

Valid or Just Viral? Post-Purchase Consumer Sentiment and Regulatory Discourse on Imported Vitamins on Shopee

Rusda^{1*}, Syamsu Alam², Ayu Latifah Alfisyahrin³, Ainul Fatha Isman⁴

¹Department of Islamic Economics and Business, Faculty of Economics and Business, Universitas Hasanuddin, Makassar, 90245, Indonesia

²Department of Digital Business, Faculty of Economics and Business, Universitas Negeri Makassar, Makassar, Indonesia

³Department of Islamic Economics and Business, Faculty of Economics and Business, Universitas Hasanuddin, Makassar, 90245, Indonesia

⁴Department of Islamic Economics and Business, Faculty of Economics and Business, Universitas Hasanuddin, Makassar, 90245, Indonesia

ABSTRACT

Purpose – This study examines sentiment and recurring topics in post-purchase reviews of imported vitamin products whose Shopee listings did not clearly display halal certification or Badan Pengawas Obat dan Makanan (BPOM; Indonesian Food and Drug Authority) registration information. It assesses whether halal and BPOM appear explicitly in consumer-generated reviews.

Design/methodology/approach – A quantitative descriptive design combined manual reference labelling, supervised sentiment classification, and descriptive lexical analysis. One researcher labelled 250 verified-buyer reviews from seven products and coded recurring keywords. Raw text was divided through a stratified 80:20 train-test split before term frequency-inverse document frequency (TF-IDF) with unigram and bigram features was fitted on the training text and applied to the held-out text. Multinomial Naive Bayes, Logistic Regression, Linear Support Vector Classifier (LinearSVC), and Random Forest were evaluated using overall, macro, weighted, class-specific, and confusion-matrix measures.

Findings – Neutral labels predominated (58.0%), followed by positive (31.6%) and negative (10.4%) labels. Random Forest achieved the strongest performance within the experiment (accuracy = 0.80; macro F1 = 0.65), followed by LinearSVC (accuracy = 0.78; macro F1 = 0.63), although both detected only one of five negative test cases. Recurring themes concerned authenticity, packaging, delivery, expiration dates, and product condition. “halal” did not occur, and one BPOM-related expression was identified.

Originality/value – The study complements intention-based halal consumer research by analysing naturally occurring post-purchase discourse while distinguishing review content from pre-purchase decision-making and official regulatory status. It also separates classifier performance from the substantive label distribution in a small, imbalanced dataset.

ARTICLE INFO

Keywords:

BPOM;
E-commerce;
Halal certification;
Machine learning;
Post-purchase review;
Sentiment analysis.

Article Information:

Received: 09/06/2026
Revise: 29/08/2026
Accepted: 13/09/2026

ISSN:

2985-3168 (Online)
2985-3222 (Print)

*Corresponding Author at:

Department of Islamic Economics and Business, Faculty of Economics and Business, Universitas Hasanuddin, Jalan Perintis Kemerdekaan No.10, Makassar 90245, Indonesia.

E-mail address: rusdaekbis@unhas.ac.id (Rusda)

The work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)



1. Introduction

E-commerce has changed how consumers encounter, evaluate, and discuss products. Alongside formal information from regulators and certification bodies, marketplace ratings and buyer reviews provide consumer-generated information about product and transaction experiences. Online reviews therefore complement product descriptions and other marketplace communications, although the content of a review should not be treated automatically as evidence of the considerations that preceded a purchase (Chen & Xie, 2008; Mudambi & Schuff, 2010).

This distinction is particularly important for vitamins and health supplements. These products involve safety, ingredient, dosage, legality, and halal-related considerations, while imported products promoted through marketplaces and social media may achieve high visibility before regulatory information is consistently visible to consumers. International research has documented risks associated with online dietary supplements, including inaccurate claims, adulteration, undeclared ingredients, uneven oversight, and challenges in controlling imported supplements sold through online marketplaces (Crawford et al., 2022; Golovan, 2024; Rathmann & Seifert, 2024; White, 2022). In Indonesia, the Indonesian Food and Drug Authority has also reported extensive illegal online links for drug and food products, underscoring the relevance of regulatory visibility in digital markets (Badan Pengawas Obat dan Makanan Republik Indonesia, 2025).

In Indonesia, Badan Pengawas Obat dan Makanan (BPOM; Indonesian Food and Drug Authority) registration and the halal assurance system serve different but complementary governance functions. The public BPOM product-registration database provides a formal channel for checking registered products (Badan Pengawas Obat dan Makanan Republik Indonesia, 2023), while Badan Penyelenggara Jaminan Produk Halal (BPJPH; Halal Product Assurance Organizing Agency) administers halal certification and reports the continuing expansion of certified products (Badan Penyelenggara Jaminan Produk Halal, 2025). Halal also operates as a market and trust-related attribute whose meaning extends beyond a product logo (Wilson & Liu, 2010). For health supplements, the study period of September–December 2024 fell within the transition toward the mandatory halal-certification deadline of 17 October 2026. The study therefore examines information visibility during a transitional regulatory period rather than presuming non-compliance.

The empirical issue is not whether the selected products definitively lacked registration or certification, but whether such information was clearly visible in the observed listings and whether consumers subsequently mentioned it in reviews. This boundary matters because an absent listing field, a consumer claim of authenticity, or a barcode scan cannot independently establish official regulatory status.

Shopee provides seller-facing fields through which halal certificate numbers and BPOM registration information may be entered. Figure 1 illustrates this platform feature. Its availability creates an information channel, but does not guarantee that every listing contains complete or verifiable regulatory information.

Figure 1. Seller Center Feature for Entering halal Certification Numbers and BPOM Registration Numbers on Shopee

The screenshot displays the 'Spesifikasi' (Specifications) page in the Shopee Seller Center. The page is divided into a sidebar on the left and a main content area on the right. The sidebar contains a 'Kualitas Informasi Produk' (Product Information Quality) section with a 'Baik' (Good) status and a 'Sangat Baik' (Very Good) button. Below this is a 'Tips' (Tips) section with a lightbulb icon and a list of tips for improving product quality. The main content area is titled 'Spesifikasi' and shows a progress bar indicating 'Lengkap: 2 / 16' (Completed: 2 / 16). It lists various product attributes to be filled out, including 'Merek' (Brand), 'Masa Penyimpanan' (Shelf Life), 'Kondisi Penyimpanan' (Storage Condition), 'Volume', 'No. Izin Edar (BPOM, PIRT)' (BPOM, PIRT License Number), 'Rasa' (Taste), 'Jenis Susu' (Milk Type), 'Tanggal Kedaluwarsa' (Expiration Date), 'Negara Asal' (Country of Origin), 'Ingredient', 'Berat Produk' (Product Weight), 'Tipe Paket' (Package Type), 'Ukuran Per Produk' (Product Size), 'Jumlah Produk Dalam Kemasan' (Number of Products in Packaging), 'Menu Makanan Khusus' (Special Food Menu), and 'No. Sertifikasi (Halal)' (Halal Certification Number). The 'No. Sertifikasi (Halal)' field is highlighted with a red box, and a message below it says 'Belum memiliki Sertifikat Halal? Daftar sekarang' (Don't have a Halal Certificate? Register now). At the bottom of the page, there are three buttons: 'Kembali' (Back), 'Arsipkan' (Archive), and 'Update'.

Source: Shopee, 2025

Accordingly, this study treats the selected listings as cases in which certification or registration information was not clearly visible across the observed digital sources. The classification is limited to information visibility during data collection and is not an official determination of halal or BPOM status.

Prior studies commonly examine purchase intention, halal awareness, religiosity, trust, electronic word-of-mouth (e-WOM), and social commerce through surveys, behavioural models, and meta-analytic approaches (Ajzen, 1991; Tedjakusuma et al., 2023; Bhutto et al., 2024; Khan et al., 2026; Handoyo, 2024; Mou & Benyoucef, 2021). Online-review research, by contrast, treats reviews as observable post-purchase communications whose content, depth, and valence can inform other consumers and marketplace actors (Chen & Xie, 2008; Mudambi & Schuff, 2010; Filieri, 2015). Meta-analytic evidence also shows that reviews can be associated with market outcomes, but such associations do not allow an individual review corpus to reveal unobserved purchase motivations (Floyd et al., 2014).

The study contributes by examining whether formal regulatory attributes specifically halal certification and BPOM registration appear in naturally occurring post-purchase reviews of highly visible imported vitamin products. The analysis does not test attitudes, purchase intentions, or causal determinants. Instead, it combines reference-label distributions, classifier comparison, and descriptive lexical analysis to identify what consumers voluntarily communicated after a transaction.

The objectives are therefore to: (1) describe the distribution of positive, neutral, and negative reference labels; (2) compare four supervised classifiers on the same held-out test set; (3) identify recurring keywords and n-grams in the review corpus; and (4) assess the explicit visibility of halal- and BPOM-related discourse. This design positions post-purchase review mining as a complement to, rather than a substitute for, intention-based consumer research or formal regulatory verification.

2. Methodology

2.1 Location and Time of the Study

The study used secondary data from publicly accessible verified-buyer reviews on Shopee. Seven imported vitamin products marketed through 12 stores were selected because each listing had a rating of at least four stars and more than 100 reviews, while halal certification or BPOM registration information was not clearly displayed during the observation period. Reviews were collected from September to December 2024, when health supplements remained within Indonesia's halal-certification transition period ending on 17 October 2026.

A single researcher manually assigned the reference labels after reading the complete text of each of the 250 reviews. Positive labels represented predominantly favourable evaluations, satisfaction, approval, recommendation, or appreciation. Neutral labels represented primarily descriptive, factual, informational, expectational, or uncertain statements without a dominant evaluative direction. Negative labels represented complaints, dissatisfaction, distrust, disappointment, concern, or other predominantly unfavourable evaluations. These reference labels served both as the descriptive sentiment distribution and as the target classes for supervised classification.

For mixed reviews, the label followed the dominant evaluative orientation of the complete review; if no orientation clearly dominated, the review was coded as neutral. Expressions such as “semoga asli” and “semoga bermanfaat” were interpreted in their full review context and were not assigned polarity from isolated words. Because one researcher performed the coding, no inter-coder agreement statistic was available. This single-coder design is reported explicitly and considered in the limitations.

2.2 Research Subjects

The subjects of this study are consumer reviews located on the product pages of viral imported vitamin products on the Shopee platform. The determination of the research subjects was based on several criteria, namely :

1. The products are imported vitamin supplements with a minimum rating of four stars and more than 100 consumer reviews.
2. At the time of data collection, the product pages did not visibly display halal certification information or BPOM registration numbers. Visibility was checked using (i) Shopee product pages, (ii) seller presentations and responses observed during TikTok Live sessions, and (iii) searches of the public BPOM and BPJPH databases using the available product identifiers. Each stage documented whether information was visible, mentioned, confirmed, or identifiable; it did not establish an official regulatory determination.
3. The reviews were written by verified buyers who had completed transactions on the Shopee platform

Based on these criteria, the research sample comprises seven imported vitamin products that are most frequently purchased and reviewed on Shopee from 12 stores, namely:

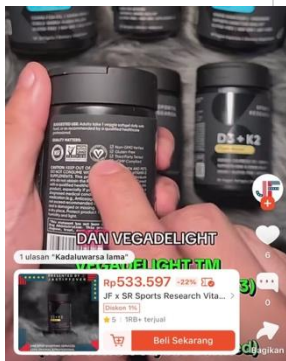
1. SR Sports Research K2 100 mcg D3 5000 IU MK7 Vegan Softgels Plant-Based;
2. Naturebell Astaxanthin 12 mg – 120 & 240 Softgels;
3. Naturebell Astaxanthin 12 mg – 240 Softgels Original USA;
4. Eagleshine Vitamin D3 5000 IU K2 100 mcg MK7 Deal Supplement 250 Softgels;

5. Deal Supplement Vitamin D3 K2 2-in-1 Complex 5000 IU & MK7 – 180 & 250 Softgels;
6. Sports Research SR Vitamin D3 + K2 5000 IU 160 Softgels;
7. SR Sports Research Vitamin D3 + K2 Plant-Based with Coconut MCT Oil 125 mcg – 30, 60, and 120 Softgels.

2.3 Product Verification Procedure

Before scraping, the researchers applied a multi-source verification procedure to ensure that the selected listings met the information-visibility criterion. The Shopee listing and TikTok Live observation recorded whether halal or BPOM information was displayed. Seller communication recorded whether the information was explicitly confirmed during the observed session. Searches of the public BPOM and BPJPH databases recorded whether a match could be identified from the product information available to the researchers (Badan Pengawas Obat dan Makanan Republik Indonesia, 2023; Badan Penyelenggara Jaminan Produk Halal, 2025). Table 1 reports the outcome of each stage and distinguishes an information search from an official-status determination.

Table 1. Multi-Source Verification of halal Certification and BPOM Information for Selected Imported Vitamin Products

No	Product	Shopee Listing	TikTok Live Observation	Seller Confirmation During Live	BPOM Search Result	BPJPH Search Result	Final Classification
1	SR Sports Research K2 100 mcg D3 5000 IU MK7 Vegan Softgels Plant-Based	Not displayed	Not displayed	 Not confirmed in live session.	No identifiable match	No identifiable match	Not visible; official status undetermined
2	Naturebell Astaxanthin 12 mg – 120 & 240 Softgels	Not displayed	Not displayed	 Not confirmed in live session.	No identifiable match	No identifiable match	Not visible; official status undetermined

No	Product	Shopee Listing	TikTok Live Observation	Seller Confirmation During Live	BPOM Search Result	BPJPH Search Result	Final Classification
3	Naturebell Astaxanthin 12 mg – 240 Softgels Original USA	Not displayed	Not displayed	 Tidak terkonfirmasi dalam sesi live	No identifiable match	No identifiable match	Not visible; official status undetermined
4	Eagleshine Vitamin D3 5000 IU K2 100 mcg MK7 Deal Supplement 250 Softgels	Not displayed	Not displayed	 Tidak terkonfirmasi dalam sesi live	No identifiable match	No identifiable match	Not visible; official status undetermined
5	Deal Supplement Vitamin D3 K2 2-in-1 Complex 5000 IU & MK7 – 180 & 250 Softgels	Not displayed	Not displayed	 Tidak terkonfirmasi dalam sesi live	No identifiable match	No identifiable match	Not visible; official status undetermined

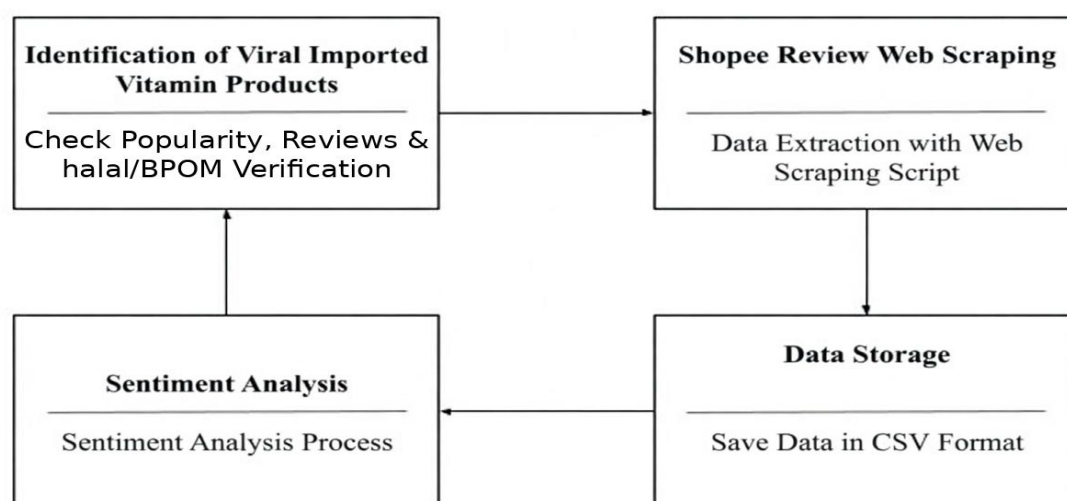
No	Product	Shopee Listing	TikTok Live Observation	Seller Confirmation During Live	BPOM Search Result	BPJPH Search Result	Final Classification
6	Sports Research SR Vitamin D3 + K2 5000 IU 160 Softgels	Not displayed	Not displayed	 Not confirmed in live session.	No identifiable match	No identifiable match	Not visible; official status undetermined
7	SR Sports Research Vitamin D3 + K2 Plant-Based with Coconut MCT Oil	Not displayed	Not displayed	 Not confirmed in live session.	No identifiable match	No identifiable match	Not visible; official status undetermined

Products were retained when certification or registration information was not clearly visible across the observed listing and live-session sources. Database results were constrained by the identifiers available for each product. The resulting classification therefore describes the visibility and traceability of information during the observation period; it does not demonstrate that a product lacked official registration or certification.

2.4 Data Description

Review data were collected through a web-scraping script adapted to the public structure of the selected Shopee pages. The resulting file contained the product identifier, review text, and manually assigned sentiment label. The final analytical dataset comprised 250 verified-buyer reviews from seven products sold through 12 stores.

Figure 2. Web Scraping Method Process



Source: Data processed by the researcher

Figure 2 summarizes the web-scraping process used after product selection. Listings were screened against the rating, review-volume, imported-product, and information-visibility criteria; the selected review texts were then extracted and stored in comma-separated values (CSV) format. The analysis used only the review text, product identifier, and reference label required for the stated objectives.

The reference-label distribution was 145 neutral reviews (58.0%), 79 positive reviews (31.6%), and 26 negative reviews (10.4%). These labels were analysed descriptively and used as the target classes for supervised model evaluation; they were not predictions produced by any one classifier.

The same single researcher examined all 250 reviews for the manual keyword analysis. Recurrent Indonesian expressions, abbreviations, spelling variants, and semantically equivalent forms were grouped under study-relevant terms such as authenticity, packaging, delivery, expiration date, safety, halal, and BPOM. Ambiguous expressions were interpreted within the complete review rather than from an isolated token. The resulting frequencies describe explicit lexical occurrence and contextual co-occurrence; they are not estimates of causal importance or aspect-level sentiment generated by a conventional aspect-based sentiment analysis (ABSA) model.

2.5 Data Analysis Techniques

The analysis combined three complementary components. First, descriptive statistics summarized the reference labels. Second, supervised classifiers evaluated how accurately the manually assigned labels could be reproduced on held-out reviews. Third, keyword, word cloud, and n-gram analyses described recurring lexical patterns. These components answer different questions and were interpreted separately.

Particular attention was given to expressions concerning authenticity, packaging, delivery, product condition, expiration dates, safety, halal certification, and BPOM registration. Frequencies and n-grams were used to identify recurring language within labelled subsets, not to infer determinants of satisfaction, trust, or purchase decisions.

Figure 3. Research Workflow for Consumer Review Collection, Sentiment Classification, Model Comparison, and Textual Analysis



Source: Data processed by the researcher

2.6 Text Pre-processing and Model Development

Review text was converted to lowercase and non-alphabetic characters were removed. The raw texts and reference labels were then divided using a stratified 80:20 train-test split with random state 42, producing 200 training reviews and 50 held-out test reviews. Stratification preserved the relative distribution of negative, neutral, and positive labels.

After the split, term frequency-inverse document frequency (TF-IDF) with unigram and bigram features and a maximum of 5,000 features was fitted only on the training text. The fitted vectorizer was subsequently used to transform both the training and held-out test texts. This sequence prevented information from the test set from contributing to feature estimation.

Four classifiers were trained on the same TF-IDF training matrix: Multinomial Naive Bayes, Logistic Regression, Linear Support Vector Classifier (LinearSVC), and Random Forest. Logistic Regression used a maximum of 1,000 iterations, Random Forest used 200 trees, and random state 42 was applied where relevant. Evaluation included accuracy, class-specific precision, recall and F1-score, macro and weighted averages, and confusion matrices. Foundational and recent sentiment-analysis literature supports comparing multiple algorithms and interpreting performance beyond accuracy alone (Pang et al., 2002; Medhat et al., 2014; Liu, 2020; Abighail et al., 2023; Mao et al., 2024; Sharma et al., 2025).

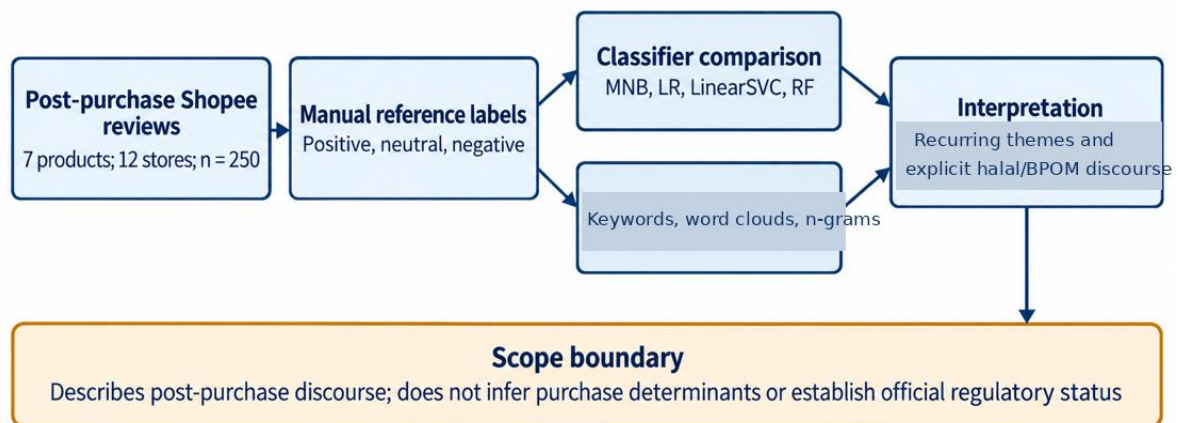
2.7 Research Procedure

The procedure comprised product selection and visibility checks, review collection, manual reference labelling, stratified splitting of raw text, training-only TF-IDF fitting, classifier training and held-out evaluation, descriptive lexical analysis, and integrated interpretation. Figure 3 displays the corrected sequence. The raw review text was split before feature extraction. TF-IDF was fitted on the 200 training reviews and applied to the 50 held-out reviews, ensuring that test-set vocabulary statistics did not inform model development.

Multinomial Naive Bayes, Logistic Regression, LinearSVC, and Random Forest were evaluated on the same test observations. Model comparison used overall, macro-averaged, weighted, class-specific, and confusion-matrix measures because the sentiment classes were imbalanced. Keyword, word cloud, and n-gram analyses were then used descriptively to identify recurring expressions within the reference-labelled corpus.

The final interpretation integrated label distribution, classifier performance, and lexical recurrence while maintaining two scope boundaries: post-purchase reviews do not reveal unobserved pre-purchase determinants, and consumer discourse does not establish official regulatory status. Figure 4 summarizes this analytical framework and the relationship among the study components.

Figure 4. Analytical Framework for Post-Purchase Review Analysis



Source: Developed by the researchers

3. Result and Discussion

3.1 Result

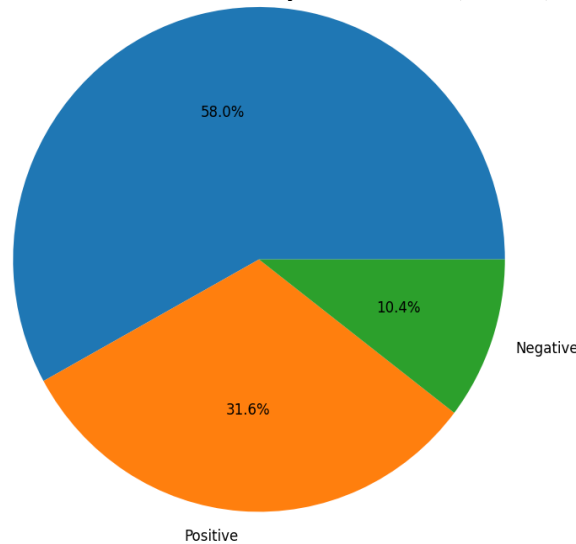
The results are reported in three parts: the distribution of manually assigned reference labels, recurring lexical patterns, and held-out classifier performance. The first and third parts are

analytically distinct: the label distribution describes the dataset, whereas model metrics evaluate whether algorithms reproduce those labels.

3.1.1 General Sentiment Distribution Analysis

Figure 5 presents the distribution of negative, neutral, and positive reference labels in the 250-review dataset.

Figure 5. Distribution of Sentiment Labels in the Analytical Dataset (n = 250)



Source: Data processed by the researchers from the labelled analytical dataset.

Neutral reviews formed the largest category (145; 58.0%), followed by positive reviews (79; 31.6%) and negative reviews (26; 10.4%). Neutral labels were assigned mainly to descriptive, informational, expectational, or uncertain statements without a dominant favourable or unfavourable orientation. The imbalance is relevant to model evaluation because the held-out test set contained 29 neutral, 16 positive, and only five negative observations. Performance for the negative class must therefore be interpreted from class-specific recall and F1-score rather than from accuracy alone.

3.1.2 Word Distribution and Frequency

Target-word frequencies identify expressions that recurred in the review corpus. They show what was mentioned and in which reference-labelled reviews the terms occurred; they do not establish that a term caused satisfaction, dissatisfaction, trust, or purchase behaviour.

Table 2. Target Words and Word Frequency Table for Each Sentiment

Target Word	Meaning (English)	Word Frequency	Most Common Words Before	Meaning (English)	Most Common Words After	Meaning (English)	Positive	Negative
Asli	Authentic	45	{'keaslian': 16, 'semoga': 7, 'barang': 3, 'kelihatan': 2, 'terjamin': 2}	{'authenticity': 16, 'hopefully': 7, 'item': 3, 'looks': 2, 'guaranteed': 2}	{'ya': 3, 'expired': 2, 'bersebel': 1, 'semoga': 1, 'cocok': 1}	{'yes': 3, 'expired': 2, 'sealed': 1, 'hopefully': 1, 'suitable': 1}	18	7

Target Word	Meaning (English)	Word Frequency	Most Common Words Before	Meaning (English)	Most Common Words After	Meaning (English)	Positive	Negative
Ori	Original (slang)	22	{'produk': 3, 'Allah': 2, 'ini': 2, 'barang': 1, 'dan': 1}	{'product': 3, 'God': 2, 'this': 2, 'item': 1, 'and': 1}	{'dan': 3, 'ya': 2, 'pengiriman': 2, 'bukan': 2, 'next': 1}	{'and': 3, 'yes': 2, 'delivery': 2, 'not': 2, 'next': 1}	8	2
Aman	Safe	22	{'packing': 5, 'dengan': 2, 'dan': 2, 'paket': 1, 'packaging': 1}	{'packing': 5, 'with': 2, 'and': 2, 'package': 1, 'packaging': 1}	{'dan': 3, 'sampai': 2, 'produk': 2, 'ada': 2, 'pengiriman': 2}	{'and': 3, 'arrived': 2, 'product': 2, 'available': 2, 'delivery': 2}	13	0
Original	Original	14	{'produk': 2, 'mudah': 1, 'jaminan': 1, 'barang': 1, 'yakin': 1}	{'product': 2, 'hopefully': 1, 'guarantee': 1, 'item': 1, 'sure': 1}	{'pricey': 1, 'barang': 1, 'segel': 1, 'kalian': 1, 'dari': 1}	{'pricey': 1, 'item': 1, 'seal': 1, 'you guys': 1, 'from': 1}	5	5
Packing	Packaging	13	{'cepat': 2, 'pesanan': 1, 'pengiriman': 1, 'asli': 1, 'tubuh': 1}	{'fast': 2, 'order': 1, 'delivery': 1, 'authentic': 1, 'body': 1}	{'aman': 5, 'rapi': 3, 'rapih': 3, 'bagus': 1, 'nya': 1}	{'safe': 5, 'neat': 3, 'tidy': 3, 'good': 1, 'suffix': 1}	12	0
Kualitas	Quality	5	{'pesanan': 1}	{'order': 1}	{'semoga': 1, 'good': 1, 'baik': 1, 'bgus': 1, 'produk': 1}	{'hopefully': 1, 'good': 1, 'good': 1, 'good': 1, 'product': 1}	1	0
Kardus	Box / Carton	3	{'ada': 1, 'kotak': 1, 'dan': 1}	{'there is': 1, 'box': 1, 'and': 1}	{'kemasan': 1, 'deskripsi': 1, 'seller': 1}	{'packaging': 1, 'description': 1, 'seller': 1}	3	0
Palsu	Fake	2	{'atau': 1, 'produk': 1}	{'or': 1, 'product': 1}	{'hanya': 1, 'sdh': 1}	{'only': 1, 'already': 1}	1	0
Abal abal	Low-quality / Counterfeit	1	{'yang': 1}	{'that': 1}	{'beredar': 1}	{'circulating': 1}	0	1
Pembelian ke berapa	Purchase count	0	{}	{}	{}	{}	0	0
BPOM	Indonesian FDA	1	{BPOM': 1}	{}	{}	{}	0	0
halal	halal	0	{}	{}	{}	{}	0	0
Viral	Viral	0	{}	{}	{}	{}	0	0

Source: Data processed by the researchers using Python-based textual analysis, 2025.

The classification in the Positive column refers to the overall sentiment label assigned to the review containing the BPOM-related expression; it should not be interpreted as positive sentiment toward BPOM status itself.

Table 2 shows that authenticity-related terms were the most recurrent among the selected targets: “asli” occurred 45 times, “ori” 22 times, and “original” 14 times. “Aman” occurred 22 times and frequently co-occurred with packaging or delivery language, while “packing” occurred 13 times. Less frequent terms included “kualitas” (5), “kardus” (3), “palsu” (2), and “abal-abal” (1). The distribution across reference labels indicates context variation rather than a fixed polarity for any isolated term.

Contextual interpretation of target words

Authenticity expressions occurred in positive, neutral, and negative-labelled reviews. In positive reviews, they often accompanied confirmation or satisfaction; in neutral reviews, they frequently appeared in questions, descriptions, or hopes; and in negative reviews, they occurred with doubt or dissatisfaction. Likewise, “aman” and “packing” commonly appeared in positive-labelled reviews, but the frequency evidence supports only their recurrence within those reviews, not their status as determinants of satisfaction.

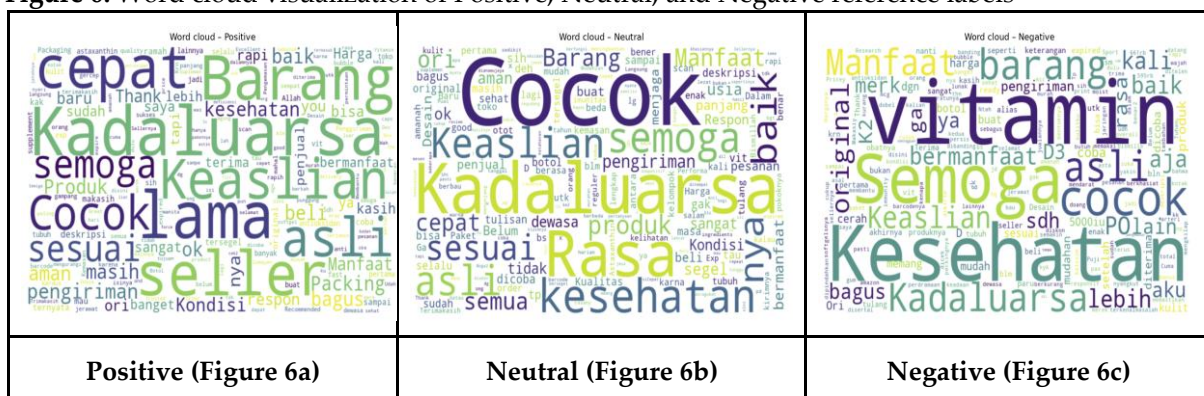
Regulation-related terms

The exact term “halal” was not identified in the 250 reviews. One review contained a BPOM-related expression, “non BPOM.” The regulatory terminology was therefore exceptionally infrequent relative to authenticity, packaging, delivery, expiration-date, and product-condition language. This finding concerns explicit occurrence in the review corpus and does not measure consumers' regulatory awareness or pre-purchase considerations.

3.1.3 Word Cloud Analysis Results

Word clouds provide a visual summary of relative word frequency within each reference-labelled subset. Figure 6 presents the Positive, Neutral, and Negative word clouds as descriptive aids, which are interpreted together with the underlying review context.

Figure 6. Word cloud visualization of Positive, Neutral, and Negative reference labels



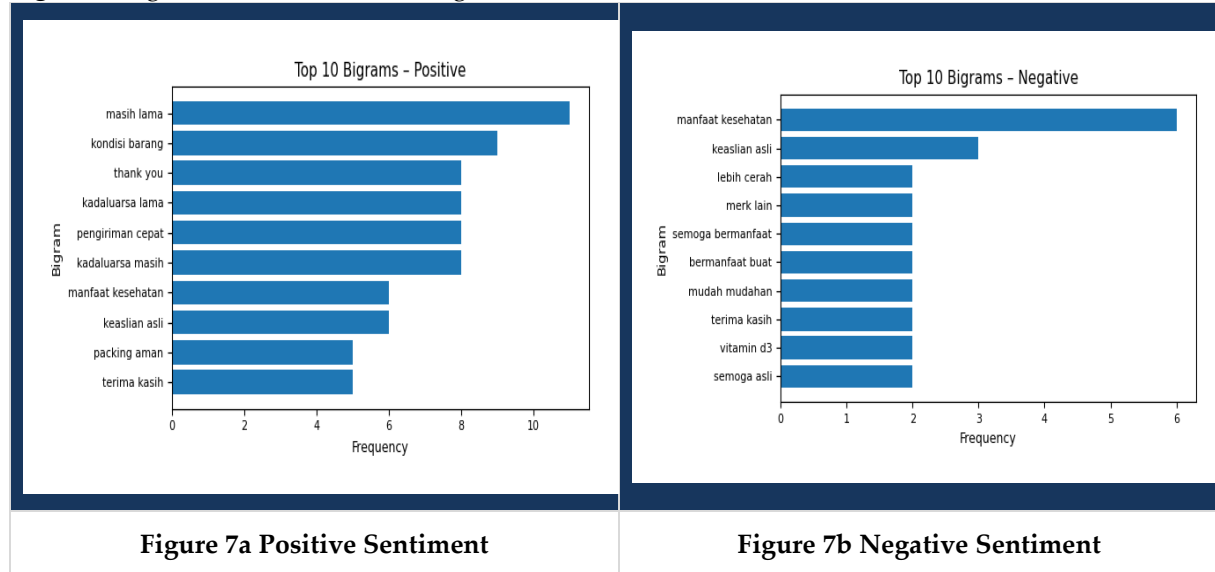
Source: Data processed by the researchers using Python-based textual analysis, 2025.

In positive-labelled reviews, prominent expressions concerned products, authenticity, delivery speed, packaging, suitability, and expiration dates. These terms recurred in favourable review contexts, but their visual prominence does not demonstrate that they independently produced positive sentiment.

Neutral-labelled reviews prominently contained suitability and expiration-related expressions, consistent with factual, expectational, or uncertain descriptions. In negative-labelled reviews, words related to vitamins, health benefits, authenticity, and expiration appeared within broader comments containing doubt, comparison, or dissatisfaction.

Bigram analysis identifies pairs of words that occurred consecutively within the labelled subsets. As with individual keywords, a bigram's meaning depends on the surrounding review and should not be assigned a fixed sentiment from the phrase alone. Figure 7 presents the most frequent bigrams in the Positive- and Negative-labelled reviews.

Figure 7. Bigrams in Positive- and Negative-Labelled Reviews



Source: Data processed by the researchers using Python-based textual analysis, 2025.

Positive-labelled reviews contained recurrent bigrams such as “masih lama,” “kadaluarsa lama,” “pengiriman cepat,” “packing aman,” “thank you,” and “terima kasih.” These phrases describe expiration periods, delivery, packaging, authenticity, or appreciation within reviews carrying positive reference labels.

Negative-labelled reviews contained phrases including “manfaat kesehatan,” “keaslian asli,” “lebih cerah,” “merk lain,” “semoga bermanfaat,” “mudah mudahan,” and “semoga asli.” Several are semantically neutral or hopeful in isolation. Their inclusion means only that they occurred in reviews labelled negative on the basis of the complete comment; it does not establish sarcasm, disappointment, or concern without the surrounding text.

Table 3. Negative Sentiment Bigram Analysis

No	Bigram	General Meaning	Observed Context Within Negative-Labelled Reviews
1	manfaat kesehatan	Related to product health claims	Occurred in a negative-labelled review; the phrase itself is not inherently negative
2	keaslian asli	About product authenticity	Co-occurred with authenticity-related language in a negative-labelled review
3	lebih cerah	Usually a usage result claim	Co-occurred with an outcome-related statement in a negative-labelled review

No	Bigram	General Meaning	Observed Context Within Negative-Labelled Reviews
4	merk lain	Comparison with other brands	Co-occurred with a comparison to another brand
5	semoga bermanfaat	Expression of hope	Expression of hope occurring in a negative-labelled review; no fixed polarity assigned
6	bermanfaat buat	Related to expected effectiveness	Benefit-related expression occurring in a negative-labelled review
7	mudah mudahan	Hope for results	Expression of hope or uncertainty occurring in a negative-labelled review
8	terima kasih	Polite thank you	Polite closing occurring in a negative-labelled review
9	vitamin d3	Ingredient/product name	Product-name expression occurring in a negative-labelled review
10	semoga asli	Hope for authenticity	Expression of hoped-for authenticity occurring in a negative-labelled review

Source: Data processed by the researcher

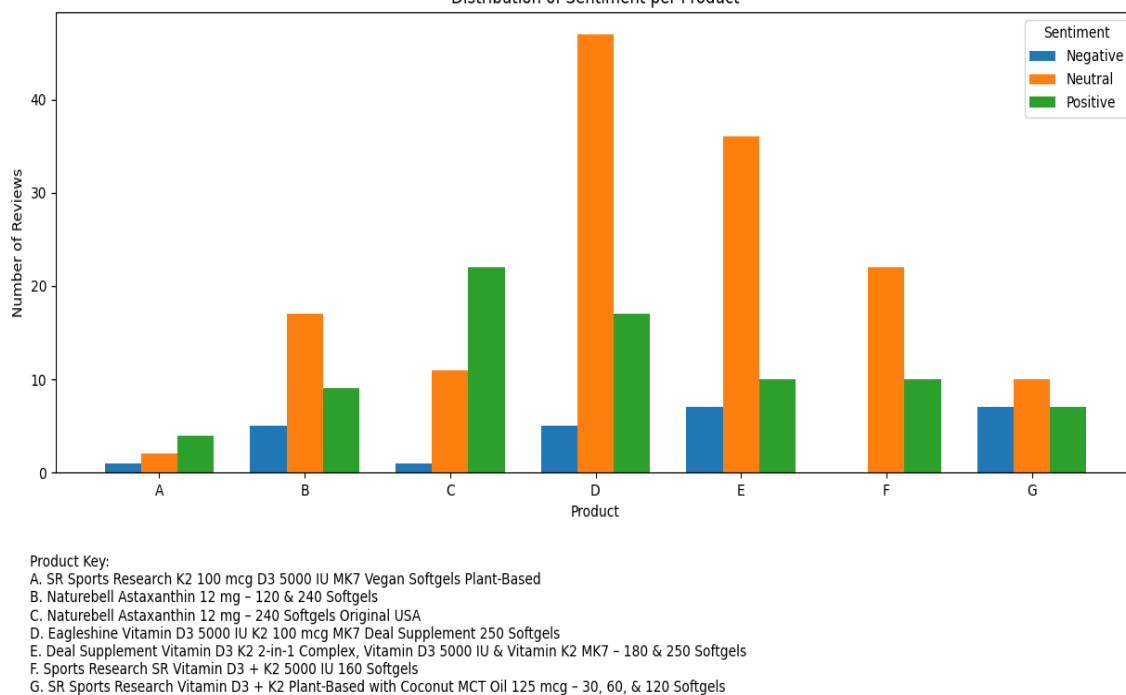
Table 3 therefore uses cautious contextual descriptions. The bigrams identify themes that co-occurred with negative reference labels, including expected benefits, authenticity, comparisons, product names, and expressions of hope or politeness. The table does not treat the phrases themselves as intrinsically negative.

3.1.4 Keyword-Based Aspect Analysis

Keyword-based aspect analysis grouped selected expressions into descriptive themes. It differs from conventional ABSA, which typically identifies aspects and estimates sentiment polarity for those aspects; recent reviews emphasize the diversity of datasets, domains, and methods used in ABSA research (Chebolu et al., 2023; Hua et al., 2024). The present study did not automatically extract aspects or estimate an independent sentiment polarity for each aspect.

Across the lexical analyses, authenticity, packaging, delivery, expiration dates, product condition, and seller-related language recurred more visibly than explicit regulatory terminology. These results describe the content of post-purchase communication and do not identify causal determinants of consumer evaluation.

The absence of “halal” and the single BPOM-related expression indicate limited explicit regulatory discourse in this dataset. Consumers may nevertheless have obtained or considered regulatory information through sources not reproduced in their reviews.

Figure 8. Distribution Of Sentiment Per Product

Source: Data processed by the researchers using Python-based textual analysis, 2025.

Figure 8 shows variation in the reference-label distribution across the seven products. Neutral labels were most numerous for Products D and E, while Product C had the largest positive component. Products E and G had comparatively more negative-labelled reviews than the other products. These differences describe how the 250 labelled observations were distributed across products. Review texts associated with the distributions mentioned authenticity, effectiveness, formulation, packaging, and transaction experiences, but the analysis does not establish those themes as causes of the observed labels.

3.1.5 Model Validation

Tables 4 and 5 summarize the model comparison: Table 4 reports overall and macro/weighted performance, whereas Table 5 presents class-specific precision, recall, F1-score, and support. The comparative evaluation demonstrates substantial variation in performance across the four classifiers. Multinomial Naive Bayes showed a strong tendency to assign observations to the neutral class. Its confusion matrix indicates that all five negative observations and twelve of the sixteen positive observations were misclassified as neutral. Although the model achieved an overall accuracy of 0.66, its macro F1-score was only 0.39, indicating weak performance across the three sentiment classes.

Logistic Regression produced a modest improvement, achieving an accuracy of 0.68 and a macro F1-score of 0.43. Nevertheless, the model also failed to correctly classify any of the five negative observations and continued to exhibit a strong tendency toward the neutral class.

LinearSVC achieved an accuracy of 0.78 and a macro F1-score of 0.63. It correctly identified 27 of 29 neutral reviews and 11 of 16 positive reviews, but only one of five negative reviews. The minority negative class therefore remained difficult to detect.

Random Forest achieved the strongest overall results, with an accuracy of 0.80, a macro F1-score of 0.65, and a weighted F1-score of 0.78. It correctly classified 28 of 29 neutral reviews

and 11 of 16 positive reviews. Similar to LinearSVC, however, Random Forest correctly identified only one of five negative reviews. Its negative-class precision of 1.00 should therefore be interpreted cautiously: the high precision results from the correctness of the limited negative predictions made by the model, while its recall of 0.20 demonstrates that most actual negative observations remained undetected. Random Forest and LinearSVC provided the strongest balanced performance within this experiment. However, their negative-class recall of 0.20 and the small number of negative test observations limit the stability of that comparison.

Table 4. Comparative Performance of Sentiment Classification Models

Classification Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted F1-score
Multinomial Naive Bayes	0.66	0.54	0.42	0.39	0.58
Logistic Regression	0.68	0.47	0.45	0.43	0.62
LinearSVC	0.78	0.84	0.61	0.63	0.76
Random Forest	0.80	0.87	0.62	0.65	0.78

Source: Data processed by the researcher

Table 5. Class-Specific Performance of Sentiment Classification Models

Model	Sentiment	Precision	Recall	F1-score	Support
Multinomial Naive Bayes	Negative	0.00	0.00	0.00	5
	Neutral	0.63	1.00	0.77	29
	Positive	1.00	0.25	0.40	16
Logistic Regression	Negative	0.00	0.00	0.00	5
	Neutral	0.67	0.97	0.79	29
	Positive	0.75	0.38	0.50	16
LinearSVC	Negative	1.00	0.20	0.33	5
	Neutral	0.79	0.93	0.86	29
	Positive	0.73	0.69	0.71	16
Random Forest	Negative	1.00	0.20	0.33	5
	Neutral	0.78	0.97	0.86	29
	Positive	0.85	0.69	0.76	16

Source: Data processed by the researcher

The comparison confirms that overall accuracy can obscure minority-class weakness. Random Forest is interpreted as the strongest model in the present split, closely followed by LinearSVC, rather than as a universally superior classifier.

3.2 Discussion

The predominance of neutral labels shows that Shopee's review area was frequently used for descriptive, informational, expectational, or uncertain communication. This pattern is consistent with the broader role of reviews as consumer-to-consumer information that

supplements seller-provided content (Chen & Xie, 2008; Mudambi & Schuff, 2010; Filieri, 2015). In this corpus, that information commonly concerned authenticity, condition, packaging, delivery, and expiration dates.

Authenticity-related expressions appeared across all three sentiment categories, demonstrating why isolated words cannot be assigned a fixed polarity. The same term could confirm authenticity, request reassurance, or express doubt depending on the complete review. Keyword, word cloud, and n-gram outputs are therefore interpreted as evidence of lexical recurrence and co-occurrence, not as determinants of consumer evaluation.

The exceptionally limited occurrence of explicit halal and BPOM terminology is the central substantive result. It indicates that formal regulatory attributes were rarely verbalized in the available post-purchase reviews compared with practical product and transaction themes. It does not show that consumers were unaware of, indifferent to, or uninfluenced by regulatory information, because those propositions require survey, interview, experimental, or behavioural data.

This scope differentiates the study from intention-based halal consumer research. Studies grounded in the Theory of Planned Behavior or related frameworks examine reported attitudes, knowledge, religiosity, trust, or intended behaviour (Ajzen, 1991; Ali et al., 2018; Izza & Akbar, 2023; Tedjakusuma et al., 2023; Bhutto et al., 2024). The present study instead analyses communication generated after verified transactions. These approaches are complementary because they address different empirical questions.

The model comparison adds a methodological finding: Random Forest and LinearSVC performed substantially better than Multinomial Naive Bayes and Logistic Regression on the same held-out observations, but all models were constrained by class imbalance. With only five negative test cases, each classification decision materially changes negative-class recall. Macro averages and class-specific metrics are therefore more informative than accuracy alone.

The findings have bounded practical implications. Marketplaces and sellers may improve the visibility, consistency, and traceability of halal and BPOM fields on product pages, particularly for imported health products. Regulators may use aggregated review mining as a supplementary means of observing what consumers discuss, while retaining official databases and surveillance as the basis for compliance decisions. These recommendations address information visibility and monitoring; the study does not test whether they would change trust or purchase intention.

More broadly, consumer reviews should not be treated as substitutes for formal verification. Positive sentiment, authenticity claims, intact packaging, or barcode information may describe consumer experience, but cannot establish product safety, legality, or certification. This distinction is consistent with research on risks in online supplement markets and with the need to preserve halal integrity and traceability across product information and supply-chain processes (White, 2022; Crawford et al., 2022; Rathmann & Seifert, 2024; Wilson & Liu, 2010; Susanty et al., 2025).

4. Conclusion

This study analysed 250 post-purchase reviews of seven imported vitamin products whose Shopee listings did not clearly display halal certification or BPOM registration information during the observation period. Neutral labels accounted for 58.0% of the dataset, positive

labels for 31.6%, and negative labels for 10.4%. Authenticity, packaging, delivery, expiration dates, product condition, and seller-related experiences recurred more often than explicit regulatory terminology; “halal” did not occur and one BPOM-related expression was identified. Random Forest achieved the strongest performance within the held-out experiment (accuracy 0.80; macro F1 0.65), followed by LinearSVC (accuracy 0.78; macro F1 0.63). Both nevertheless achieved only 0.20 recall for the five negative test cases. The comparison demonstrates the importance of macro and class-specific measures in small, imbalanced review datasets.

The contribution lies in separating post-purchase communication from two different questions that the data cannot answer: what caused the original purchase and whether a product holds official regulatory status. The findings show limited explicit regulatory discourse in the observed reviews, not limited consumer concern. Review mining can therefore complement intention-based research and formal oversight without replacing either.

5. Limitations and Future Research

Limitations

The findings are bounded by the small purposive dataset of 250 reviews from seven products, 12 stores, and one marketplace. They describe the analytical corpus and should not be generalized as population estimates for Shopee consumers, other marketplaces, or other health-product categories. The label distribution was imbalanced and the held-out test set contained only five negative cases. Model evaluation also relied on a single stratified 80:20 split. Although raw text was split before training-only TF-IDF fitting, the small test set makes comparative metrics sensitive to individual observations. Repeated stratified cross-validation and imbalance-handling methods would provide more stable estimates.

One researcher assigned both the sentiment labels and manual keyword groupings. Explicit coding rules and full-review contextual interpretation improved procedural transparency, but no independent coder or agreement statistic was available. The reference labels and thematic groupings may therefore contain coder-specific judgement, particularly for mixed, informal, abbreviated, or uncertain expressions. Finally, explicit keyword occurrence cannot capture every implicit regulatory concern, and post-purchase reviews cannot reveal unexpressed pre-purchase considerations. The product-verification procedure assessed whether information was visible or identifiable from the observed sources; it did not independently determine legal, safety, BPOM, or halal status.

Future Research

Future research should use larger, more balanced, independently coded datasets and report inter-coder agreement. Repeated stratified cross-validation, class weighting, resampling, and contextual language models could be compared with the conventional classifiers used here, particularly for minority negative expressions. Mixed-method research could combine review mining with surveys, interviews, experiments, or regulatory database studies to examine halal awareness, religiosity, perceived health risk, digital literacy, trust, and purchase processes directly. Cross-platform and longitudinal designs could also test whether regulatory information becomes more visible as certification and marketplace practices evolve.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the revision of this manuscript, the authors used ChatGPT (OpenAI) to improve language, readability, terminology consistency, and presentation. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgement

The authors express their deepest gratitude to the Research Institute for Research and Community Service (LPPM) of Hasanuddin University, Makassar, Indonesia, for their support and facilities. We also thank the anonymous reviewers for their insightful comments and constructive feedback throughout the review process. This research would not have been possible without the availability of consumer review data on the Shopee platform, and we therefore extend our appreciation to all those who contributed to this data.

Reference

- Abighail, B. M. D., Fachrifansyah, Firmanda, M. R., Anggreainy, M. S., Harvianto, & Gintoro. (2023). Sentiment analysis e-commerce review. *Procedia Computer Science*, 227, 1039–1045. <https://doi.org/10.1016/j.procs.2023.10.613>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ali, A., Xiaoling, G., Sherwani, M., & Ali, A. (2018). Antecedents of consumers' Halal brand purchase intention: An integrated approach. *Management Decision*, 56(4), 715–735. <https://doi.org/10.1108/MD-11-2016-0785>
- Badan Pengawas Obat dan Makanan Republik Indonesia. (2023). *Cek produk: Database registrasi produk*. Retrieved from <https://cekbpom.pom.go.id>
- Badan Pengawas Obat dan Makanan Republik Indonesia. (2025, April 30). *BPOM ungkap 39.642 tautan ilegal produk obat dan makanan di media daring*. Retrieved from <https://siber.pom.go.id/berita/bpom-ungkap-39-642-tautan-ilegal-produk-obat-dan-makanan-di-media-daring>
- Badan Penyelenggara Jaminan Produk Halal. (2025). *Produk bersertifikat halal tembus 9,6 juta, BPJPH terus intensifkan kolaborasi dalam & luar negeri*. Retrieved from <https://bpjph.halal.go.id/read/produk-bersertifikat-halal-tembus-9-6-juta-bpjph-terus-intensifkan-kolaborasi-dalam-and-luar-negeri>
- Bhutto, M. Y., Rütelioné, A., & Vienažindienė, M. (2024). Investigating EWOM and halal product knowledge on Gen Z's halal cosmetics purchase intentions in Pakistan. *Journal of Islamic Marketing*, 15(11), 3266–3281. <https://doi.org/10.1108/JIMA-09-2023-0292>
- Chebolu, S. U. S., Dernoncourt, F., Lipka, N., & Solorio, T. (2023). A review of datasets for aspect-based sentiment analysis. In *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics* (pp. 611–628). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.ijcnlp-main.41>

- Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, 54(3), 477–491. <https://doi.org/10.1287/mnsc.1070.0810>
- Crawford, C., Avula, B., Lindsey, A. T., Walter, A., Katragunta, K., Khan, I. A., & Deuster, P. A. (2022). Analysis of select dietary supplement products marketed to support or boost the immune system. *JAMA Network Open*, 5(8), e2226040. <https://doi.org/10.1001/jamanetworkopen.2022.26040>
- Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, 68(6), 1261–1270. <https://doi.org/10.1016/j.jbusres.2014.11.006>
- Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, 90(2), 217–232. <https://doi.org/10.1016/j.jretai.2014.04.004>
- Golovan, T. V. (2024). The on-line turnover of biologically active additives. Part 2. Customs control for the turnover of imported dietary supplements on marketplaces as a measure to protect the life and health of citizens. *Problemy Sotsial'noi Gigieny, Zdravookhraneniia i Istorii Meditsiny*, 32(3), 363–369. <https://doi.org/10.32687/0869-866X-2024-32-3-363-369>
- Handoyo, S. (2024). Purchasing in the digital age: A meta-analytical perspective on trust, risk, security, and e-WOM in e-commerce. *Heliyon*, 10(8), e29714. <https://doi.org/10.1016/j.heliyon.2024.e29714>
- Hua, Y. C., Denny, P., Wicker, J., & Taskova, K. (2024). A systematic review of aspect-based sentiment analysis: Domains, methods, and trends. *Artificial Intelligence Review*, 57, Article 296. <https://doi.org/10.1007/s10462-024-10906-z>
- Izza, N. N., & Akbar, N. (2023). Uncovering halal label food purchase intention behavior in Indonesia: A Twitter analytics and SEM-based investigation. *Jurnal Ekonomi Syariah Teori Terapan*, 10(6), 551–570. <https://doi.org/10.20473/vol10iss20236pp551-570>
- Khan, Z., Khan, A., Nazish, M., & Arwab, M. (2026). A fusion of TAM and TPB with halal product knowledge and religiosity to understand online purchase intentions of halal cosmetics among Muslims Gen Z. *Journal of Islamic Marketing*, 17(4), 1415–1437. <https://doi.org/10.1108/IIMA-10-2024-0456>
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions* (2nd ed.). Cambridge University Press.
- Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University – Computer and Information Sciences*, 36, Article 102048. <https://doi.org/10.1016/j.jksuci.2024.102048>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Mou, J., & Benyoucef, M. (2021). Consumer behavior in social commerce: Results from a meta-analysis. *Technological Forecasting and Social Change*, 167, Article 120734. <https://doi.org/10.1016/j.techfore.2021.120734>

- Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34(1), 185–200. <https://doi.org/10.2307/20721420>
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* (pp. 79–86). Association for Computational Linguistics. <https://doi.org/10.3115/1118693.1118704>
- Rathmann, A.-M., & Seifert, R. (2024). Vitamin A-containing dietary supplements from German and US online pharmacies: Market and risk assessment. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 397(9), 6803–6820. <https://doi.org/10.1007/s00210-024-03050-6>
- Sharma, N. A., Shawkat Ali, A. B. M., & Kabir, M. A. (2025). A review of sentiment analysis: Tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics*, 19(3), 351–388. <https://doi.org/10.1007/s41060-024-00594-x>
- Shopee. (2025). *Seller center feature for entering halal certification numbers and BPOM registration numbers*. Retrieved from <https://seller.shopee.co.id>
- Susanty, A., Puspitasari, N. B., Jie, F., Akhsan, F. A., & Jati, S. (2025). Consumer acceptance of halal food traceability systems: A novel integrated approach using modified UTAUT and DeLone & McLean models to promote sustainable food supply chain practices. *Cleaner Logistics and Supply Chain*, 15, Article 100226. <https://doi.org/10.1016/j.clscn.2025.100226>
- Tedjakusuma, A. P., Au Yong, H. N., Andajani, E., & Mohamad, Z. Z. (2023). Intention to purchase halal health supplement online: Lessons learned from the health crisis. *Heliyon*, 9(9), e19840. <https://doi.org/10.1016/j.heliyon.2023.e19840>
- White, C. M. (2022). Continued risk of dietary supplements adulterated with approved and unapproved drugs: Assessment of the US Food and Drug Administration's Tainted Supplements Database 2007 through 2021. *The Journal of Clinical Pharmacology*, 62(8), 928–934. <https://doi.org/10.1002/jcph.2046>
- Wilson, J. A. J., & Liu, J. (2010). Shaping the Halal into a brand? *Journal of Islamic Marketing*, 1(2), 107–123. <https://doi.org/10.1108/17590831011055851>