A Classification of Airline Reviews and Mobile Reviews.

Malhar Kakade, Sumit Chougale, Niraj Kotve, Shreyash Bhatt, Aditya Acharya

Department of Computer Engineering and Technology
Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India

malharkakade942@gmail.com
sumit.chougale09@gmail.com
neerajkotve19@gmail.com
shreyashbhatt21@gmail.com
adityaacharya.2612@gmail.com

I. Objectives

Classification of Airline Reviews:

1. Sentiment Analysis:
To understand consumer satisfaction and find areas for development, determine the sentiment of airline reviews (positive, negative, or neutral).

2. Competitive Analysis:
Compare the performance of various airlines by categorizing reviews for each, revealing strengths and weaknesses in comparison to competitors.

3. Customer Experience Improvement:
Determine which areas of the customer experience need to be improved, allowing airlines to make data-driven decisions to improve services and overall satisfaction.

4. Marketing Insights:
Learn which components of the airline's service are most well-liked by customers, which can help drive marketing plans and promotional initiatives.

Classification of Mobile Phone Reviews:

1. Feature Evaluation:
Determine the sentiment of Mobile reviews (positive, negative, or neutral) to assess user preferences and prioritize feature development.

2. Market Trends and Competitor Analysis:
Examine reviews for various mobile phone models to detect market trends and do comparative analysis against competitors, which can help guide strategic decisions.

3. User Experience Enhancement:
Classify evaluations based on their relevance to the entire user experience in order to discover pain points and areas for improvement, directing user interface and experience design decisions.

4. Brand Perception:
Analyze reviews to learn how people perceive the brand, allowing businesses to make more educated decisions about branding and marketing initiatives.

These goals can be met by automating the classification of reviews using natural language processing (NLP) and machine learning techniques. This makes it possible to analyze big datasets quickly and derive useful insights.

II. Need

Whether evaluations are for mobile phones or airplanes, categorizing them fulfills a number of vital functions that are advantageous to companies, customers, and other stakeholders. It gives companies the ability to make data-driven decisions, which helps them improve operations, respond to customer input, and stay competitive in the market. It facilitates comprehension of:

1. Customer Sentiment.
2. Operational Improvements.
3. Competitor Benchmarking.
4. Brand Perception and Reputation Management.
5. Market Research and Trends.
6. Customer Relationship Management.

III. Description

The issue at hand is the analysis of unstructured textual data in the form of reviews, with a particular emphasis on those pertaining to mobile phones and airlines. Finding *relevant information from a plethora of subjective and varied client feedback presents a difficulty. The task entails using classification algorithms to group evaluations according to emotion in order to address this. The goal for airlines is to improve overall service quality and comprehend client satisfaction. When it comes to smartphones, the goal is to assess customer input and guarantee a satisfying user experience. Businesses can better serve their customers, increase operational effectiveness, and maintain their competitiveness in their particular industries by effectively classifying and analyzing feedback.*

IV. List of Functional Requirements

System Input and Data Processing:

1. Data Ingestion:
Text files or databases should be able to provide the system with access to a vast amount of textual data, including reviews of mobile phones and airlines.
2. Data Preprocessing:
Use preprocessing methods, such as stop word removal, to clean and standardize the textual data.

Classification and Analysis:
3. Sentiment Analysis:
Create a sentiment analysis module to determine consumer sentiment by categorizing reviews as neutral, positive, or negative.

User Interface and Reporting:
4. Interactive Dashboard:
Create an intuitive and interactive dashboard that enables users to view categorization results, enter criteria, and visually explore findings.

Automation and Scalability:
5. Automated Classification:
Utilize machine learning techniques to automate the classification process and assess new reviews as they come in.

Performance and Accuracy:
6. Performance Metrics:
Use performance measurements (precision, recall, and accuracy) to evaluate the classification model's efficacy and make ongoing improvements.

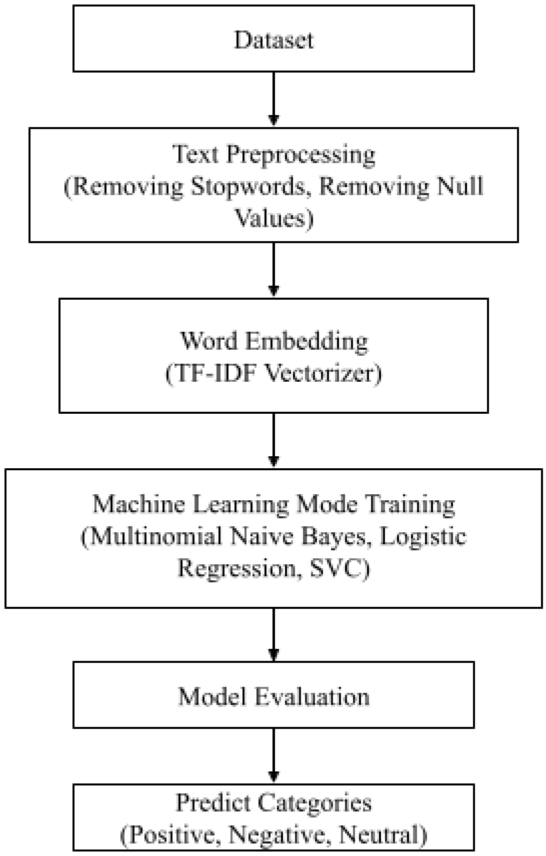
These functional criteria are meant to serve as a roadmap for building a solid system that can efficiently categorize and examine evaluations from mobile phones and airplanes, offering insightful information for attempts to improve and make decisions.

V. Introduction

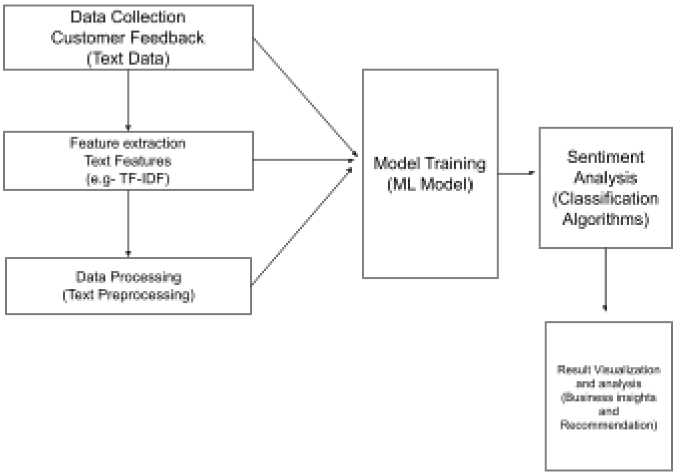
The project's main focus is using machine learning and natural language processing to categorize evaluations from mobile phones and airlines. The objective is to derive valuable insights from the analysis of unstructured textual data, such customer reviews. Sentiment analysis is the main focus for airlines, but improving the entire user experience and assessing user feedback are the goals for mobile phones. The goal of the classification process is to give organizations in these sectors organized, useful insights that will help them make data-driven decisions.

VI. System Design

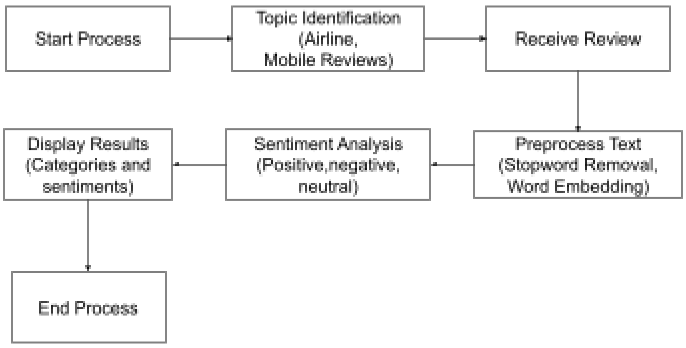
1. Simple Block Diagram:



2. System Architecture Diagram:



3. UML Activity Diagram:



VII. Literature Survey

No.	Title of Paper	Year	Methods Used	Dataset Used	Results Achieved	Research Gaps/Improvement
1	On The Feature Extraction For Sentiment Analysis of Movie Reviews Based On SVM	2020	Preprocessing: Remove Punctuation Case Folding Tokenization StopWord Removal Lemmatization Feature Extraction: Term Frequency - Inverse Document Frequency (TF-IDF) Classification: Support Vector Machine (SVM)	The Internet Movie Dataset (IMDB) 2000 Movie Review 1000-Positive 1000-Negative	This research produced the best performance on a combination of TF-IDF and LDA, with 240 topics and 29792 features, which is 82.16%.	Using two feature extraction methods increases computational time compared to using only TF-IDF. Multinomial Naive Bayes (MNB) can be used here since it performs better on snippets, whereas Support Vector Machine (SVM) is more suitable for full-length content. Also, SVM is considered one of the best text classification algorithms, but it's a little slower than MNB.
2	Twitter Sentiment Analysis with Machine Learning	2022	Preprocessing: Tokenization Lemmatization Stemming Feature Extraction: Term Frequency - Inverse Document Frequency (TF-IDF) Classification: Support Vector Machine (SVM) Random Forest Classifier logistic regression Decision tree classifier XGB Classifier	Twitter Dataset	The performance of various machine learning based classifiers are compared with Twitter data in this study. Additionally, various methods of machine learning algorithms are used to predict results with different levels of accuracy.The highest accuracy being 95% of both Random Forests and SVM	The accuracy shown by all the methods is too high thus there is a possibility of overfitting.We can prevent that by giving more importance to reduce data leakage

3	A Study on Sentiment Analysis of Product Reviews	2018	DataProcessing: Tokenization Lemmatization Feature Extraction: Term Frequency-Inverse Document Frequency (TF-IDF) Classification: Lexicon approach Naïve Bayes Decision Tree Random Forest Support Vector Machine	Product Reviews Dataset	The results achieved in these papers include the development of different approaches and techniques for sentiment analysis, such as machine learning-based and lexicon-based approaches, and the comparison of different techniques for sentiment analysis. The papers also discuss the importance of sentiment analysis in various fields, such as business intelligence, recommender systems, managing brand reputation, politics, and government intelligence. Overall, the research papers provide valuable insights into sentiment analysis and its applications.	study on detecting fake reviews using machine learning techniques" discusses the need for more accurate techniques to detect fake reviews, which can be a major challenge in sentiment analysis
4	Analysis of Twitter Sentiments using Machine Learning to Identify Polarity	2022	Preprocessing: Tokenization Stop Words Lemmatization Remove Punctuation Whitespace Removal Sentiment Classification: Confidence score Classification: Naive-Bayes	Dataset not mentioned	1)The study observed a massive number of improvements with respect to prior technique approaches, which closes the gap with supervised feature learning. 2)The efficiency of models has increased considerably, reaching almost 85%-90% in instances	The study concluded that the approach is a good classifier, but on the other hand, it is a weak computer, which is the fundamental drawback of the NB Algorithm. Therefore, future research can focus on improving the accuracy of the model and overcoming the limitations of the NB Algorithm.

5	Sentiment Analysis of Yelp Reviews By Machine Learning	2019	Preprocessing: Tokenization Stop Words Lemmatization Remove Punctuation Whitespace Removal Classification: Naive-Bayes (Multinomial Naive Bayes,Bernoulli Naive Bayes) Regression (Logistic Regression) Linear SVC (Support Vector Clustering)	Yelp Reviews Dataset	1) Analysis of the results shows that we have successfully been able to get a satisfactory level of accuracy in the classification of Yelp reviews using supervised learning. 2) Satisfactory Results have been obtained for simple English sentences as mode of the outputs of the algorithms are used for deciding sentiment. 3) Long and Descriptive sentences are encouraged in order to get better and more accurate results.	Some limitations of the project include accuracy and less diversity of categorization.The training set can never be perfect as it is not an easy task to separate 20,000 reviews into good or bad just by reading them.The classification is binary and gives an overall result of the goodness or badness of the review.The future scope is to rate businesses based on different features,to use better and more dataset to train and to detect sarcasm.

VIII. Dataset Used

Flight Reviews:

Dataset taken from github repository.

Scrapped Review Data of Airline:

Created a Dataset from the scraped data collected from airlinequality.com.

Mobile Reviews:

Cell Phones Reviews Sentiment Analysis Dataset taken from kaggle.com.

A) Data Visualization

Flight Reviews:

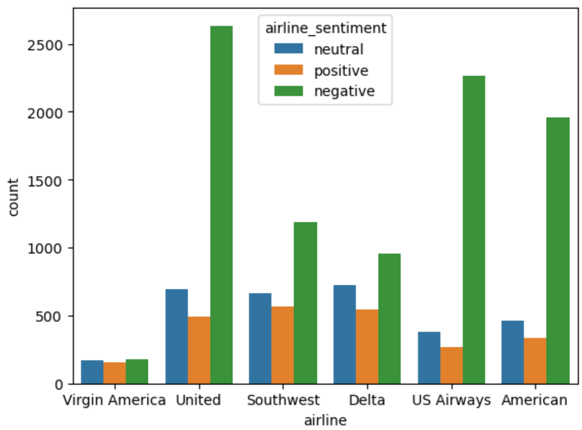


Fig. 1

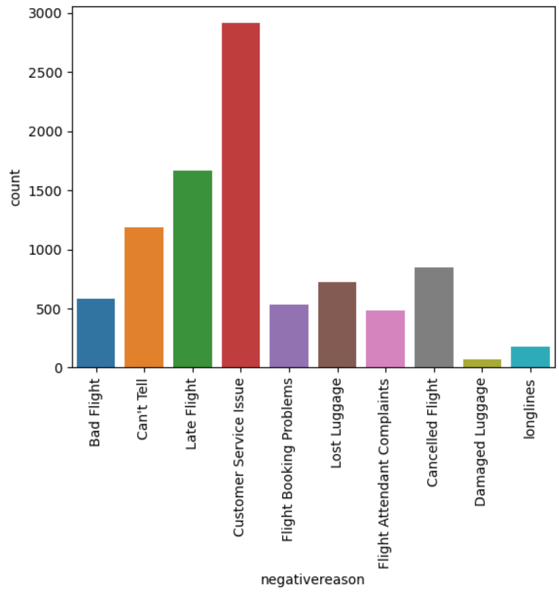


Fig. 2

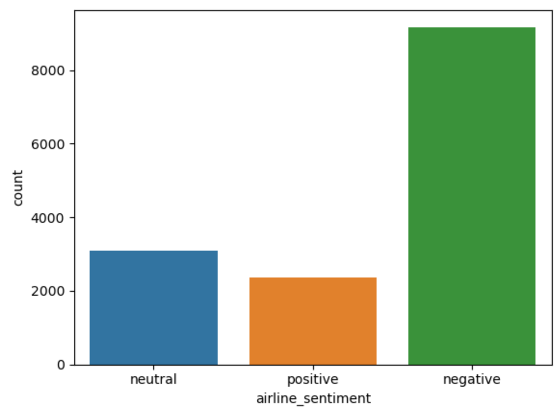


Fig. 3

Scrapped Review Data of Airline:

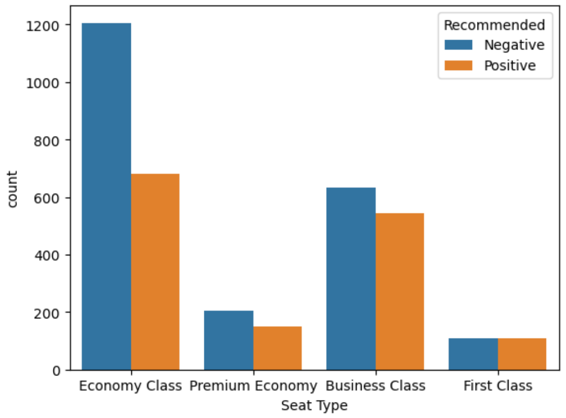


Fig. 4

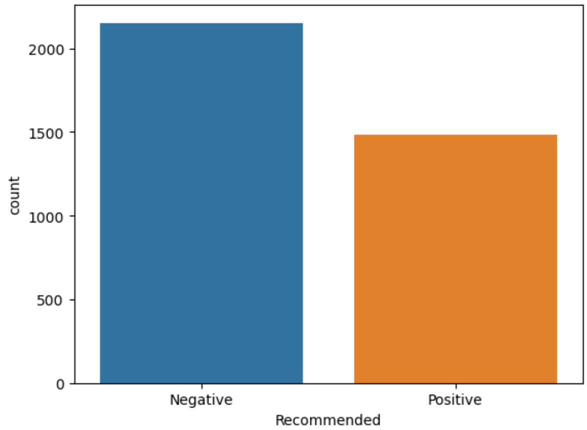


Fig. 5

Mobile Review:

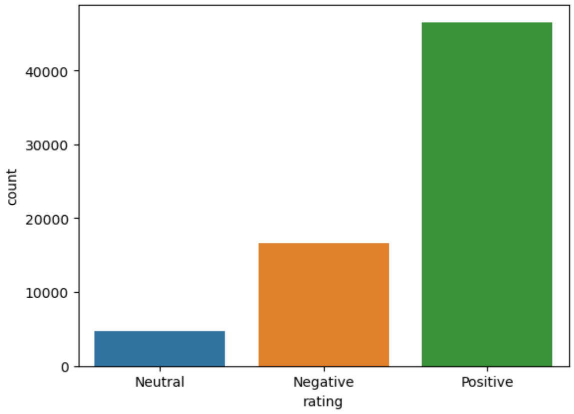


Fig. 6

B) PREPROCESSING METHODS USED

Removing Null Values (Dropna):

Using Pandas' "Dropna" method to remove null values from a dataset has various advantages that improve data quality, analysis, and modeling procedures in general. This is another place to utilize "Fillna."

Elimination of stop words and conversion of unprocessed data into numerical representation appropriate for machine learning algorithms (TF-IDF Vectorizer):

A popular feature extraction method in text mining and natural language processing (NLP) is the TF-IDF Vectorizer. The TF-IDF Vectorizer is a tool that converts a set of unprocessed documents into a numerical representation appropriate for machine learning methods. TF-IDF stands for Term Frequency-Inverse Document Frequency. Additionally, stop words in Common English (such as "the," "and," and "is") are not included in the tokenization process. Stop words are frequently eliminated in favor of more significant words that help

IX. ARE YOU CREATING A CORPUS FROM SCRATCH

Yes, we started from zero while creating the airline reviews dataset. For site scraping, we used the Beautiful Soup Python library. It is employed to extract data from XML and HTML files. In order to scrape and extract data from websites, Beautiful Soup is frequently used in conjunction with other libraries such as Requests. It gives developers an easier way to navigate and work with HTML and XML documents using Python, which makes web scraping chores more manageable.

Challenges Faced: We faced difficulties in web scraping the mobile reviews data. Mostly the data available was unlabelled. Also, it was very difficult to web scrape all the reviews of various different mobiles, as in amazon we were able to access reviews of only one mobile at a time. Thus to web scrape the data of all different mobiles was difficult.

X. METHODS USED FOR CLASSIFICATION

Three machine learning models are used to categorize the reviews. These models are essential for sentiment analysis since they help analyze textual data from reviews. By using a variety of models, it is possible to identify the optimal model for classification and improve the process's efficacy and accuracy. The following machine learning algorithms are employed:

1. MULTINOMIAL NAIVE BAYES:

Naive Multinomial Based on the Bayes theorem, the Bayes algorithm is a probabilistic categorization tool. It works especially effectively for text categorization jobs like sentiment analysis and spam detection. It makes the assumption that terms are conditionally independent and models the probability of observing a term given its class in a text. The technique computes probabilities and generates predictions based on the frequency of phrases, which are commonly represented as a bag-of-words.

2. *Logistic Regression:*

One statistical technique for binary classification problems is logistic regression. It simulates the likelihood that an instance is a member of a specific class. The approach converts a linear combination of input information into a range between 0 and 1, which represents the probability of belonging to the positive class. This is done by applying a logistic (sigmoid) function to the input features.

3. *Support Vector Classification (SVC):*

A machine learning technique called support vector classification, or support vector machines (SVM) for classification, is utilized for both binary and multi-class classification tasks. Because of its adaptability and strong performance, it is frequently used for jobs like text categorization, picture classification, and other classification issues.

XI. Results

For sentiment analysis, three machine learning models were employed. The results of these models demonstrated that Support Vector Classification (SVC) outperformed the other two machine learning models in its ability to classify the reviews. The following lists the specifics of the different evaluation parameters:

Airline Reviews:

Multinomial NB : 67.31%

Classification Report:

NB MODEL

	precision	recall	f1-score	support
negative	0.66	0.99	0.79	1817
neutral	0.79	0.15	0.26	628
positive	0.89	0.14	0.24	483
accuracy			0.67	2928
macro avg	0.78	0.43	0.43	2928
weighted avg	0.73	0.67	0.59	2928

0.673155737704918

Fig. 7

Confusion Matrix:

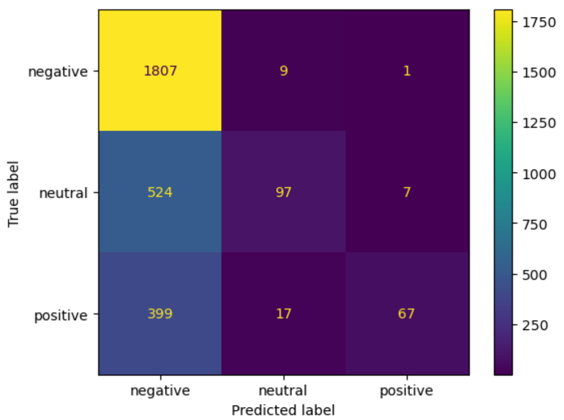


Fig. 8

Linear Regression : 77.45%

Classification Report:

Logistic Regression

	precision	recall	f1-score	support
negative	0.80	0.93	0.86	1817
neutral	0.63	0.47	0.54	628
positive	0.82	0.58	0.68	483
accuracy			0.77	2928
macro avg	0.75	0.66	0.69	2928
weighted avg	0.77	0.77	0.76	2928

0.7745901639344263

Fig. 9

Confusion Matrix:

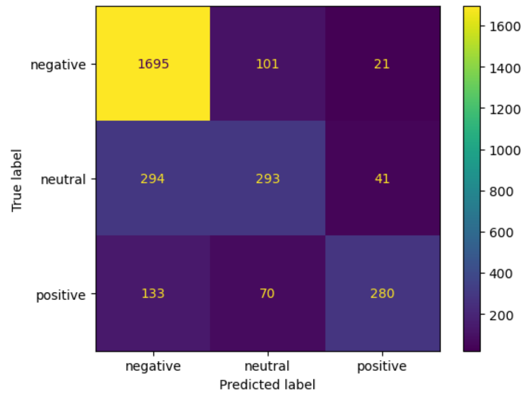


Fig. 10

SVC : 77.15%

Classification Report:

SVC

	precision	recall	f1-score	support
negative	0.82	0.89	0.86	1817
neutral	0.59	0.52	0.55	628
positive	0.76	0.64	0.69	483
accuracy			0.77	2928
macro avg	0.73	0.68	0.70	2928
weighted avg	0.76	0.77	0.77	2928

0.7715163934426229

Fig. 11

Confusion Matrix:

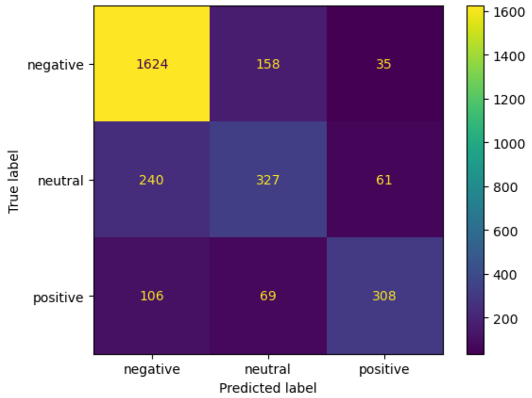


Fig. 12

Scrapped Review Data of Airline:

Multinomial NB : 82.96%

Classification Report:

```
NB MODEL
              precision    recall  f1-score   support

           0       0.79      0.97      0.87       437
           1       0.94      0.61      0.74       291

    accuracy                           0.83       728
   macro avg       0.87      0.79      0.81       728
weighted avg       0.85      0.83      0.82       728

0.8296703296703297
```

Fig. 13

Confusion Matrix:

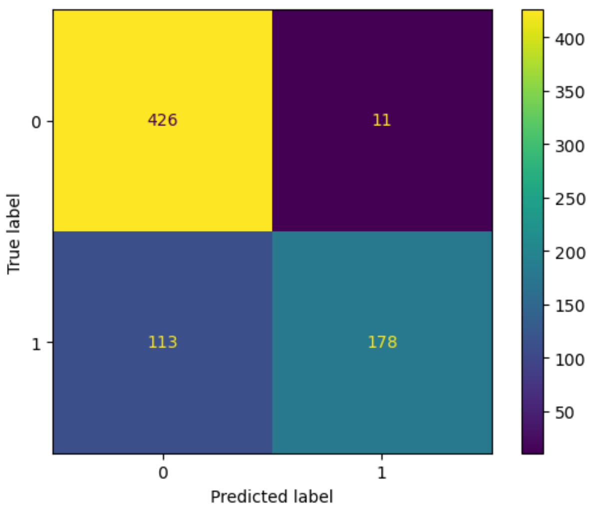


Fig. 14

AUC-ROC Curve:

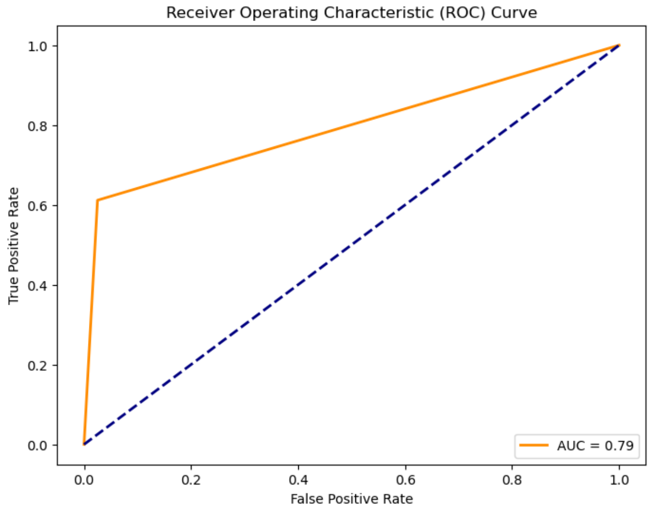


Fig. 15

Linear Regression : 87.08%

Classification Report:

```
Logistic Regression
              precision    recall  f1-score   support

           0       0.86      0.93      0.90       437
           1       0.89      0.78      0.83       291

    accuracy                           0.87       728
   macro avg       0.87      0.86      0.86       728
weighted avg       0.87      0.87      0.87       728

0.8708791208791209
```

Fig. 16

Confusion Matrix:

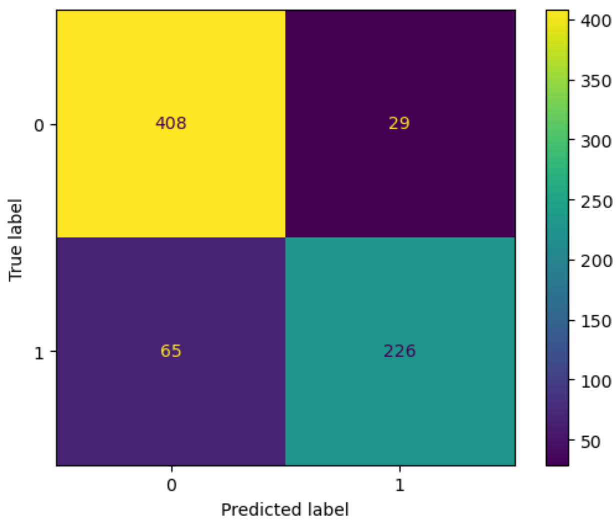


Fig. 17

AUC-ROC Curve:

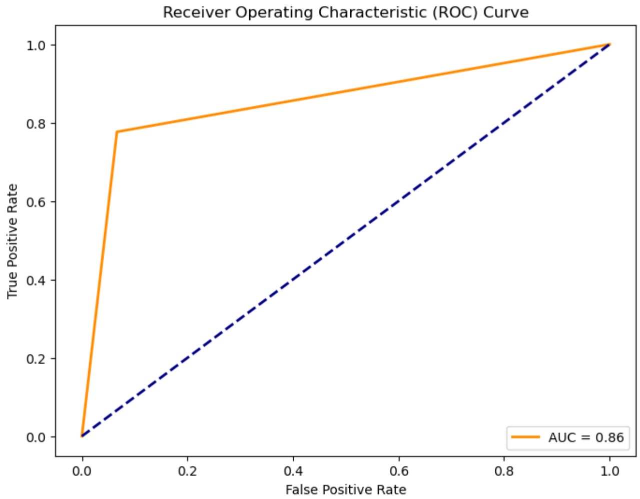


Fig. 18

SVC : 88.59%

CLASSIFICATION REPORT:

```
SVC
              precision    recall  f1-score   support

           0       0.88      0.93      0.91       437
           1       0.89      0.82      0.85       291

    accuracy                           0.89       728
   macro avg       0.89      0.87      0.88       728
weighted avg       0.89      0.89      0.89       728

0.885989010989011
```

Fig. 19

Confusion Matrix:

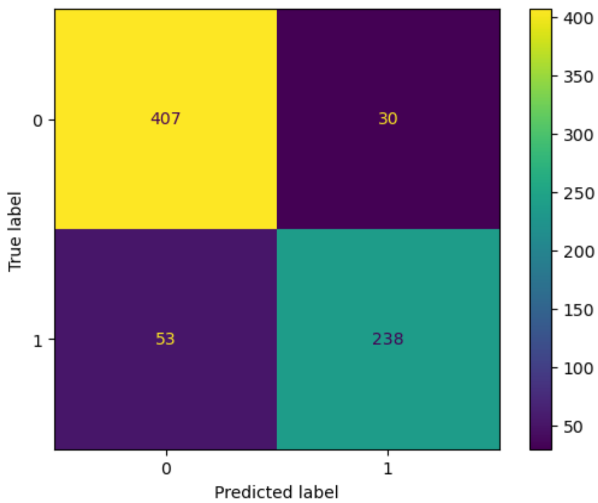


Fig. 20

AUC-ROC CURVE:

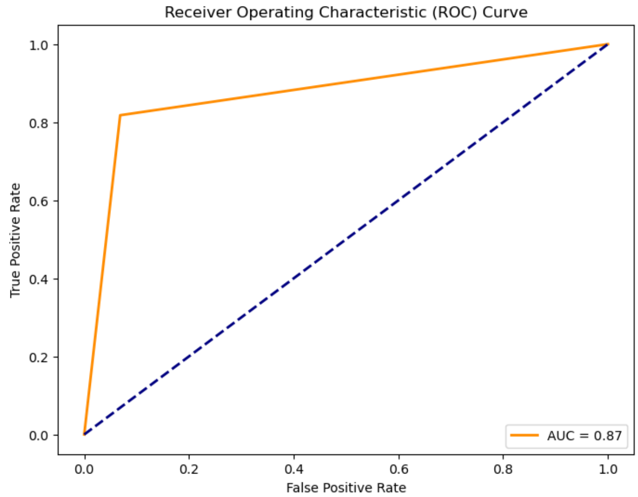


Fig. 21

MOBILE REVIEWS:

MULTINOMIAL NB : 80.52%

CLASSIFICATION REPORT:

```
NB MODEL
              precision    recall  f1-score   support

    Negative       0.86      0.53      0.65      3325
     Neutral       0.00      0.00      0.00       954
    Positive       0.80      0.99      0.88      9312

    accuracy                           0.81     13591
   macro avg       0.55      0.50      0.51     13591
weighted avg       0.76      0.81      0.76     13591

0.8052387609447429
```

Fig. 22

Confusion Matrix:

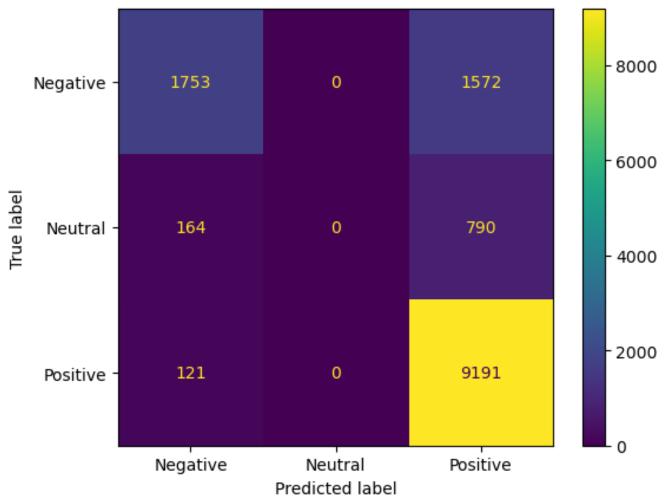


Fig. 23

LINEAR REGRESSION : 85.66%

CLASSIFICATION REPORT:

```
Logistic Regression
              precision    recall  f1-score   support

    Negative       0.78      0.81      0.80      3325
     Neutral       0.35      0.03      0.06       954
    Positive       0.89      0.96      0.92      9312

    accuracy                           0.86     13591
   macro avg       0.67      0.60      0.59     13591
weighted avg       0.82      0.86      0.83     13591

0.8566698550511368
```

Fig. 24

Confusion Matrix:

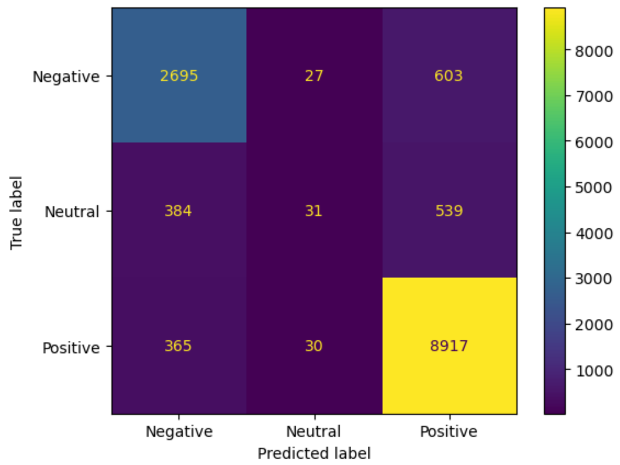


Fig. 25

SVC : 85.57%

CLASSIFICATION REPORT:

```
SVC
              precision    recall  f1-score   support

    Negative       0.77      0.82      0.79      3325
     Neutral       0.43      0.07      0.12       954
    Positive       0.89      0.95      0.92      9312

    accuracy                           0.86     13591
   macro avg       0.70      0.61      0.61     13591
weighted avg       0.83      0.86      0.83     13591

0.8557133397101023
```

Fig. 26

Confusion Matrix:

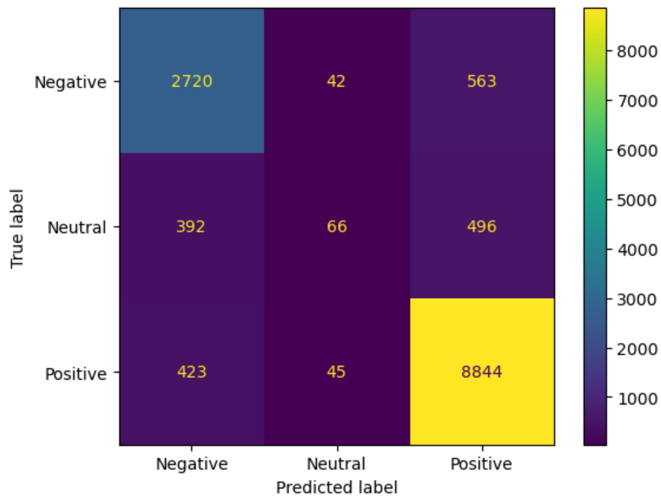


Fig. 28

XII. INDIVIDUAL CONTRIBUTION

NAME	*CONTRIBUTION*
MALHAR KAKADE	Mobile Review Classification , Data Visualization , Report
NIRAJ KOTVE	Data Scraping, Scrapped Airline Review Classification , Report
SHREYASH BHATT	Airline Review Classification , Report
SUMIT CHOUGALE	Airline Review Classification , Data Visualization , Report
ADITYA ACHARYA	Scrapped Airline Review Classification , Report

XIII. REFERENCES

[1] F. E. Cahyanti, Adiwijaya and S. A. Faraby, "On The Feature Extraction For Sentiment Analysis of Movie Reviews Based on SVM," 2020 8th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 2020.

[2] L. Mandloi and R. Patel, "Twitter Sentiments Analysis Using Machine Learninig Methods," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020.

[3] A. S. Parihar and Bhagyanidhi, "A Study on Sentiment Analysis of Product Reviews," 2018 International Conference on Soft-computing and Network Security (ICSNS), Coimbatore, India, 2018.

[4] P. K, N. S, S. RJ and N. P, "Analysis of Twitter Sentiments Using Machine Learning to Identify Polarity," 2022 1st International Conference on Computational Science and Technology (ICCST), CHENNAI, India, 2022.

[5] H. S. and R. Ramathmika, "Sentiment Analysis of Yelp Reviews by Machine Learning," 2019 International Conference on Intelligent

Computing and Control Systems (ICCS), Madurai, India, 2019.