INFORMATICS AND DIGITAL EXPERT (INDEX) - VOL. 7 NO. 1 (2025) 77-85
Terbit *online* pada laman web jurnal : https://e-journal.unper.ac.id/index.php/informatics
| ISSN (Print) 2775-2208 | ISSN (Online) 2715-0453 |

Beauty Product Recommendation from Customer Reviews Based on Multinomial Naïve Bayes Algorithm

Claretty Therence [1], Jimmy Tjen[2]

[1]*Digital Business, Faculty of Information Technology, Universitas Widya Dharma Pontianak, Pontianak, Indonesia*
[2] *Informatics, Faculty of Information Technology, Universitas Widya Dharma Pontianak, Pontianak, Indonesia*
*e-mail: 22430149@widyadharma.ac.id[*1], jimmy.tjen@mathmods.eu[2]*

ARTICLE INFORMATION

Sejarah Artikel:
Diterima Redaksi : 30 April 2025
Revisi Akhir : 20 Mei 2025
Diterbitkan *Online* : 31 Mei 2025

Kata Kunci:
Customer Sentiment, Concealer, Sentiment Analysis, Naïve Bayes Classification, Product Development.

Korespondensi :
Telepon / Hp : +6285251522500
E-mail : **22430149@widyadharma.ac.id**

A B S T R A C T

The cosmetic industry's increasing dependence on online platforms makes understanding customer sentiment essential for brand success. The study addresses the challenge of understanding customer sentiment regarding concealer brands on the Shopee platform, a critical aspect for brand managers and consumers alike. The problem stems from the vast amount of reviews that can be overwhelming for stakeholders looking to extract actionable insights. To solve this, we applied a Naïve Bayes classification approach with a Bag of Words (BoW) model to analyze a dataset containing 3,920 customer reviews. This dataset was divided into 1,120 training samples and 280 testing samples of each of 10 brands, following an 80:20 ratio. The analysis yields recurring positive and negative sentiment themes, achieving test accuracy of 87.95% and a training accuracy of 94.31%. Finding reveal key consumer preferences and lead to specific product recommendations, such as Mad For Makeup and Luxcrime for high coverage, Guele for lightweight formulas. Additionally, tailored marketing strategies like enhancing packaging and engaging with consumers through social media are suggested. This research provides actionable insights for brand managers, contributing to sentiment analysis literature in the cosmetics sector.

1 INTRODUCTION

The rapid growth of e-commerce has reshaped consumer interactions particularly in the cosmetics industry [1]. Platforms such as Shopee in Indonesia have enabled users to share detailed product reviews, significantly influencing purchase decisions and brand reputation [2]. Understanding buyer perceptions through reviews is crucial for companies seeking actionable insights to drive product development and strengthen their competitive edge [3], [4]. Sentiment analysis—an application of text mining and Natural Language Processing (NLP) [5], [6]—plays a crucial role in interpreting reviews to uncover consumer preferences and market trends. However, traditional sentiment analysis approaches face challenges in processing massive and nuanced content.

Machine learning algorithms, such as Naïve Bayes (NB), offer a robust solution by combining high accuracy with simplicity, making them especially effective in resource-constrained environments [7]. This makes NB well-suited for extracting actionable insights from product reviews to inform cosmetic brand strategies. Naïve Bayes is a probabilistic classifier based on Bayes' theorem [8], which utilizes prior probability distributions to estimate the posterior probability of a sample belonging to a specific class [9]. In sentiment analysis, the Multinomial Naïve Bayes (MNB) variant is particularly suitable, as it considers word frequency rather than treating words as mere binary occurrences [10]. MNB simplifies computation and effectively handles large datasets by assuming conditional independence among features [11]. Its fast classification time and competitive accuracy make it particularly well-suited for processing high-volume textual data [12].

Numerous studies have applied MNB in sentiment analysis. For example, [13] utilized 50,000 IMDb movie reviews using Bag of Words (BoW) and TF-IDF, achieving 89% accuracy. In [14] The sentiment classification of tweets about government policies in English and Filipino achieved an accuracy of 81.77%. Study [15] analyzed sentiment in Lazada app reviews using TF-IDF and ensemble learning, finding that MNB outperformed other models with 89.1% accuracy. In [16], sentiment analysis of beauty product reviews combined Naïve Bayes with Term Frequency-Inverse Document Frequency (TF-IDF) and Chi-Square for feature selection, achieving 80.18% accuracy. Lastly, [17] applied sentiment analysis to assist customers in selecting natural skincare products, achieving nearly 80% accuracy.

This study extends our previous work [18] on sentiment analysis of concealer products by expanding the dataset and adopting a more detailed, brand-specific approach. The prior study analyzed five anonymized concealer brands, achieving an accuracy of 73.75%. However, it did not differentiate between positive and negative word frequencies, incorporate word cloud visualizations, or assess the distinctive strengths of each brand. Its analysis was limited to identifying frequently mentioned product attributes without providing detailed brand recommendations based on customer needs or actionable marketing recommendations.

In contrast, the present study investigates ten explicitly named concealer brands available on Shopee Indonesia, utilizing a dataset comprising 3,920 product reviews. Of these, 2,800 reviews—280 per brand—are allocated for testing. Meanwhile, the remaining 1,120 reviews serve as the training dataset for building sentiment analysis models across all ten brands. This 80:20 split between training and testing datasets ensures balanced and reliable performance evaluation. The study analyzes word frequencies by sentiment polarity to reveal each brand's strengths and weaknesses, providing strategic insights to support data-driven product development and brand positioning.

This study addresses three main questions: (1) How do customer reviews on Shopee reflect sentiment toward specific concealer brands? (2) What recurring positive and negative themes can be identified through word frequency and sentiment polarity analysis? (3) How can these insights inform strategic recommendations for product development and marketing, while also providing customers with product recommendations based on their needs? The study contributes to the literature on sentiment analysis in the cosmetic sector and offers practical implications for consumers and brand managers.

This paper is structured as follows: the introduction outlines the study's background and novelty; the methodology describes the dataset and preprocessing procedure; the results and discussion present the findings; and the conclusion summarizes the key insights, highlights contributions, and suggests future research directions.

2 METHODOLOGY

This research employs a text classification approach with the BoW method to analyze sentiment across ten concealer brands. The process involves preprocessing, feature extraction, model evaluation, and insights generation, as shown in Figure 1.

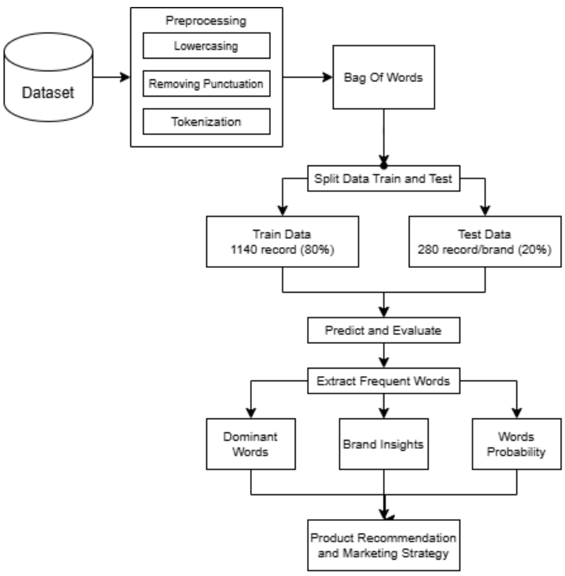

Figure 1. Research Workflow

2.1. Dataset

The dataset for this study consists of 3,920 text reviews collected from Shopee, manually labeled for sentiment brand categories. As shown in Figure 2, the data were split into two subsets using an 80:20 ratio: 1.120 entries for training and 280 reviews for testing, assigned to each of the ten brands: BLP, Luxcrime Guele, Esqa, L'Oréal, Mad For Makeup, Make Over, Maybelline, Rose All Day, and Secondate. This structure ensures consistency in training while allowing for brand-specific performance analyses and nuanced sentiment insights.

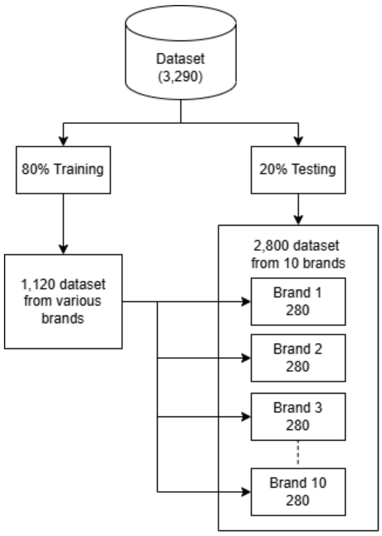

Figure 2. Dataset Partitioning and Allocation for Multi-Brand Sentiment Analysis

2.2. Preprocessing

Before modeling, the test data underwent a series of preprocessing steps designed to improve data quality and enhance model accuracy. These steps included thorough cleaning and preparation to ensure optimal performance of the sentiment classification model [19]. As shown in Table 1 and Table 2:

Table 1. Preprocessing Steps

Step	Description
Lowercasing	All text was converted to lowercase to ensure consistency.
Punctuation Removal	Special characters and punctuation marks were eliminated to reduce noise during tokenization.
Tokenization	Sentences were split into individual tokens (words) to enable detail.

Table 2. Example of Preprocessing Steps

Step	Example
Original Text	"Concealernya bagus bgt!!? Shadenya pas 😍"
Lowercasing	"concealernya bagus bgt!!? shadenya pas 😍"
Punctuation Removal	"concealernya bagus bgt shadenya pas"
Tokenization	['concealernya', 'bagus', 'bgt', 'shadenya pas']

2.1. Feature Extraction (Bag of Words)

After preprocessing, the cleaned data were converted into numerical representations using the BoW model. This approach converts each review into a vector that reflects the frequency of its words. [20], enabling the machine learning model to identify patterns. For instance, the review 'concealernya bagus bgt shadenya pas' is processed into the tokens ['concealernya', 'bagus', 'bgt', 'shadenya pas'] and converted into a word-frequency dictionary: ['concealernya':1, 'bagus':1, 'bgt': 1, 'shadenya':1, pas':1]. Across multiple reviews, a word dictionary is created, and each review is represented as a vector indicating word frequency, as shown in Table 3.

Tabel 3. Example of Bag of Words (BoW) process

Review 1	"Concealernya bagus bgt shadenya pas"
Review 2	"Coverage concealernya oke, shadenya pas"

Word	Review 1	Review 2
bagus	1	0
bgt	1	0
concealernya	1	1
coverage	0	1
oke	0	1
pas	1	1
shadenya	1	1

2.4. Naive Bayes Classifier

After vectorization, the training data were used to train a Naive Bayes classifier for sentiment analysis. This model was then applied to each brand's testing dataset to evaluate performance across various review contexts. In maintaining uniformity, the training dataset remained the same for all brand evaluations. Performance was assessed using standard confusion matrices, as shown in Table 4.

Tabel 4. Confusion Matrix

	Predicted Negative	**Predicted Positive**
Actual Negative	True Negative (TN)	False Positive (FP)
Actual Positive	False Negative (FN)	True Positive (TP)

We trained the sentiment classification model using a unified training dataset and tested it on data from each brand to ensure consistent evaluation. To measure the effectiveness, we focused on accuracy, defined in Equation (1):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

Where:

1. True Positive (TP) represents the number of positive sentiment instances correctly classified as positive.
2. True Negative (TN) refers to negative sentiments that were accurately classified as negative.
3. False Positive (FP) refers to negative sentiments correctly classified as positive, and
4. False Negative (FN) corresponds to positive sentiments mistakenly predicted as negative.

These four components form the basis of the accuracy calculation used to evaluate model performance [21]. We also analyzed correctly and incorrectly classified samples to gain insights into the model's strengths and limitations, identifying areas for improvement in our sentiment analysis approach.

3. RESULTS AND DISCUSSIONS

This section presents the results of the sentiment classification model, followed by discussions on its performance across various evaluation scenarios.

3.1. Model Evaluation on Training Dataset

The sentiment classification model was initially evaluated using the training dataset of 1,120 reviews, split into 80% training (896 records) and 20% internal testing (224 records). This dataset comprised 632 positive and 488 negative reviews, as shown in Figure 3. The model achieved an accuracy of 94.31% on the training data and 87.95% on the internal test set, as presented in Table 5.

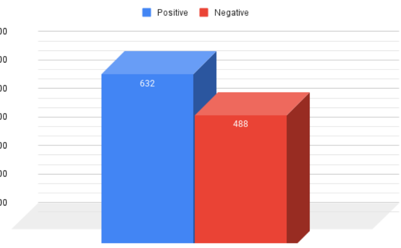

Figure 3. Distribution of Positive and Negative Reviews in the Training Dataset

Table 5. Training Data Accuracy

Evaluation	Value
Test Accuracy	87.95%
Train Accuracy	94.31%

The confusion matrix in Figure 4 indicates that out of 139 positive reviews, the model accurately identifies 132 as positive, while out of 75 negative reviews, it correctly classifies 62 as negative. These results reflect the model's strong ability to learn sentiment patterns from the training data, minimizing misclassifications.

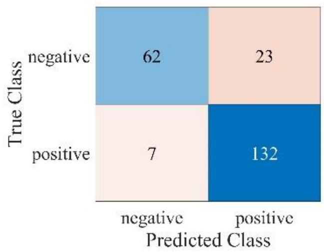

Figure 4. Training Confusion Matrix

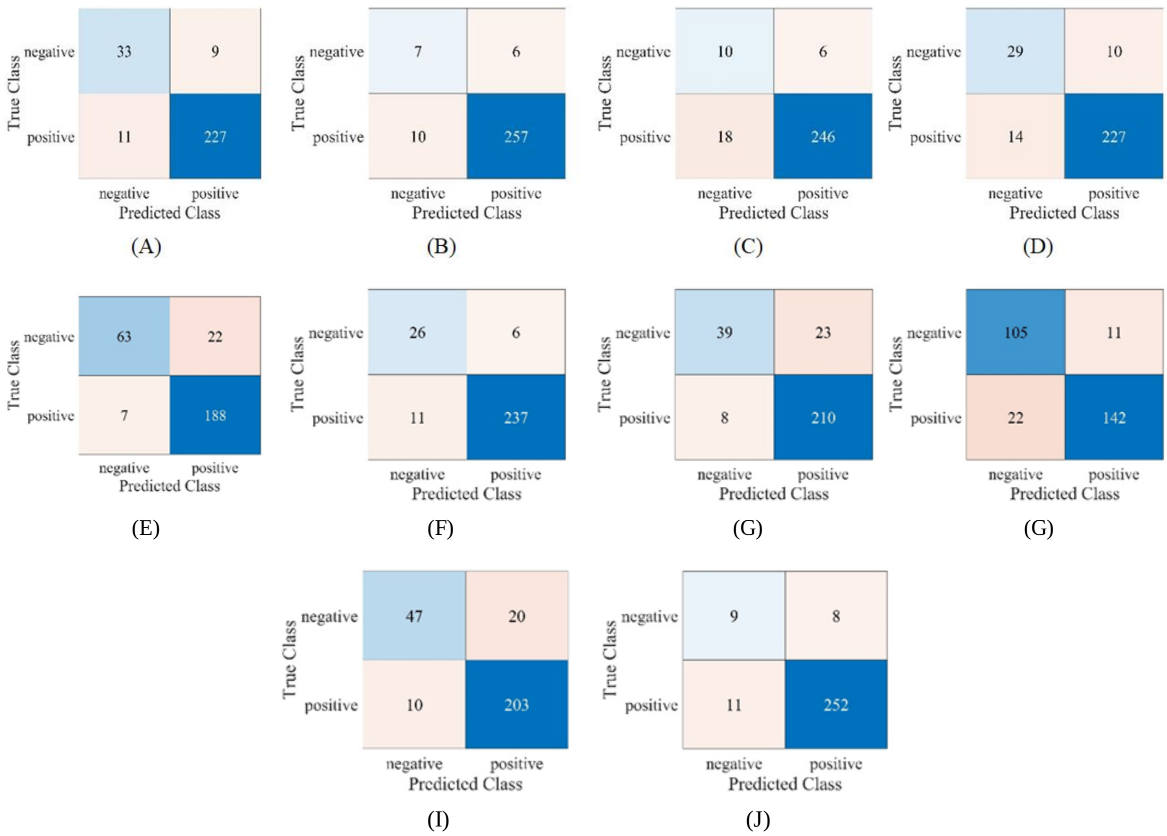

Figure 5. Confusion Matrix across brands:
(a) BLP, (b) Luxcrime, (c) Guele, (d) Esqa, (e) L'Oréal, (f) Mad For Makeup, (g) Make Over, (h) Maybeline, (i) Rose All Day, and (j) Secondate

3.2. Model Evaluation on Brand-Specific Sets

After training, the trained model was tested on external data from ten concealer brands, each contributing 280 reviews, totaling 2,800 brand-specific test samples. This phase evaluated the model's ability to generalize across varying brand contexts. As shown in Table 6, classification accuracy ranged from 88.21% for Maybelline to 94.28% for Luxcrime. All brands exceeded the 88% accuracy threshold, demonstrating consistent and reliable performance. These findings indicate that the model generalizes effectively to unseen brand-specific inputs, reinforcing its suitability for multi-brand sentiment analysis.

Figure 5 presents confusion matrices for the 10 brands, providing a more detailed view of how well the model distinguishes between positive and negative sentiments. For instance, in Figure 5a (BLP) and 5b (Luxcrime), the model shows strong performance in correctly identifying positive reviews, as well as negative reviews, while maintaining relatively low misclassification. This reinforces the model's effectiveness in real sentiment classification tasks, especially for brands with high review volume, to gain insights about their products' performance in the customer's eye. Figure 6 shows the contribution of positive and negative reviews to each brand, showing that Luxcrime has the highest positive probability, followed by Guele and Secondate.

Table 6. Accuracy of Each Brand

Brand	Accuracy
BLP	92.85%
Luxcrime	94.28%
Guele	91.42%
Esqa	91.42%
L'Oréal	89.64%
Mad For Makeup	93.92%
Make Over	88.92%
Maybeline	88.21%
Rose All Day	89.28%
Secondate	93.21%

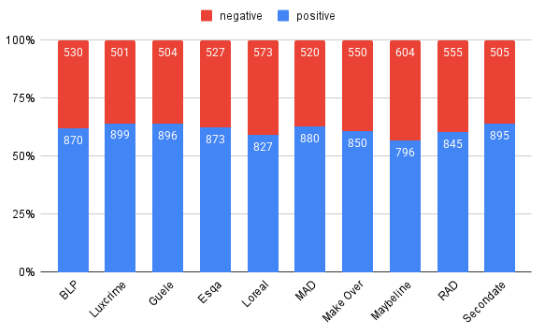

Figure 6. Sentiment Polarity Across Brands

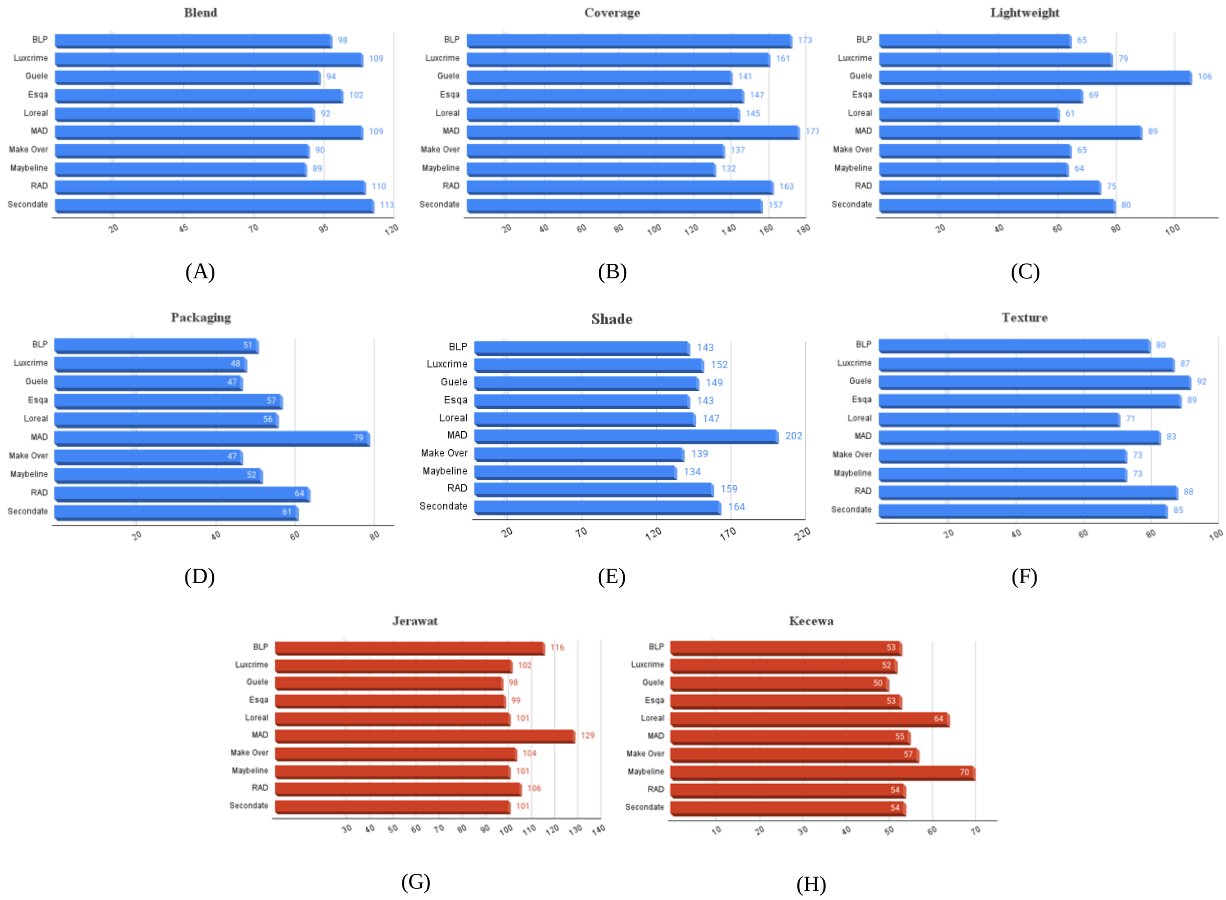

Figure 7. Word Frequency Analysis Across Brands
(a) Blend, (b) Coverage, (c) Lightweight (d) Packaging, (e) Shade, (f) Texture, (g) Acne or "Jerawat", and (h) Disappointment or "Kecewa"

3.3. Word Frequency Analysis

Analyzing word frequency in customer reviews provided deeper insights into how each brand is perceived. This approach identified commonly mentioned terms, offering clues into which product attributes (e.g., coverage, blend, shade, texture) matter most to customers.

In Figure 7, the distribution of keyword mentions illustrates how each brand performs across the most frequently discussed concealer attributes. This visual representation complements the frequency table by offering a clearer comparison of brand performance per attribute, making it easier to identify standout qualities and areas that require improvement.

The frequency data reflects which product attributes are most frequently mentioned by customers while also indicating how well each brand meets those expectations. A higher frequency of certain keywords in a brand's review dataset suggests that the attribute, whether coverage,

Table 7. Frequency Interpretation of Customer Review

Word	Interpretation of Frequency
Blend	Customers care about the ease of application. High frequency = users often discuss how easily it blends.
Coverage	High concern about how well imperfections are covered. High frequency = brand known for covering power.
Packaging	Packaging matters aesthetically/functionally. High = customers care about designs and practicality.
Shade	Customers are sensitive to shade variety/accuracy. High = expectations or attention to match tones.
Texture or "Tekstur"	People care about the feel on the skin. High = users talk about it being light, creamy, or cakey.
Lightweight or "Ringan"	People want light-feel concealer. High = interest in comfort and breathable formula
Acne or "Jerawat"	Indicated user concern with acne-safe formula. High = many are looking for non-comedogenic products.
Disappointment or "Kecewa"	Shows dissatisfaction. High = more users had a bad experience with the product.

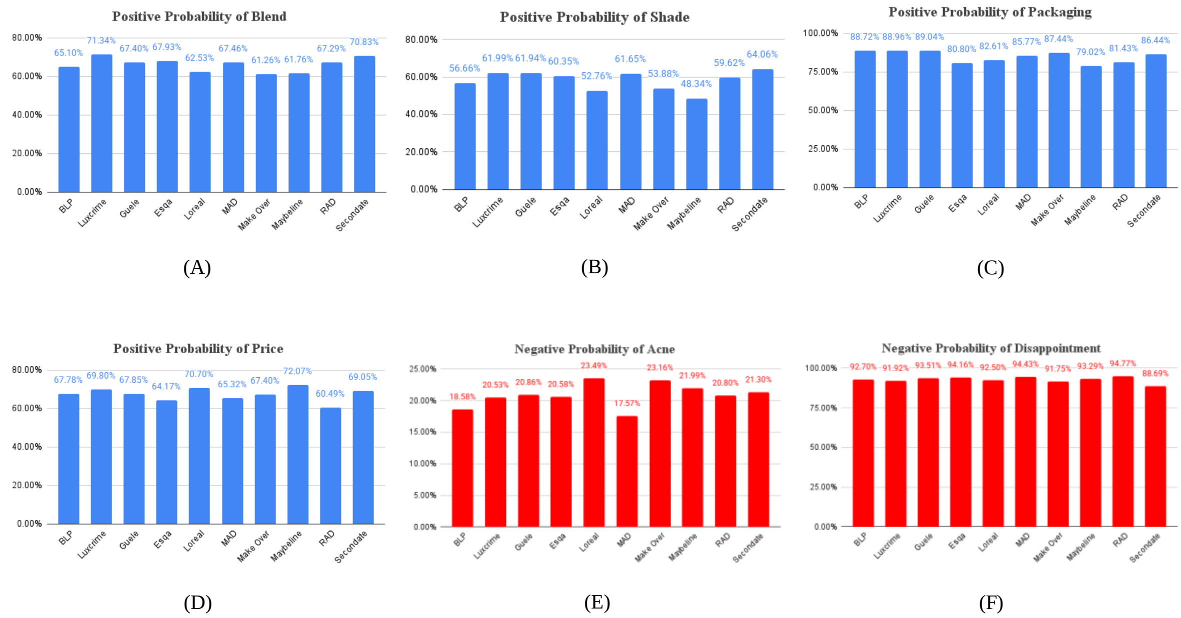

Figure 8. Word Positive Probability Across Brands
(a) Shade, (b) Shade, (c) Packaging, (d) Price, (e) Acne, (f) Disappointment

smoothness in blending, shade variety, or others, stands out in customers' minds. To better understand each brand's positioning and perceived strengths, Table 7 shows which brands lead in each key area as indicated by the frequency results.

While frequency analysis reveals commonly occurring terms across brands, it may need more contextual nuances. To address this, the subsequent section applies probability analysis to uncover deeper insights, particularly for words with dual meanings that cannot be distinctly classified by frequency alone

3.4 Word Probability

To better understand customer needs for concealer products, we present word probability to provide more detailed insights. Figure 8 illustrates the positive probability of "Blend", "Shade", and "Packaging", as well as the negative probability of "Acne" and "Disappointment". This probability analysis offers deeper insights into the potential for double meaning in certain words, extending beyond what frequency alone can reveal. For instance, the word "Acne" is mentioned most frequently in the brand Mad For Makeup, and it has the lowest negative probability. This suggests that customers aren't worried about acne when considering this brand; instead, they perceive Mad For Makeup as having a formula that is friendly for acne-prone skin. This highlights how word probability helps uncover sentiment nuances that may be overlooked when relying solely on frequency.

3.5 Brand Insights

Combining word frequency and word probability analyses reveals each brand's key traits, including standout strengths and areas needing improvement, visualized in Figure 7 and Figure 8, summarized in Table 8.

Table 8. Brand Features and Suggestion for Improvement

Brand	Best Features	Suggested Improvement
BLP	Coverage (2nd highest	Packaging, Shade
Luxcrime	Texture, Blend	Packaging
Guele	Texture, Lightweight, Packaging	Shade,
Esqa	Texture	Packaging
L'Oréal	Moderate for Packaging	Shade, Disappointment
Mad For Makeup (MAD)	Shade, Coverage, Packaging, Acne-friendly	Texture
Make Over	Packaging (4th highest probability)	Disappointment
Maybelline	Price	Highest in Disappointment
Rose All Day (RAD)	Blend, Packaging, Shade	Texture
Secondate	Blend, Shade, Texture	Coverage

3.6 Brand Recommendation

Table 9. Brand Recommendation Based on Customer Needs

Customer Needs	Brand Recommendations	Reason	
		Frequency	Probability
High Coverage for blemishes, redness	MAD, BLP, Luxcrime	MAD = 1^{st} coverage frequency (177) BLP = 2^{nd} coverage frequency (173) Luxcrime = 4^{th}coverage frequency (161)	MAD = 4^{th} coverage positivity (74.81%) BLP = 2^{nd} coverage positivity (76.08%) Luxcrime = 1^{st} coverage positivity (76.58%)
Lightweight and natural finish	Guele, MAD Secondate	Guele = 1^{st} lightweight frequency (106) MAD = 2^{nd} lightweight frequency (89) Secondate = 3^{rd} lightweight frequency (80)	Guele = 1^{st} lightweight positivity (92.58%) MAD = 2^{nd} lightweight positivity (89.36%) Secondate = 4th lightweight positivity (89.6%)
Excellent Shade Range	MAD, Secondate, RAD	MAD = 1^{st} shade frequency (202) Secondate = 2^{nd} shade frequency (164) RAD = 3^{rd} shade frequency (159)	MAD = 4^{st} shade positivity (61.65%) Secondate = 1st shade positivity (64.06%) RAD = 6^{th} shade positivity (59.62%)
Smooth Texture	Guele, Esqa, RAD	Guele = 1^{st} texture frequency (92) Esqa = 2^{nd} texture frequency (89) RAD = 3^{rd} texture frequency (88)	Guele = 1^{st} texture positivity (82.72%) Esqa = 2^{nd} texture positivity (81.33%) RAD = 5^{th} texture positivity (79.58%)
Effective Packaging	MAD, Secondate	MAD = 1^{st} packaging frequency (79) Secondate = 3^{rd} packaging frequency (61)	MAD = 6^{th} packaging positivity (85.77%) Secondate = 5^{th} packaging positivity (86.44%)
Acne-friendly Formula	MAD, BLP	MAD = 1^{st} acne frequency (129) BLP = 2^{nd} acne frequency (116)	MAD = 10^{th} acne negativity (17.57%) BLP = 9^{th} acne negativity (18.58%)
Price Sensitivity	Maybelline, L'Oréal	Maybelline = 1^{st} price positivity (72.07%) L'Oréal = 2^{nd} price positivity (70.7%)	

Based on the analysis of word frequency and word probability across brands, customer needs can be strategically matched with the most suitable products. Table 9 provides brand recommendations tailored to specific consumer priorities by presenting the underlying rationale for each recommendation. This approach ensures that recommendations are based on evidence and centered around customer needs, combining statistical insights with practical product features to better align with consumer expectations. In conclusion, while a high frequency of a word may indicate its relevance to a brand's offerings, it is essential to consider the associated positive or negative probability to gauge the significance of that term for each brand. This dual approach enhances the robustness of our recommendations, ensuring they are frequent, reflective of positive customer sentiment, and aligned with consumer needs.

3.7 Brand Strategy Insights

To better support brand growth, we have identified specific areas for improvement and corresponding marketing strategies as summarized

Table 10. Areas for Improvement and Marketing Strategies for Brands

Brand	Areas For Improvement	Marketing Strategy
BLP	Improve packaging and overall appeal	Highlights strengths in coverage and texture in social media campaign; offers customer feedback opportunities for packaging redesigns.
Luxcrime	Improve packaging and increase shade representation	Focus on redesigning packaging for a more appealing look while promoting the brand's shade range. Utilize customer testimonials to highlight both improvements and the variety of shades available.
Guele	Enhance shade variety representation	Create campaigns that spotlight the diversity of shades and depict real customers using various shades, fostering relatability.
Esqa	Improve texture consistency	Launch influencer partnerships to demonstrate the product's benefits, focusing on texture and how it enhances skin appearance.
L'Oréal	Tackle negative perception around acne	Develop a marketing strategy highlighting acne-friendliness, backed by expert endorsements or before-and-after customer testimonials.
Mad For Makeup	Leverage strengths in acne-friendliness	Promote through targeted advertising focusing on the brand's effectiveness for acne-prone individuals; share success stories on social media.
Make Over	Refine texture and reduce disappointment	Engage in product reformulation based on customer feedback and communicate these improvements actively in marketing campaigns.
Maybelline	Increase customer engagement	Foster customer interaction through online polls about desired product features and launch new ranges based on direct feedback.
Rose All Day	Enhance overall satisfaction and perceived value	Re-evaluate pricing strategy and improve product consistency. Utilize marketing communications to express value and quality, backed by customer testimonials.
Secondate	Highlight natural ingredients	Use sustainable marketing campaigns centered on natural and safe ingredients, aligning with health-conscious consumer trends.

in Table 10. Addressing these areas for improvement through targeted marketing strategies strengthens brand positioning and significantly improves customer satisfaction. By leveraging the unique strengths and addressing the weaknesses of each brand, the proposed initiative aims to foster deeper connections with customers. This strategic approach drives brand growth, establishes loyalty, and ensures long-term success in a competitive market. Ultimately, these efforts will create a more engaged customer base, increasing brand advocacy and origin reach.

4. CONCLUSION

This study employs sentiment analysis using the Naive Bayes algorithm and the Bag of Words method to classify customer reviews of ten concealer brands on Shopee. The model achieved an accuracy of 87.95%, effectively extracting insights from 3,920 reviews, including ten known concealer brands in Indonesia. By combining word frequency and sentiment probability, the analysis identified each brand's strengths, such as blendability, coverage, packaging, and areas for improvement, including shade inclusivity, acne friendliness, and disappointment rates.

The findings led to targeted brand recommendations based on customer needs, ranging from full coverage and lightweight textures to safer formulations for acne-prone skin. Additionally, strategic suggestions were provided to help brands enhance their market position by amplifying strong attributes and addressing customer pain points. Overall, this research highlights the value of sentiment analysis as a tool for customer-centered product and marketing development in the cosmetic industry.

REFERENCE

[1] Y. Pei, J. Niu, and K. Min, "Value co-destruction in live-commerce ecosystem: an evolutionary game approach with insights from China," *Electronic Commerce Research*, 2025, doi: 10.1007/s10660-025-09957-3.

[2] A. S. Sridhar and S. Nagasundaram, "Sentiment classification for review rating prediction based on hybrid reaction diffusion support vector neural network," *Soc Netw Anal Min*, vol. 14, no. 1, p. 240, 2024, doi: 10.1007/s13278-024-01397-5.

[3] A. Daza, N. D. González Rueda, M. S. Aguilar Sánchez, W. F. Robles Espíritu, and M. E. Chauca Quiñones, "Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works," *International Journal of Information Management Data Insights*, vol. 4, no. 2, p. 100267, 2024, doi: https://doi.org/10.1016/j.jjimei.2024.100267.

[4] A. Ray *et al.*, "Sentiment Analysis of Amazon Product Reviews: A Comprehensive Evaluation Using Naïve Bayes Classifiers," in *Computing and Machine Learning*, J. C. Bansal, S. Borah, S. Hussain, and S. Salhi, Eds., Singapore: Springer Nature Singapore, 2024, pp. 347–358.

[5] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl Based Syst*, vol. 226, p. 107134, 2021, doi: https://doi.org/10.1016/j.knosys.2021.107134.

[6] Q. Lu, X. Sun, Y. Long, Z. Gao, J. Feng, and T. Sun, "Sentiment Analysis: Comprehensive Reviews, Recent Advances, and Open Challenges," *IEEE Trans Neural Netw Learn Syst*, vol. 35, no. 11, pp. 15092–15112, 2024, doi: 10.1109/TNNLS.2023.3294810.

[7] M. Al-Fayoumi and Q. Abu Al-Haija, "Capturing low-rate DDoS attack based on MQTT protocol in software Defined-IoT environment," *Array*, vol. 19, p. 100316, 2023, doi: https://doi.org/10.1016/j.array.2023.100316.

[8] A. H. Aljemely, J. Xuan, L. Xu, F. K. J. Jawad, and O. Al-Azzawi, "Wise-local response convolutional neural network based on Naïve Bayes theorem for rotating machinery fault classification," *Applied Intelligence*, pp. 1–19, 2021.

[9] T. Ige, C. Kiekintveld, A. Piplai, A. Waggler, O. Kolade, and B. H. Matti, "An investigation into the performances of the Current state-of-the-art Naive Bayes, Non-Bayesian and Deep Learning Based Classifier for Phishing Detection: A Survey," *ArXiv*, vol. abs/2411.16751, 2024, [Online]. Available: https://api.semanticscholar.org/CorpusID:274280486

[10] S. Kumar, A. Sharma, B. K. Reddy, S. Sachan, V. Jain, and J. Singh, "An intelligent model based on integrated inverse document frequency and multinomial Naive Bayes for current affairs news categorisation," *International Journal of System Assurance Engineering and Management*, vol. 13, no. 3, pp. 1341–1355, 2022, doi: 10.1007/s13198-021-01471-7.

[11] M. Schonlau, "The Naive Bayes Classifier," in *Applied Statistical Learning: With Case Studies in Stata*, M. Schonlau, Ed., Cham: Springer International Publishing, 2023, pp. 143–160. doi: 10.1007/978-3-031-33390-3_8.

[12] A. Angdresey, L. Sitanayah, and I. L. H. Tangka, "Sentiment Analysis for Political Debates on YouTube Comments using BERT Labeling, Random Oversampling, and Multinomial Naïve Bayes," *Journal of Computing Theories and Applications*, vol. 2, no. 3, pp. 342–354, 2025.

[13] N. L. Adam, N. H. Rosli, and S. C. Soh, “Sentiment Analysis on Movie Review using Naïve Bayes,” in *2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, 2021, pp. 1–6. doi: 10.1109/AiDAS53897.2021.9574419.

[14] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J.-H. Jeng, and J.-G. Hsieh, “Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes,” *Information*, vol. 12, no. 5, p. 204, 2021.

[15] K. A. Hashim, Y. Sibaroni, and S. S. Prasetyowati, “The Effectiveness of the Ensemble Naive Bayes in Analyzing Review Sentiment of the Lazada Application on Google Play,” in *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, 2024, pp. 1–5. doi: 10.1109/ICETSIS61505.2024.10459562.

[16] F. Novitasari and M. Dwifebri Purbolaksono, “Sentiment Analysis Aspect Level on Beauty Product Reviews Using Chi-Square and Naïve Bayes,” *JOURNAL OF DATA SCIENCE AND ITS APPLICATIONS*, vol. 4, no. 1, pp. 18–30, 2021.

[17] D. Marlina, K. Tri Basuki, Z. Mohd Zaki, and A. Siti Farahnasihah, “Sentiment Analysis on Natural Skincare Products,” *Journal of Data Science*, vol. 2022, no. 12, pp. 1–17, 2022.

[18] C. Therence and J. Tjen, “Analysis of Cosmetic Product Opinions on E-commerce Based on Naïve Bayes Classifier,” in *International Conference on Digital Business Innovation and Technology Management (ICONBIT)*, 2024.

[19] N. Suryawanshi, “Sentiment analysis with machine learning and deep learning: A survey of techniques and applications,” *International Journal of Science and Research Archive*, vol. 12, pp. 5–15, Jun. 2024, doi: 10.30574/ijsra.2024.12.2.1205.

[20] E. De Santis, A. Martino, F. Ronci, and A. Rizzi, “From Bag-of-Words to Transformers: A Comparative Study for Text Classification in Healthcare Discussions in Social Media,” *IEEE Trans Emerg Top Comput Intell*, vol. 9, no. 1, pp. 1063–1077, 2025, doi: 10.1109/TETCI.2024.3423444.

[21] S. Juanita, A. H. Daeli, M. Syafrullah, W. Anggraeni, and M. H. Purnomo, “A machine learning approach for text pattern diagnosis in mental health consultations,” *Decision Analytics Journal*, p. 100572, 2025, doi: https://doi.org/10.1016/j.dajour.2025.100572.