDb movie reviews." 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2016.

[15] Haque, Tanjim Ul, Nudrat Nawal Saber, and Faisal Muhammad Shah. "Sentiment analysis on large scale Amazon product reviews." 2018 IEEE international conference on innovative research and development (ICIRD). IEEE, 2018.