IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 4, July-August 2025

Impact Factor: 8.152

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

Sentiment Analysis and Prediction of Consumer Reviews of Clothing Product

Anamika Singh, Anusha B Y, Anita Seetaram Bhat

Dept. of Computer Applications, CMR Institute of Technology, Bengaluru, India

Dept. of Computer Applications, CMR Institute of Technology, Bengaluru, India

Dept. of Computer Applications, CMR Institute of Technology, Bengaluru, India

ABSTRACT: Sentiment analysis is a key component of Natural Language Processing (NLP), which is essential for comprehending client sentiment and enhancing commercial tactics. This study aims to forecast the tone of consumer reviews of clothing items in order to provide businesses with relevant information. To achieve the best results, this research uses a range of machine learning models, such as Random Forest, Logistic Regression, Naïve Bayes, LSTM, KNN, Decision Tree, XG Boost, ADA Boost, and Ensemble Voting Classifier. Significant patterns and trends were found after extensive data preprocessing, which included feature engineering, data cleaning, and exploratory data analysis (EDA). Among some of the visual aids, word clouds, sentence sentiment distribution graphs, and correlation analysis where obtained and found very helpful and improved knowledge of customer attitude. To ensure that the models are accurate, precision, recall, and F1-score, performance criteria were used to test the models. These findings can help companies to advance their products or service, enhanced customer satisfaction, etc. and also get to know more of their customers preferences. Consequently, they will be in a position to enhance their selling tactics and achieve better choices.

KEYWORDS: Ensemble Voting Classifier, EDA, Sentiment Analysis, ADA Boost, Correlation Analysis, and Customer Insights

I. INTRODUCTION

Due to the tremendous change in the way people consume clothing introduced by the internet, it has become important to understand the customer. Due to the rapid growth of e-commerce, internet reviews now play a significant role in decision-making. Sentiment analysis, a subfield of natural language processing (NLP), is tasked with examining how customers think and feel in order to make conclusions based on their assessments. Companies can use machine learning and deep learning techniques to enhance product recommendations that are provided to them. Boost general customer satisfaction and learn about customer preferences. The current project proposal is expected to study and forecast the buyer sentiment in clothing product reviews with the help of a variety of Machine and deep learning models. Through their performance we determine the quality with which different classification methods predict the customer attitude. Valuable issues such as how to work with imbalanced datasets, how to classify sentiments in a more accurate manner, and how to compare them are addressed. study deals with deep learning techniques in combination with traditional machine learning models. The following research questions were also answered.

1. Which deep learning or machine learning models are best at forecasting customer sentiment in evaluations of apparel products?
2. What information can sentiment analysis give companies to improve their marketing and product offerings?

Section 2 gives a summary of pertinent research on sentiment analysis of product reviews. The dataset, pre-processing procedures, and sentiment categorization methodology are described in depth in Section 3. The experimental results are shown in Section 4, where we assess and contrast several classification methods. The study is concluded with recommendations for further research in Section 5.

II. RELATED WORK

Scholars have been investigating a wide range of machine learning and deep learning approaches in order to apply sentiment analysis to reviews of clothing products. For sentiment classification and customer opinion prediction, traditional machine learning models such as Naïve Bayes, Random Forest, Support Vector Classifier (SVC), and

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

Extreme Gradient Boosting (XGB) have been used extensively [1]. In addition to these approaches, fashion product reviews have been successfully categorized using clustering algorithms, especially k-means [2]. Bidirectional recurrent neural networks (RNNs) with long short-term memory (LSTM) units are examples of deep learning models that have continuously produced good F1 scores in tasks involving recommendation and sentiment categorization [3].Additionally, semantic orientation techniques, which include context-based mining tactics and the stop word removal, have improved aspect-based sentiment classification [4]. All of these studies highlight how important sentiment analysis is to increasing consumer satisfaction, marketing strategy, and product planning in the fashion e-commerce industry. improving marketing strategies and customer happiness in the context of online fashion sales. Despite these advancements, problems persist, particularly when handling diverse datasets and exams covering multiple topics. Researchers have employed a variety of techniques to solve these issues, including the Synthetic Minority Oversampling Technique (SMOTE) for data balancing and negation handling to improve sentiment interpretation [1] [2]. One notable study extracted sentiment patterns from online reviews of an ecommerce platform using k-means clustering to identify removal of the stop word of the results and review text. recurrent themes and business development potential [5]. In another study, Word2Vec embeddings and the convolutional neural networks (CNNs) were combined in order to develop an information deep learning framework that performed better than traditional machine learning structures in sentiment classification accuracy [6].Sentiment analysis has been applied to many different areas but some researchers have conducted medicine reviews with more vigor in comparison to the other areas than reviews apparel in priority [7].Instead, the analysis of the user reviews of e-commerce websites selling women's clothes would be conducted based on machine learning algorithms to classify reviews either as positive or as negative, which gives details about what customers want via online compared to in store shopping [8].Besides, supervised learning models work well in denoting the right attitude of the consumers to an extensive body of sentiment polarization work upon data sets of Amazon product ratings [9].In a similar vein to categorize the opinion of the users based on reviews of the Indonesian online stores, it is possible to use a sentiment analysis technique. The Naive Bayes based approach was developed. It is based on the aspect approach, and can distinguish neutral, negative and positive attitudes[10]. Rating of sentiment may be complicated and require certain support along with comprehensive analysis and contextual knowledge although they present useful information. The perceptions of the customers may sometimes be prompted by expectations over the prices and quality of the goods. Sentiment analysis can thus be adopted as a helpful technique to businesses that seek to evaluate whether their prices meet market expectations. An instance is where buyers that hold a negative perception regarding the costly products would feel the latter are over priced but on the other hand others who hold a favourable view on lower priced products may feel they are giving good deals using large language models (LLMs) based on transformer-like deep learning architectures along with pre trained models [11] has made the sentiment analysis techniques even better recently. By capturing complex linguistic nuances, empirical investigations show that these models greatly improve sentiment categorization accuracy [12][13]. Furthermore, a comparison study has revealed how well various sentiment analysis approaches can anticipate the emotions of product reviews for e-commerce applications [14].

III. METHODOLOGY

The consumer-review-of-clothing-product dataset includes sales information from several retailers in various places. It includes facts like the year of establishment, size, location type, and outlet type, and also product-specific and store-specific statistics like weight, fat content, visibility, product type, and MRP. As year of establishment, size, locational type and outlet type. Preparing data was a precondition to the training of the machine learning model. The first step was mainly involved with how to handle missing values. An example is where by normally missing value in numerical columns such as Age was replaced with the mean or median. Rating to have uniformity throughout the data set. The most frequently occurring was the one that was utilized to manage categorical fields such as Department Name and Division Name to ensure significant data is maintained. After that, they were not at all suggestive of sentiment analysis, so they only served as mere identifiers, and thus were removed. such columns as Clothing ID were eliminated. Division Name, Department Name, Class Name and category data such as these can be used depending upon the situation. Review Title were converted and put in numerical form either through label encoding, or one-hot encoding. The regular expressions were applied in the cleaning, lemmatization, tokenization, and As pre-processing was completed, the data has been divided into training and testing portions of 80 and 20 percent respectively toward testing. Last but not least, exploratory data analysis, or EDA, was done to examine the distribution of the sentiments and occurrence of which latent trends or anomalies within that data set. The Random Forest Classifier, XGBoost Classifier, AdaBoost Classifier, Naïve Bayes, Decision Tree Classifier, K Nearest Neighbors (KNN), Ensemble Voting Classifier, and Long Short Term Memory (LSTM) were among the classification models that were used on the dataset. Three important indicators were employed to assess the models' performance. Recall, which stresses identifying all actual positive instances and is useful when failing to detect a positive sentiment carries a higher cost; F1 Score, which balances precision and recall and is ideal for handling imbalanced sentiment classes; and accuracy, which gauges the model's overall correctness and is best suited for balanced datasets. In this case, a more robust and dependable model for predicting consumer sentiment from review data was indicated by lower accuracy values along with better F1 scores and recall.

IV. RESULTS AND DISCUSSION

We looked into a number of the machine learning algorithms discussed in the previous section in order to address the first research question, and the findings are shown in Table 1, With the highest accuracy of 0.89 and F1-score of 0.94 among the models that were examined, the Gradient Boost Classifier and Ensemble Voting Classifier were the most successful in predicting sentiment in customer reviews of clothes. The ability of these models to detect both positive and negative attitudes in text data is demonstrated by their high recall scores (0.97 and 0.96, respectively). Gradient Boost's ability to use regularization to limit overfitting and boost iteratively to lower classification errors makes it a successful technique. High-scoring models included Naïve Bayes (accuracy 0.88, F1-score 0.93) and Random Forest Classifier (accuracy 0.87, F1-score 0.93). Their simplicity and ensemble nature, respectively, allowed them to produce reliable outcomes. Models such as ADA Boost (accuracy range 0.84–0.87), KNearest Neighbors (KNN) (F1-score 0.92), Decision Tree Classifier (F1-score 0.90), and LSTM (F1-score 0.90) performed somewhat worse. Even though they continued to generate correct predictions, their handling of sentiment shifts in review text was less trustworthy. According to these findings, ensemble techniques that can control feature complexity, enhance classification accuracy, and generalize well across data—like Gradient Boost and Voting Classifier— are more appropriate for sentiment analysis in textual customer evaluations. Similarly, because it integrates the advantages of several models to generate predictions that are more dependable and broadly applicable, the Ensemble Voting Classifier did well. Classifier Model Accuracy Recall F1-Score Random Forest 0.87 0.99 0.93 LSTM 085 091 090, K-Nearest Neighbors (KNN) 085 097 092, Decision Tree Classifier 084 091 091, Ensemble Voting Classifier 089 096 094, ADA Boost 087 096 093, and Naïve Bayes 088 094 093

Table1. Results

Model	Accuracy	Recall	F1-Score
Random Forest Classifier	0.87	0.99	0.93
Naïve Bayes	0.88	0.94	0,93
Gradient Boost	0.89	0.97	0.94
LSTM	0.85	0.91	0.90
KNN	0.85	0.97	0.92
Decision Tree classifier	0.84	0.91	0.91
Ensemble Voting classifier	0.89	0.96	0.94
ADA Boost	0.87	0.96	0.93

Table 1 performance metrics of different classifiers

During the exploratory data analysis (EDA) phase, we employed a range of visualization techniques to extract insights from the consumer review dataset in order to address the second research question. In order to understand the distribution of the feeling classes and see the words commonly being used in the positive and negative ones evaluations, visualizations were used (word clouds and pie and bar charts).A comparatively higher percentage of positive evaluations was indicated on the emotion distribution chart, which implies that a slight imbalance. It was upon this realization that techniques of pre processing such as stratified sampling were established to ensure that balanced data obtaining training. Word clouds especially gave a clear picture of presences and preferences of the clients, and zones of dissatisfaction focused on the emotional overtones to the analyses. As negative words, poor, tight, and disappointed were very common, whereas love and perfect, as positive words, were present as well. Frequent in favorable expressions were comfortable and confortable. It was also determined to establish the most crucial textual characteristics to use in the sentiment classification significance visualizations. To assist in the reduction of dimensionality and enhancement of features selection, we demonstrated how to implement some of the techniques like the following used natural language processing (NLP) to assign heavier weighting to certain words which exerted a

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

greater influence in terms of sentiment. Due to these findings, a more complex set of features was developed increasing the models accuracy of detection minute differences in the use of language according to sentences. Also, the graphical representations help us to see whether there are any neutral? review, or outliers, which when not managed properly would skew the results.

The bias during preprocessing might reduce the accuracy of the model.These discoveries directly benefited our model selection process.The visual indicators emphasized the need to employ models that would be able to accommodate linguistic variance and unbalanced data. Models like Ensemble Voting Classifier, Gradient Boost, and Naïve Bayes have shown remarkable performance in terms of recall and F1score. We were also able to identify instances in which models were incorrectly identifying emotions by using visual comparisons between model outputs, such as confusion matrices and classification reports. The selection of classifiers that could successfully generalize on unseen data and the adjustment of hyperparameters were made possible by this recurring feedback loop between the model and visual analysis.

V. VISUALISATION

Cloth Course Distribution Visualizations as shown in Fig 1. The collection consists of a limited number of apparel types, which represent the bulk of transactions and assessments. Fig 1 breaks down the scope of each texture instruction, helping retailers spot which categories perform best and deserve more attention.

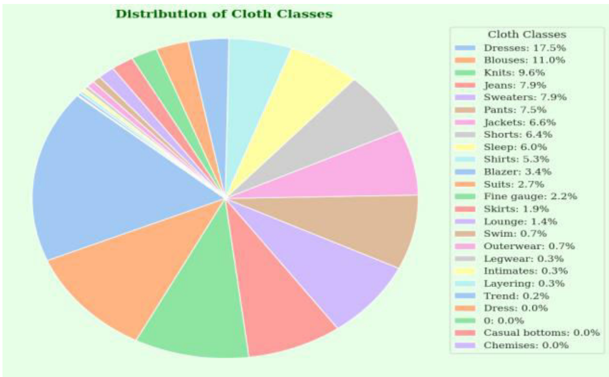

Fig1. The image shows the distribution of different cloth classes in the dataset.

The legend aids in stock and showcasing decisions by illuminating the ways in which various types of clothing contribute to the dataset. Drawbacks of Histogram Rating Conveyance. As shown in Fig 2, The histogram illustrates the recurring conveyance of unfavorable (cons) evaluations in client assessments

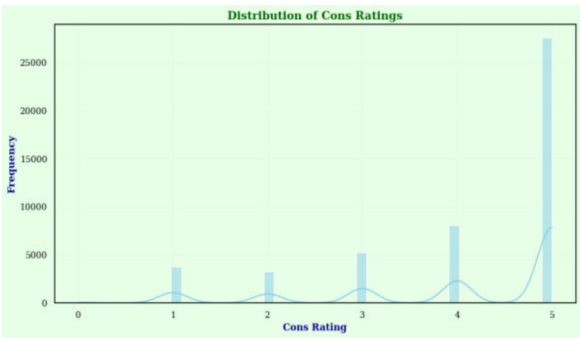

Fig2. The image shows the distribution of cons ratings in the dataset.

Findings The distribution sheds light on how frequently customers give negative reviews. When a KDE (Kernel Density Estimate) curve is present, it sheds light on the data's shape. Most ratings are either very low or around a specific number if the distribution is skewed. Box Plot Construction Ratings vs. Drawbacks (as shown in Fig 3). The purpose of this box plot is to investigate the connection between negative (cons) evaluations and the quality of apparel

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

manufacturing. Results It assists in determining whether subpar construction has a significant role in low ratings. Consumer unhappiness may be exacerbated by a building category's higher median cons ratings. The existence of outliers suggests that there were few exceptional instances in which clients gave the construction a low rating.

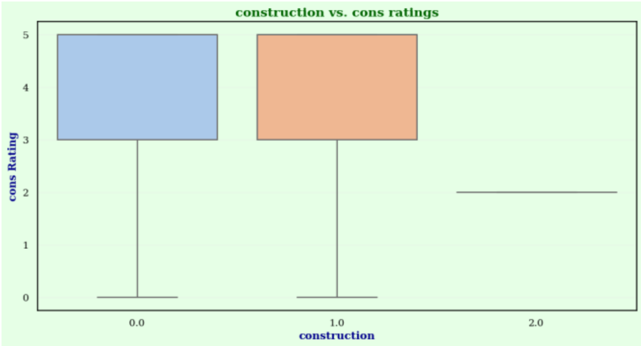

Fig3. The images show cons rating distribution and its relation to construction quality.

Color Distribution(as shown in Fig.4) The distribution of the various colors in the dataset is intended to be shown by this count plot. The results help determine which colors are most and least popular. This information can be used by retailers to offer colors that consumers prefer. The sentiment analysis results may be impacted if a specific color predominates in the sample.

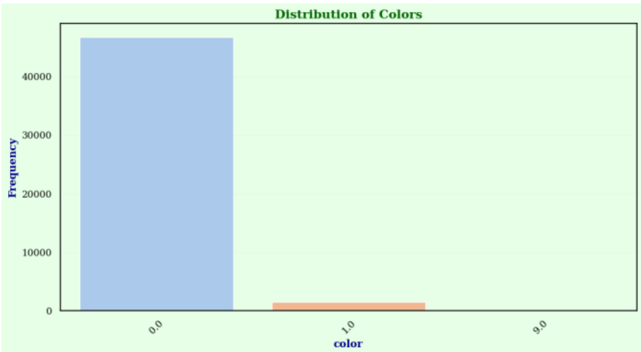

Fig4. The image shows the frequency distribution of colors in the dataset.

A pairplot of numerical variables(as shown in Fig. 5) .This Fig. 5, aims to illustrate the relationships between significant numerical attributes like cons_rating, materials, construction, finishing, and durability. Findings help determine if the correlations between variables are linear or nonlinear. determines potential correlations, such as if higher-quality materials lead to higher evaluations. The Kernel Density Estimate (KDE) provides information on the distribution of specific characteristics on diagonal plots.

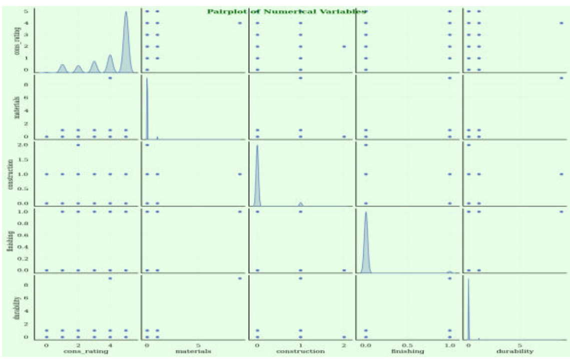

Fig5. The image shows relationships between numerical variables in the dataset.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

Word Cloud of Reviews (as shown in Fig. 6)The terms most frequently used in customer reviews are displayed in this image. Findings The increased frequency of larger terms in the sample indicates common themes in customer feedback. If the majority of the terms are negative, it suggests that there are more consumer complaints. assists businesses in identifying critical problems and improving their products in response to customer input

Fig 6. The image shows a word cloud of frequent terms from customer reviews.

Sentiment Score Distribution(as shown in Fig. 7) . This Fig. 7 shows the distribution of sentiment scores derived from customer reviews. Findings A right-skewed distribution in the dataset indicates more positive opinions. A left-skewed distribution indicates that customers are more inclined to leave negative reviews. A normal distribution suggests a balanced distribution of reviews, both good and negative.

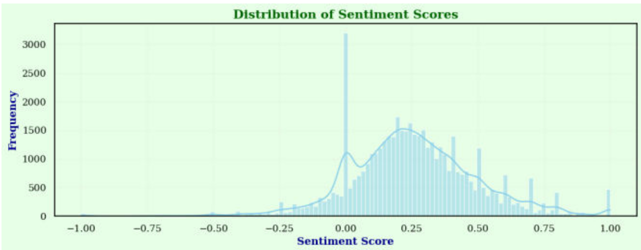

Fig 7. The image shows the distribution of sentiment scores in the dataset.

The connection Map of Heat The purpose of this heatmap is to show (in Fig. 8) how quantifiable attributes, such as sentiment, material quality, construction, and durability, are related to one another. Conclusions High correlations imply that some features have a major impact on evaluations. There is a strong link between durability and material quality, suggesting that longer-lasting items are made from higher-quality materials. Customers' perceptions of a product are significantly influenced by its features if sentiment ratings and those features are associated.

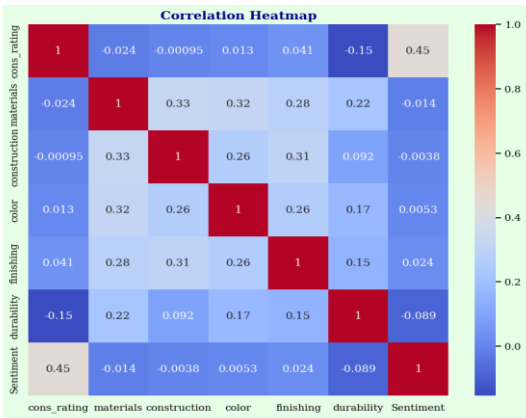

Fig 8. The image shows a correlation heatmap of numerical features in the dataset.

Clustering of Materials vs. Construction (as shown in Fig. 9) This Fig.9 objective is to group products into clusters based on construction ratings and material quality using K-means clustering. Findings The presence of diverse clusters suggests that the construction and material quality of various product types varies significantly. Overlapping clusters imply that these characteristics may not be particularly distinctive in consumer evaluations. Using clustering data, retailers may effectively target different customer categories and categorize items.

Fig 9. The image shows clustering results for materials versus construction.

VI. CONCLUSION

With a high accuracy of 89%, the Ensemble Voting Classifier and Gradient Boosting machine learning models proved to be the most successful in sentiment analysis of reviews for apparel products. Sentiment prediction was greatly impacted by important text-based characteristics such as word frequency, sentiment polarity, and review length. Retailers can use these insights to improve the entire customer experience, automate the study of customer comments, improve product design, and personalize marketing campaigns. Although the present models have good predictive power, sophisticated feature engineering, deep learning techniques like transformers (e.g., BERT), and finetuning hyperparameters could increase accuracy even further. For retail data-driven decisionmaking, the system would be more useful and practical if it included real-time sentiment analysis in addition to explainability tools like SHAP or LIME

REFERENCES

[1] C. Guru and W. Bajnaid, "Prediction of Customer Sentiment Based on Online Reviews Using Machine Learning Algorithms," International Journal of Data Science and Advanced Analytics, vol. 5, no. 5, pp. 272-279, 2023, doi: 10.69511/ijdsaa.v5i5.200.

[2] H. Kim, "Classification and Prediction of Fashion Product Online Review," Korean Journal of Human Ecology, vol. 32, no. 6, pp. 767-782, 2023, doi: 10.5934/kjhe.2023.32.6.767.

[3] A. F. Agarap, "Statistical analysis on E-commerce reviews, with sentiment classification using bidirectional recurrent neural network (RNN)," arXiv preprint arXiv:1805.03687, 2018.

[4] R. J. Sri and P. Ajitha, "Survey of product reviews using sentiment analysis," Indian Journal of Science and Technology, vol. 9, no. 21, pp. 1-4, 2016.

[5] N. Amalia, N. Royanti, I. Indrayanti, and B. Ismanto, "Analisa Sentimen Pelanggan pada Review Belanja Online Berbasis Text Mining Menggunakan Metode K-Means," Journal of Information System Research (JOSH), vol. 4, no. 4, pp. 1441-1447, 2023, doi: 10.47065/josh.v4i4.3781.

[6] J. Panthati, J. Bhaskar, T. K. Ranga, and M. R. Challa, "Sentiment Analysis of Product Reviews using Deep Learning," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, pp. 10.1109/ICACCI.2018.8554551.

[7] K. Mahboob and F. Ali, "Sentiment Analysis of Pharmaceutical Products Evaluation Based on Customer Review Mining," J Comput Sci Syst Biol, vol. 11, pp. 190-194, 2018, doi: 10.4172/jcsb.1000271.

[8] P. Sharma, R. Kumar, and M. Gupta, "Impacts of Customer Feedback for Online-Offline Shopping using Machine Learning," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, pp. 1696-1703, 2021, doi: 10.1109/ICOSEC51865.2021.9591939.

[9] T. U. Haque, N. N. Saber, and F. M. Shah, "Sentiment analysis on large-scale Amazon product reviews," 2018 IEEE International Conference on Innovative Research and Development (ICIRD), Bangkok, Thailand, pp. 1-6, 2018, doi: 10.1109/ICIRD.2018.8376299.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 4, July - August 2025 ||

DOI:10.15680/IJARETY.2025.1204062

[10]C. Fiarni, H. Maharani, and R. Pratama, "Sentiment analysis system for Indonesia online retail shop review using hierarchy Naive Bayes technique," 2016 4th International Conference on Information and Communication Technology (ICoICT), Bandung, Indonesia, pp. 1-6, 2016, doi: 10.1109/ICoICT.2016.7571912.

[11]R. Tao, Y. Luo, and G. Liu, "Sentiment analysis based on the domain dictionary: A case of analyzing online apparel reviews," International Journal of Web Engineering and Technology, vol. 13, no. 4, pp. 380-407, 2018.

[12]O. Bellar, A. Baina, and M. Ballafkih, "Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers," Mathematics, vol. 12, no. 15, p. 2403, 2024, doi: 10.3390/math12152403.

[13]P. S. Ghatora, S. E. Hosseini, S. Pervez, M. J. Iqbal, and N. Shaukat, "Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM," Big Data and Cognitive Computing, vol. 8, no. 12, p. 199, 2024, doi: 10.3390/bdcc8120199.

[14]"Three text sentiment analysis methods—and why ours is better," Bellomy Market Research, accessed March 4, 2025, [Online]. Available: https://www.bellomy.com/blog/three-text sentiment-analysis-methods-and-why-ours-better.

[15]"Consumer Review of Clothing Product" Available: https://www.kaggle.com/datasets/jocelyndumlao/consumer review-of-clothing-product

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com