JMK (Jurnal Manajemen dan Kewirausahaan)
http://ejournal.uniska-kediri.ac.id/index.php/ManajemenKewirausahaan
JMK 10 (3) 2025, 141-155
P-ISSN 2477-3166 E-ISSN 2656-0771

Data-Driven Approach to Managing Best-Selling Beauty Categories: Price, Rating, Review, and Stock

*****Indah Safira Diroatmodjo
Master of Science in Management, Faculty of Economics and Business, Padjadjaran University
indah21006@mail.unpad.ac.id

Samidi
Master of Science in Management, Faculty of Economics and Business, Padjadjaran University, Master of Computer Science, Budi Luhur University
samidiunpad@gmail.com *samidi@budiluhur.ac.id*

Abstract

The beauty industry in Indonesia is experiencing rapid growth, particularly through e-commerce platforms like *Tokopedia*. Many businesses still rely on intuition for product management, including decisions related to stock and pricing. This study develops a machine learning-based classification model to identify beauty products with high sales potential on *Tokopedia*, considering factors such as price, rating, review count, and stock availability. Ten classification algorithms are applied, including Naive Bayes, SVM, K-Nearest Neighbors, Decision Tree, Random Forest, XGBoost, LightGBM, CatBoost, Extra Trees, and Multi-Layer Perceptron (MLP). The data is processed using *Python* on *Google Colab*. The results show that ensemble algorithms, particularly Random Forest, LightGBM, and Extra Trees, provide prediction accuracy above 91% and are highly effective in predicting best-selling products. Based on this model, businesses can optimize stock and pricing management to ensure that best-selling products are always available, thereby improving operational efficiency in a highly competitive market. This research offers a data-driven solution for more strategic and evidence-based product management on e-commerce platforms.

Keywords: e-commerce, machine learning, classification, stock management, best-selling products

DOI	:	http://dx.doi.org/10.32503/jmk.
Article History	:	Article received (July 2025); revised (August 2025); accepted (September 2025)
Email Co-Author	:	*indah21006@mail.unpad.ac.id*

Introduction

The global beauty industry is rapidly expanding and is expected to reach USD 677.19 billion by 2025, with Indonesia emerging as a key market (Statista, 2025a). In Indonesia, e-commerce platforms such as *Tokopedia, Shopee, and Bukalapak* play a crucial role in distributing beauty products, with *Tokopedia* leading the sector (Statista, 2025b). However, with intense competition and a wide array of products, relying solely on intuition for business decisions is no longer sufficient. There is an increasing necessity to adopt a systematic, predictive, data-driven approach for effective strategic and competitive product management (Pocchiari et al., 2024; Poláček et al., 2024).

Currently, many sellers on e-commerce platforms like *Tokopedia* still manage inventory and pricing based on intuition, despite having access to data on ratings, reviews, and inventory on the platform. (Dabestani et al., 2025; Oancea, 2023). A data-driven approach to product category management allows companies to make more precise and evidence-based decisions by utilizing transaction data, user reviews, and product attributes. (Chaube et al., 2025; Nguyen et al., 2024; T. C. Wang et al., 2023). However, there is an urgent need to immediately utilize this data to prevent falling behind in increasingly competitive markets. (Schmitt, 2023; Valencia-Arias et al., 2024). This gap indicates that, despite the availability of data, there remains no perfect method for incorporating key variables into a systematic classification model.

This research is based on three significant gaps in the literature. First, at the input stage, no studies have simultaneously highlighted the influence of price, rating, reviews, the number of ratings, and stock on the classification of best-selling products, despite these five variables being important indicators in consumer purchasing decisions. (Nikose et al., 2022; X. Wang et al., 2024; X. Wu et al., 2024). Second, during the process stage, although various prediction models, such as ensemble learning and deep learning, have been developed, no study has systematically compared the effectiveness of ten classification algorithms —both classical and modern —in the context of beauty products in Indonesia. (Agarwal & Yadav, 2024; Sharma et al., 2022; J. Wu et al., 2024). Third, from the output perspective, most existing classification models have not generated practical managerial advice on product stock and pricing management for e-commerce businesses. (D. Li et al., 2023; Rios & Vera, 2023; Yang et al., 2024).

This study aims to develop a machine learning-based classification model to identify beauty products with the potential to become bestsellers on *Tokopedia*. To accomplish this, the study incorporates five primary predictor variables: price, rating, number of reviews, number of ratings, and stock.

Price, rating, number of reviews, and stock are key factors influencing product sales, especially in the e-commerce context (Hirsch et al., 2025; Nikose et al., 2022; X. Wang et al., 2024). Competitive pricing can increase sales volume because consumers tend to choose products with more affordable prices (Nie et al., 2024; Q. Zhang & Xiao, 2024). High ratings often indicate consumer confidence in a product's quality, which boosts the chances of it becoming a bestseller. (X. Wang et al., 2024; Zaghloul et al., 2024). Furthermore, a large number of reviews indicate that the product has been widely purchased and discussed by consumers, which increases the likelihood of its popularity and potential to become a bestseller

(Pocchiari et al., 2024; Yuhsiang & Lichung, 2024). Having sufficient stock is also crucial to ensure product availability and prevent lost sales, which can occur when consumers are unable to purchase the products they want due to stock shortages (Trapero et al., 2024; Xu et al., 2023).

Based on the integration of these variables, the novelty of this study lies in three key aspects. First, it combines five important variables into a single classification model specifically for the beauty product category on *Tokopedia*, an area that has not been extensively explored in previous research (Nikose et al., 2022; Y. Wang & Zhang, 2023; X. Wu et al., 2024). Second, comparing ten machine learning algorithms systematically in the context of local Indonesian e-commerce highlights the need for a comprehensive predictive approach (Agarwal & Yadav, 2024; Nikose et al., 2022; Y. Wang & Zhang, 2023). Third, this model was developed not only for academic purposes but also to support operational business decisions. It is designed to be directly usable as a decision-making tool at the operational level of online stores in Indonesia, especially for managing the beauty product category.

In implementing these models, machine learning-based classification methods are essential for predicting product sales in e-commerce. Several traditional algorithms, such as Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Random Forest, are commonly used to develop predictive models. At the same time, modern algorithms such as XGBoost, LightGBM, CatBoost, Extra Trees, and Multi-Layer Perceptron (MLP) offer enhanced accuracy and efficiency, particularly when handling large-scale and complex data. Combining these traditional and modern algorithms enables a comprehensive assessment of classification performance across various data contexts, including identifying the characteristics of products that have the potential to become best-sellers (Agarwal & Yadav, 2024; Hincapié-López et al., 2024; Sharma et al., 2022).

E-commerce platforms *like Tokopedia*, *Shopee*, and *Bukalapak* utilize recommendation algorithms and predictive analytics to enhance the visibility of best-selling products, making it easier for consumers to find relevant items based on their past purchasing preferences (Nguyen et al., 2024; Valencia-Arias et al., 2024). By integrating big data and machine learning analytical techniques, these platforms can track market trends, understand consumer behavior, and deliver more precise product recommendations (Nguyen et al., 2024; X. Wang et al., 2024). Applying sales prediction models, such as those using algorithms like XGBoost and LightGBM, helps optimize inventory and pricing while reducing losses from stockouts (Mahin et al., 2025; Tang et al., 2023). By utilizing sentiment analysis techniques to analyze product reviews, e-commerce platforms can gain deeper insights into consumer preferences and dynamically adjust their product offerings. (Le et al., 2024; Pocchiari et al., 2024). The implementation of this technology helps the platform predict and manage best-selling products efficiently, as well as design more effective marketing strategies (Chaube et al., 2025; J. Wu et al., 2024).

This research offers practical benefits by helping business owners identify potential best-selling products, manage inventory efficiently, and set prices based on product performance data (Nie et al., 2024; Zhuang & Xu, 2025). Additionally, the chosen algorithms were selected based on their effectiveness in managing large and imbalanced datasets, with a focus on speed, accuracy, and the interpretability of the results (Le et al., 2024; Mahin et al., 2025). This approach also aligns with the principles of data-driven business models (DDBM), which highlight the importance of data analytics in strategic decision-making and value creation in contemporary business management (Dabestani et al., 2025). To develop a strategy based on the classification results, this research draws on the strategic management process outlined by Wheelen and Hunger, which encompasses internal analysis, strategy formulation, implementation, and strategy evaluation (Khalif & Slim, 2024). Meanwhile, the proposed stock and pricing management strategy aligns with Kotler and Keller's marketing principles, as supported by Gonen et al. (2024), who show that value-based pricing and segmented product strategies enhance both profitability and customer loyalty (Gonen et al., 2024).

Methodology

This research is a quantitative and descriptive study with a predictive approach that aims to develop a data-driven classification model to identify beauty products with the potential to become bestsellers on the *Tokopedia* platform. The primary focus of this study is to assess the impact of variables such as price, rating, number of reviews, number of ratings, and stock on product sales, and to create a classification model that categorizes products into best-selling and non-best-selling groups based on historical sales data.

The research data was collected through web scraping using the Web Scraper extension on *Google Chrome*. The scraping focused on the beauty category on *Tokopedia*, encompassing various subcategories, including facial makeup, skincare, and hair products. A total of 2,790 entries were gathered, each representing a single product and including attributes like product name, price (in Rupiah), number of units sold, average user rating, number of ratings, number of reviews, and stock availability. Table 1 presents five sample entries from the raw dataset, highlighting key variables used in the analysis, including price, ratings, reviews, stock, and sales.

Table 1. Five-Row Sample of Product Data from *Tokopedia*

Product	Price	Rating	Total Rating	Total Review	Stock	Sales
Lip ICE Sheer Color Lip Balm	21.500	4.9	739	139	241	3.000
Hanasui Lip Cream Boba Edition	23.760	4.9	60	15	100	100
Nuface Nu Matte Lip Cream	20.600	4.6	88	74	47	500
Azarine Tinted Lippie Cake	38.200	4.9	149	149	23	750
Lipgloss Maybelline	75.000	5.0	62	6	990	500

Source: Author's Processed (2025)

After data collection, a cleaning and transformation process was carried out using *Python* in *Google Colab*. This step involved removing duplicate data and empty values, converting numeric data types, and applying logarithmic transformations to numeric variables with highly skewed distributions, such as price, number of ratings, number of reviews, stock, and quantity sold. These transformations aimed to stabilize variance, reduce outliers, and enhance the performance of the classification model. Additionally, normalization and standardization were performed before the modeling process.

Next, the top 25% of products with the highest sales value (Q4 of log_sold) were grouped as best-selling products (label 1), while the rest were labeled 0, following the performance-based classification segmentation method described by Vijai Kotu and the leading subset classification practice in research by Chaube et al. (2025), which sets boundaries based on ranking and the highest predictive score (Chaube et al., 2025).

To provide a complete overview of the research process, the following is a conceptual framework of the steps taken in this study.

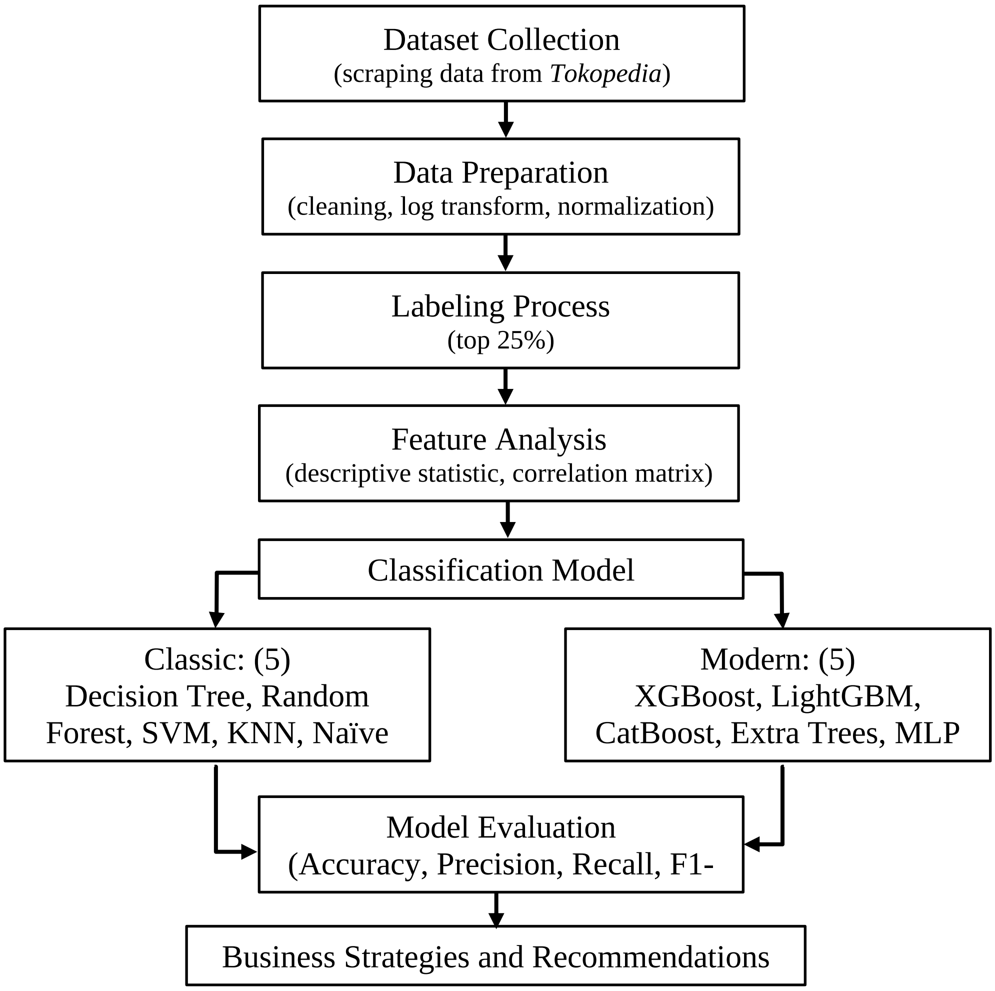

Figure 1. Process Conceptual Framework
(Source: Author's Processed)

The modeling phase was carried out solely using a classification approach. The dataset was split into two parts: 80% for training and 20% for testing. Ten classification algorithms were employed to develop and compare the models, which are categorized into two main groups: classical and modern algorithms. Classical algorithms include Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Random Forest. Modern algorithms include XGBoost, LightGBM, CatBoost, Extra Trees, and Multi-Layer Perceptron (MLP) Neural Network.

Model evaluation was conducted using accuracy, precision, recall, and F1-score metrics, focusing on performance in classifying best-selling products (class 1). Furthermore, feature importance analysis was conducted to determine which variables were most influential in predicting product sales. The best model was then used as the basis for developing strategic recommendations for businesses, including more efficient, data-driven pricing and inventory management.

Result and Discussion

Result

This research was motivated by the need for business owners to identify beauty products with the potential to become bestsellers on *Tokopedia*, thereby supporting more accurate decision-making in data-driven inventory management and pricing. To achieve this, a machine learning-based classification approach was used, considering key numeric variables such as price, rating, number of reviews, number of ratings, and stock availability.

The dataset, comprising 2,790 web-scraped products, underwent several preprocessing steps, including removing duplicates and empty values, converting data types, and applying a logarithmic transformation to numeric variables to reduce skewness. The target variables were classified into two classes based on the upper quartile (the highest 25%) of the transformed sales data. This strategy draws on the performance segmentation approach proposed by Chaube et al. (2025).

Descriptive statistics provide an overview of the distribution and variability of the key variables used in this study. Detailed data statistics are shown in Table 2.

Table 2. Descriptive Statistics

=== Descriptive Statistics ===

	price	rating	total_rating	total_review	stock
count	2.789000e+03	2789.000000	2789.000000	2789.000000	2789.000000
mean	1.281762e+05	4.863464	178.313015	97.116529	3218.313015
std	4.541574e+05	0.165286	588.356291	350.171053	20948.982915
min	0.000000e+00	3.500000	1.000000	0.000000	0.000000
25%	1.490000e+04	4.800000	12.000000	4.000000	2.000000
50%	3.590000e+04	4.900000	44.000000	17.000000	22.000000
75%	9.900000e+04	4.900000	130.000000	79.000000	150.000000
max	1.000000e+07	5.000000	12861.000000	6152.000000	666650.000000

	sales
count	2789.000000
mean	90.676632
std	121.806012
min	1.000000
25%	21.000000
50%	69.777778
75%	100.000000
max	750.000000

Source: Author's Processed (2025)

Based on Table 2, the average product price is around Rp128,000, with an extensive range ranging from Rp0 to Rp10 million. This indicates the presence of extreme values or the possibility of input errors. The average product rating is relatively high, at 4.86 on a scale of 1–5, with a low standard deviation (0.165), thus indicating relatively even customer satisfaction. However, the number of ratings and reviews exhibits significant variations, with averages of 178 and 97, respectively, and maximum values of 12,861 and 6,152. Product inventory also varies widely, ranging from 0 to 666,000 units, with an average of approximately 3,218 units. The number of product sales ranges from 1 to 750 units, with 75 percent of products selling no more than 100 units. This uneven distribution suggests the need for a logarithmic transformation to normalize the data and prepare it for modeling.

Correlation analysis revealed that the number of ratings and reviews had a positive correlation with sales, at 0.29 and 0.26, respectively. Conversely, price and ratings had a negative correlation with sales, indicating that the intensity of customer interactions, such as reviewing, had a greater impact on sales than high price or ratings alone. These correlation results are shown in Figure 2 and address the first question in the introduction regarding the relevance of predictors in explaining sales.

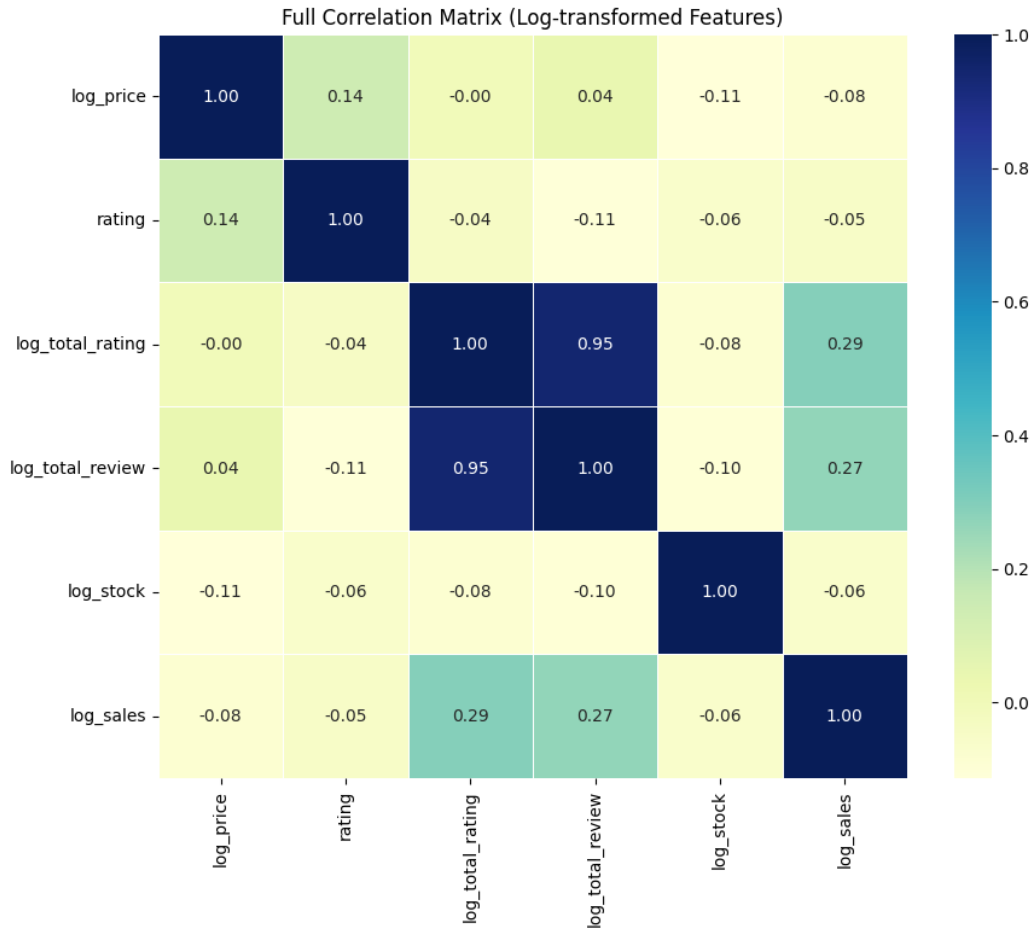

Figure 2. Correlation Matrix
(Source: Author’s Processed (2025))

The modeling process was carried out using ten classification algorithms divided into two categories: five classical algorithms (Naive Bayes, SVM, KNN, Decision Tree, and Random Forest) and five modern algorithms (XGBoost, LightGBM, CatBoost, Extra Trees, and MLP Neural Network). Evaluation was carried out on test data (20%) using accuracy, precision, recall, and f1-score metrics, specifically for the "best-selling" product class.

The analysis results show that the three algorithms with the best overall performance were Random Forest, LightGBM, and Extra Trees, each achieving an accuracy above 91% and a balanced F1 score. This finding provides an answer to the second question regarding the effectiveness of classification models in the e-commerce context. Furthermore, this finding also addresses the third question, as the model results can serve as a basis for developing more strategic stock and pricing management recommendations for businesses. A summary of the performance of all classification models is presented in Table 3.

Table 3. Performance Summary of 10 Classification Models

Model	Category	Accuracy	Precision (1)	Recall (1)	F1-Score (1)
Naive Bayes	Classic	81%	0.65	0.77	0.71
SVM	Classic	87%	0.79	0.77	0.78
K-Nearest Neighbors	Classic	86%	0.80	0.71	0.76
Random Forest	Classic	92%	0.87	0.87	0.87
Decision Tree	Classic	88%	0.79	0.83	0.81
XGBoost	Modern	91%	0.85	0.84	0.84
LightGBM	Modern	92%	0.87	0.85	0.86
CatBoost	Modern	91%	0.85	0.83	0.84
Extra Trees	Modern	91%	0.86	0.85	0.85
MLP (Neural Network)	Modern	89%	0.84	0.80	0.82

Source: Author’s Processed (2025)

Discussion

The results of the correlation analysis indicate a significant positive relationship between the number of ratings and sales, with a correlation coefficient of 0.29, and between the number of reviews and sales, with a correlation coefficient of 0.26. This finding is consistent with the research of Wu et al. (2025), which states that product reviews and ratings have a significant impact on sales, where higher customer interaction (through reviews and ratings) increases product sales. Conversely, price and ratings show a negative correlation with sales, indicating that higher prices tend to reduce sales, as found by Zhang & Yang (2023). These results suggest that, in the context of e-commerce, user interaction factors, as reflected in reviews, are stronger in influencing purchasing decisions compared to price or product ratings alone.

The modeling process using ten classification algorithms, including Naive Bayes, SVM, KNN, Decision Tree, Random Forest, XGBoost, LightGBM, CatBoost, Extra Trees, and MLP Neural Network, showed that ensemble algorithms such as Random Forest, LightGBM, and Extra Trees provided the best performance with accuracy above 91% and balanced F1-score for the "best-selling" class. This finding is in line with the results of research by Chaube et al. (2025), which emphasized that ensemble-based models, such as Random Forest and boosting models such as LightGBM, are more effective in handling complex data and interactions between variables, which are very relevant in the context of e-commerce. Wu et al. (2025) also confirmed that ensemble algorithms provide better results in sales prediction than classical models.

Furthermore, these results also show that the SVM and Naive Bayes models perform lower, especially in identifying best-selling products, with f1-scores for the "best-selling" class of 0.71 and 0.78, respectively. This suggests that classical algorithms are less effective in handling imbalanced data, which is common in e-commerce datasets (Oancea, 2023).

Based on the analysis and prediction model results, several managerial strategies can be identified to improve product management in e-commerce, particularly in the beauty product category. These strategies are formulated based on the strategic management process, as outlined by Wheelen and Hunger, which encompasses the stages of strategy analysis, formulation, implementation, and evaluation (Khalif & Slim, 2024). The following table summarizes the suggested strategies, implementation steps, the person responsible for each strategy, and the relevant monitoring and evaluation plan:

Table 4. Summary of Strategy Recommendations from Findings

Strategy Formulation	Strategy Plan	Strategy Implementation	Person In Charge	Monitoring & Evaluation Plan	Previous Journal
Price	Price adjustment for best-selling products	1. Dynamic pricing based on market and demand data	Marketing Team - Pricing Strategy	Monitoring market price changes and consumer response	(Nie et al., 2024; Poláček et al., 2024; C. Zhang et al., 2023)
		2. Implementation of price bundling with other products	Finance Team - Discount Management	Profit margin monitoring after bundling	(Hemmati et al., 2023; Nie et al., 2024; C. Zhang et al., 2023)
		3. Implementation of discount pricing strategies based on seasonality and trends	Marketing Team - Seasonal Campaigns	Measuring the impact of discounts on sales and profits	(Guo & Mai, 2024; Ni, 2023; Rios & Vera, 2023; Trapero et al., 2024)
Rating	Increase product rating with positive feedback	1. Encourage customers to leave positive reviews	Marketing Team - Customer Communications	Sentiment analysis of product reviews	(J. Li et al., 2024; Y. Wu et al., 2024; B. Zhang et al., 2024)
		2. Provide incentives for positive reviews	Marketing Team - Loyalty Programs	Evaluate the impact of incentives on rating improvement	(Y. Wu et al., 2024; B. Zhang et al., 2024; C. Zhang et al., 2023)
		3. Develop campaigns to	Communications Team - Social Media	Customer conversion and	(Henderi et al., 2023; B. Zhang et al., 2024; C.

Strategy Formulation	Strategy Plan	Strategy Implementation	Person In Charge	Monitoring & Evaluation Plan	Previous Journal
		promote high-rated reviews		engagement analysis	Zhang et al., 2023)
Total Rating	Focus on high-rated products	1. Promote high-rated products through targeted advertising	Marketing Team - Social Media & Advertising	Measuring conversion and upselling of products	(Y. Wang & Zhang, 2023; J. Wu et al., 2024; B. Zhang et al., 2024)
		2. Offer high-rated products on the main page or banner	Marketing Team - Website Management	Analyzing traffic and customer engagement	(Henderi et al., 2023; Y. Wang & Zhang, 2023; J. Wu et al., 2024)
Total Review	Increase the number of reviews to strengthen consumer trust	1. Implement features that make it easy for customers to write reviews	IT Team - Feature Development	Analyzing feature usage and its impact on reviews	(Henderi et al., 2023; Y. Wu et al., 2024; B. Zhang et al., 2024)
		2. Organize contests or giveaways to increase the number of reviews	Marketing Team - Promotion Programs	Evaluating the effectiveness of giveaways in gathering reviews	(Y. Wu et al., 2024; B. Zhang et al., 2024; C. Zhang et al., 2023; Q. Zhang & Xiao, 2024)
		3. Provide a loyalty program to encourage customers to leave reviews	Customer Service Team - Engagement	Measuring the number of new reviews from loyalty programs	(Y. Wu et al., 2024; B. Zhang et al., 2024; Q. Zhang & Xiao, 2024)
Stock	Guarantee the availability of best-selling products	1. Increase stock of best-selling products based on demand prediction	Operations Team - Stock Management	Monitoring product inventory and demand	(Mahin et al., 2025; Trapero et al., 2024; C. Zhang et al., 2023)
		2. Optimize stock management using a just-in-time system	Operations Team - Supply Chain Management	Monitoring the effectiveness of just-in-time systems	(Hou et al., 2024; Mahin et al., 2025; Yang et al., 2024; C. Zhang et al., 2023)
		3. Stock adjustment based on seasonal trend analysis and promotions	Operations Team - Inventory Management	Real-time stock analysis and its impact on sales	(Guo & Mai, 2024; Rios & Vera, 2023; Trapero et al., 2024; B. Zhang et al., 2024)

Source: Author's Processed (2025)

Based on the results of this study, several strategic recommendations can be put forward. In terms of pricing, dynamic price adjustments based on market data and demand can help improve product competitiveness. Bundling pricing and discount pricing strategies, based on seasonal trends, can also be implemented to increase sales volume, in line with the findings of Zhang & Yang (2023) and Nie et al. (2024), which demonstrate the effectiveness of bundling pricing in influencing consumer purchasing decisions. Regarding ratings and review volume, it is vital to encourage customers to leave positive reviews and utilize incentives to increase the number of reviews. Campaigns promoting high-rated reviews can also improve product visibility on the platform, as suggested by Wu et al. (2025) and Zhang et al. (2024). On the inventory side, ensuring the availability of best-selling products through efficient stock management systems, such as using just-in-time systems and seasonal trend analysis, can help reduce the risk of stockouts and improve operational efficiency (D. Li et al., 2023; Yang et al., 2024). The marketing strategy proposed in this study is based on Kotler and Keller's marketing principles and is empirically supported by Gonen et al. (2024), who highlight that value-based pricing, product portfolio alignment, and loyalty-focused offerings drive customer retention and profitability (Gonen et al., 2024).

Conclusion

This study identifies factors influencing the sales success of beauty products on *Tokopedia* using a machine learning-based classification approach. The results show that the number of ratings and reviews has a significant positive effect on sales, while price and rating show a negative correlation. Algorithms such as Random Forest, LightGBM, and Extra Trees demonstrate an accuracy of greater than 91% in predicting best-selling products, providing a basis for data-driven managerial strategies. This study recommends strategies for dynamic pricing adjustments, rating enhancements, and data-driven inventory management to improve operational efficiency and product competitiveness.

However, this study has limitations. The data is limited to the *Tokopedia* platform and the beauty product category, which restricts the generalizability of the results to other platforms or product categories. Additionally, comparisons with regression models or alternative approaches have not been conducted. External variables such as promotions and seasonal factors have not been incorporated into the model.

For future research, it is recommended to expand the scope of the analysis by including various product categories and other e-commerce platforms, as well as considering external variables that influence product sales success. The development of deep learning-based models and the use of temporal data can also improve prediction accuracy and provide valuable insights into product management in e-commerce.

References

Agarwal, N. B., & Yadav, Dr. D. K. (2024). A Comprehensive Analysis of Classical Machine Learning and Modern Deep Learning Methodologies. *International Journal of Engineering Research & Technology*, *13*(5). https://doi.org/10.17577/IJERTV13IS050275

Chaube, S., Kar, R., Gupta, S., & Kant, M. (2025). Multimodal AI framework for the prediction of high-potential product listings in e-commerce: Navigating the cold-start challenge. *Expert Systems with Applications*, *282*(March), 127524. https://doi.org/10.1016/j.eswa.2025.127524

Dabestani, R., Solaimani, S., Ajroemjan, G., & Koelemeijer, K. (2025). Exploring the enablers of data-driven business models: A mixed-methods approach. *Technological Forecasting and Social Change*, *213*(February), 124036. https://doi.org/10.1016/j.techfore.2025.124036

Gonen, L. D., Tavor, T., & Spiegel, U. (2024). Unlocking Market Potential: Strategic Consumer Segmentation and Dynamic Pricing for Balancing Loyalty and Deal Seeking. *Mathematics*, *12*(21). https://doi.org/10.3390/math12213364

Guo, Q., & Mai, Z. (2024). How do seasonal, significant events, and policies affect China's REE export prices? Based on deep learning perspective. *Resources Policy*, *96*(July), 105205. https://doi.org/10.1016/j.resourpol.2024.105205

Hemmati, M., Fatemi Ghomi, S. M. T., & Sajadieh, M. S. (2023). Separate and bundling selling strategies for complementary products in a participative pricing mechanism. *Computers and Industrial Engineering*, *177*(November 2022), 109018. https://doi.org/10.1016/j.cie.2023.109018

Henderi, H., Irawatia, R., Indra, I., Dewi, D. A., & Kurniawan, T. B. (2023). Big Data Analysis using Elasticsearch and Kibana: A Rating Correlation to Sustainable Sales of Electronic Goods. *HighTech and Innovation Journal*, *4*(3), 583–591. https://doi.org/10.28991/HIJ-2023-04-03-09

Hincapié-López, M., Vrebosch, J., Garcia-Zapirain, B., Pinzón-Reyes, E., & Pabón-Martínez, Y. V. (2024). Comparison of classical Machine Learning-based algorithms to predict Triplex Forming Oligonucleotides. *Computational and Structural Biotechnology Reports*, *1*, 100013. https://doi.org/10.1016/j.csbr.2024.100013

Hirsch, S., Guy, I., Novgorodov, S., Lavee, G., & Shapira, B. (2025). What is in a title? Characterizing product titles in e-commerce. *Expert Systems with Applications*, *287*(February), 127702. https://doi.org/10.1016/j.eswa.2025.127702

Hou, L., Nie, T., & Zhang, J. (2024). Pricing and inventory strategies for perishable products in a competitive market considering strategic consumers. *Transportation Research Part E: Logistics and Transportation Review*, *184*(February), 103478. https://doi.org/10.1016/j.tre.2024.103478

Khalif, M. A., & Slim, C. (2024). The Role of Strategic Management in Achieving Organizational Resilience. *International Journal of Religion*, *5*(8), 658–672. https://doi.org/10.61707/xqqf3c92

Le, H. S., Do, T. V. H., Nguyen, M. H., Tran, H. A., Pham, T. T. T., Nguyen, N. T., & Nguyen, V. H. (2024). Predictive model for customer satisfaction analytics in E-commerce sector using machine learning and deep learning. *International Journal of Information Management Data Insights*, *4*(2), 100295. https://doi.org/10.1016/j.jjimei.2024.100295

Li, D., Gu, F., Li, X., Du, R., Chen, D., & Madden, A. (2023). Dynamic sales prediction with auto-learning and elastic-adjustment mechanism for inventory optimization. *Information Systems*, *119*, 102259. https://doi.org/10.1016/j.is.2023.102259

Li, J., Fan, L., Wang, X., Sun, T., & Zhou, M. (2024). Product Demand Prediction with Spatial Graph Neural Networks. *Applied Sciences (Switzerland)*, *14*(16). https://doi.org/10.3390/app14166989

Mahin, R. M. P., Shahriar, M., Das, R. R., Roy, A., & Reza, A. W. (2025). Enhancing Sustainable Supply Chain Forecasting Using Machine Learning for Sales Prediction. *Procedia Computer Science*, *252*, 470–479. https://doi.org/10.1016/j.procs.2025.01.006

Nguyen, D. N., Nguyen, V. H., Trinh, T., Ho, T., & Le, H. S. (2024). A personalized product recommendation model in e-commerce based on retrieval strategy. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(2), 100303. https://doi.org/10.1016/j.joitmc.2024.100303

Ni, G. (2023). Optimal decisions on price and inventory for a newsboy-type retailer with identifiable information and discount promotion. *PLoS ONE*, *18*(7 July), 1–20. https://doi.org/10.1371/journal.pone.0288874

Nie, T., Song, B., & Zhang, J. (2024). Sales pricing models based on returns: Bundling vs. add-on. *Omega (United Kingdom)*, *125*(January), 103038. https://doi.org/10.1016/j.omega.2024.103038

Nikose, A., Mungale, T., Shelke, M., Shelote, R., & Solanke, P. (2022). Best Selling Product and Category Prediction Using Sales Analysis. *International Journal of Advanced Research in Science, Communication and Technology*, *2*(2), 805–811. https://doi.org/10.48175/ijarsct-2970

Oancea, B. (2023). Automatic Product Classification Using Supervised Machine Learning Algorithms in Price Statistics. *Mathematics*, *11*(7). https://doi.org/10.3390/math11071588

Pocchiari, M., Proserpio, D., & Dover, Y. (2024). Online reviews: A literature review and roadmap for future research. *International Journal of Research in Marketing*, *42*(2), 275–297. https://doi.org/10.1016/j.ijresmar.2024.08.009

Poláček, L., Ulman, M., Cihelka, P., & Šilerová, E. (2024). Dynamic Pricing in E-commerce: Bibliometric Analysis. *Acta Informatica Pragensia*, *13*(1), 114–133. https://doi.org/10.18267/j.aip.227

Rios, J. H., & Vera, J. R. (2023). Dynamic pricing and inventory control for multiple products in a retail chain. *Computers and Industrial Engineering*, *177*(December 2022), 109065. https://doi.org/10.1016/j.cie.2023.109065

Schmitt, M. (2023). Automated machine learning: AI-driven decision making in business analytics. *Intelligent Systems with Applications*, *18*(June 2022), 200188. https://doi.org/10.1016/j.iswa.2023.200188

Sharma, D., Kumar, R., & Jain, A. (2022). Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning. *Measurement: Sensors*, *24*(November), 100560. https://doi.org/10.1016/j.measen.2022.100560

Statista. (2025a). *Beauty & Personal Care*. Https://Www-Statista-Com.Unpad.Idm.Oclc.Org/Outlook/Cmo/Beauty-Personal-Care/Worldwide.

Statista. (2025b). *eCommerce*. Https://Www-Statista-Com.Unpad.Idm.Oclc.Org/Outlook/Emo/Ecommerce/Indonesia.

Tang, Y. M., Chau, K. Y., Lau, Y. Y., & Zheng, Z. (2023). Data-Intensive Inventory Forecasting with Artificial Intelligence Models for Cross-Border E-Commerce Service Automation. *Applied Sciences (Switzerland)*, *13*(5). https://doi.org/10.3390/app13053051

Trapero, J. R., de Frutos, E. H., & Pedregal, D. J. (2024). Demand forecasting under lost sales stock policies. *International Journal of Forecasting*, *40*(3), 1055–1068. https://doi.org/10.1016/j.ijforecast.2023.09.004

Valencia-Arias, A., Uribe-Bedoya, H., González-Ruiz, J. D., Santos, G. S., & Ramírez, E. C. (2024). Artificial intelligence and recommender systems in e-commerce. Trends and research agenda. *Intelligent Systems with Applications*, *24*(July). https://doi.org/10.1016/j.iswa.2024.200435

Wang, T. C., Guo, R. S., & Chen, C. (2023). An Integrated Data-Driven Procedure for Product Specification Recommendation Optimization with LDA-LightGBM and QFD. *Sustainability (Switzerland)*, *15*(18). https://doi.org/10.3390/su151813642

Wang, X., Zhang, C., & Xu, Z. (2024). A product recommendation model based on online reviews: Improving PageRank algorithm considering attribute weights. *Journal of Retailing and Consumer Services*, *81*(June), 104052. https://doi.org/10.1016/j.jretconser.2024.104052

Wang, Y., & Zhang, Y. (2023). Multivariate SVR Demand Forecasting for Beauty Products Based on Online Reviews. *Mathematics*, *11*(21). https://doi.org/10.3390/math11214420

Wu, J., Liu, H., Yao, X., & Zhang, L. (2024). Unveiling consumer preferences: A two-stage deep learning approach to enhance accuracy in multi-channel retail sales forecasting. *Expert Systems with Applications*, *257*(July), 125066. https://doi.org/10.1016/j.eswa.2024.125066

Wu, X., Liao, H., Lev, B., & Ding, W. (2024). A Classification-Based Product Selection Method Based on Online Reviews on Multifaceted Attributes. *IEEE Transactions on Computational Social Systems*, *12*(1), 11–24. https://doi.org/10.1109/TCSS.2024.3485009

Wu, Y., Chen, L., Ngai, E. W. T., & Wu, P. (2024). Stimulating positive reviews by combining financial and compassionate incentives. *Internet Research*, *62001314*. https://doi.org/10.1108/INTR-01-2023-0062

Xu, S., Tang, H., & Huang, Y. (2023). Inventory competition and quality improvement decisions in dual-channel supply chains with data-driven marketing. *Computers and Industrial Engineering*, *183*(July), 109452. https://doi.org/10.1016/j.cie.2023.109452

Yang, B., Xu, X., Gong, Y., & Rekik, Y. (2024). Data-driven optimization models for inventory and financing decisions in online retailing platforms. *Annals of Operations Research*, *339*(1–2), 741–764. https://doi.org/10.1007/s10479-023-05234-4

Yuhsiang, L., & Lichung, J. (2024). The impact of consumer heterogeneity in the product life cycle on the diffusion patterns of user reviews and sales. *Journal of Retailing and Consumer Services*, *76*(September 2023), 103558. https://doi.org/10.1016/j.jretconser.2023.103558

Zaghloul, M., Barakat, S., & Rezk, A. (2024). Predicting E-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches. *Journal of Retailing and Consumer Services*, *79*(March), 103865. https://doi.org/10.1016/j.jretconser.2024.103865

Zhang, B., Zhang, Z., Lai, K. H., & Zhang, Z. (2024). Incentive hierarchies intensify competition for attention: A study of online reviews. *Decision Support Systems*, *185*(October 2023). https://doi.org/10.1016/j.dss.2024.114293

Zhang, C., Yang, F., & Zhang, X. (2023). A new inclusion measure-based clustering method and its application to product classification. *Information Sciences*, *626*, 474–493. https://doi.org/10.1016/j.ins.2023.01.061

Zhang, Q., & Xiao, T. (2024). Incentive strategies of an e-tailer considering online reviews: Rebates or services. *Electronic Commerce Research and Applications*, *68*(September), 101453. https://doi.org/10.1016/j.elerap.2024.101453

Zhuang, Y., & Xu, X. (2025). Flash sale or continuing Sale? Examining the timing flow of E-tailers' promotion effects. *Journal of Retailing and Consumer Services*, *86*(May), 104322. https://doi.org/10.1016/j.jretconser.2025.104322