# Sentiment Analysis in Product Reviews Classification

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

**Bachelor of Technology**

In

**Computer Science and Engineering**

**School of Engineering and Sciences**


**SUBMITTED BY**

**V.Venkat Raja - AP19110010420**

**P.Sai Janith-AP19110010321**

**M.Niteesh-AP19110010326**


Under the Guidance of

**Dr.** Satya Krishna Nunna


**SRM University–AP**

**Neerukonda, Mangalagiri, Guntur**

**Andhra Pradesh – 522 240**


**[May - 2023]**

# Certificate

Date: 28-Apr-23

This is to certify that the work present in this Project entitled "**Sentiment Analysis in Product Reviews Classification**" has been carried out by M.**Nitheesh, V.Venkat Raja, P.Sai Janith** under my/our supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology/Master of Technology in **School of Engineering and Sciences**.

**Supervisor**

(Signature)

**Dr.Satya Krishna Nunna**

Designation,

Affiliation.

# Acknowledgements

We would like to express our sincerest gratitude and appreciation to our mentor, **Dr.Satya Krishna Nunna**, for his valuable guidance, unwavering support, and exceptional mentorship throughout our final semester project. His constant encouragement, insightful suggestions, and constructive feedback have significantly contributed to the growth of our technical skills and understanding of the subject matter. His guidance and expertise have played a vital role in helping us navigate challenges and ensuring the project's smooth progress.

We would also like to extend our appreciation to the Head of the Department, **Dr. Jatindra Kumar Dash**, for his support and encouragement throughout our final semester.

Lastly, we acknowledge that without the cooperation and combined efforts of everyone involved, this project would not have been possible.

As we conclude our final semester, we carry with us the knowledge, skills, and experiences gained from this project. We are grateful for the opportunities provided and remain committed to applying our learning in our future endeavors.

# TABLE OF CONTENTS

CERTIFICATE ..... I

ACKNOWLEDGEMENTS ..... II

TABLE OF CONTENTS ..... III

ABSTRACT ..... IV

ABBREVIATIONS ..... VI

LIST OF TABLES ..... VII

LIST OF FIGURES ..... VIII

LIST OF EQUATIONS ..... IX

1. INTRODUCATION ..... 1
2. LITERATURE REVIEW ..... 5
3. PRELIMINERIES ..... 7
4. METHODOLOGY ..... 9
    - EXPERIMENTAL SETUP ..... 9
    - DATA COLLECTION ..... 10
    - DATA PRE PROCESSING ..... 10
    - FEATURE SELECTION ..... 10
    - MODEL BUILDING ..... 11
    - SENTIMENT CLASSIFICATION ..... 11
    - PERFORMANCE EVALUATION ..... 11
5. RESULTS AND DISCUSSIONS ..... 31
6. CONCLUDING REMARKS ..... 35
7. FUTURE WORK ..... 36
8. REFERENCES ..... 37

# Abstract

The advent of e-commerce platforms and the proliferation of online shopping have significantly transformed the consumer landscape. With the vast amount of product information available on the internet, consumers heavily rely on product reviews to make informed purchasing decisions. As a result, sentiment analysis has emerged as a valuable tool for businesses to extract meaningful insights from the vast volume of textual data generated through product reviews.

This report focuses on sentiment analysis specifically applied to product reviews, aiming to explore the efficacy of this technique in understanding customer sentiments, identifying key factors driving positive or negative feedback, and extracting actionable intelligence for product improvement and marketing strategies. By analyzing sentiment patterns and opinions expressed in product reviews, businesses can gain a deeper understanding of customer satisfaction, identify areas for improvement, and optimize their products and services accordingly.

The report begins by providing an overview of sentiment analysis and its significance in the context of product reviews. We discuss the various methodologies and techniques employed in sentiment analysis, including rule-based approaches, lexicon-based methods, and machine learning-based algorithms. Special emphasis is placed on the application of machine learning models, such as supervised and unsupervised learning, in training sentiment analysis systems on annotated datasets.

Furthermore, we delve into the challenges and complexities of sentiment analysis on product reviews. We explore the impact of factors such as language ambiguity, sarcasm, and contextual understanding, and discuss techniques employed to address these challenges, such as domain adaptation and sentiment lexicon expansion. Ethical considerations, including privacy concerns and bias mitigation, are also examined to ensure responsible usage of sentiment analysis in the realm of product reviews.

The report also highlights the practical implications of sentiment analysis on product reviews across various industries. We explore how sentiment analysis can be leveraged for reputation management, brand monitoring, competitive analysis, and customer feedback analysis. Furthermore, we delve into the role of sentiment analysis in understanding market trends, predicting consumer behavior, and supporting decision-making processes for product development and marketing campaigns.

To validate the effectiveness of sentiment analysis on product reviews, we present case studies and empirical evidence from real-world applications. By showcasing successful implementations and their outcomes, we demonstrate how sentiment analysis can uncover invaluable insights and drive tangible business benefits. Moreover, we discuss limitations and

future directions, exploring emerging techniques and trends that have the potential to further enhance the accuracy and applicability of sentiment analysis on product reviews.

In conclusion, sentiment analysis applied to product reviews offers a powerful means of extracting customer sentiments and opinions. By analyzing and understanding the sentiments expressed in product reviews, businesses can make data-driven decisions, optimize their products, and create more satisfying customer experiences. This report serves as a comprehensive guide, providing insights, methodologies, and practical applications of sentiment analysis on product reviews, enabling organizations to harness the wealth of information within customer feedback and gain a competitive edge in the marketplace.

# Abbreviations

| | |
|---|---|
| CNN | Computational Neural Network |
| KNN | K-Nearest Neighbors |
| RNN | Recurrent Neural Network |
| RF | Random Forest |

# List of Tables

Table 1. Description of Dataset 1 ..... 31
Table 2. Description of Dataset 2 ..... 31
Table 3. Description of Dataset 3 ..... 32
Table 4. Description of Dataset 4 ..... 32

# List of Figures

Figure 1. Precision Recall Curve for dataset 1……....13

Figure 2. Histogram of Predicted Probabilities for dataset 1…..... 13

Figure 3. Receiver Operating Characterstic for dataset 1 ..... 14

Figure 4. Cross Validation Scores for dataset 2 ..... 15

Figure 5. Model Performance for dataset 2..... 16

Figure 6. Top 10 Most Common Words in Sentiment for dataset 3..... 20

Figure 7. Histogram of Product Price for dataset 3 ..... 20

Figure 8. Count of Unique Values in Rate ..... 21

Figure 9. Precision, Recall and F1-score..... 21

# List of Equations

Equation 1. Accuracy ..... 22
Equation 2. Precision ..... 22
Equation 3. Recall ..... 22
Equation 4. F1-Score ..... 22
Equation 5. Misclassification rate… ..... 22
Equation 6. Specificity ..... 22

# 1.Introduction

In today's data-driven world, where vast amounts of information are generated every second, understanding human sentiment has become increasingly vital for various industries. Sentiment analysis, also known as opinion mining, is a fascinating field at the intersection of natural language processing, machine learning, and data analytics. It focuses on extracting and analyzing emotions, opinions, and attitudes from textual data, enabling organizations to gain valuable insights into public perception.

Sentiment analysis has gained significant attention in recent years due to its wide-ranging applications across different domains. From business and marketing to social media monitoring and customer service, organizations are leveraging sentiment analysis techniques to better understand how people feel about their products, services, or brand.

At its core, sentiment analysis aims to determine the sentiment expressed in a given text, whether it is positive, negative, or neutral. This task involves processing and interpreting unstructured textual data, such as customer reviews, social media posts, news articles, and survey responses. By automatically analyzing these vast volumes of textual data, sentiment analysis allows businesses and researchers to gauge public opinion on specific topics, track trends, identify emerging issues, and make informed decisions.

The advancements in natural language processing and machine learning algorithms have revolutionized sentiment analysis, enabling more accurate and efficient sentiment classification. Traditional approaches relied on rule-based systems and lexical analysis, which involved the use of predefined dictionaries and linguistic rules. However, the advent of machine learning techniques, particularly supervised learning models like support vector machines, decision trees, and deep learning models such as recurrent neural networks and transformers, has led to significant improvements in sentiment analysis accuracy.
Beyond sentiment polarity classification (positive, negative, neutral), sentiment analysis has evolved to capture more nuanced aspects of emotions. Fine-grained sentiment analysis goes beyond simple binary classification and focuses on aspects like sentiment intensity, emotions (such as joy, sadness, anger), and even aspect-based sentiment analysis, where sentiments are associated with specific aspects or features of a product or service.

The potential applications of sentiment analysis are wide-ranging. In marketing, sentiment analysis helps companies evaluate customer feedback, identify customer preferences, and tailor their marketing strategies accordingly. Customer service departments can use sentiment analysis to track customer sentiment in real-time, enabling prompt responses and effective issue resolution. In financial markets, sentiment analysis can provide valuable insights into investor sentiment and predict market trends. It can also aid in

political analysis, reputation management, brand monitoring, and identifying emerging trends and sentiments in social media.

This report aims to provide an in-depth exploration of sentiment analysis, its methodologies, challenges, and applications. By delving into the core principles and techniques of sentiment analysis, we will uncover how this powerful tool unlocks hidden value in textual data, providing actionable insights for decision-makers. Additionally, we will discuss the limitations and ethical considerations associated with sentiment analysis, highlighting the need for careful interpretation and responsible usage.

By shedding light on sentiment analysis, this report seeks to demonstrate its potential to revolutionize the way we understand human emotions, enhance customer experiences, and drive data-informed decision-making across industries.
In an increasingly interconnected and digitized world, the explosion of user-generated content on various online platforms has provided an unprecedented opportunity to gain insights into public opinion and sentiments. Sentiment analysis, also known as opinion mining, has emerged as a vital field within the domain of natural language processing (NLP) and artificial intelligence (AI). By leveraging computational algorithms and machine learning techniques, sentiment analysis enables us to automatically classify and quantify the subjective information embedded within textual data, revealing valuable insights into the emotional states of individuals or communities.

The ability to discern and understand human sentiments from vast amounts of unstructured text data has immense applications across numerous domains. From market research and brand management to political analysis and customer feedback analysis, sentiment analysis has become a powerful tool for organizations to extract actionable intelligence and make data-driven decisions. By gauging the polarity of sentiments expressed in social media posts, product reviews, news articles, and other textual sources, sentiment analysis empowers businesses, researchers, and policymakers with the ability to gauge public opinion, detect emerging trends, and respond proactively to sentiment shifts.
The impact of sentiment analysis extends beyond the realm of business and marketing. It has also found applications in social sciences, political science, psychology, and healthcare. By analyzing public sentiments towards political candidates, policies, or social issues, sentiment analysis can provide valuable insights for election forecasting, policy evaluation, and public opinion monitoring. In healthcare, sentiment analysis of patient feedback and social media posts can aid in monitoring mental health, identifying patterns of distress, and improving healthcare services.

As sentiment analysis continues to evolve, its potential for societal impact and commercial value will only grow. The ability to accurately gauge and understand human emotions from text data can empower organizations, governments, and individuals to make informed decisions, tailor communication strategies, and respond promptly to emerging trends and concerns. However, it is crucial to address ethical considerations, such as

privacy concerns, bias mitigation, and responsible use of sentiment analysis outputs, to ensure its ethical and responsible implementation.

In this report, we delve into the intricacies of sentiment analysis, exploring its methodologies, challenges, and applications across various domains. We examine the latest advancements in machine learning and deep learning techniques, discuss the limitations and future directions of sentiment analysis, and highlight the ethical considerations that accompany its implementation. By shedding light on the significance and potential of sentiment analysis, we aim to contribute to the understanding and utilization of this powerful tool in extracting valuable insights from the ever-expanding realm of textual data.

Recent advancements in sentiment analysis have been fueled by the availability of large-scale datasets and the continuous progress of computing power, enabling more accurate and efficient sentiment classification. Social media platforms, in particular, have become a treasure trove of real-time, user-generated content that reflects people's opinions, experiences, and emotions. The ability to analyze sentiments expressed on platforms like Twitter, Facebook, and Instagram has opened up new avenues for understanding public sentiment on a global scale.

One of the key factors driving the popularity of sentiment analysis is its potential to revolutionize customer experience management. By automatically processing and analyzing customer feedback, sentiment analysis allows businesses to gain insights into the satisfaction levels, preferences, and concerns of their customers. This information can be invaluable in improving products and services, tailoring marketing strategies, and building stronger customer relationships. Sentiment analysis also enables companies to detect and address potential issues before they escalate, leading to better reputation management and crisis response.

In the era of information overload, sentiment analysis serves as a powerful filtering mechanism, enabling individuals to navigate through vast amounts of textual data and identify relevant and meaningful information. By automatically categorizing content as positive, negative, or neutral, sentiment analysis can save time and effort in information retrieval, decision-making processes, and content curation. This capability is especially beneficial for media organizations, researchers, and content creators who can leverage sentiment analysis to identify trending topics, gauge public opinion on specific issues, and tailor their content to meet the preferences of their target audience.

However, sentiment analysis is not without its limitations and challenges. The nuances of human emotions and the complexity of language make it difficult to achieve perfect accuracy in sentiment classification. Ambiguities, sarcasm, cultural differences, and context-dependent sentiments pose hurdles for sentiment analysis algorithms. Additionally, the inherent subjectivity of sentiment analysis introduces a level of bias that needs to be carefully addressed and mitigated to ensure fair and unbiased analysis.

As sentiment analysis progresses, researchers are exploring innovative avenues to enhance its capabilities. The integration of multimodal data, such as combining textual analysis with image or video analysis, holds promise in capturing a more comprehensive understanding of sentiments expressed across different modalities. Additionally, advancements in deep learning, natural language generation, and explainability techniques are poised to improve the interpretability and reliability of sentiment analysis models.

In conclusion, sentiment analysis is a rapidly evolving field that holds immense potential for understanding human emotions, opinions, and attitudes from textual data. Its applications span across industries and disciplines, empowering organizations and individuals to make data-driven decisions, anticipate trends, and enhance customer experiences. By addressing the challenges and advancing the techniques of sentiment analysis, we can unlock its full potential and pave the way for a more insightful and emotionally intelligent future.

# 2.Literature Review

Sentiment analysis of product reviews has gained significant attention in the field of natural language processing and machine learning. In recent years, with the proliferation of e-commerce platforms and online shopping, the ability to automatically analyze the sentiment expressed in product reviews has become crucial for businesses to gain insights into customer opinions and make data-driven decisions. Numerous studies have focused on developing and improving sentiment analysis techniques to accurately classify product reviews as positive, negative, or neutral. These approaches employ various machine learning algorithms, such as support vector machines, naive Bayes, and deep learning models like recurrent neural networks and convolution neural networks. Additionally, researchers have explored different feature extraction methods, including bag-of-words, n-grams, and word embeddings, to capture the contextual information and semantic meaning of the text. The evaluation of sentiment analysis models often involves benchmark datasets like the Amazon Product Reviews dataset, Yelp dataset, or IMDB dataset, which enable researchers to compare the performance of different algorithms. Furthermore, advancements in sentiment analysis have led to the development of domain-specific models, targeting specific product categories such as electronics, books, or clothing, to improve the accuracy of sentiment classification. Overall, sentiment analysis of product reviews has proven to be a valuable tool for businesses, enabling them to understand customer sentiments, identify areas of improvement, and enhance customer satisfaction.

In addition to the advancements in sentiment analysis techniques, researchers have also explored various aspects related to product reviews. One area of focus has been aspect-based sentiment analysis, where the goal is to identify and analyze the sentiment associated with specific aspects or features of a product mentioned in the reviews. This fine-grained analysis provides deeper insights into customer preferences and helps businesses address specific concerns or highlight standout features.

Another important aspect that researchers have investigated is the effect of sentiment analysis on recommendation systems. By incorporating sentiment analysis into recommendation algorithms, personalized product recommendations can be generated based not only on the users' historical preferences but also on their sentiment towards certain products or features. This integration enhances the accuracy and relevance of the recommendations, leading to improved customer satisfaction and potentially increased sales.

Furthermore, researchers have examined the challenges and limitations of sentiment analysis in the context of product reviews. One common challenge is dealing with the presence of sarcasm, irony, or subjective language, which can affect the accuracy of sentiment classification. Additionally, the handling of noisy or unstructured text, spelling errors, and grammatical inconsistencies pose additional difficulties in accurately capturing sentiment. Researchers have explored techniques such as incorporating contextual information,

leveraging sentiment lexicons, and utilizing ensemble models to address these challenges and improve the overall performance of sentiment analysis systems.

Overall, sentiment analysis of product reviews continues to be a vibrant and evolving research area. The advancements in machine learning algorithms, feature extraction methods, and the integration of sentiment analysis with recommendation systems have paved the way for more accurate and effective sentiment analysis models. As businesses increasingly rely on customer feedback for decision-making, sentiment analysis of product reviews will continue to play a pivotal role in understanding customer sentiments, improving products and services, and driving customer satisfaction.

1. Sentiment analysis in multilingual environments: While sentiment analysis techniques have primarily been developed and evaluated on English text, researchers have also explored sentiment analysis in multilingual settings. This involves adapting models to handle different languages, cultural nuances, and sentiment expressions specific to each language. The challenges in cross-lingual sentiment analysis include limited labeled data, translation discrepancies, and variations in sentiment analysis rules across languages.

2. Aspect-based sentiment analysis for fine-grained insights: Aspect-based sentiment analysis goes beyond overall sentiment classification and focuses on extracting sentiments associated with specific aspects or features of a product. This level of granularity provides businesses with valuable insights into strengths and weaknesses of different product attributes, allowing for targeted improvements and better decision-making.

3. Comparative sentiment analysis: Comparative sentiment analysis involves analyzing the sentiment expressed in product reviews by comparing multiple products or brands within the same review. This approach helps identify preferences, strengths, and weaknesses across different options, aiding businesses in understanding market dynamics, competitor analysis, and product positioning strategies.

# 3. Preliminaries

This section will discuss various supervised machine learning and deep learning methods that we have studied and utilized.

**i.**Logestic Regression:

- A linear classifier used for binary classification issues is logistic regression.
- The logistic function is used to model the relationship between the features and the likelihood of a specific result.
- The assumption behind logistic regression is that the characteristics and the log-odds of the outcome are linearly related.
- It is rather quick and easy to understand, and it can handle both categorical and numerical information.
- For binary classification tasks, the logistic regression baseline model is frequently utilised.

**ii.** Multinomial Naive Bayes (Multi NB):

- A popular probabilistic classifier for text classification tasks like sentiment analysis is Multinomial Naive Bayes.
- The Bayes theorem is applied under the presumption of feature independence.
- Particularly useful for discrete features are word frequencies or counts, multinomial naive bayes.
- Large feature spaces may be used with ease, and it is computationally effective.
- Multinomial Naive Bayes can frequently produce good performance in practice despite its simplistic assumptions.

**iii.** Support Vector Machine (SVC):

- Support A potent supervised learning technique used for classification and regression tasks is the vector machine.
- In the feature space, it locates a hyperplane that maximally divides the classes.
- High-dimensional feature spaces may be handled by SVC with ease, and it is resistant to overfitting.
- It creates a higher-dimensional space from the data using a kernel function, enabling non-linear decision bounds.
- Both linearly separable and non-linearly separable data can be handled using SVC. iv.K-nearest

neighbor (KNN):

- A non-parametric classification approach called K-nearest neighbour gives a data point a class label based on the majority class label of its k nearest neighbours.
- It utilises distance metrics, such as Euclidean distance, to determine how close together the data points are.
- KNN can handle both classification and regression jobs and is easy to build.
- As a result of not explicitly learning a model from the training data, it is regarded as an instance-based learning algorithm.
- The selection of k and the distance metric may have an impact on how well KNN performs.

V. Random Forest:

- An ensemble learning technique called Random Forest uses several decision trees to produce predictions.
- It creates a number of decision trees and averages or votes on the predictions made by each one.
- By injecting randomization throughout the tree-building process, for as by employing random subsets of features, Random Forest lessens overfitting.
- It is resilient to outliers and noisy data, and it can handle high-dimensional data.
- The performance, interpretability, and capacity of Random Forest to capture complicated relationships are well established.

VI. Convolutional Neural Network (CNN):

- A deep learning architecture frequently utilised for text and picture classification applications is the convolutional neural network.
- Convolutional layers are used, and these layers automatically identify hierarchical characteristics from the input data.
- CNNs are good at identifying temporal or spatial patterns in the data.
- CNNs can be used to learn local features from word embeddings or character-level representations for text classification.
- Sentiment analysis is one natural language processing task where CNNs have demonstrated good performance.

VII. Long Short-Term Memory (LSTM):

- A recurrent neural network (RNN) called Long Short-Term Memory is made to handle sequence data, like text or time series.
- A memory cell and several gating mechanisms are used by LSTM to learn long-term dependencies.
- It fixes the issue with traditional RNNs' vanishing gradients, making it possible to effectively model sequential data.
- Contextual data can be captured and the semantic structure can be understood using LSTM.

# 4.Methodology

## Experimental setup:

In this, four datasets are used to build all the machine learning models like Multi NB,SVC,KNeighbors Classifier,Logestic Regression,Random Forest and other deep learning models like cnn and lstm.Here all the four dataset are taken from kaggle. The first dataset consists of 10,000 reviews from Amazon customers for each of the following categories: video games, books, kitchen appliances, and electronics. The dataset's goal is to offer a text classification job in which each review's sentiment is categorised as either positive or negative.Each review is classified as either favourable or unfavourable. The reviews are evenly distributed, with 50% of them being classified as favourable and the other 50% as unfavourable.

A collection of laptop and dining establishment reviews from the SemEval 2014 competition makes up the second dataset. The dataset is intended for aspect-based sentiment analysis, whose objective is to determine the sentiment held towards different features of a good or service. Reviews from numerous websites, including Yelp and TripAdvisor, are included in the dataset. 3,118 reviews of laptops and 3,276 reviews of restaurants are included in the dataset. The sentiment towards each feature of the good or service, such as the cost or the standard of the cuisine, is noted in the reviews. Each aspect is assigned one of three labels: good, negative, or neutral.

A collection of reviews for women's clothes from an online retailer makes up the third dataset. 23,486 reviews of women's apparel products are included in the dataset. With each review classified as good, negative, or neutral, the dataset's goal is to perform text classification for sentiment analysis. The labels on the reviews reflect the general opinion that customers have about the product. 82% of the samples are deemed positive, 10% are deemed negative, and 8% are deemed neutral.

## Methodology Used:

Our work involved various steps such as data collection, data prepossessing feature selection,text Tokenization,Feature Extraction, model building,Sentiment Classification and performance evaluation.

The implementation of supervised machine learning algorithms (Logistic Regression, Support Vector Machine (SVM), K-nearest neighbor (KNN), Naive Bayes, Decision Tree, Random Forest) for sentiment analysis of product reviews involved utilizing the Python programming language and the scikit-learn library. Supervised learning algorithms are preferred over unsupervised algorithms because they have the ability to learn from labeled training data, where feedback in the form of correct labels is provided during the training process. This enables the algorithms to make predictions based on the learned patterns and relationships from the labeled data.

**Data Collection:**

The first step in conducting sentiment analysis of product reviews is to gather a comprehensive dataset. This dataset should comprise a diverse range of product reviews, covering various domains, products, and sentiments. Data can be collected from e-commerce websites, social media platforms, or specialised review websites. Ensuring a balanced representation of positive, negative, and neutral reviews is essential for training accurate sentiment classification model.

**Data Preprossessing:**

One of the crucial stages in the life cycle of machine learning is data preprocessing, which facilitates data analysis and improves the efficiency and speed of ML algorithms.Clean and preprocess the collected data to prepare it for analysis. This step often includes removing irrelevant information, such as HTML tags or special characters, and handling noise, such as spelling errors or abbreviations.

**Feature Selection:**

Attributes of a dataset are a property that is utilized by the system. For the ML model analysis purposes, various attributes like rate,id,age,polarity.Literature review reveals that the features which are highly correlated with the target values are more likely to be important in predicting the outcome. Therefore, based on the literature review, we have finalized 14 features that are highly correlated with the target value and are selected for building all the 6 supervised machine learning models(Logistic Regression, Support vector machine(SVM) K-nearest neighbor(KNN), Naive Bayes, Decision Tree, Random Forest).

Here we had Found and extract from the reviews the pertinent features. This can be accomplished using simple methods like bag-of-words, where each review is represented by

a vector of word frequencies, or more sophisticated strategies like word embeddings, which identify the semantic meaning of words.

**Model building:**

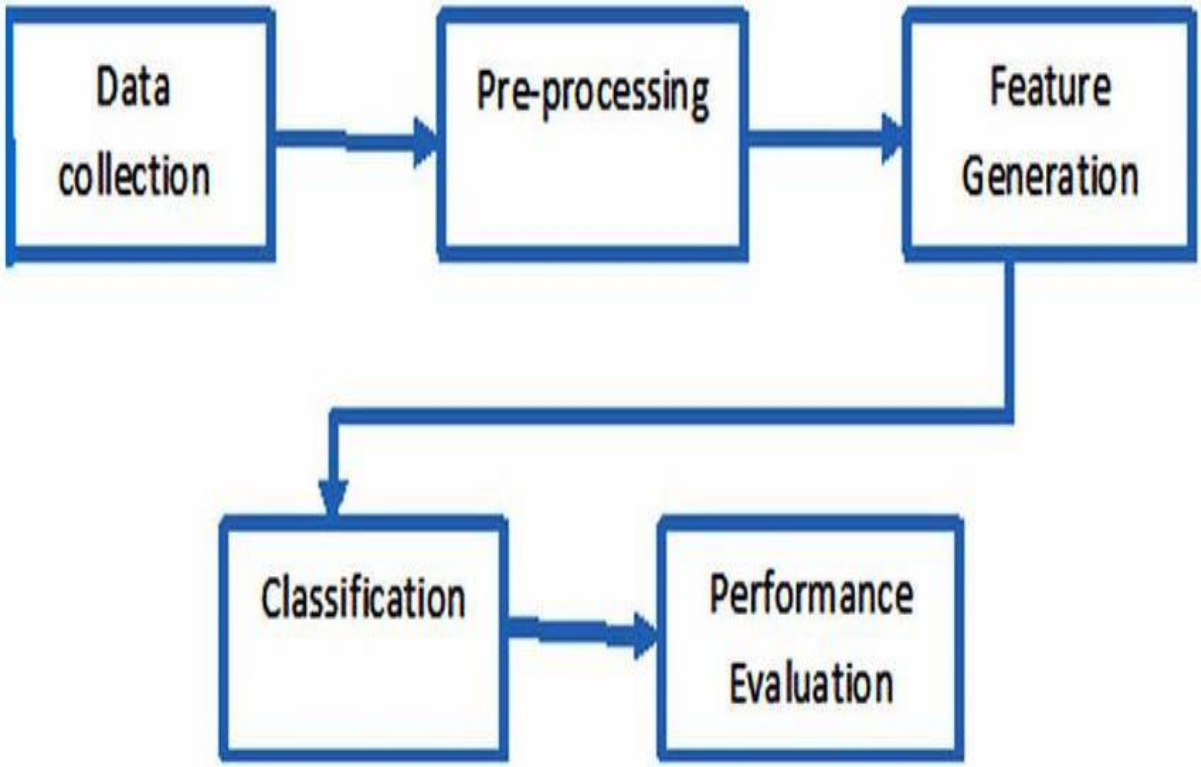
Here we have chosen the appropriate algorithms and the model learns the underlying patterns and relationships between the textual features and the corresponding sentiment labels. Fine-tune the model by adjusting its hyperparameters, such as learning rate, regularization parameters, or maximum tree depth. This process helps optimize the model's performance by finding the best combination of hyperparameters.And thn we Trained the selected model using the labeled training data. The model learns the underlying patterns and relationships between the textual features and the corresponding sentiment labels.

**Sentiment Classification:**

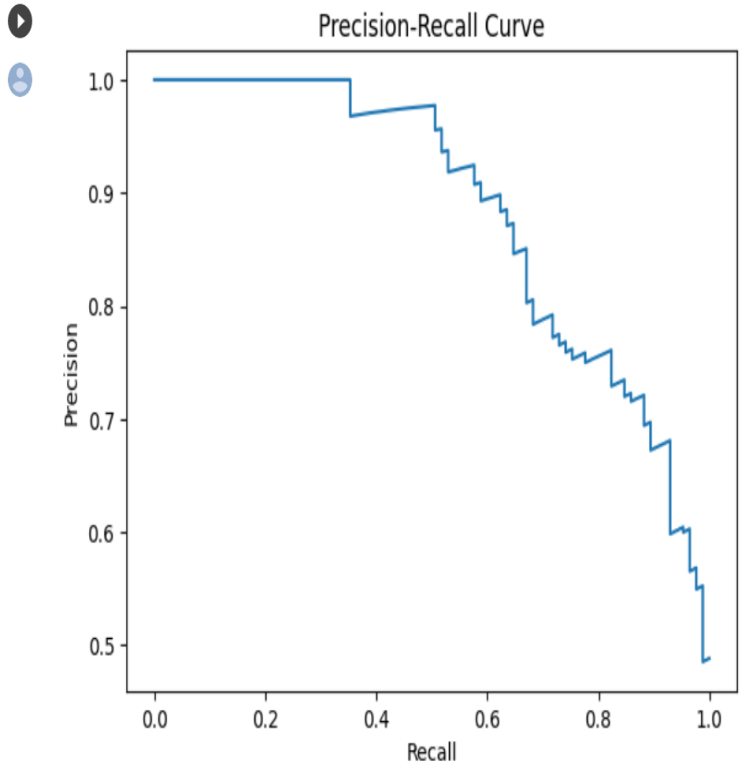
To categorise the reviews into categories of good, negative, or neutral sentiment, train a machine learning or deep learning model. In this step, labelled training data are often used to create a predictive model that can effectively generalise to unreviewed reviews.

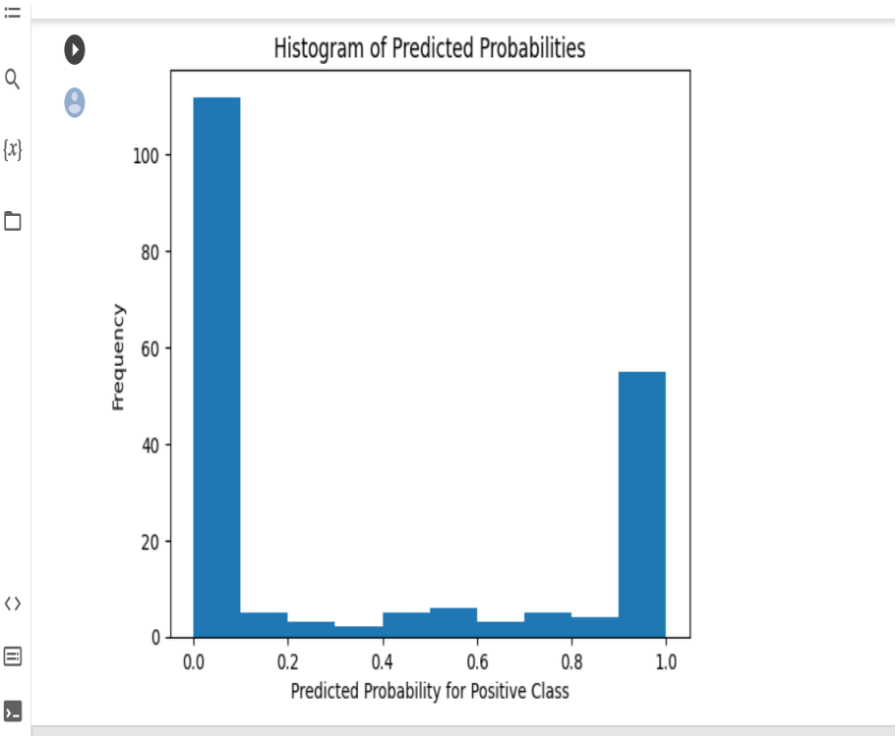
**Performance Evaluation:**

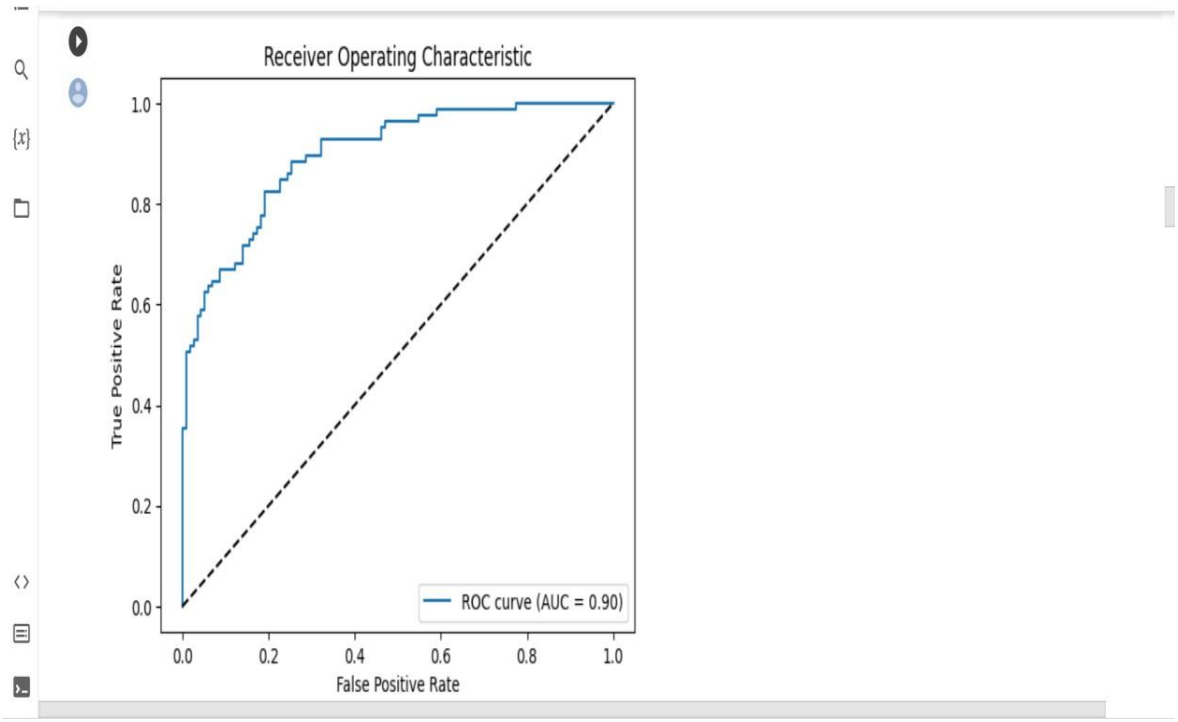
Later, we evaluated the sentiment classification model's performance using metrics like accuracy, precision, recall, or F1-score. This step assists in assessing the model's performance and determining whether any modifications or enhancements are required.

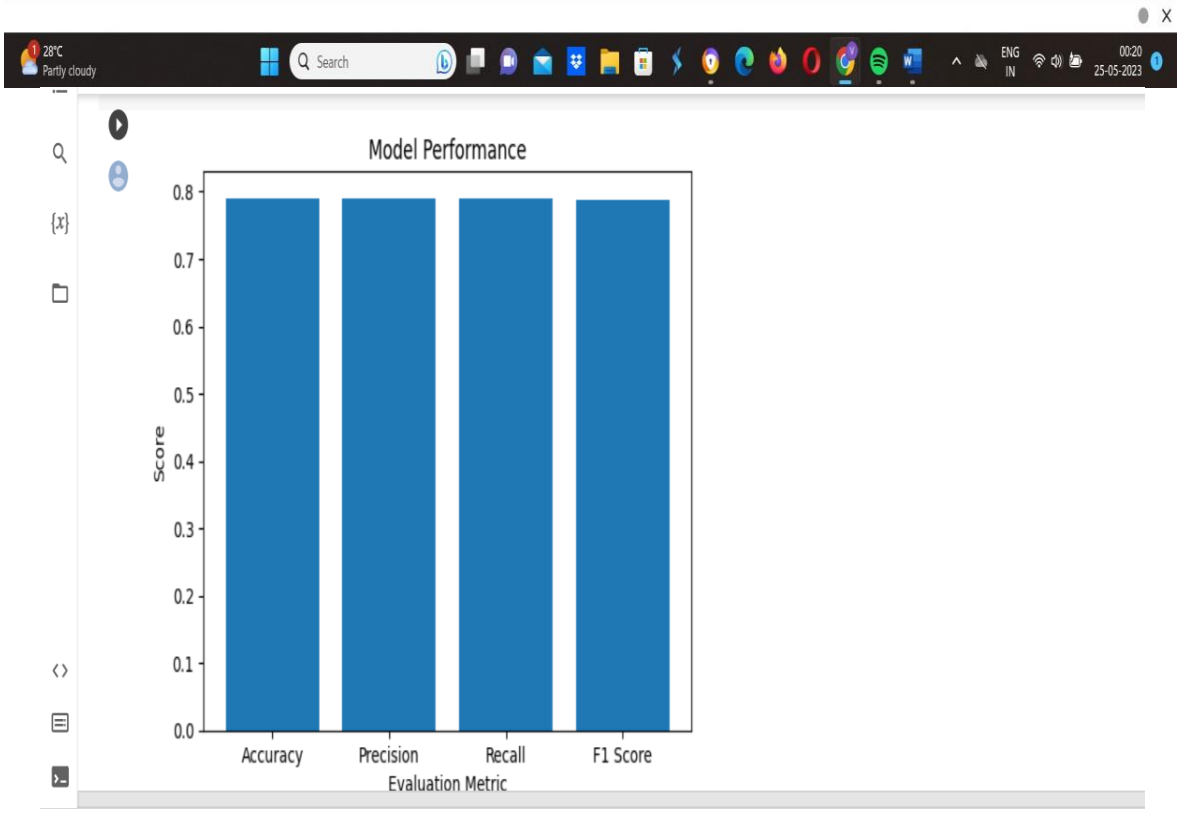

Data collection
Pre-processing
Feature Generation
Classification
Performance Evaluation

# DATA SET-1

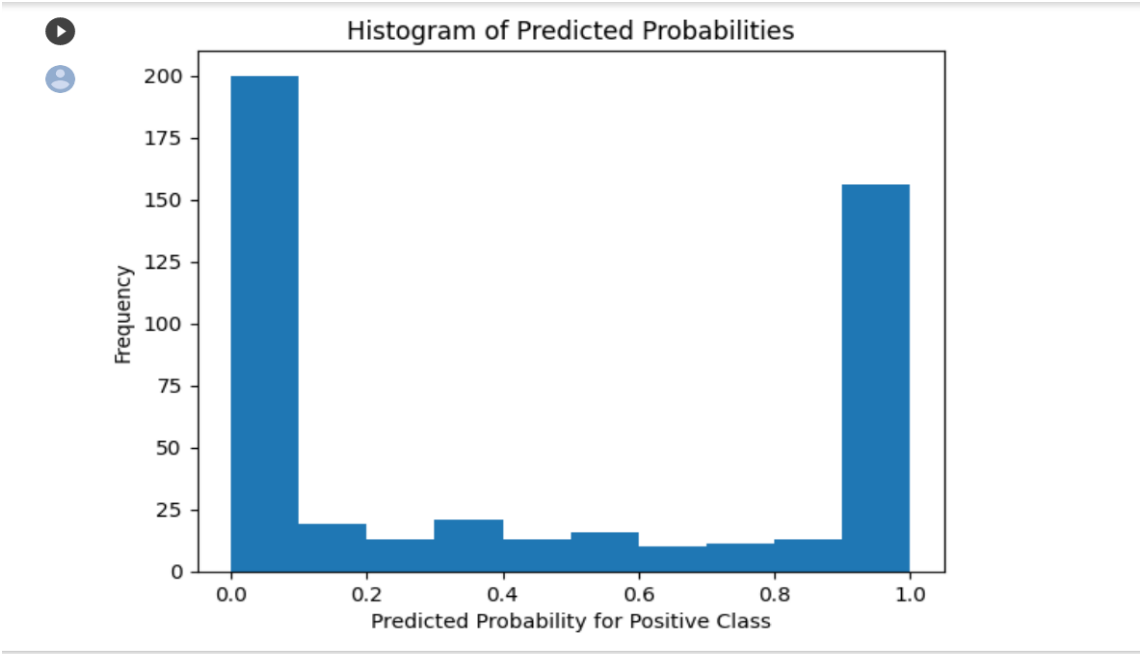

Precision-Recall Curve
Precision
Recall
1.0
0.9
0.8
0.7
0.6
0.5
0.0
0.2
0.4
0.6
0.8
1.0

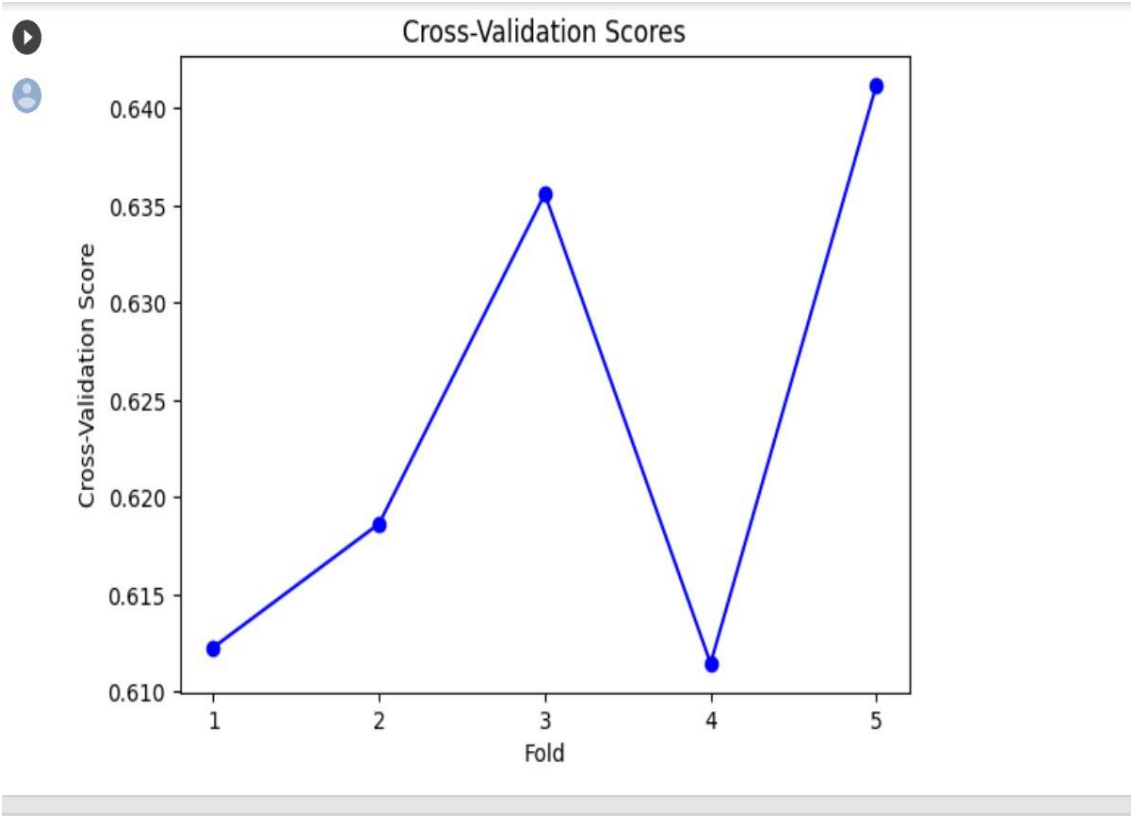

Histogram of Predicted Probabilities
Frequency
Predicted Probability for Positive Class
100
80
60
40
20
0
0.0
0.2
0.4
0.6
0.8
1.0

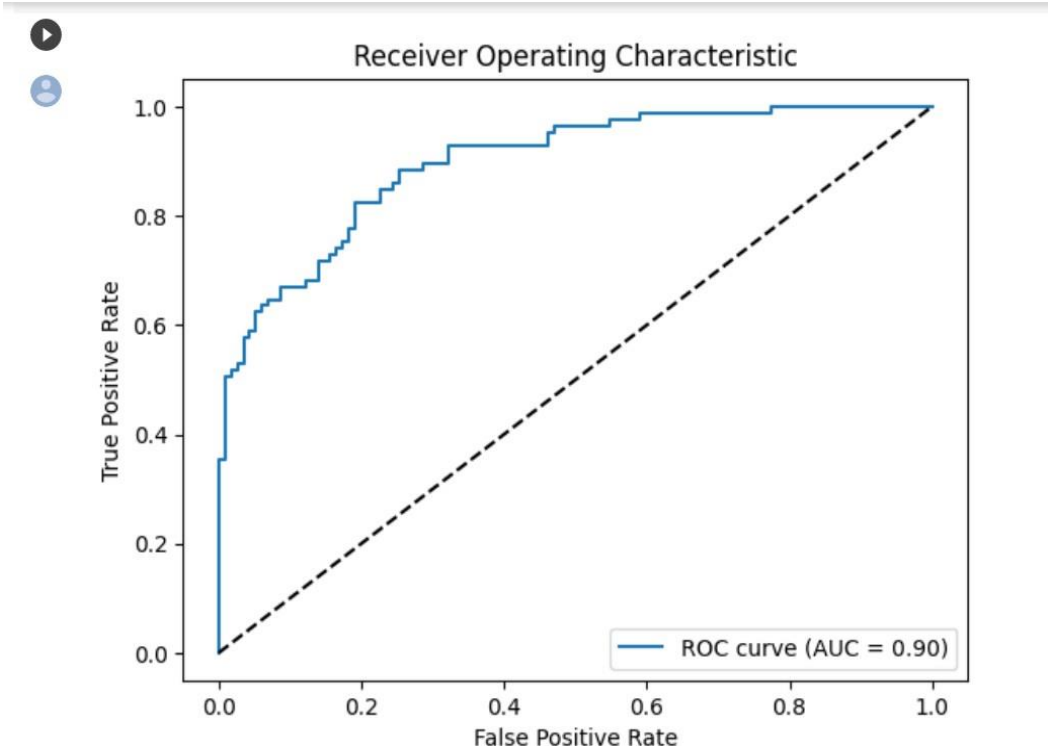

Receiver Operating Characteristic
True Positive Rate
False Positive Rate
ROC curve (AUC = 0.90)
0.0
0.2
0.4
0.6
0.8
1.0

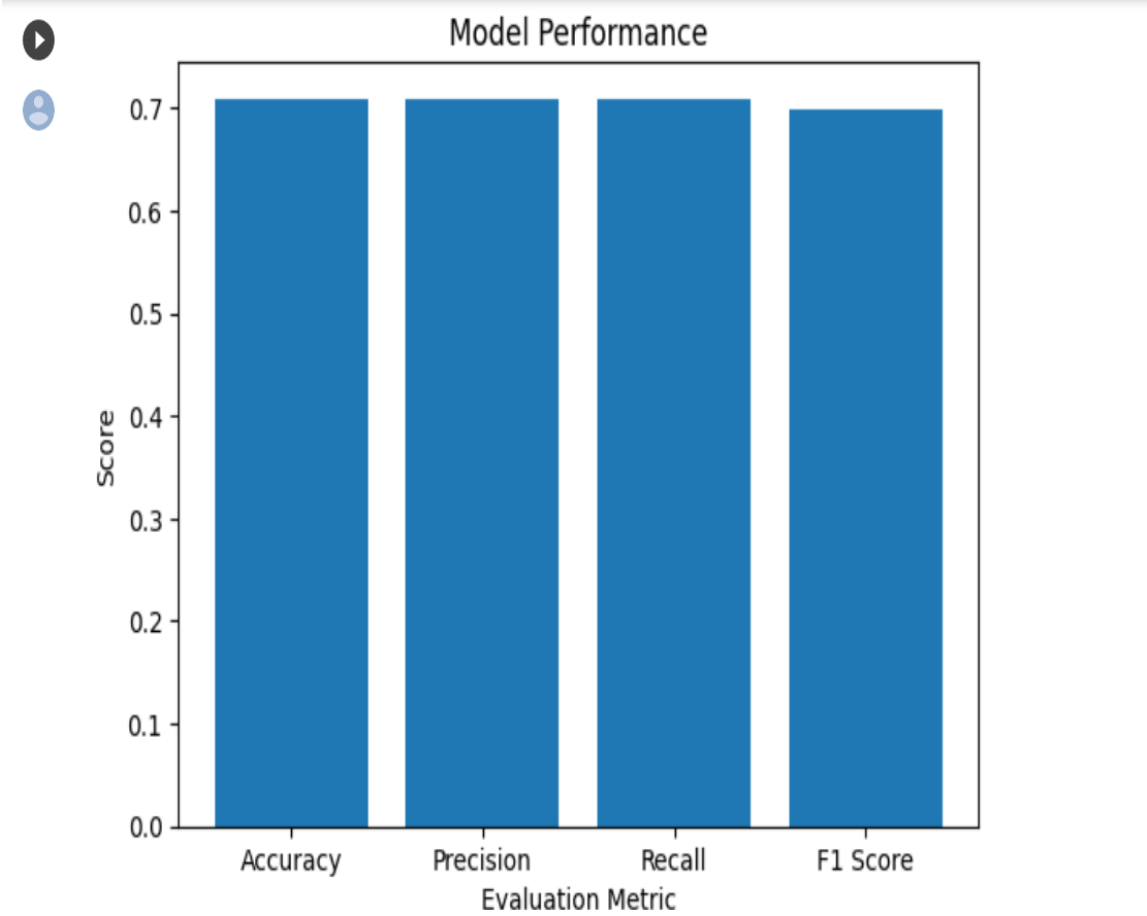

Model Performance
Score
Accuracy
Precision
Recall
F1 Score
Evaluation Metric
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

# Data set -2

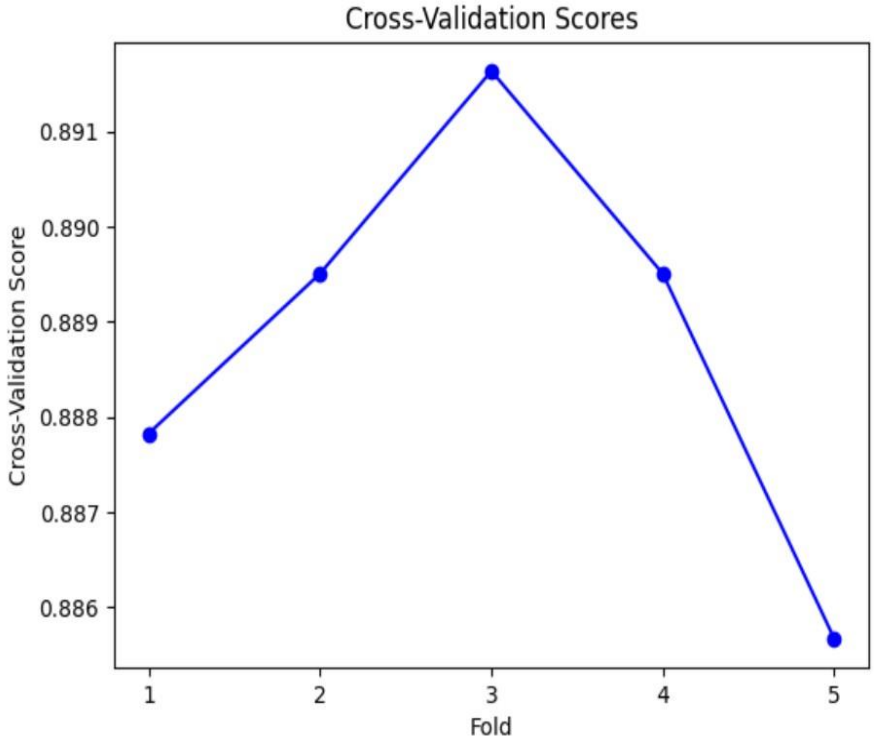

Histogram of Predicted Probabilities
Frequency
200
175
150
125
100
75
50
25
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted Probability for Positive Class

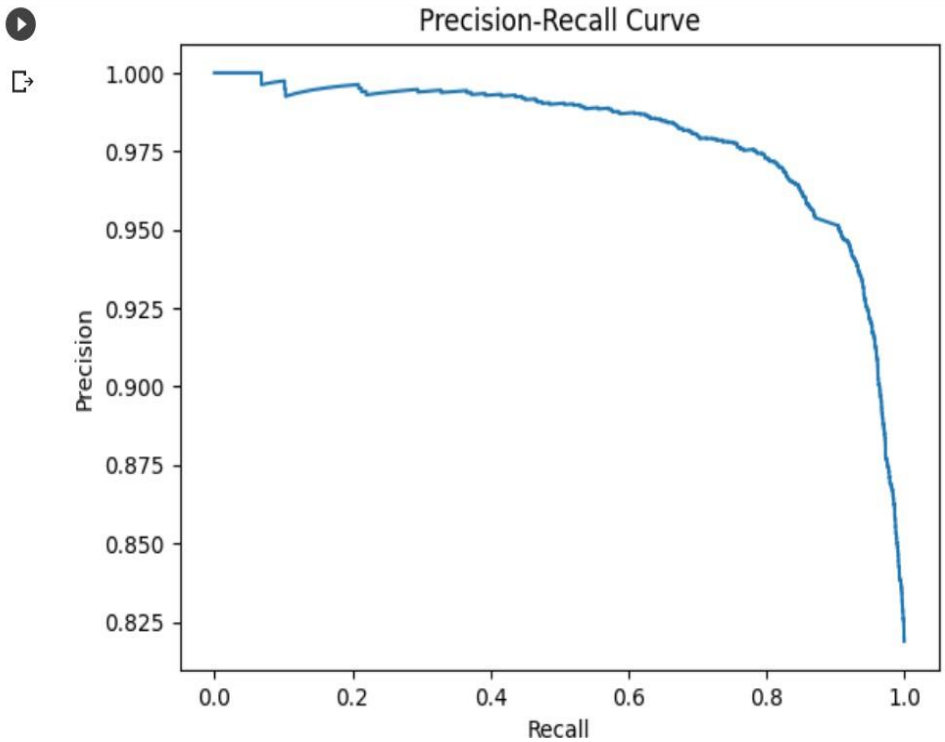

Cross-Validation Scores
Cross-Validation Score
0.640
0.635
0.630
0.625
0.620
0.615
0.610
1
2
3
4
5
Fold

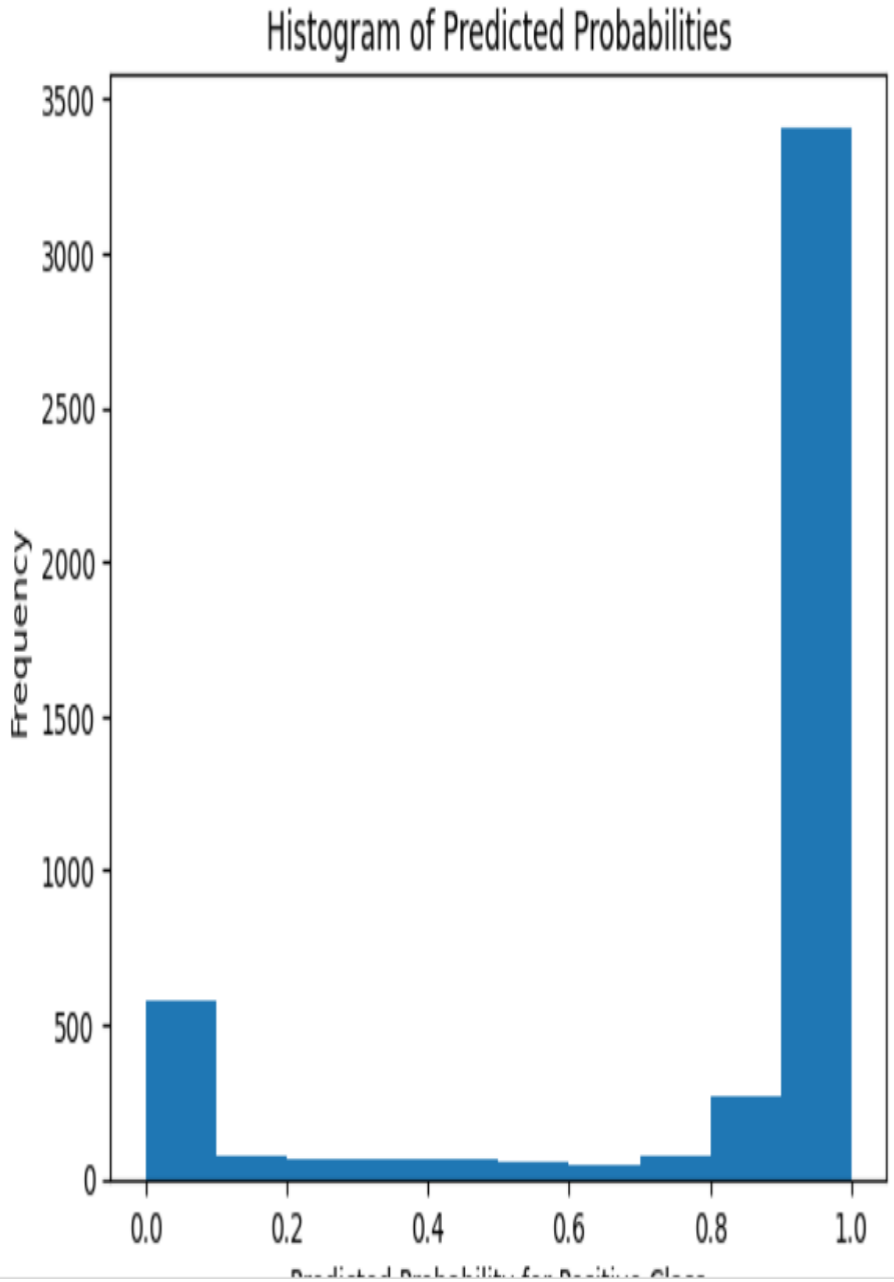

Receiver Operating Characteristic
True Positive Rate
False Positive Rate
ROC curve (AUC = 0.90)
0.0
0.2
0.4
0.6
0.8
1.0

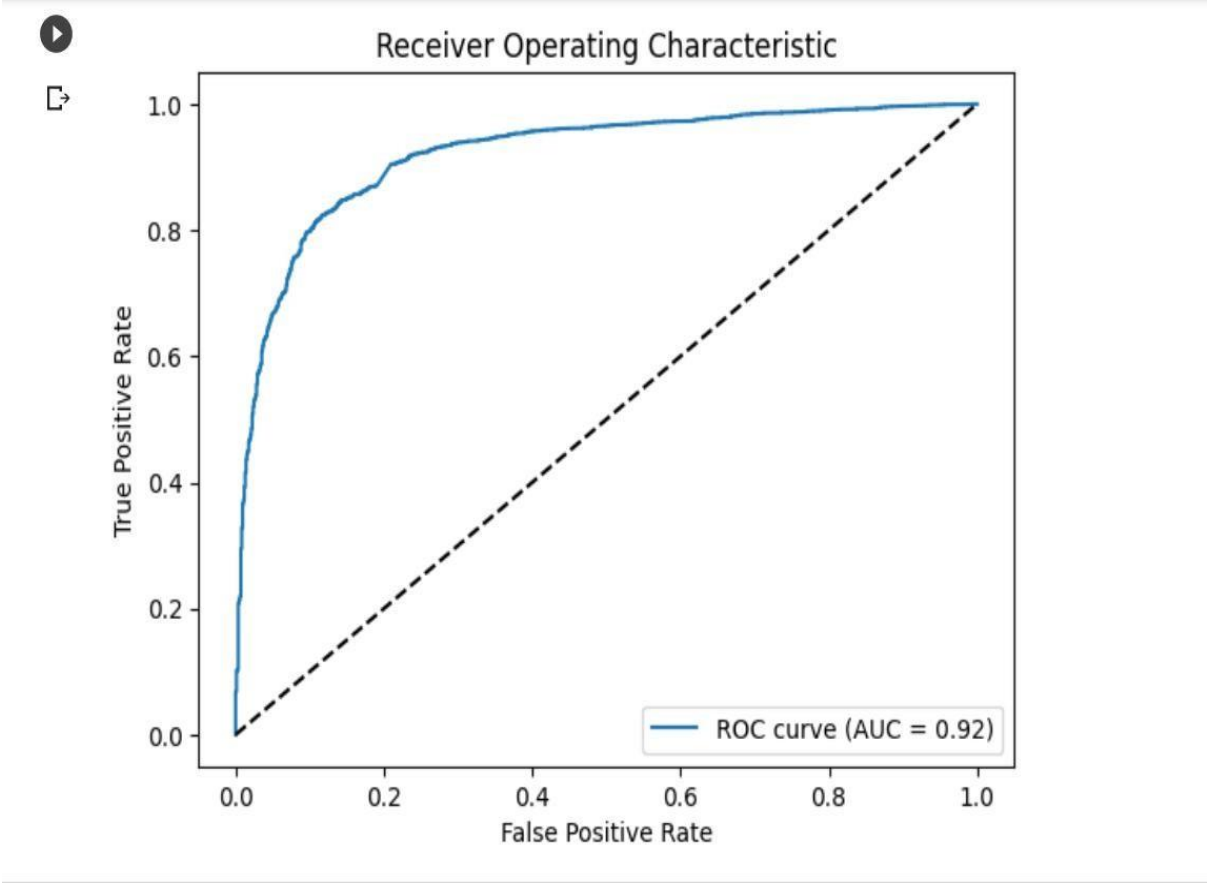

Model Performance
Score
Evaluation Metric
Accuracy
Precision
Recall
F1 Score
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

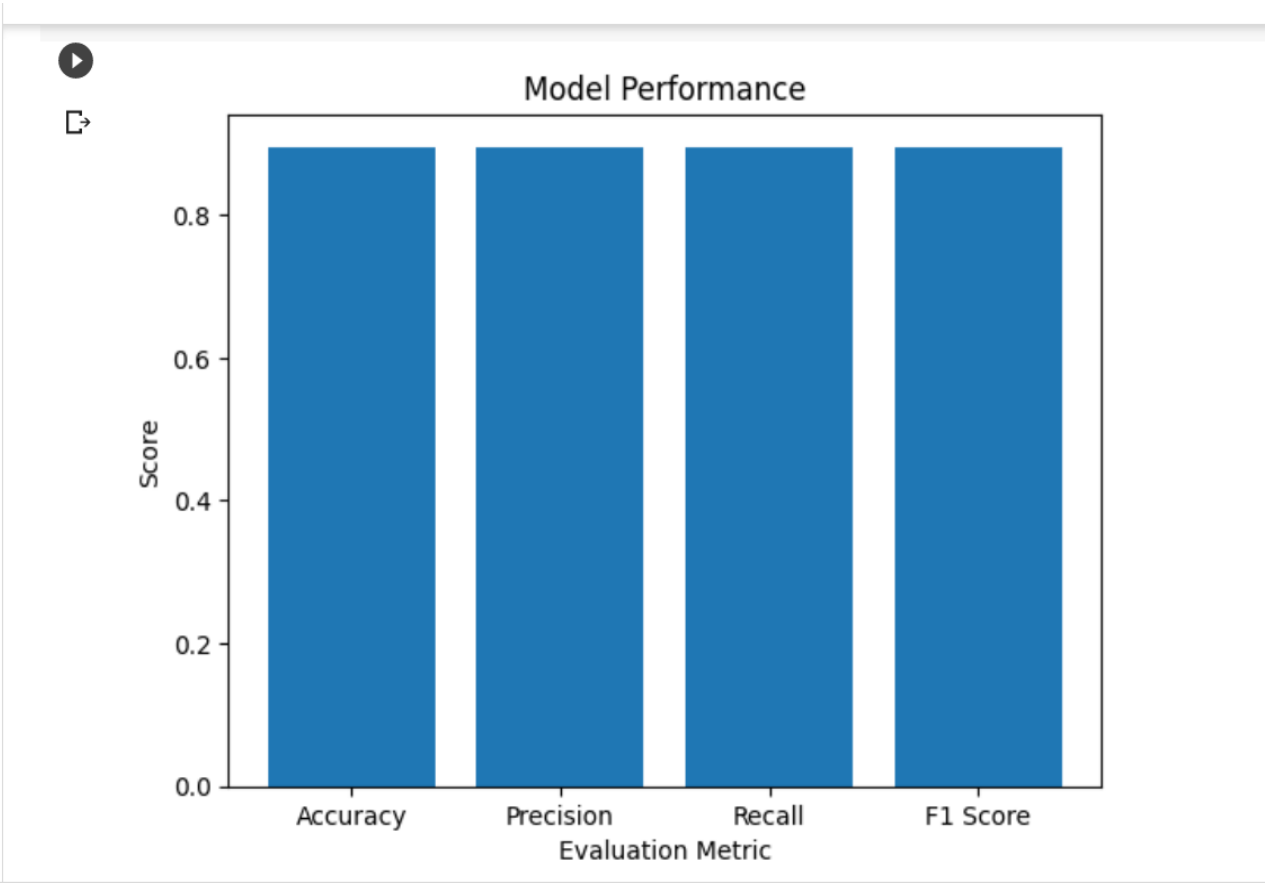
## Data set -3


Cross-Validation Scores
Cross-Validation Score
0.891
0.890
0.889
0.888
0.887
0.886
1
2
3
4
5
Fold

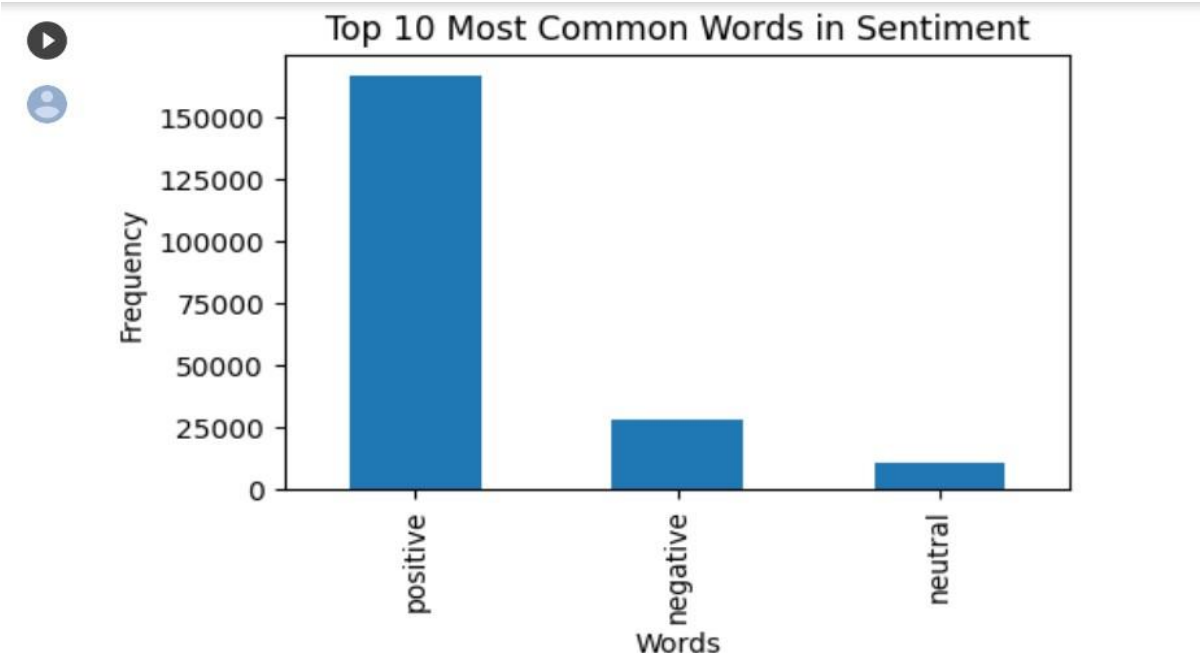

Precision-Recall Curve
Precision
1.000
0.975
0.950
0.925
0.900
0.875
0.850
0.825
0.0
0.2
0.4
0.6
0.8
1.0
Recall

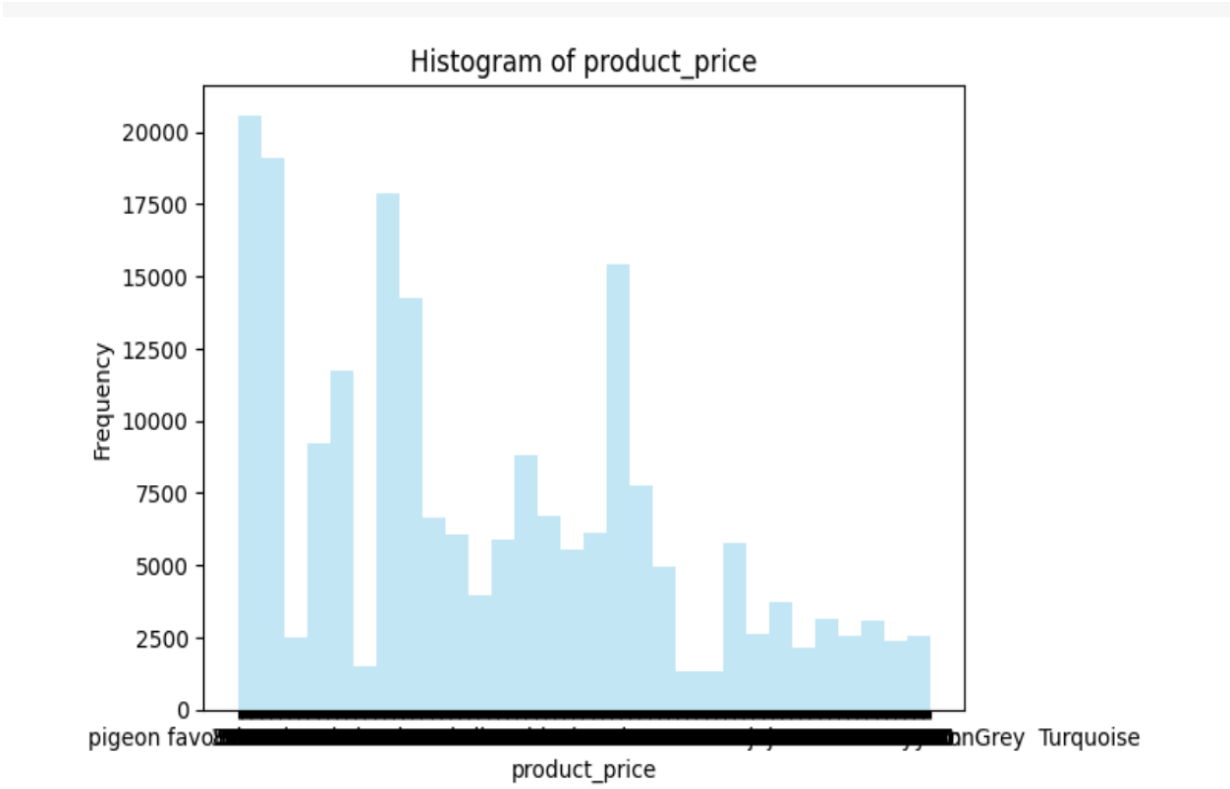

Histogram of Predicted Probabilities
Frequency
3500
3000
2500
2000
1500
1000
500
0
0.0
0.2
0.4
0.6
0.8
1.0

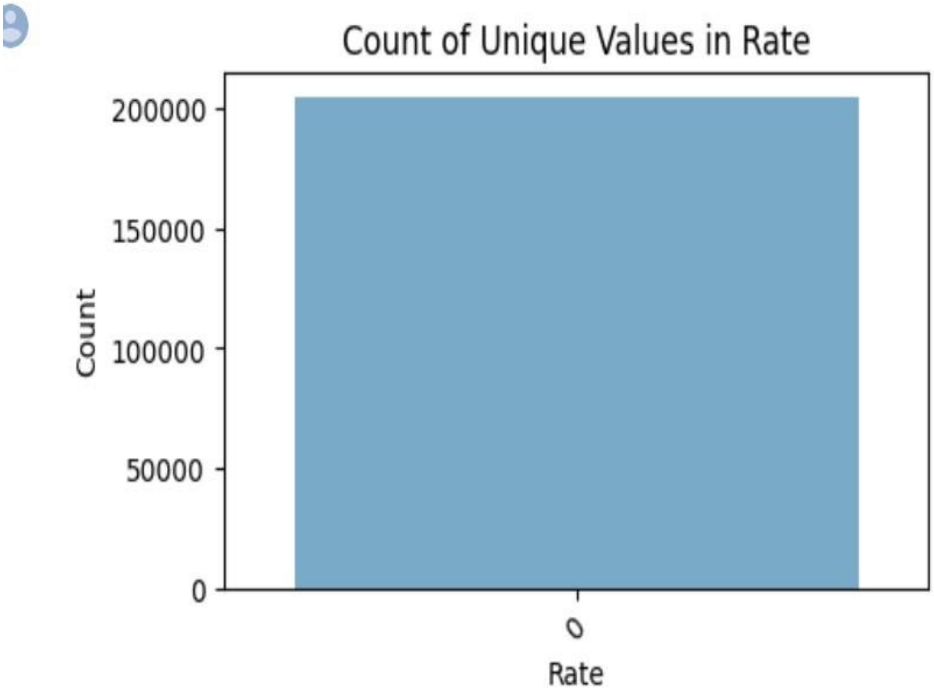

Receiver Operating Characteristic
True Positive Rate
False Positive Rate
ROC curve (AUC = 0.92)
0.0
0.2
0.4
0.6
0.8
1.0

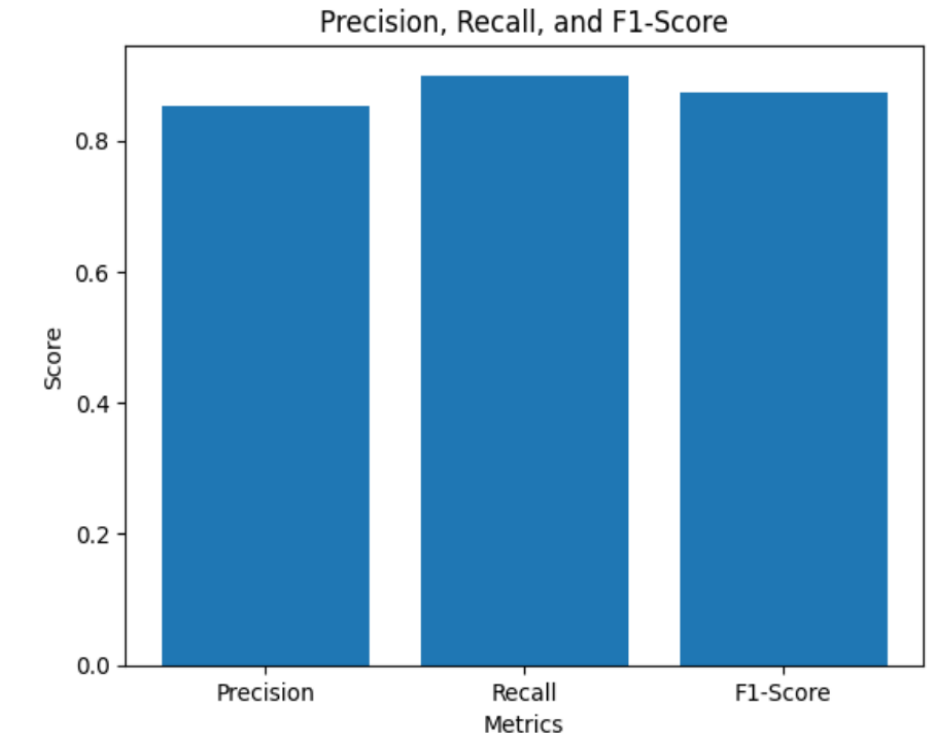

Model Performance
Score
Evaluation Metric
Accuracy
Precision
Recall
F1 Score
0.0
0.2
0.4
0.6
0.8

## Data set-4


Top 10 Most Common Words in Sentiment
Frequency
150000
125000
100000
75000
50000
25000
0
positive
negative
neutral
Words


Histogram of product_price
Frequency
20000
17500
15000
12500
10000
7500
5000
2500
0
pigeon favo
Turquoise
product_price


Count of Unique Values in Rate
200000
150000
100000
50000
0
Count
0
Rate


Precision, Recall, and F1-Score
0.8
0.6
0.4
0.2
0.0
Score
Precision
Recall
F1-Score
Metrics

# Result and Discussion

In this section, the result of the work is discussed. Accuracy, Precision, Recall, F1-Score, Misclassification Rate, and Specificity are used as performance metrics to evaluate the performance of the ML models. Accuracy, Precision, Recall, F1-Score, Miss classification Rate, and Specificity is defined using Eq. 1, Eq. 2, Eq. 3, Eq. 4, Eq. 5, and Eq. 6 respectively.

$$Accuracy = \frac{TruePositive + True\ Negative}{Total\ Number\ of\ Predictions} \quad (1)$$

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (2)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (3)$$

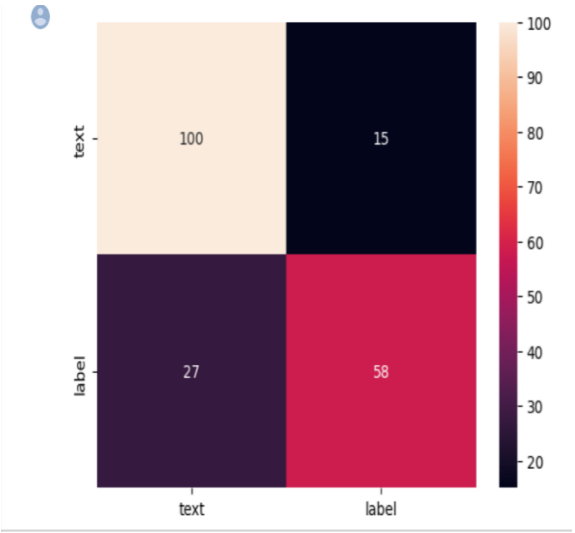
$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

$$Misclassification\ Rate = 1 - Accuracy \quad (5)$$

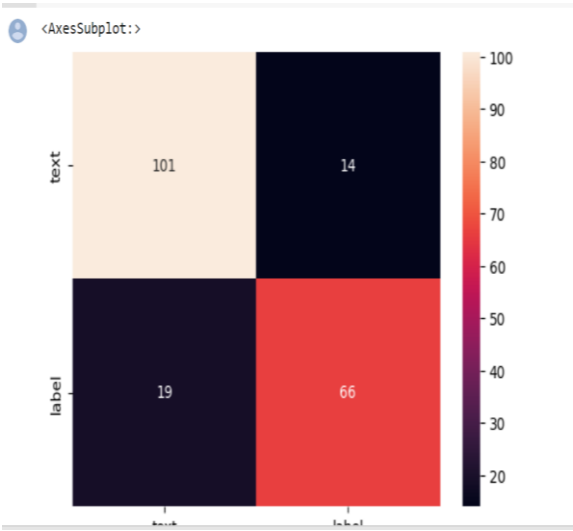
$$Specificity = \frac{TrueNegative}{TrueNegative + FalsePositive} \quad (6)$$

## DATASET-1

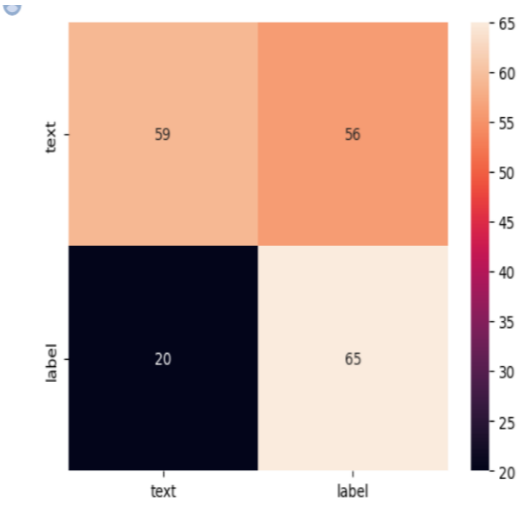
### i. Multi Naive bayes



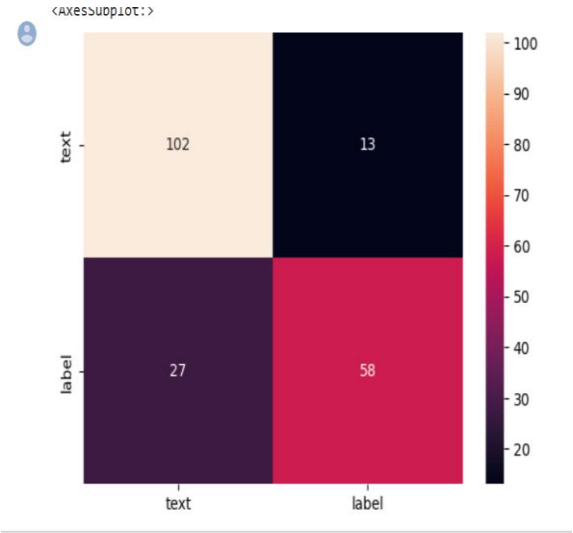

### ii. Logestic Regression



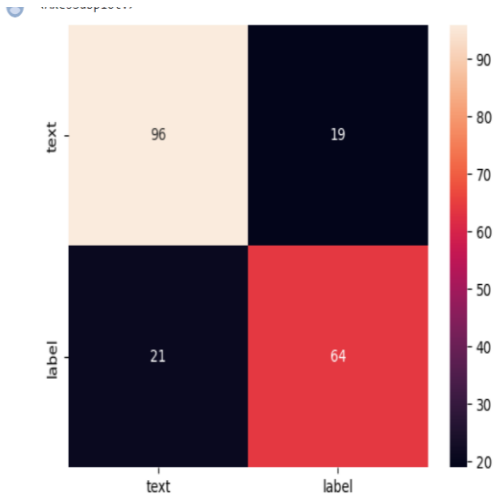

### iii. Random Forest



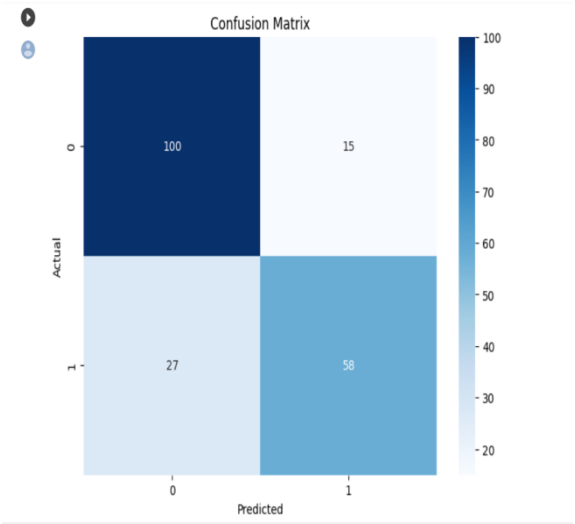

### iv. K-Neighbor



**V.SVC**

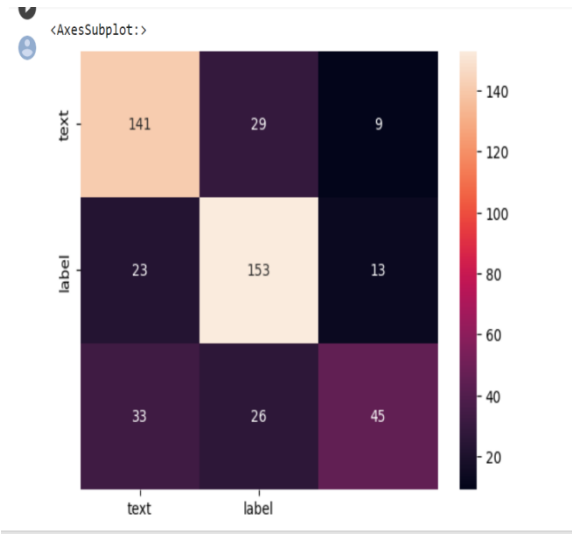

96
19
21
64
text
label
text
label
90
80
70
60
50
40
30
20

**Vi.CNN**

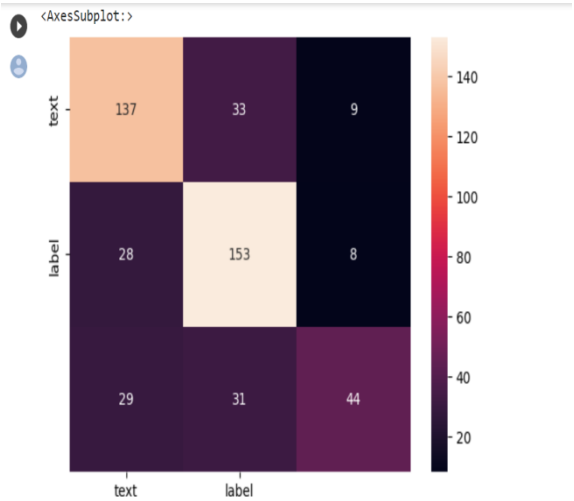

Confusion Matrix
100
15
27
58
Actual
0
1
Predicted
0
1
100
90
80
70
60
50
40
30
20

## DATASET-2

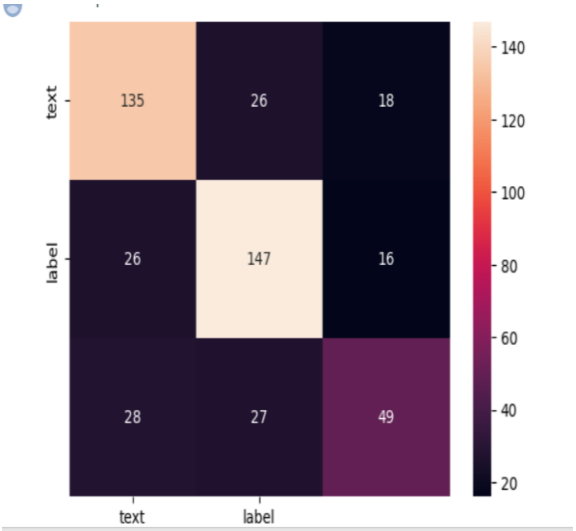
### i. Random Forest



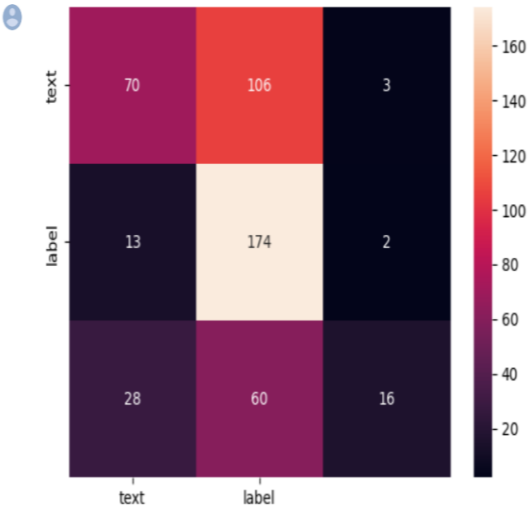

### ii. Multi Naive Bias



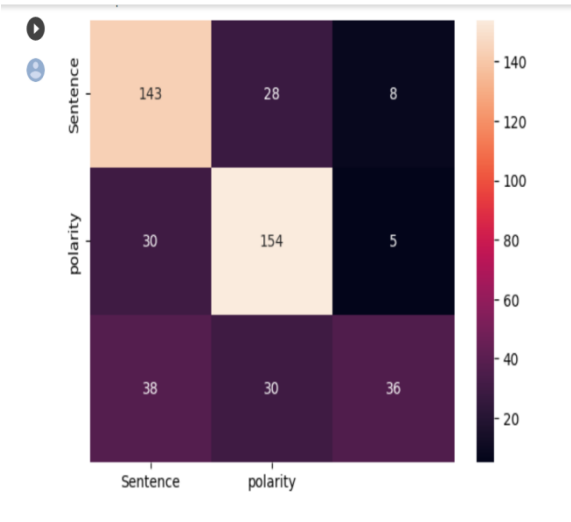

### iii. Logistic Regression



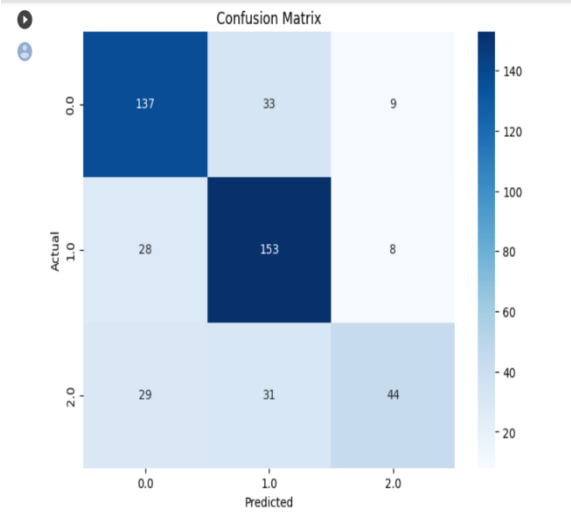

### iv.K-Neighbor



## V. SVC

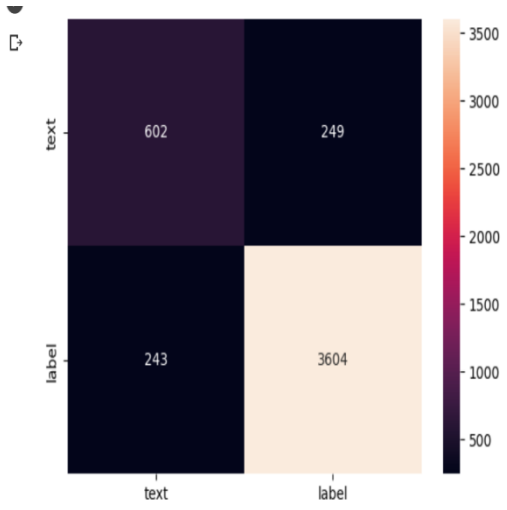

Sentence
polarity
143
28
8
30
154
5
38
30
36
Sentence
polarity
140
120
100
80
60
40
20

## Vi.CNN

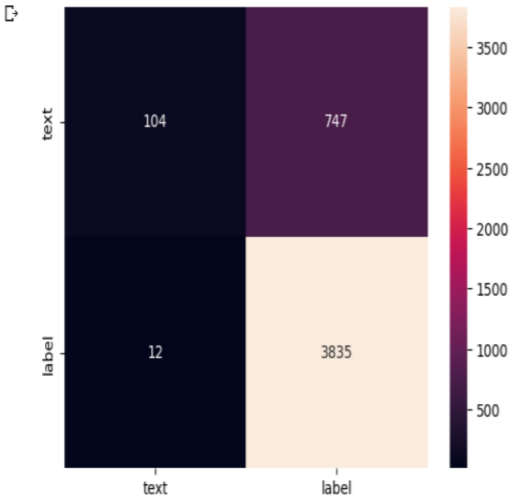

Confusion Matrix
Actual
0.0
1.0
2.0
137
33
9
28
153
8
29
31
44
0.0
1.0
2.0
Predicted
140
120
100
80
60
40
20

## DATASET-3

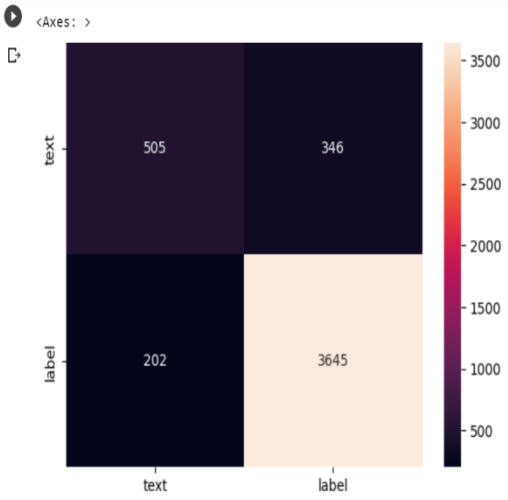
### i. Multinomial NB




### ii. Random Forest

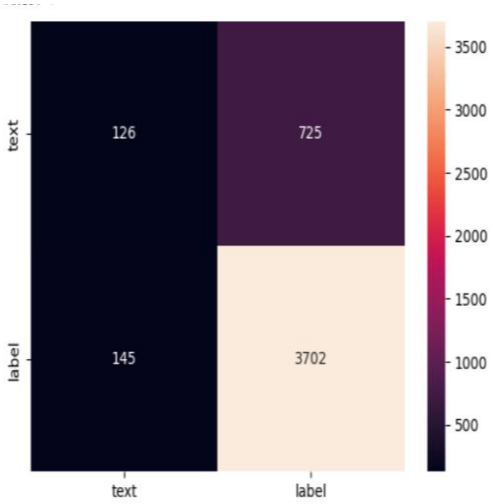



### iii. Logistic Regression

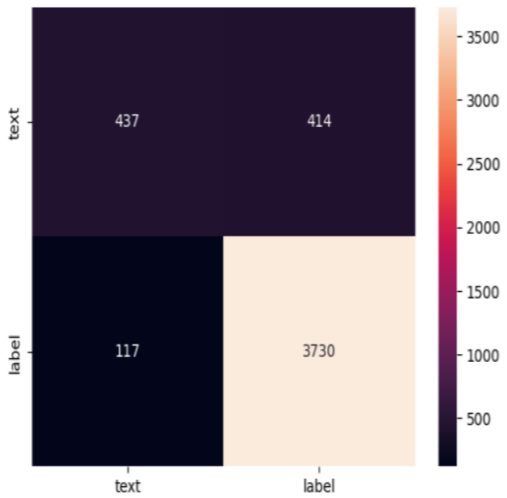




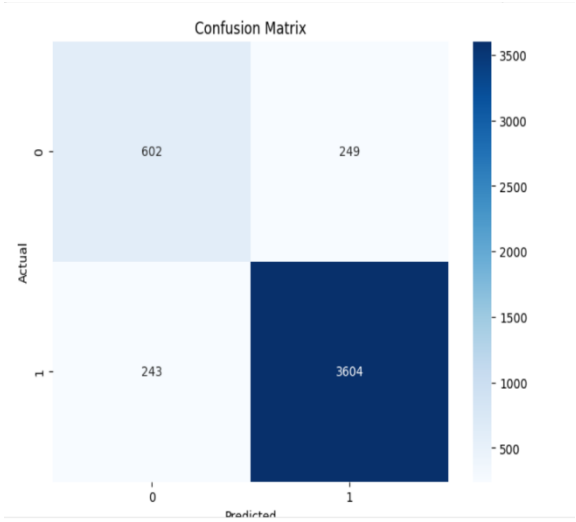

### iv. K Neighbour



**v.SVC**

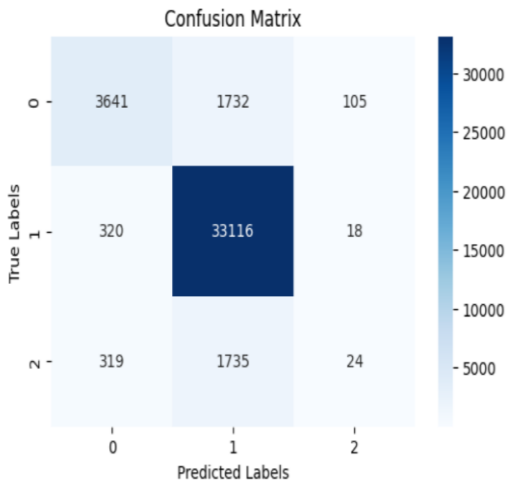

437
414
117
3730
text
label
text
label
3500
3000
2500
2000
1500
1000
500

**vi.CNN**

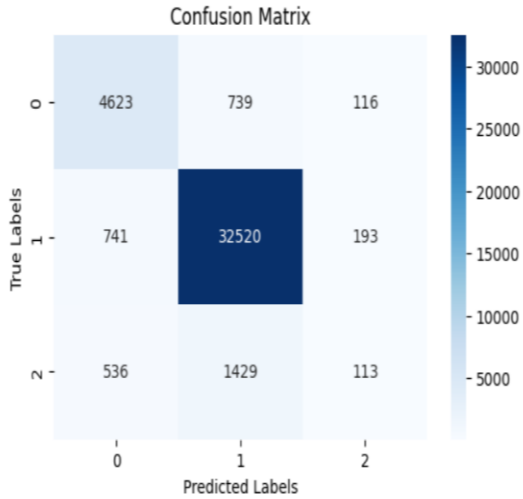

Confusion Matrix
602
249
243
3604
Actual
0
1
0
1
3500
3000
2500
2000
1500
1000
500

## DATASET-4

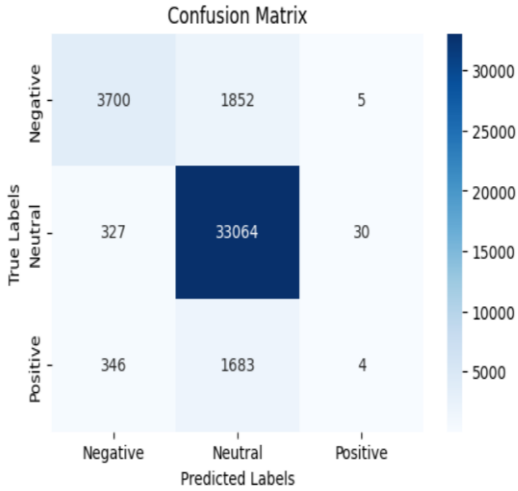
### i. K-Neighbor



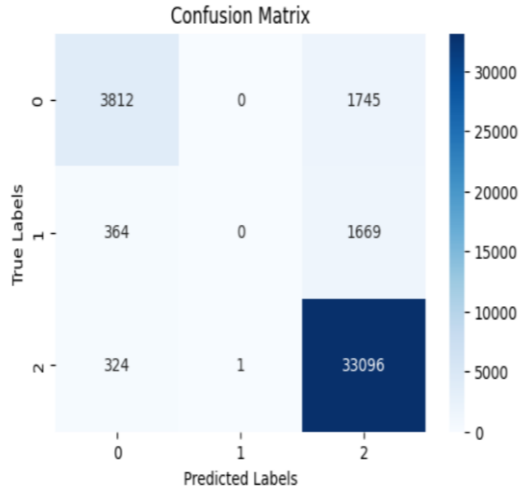

### ii. CNN



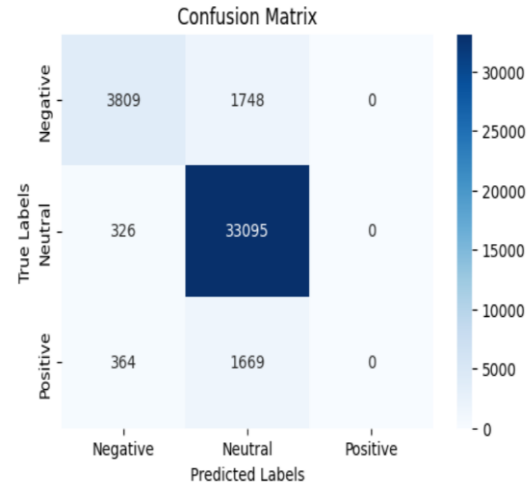

### iii. Multi NB



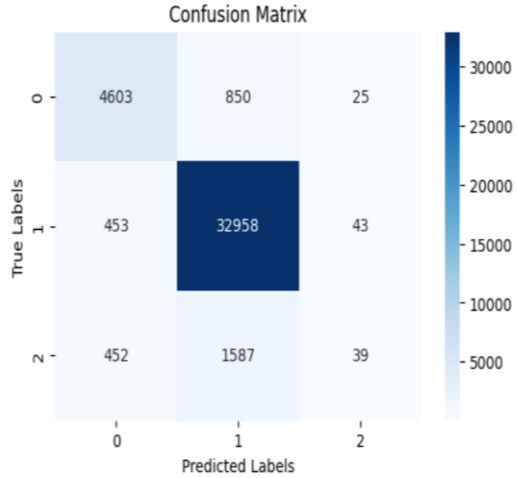

### iV SVC



## v. Logistic Regression




## vi.Random Forest



# Result and Discussion

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | | DATA SET -1 | | | | |
| 2 | | precision | Recall | F1 score | Accuracy | | |
| 3 | Logistic Reg | 0.825 | 0.77 | 0.8 | 0.835 | | |
| 4 | Multi NB | 0.794 | 0.68 | 0.73 | 0.79 | | |
| 5 | SVC | 0.77 | 0.75 | 0.761 | 0.8 | | |
| 6 | KNN | 0.53 | 0.76 | 0.63 | 0.62 | | |
| 7 | Random Forest | 0.775 | 0.812 | 0.611 | 0.697 | | |
| 8 | CNN | 0.79 | 0.68 | 0.73 | 0.84 | | |
| 9 | LSTM | 0.696 | 0.729 | 0.71 | 0.75 | | |
| 10 | | | | | | | |
| 11 | | | | | | | |
| 12 | | | DATA SET -2 | | | | |
| 13 | | precision | Recall | F1 score | Accuracy | | |
| 14 | Logistic Reg | 0.695 | 0.701 | 0.696 | 0.701 | | |
| 15 | Multi NB | 0.709 | 0.707 | 0.697 | 0.707 | | |
| 16 | SVC | 0.709 | 0.7055 | 0.689 | 0.7055 | | |
| 17 | KNN | 0.611 | 0.55 | 0.502 | 0.55 | | |
| 18 | Random Forest | 0.713 | 0.718 | 0.709 | 0.718 | | |
| 19 | CNN | 0.538 | 0.51 | 0.497 | 0.514 | | |
| 20 | LSTM | 0.469 | 0.5656 | 0.495 | 0.5656 | | |
| 21 | | | | | | | |
| 22 | | | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 22 | | | | | |
| 23 | | | | | |
| 24 | | | DATA SET - 3 | | |
| 25 | | precision | Recall | F1 score | Accuracy |
| 26 | Logistic Reg | 0.91 | 0.947 | 0.93 | 0.883 |
| 27 | Multi NB | 0.935 | 0.936 | 0.936 | 0.895 |
| 28 | SVC | 0.9 | 0.96 | 0.933 | 0.886 |
| 29 | KNN | 0.83 | 0.96 | 0.89 | 0.81 |
| 30 | Random Forest | 0.835 | 0.996 | 0.908 | 0.836 |
| 31 | CNN | 0.836 | 0.996 | 0.909 | 0.89 |
| 32 | LSTM | 0.908 | 0.951 | 0.929 | 0.882 |
| 33 | | | | | |
| 34 | | | | | |
| 35 | | | Data Set-4 | | |
| 36 | | precision | Recall | F1 score | Accuracy |
| 37 | Logistic Reg | 0.853 | 0.899 | 0.87 | 0.899 |
| 38 | Multi NB | 0.86 | 0.9 | 0.87 | 0.896 |
| 39 | SVC | 0.85 | 0.9 | 0.87 | 0.9 |
| 40 | KNN | 0.88 | 0.91 | 0.89 | 0.908 |
| 41 | Random Forest | 0.89 | 0.92 | 0.89 | 0.916 |
| 42 | CNN | 0.79 | 0.82 | 0.81 | 0.815 |
| 43 | LSTM | 0.856 | 0.87 | 0.85 | 0.873 |
| 44 | | | | | |
| 45 | | | | | |
| 46 | | | | | |
| 47 | | | | | |
| 48 | | | | | |
| 49 | | | | | |
| 50 | | | | | |

For the dataset -1 The CNN model, which has an accuracy of 0.84 based on the information presented, has the highest accuracy. The Logistic Regression model, which has an accuracy of 0.835, is the most accurate classical machine learning model, followed by the Support Vector Classifier (SVC) model, which has an accuracy of 0.8.On the other hand, the Random Forest model achieves an accuracy of 0.697, while the K-Nearest Neighbour (KNN) model earns the lowest accuracy of 0.62. The accuracy of the Long Short-Term Memory (LSTM) model is 0.75, which is lower than the accuracies attained by the best models.

The characteristics of the dataset and the model's algorithmic complexity can be used to explain the models' performance. The dataset includes reviews of Amazon products, and sentiment analysis is a logical fit for this kind of information. Convolutional neural networks are used in the CNN model, a deep learning algorithm that works well for text categorization tasks. Its excellent accuracy in comparison to conventional machine learning models reflects this.Among the conventional machine learning methods, linear models like Logistic Regression and SVC are easy and quick to train. However, non-linear models like KNN and Random Forest can be more computationally expensive and demand careful hyperparameter adjustment. This could account for why they performed poorly on this particular task.Whereas for the dataset- 2 the Random Forest model has the best accuracy (0.718), closely followed by the Logistic Regression, Multi Naive Bayes, and SVC models (accuracies ranging from 0.7055 to 0.707). On the other side, the LSTM model has the highest accuracy of 0.5656, followed by the K-Nearest Neighbour (KNN) model, which has the lowest accuracy of 0.55.

The characteristics of the dataset and the model's algorithmic complexity can be used to explain the models' performance. Aspect-based sentiment analysis of customer evaluations is included in the dataset with the aim of determining the sentiment of particular features of a good or service. Multiple decision trees are combined in the Random Forest model, an ensemble learning technique, to enhance performance and decrease overfitting. It is appropriate for this kind of dataset since it can manage both categorical and continuous characteristics.

And for the dataset-3 The Multi Naive Bayes model has the best accuracy (0.895), closely followed by the Logistic Regression and SVC models, whose accuracies range from 0.883 to 0.886. The Random Forest model has an accuracy of 0.836, followed by the K-Nearest Neighbour (KNN) model with the lowest accuracy of 0.81. The characteristics of the dataset and the model's algorithmic complexity can be used to explain the models' performance. Reviews of women's clothing products are included in the dataset, and the intention is to forecast the reviewer's mood. The Multi Naive Bayes model is a probabilistic model that can handle categorical and continuous information and assumes independence between features. It is appropriate for this kind of dataset because it is known to perform well on text data. With an accuracy of 0.882, the LSTM model performs comparatively well. LSTMs are

known to work well on sequential data and can capture long-term dependencies, which could explain its relatively good performance. The CNN model outperforms the majority of conventional machine learning models in accuracy (0.89), but its implementation is not disclosed.

Dataset 4 shows consistent accuracy ratings across models, with the Random Forest model attaining the best accuracy of 0.916, which is obtained from the Flipkart Product Customer Reviews dataset. With accuracy levels exceeding 0.89, the Logistic Regression, Multi NB, and SVC models also perform admirably. The CNN model is less accurate, with a 0.815 accuracy rating.

There are various reasons for the differences in accuracy across datasets and models. First and foremost, the nature and quality of the datasets themselves are very important. Datasets with descriptive and unambiguous attributes typically produce results with higher accuracy. The performance of the models is also impacted by the ambiguity and complexity of the data. Some datasets could have more complex patterns, which makes it more difficult for some models to anticipate outcomes correctly. Accuracy can also be impacted by how well the dataset's properties match those of the modelling methods. Each model has its own assumptions and constraints, and the degree to which the dataset supports these assumptions will determine the accuracy of the results.

# Concluding Remarks

Considering all the four datasets there is clear conclusion has been made it portrays that for the first The performance of CNN, SVC, Multinomial Naive Bayes, and Logistic Regression was comparatively high.The effectiveness of Random Forest and K-nearest Neighbour (KNN) was lower.The F1 score, accuracy, and precision achieved via logistic regression were the greatest. On this dataset, logistic regression performed the best overall.For the dataset-2 (Aspect-Based Sentiment Analysis, SemEval 2014 Task 4) Better results were demonstrated using Random Forest, Logistic Regression, and Multinomial Naive Bayes. KNN, CNN, and LSTM all had subpar performance. The best F1 score, recall, and precision were attained by Random Forest. On this dataset, Random Forest performed the best overall.And considering Women's E-commerce Clothing Reviews Dataset 3 SVC, Multinomial Naive Bayes, and Logistic Regression all worked nicely.Results from Random Forest, CNN, and LSTM were equivalent.While Multinomial Naive Bayes had the highest precision and F1 score, Random Forest had the highest recall.And for the last one (Flipkart Product Customer Reviews) Dataset 4 The performance of KNN, Random Forest, and Logistic Regression was good.Lower scores were attained by CNN, LSTM, and Multinomial Naive Bayes. The precision, recall, and F1 score of Random Forest were the greatest.Random Forest outperformed all other algorithms on this dataset.Random Forest consistently performed well when taking into account the performance across the four datasets as a whole, achieving good precision, recall, and F1 scores. The effectiveness of various models varied across datasets, highlighting the impact of dataset properties and domain-specific features, although it's vital to remember that.

.

# 4.Future Work

In the realm of product reviews, the future holds immense potential for advancements in sentiment analysis.Researchers and developers will probably concentrate on enhancing and extending the capabilities of sentiment analysis algorithms as technology continues to advance. The creation of more precise and context-sensitive sentiment classification models will be one of the main areas for advancement. With the help of enormous volumes of labelled data, these models will be trained to recognise the subtleties and complexities of human language, such as sarcasm, irony, and cultural allusions.

Additionally, sentiment analysis won't merely be used on text-based reviews.Multimodal sentiment analysis, which includes visual and aural clues including images, videos, and voice recordings to provide a more thorough knowledge of the sentiments conveyed by consumers, will become increasingly prevalent in the future. As a result, product reviews will be examined more comprehensively, taking into account not only the textual information but also the feelings expressed through different modalities.

In addition, with the growth of social media and user-generated content, sentiment analysis will encompass websites like Twitter, Instagram, and YouTube in addition to traditional review sites. This growth calls for the creation of specialised algorithms that can work with multiple data sources and extract sentiments from different forms. Finally, the use of sentiment analysis in product reviews in the future promises to revolutionise how we comprehend and analyse customer feedback, empowering companies to make data-driven decisions and improve their goods and services accordingly.

Furthermore, future work will explore the integration of sentiment analysis with other data sources, such as social media trends, user demographics, and purchase history. By combining sentiment analysis with additional data points, businesses can uncover deeper insights into customer preferences and behavior, enabling them to make more informed decisions regarding product development, marketing strategies, and customer support.

Overall, the future of work in product reviews using sentiment analysis holds immense potential for businesses to gain valuable insights into customer sentiments, improve their products and services, and enhance the overall customer experience. With continued advancements in natural language processing and machine learning, we can expect more sophisticated and accurate sentiment analysis models that revolutionize the way companies analyze and leverage customer feedback.

# References:


Zhang, X., Zhang, X., & Sun, Y. (2016). Mining online reviews for predicting patient satisfaction with hospitals. Computers in Human Behavior, 60, 188-199.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational Linguistics, 37(2), 267-307.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

Cambria, E., & Hussain, A. (2012). Sentic computing: Techniques, tools, and applications. Springer.

Xu, J., Mukherjee, A., & Liu, B. (2016). Mining twitter for sarcasm: A review of current methodologies. In Handbook of Natural Language Processing (Second Edition) (pp. 579-602). CRC Press.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113.

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. Eighth International Conference on Weblogs and Social Media.

Poria, S., Cambria, E., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. Information Fusion, 37, 98-125.

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. IEEE Transactions on Knowledge and Data Engineering, 23(10), 1498-1512.

Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion word expansion and target extraction through double propagation. Computational Linguistics, 37(1), 9-27.

Lu, B., & Liu, Y. (2012). The role of textual and non-textual information in online product reviews: A multimodal neural approach. Proceedings of the 20th ACM International Conference on Information and Knowledge Management, 158-167.

Wang, C., Zhu, F., & Zhu, W. (2011). An unsupervised approach for product feature extraction from online consumer reviews. Expert Systems with Applications, 38(10), 12419-12427.

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. Proceedings of the Seventh International Conference on Language Resources and Evaluation, 2200-2204.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational Linguistics, 37(2), 267-307.

Wang, C., Huang, L., Zhu, F., & Zhu, W. (2012). Mining consumer opinions from reviews: A comparison of pivot-based and graph-based methods. Decision Support Systems, 53(4), 759-767.

Kim, Y., & Zhang, X. (2014). A comparison of consumer review and eWOM sentiment analysis: Evidence from Chinese hotels. International Journal of Hospitality Management, 42, 65-71.

Li, L., Li, Y., Liu, Y., Huang, Y., & Li, X. (2018). Feature extraction for Chinese online reviews based on aspect and sentiment co-occurrence analysis. Knowledge-Based Systems, 156, 106-117.

Sun, X., Xu, L., Li, H., Liu, X., & Zhou, J. (2020). A deep joint aspect sentiment model for product review analysis. Journal of Intelligent Manufacturing, 31(2), 431-441.

X. Fang and J. Zhan, “Sentiment analysis using product review data,”
J. Big Data, 2015, doi: 10.1186/s40537-015-0015-2.

K. Tamara and N. Milićević, “Comparing Sentiment Analysis and Document Representation Methods of Amazon Reviews,” SISY 2018 - IEEE 16th Int. Symp. Intell. Syst. Informatics, Proc., pp. 283–288, 2018, doi: 10.1109/SISY.2018.8524814.

J. Sadhasivam and R. B. Kalivaradhan, “Sentiment analysis of Amazon products using ensemble machine learning algorithm,” Int. J. Math. Eng. Manag. Sci., vol. 4, no. 2, pp. 508–520, 2019, doi: 10.33889/ijmems.2019.4.2-041.

W. Tan, X. Wang, and X. Xu, “Sentiment Analysis for Amazon Reviews,” Reseaechgate, no. March, pp. 0–9, 2019, doi: 10.13140/RG.2.2.13939.37920.

R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh, Sentiment Analysis on Product Reviews Using Machine Learning Techniques : Proceeding of CISC 2017 Sentiment Analysis on Product Reviews Using Machine Learning Techniques, no. January. Springer Singapore, 2019.

A. Singh, A. Agarwal, and P. Dimri, “ScienceDirect ScienceDirect Comparative Study of Machine Learning Approaches for Amazon Reviews,” Procedia Comput. Sci., vol. 132, pp. 1552–1561, 2018, doi: 10.1016/j.procs.2018.05.119.

C. Rain, “Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning,” Swart. Coll., 2013, doi: 10.1097/IOP.0b013e318213f5d9.

H. Nguyen, R. Al, and K. Academy, “Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches,” SMU Data Sci. Rev., vol. 1, no. 4,2018.

S. A. Aljuhani and N. S. Alghamdi, “A comparison of sentiment analysis methods on Amazon reviews of Mobile Phones,” Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 6, pp. 608–617, 2019, doi: 10.14569/ijacsa.2019.0100678.

K. L. Santhosh Kumar, J. Desai, and J. Majumdar, “Opinion mining and sentiment analysis on online customer review,” 2016 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2016, 2017, doi: 10.1109/ICCIC.2016.7919584.

B. Yuen, X. Dong and T. Lu, "Inter-Patient CNN-LSTM for QRS Complex Detection in Noisy ECG Signals," in IEEE Access, vol. 7, pp. 169359-169370, 2019, doi: 10.1109/ACCESS.2019.2955738

https://doi.org/10.1016/j.ijcce.2021.11.003

https://doi.org/10.1007/978-981-99-0047-3