Unveiling Consumer Sentiments: AI-Powered Sentiment Analysis of Amazon Reviews Using Logistic Regression

Jivites Damodar*, Ramkumar K R*, Sukanthan S*, Mehan Ranka*, Lekshmi C. R.*
*Amrita School of Artificial Intelligence, Coimbatore, Amrita Vishwa Vidyapeetham, India.
Emails: {cb.sc.u4aie24020, cb.sc.u4aie24042, cb.sc.u4aie24056, cb.sc.u4aie24031}@cb.students.amrita.edu, cr_lekshmi@cb.amrita.edu

***Abstract*—This study applies sentiment analysis using natural language processing (NLP) to evaluate customer opinions from Amazon product reviews. The methodology begins with text preprocessing, including stopword and punctuation removal, lowercase conversion, and transformation into numerical features using TF-IDF (Term Frequency-Inverse Document Frequency). Logistic regression, a widely used and interpretable classification model, is then employed to classify reviews as positive, negative, or neutral due to its efficiency, robustness, and suitability for text-based sentiment analysis. A key contribution is the integration of real-time review extraction, fake review detection, and review summarization, enhancing the reliability of sentiment insights. The model's performance is assessed using accuracy, precision, recall, and F1-score. This approach helps businesses understand consumer sentiment, refine marketing strategies, and improve product recommendations. Beyond e-commerce, its applications extend to social media monitoring, brand reputation analysis, and automated customer feedback processing.**

***Index Terms*—Sentiment analysis, Natural language processing Logistic regression, stop words, Accuracy, TF-IDF**

I. Introduction

The rise of online shopping has significantly changed how consumers make purchasing decisions, with platforms like Amazon offering a vast range of products. Customer reviews play a key role in evaluating product quality, usability, and overall satisfaction. However, the large volume of reviews can make it difficult for buyers to analyze feedback effectively. Additionally, some reviews may be biased, misleading, or contradictory, making the decision-making process more complex.

To tackle these challenges, sentiment analysis can help classify product reviews into positive, negative, or neutral categories. A Random Forest classifier, known for its reliability and accuracy, can analyze review text to determine sentiment. Additionally, by incorporating real-time review extraction, users can input an Amazon product URL to retrieve the latest customer feedback instantly. This allows for a more efficient evaluation of product sentiment, helping consumers make well-informed choices.

By automating sentiment analysis, consumers can quickly access relevant insights without manually sifting through extensive reviews. This approach also helps identify unreliable or manipulated feedback, leading to more transparent and trustworthy purchasing decisions. As online shopping continues to evolve, improving the way product reviews are analyzed enhances the overall shopping experience.

*Corresponding author: Lekshmi C. R. (Email: cr_lekshmi@cb.amrita.edu)

II. Related Work

In the e-commerce space, sentiment analysis has drawn a lot of interest, especially when it comes to examining customer reviews to glean insightful information. To improve the accuracy of consumer sentiment classification, researchers have investigated various kinds of machine learning and deep learning approaches. With an emphasis on various approaches, feature extraction strategies, and classification models, this section highlights significant research that has advanced sentiment analysis.

Xiao et al. [20] used natural language processing (NLP) techniques to analyze Amazon product reviews for sentiment. The Amazon Customer Reviews Dataset, which was accessed via Amazon Simple Storage Service (Amazon S3), was used in their study. Several machine learning models, including Random Forest, were used to improve classification accuracy. Metrics like accuracy, precision, recall, and F1-score were taken into account during the evaluation process [2], [4]. In order to improve the accuracy of sentiment classification, the study further took user behavior patterns and review authenticity into account.

Saaqib et al. [16] used supervised machine learning algorithms to study sentiment prediction in Amazon product reviews for musical instruments. The study used Logistic Regression for classification and TF-IDF for feature extraction. The dataset was divided into 25% testing data and 75% training data. The precision, recall, and F1-score of several classifiers were assessed, including Logistic Regression, Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Naïve Bayes. The most accurate of these was Logistic Regression.

Tan et al. [18] investigated deep learning methods like Recurrent Neural Networks (RNNs) and conventional machine learning methods like Naïve Bayes, SVM, and KNN for sentiment analysis tasks. In a similar vein, Elli et al. [2] used

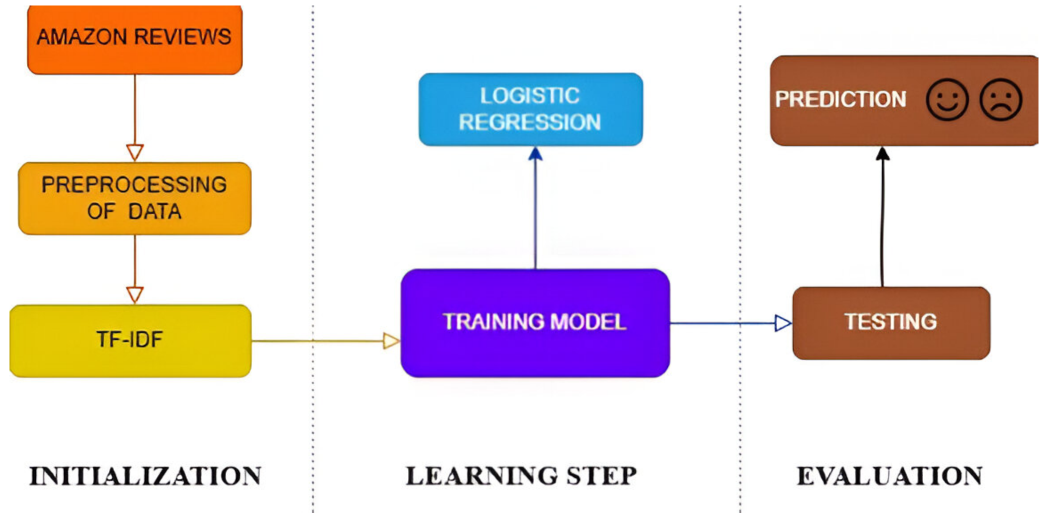


Fig. 1. Block diagram of the proposed sentiment analysis method using Logistic Regression

machine learning techniques to identify fraudulent reviews and analyze trends in business applications. Their study shed light on the ways sentiment analysis can be applied to spot phony reviews and trends that undermine customer confidence.

Using the Amazon review dataset, Rain et al. [15] conducted sentiment analysis, classifying reviews as either positive or negative using Naïve Bayes and decision list classifiers. The dataset mostly consisted of Amazon Kindle reviews and book reviews. In the meantime, Shaikh et al. [17] looked into a number of feature extraction and selection methods for sentiment analysis. Amazon reviews for books, music, and cameras were included in their dataset. Preprocessing steps included filtering out special characters and removing stopwords. Phrase-level features performed far better than single-word, multi-word, and phrase-level feature selection methods, according to the study. The Naïve Bayes classifier was the only one used, which prevented more extensive comparisons with other classification models. Additional research in sentiment analysis includes studies by Yue et al. [22], who examined Amazon reviews using machine learning techniques, and Fang et al. [4], who focused on product review sentiment methodologies. Jiang et al. [7] explored advanced opinion mining techniques, while Liu et al. [14] introduced cross-domain sentiment-aware word embeddings to enhance classification accuracy. Sentiment Analysis using Improved Vader and Dependency Parsing is done in [19].

Despite significant advancements in sentiment analysis, several challenges remain. Contextual understanding, multilingual adaptability, and aspect-based sentiment analysis require further development to improve the precision and scalability of insights. Additionally, limited research on fake review detection and comparative analyses between traditional and modern approaches hinders the creation of more robust sentiment analysis models. Addressing these gaps can lead to more accurate, reliable, and context-aware sentiment analysis systems, expanding their applications across various domains.

III. METHODOLOGY

The general structure of the suggested sentiment analysis method is shown in Figure 1. The first step in the process is gathering a dataset of product reviews that have been classified as neutral, negative, or positive. Preprocessing steps include changing all text to lowercase for consistency and eliminating extraneous components like punctuation and stopwords.

Term Frequency-Inverse Document Frequency (TF-IDF) is used to transform textual data into a numerical format appropriate for machine learning. By emphasizing the importance of words in a review, this method makes sure that commonly used but uninformative words don't take over the analysis. Because it is effective at handling textual data, a Logistic Regression model is used for sentiment classification. To enable the model to identify patterns in the data, the dataset is divided into training and testing sets. Model performance is measured using accuracy, precision, recall, and F1-score to ensure a comprehensive assessment [2].To enhance the effectiveness of sentiment classification, we incorporated extractive text summarization using a TextRank-based approach. Inspired by the PageRank algorithm, this method constructs a graph where sentences are nodes and edges represent similarity. By computing a ranking score for each sentence, the most informative ones are extracted to form a concise summary, helping the model focus on key textual elements while reducing noise.

IV. SYSTEM DESCRIPTION

Customer opinions are collected and interpreted from product reviews by the proposed sentiment analysis system. Reviews in the dataset are divided into three sentiment classes: neutral, negative, and positive. Preprocessing methods like text normalization, stopword removal, and punctuation removal are used to improve the quality of textual data. The processed text is then represented numerically using TF-IDF, highlighting the significance of important terms. These features are used to train a Logistic Regression model for sentiment classification.

Accuracy, precision, recall, and F1-score are used to assess performance after the dataset is split into training and testing sets.

The structure of the paper is as follows: Section V presents the details of the proposed methodology, Section VI discusses the results, and Section VII provides the conclusion.

V. Proposed Methodology

To process product reviews effectively, the sentiment analysis framework includes a number of crucial steps. The methodology is outlined by the following crucial steps:

- **Data Collection:** Compiling a dataset of product reviews along with sentiment labels for each review.
- **Text Preprocessing:** Normalizing and cleaning text by converting words to lowercase and reducing punctuation and stopwords.
- **Feature Extraction:** utilizing TF-IDF to convert textual data into numerical features.
- **Model Training:** putting in place a classifier using logistic regression to group reviews according to sentiment. Performance Evaluation:
- evaluating the classifier's performance by measuring its accuracy, precision, recall, and F1-score.
- **Text Summarization :**Text Summarization: Applying extractive summarization techniques such as TextRank to condense lengthy reviews into concise, meaningful summaries while preserving key information.

Thus sentiment analysis model seeks to improve the ability to effectively interpret customer feedback by offering a dependable classification system for product reviews through the use of this structured approach. The steps are explained below.

A. Dataset

1) Data acquisition and preprocessing: This study makes use of the Stanford Network Analysis Project (SNAP)-sourced Amazon reviews dataset from Kaggle [1], which comprises more than 34 million reviews over 18 years. [2] User reviews with ratings, textual comments, and product details are included in the dataset. Reviews with ratings of 1 or 2 are categorized as negative, while those with ratings of 4 or 5 are classified as positive. Reviews with a score of 3 are left out in the sentiment classification process. This dataset's diversity improves model generalizability in contrast to previous research that concentrates on particular product categories.

- **Polarity** - 1 for negative and 2 for positive.
- **Title** - Review heading.
- **Text** - User-provided review body.

2) Data Preprocessing: To ensure data quality, preprocessing techniques are applied, including handling missing values, eliminating unnecessary information, and transforming text into a structured format suitable for analysis.

[1]https://www.kaggle.com/datasets/kritanjalijain/amazon-reviews

[2]https://snap.stanford.edu/data/web-Amazon.html

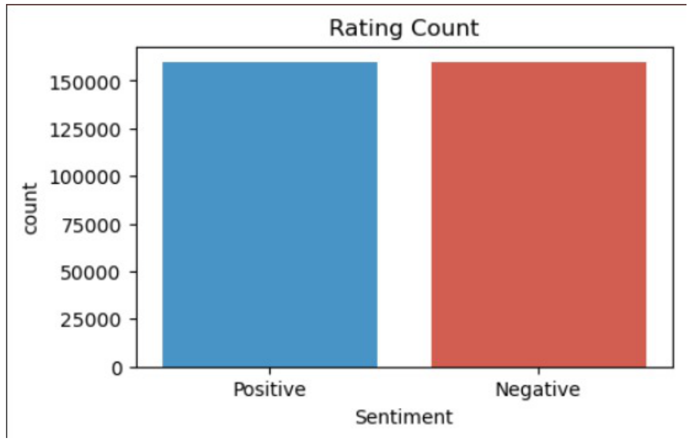


Fig. 2. Rating Distribution in Dataset

- **Lowercasing** - Converts all text to lowercase to maintain uniformity and prevent duplicate word representations.
- **Removing Special Characters and Punctuation** - Eliminates non-alphanumeric symbols (e.g., @,!,#) that do not contribute to sentiment analysis.
- **Stopword Removal** [1] - Removes common words such as "the," "is," and "and" to enhance focus on meaningful terms.
- **Lemmatization** [1] - Converts words to their root form, ensuring consistency across variations of the same term.

3) Data Transformation: The cleaned text data is converted into numerical features for machine learning models.

- **TF-IDF Vectorization** [21] - Term Frequency-Inverse Document Frequency (TF-IDF) quantifies the importance of words within a review. It is calculated as follows: [2]

$$TF(T,D) = \frac{\text{Number of times } T \text{ appears in } D}{\text{Total terms in } D} \tag{1}$$

$$IDF(T) = \log\left(\frac{\text{Total documents}}{\text{Number of documents containing } T}\right) \tag{2}$$

- **Label Encoding** - Converts sentiment labels into numerical values, where 1 represents positive sentiment and 2 represents negative sentiment [2].
- **Training and Testing Split** - The dataset is divided into 80% for training and 20% for testing to evaluate model performance effectively.

B. Classification Model: Logistic Regression

The model we have chosen for sentiment analysis on Amazon reviews is Logistic Regression. Logistic regression [23] is a widely used machine learning algorithm for binary or multiclass classification tasks [6], such as predicting positive and negative sentiments. It works by estimating the probability of a class using a logistic function and is efficient for text data when combined with TF-IDF features. Logistic regression can handle large datasets well. Its speed and simplicity make it a good choice for sentiment analysis tasks.

The sigmoid function, also known as the logistic function, is used to map the predicted values to probabilities. It transforms any real value into a range between 0 and 1, forming an S-shaped curve [23]:

$$P(y=1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \tag{3}$$

where:

- $P(y=1|X)$ represents the probability that a given input X belongs to class 1 (positive sentiment).
- β_0 is the bias term.
- $\beta_1, \beta_2, \ldots, \beta_n$ are the model coefficients learned during training.
- $X_1, X_2, \ldots, X_n$ are the input features extracted from the text using TF-IDF.

By applying *maximum likelihood estimation (MLE) [13]*, the model learns optimal parameters that best separate positive and negative reviews.

Logistic Regression provides interpretable coefficients, making it easy to understand sentiment classification. It is efficient for large-scale datasets, trains quickly with TF-IDF features, and performs well without requiring extensive computational resources or large training data, making it ideal for real-world applications.

C. Performance Evaluation

Performance evaluation measures how well the model performs on the testing dataset. For sentiment analysis, commonly used metrics include Accuracy, Precision, Recall, and F1-score [1].

- **Accuracy** measures how often the classifier makes the correct prediction. It is the ratio of correct predictions to the total number of predictions [2].

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Data Points}} \tag{4}$$

- **Precision** quantifies how many of the predicted positive instances are actually correct [2].

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{5}$$

- **Recall** measures how many of the actual positive instances in the dataset were correctly predicted by the model [2].

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{6}$$

- **F1-score** combines precision and recall into a single metric, representing their harmonic mean.

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

The dataset of 399,999 reviews (200,000 negatives, 199,999 positive) was split, keeping 80% (319,999 records) for model development. Using train_test_split, 75% was allocated for training and 25% for testing. This approach ensured a robust evaluation, allowing the model to learn patterns from a large training set while being tested on an unseen subset to assess its generalization performance accurately. The various hyper-parameters we used in the proposed framework are shown in Table I

TABLE I
HYPERPARAMETER CONFIGURATION FOR SENTIMENT ANALYSIS PIPELINE

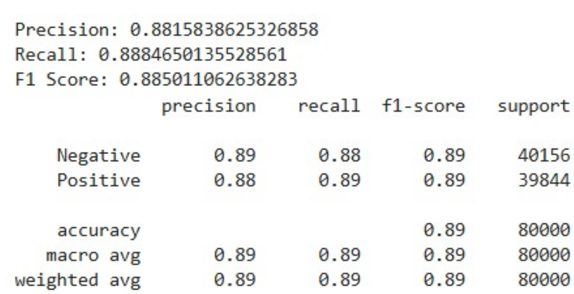
Hyperparameter	**Value**	**Used In**
max_features	5000	TF-IDF Vectorizer
ngram_range	(1,1)	TF-IDF Vectorizer
n_splits	1	StratifiedShuffleSplit
test_size	0.2 (20% data ignored)	StratifiedShuffleSplit
train_size	0.8 (80% data used)	StratifiedShuffleSplit
test_size	0.25	train_test_split
train_size	0.75	train_test_split
random_state	0	train_test_split

VI. RESULTS AND ANALYSIS

A. Overall Performance Evaluation

The sentiment analysis model demonstrated strong performance based on key classification metrics (ref Fig3). The overall accuracy of the model was 89%, meaning that it correctly predicted sentiments for the majority of the reviews. This high accuracy indicates that the model effectively distinguishes between positive, negative, and neutral reviews.

With a precision score of 88.16%, the model was accurate 88.16% of every instance when classifying a review as positive. This indicates a high degree of correctness in identifying favorable sentiments while reducing inaccurate categorizations. The model successfully identified 88.85% of all real positive reviews, according to the recall score of 88.85%. A high recall indicates that the model is effective at identifying the majority of positive reviews, preventing a greater number of positive sentiments from being incorrectly categorized as neutral or negative. Lastly, the F1-score, which strikes a balance between recall and precision, was 88.50%. This measure demonstrates that the model is a well-rounded method for sentiment classification since it consistently performs well across various sentiment categories.


```
Precision: 0.8815838625326858
Recall: 0.8884650135528561
F1 Score: 0.885011062638283
              precision    recall  f1-score   support

    Negative       0.89      0.88      0.89     40156
    Positive       0.88      0.89      0.89     39844

    accuracy                           0.89     80000
   macro avg       0.89      0.89      0.89     80000
weighted avg       0.89      0.89      0.89     80000
```

Fig. 3. Performance Metrics of Proposed Method

These findings show that the Logistic Regression model with TF-IDF maintains efficiency and interpretability while offering powerful predictive capabilities.

B. Word Cloud Analysis

A word cloud displaying the most often used terms was created from Amazon reviews in order to obtain a better understanding of customer sentiments. Common themes like product quality, cost, and general customer satisfaction are highlighted in this graphic. Aspects that customers often mention in their reviews are indicated by words that are larger in the word cloud.

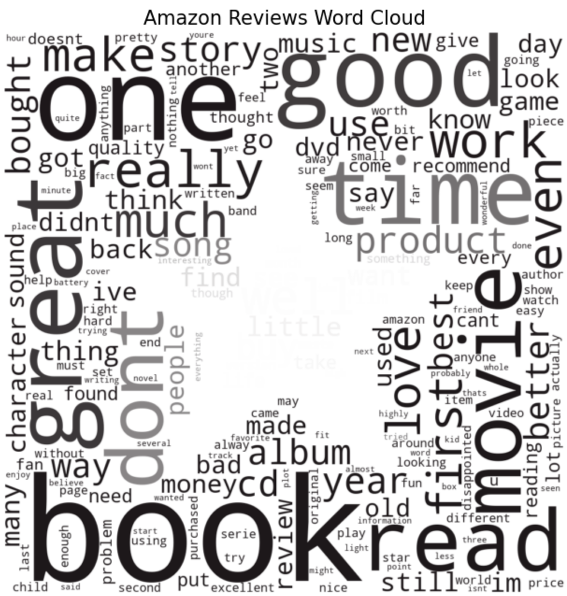


Fig. 4. Word Cloud Representation of Amazon Reviews

By means of word cloud analysis, we found that the terms ”quality,” ”price,” ”durability,” and ”delivery” were used a lot, highlighting their significance in customer feedback. Words like ”great,” ”excellent,” and ”value” were frequently used in positive reviews, whereas ”poor,” ”broken,” and ”slow” were frequently used in negative reviews. Businesses and sellers can better understand customer expectations and areas for improvement by identifying these keywords.

Monitoring the occurrence of particular terms can also highlight new trends in customer reviews. For instance, an increase in terms like ”refund” or ”defective” could indicate escalating discontent and lead companies to look into possible problems. E-commerce platforms can improve customer satisfaction, product quality, and strategy with the help of these insights.

C. Text Summarization

In Natural Language Processing (NLP), text summarization refers to the process of condensing long texts into shorter, concise versions while retaining the most essential information and meaning. One popular approach is extractive summarization, which involves selecting and combining key phrases or sentences directly from the source text to form a summary. The TextRank algorithm, inspired by the PageRank algorithm originally used to rank web pages, is widely used for this purpose. In PageRank, a score is computed to represent the likelihood of a user visiting a page by creating a square transition matrix M, where each entry represents the probability of moving from page i to page j. If a page has no outgoing links (a dangling page), it is assumed that the user can transition to any page with equal probability. TextRank adapts this idea by treating sentences instead of web pages as nodes, and using the similarity between sentences—rather than hyperlinks—as the transition weights. These similarity scores are stored in a matrix analogous to the PageRank matrix. TextRank is an extractive, unsupervised method that ranks sentences based on importance and selects the top ones for the final summary. Additionally, techniques like GloVe (Global Vectors for Word Representation), developed by Stanford, provide pre-trained word embeddings that can enhance sentence similarity calculations by representing words as vectors in a continuous vector space.

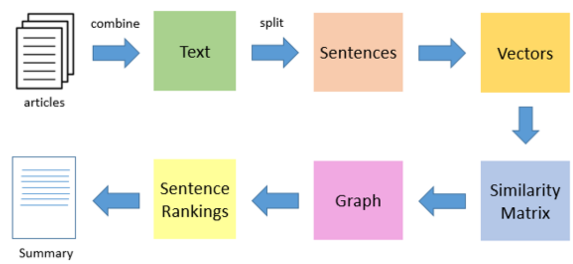


Fig. 5. Text summarization Work flow

D. Comparison of Sentiment Analysis Methods

We assessed the accuracy of our method by contrasting it with earlier sentiment analysis research on Amazon reviews. The performance of several models used in sentiment classification is compiled in Table II.

The comparison shows that our strategy, which combined TF-IDF and Logistic Regression, achieved an accuracy of 89%, placing it competitively with earlier approaches. Our approach provides a balance between classification performance and computational efficiency, even though deep learning models have occasionally demonstrated higher accuracy.

E. Confusion Matrix Analysis

To assess the classification performance more thoroughly, a confusion matrix was created. It shows possible areas for improvement and offers insights into how well the model can identify sentiments.

The reliability of the model was confirmed by our analysis of the confusion matrix in Fig. 6, which showed that the majority of reviews, both positive and negative, were correctly classified. There was very little misclassification, as evidenced by the low number of false positives and false negatives. Furthermore, the model performed marginally better at forecasting positive sentiments than negative ones, which may have been impacted by the dataset’s sentiment class distribution. Thus our model is a reliable method for sentiment classification in

TABLE II
COMPARISON OF SENTIMENT ANALYSIS METHODS ON AMAZON REVIEWS

Paper Title	Year	Dataset	Accuracy
Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning [15]	2013	Reviews of books	84.44%
Amazon Reviews, Business Analytics with Sentiment Analysis [2]	2016	Reviews of cellphones & accessories	80.11%
Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews [17]	2016	Reviews of books, music, and cameras	70%, 80%, 80%
Analysis of Sentiment on Amazon Product Reviews [16]	2023	Reviews of musical instruments	94%
Our Approach: Logistic Regression + TF-IDF	2025	Mixed-category Amazon reviews	89%

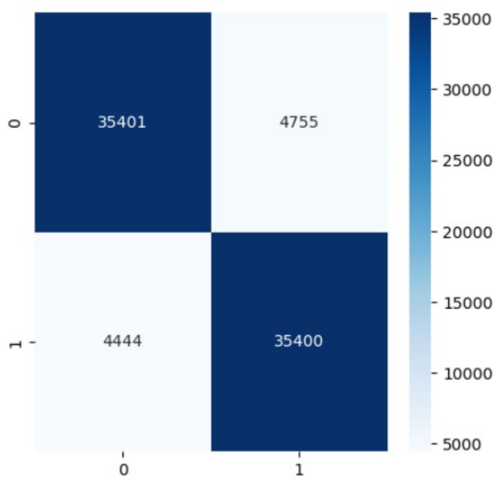


Fig. 6. Confusion Matrix

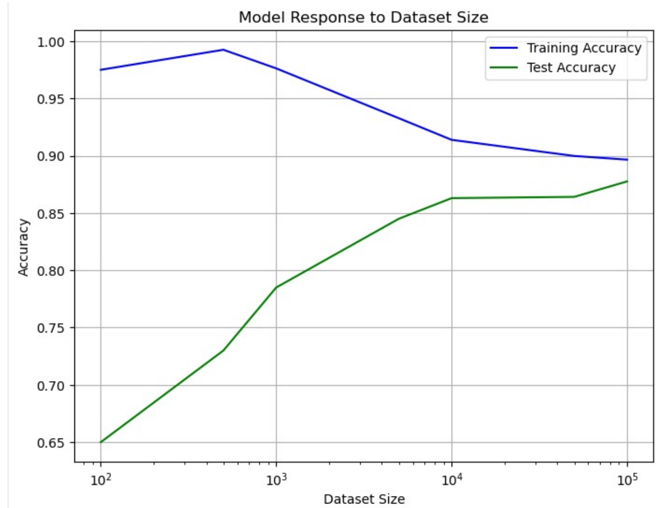


Fig. 7. Model Response to Dataset size

e-commerce applications because it generally achieved high accuracy with a low error rate.

F. Comparison with Existing Approaches

Table II presents a comparative analysis of sentiment analysis models. While deep learning-based methods achieve higher accuracy, they require extensive data and computational resources. Our approach, using Logistic Regression with TF-IDF, balances accuracy and efficiency, making it practical for large-scale sentiment analysis. Compared to probabilistic and feature-based models, our method demonstrates improved accuracy due to optimized feature selection. Additionally, it offers better interpretability than black-box deep learning models, making it a suitable choice for real-world applications [3].

G. Impact of Dataset Size on Model Performance

The graph 7 demonstrates the influence of dataset size on model performance, focusing on training and test accuracy. Training accuracy starts near 100% with smaller datasets, reflecting overfitting, but gradually declines and stabilizes around 90% as dataset size increases, indicating improved generalization. Test accuracy, initially low with smaller datasets, steadily improves with larger datasets and stabilizes at approximately 88–90%, showing enhanced performance on unseen data. The narrowing gap between training and test accuracy highlights how larger datasets reduce overfitting and improve generalization. This analysis underscores the importance of sufficient data in building robust and reliable machine learning models.

H. Discussion and Future Scope

Our sentiment analysis approach using Logistic Regression with TF-IDF offers notable advantages, including computational efficiency, making it suitable for real-time applications, interpretability through clear feature weights that identify sentiment-driving words, and scalability, as it handles large datasets effectively without significant computational overhead. However, the model faces limitations, such as struggling with contextual sentiment understanding and failing to detect sarcasm, irony, or complex linguistic expressions, which are better handled by transformer-based models.

Future enhancements can address these limitations and improve sentiment analysis further. Transformer-based models like BERT and RoBERTa can enhance the contextual understanding of sentiments while expanding the model to support multi-lingual analysis will make it more inclusive for Amazon's global user base [5]. Aspect-based sentiment Analysis (ABSA) can provide detailed insights into specific product features, and integrating fake review detection mechanisms will ensure greater credibility. By adopting these advancements, sentiment analysis systems can become more robust, context-aware, and effective in addressing diverse real-world applications [8]–[12].

VII. Conclusion

This research successfully implements sentiment analysis using logistic regression with TF-IDF, achieving an accuracy of 89%. The model offers an efficient and interpretable approach, balancing performance and computational efficiency, making it suitable for large-scale sentiment classification. Compared to existing methods, our approach demonstrates strong accuracy while maintaining simplicity. Future enhancements, including deep learning and aspect-based sentiment analysis, can further refine sentiment classification and broaden its applicability.

References

[1] Gitanshu Chauhan, Akash Sharma, and Nripendra Dwivedi. Amazon product reviews sentimental analysis using machine learning. In *2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)*, volume 5, pages 1202–1206, 2024.

[2] Maria Soledad Elli, Yi-Fan Wang, and Abdulaziz Elwalda. Amazon reviews, business analytics with sentiment analysis. *Elwalda, Abdulaziz, et al."Perceived Derived Attributes of Online Customer Reviews*, 2016.

[3] Lixin Fan, Chang Liu, Yuhao Zhou, Tianyu Zhang, and Qiang Yang. Interpreting and evaluating black box models in a customizable way. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 5435–5440, 2020.

[4] Xing Fang and Justin Zhan. Sentiment analysis using product review data. *Journal of Big data*, 2:1–14, 2015.

[5] Xinli Guo. Sentiment analysis based on roberta for amazon review: An empirical study on decision making. *arXiv preprint arXiv:2411.00796*, 2024.

[6] Jackson Isaac and Sandhya Harikumar. Logistic regression within dbms. In *2016 2nd International conference on contemporary computing and informatics (IC3I)*, pages 661–666. IEEE, 2016.

[7] Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang. An approach based on tree kernels for opinion mining of online product reviews. In *2010 IEEE International Conference on Data Mining*, pages 256–265. IEEE, 2010.

[8] B Jin and X Xu. Carbon emission allowance price forecasting for china guangdong carbon emission exchange via the neural network. global finance review. 2024; 6 (1): 3491, 2016.

[9] Bingzi Jin and Xiaojie Xu. Machine learning predictions of regional steel price indices for east china. *Ironmaking & Steelmaking*, page 03019233241254891, 2024.

[10] Bingzi Jin and Xiaojie Xu. Pre-owned housing price index forecasts using gaussian process regressions. *Journal of Modelling in Management*, 19(6):1927–1958, 2024.

[11] Bingzi Jin and Xiaojie Xu. Wholesale price forecasts of green grams using the neural network. *Asian Journal of Economics and Banking*, (ahead-of-print), 2024.

[12] Bingzi Jin, Xiaojie Xu, and Yun Zhang. Thermal coal futures trading volume predictions through the neural network. *Journal of Modelling in Management*, 20(2):585–619, 2025.

[13] Rounak A. Kharait, Jeff Newmiller, Jackson Moore, and Cegeon Chan. Maximum likelihood estimation (mle) approach for determining most representative solar resource data set for united states. In *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, pages 1240–1243, 2021.

[14] Jun Liu, Shuang Zheng, Guangxia Xu, and Mingwei Lin. Cross-domain sentiment aware word embeddings for review sentiment analysis. *International Journal of Machine Learning and Cybernetics*, 12:343–354, 2021.

[15] Callen Rain. Sentiment analysis in amazon reviews using probabilistic machine learning. *Swarthmore College*, 42, 2013.

[16] Najam Ul Saaqib, Harsh Kumar Verma, et al. Analysis of sentiment on amazon product reviews. In *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pages 697–702. IEEE, 2023.

[17] Tahura Shaikh and Deepa Deshpande. Feature selection methods in sentiment analysis and sentiment classification of amazon product reviews. *International Journal of Computer Trends and Technology (IJCTT)*, 36(4):225–230, 2016.

[18] Wanliang Tan and Xinyu Wang. Analysis for amazon reviews. 2018.

[19] G Veena, Aadithya Vinayak, and Anu J Nair. Sentiment analysis using improved vader and dependency parsing. In *2021 2nd global conference for advancement in technology (GCAT)*, pages 1–6. IEEE, 2021.

[20] Yuanhang Xiao, Chengbin Qi, and Hongyong Leng. Sentiment analysis of amazon product reviews based on nlp. In *2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, pages 1218–1221. IEEE, 2021.

[21] En Yang and Zhaohua Long. Research on the weighting method based on tf-idf and apriori algorithm. In *2023 IEEE 6th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pages 1003–1005, 2023.

[22] Lin Yue, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin. A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 60:617–663, 2019.

[23] Xiaonan Zou, Yong Hu, Zhewen Tian, and Kaiyuan Shen. Logistic regression model optimization and case analysis. In *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, pages 135–139, 2019.