A Sales Prediction Model for Live Commerce on Douyin (TikTok): Using Streaming Strategy Data and Viewer Comment Data

Yi Mei[1], Kohei Otake[2,*]

[1]Graduate School of Economics, Sophia University, Tokyo, Japan

[2]Faculty of Economics, Sophia University, Tokyo, Japan

Received 30 December 2025; received in revised form 12 April 2026; accepted 20 April 2026

DOI: https://doi.org/10.46604/aiti.2026.16033

Abstract

This study develops a sales prediction model for Douyin live commerce, focusing on five cosmetics influencers. Using sales revenue as the outcome, it examines whether quantitative operational variables and viewer comment features can explain sales performance. The dataset includes cumulative viewers, streaming duration, start time, preview video characteristics, and comment-derived features based on sentiment-related measures and keywords. Multiple regression and random forest regression are used to compare predictive performance, and the results are interpreted through the stimulus-organism-response (S-O-R) framework. The random forest model using only quantitative variables generally achieves the highest R^2, indicating that operational variables are the main drivers of sales prediction. The effects of individual variables differ across influencers, suggesting that effective sales strategies are influencer-specific. In contrast, linguistic features extracted from comments add limited predictive value, likely because the comments are brief, repetitive, and dominated by emojis or stamps.

Keywords: TikTok, live commerce, sales prediction, natural language processing, influencer marketing

1. Introduction

In recent years, the rapid global expansion of social media use has significantly transformed the structure of consumer information search and purchasing behavior. As of April 2025, the number of social media users worldwide has reached 5.31 billion, accounting for 64.7% of the global population [1]. Against this backdrop of social penetration, content formats centered on short videos and live streaming have emerged, with TikTok strengthening its global influence by reporting 1.59 billion advertising-reachable users as of January 2025. This environment can be characterized as a complex industrial information system where human agents and autonomous algorithms interact to drive commercial outcomes [2].

Meanwhile, live commerce is experiencing rapid market expansion internationally. Grand View Research [3] predicts that the global live commerce market reached $128.4 billion in 2024 and will maintain an annual average growth rate of 39.9% from 2025 to 2033. Furthermore, the Asia-Pacific region accounts for 66.0% of this market, indicating that the development of live commerce is centered in East Asia. In particular, annual statistics on China's live streaming and short-video industry report that approximately 29,000 multi-channel networks (MCNs) exist domestically from 2024 to 2025, with related market size reaching 63.6 billion yuan [4]. MCNs are management organizations that bundle multiple streamers and handle content production support, streaming operations, and advertising and sales intermediation, thereby serving as key supply-side entities in China's live commerce.

* Corresponding author. E-mail address: k-otake@sophia.ac.jp

Therefore, in China, not only individual streamers but also operational entities such as MCNs play a critical role in support market expansion. Within this complex industrial information system, MCNs must navigate trade-offs in operational transparency, balanceing algorithmic traffic data with human-centered content strategies. Wang and Zhang [2] emphasized that ensuring stakeholders perceive the "truth" of product claims, rather than just the "hype" of the stream is crucial for maintaining responsible retail operations. This perspective is relevant to live commerce, where viewers’ trust in product claims affect platform effectiveness. Furthermore, the E-Commerce Research Center of China [5] identifies live commerce as a major growth area for China's overall e-commerce. Additionally, P&S Intelligence [6] estimates that China's e-commerce market has reached approximately $885.5 billion in 2024, indicating the continued expansion of China's overall digital consumer market.

Under these circumstances, research on live commerce has grown rapidly. Existing research can be broadly classified into the following three perspectives. The first perspective focuses on the streaming environment of live commerce and the uncertainty of sales volume. Wang and Zhang [7] investigated the fundamental characteristics of live streaming. They showed that real-time nature, visual appeal, and immediate responsiveness improve purchasing efficiency. However, their study also demonstrated that live sales volume is difficult to predict due to significant fluctuations caused by numerous external factors. Additionally, Jiang [8] focused on the interactivity of live streaming, revealing that interaction between viewers and streamers strengthens customer engagement and influences the formation of purchase intent. From another perspective, Kanatani and Wang [9] viewed live commerce as decision-making under time constraints, demonstrated that consumers tend to make rapid judgments using word-of-mouth and existing knowledge as cues within limited time.

The second perspective focuses on the attributes of social media influencers and streamers. Prior studies showed that attributes such as the credibility and attractiveness of influencers affect purchase intentions through viewers' cognition and emotions [10-12]. This indicates that the characteristics of individual streamers play an important role in purchasing behavior and sales formation. Furthermore, the "human" element of the streamer—specifically their creativity in presenting products and their technological competence in leveraging platform features—represents a distinct variable that can significantly impact marketing performance beyond simple operational metrics [13].

The third perspective concerns research on consumer behavior analysis and sales prediction using textual information such as reviews and comments. Xu et al. [14] proposed a method for sentiment classification in review texts, demonstrating that textual information is useful for analyzing consumer behavior. Additionally, Chen et al. [15] clarified that incorporating features from product descriptions and review texts improves the accuracy of sales prediction models. However, Wang and Sun [16] proposed that in AI-mediated environments, high volumes of information may lead to "information redundancy," which can negatively affect user acceptance and behavior. In high-traffic live streams, comments may become noise rather than meaningful signals. This leads to an information overload that does not correlate with purchase intent.

Prior research demonstrated the uncertainty of the streaming environment in live commerce, the importance of influencers, and the usefulness of textual information. Nevertheless, research that integrates these factors under a unified framework and empirically examines their impact on sales remains insufficient.

Based on the above background and issues in existing research, this study aims to clarify how quantitative operations data influence live commerce sales on Douyin (the Chinese version of TikTok). These The research also explores text data contained in viewer comments (particularly emotional expressions). Drawing on the logic of information redundancy [16], the analysis treats this text data strictly as auxiliary variable. Specifically, it investigates why massive comment volumes may fail to enhance predictive accuracy compared to cleaner operational signals.

Douyin is selected as the research platform because it integrates short videos, live streaming, e-commerce functions, and payment functions. This integrated structure allows users to complete the entire consumption cycle—from information acquisition to payment—within the app. Consequently, the platform contains richer data for understanding consumer behavior compared to other social media. In this study, sales prediction models are constructed using quantitative data to examine predictive accuracy and key influencing factors. Textual data are then added to test for any accuracy improvement.

This study seeks to complement conventional single-perspective research by capturing more comprehensive determinants of sales performance. The analysis relies primarily on quantitative metrics and the text data provides a secondary exploratory test. The findings provide practical implications for live streamers, MCNs, and companies in managing the integration of traffic data and content strategy [2], while a theoretical framework is constructed that contributes to the development of live commerce research.

2. Literature Review

This section reviews prior research related to this study from three perspectives: (i) the streaming environment unique to live commerce and the uncertainty of purchasing decisions, (ii) the influence of influencer and streamer attributes on purchasing behavior, and (iii) the potential for consumer behavior analysis and sales prediction using textual information.

2.1. Streaming environment and characteristics of purchasing decisions in live commerce

Live commerce is a transaction format that enables product demonstrations and immediate interaction through real-time streaming; this forms a purchasing environment different from traditional e-commerce.

Wang and Zhang [7] pointed out that while characteristics such as real-time nature, visual appeal, and immediate responses are closely related to purchasing behavior, sales are strongly influenced by streaming conditions and external factors, making prediction difficult. Additionally, Jiang [8] focused on the interactivity, demonstrating that interaction between streamers and viewers enhances engagement and influences the formation of purchase intent. Furthermore, Kanatani and Wang [9] viewed live commerce as a decision-making environment with time constraints. They revealed that viewers tend to make rapid judgments using word-of-mouth and existing knowledge as cues within limited streaming time.

Beyond these operational characteristics, recent research suggests that such ecosystems effectively function as complex industrial information systems. In these systems, autonomous AI agents synthesize operational data to facilitate real-time decisions in retail environments. This perspective emphasizes that operational transparency—defined as how viewers perceive the "truth" of product claims—is a critical factor in responsible and effective commercial operations [2]. These studies suggest that live commerce is a purchasing environment with both high uncertainty and immediacy, where the process of purchasing decisions and sales formation is complex.

2.2. Influencer and streamer attributes and purchasing behavior

In social media environments, influencers and streamers serve as key information sources and are widely studied in relation to viewers' attitudes and purchasing behavior. For example, Lou and Yuan [10] demonstrated that in branded content, message value and source credibility are involved in consumer trust formation, which subsequently leads to purchase-related responses. Lin and Minamikawa [11] focused on a framework showing that influencer such as trustworthiness and attractiveness, affect purchase intentions through cognitive and emotional responses. Furthermore, Ilieva et al. [12] empirically examined the impact of social media influencers on consumer attitudes and purchasing behavior, revealing the importance of influencers' role as information sources.

In addition to perceived attributes, Wang and Zhang [13] highlighted that human agents' creativity and technological competence can function as distinct drivers of marketing performance beyond operational metrics. In live commerce, this suggests that the streamer's interactive quality may represent an important factor not fully captured by the current model. These studies suggested the possibility that streamer and influencer attributes influence purchasing behavior through viewers' psychological responses.

2.3. Textual information and sales prediction research

In recent years, research has been actively conducted to analyze textual information, particularly represented by reviews and word-of-mouth among consumer-generated content, and to understand and predict purchasing behavior and sales. For example, Xu et al. [14] proposed a method for sentiment classification targeting reviews on e-commerce platforms, demonstrating that textual information is useful for capturing consumer evaluations and responses. Chen et al. [15] empirically demonstrated that using features of expressions contained in product descriptions and review texts improves the performance of sales prediction models. Wang and Zhang [7] showed the necessity of handling diverse information in an integrated manner in sales prediction using live streaming data, indicating that prediction based solely on single numerical indicators is insufficient. These studies suggest that textual information is a vital source for understanding consumer behavior and can enhance predictive analysis.

Wang and Sun [16] argued that excessive information volume may lead to redundancy, causing cognitive overload and reducing effective information processing. This theoretical logic implies that in high-traffic live streams, massive comment volumes may function as noise rather than signals. Such overload does not necessarily correlate with purchase intent.

Prior research clarifies three key areas: (i) the unique streaming environment and purchasing decisions, (ii) the influence of streamer attributes on purchasing behavior, and (iii) the potential for consumer behavior analysis and sales prediction using textual information, respectively. However, research combining these factors remains limited. To address this gap, this study builds a framework based on quantitative operational data. The framework adopts the stimulus-organism-response (S-O-R) framework to interpret the integration of these factors, using textual information as an auxiliary test.

3. Analytical Procedures and Dataset

This section describes the analytical procedures, data acquisition, and dataset construction. This study consists of three steps: (i) data acquisition, (ii) extraction of sentiment words from viewer comments and feature engineering, and (iii) dataset construction and model building, as shown in Fig. 1.

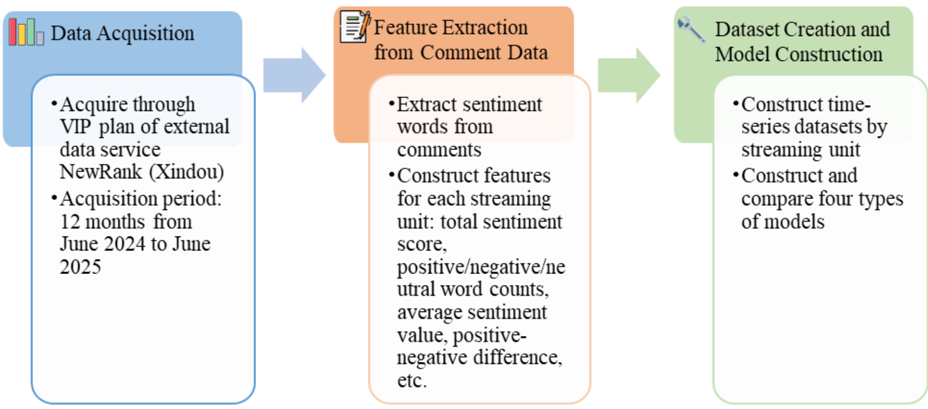

Fig. 1 Analytical procedures of this study

The data used in this study are obtained through the VIP plan of NewRank (Xindou), an external data service that collects live streaming data from Douyin. The acquisition period spans 12 months from June 2024 to June 2025, during which operational data and user comment are collected.

The analysis targets of this study were five influencers belonging to the cosmetics category on Douyin. These influencers all belong to different MCNs and are characterized by being in competitive relationships with each other. The reason for selecting the cosmetics category is that the beauty and skincare sector shows particularly high gross merchandise value (GMV). Furthermore, viewer participation rates in the live commerce market, and because viewer comments frequently generate opinions about product features, usability, and prices, providing abundant data suitable for text analysis.

For each influencer, the number of streaming days and comments available for acquisition differed, resulting in variations in the amount of data (number of records) used for analysis. Specifically, the analysis targeted 58 streaming unit data for Influencer A, 38 for B, 77 for C, 44 for D, and 56 for E.

For data acquisition, operational data per stream, related preview video data, and the occurrence counts of the top 20 keywords from live comments are gathered. All data are organized by streaming unit to construct time-series datasets for each streamer. When multiple streams are conducted on a single day, the values for that day are aggregated to ensure analytical consistency. In this study, the total sales revenue generated through live streaming on that day is used as the target variable. The explanatory variables are classified into two categories: (i) quantitative data related to streaming strategies (Table 1), and (ii) sentiment data features, including eight sentiment variables and occurrence counts of each sentiment keyword created from comment data (Table 2).

Table 1 Quantitative data related to streaming strategies used for model construction and their overview

Variable (Daily basis)		Variable name	Data type	Description
Quantitative data	Common data	Holiday/Weekend/Weekday	Binary	Dummy variable indicating whether the streaming day is a holiday, weekend, or weekday
		Event period	Binary	Dummy variable indicating promotional periods or major sale days
		Season (Spring/Summer/Fall/Winter)	Binary	Variable reclassifying months into quarters (March-May, June-August, September-November, December-February)
		Number of followers	Numerical	Total number of followers of the streamer's account on that day
		Follower increase	Numerical	Difference in followers compared to the previous day
	Preview video data (Previous Day + Current Day)	Number of preview video posts	Numerical	Number of preview videos posted from the day before to the current day of streaming
		Preview video interactions (saves, shares, comments, likes)	Numerical	Total viewer engagement for each preview video (calculated separately for each metric)
		Preview video duration (seconds)	Numerical	Duration of preview videos in seconds
		Preview video post promotion	Binary	1 if the preview video contains promotional elements (advertising copy, sales navigation)
		Preview video post time (Morning/Noon/Evening/Night)	Numerical	Variable indicating preview video posting time slot (before 11:00, 11:00-15:00, 15:00-19:00, after 19:00)
	Live stream data	Number of live streams	Numerical	Number of live streams conducted on that day
		Number of live products	Numerical	Number of different products handled in that day's live stream
		Live promotion	Binary	1 if promotional activities are implemented during the stream
		Live streaming duration (minutes)	Numerical	Duration of each live stream in minutes
		Live cumulative viewers	Numerical	Total number of viewers who participated in the stream
		Follower increase from live	Numerical	Difference in followers before and after streaming
		Live start time (Morning/Noon/Evening/Night)	Numerical	Variable indicating streaming start time slot (before 11:00, 11:00-15:00, 15:00-19:00, after 19:00)

Table 2 Sentiment data related to comment data used for model construction and their overview

Variable	Variable name	Data type	Description
Comment data	Occurrence counts of keywords	Numerical	Occurrence counts of keywords (Top 20) appearing in real-time comments of each live stream on that day
	Total sentiment score	Numerical	An indicator comprehensively shows the emotional direction and intensity of the entire stream. For calculation, word scores are first determined based on the polarity (positive/negative) and intensity of each sentiment word. This value is obtained by multiplying the word-level score by the occurrence count within the stream and summing them, where a larger positive value indicates dominant positive or negative emotions.
	Positive word count	Numerical	The total number of occurrences of words with positive polarity in the sentiment dictionary within the stream.
	Negative word count	Numerical	The total number of occurrences of words with negative polarity in the sentiment dictionary within the stream.
	Neutral word count	Numerical	The total number of occurrences of words judged as neutral (word score of 0) in the sentiment dictionary within the stream.
	Number of sentiment word types	Numerical	The number of different sentiment words (types) used within the stream. Used as an indicator of the diversity of emotional expression vocabulary.
	Average word score	Numerical	The value obtained by dividing the total sentiment score by the total number of sentiment words detected in that stream. An indicator showing the average emotional intensity per word, excluding the influence of word count.
	Net sentiment score	Numerical	The value obtained by subtracting the total negative word count from the total positive word count. Defined as an indicator showing the quantitative balance between positive and negative expressions within the stream.
	Positive sentiment ratio	Numerical	The proportion of positive words among polarized sentiment words (positive and negative words) excluding neutral words. An indicator of emotional bias normalized to a range from 0 to 1.

4. Construction and Evaluation of Sales Prediction Models for Live Commerce

This section first describes the datasets constructed from the data obtained in Section 3, including data preprocessing and feature design. It then explains the development of sales prediction models based on these datasets, and finally presents the evaluation metrics used to assess model performance and predictive accuracy in a systematic manner.

4.1. Dataset creation

In this study, four types of datasets are prepared for each of the five influencers: (i) a dataset containing only quantitative data related to streaming strategies; (ii) a dataset integrating quantitative data with the eight variables derived from comment data and occurrence counts of each sentiment keyword; (iii) a dataset consisting only of the four comment-derived variables: "total sentiment score," "positive word count," "negative word count," and "average word score,"; and (iv) a dataset adding only the above four variables to the quantitative data. The quantitative dataset serves as the main model, while the datasets incorporating sentiment variables are used as exploratory extensions.

4.2. Sales prediction model construction

Considering the nature of live streaming data, this study constructs prediction models using both linear models (multiple regression) and nonlinear models (random forest regression). Multiple regression is a traditional statistical method premised on linear relationships between explanatory variables and dependent variables; it is characterized by its ability to explicitly capture direct linear effects. In contrast, Random forest is a nonlinear learning method that constructs multiple decision trees as an ensemble, excelling at capturing complex interactions and nonlinear structures among data. Additionally, it maintains high stability even with smaller datasets, making it highly suitable for data with high volatility and nonlinear feature effects, such as live streaming.

4.3. Model evaluation methods

For the constructed models, prediction performance is evaluated based on the following criteria. First, all collected data were split into training and test sets at an 8:2 ratios, using 80% for training and 20% for testing to evaluate the model's generalization performance. Mean absolute error (MAE) and R^2 are used as evaluation metrics for the models.

MAE is an indicator showing the average of absolute errors between predicted and actual values, allowing for an intuitive interpretation of error magnitude:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{1}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of samples.

R^2 is an indicator measuring the extent to which independent variables explain the variance of the dependent variable, showing the overall explanatory power of the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2} \tag{2}$$

where $\bar{y}$ is the mean of the actual values.

Additionally, SHapley additive explanations (SHAP) values are used to quantify the contribution of each feature to the model’s predictions [17]. SHAP enables visualize the direction and strength of feature influence, providing explanatory power beyond traditional model coefficients. By using these evaluation metrics in combination, model performance is comprehensively judged from three perspectives: prediction accuracy, explanatory power, and interpretability.

5. Analysis Results and Discussion

This section first compares and verifies the accuracy of multiple constructed prediction models and selects the analytical model for discussion. Subsequently, detailed discussion of the sales determinants for each influencer is conducted using the selected model. To explain these results, the study adopts the S-O-R framework. In this framework, live streaming strategies and operational data act as the stimulus (S). These stimuli affect the internal cognitive and emotional states of viewers, which represent the organism (O). These internal states finally drive the purchasing response (R), measured as sales revenue.

5.1. Comparison of model performance

As described in Section 4.1, models are created using multiple regression and random forest regression for these four types of datasets: (i) quantitative data related to streaming strategies, (ii) quantitative data integrated with the eight variables derived from comment data and sentiment keyword counts, (iii) four comment-derived variables only: "total sentiment score," "positive word count," "negative word count," and "average word score," and (iv) quantitative data supplemented with these four variables. The prediction performance on test data was compared in Table 3. Note that the values in the table show R^2, and the values in parentheses show MAE.

Table 3 Performance comparison of eight models for five influencers

	Model \ Dataset	Quantitative data only	Quantitative + Sentiment data	4 Sentiment indicators only	Quantitative + 4 Sentiment indicators
A	Multiple Regression	0.47 (2793434)	0.35 (3121300)	-1.26 (3018682)	0.46 (2751760)
	Random Forest	0.900 (991843)	0.890 (1160215)	-1.49 (3139294)	0.895 (1102127)
B	Multiple Regression	0.54 (672107)	-3738.00 (27442481)	-2.31 (1519125)	-5.56 (1882705)
	Random Forest	0.59 (664121)	0.57 (655748)	-0.16 (989896)	0.58 (656780)
C	Multiple Regression	-0.27 (20110636)	-0.56 (25468556)	-0.24 (22042102)	-0.15 (18503534)
	Random Forest	0.29 (14021375)	0.25 (13940081)	-0.14 (20984823)	0.20 (16735883)
D	Multiple Regression	0.67 (4326737)	0.46 (5404667)	0.50 (4861537)	0.05 (7212132)
	Random Forest	0.48 (4226250)	0.35 (5172624)	0.17 (5280398)	0.49 (4232181)
E	Multiple Regression	-1.15 (894928)	-8.84 (1439296)	-0.61 (874723)	-3.44 (1062055)
	Random Forest	0.55 (439439)	0.48 (455081)	-0.61 (870684)	0.52 (444320)

From Table 3, as an overall trend, the random forest model using only quantitative data shows the highest prediction accuracy. On the other hand, models that added sentiment features derived from comments failed to outperform models using only quantitative data in any method. The contribution of comment data is limited. This limitation arises because the extracted features were simple, consisting mainly of emojis and stamps. Such features struggle to capture complex purchase intentions. Additionally, massive comment volumes in high-traffic streams create information redundancy, which acts as noise rather than useful signals. Therefore, this study uses random forest results with quantitative data to discuss feature importance, except for influencer D, where multiple regression model is employed.

5.2. Discussion of sales determinants for each influencer

(1) Discussion of model for influencer A

For influencer A, the relationship between predicted and actual values for the random forest model using only quantitative data, which shows the highest prediction accuracy, is presented in Fig. 2. From this figure, a clear positive correspondence is confirmed between actual and predicted values, suggesting that the model possesses consistent prediction performance across the entire sales level.

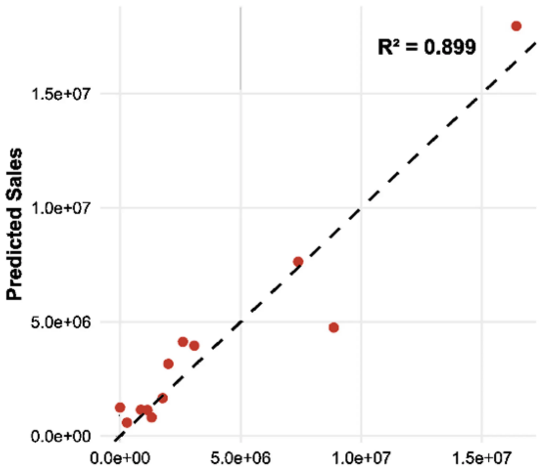

Fig. 2 Relationship between predicted and actual sales for influencer A

The results of visualizing the feature importance of influencer A using SHAP are shown in Fig. 3. The colors of the bars in Fig. 3 signify the direction of each variable's contribution to sales prediction. Green bars mean that higher variable values contribute to increasing sales (positive direction), while red bars represent that higher variable values contribute to decreasing sales (negative direction). The numbers on the left side of each row (green and red numbers) indicate the direction of the average SHAP value, and the black values on the right side are the average absolute SHAP values, indicating the magnitude of variable importance. A larger average absolute SHAP value implies a greater contribution to the model's prediction.

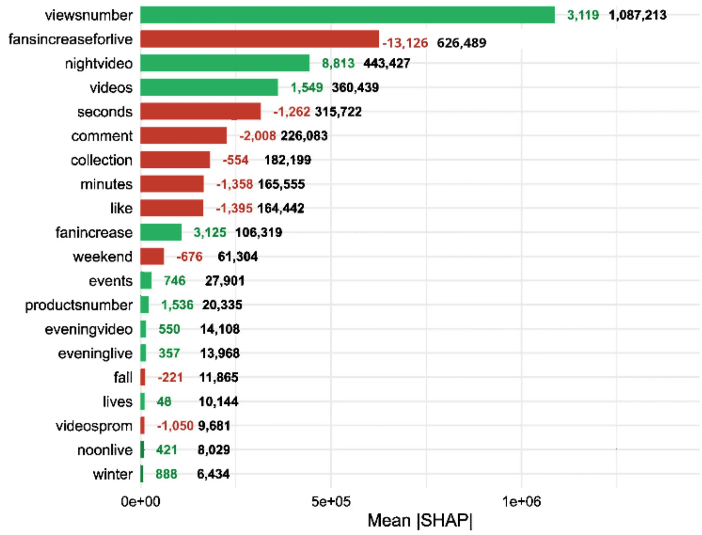

Fig. 3 Feature importance by SHAP in the random forest regression model using only quantitative data for influencer A (Top 20 features)

Using Fig. 3 to discuss the feature importance for Influencer A, the study applies the S-O-R framework to interpret the results. First, focusing on cumulative viewers, this feature shows the maximum positive contribution to sales prediction. As a significant stimulus (S), the expansion of viewer scale is a primary factor in enhancing viewer interest (O) and inducing purchases, reconfirming that traffic in live commerce acts as a driver for the purchasing response (R). Additionally, the number of preview videos posted before live streaming show a positive contribution to sales. Furthermore, the results reveal that preview videos posted after 19:00 have a stronger effect. These timely stimuli (S) effectively capture viewer attention (O) during peak platform activity, leading to a stronger sale response (R) by enhancing the viewer-guiding effect.

Conversely, follower increase shows a negative contribution, confirming a tendency for predicted sales to decrease as followers increase. This phenomenon is considered to originate from incentivized following campaigns implemented by this influencer. For example, through measures such as "receive prizes by following" or "only users who follow can obtain benefits". These measures act as a stimulus (S) that attracts viewers with only temporary interest. Such viewers may experience a state of low purchase intent (O), which does not lead to a purchasing response (R). Consequently, a reversal phenomenon occurred where follower increase does not necessarily lead to immediate sales increase.

Additionally, in contrast to the "number of posts" of preview videos, the "length" of videos showed a negative contribution. Excessive duration (S) may cause viewer fatigue or a decline in freshness (O). This negative internal state hinders the transition to a purchasing response (R), potentially damaging expectations for the live stream. Therefore, while the number of preview videos is important, it is suggested that excessive length can be counterproductive.

(2) Discussion of model for influencer B

For influencer B, Fig. 4 shows the relationship between predicted and actual values for the random forest model using only quantitative data, which demonstrates the highest prediction accuracy. From Fig. 4, a positive correlation is confirmed between actual and predicted values, suggesting that the model possesses consistent prediction performance across the entire sales level.

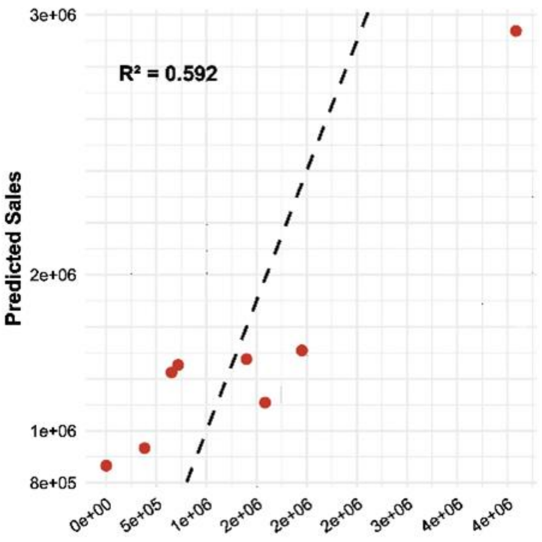

Fig. 4 Predicted and actual sales for influencer B

Using Fig. 5 to discuss the feature importance for influencer B, the study again utilizes the S-O-R framework. First, cumulative viewers show the maximum negative contribution to sales prediction. This indicates that a massive influx of viewers as a stimulus (S) may only attract users with casual interest (O). In this case, the internal state of the viewers (O) does not translate into a purchasing response (R). That is, on days when many users whose purpose is only viewing are present, the conversion rate to sales tends to be low.

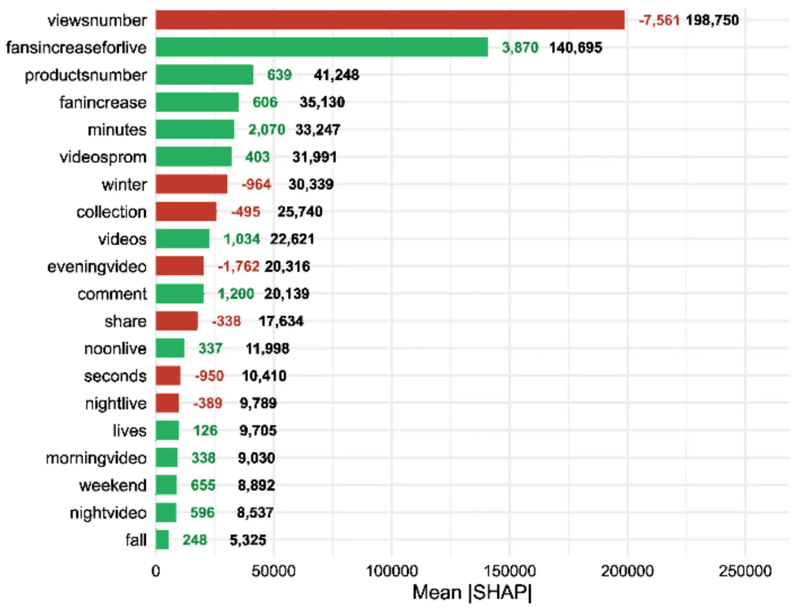

Fig. 5 Feature importance by SHAP in the random forest regression model using only quantitative data for influencer B (Top 20 features)

In contrast, follower increase during live streaming shows a significant positive contribution. This result is opposite to the case of influencer A, where follower increase showed a negative contribution. For influencer B, the acquisition of new followers during the stream acts as a stimulus (S) that reflects high-intent engagement (O). This positive internal state directly leads to a purchasing response (R). This indicates that for this influencer, follower growth serves as a strong signal of high-intent engagement from users who are likely to purchase.

While cumulative viewers and follower increase are the main factors, other live operational elements have relatively small contributions. This reveals that in this influencer's sales model, a specific small number of stimuli (S) greatly influence the purchasing response (R).

(3) Discussion of model for influencer C

For influencer C, Fig. 6 shows the relationship between predicted and actual values for the random forest model using only quantitative data, which demonstrates the highest prediction accuracy. From Fig. 6, a certain positive correlation is confirmed between actual and predicted values. This result indicates that while the model explains part of the sales variation, the prediction accuracy remains relatively low compared to influencers A and B. This suggests that influencer C's sales structure is relatively complex.

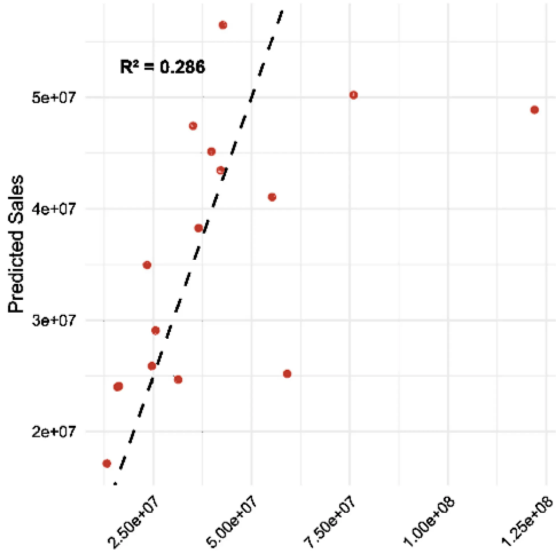

Fig. 6 Predicted and actual sales for influencer C

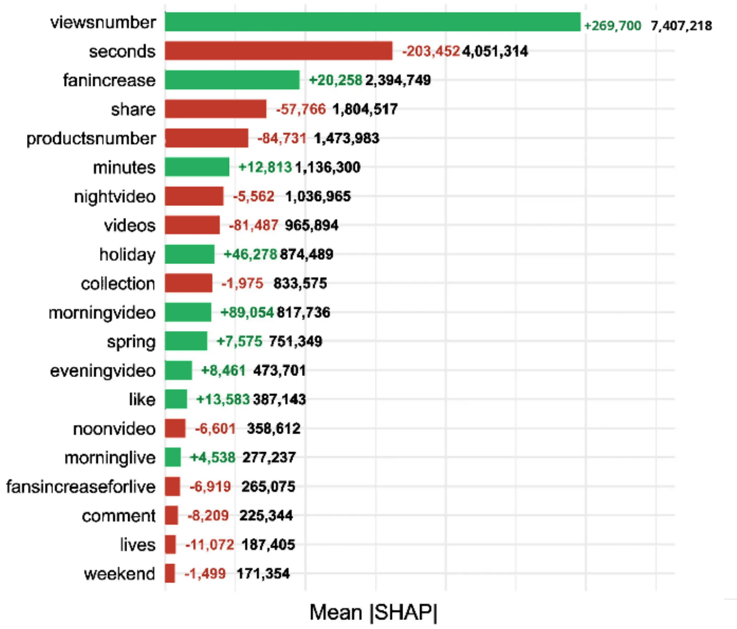

Fig. 7 Feature importance by SHAP in the random forest regression model
using only quantitative data for influencer C (Top 20 features)

Fig. 7 presents the results of visualizing the feature importance for Influencer C using SHAP. The results indicate that for influencer C, a larger audience scale as a stimulus (S) effectively enhances the customer attraction level (O). This positive internal state leads to a higher purchasing response (R), revealing a structure where high customer attraction translates into positive sales outcomes.

In contrast, the length of preview videos (seconds) exhibits a negative contribution. Longer preview content (S) may fail to maintain high-intent interest (O) in the products. Shorter, more concise preview videos may be more effective Stimuli (S) in driving the purchasing response (R).

Daily follower increase also shows a significant positive contribution to sales prediction. This steady growth (S) serves as a stable indicator of the account's popularity (O). This general reach supports a consistent purchasing response (R) and reflects the sales potential for this influencer.

Furthermore, the number of preview post shares also demonstrates a negative contribution. While wide sharing of preview posts itself (S) is considered to lead to recognition expansion, it may not directly influence the cognitive state required for immediate purchasing (O). That is, a certain gap exists between sharing behavior and purchasing behavior (R).

In addition, the number of products sold indicates a negative contribution. This suggests that as the number of products increases (S), it may disperse viewer choices or create a high cognitive load (O). This internal state reduces purchasing efficiency, leading to a lower sale response (R). It is conceivable that multi-product strategy disperses viewer choices, resulting in decreased purchasing efficiency for individual products.

Overall, the results suggest that influencer C's sales structure depends on viewer scale and follower growth (S) to drive the response. However, internal streaming elements such as preview length and product act as stimuli (S) that can limit potential sales. In this respect, influencer C exhibits a hybrid sales structure in which both traffic-driven and content-related factors jointly shape performance.

(4) Discussion of model for influencer D

For influencer D, unlike other research subjects, the highest coefficient of determination (R^2) is obtained in multiple regression analysis using only quantitative data. This is attributed to the relatively small amount of target data; while nonlinear models like random forest fail to demonstrate sufficient generalization performance, multiple regression analysis accurately captures the linear relationships of the main quantitative factors determining sales.

Regarding the random forest model using only quantitative data, Fig. 8 shows the relationship between predicted and actual values. From Fig. 8, a certain positive correspondence is confirmed between actual and predicted values, with a coefficient of determination of R^2=0.481. This result indicates that for Influencer D, quantitative streaming indicators relatively well explain sales variation.

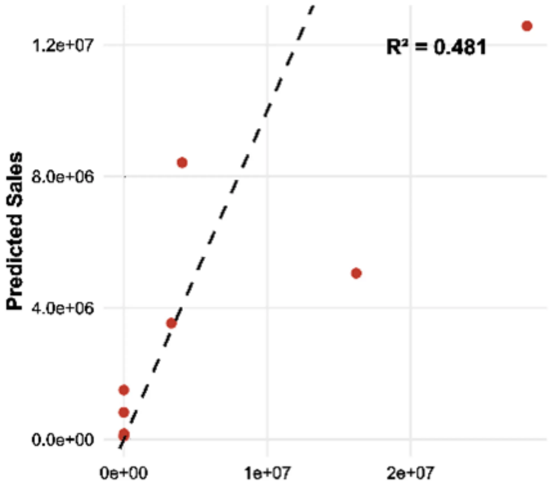

Fig. 8 Predicted and actual sales for influencer D

Fig. 9 illustrates the results of visualizing feature importance using SHAP. Based on Fig. 9, daily follower increase shows the highest importance, confirming that it has a strong negative impact on sales prediction. This indicates that general follower growth as a stimulus (S) does not directly lead to immediate purchasing behavior (R). Similarly, follower increase from live also exhibits a negative contribution. This suggests that follow promotion measures (S) may disperse viewer attention from purchasing to participation (O), potentially suppressing short-term sales.

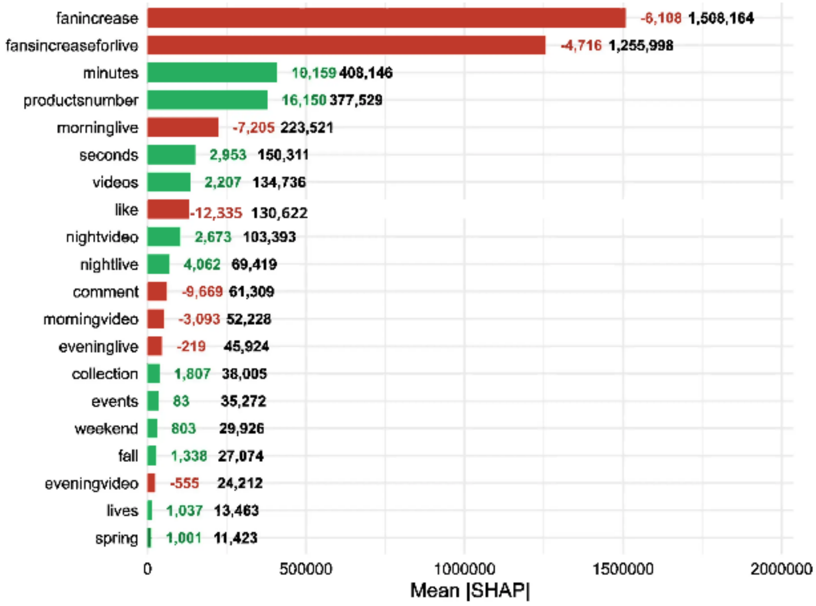

Fig. 9 Feature importance by SHAP in the random forest regression model using only quantitative data for influencer D (Top 20 features)

Conversely, streaming minutes demonstrates a positive contribution. Longer streaming sessions (S) provide more opportunities for viewer engagement and interest (O), which leads to higher sales responses (R). Additionally, the number of products handled in the stream also shows a positive contribution. A wider variety of product offerings (S) encourages a positive purchasing intent (O) and promotes higher sales for Influencer D.

While daily follower increase, follower increase from live, streaming minutes, and the number of products are the main factors in the random forest model, the multiple regression model is also examines for comparison. The results of the multiple regression model using only quantitative data are presented in Table 4. Consistent with the random forest model, both daily follower increase and the number of products showed significant impacts on sales, with negative and positive coefficients, respectively.

Table 4 Partial regression coefficients and p-values of the multiple regression model using only quantitative data for influencer D

	Partial regression coefficient	P-value
Intercept	-3.076e+06	0.015 *
Follower increase	-2.023e+02	0.000 ***
Number of preview video posts	2.287e+05	0.175
Preview video interaction (Saves)	1.089e+01	0.146
Preview video post promotion	-1.090e+06	0.179
Preview video duration (seconds)	3.658e+03	0.142
Preview video post time (Morning)	1.361e+06	0.027 *
Preview video post time (Evening)	-9.176e+05	0.038 *
Number of live streams	2.571e+06	0.003 **
Number of live products	3.119e+04	0.000 ***
Live start time (Morning)	-4.801e+06	0.000 ***
Live start time (Noon)	-2.954e+06	0.011 *
Live start time (Night)	-2.177e+06	0.001 **
Holiday	1.619e+07	0.000 ***

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$

Additionally, regarding streaming start time slots, morning (starting before 11:00) showed a significant negative coefficient. This indicates that the morning time slot (S) may not align with the high-intent shopping state of viewers (O), resulting in a lower sale response (R). Furthermore, holiday streaming also showed a significant positive impact. Holidays act as a stimulus (S) where viewers have higher shopping readiness and available time (O), confirming a tendency for live streams conducted on holidays to have higher sales compared to regular days.

From the above results, it is revealed that daily follower increase and the number of products are the most important stimuli (S) driving the purchasing response (R) for Influencer D's sales, as their significance was consistently confirmed in both models. Other operational elements, such as streaming timing (morning time slot and holidays), are considered secondary factors influencing sales fluctuations.

(5) Discussion of model for influencer E

For influencer E, Fig. 10 presents the relationship between predicted and actual values in the random forest model using only quantitative data. From Fig. 10, a clear positive correspondence is confirmed between actual and predicted values, with a coefficient of determination of R^2=0.546.

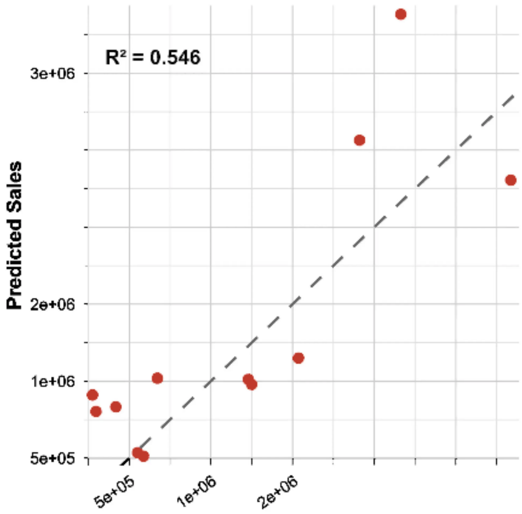

Fig. 10 Predicted and actual sales for influencer E

Fig. 11. Illustrates the results of visualizing the feature importance for influencer E using SHAP. Based on Fig. 11, cumulative viewers show the highest importance with a negative contribution. This suggests that for influencer E, a large audience scale as a stimulus (S) may attract many casual viewers with low purchase intent (O). As a result, this stimulus (S) does not lead to a strong purchasing response (R). This result confirms a tendency where the expansion of viewing scale does not necessarily lead to a direct increase in sales for this influencer. On the other hand, the number of products sold showed a positive contribution. A wider variety of items (S) provides a stable Stimulus (S) that effectively drives the purchasing response (R) in sales prediction.

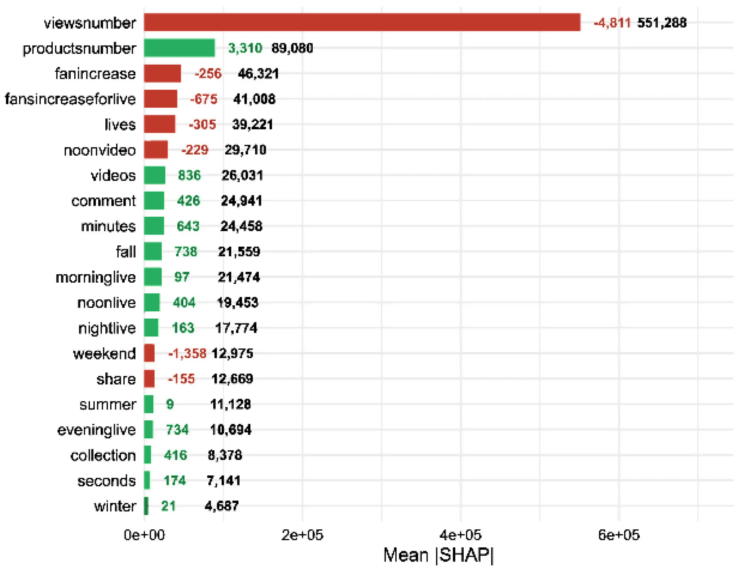

Fig. 11 Feature importance by SHAP in the random forest regression model using only quantitative data for influencer E (Top 20 features)

(6) Overall Summary

The discussion results of models for the five influencers are summarized in Table 5. From this table, it is evident that the factors affecting sales differ depending on the subject. First, focusing on influencers D and E, the "number of live products" acts as a positive stimulus (S) for sales. This indicates that for these influencers, a wider variety of product offerings supports a positive internal state in viewers (O) and revenue growth. Conversely, for influencer C, the number of live products shows a negative impact. As a stimulus (S), when too many products are handled, viewer choices may become dispersed, increasing cognitive load (O) and potentially decreasing sales efficiency (R).

Table 5 Comparison of sales influencing factors for the five influencers

Influencer	Positive influencing factors	Negative influencing factors
A	・Live cumulative viewers ・Preview video post time (night) ・Number of preview video posts	・Follower increase from live ・Preview video duration (seconds)
B	・Follower increase from live	・Live cumulative viewers
C	・Live cumulative viewers ・Follower increase	・Preview video duration (seconds) ・Preview video interactions (shares) ・Number of live products
D	・Live streaming duration (minutes) ・Number of live products	・Follower increase ・Follower increase from live
E	・Number of live products	・Live cumulative viewers

Second, the impact of "live cumulative viewers" as a stimulus (S) is not uniform. For influencers A and C, sales increased as viewer scale grew. However, for influencers B and E, the opposite result is obtained. For them, an increase in viewers has a negative impact on sales. This suggests that simply attracting a large number of viewers (S) does not necessarily lead to a high-intent state in the Organism (O), as many viewers may lack purchase intent.

Third, "follower increase from live" also shows different effects as a Stimulus (S). It is a positive influencing factor for Influencer B. In contrast, it has a negative impact for Influencers A and D. This implies that for some influencers, new followers acquired during a stream may not connect to the immediate internal intent required for a purchasing response (R).

Summarizing these findings, sales structures are broadly divided into different types. Some influencers rely on viewer scale and specific timing (S), while others depend more on product variety and engagement signals (S). Simply gathering people (S) does not guarantee sales (R). It is necessary to consider specific sales strategies and internal relationships with fans (O) according to each influencer's characteristics.

6. Conclusion and Future Research

This study empirically analyzed the impact of streaming strategies and viewer comments on revenue outcomes in live commerce on Douyin (TikTok). To identify variables associated with sales, prediction models were constructed using information related to streaming strategies and viewer comments, and evaluated their predictive accuracy. The analysis clarified key factors influencing sales performance in live commerce. The specific conclusions drawn from this study are summarized as follows:

(1) The comparison of model accuracy showed that the random forest model using only quantitative data generally achieved the highest R^2. This study further interpreted the results using the S-O-R framework, organized the characteristics and common tendencies of each influencer, and discussed sales factors in live commerce.

(2) Across most subjects, quantitative data related to streaming strategies acting as stimuli (S), such as cumulative viewers, streaming duration, and start time, were the main sales-determining factors. On the other hand, follower increase during live streaming did not consistently lead to a positive purchasing response (R), and in some cases it was negatively associated with sales. While measures aimed at increasing follower count attracted short-term interest, they did not necessarily increase viewers with high purchase intent (O) and could even have negative associations with sales in some cases.

(3) Clear differences appeared depending on the influencer. While the impact of specific variables varied across individuals, the analysis suggested that strategic factors such as product composition and internal fan relationships (O) may often play a more important role than audience size. This suggests that effective sales strategies are highly dependent on influencer-specific characteristics.

(4) In the models constructed in this study, the predictive contribution of linguistic expressions remained limited. This is because the comment data mainly consisted of simple features such as emojis and stamps, which were insufficient to capture complex purchase intentions. This result suggests that the effectiveness of sentiment features depends on data quality and vocabulary suitability.

Future research should further refine the analysis of viewer comments. In particular, future studies should develop sentiment vocabularies that better reflect expressions frequently appearing in modern live commerce and examine whether richer linguistic features can improve predictive performance.

Funding

This work was supported by JSPS KAKENHI (grant number 24K05159).

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] DataReportal, “Digital 2025: Global Overview Report,” https://www.datareportal.com, accessed in 2025.

[2] S. Wang and H. Zhang, “Industrial information integration through autonomous AI agents: Paradoxical effects on transparency, dehumanization, and responsible retail operations,” Journal of Industrial Information Integration, ,vol. 51 article no. 101089, 2026.

[3] Grand View Research, “Live Commerce Market Size, Share & Trends Analysis Report,” https://www.grandviewresearch.com, accessed in 2024.

[4] China Association of Performing Arts, “Annual Report on China’s Live Streaming and Short Video Industry,” https://www.capa.com.cn/, accessed in 2025.

[5] E-Commerce Research Center of China, “China Live Commerce Development Report,” http://www.100ec.cn, accessed in 2024.

[6] P&S Intelligence, “China E-Commerce Market Size, Share, Growth and Analysis Report,” https://www.psmarketresearch.com/market-analysis/china-e-commerce-market-report, accessed in 2024.

[7] L. Wang and X. Zhang, “Livestream Sales Prediction Based on an Interpretable Deep-Learning Model,” Scientific Reports, vol. 14, article no. 20594, 2024.

[8] Y. Jiang, “The Impact of Interaction on Customer Engagement in Live Commerce,” Shogaku Kenkyu Ronshu, no. 55, pp. 39-55, 2021. http://hdl.handle.net/10291/22001. (in Japanese)

[9] K. Kishiya and J. Wang, “A Study on Impulse Buying Under Time Constraints in Live Commerce: Effects of Word-of-Mouth and Consumer Knowledge,” The Business Review of Kansai University, vol. 68, no. 4, pp. 37-53, 2024. (in Japanese).

[10] C. Lou and S. Yuan, "Influencer Marketing: How Message Value and Credibility Affect Consumer Trust of Branded Content on Social Media," Journal of Interactive Advertising, vol. 19, no. 1, pp. 58-73, 2019.

[11] Y. Lin and K. Minamikawa, "The Influence of Influencers on Consumer Behavior in Live Commerce," Nanzan Management Review, vol. 37, no. 2, pp. 167-185, 2022. https://nanzan-u.repo.nii.ac.jp/records/4226. (in Japanese)

[12] G. Ilieva, T. Yankova, M. Ruseva, Y. Dzhabarova, S. Klisarova-Belcheva, and M. Bratkov, "Social Media Influencers: Customer Attitudes and Impact on Purchase Behaviour," Information, vol. 15, no. 6, article no. 359, 2024.

[13] S. Wang and H. Zhang, "Generative AI in International Hotel Marketing: Impacts on Employee Creativity and Performance," International Journal of Contemporary Hospitality Management, vol. 37, no. 8, pp. 2601-2626, 2025.

[14] F. Xu, Z. Pan, and R. Xia, "E-Commerce Product Review Sentiment Classification Based on a Naïve Bayes Continuous Learning Framework," Information Processing & Management, vol. 57, no. 5, article no. 102221, 2020.

[15] S. Chen, S. Ke, S. Han, S. Gupta, and U. Sivarajah, "Which Product Description Phrases Affect Sales Forecasting? An Explainable AI Framework by Integrating WaveNet Neural Network Models with Multiple Regression," Decision Support Systems, vol. 176, article no. 114065, 2024.

[16] S. Wang and Z. Sun, "Roles of artificial intelligence experience, information redundancy, and familiarity in shaping active learning: Insights from intelligent personal assistants," Education and Information Technologies, vol. 30, no. 2, pp. 2525-2546, 2025.

[17] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," Advances in Neural Information Processing Systems, vol. 30, pp. 4768-4777, 2017.

[18] L. Xu, H. Lin, Y. Pan, H. Ren, and J. Chen, "Construction of an Emotion Lexicon Ontology," Journal of the China Society for Scientific and Technical Information, vol. 27, no. 2, pp. 180–185, 2008. https://ir.dlut.edu.cn/info/1013/1142.htm. (in Chinese)