Aspect Based Sentiment Analysis of Beauty Product Review using Machine Learning Algorithm

Kajal D. Dolar[1], Ruapl Shilu[2], and Tosal M. Bhalodia[3]
[1]*Lecturer, Atmiya University*
[2,3]*Assistant Professor, Atmiya University*

***Abstract*—The proliferation of online beauty product reviews has given rise to the need for more nuanced sentiment analysis approaches. This study focuses on aspect-based sentiment analysis (ABSA) of skincare product reviews using the XGBoost machine learning algorithm. Utilizing the Sephora Skincare Reviews dataset, the analysis targets the 10 most-reviewed products to extract key aspects such as packaging, price, scent, texture, and effectiveness. Each aspect is then classified according to sentiment polarity: positive, negative, or neutral. Unlike traditional sentiment analysis, ABSA provides fine-grained insights by associating sentiments with specific product attributes, enabling better decision-making for both consumers and businesses. XGBoost was selected for its efficiency and high performance in handling structured data and imbalanced class distributions. The system pipeline includes data preprocessing, aspect extraction using rule-based and frequency-based methods, sentiment labeling, and classification. Results indicate that XGBoost, when tuned with optimal hyperparameters and trained on selected aspects, achieves strong performance across multiple sentiment classes. Accuracy achived 83%.This approach demonstrates the potential of using advanced machine learning models for detailed opinion mining in e-commerce domains, especially for consumer-centric industries such as beauty and skincare. Future work will explore deep learning approaches and hybrid models for improved accuracy in aspect detection and sentiment classification.**

***Index Terms*—Aspect-Based Sentiment Analysis (ABSA), Natural Language Processing (NLP), Rule-Based Aspect Extraction, XGBoost.**

I. INTRODUCTION

In recent years, the beauty and skincare industry has witnessed exponential growth in online consumer engagement, especially through product reviews. These reviews offer valuable insights into consumer satisfaction and perception, particularly when analyzed through the lens of sentiment analysis. Traditional sentiment analysis typically classifies entire reviews as positive, negative, or neutral. However, such coarse analysis overlooks the multifaceted nature of product feedback. Aspect-Based Sentiment Analysis (ABSA) addresses this limitation by associating sentiments with specific product features such as scent, packaging, or price, product and effectiveness.

This study investigates aspect-based sentiment analysis of skincare product reviews, focusing on the 10 most-reviewed items in the Sephora Skincare Reviews dataset. The goal is to identify key aspects mentioned by users and evaluate the sentiment polarity associated with each. Unlike the base paper, which used SVM and Word2Vec for aspect-level sentiment classification on the Female Daily dataset, this research leverages the XGBoost algorithm due to its effectiveness in handling structured data and superior performance in classification tasks.

By applying ABSA with XGBoost, this work aims to provide more actionable insights for consumers, marketers, and product developers in the beauty industry. The analysis highlights the importance of each aspect in shaping overall sentiment and guides future improvements in product design and recommendation systems.

II. LITERATURE REVIEW

Aspect-Based Sentiment Analysis (ABSA) provides more fine-grained sentiment classification by connecting opinions to specific product attributes. Various studies have implemented ABSA using machine learning and natural language processing techniques across domains such as beauty, movies, electronics, and music.

Clara et al. [1] conducted ABSA on beauty product reviews using Random Forest with TF-IDF and n-gram techniques, focusing on aspects such as toner and serum. Their model achieved a notable accuracy of 90.48%, demonstrating the effectiveness of ensemble learning methods in capturing aspect-level sentiments.

Fazri et al. [2] applied Support Vector Machine (SVM) with Word2Vec embeddings to classify sentiments on beauty product aspects like aroma and price. They highlighted the importance of preprocessing techniques such as stemming and dimensional tuning, achieving an F1-score of 81.93%.

Kothalawala and Thelijjagoda [3] introduced a rule-based ABSA system for hair care reviews and reached an overall accuracy of 85%, showcasing the utility of structured pipelines in sentiment analysis.

In the movie domain, Wong and Joseph [4] used Logistic Regression and Decision Tree for ABSA, concluding that Decision Tree performed best for aspect identification (98% accuracy), while Logistic Regression was better for sentiment polarity classification (93%).

Kumar et al. [5] explored ABSA on Amazon musical instrument reviews using eight machine learning models. SVM achieved the highest accuracy of 97.7%, affirming its reliability in text classification.

Chauhan et al. [6] classified Amazon product reviews using SVM and Naïve Bayes, employing feature extraction techniques such as TF-IDF. Although the study focused on sentiment classification rather than ABSA, it established strong model baselines.

Horesh Kumar et al. [7] analyzed Amazon electronics reviews using models including SVM, Logistic Regression, and Random Forest. Their work highlighted the comparative performance of classifiers in sentiment tasks, where SVM again stood out.

Pratama et al. [8] focused on binary sentiment classification of beauty product reviews using SVM, attaining 80.06% accuracy and validating its application in e-commerce review analysis.

Kiran Kumar et al. (2023) explored various machine learning (ML) techniques for sentiment classification of Amazon and Flipkart product reviews, focusing on a comparative study using algorithms such as Naive Bayes, SVM, and Logistic Regression. Their work highlights the superior performance of SVM with an accuracy of 71.8% and emphasizes the need for robust preprocessing and feature extraction to enhance model performance.

In another study, Automated Sentiment Classification of Amazon Product Reviews Using LSTM and Bi-LSTM, the authors employed deep learning models—specifically LSTM and BiLSTM—demonstrating that BiLSTM outperforms traditional LSTM due to its ability to capture contextual dependencies in both forward and backward directions. This deep learning-based approach significantly improved the classification accuracy compared to conventional machine learning models.

The study by Nasrin et al. (2022), titled Bangla E-Commerce Sentiment Analysis Optimization Using Tokenization and TF-IDF, adapted sentiment analysis methods for Bangla language e-commerce reviews. It used TF-IDF and tokenization techniques to optimize the sentiment classification process, illustrating the importance of language-specific preprocessing techniques in multilingual sentiment analysis.

A comprehensive survey by Rajan et al. (2021) in Aspect-Based Sentiment Analysis: An Extensive Study of Techniques, Challenges and Applications systematically reviewed various techniques in aspect-based sentiment analysis (ABSA). They categorized existing methods into lexicon-based, supervised learning, and deep learning approaches. The paper also addressed challenges such as aspect extraction, sentiment polarity detection, and multilingual sentiment analysis, indicating that ABSA holds significant potential in fine-grained opinion mining.

Despite these developments, XGBoost remains underutilized in ABSA for beauty reviews. Given its strengths in handling structured and imbalanced data, this study investigates XGBoost as a powerful alternative for multi-aspect sentiment classification in beauty product reviews.B. Preprocessing: This is a critical step for cleaning raw text data. The right-hand side of the figure breaks this down into sub-steps, Removing Punctuations, Numbers & Symbols, Converting to Lowercase, Tokenization, Removing Stop words, and Lemmatizing.

III. PROPOSED WORK

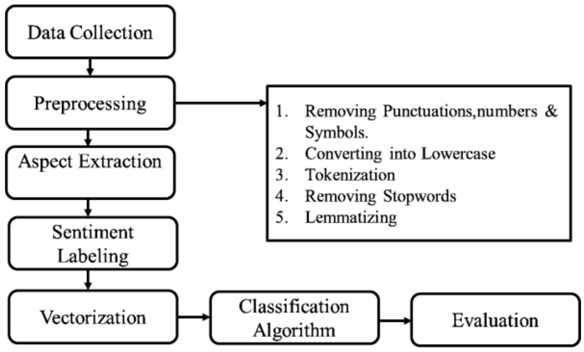

Figure: 1 A typical sentiment analysis model.

The proposed system for Aspect-Based Sentiment Analysis (ABSA) of skincare product reviews involves a multi-stage pipeline, encompassing data preprocessing, aspect extraction, sentiment classification, feature vectorization, and classification using XGBoost. The dataset used is a filtered version of the Sephora skincare product reviews, focusing on the top 10 most reviewed products.

1. Data Collection and Preprocessing:
The raw review data was initially cleaned by removing irrelevant columns and handling missing values. Text preprocessing steps included:
Converting text to lowercase, Removing numbers, punctuation, and special characters using regular expressions, Tokenizing the text using NLTK, Removing English stop words, Lemmatizing each token using WordNetLemmatizer. The cleaned tokens were then recombined into processed review text for downstream tasks.

2. Aspect Extraction
To detect relevant product aspects, a keyword-based matching approach was used. Five core aspects were selected based on frequency and relevance in reviews: Price, Product effectiveness, Moisturization, Smell/Aroma, and Packaging. Each review was scanned for predefined keywords representing these aspects. A binary flag was assigned to each aspect column to indicate presence.

3. Sentiment Labeling
Aspect-specific sentiment was determined using a hybrid of two rule-based sentiment analyzers:
VADER (Valence Aware Dictionary and sEntiment Reasoner): Ideal for analyzing short and informal review texts. TextBlob: Used for sentiment polarity scoring. The final sentiment polarity for each detected aspect was derived by averaging VADER's compound score and TextBlob's polarity score. Labels were assigned as:
Positive if score ≥ 0.05
Negative if score ≤ -0.05
Neutral otherwise

4. Feature Vectorization
The processed review text was transformed using the TF-IDF (Term Frequency-Inverse Document Frequency) vectorization method. This converted each review into a numerical feature vector capturing the importance of each word relative to the corpus. To address potential class imbalance in dataset (e.g., more positive reviews than negative), apply the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE generates synthetic examples of the minority class to balance the dataset, which can improve the performance of classifiers. SMOTE is a well-established method for dealing with imbalanced datasets in machine learning.

5. Classification Using XGBoost
The XGBoost classifier, known for its efficiency and performance in classification tasks. To optimize the model, use RandomizedSearchCV for hyper parameter tuning, which searches over specified parameter values to find the best combination. XGBoost is a scalable and accurate implementation of gradient boosting machines, widely used in machine learning competitions and real-world applications.

6. Evaluation
Performance metrics such as accuracy, precision, recall, and F1-score were computed for each aspect-based classifier to evaluate the system's effectiveness in identifying sentiment across different product features.

Table I CONFUSION MATRIX

Class Prediction		
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

According to the, some of the terms used are as follows:
True Positive (TP): Positive data is accurately Detected.
True Negative (TN): Negative data is accurately Detected.
False Positive (FP): Negative data is detected as Positive data.
False Negative.

Table II. RESULT ANALYSIS

	Predicted: Negative	Predicted: Neutral	Predicted: Positive	Total Actual
Actual: Negative	167	46	101	314
Actual: Neutral	230	5676	393	6299
Actual: Positive	919	966	7283	9168
Total Predicted	1316	6688	7777	15781

Table III. CLASSIFICATION REPORT

Label	Precision	Recall	F1-Score	Support
Negative	0.13	0.53	0.2	314
Neutral	0.85	0.9	0.87	6299
Positive	0.94	0.79	0.86	9168
Accuracy			0.83	15781
Macro Avg	0.64	0.74	0.65	15781
Weighted Avg	0.89	0.83	0.85	15781

The XGBoost classifier, optimized with Randomized SearchCV and trained on a SMOTE-balanced dataset, demonstrates strong performance in classifying aspect-based sentiment, particularly for the Positive and Neutral classes. The model correctly classified 7283 Positive, 5676 Neutral, and 167 Negative reviews. However, there is a significant imbalance in the classification accuracy across classes. The Negative class is underrepresented and was often misclassified as Positive (101 instances) or Neutral (46 instances), likely due to inherent class imbalance in the original dataset and subtle linguistic features that differentiate negative sentiments. Despite using SMOTE to synthetically balance the training data, the model still struggles with capturing the Negative sentiment effectively. On the other hand, the Neutral class exhibits the best overall accuracy, with only a small portion misclassified, indicating that the model captures the neutral tone in reviews well. The Positive class, while having the highest number of correct predictions, shows some confusion with Neutral and Negative reviews, suggesting that borderline sentiments can cause overlap in feature representation. Overall, the model performs well with macro-averaged F1-scores likely skewed by the imbalance, and future improvements could focus on more refined feature engineering or incorporating transformer-based embeddings to enhance semantic understanding.

V.CONCLUSION AND FUTURE WORK

This research illustrates that XGBoost, optimized through RandomizedSearchCV and trained on a dataset balanced with SMOTE, is capable of effectively conducting aspect-based sentiment analysis on beauty product reviews. The model shows strong performance in identifying Positive and Neutral sentiments, reflecting its ability to generalize well in recognizing commonly expressed opinions. Nonetheless, its performance in classifying Negative sentiments is comparatively weak, indicating challenges in managing minority class data or subtle negative expressions, even with synthetic oversampling. While TF-IDF features provide a lightweight representation, they may not fully capture the contextual and semantic subtleties inherent in natural language reviews.

Although SMOTE contributed to balancing the dataset, employing more sophisticated methods such as ensemble-based resampling, cost-sensitive learning, or focal loss could enhance the classification of less represented categories like negative sentiments. Additionally, the aspect extraction process could be improved by utilizing syntactic parsing or neural models that more effectively identify pertinent terms and phrases, thereby increasing the accuracy of sentiment classification for each aspect.

REFERENCES

[1] A. Y. Clara, A. Adiwijaya, and M. D. P. Purbolaksono, "Aspect Based Sentiment Analysis on Beauty Product Review Using Random Forest," J. Data Sci. Appl., vol. 3, no. 2, pp. 67–77, 2020.

[2] A. Fazri, M. D. P. Purbolaksono, and W. Astuti, "Aspect-Based Sentiment Analysis of Beauty Product Reviews on Female Daily Website Using Support Vector Machine and Word2Vec," in Proc. 2024 Int. Conf. on Data Science and Its Applications (ICoDSA), IEEE, 2024.

[3] M. Kothalawala and S. Thelijjagoda, "Aspect-based Sentiment Analysis on Hair Care Product Reviews," in Smart Computing and Systems Engineering (SCSE), 2020.

[4] B. Wong and M. H. Joseph, "Aspect-Based Sentiment Analysis on Movie Reviews," in 2023 15th Int. Conf. on Developments in eSystems Engineering (DeSE), IEEE, 2023.

[5] A. Kumar, T. Jain, P. Tiwari, and R. Sharma, "Opinion Mining on Amazon Musical Product Reviews using Supervised Machine Learning Techniques," in 2023 11th Int. Conf. on Internet of Everything (IEMECON), IEEE, 2023.

[6] G. Chauhan, A. Sharma, and N. Dwivedi, "Amazon Product Reviews Sentimental Analysis using Machine Learning," in 2024 Int. Conf. on Computing, Power, and Communication Technologies (IC2PCT), IEEE, 2024.

[7] H. Kumar et al., "Sentiment Analysis on Amazon Electronics Product Reviews using Machine Learning Techniques," in 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), IEEE, 2023.

[8] M. R. Pratama et al., "Sentiment Analysis of Beauty Product E-Commerce Using Support Vector Machine Method," Jurnal RESTI, vol. 6, no. 2, pp. 269–274, 2022.

[9] Kiran Kumar, S. K. Singh, A. Kumar, A. S. Singh. (2023). Sentimental Analysis of Customer’s Reviews Using Machine Learning. 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), IEEE.

[10] Authors Unknown. (n.d.). Automated Sentiment Classification of Amazon Product Reviews Using LSTM and Bi-LSTM.

[11] Nasrin, F., & Others. (2022). Bangla E-Commerce Sentiment Analysis Optimization Using Tokenization and TF-IDF.

[12] Rajan, R., Balamurali, A. R., & Nair, B. B. (2021). Aspect-Based Sentiment Analysis: An Extensive Study of Techniques, Challenges and Applications.