Trends in Social Sciences and Humanities Research
Print ISSN: 2959-992X
Online ISSN: 2959-9938
DOI: https://doi.org/10.61784/tsshr3176

A DEEP LEARNING DRIVEN ANALYSIS OF THE NON-LINEAR AND INTERACTIVE EFFECTS OF TITLE EMOTION AND VIDEO LENGTH ON DEPTH OF COMMUNICATION IN BILIBILI

YaWei He*, LinYi Bao
School of Media and Communication, Shanghai Jiao Tong University, Shanghai 200240, China.
Corresponding Author: YaWei He, Email: heyawei1992@sjtu.edu.cn

Abstract: In the competitive digital landscape of online video platforms, understanding the drivers of user engagement is paramount for content creators and platform strategists. This study investigates the complex, non-linear relationships between video title sentiment intensity, video duration, and user engagement, measured by like counts, within the automotive content niche on the Chinese platform Bilibili. Drawing on a dataset of 892 videos collected via web scraping, this research employs a multi-method analytical approach, combining OLS regression with advanced techniques including quantile regression and SHAP analysis based on an XGBoost model. The findings reveal that both title sentiment intensity and video duration have significant, non-linear (U-shaped and quadratic, respectively) effects on the natural logarithm of like counts. Specifically, videos with either very low (neutral) or very high (emotional) title sentiment intensity tend to receive more likes than those with moderate intensity. Furthermore, a significant interaction effect is uncovered, with the Johnson-Neyman analysis indicating that the effect of sentiment intensity is significantly moderated by video duration. Quantile regression results show that these effects are heterogeneous across different levels of video popularity, suggesting that the drivers of engagement for viral content differ from those for average videos. While the overall model explains a modest portion of the variance, the identified non-linear and interactive patterns challenge simplistic linear assumptions and provide nuanced, actionable insights. The study contributes to the field of computational communication by demonstrating a sophisticated analytical framework for dissecting engagement metrics and offers practical guidance for content strategy in vertical interest communities.
Keywords: User engagement; Sentiment intensity; Video duration; Bilibili; Computational communication; Non-linear effects; Interaction effects; SHAP

1 INTRODUCTION

The proliferation of online video platforms has fundamentally reshaped information dissemination and user interaction, establishing itself as a core battleground for audience attention [1]. Platforms such as YouTube, TikTok, and Bilibili have evolved into complex ecosystems where billions of users consume, create, and engage with content, making user engagement a critical currency for creators, marketers, and the platforms themselves. Engagement, often operationalized through metrics like views, likes, comments, and shares, not only signifies audience appreciation but also directly influences content visibility through algorithmic amplification [2]. Consequently, identifying the determinants of user engagement has become a central pursuit in digital communication research.

This study focuses on Bilibili, a unique and influential platform in the Chinese digital media landscape. Unlike its global counterparts, Bilibili has cultivated a distinct community culture, primarily centered around its "Generation Z" user base [3]. Several characteristics make Bilibili a compelling context for this research. First is its signature "Danmu" (or "bullet comments") system, where real-time user comments float across the screen, fostering a sense of co-viewing and collective experience that enhances participation [4]. Second, the platform's ecosystem is highly dependent on the relationship between content creators, known as "UPzhu" (literally "uploader"), and their fans, creating strong community ties and high user stickiness [5]. Finally, Bilibili hosts a vast array of deep, vertical-interest content, moving beyond general entertainment to specialized fields. The automotive niche, in particular, represents a highly competitive and valuable market where creators produce in-depth reviews, test drives, and technical analyses, demanding sophisticated content strategies to capture audience engagement [6].

Within this context, content creators face strategic dilemmas regarding content production and presentation. Two of the most fundamental and controllable elements are the video's duration and its title. Video length is a constant balancing act between providing substantive content and retaining audience attention in an era of declining attention spans [7]. Video titles, as the primary textual gateway to the content, serve to manage expectations, convey information, and evoke emotion to entice clicks and engagement. The emotional tone, or sentiment, of a title is a powerful tool in this regard [8].

While a body of literature has explored the effects of sentiment and content length on user engagement, research has often relied on linear models and has underexplored the nuanced, potentially non-linear and interactive relationships between these variables. It is plausible, for instance, that the effect of a highly emotional title is not uniform but depends on the length of the video it represents—a long-form documentary may benefit from a different titling strategy than a short, punchy clip. Moreover, the nature of sentiment itself is complex; much research has focused on valence (positive vs.

negative), while the impact of emotional intensity (arousal) is less understood. A neutral, objective title and a highly emotional one may be more effective than a mildly emotional one, suggesting a U-shaped or curvilinear relationship [9].

To address these gaps, this study undertakes a deep computational analysis of 892 automotive videos from Bilibili. It moves beyond traditional linear assumptions to investigate the following research questions:

RQ1: What are the individual effects of video duration and title sentiment intensity on user engagement (measured by like counts)?

RQ2: Are the relationships between video duration, title sentiment intensity, and user engagement non-linear?

RQ3: Does video duration moderate the relationship between title sentiment intensity and user engagement?

To answer these questions, this study employs a multi-faceted analytical approach. We begin with descriptive and bivariate analyses, followed by a hierarchical OLS(ordinary least squares) regression to model the main, non-linear, and interactive effects. We then probe the significant interaction using the Johnson-Neyman technique. To further enhance the robustness and depth of our findings, we utilize quantile regression to explore how these effects vary across different levels of video popularity and employ SHAP (SHapley Additive exPlanations) on a trained XGBoost model to provide machine learning-based insights into feature importance and interactions [10].

This research aims to make several contributions. Theoretically, it challenges simplistic linear models of user engagement and provides empirical evidence for the complex, curvilinear, and interactive nature of content characteristics' effects. Methodologically, it showcases a comprehensive workflow combining traditional statistical inference with advanced computational techniques to yield a more holistic understanding. Practically, the findings offer nuanced, data-driven insights for content creators in the automotive niche and beyond, helping them optimize their titling and content length strategies to maximize user engagement on platforms like Bilibili.

2 LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

2.1 User Engagement on Digital Media Platforms

User engagement is a cornerstone concept in digital communication, broadly defined as the cognitive, affective, and behavioral investment a user makes in their interactions with a media object or platform [11]. In the context of social video platforms, engagement is typically operationalized through a suite of observable behavioral metrics, including views, likes, comments, and shares [12]. While views may indicate reach, "thicker" engagement metrics like likes and comments are often interpreted as more active forms of audience feedback and approval [13]. Likes, in particular, serve as a low-cost, immediate signal of positive reception, and their aggregation functions as a powerful social proof heuristic that can influence subsequent viewers' perceptions and engagement [14]. Furthermore, these engagement signals are critical inputs for platform recommendation algorithms, which prioritize and promote content that demonstrates a high potential for user interaction, creating a feedback loop where engagement begets visibility [15].

2.2 The Role of Content Characteristics: Duration and Titling

The effect of video duration on engagement is complex and contested. Some research suggests that in an economy of attention, shorter content is more effective, capturing users before their attention wanes [16]. This is exemplified by the rise of short-form video platforms like TikTok. However, other studies find a positive relationship between video length and certain engagement outcomes, particularly on platforms like YouTube that reward watch time [17]. Longer videos may afford creators the opportunity to build a more compelling narrative, provide more in-depth information, and foster a stronger parasocial relationship with the viewer, leading to higher overall satisfaction and engagement. This suggests that the optimal video length is likely context- and platform-dependent. In this study, we hypothesize a positive but potentially non-linear relationship, where engagement increases with duration up to a certain point before plateauing or declining as videos become excessively long [18]. This is reflected in our decision to test for a quadratic effect of duration.

As the primary textual cue, the video title is instrumental in setting expectations and driving click-through behavior [12]. The emotional content of text, or sentiment, has been shown to be a potent driver of information diffusion and engagement. Research on textual sentiment often focuses on valence (the positivity or negativity of the emotion). Studies have found evidence for both a "positivity bias," where positive content is shared more widely, and a "negativity bias," where negative content can be more arousing and attract more attention [8].

However, a focus solely on valence may overlook the role of emotional arousal or intensity—the degree of emotional activation a piece of content evokes, regardless of its positive or negative direction. High-arousal emotions (e.g., awe, anger, anxiety) have been found to be more viral than low-arousal emotions (e.g., sadness, contentment) [19]. This study operationalizes sentiment in terms of intensity, arguing that in the crowded attention market of video platforms, the primary function of a title's emotion is to be arresting and arousing. We hypothesize a non-linear, U-shaped relationship between sentiment intensity and engagement. Titles with very low intensity may succeed by appearing objective, informative, and professional, while titles with very high intensity may succeed by being emotionally provocative and acting as effective "clickbait." Titles with moderate, lukewarm emotionality may fail to stand out, thus receiving lower engagement. This is represented by our first hypothesis:

H1: The relationship between title sentiment intensity and like counts will be non-linear (U-shaped).

2.3 Interaction Effects and a Moderated Model

The effects of individual content characteristics do not operate in a vacuum. It is crucial to consider how they interact to jointly influence user engagement. The effectiveness of a particular titling strategy, for example, may be contingent on the nature of the content it represents. An intensely emotional title might seem appropriate for a short, dramatic clip but could feel mismatched or exhausting for a 30-minute technical deep-dive. Conversely, the investment required to watch a long video might be better justified by a title that promises high informational value (low sentiment intensity) or high emotional payoff (high sentiment intensity).

This leads to the concept of moderation, where the relationship between a predictor (sentiment intensity) and an outcome (like counts) is altered by the level of a third variable (video duration). We propose that video duration moderates the effect of sentiment intensity on engagement. This aligns with a contingency perspective on media effects, suggesting that there is no single "best" content strategy, but rather that optimal strategies depend on the interplay of multiple factors. Therefore, we posit our second hypothesis:

H2: Video duration will moderate the relationship between title sentiment intensity and like counts.

By testing these hypotheses, this study aims to provide a more nuanced and comprehensive model of user engagement, accounting for the non-linear and interactive ways in which content characteristics shape audience behavior in a modern digital media environment.

3 METHODS

3.1 Data Collection and Sample

The data for this study were collected from Bilibili (bilibili.com), a prominent video-sharing platform in China. A dataset was compiled by executing a custom web scraper using Python, targeting videos within the automotive content category. The data collection period spanned from July 2019 to December 31, 2024, to ensure a wide representation of videos over time.

The initial sample consisted of videos primarily focused on automotive topics, including vehicle reviews, test drive experiences, model comparisons, and brand events. A data cleaning process was then implemented to ensure data quality. Videos were excluded if they had poor image quality, lacked essential information about the vehicle or presenter, or had audio-video synchronization issues. After this filtering process, the final sample consisted of N = 892 unique videos. The sample includes content from creators of varying sizes and audience levels, enhancing the generalizability of the findings within the Bilibili automotive community.

3.2 Measures

3.2.1 Dependent variable

User engagement was operationalized as the total number of "likes" a video received. The raw distribution of like counts was found to be heavily right-skewed, a common characteristic of engagement metrics where a small number of videos achieve viral popularity. This positive skewness violates the assumptions of ordinary least squares (OLS) regression. As shown in Figure 1 the mean like count (31,614) was substantially higher than the median (19,507), confirming the presence of high-value outliers. To normalize the distribution, a natural logarithm transformation was applied to the like count variable (ln(Like Count + 1) to handle any cases of zero likes). The resulting log-transformed variable, Log_Like_Count, was approximately normally distributed (Mean = 9.88, Median = 9.88, Skewness = 0.10) and served as the dependent variable in all regression analyses.

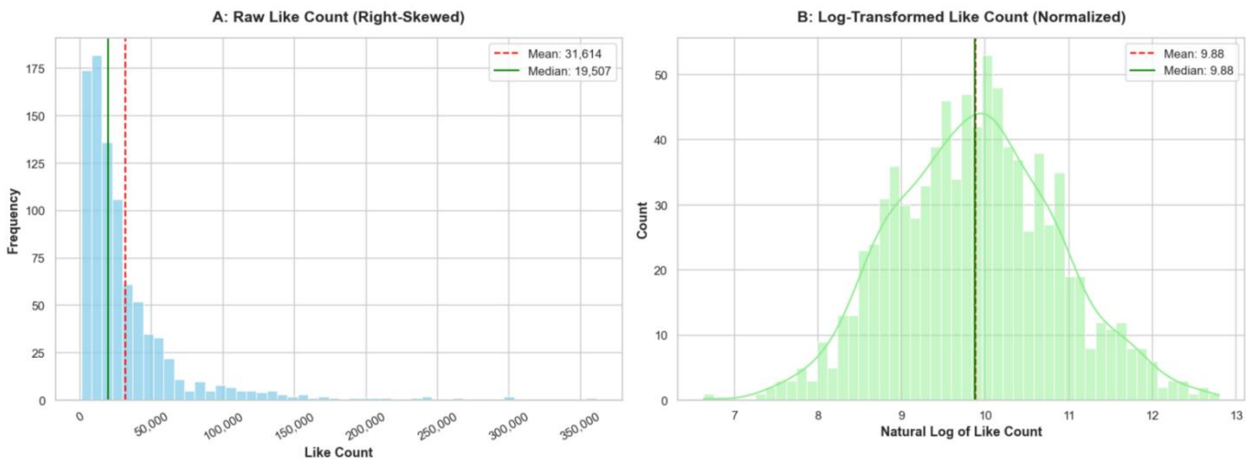

Figure 1 Like Count Distribution Before and After Log Transformation: (A) the raw, right-skewed distribution of like counts, and (B) the normalized, bell-shaped distribution after log transformation

3.2.2 Independent variables

Video Duration. This was measured in total seconds, from the start to the end of the video. The duration of videos in the sample ranged from 26 seconds to 1,436 seconds, with a mean of 649.57 seconds (approx. 10.8 minutes). The distribution of video duration was slightly right-skewed, as depicted in Figure 2.

Title Sentiment Intensity. This key variable was derived through a two-step process to measure the degree of emotional arousal in a video's Chinese title. First, a pre-trained multilingual BERT (Bidirectional Encoder Representations from Transformers) sentiment classification model was employed. This model, trained on a five-star rating system, analyzed each video title to produce a raw sentiment valence score (Sraw) on a continuous scale from [0, 1], where 0 represented the most negative sentiment, 1 represented the most positive sentiment, and 0.5 represented neutrality. Second, to capture emotional intensity irrespective of valence, the raw valence score was transformed using the following formula: Sintensity = 2 × |Sraw - 0.5|. This calculation converts the U-shaped valence scale into a linear intensity scale. Scores near the neutral midpoint of 0.5 on the valence scale result in an intensity score near 0 (low arousal), while scores at the extremes (near 0 or 1) result in an intensity score near 1 (high arousal). This final Sintensity score was used in the analysis.

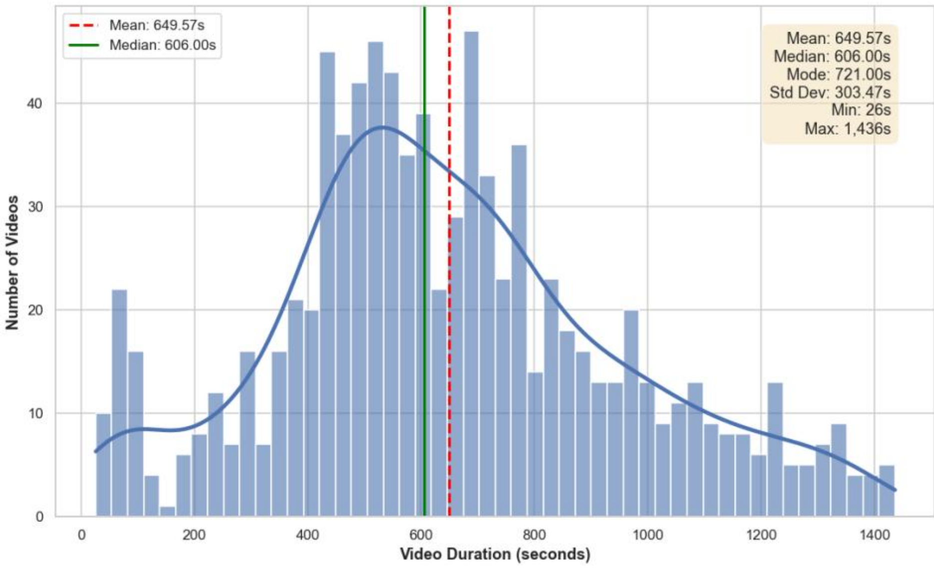

Figure 2 Distribution of Video Duration

3.3 Analytical Strategy

The data analysis was conducted in several stages. First, descriptive and bivariate analyses were conducted. The initial bivariate relationship between video duration and Log_Like_Count was assessed using a Pearson correlation coefficient, visualized in Figure 3. Videos were then grouped into bins based on duration and sentiment intensity to visualize average like counts, as shown in Figure 4, Figure 5 and Figure 6. Further visualizations, including a hexbin density plot (see Figure 7) and a binned heatmap (see Figure 8), were used to explore the joint distribution of the independent variables.

Second, to formally test the hypotheses, a hierarchical OLS regression analysis was performed. The final model (Model D) included main effects, quadratic terms to test for non-linear relationships (H1), and an interaction term to test for moderation (H2), as visualized in Figure 9. Model diagnostics were performed (see Figure 10). Third, a significant interaction was probed using the Johnson-Neyman technique (see Figure 11). Finally, advanced supplementary analyses, including quantile regression (see Figure 12) and SHAP analysis on an XGBoost model (see Figure 13 and Figure 14), were used to add depth and robustness.

4 RESULTS

4.1 Descriptive and Preliminary Analyses

Preliminary analysis revealed a weak but statistically significant positive linear correlation between video duration and the log-transformed like count ($r = .071$, $p = .035$), as shown in Figure 3. This suggests that, overall, longer videos tend to garner slightly more likes.

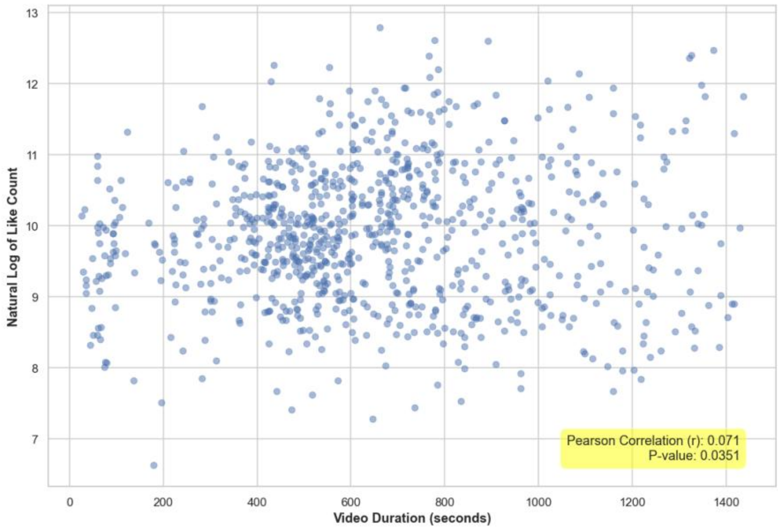

Figure 3 Relationship between Video Duration and Log-Transformed Like Count

A more detailed examination using binned analysis provided richer insights. As shown in Figure 4, there is a clear monotonic trend of increasing average like counts as video duration increases. Videos shorter than 1 minute had the lowest average likes (11,458), while those longer than 20 minutes had the highest (43,281). However, when considering engagement efficiency, this trend reverses. Figure 5 demonstrates that shorter videos are far more efficient at accumulating likes per minute. "Shorts" (< 1 min) averaged 19,618 likes per minute, a figure that drops sharply to 1,978 for videos over 20 minutes long.

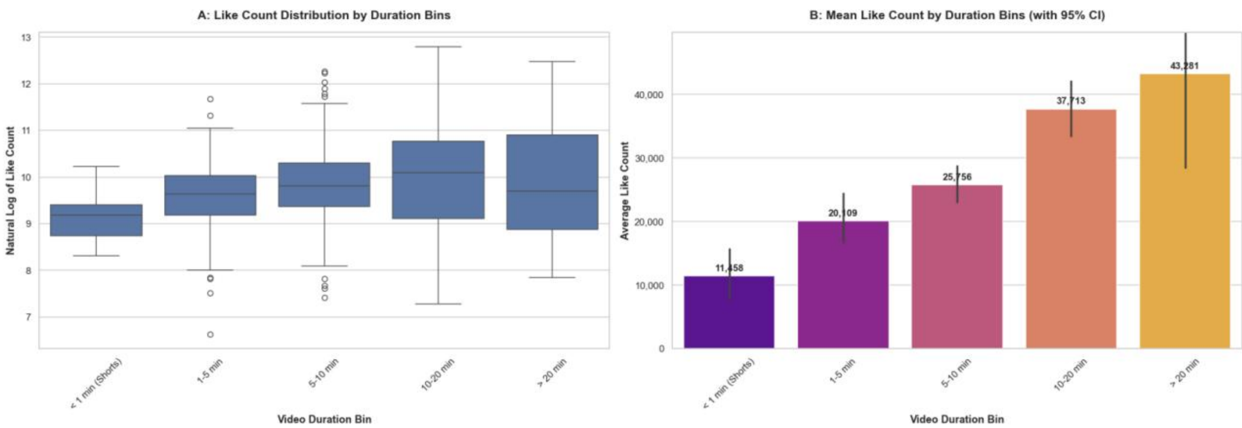

Figure 4 Like Count Analysis by Binned Video Duration

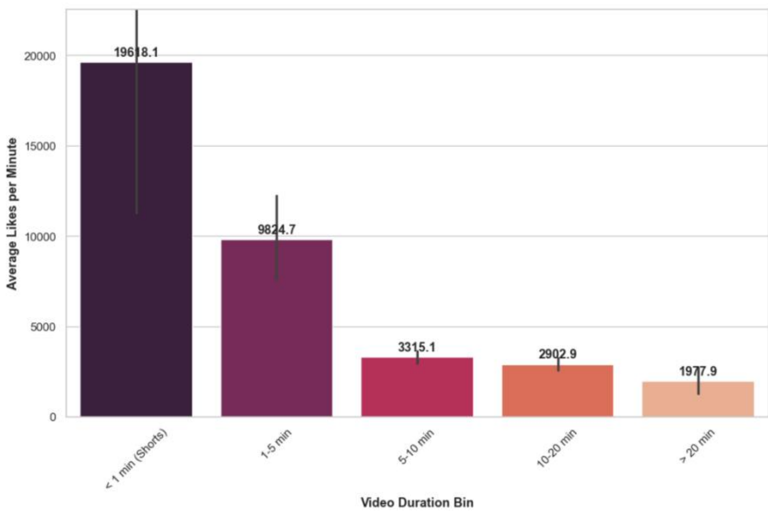

Figure 5 Like Efficiency by Video Duration

The analysis of sentiment intensity bins (see Figure 6) revealed a non-linear pattern. Average like counts were highest at the extremes: 34,351 for the "Very Low" intensity bin (0-0.2) and 41,026 for the "Very High" intensity bin (0.8-1.0), providing initial support for the U-shaped hypothesis (H1). However, a one-way ANOVA indicated these differences were not statistically significant, $F(4, 887) = 1.53$, $p = .192$.

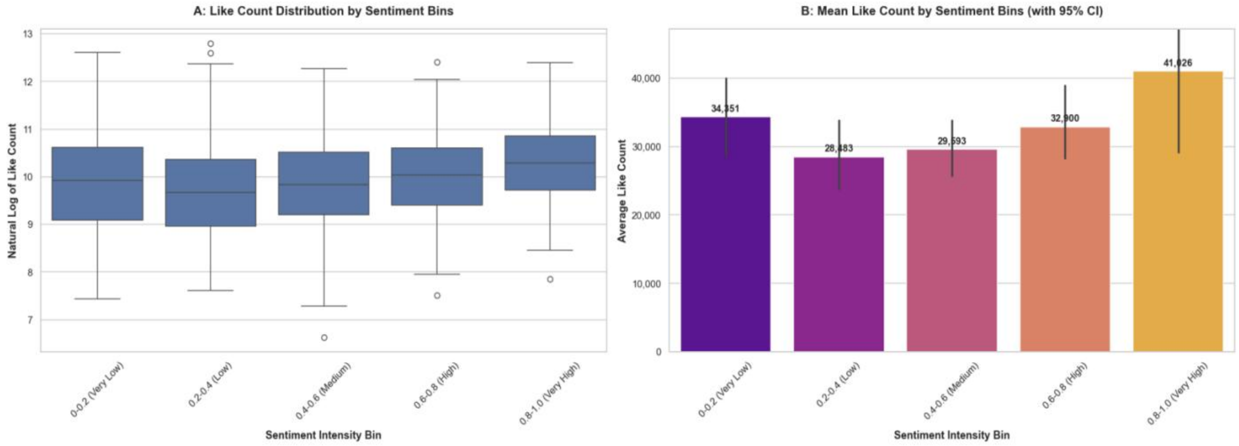

Figure 6 Like Count Analysis by Binned Sentiment Intensity

To visually explore the joint relationship and potential interaction between the independent variables, a hexbin density plot and a binned heatmap were generated. The hexbin plot (Figure 7) maps the joint distribution of sentiment intensity and video duration, with color indicating the mean log-transformed like count. It reveals that the highest concentrations of engagement (brighter colors) are most frequently observed for videos with a moderate duration (approximately 200-800 seconds) and a moderate-to-high sentiment score (approximately 0.4-0.7), suggesting a complex, non-linear relationship rather than a simple trend toward the extremes.

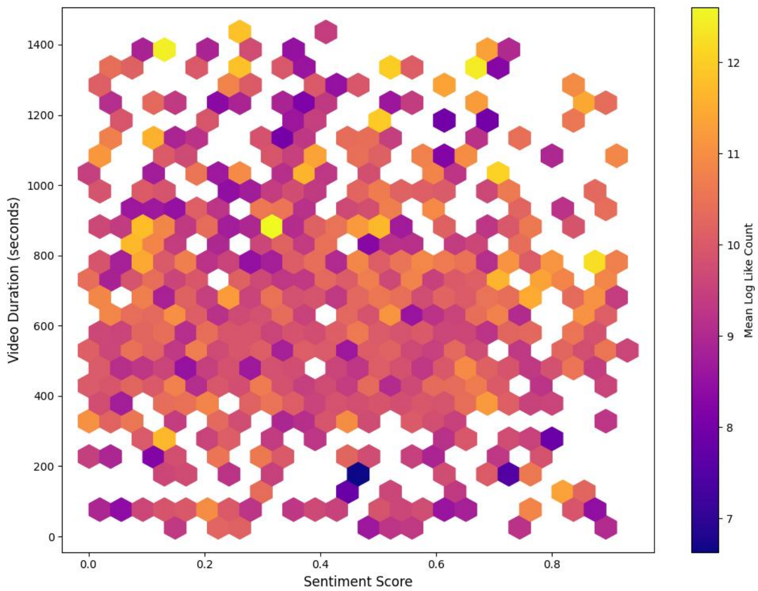

Figure 7 Hexbin Density Plot

To clarify this interaction with raw engagement numbers, the binned heatmap (Figure 8) displays the average actual like count for discrete bins of both variables. This visualization powerfully illustrates a non-monotonic interaction pattern, showing that the effect of sentiment intensity is highly conditional on video length. For instance, the single highest-performing combination was for videos longer than 20 minutes with a high sentiment score (0.6-0.8), garnering an average of approximately 65,910 likes. In stark contrast, the same high sentiment score in the shortest videos (< 3 minutes) resulted in one of the lowest outcomes, with only 12,698 average likes. This strong visual evidence that the effectiveness of a titling strategy depends fundamentally on video duration motivates the formal moderation analysis in the subsequent regression model.

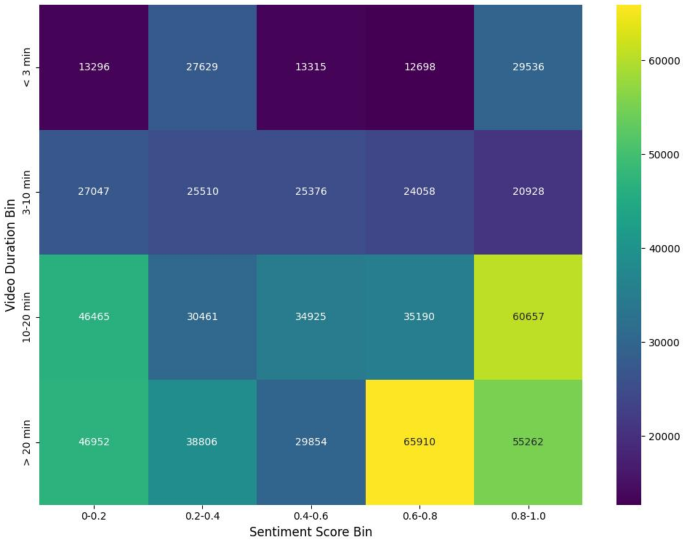

Figure 8 Binned Heatmap (Actual Mean Like Count)

4.2 Hierarchical Regression Analysis

A hierarchical OLS regression was performed to formally test the hypotheses. The final model (Model D) provided the best fit, with an Adjusted R^2 of .033. While modest, this indicates the model explains approximately 3.3% of the variance in log-transformed like counts. Diagnostic checks (Figure 9) confirmed the model assumptions were reasonably met.

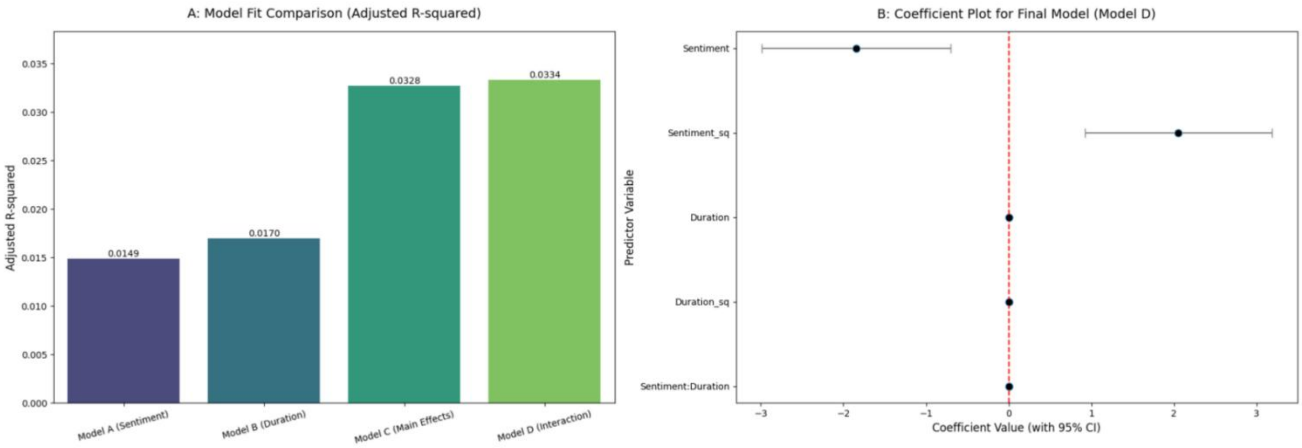

Figure 9 Hierarchical Regression Results

The coefficients for the final model (see Figure 9) revealed significant findings. The coefficients for 'Sentiment_Intensity'($b = -1.846$, $p = .002$) and 'Sentiment_Intensity²' ($b = 2.054$, $p < .001$) were both significant, describing a convex, U-shaped curve and confirming H1. Similarly, the coefficients for 'Duration' ($b = 0.0013$, $p = .002$) and 'Duration²' ($b \approx -0.0000$, $p < .001$) were significant, describing an inverted U-shape. Crucially, the interaction term 'Sentiment_Intensity' × 'Duration' was statistically significant ($b = 0.0006$, $p < .05$), supporting H2.

To ensure the validity of the OLS regression results, a series of model diagnostics were performed, as shown in Figure 10. The Residuals vs. Fitted plot (Panel A) and the Scale-Location plot (Panel C) show that the residuals are randomly scattered around the horizontal line without any obvious patterns, satisfying the assumptions of linearity and homoscedasticity. Furthermore, the Normal Q-Q plot (Panel B) reveals that the residuals closely follow the theoretical diagonal line, indicating that the assumption of normality is meet. While the Cook's Distance plot (Panel D) identifies 51 observations with potential influence (Cook's distance > 4/n), none have a Cook's distance greater than 1, suggesting that no single data point excessively biases the model. Overall, the diagnostic plots confirm that the final regression model (Model D) is robust and its assumptions are well-satisfied.

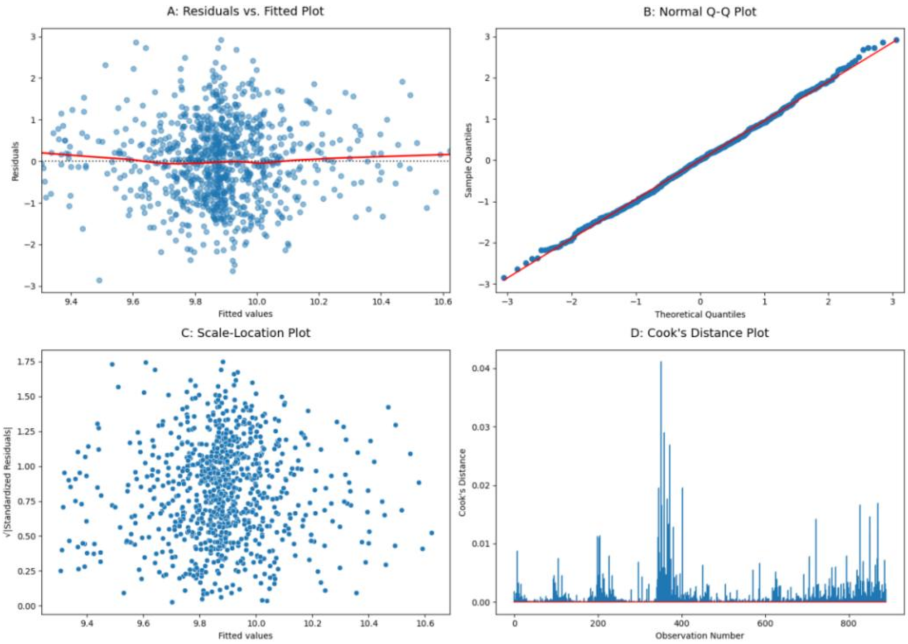

Figure 10 Diagnostic Plots for Final Model (Model D)

4.3 Probing the Interaction: Johnson-Neyman Analysis

The Johnson-Neyman (J-N) analysis was conducted to interpret the significant moderation effect (Figure 11). The analysis revealed that for videos with a duration between 26 seconds and approximately 1,209 seconds (~20 minutes), the marginal effect of sentiment intensity is significantly negative. This indicates that for most videos, as sentiment intensity increases, the predicted like count initially decreases (the downward slope of the U-curve). The positive moderating effect means this negative slope becomes less severe as duration increases. For videos longer than 1,209 seconds, the effect was not significant.

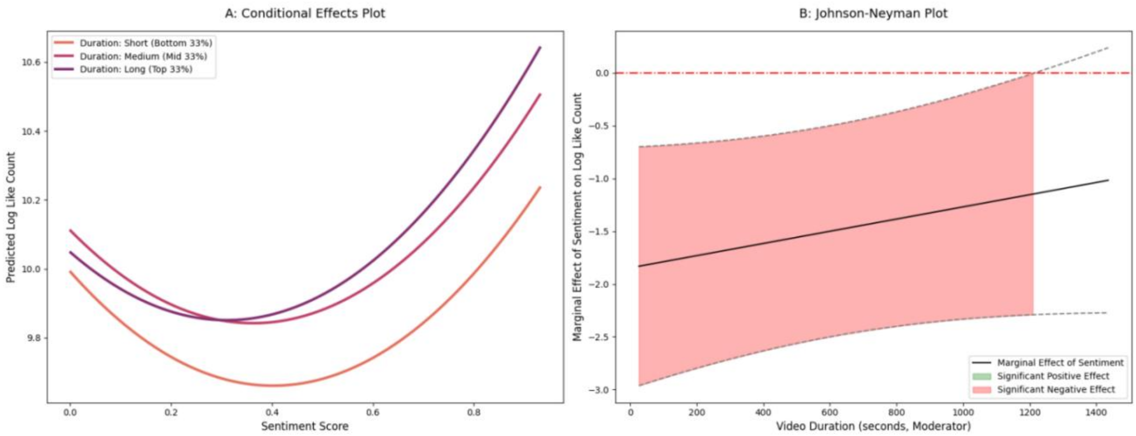

Figure 11 Interaction Visualization and Region of Significance

4.4 Advanced Analyses: Quantile Regression and SHAP

To add depth and robustness to the OLS findings, two advanced analytical techniques employed. First, quantile regression used to investigate whether the interaction effect was consistent across videos with varying levels of popularity. The results, shown in Figure 12, reveal that the effect is highly heterogeneous and not stable across the popularity distribution. For median-performing videos (at the 50th percentile), the interaction coefficient was significantly positive, indicating that for this group, longer duration amplifies the positive impact of extreme sentiment. However, in a striking reversal, the coefficient becomes significantly negative for slightly more popular videos (at the 60th percentile), suggesting that the strategic interplay of sentiment and duration that helps a video achieve average popularity may be different from the strategy needed to push it into a higher tier of engagement.

Second, to complement the statistical inference from OLS, an explainable AI approach used by training a high-performance XGBoost model and interpreting it with SHAP values. The SHAP summary plot (Figure 13) provides a global ranking of feature importance. Consistent with the regression model, it identified Sentiment (mean absolute SHAP value = 0.3654) as the marginally most influential feature in driving predictions, closely followed by Duration (mean absolute SHAP value = 0.3579). The wide horizontal spread of SHAP values for both features visually confirms their substantial and complex impact. Furthermore, the SHAP interaction dependence plot (Figure 14) directly visualized the synergy between the two variables. It confirmed a slight positive average interaction effect (mean interaction SHAP value $\approx$ 0.0087), corroborating the core moderation hypothesis from a machine learning perspective. More specifically, the plot revealed that this positive synergy is most pronounced for long videos (represented by red points) that also have high-intensity titles, providing convergent evidence for the conditional nature of these effects.

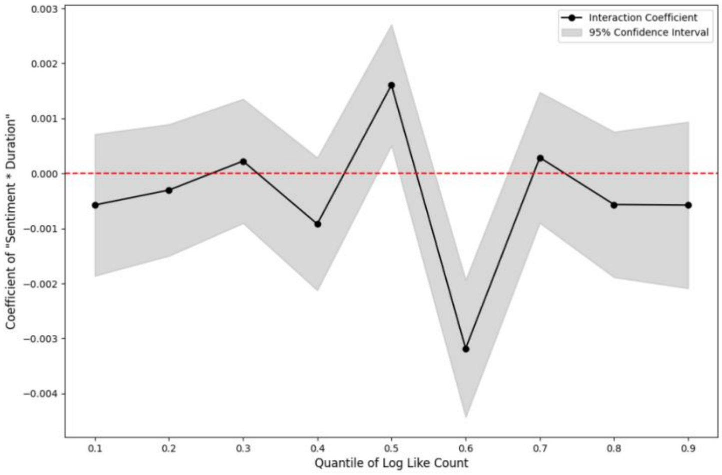

Figure 12 Quantile Regression Coefficients for the Interaction Term

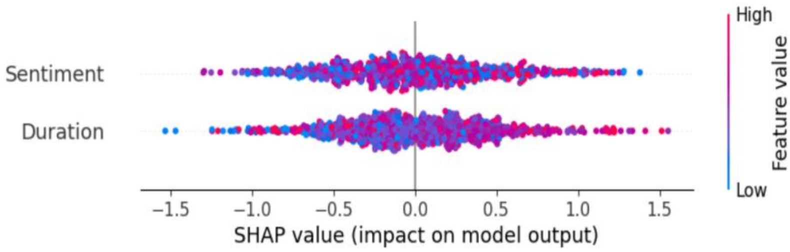

Figure 13 SHAP Summary Plot

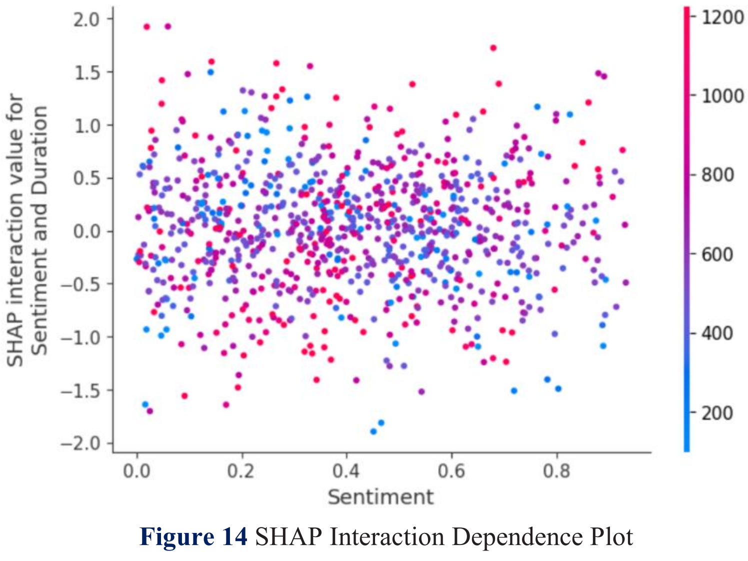

Figure 14 SHAP Interaction Dependence Plot

5 DISSCUSSION

5.1 Summary and Interpretation of Key Findings

This study set out to unravel the complex relationships between video title sentiment intensity, video duration, and user engagement on Bilibili. First, the results robustly demonstrate that the relationships of both title sentiment intensity and video duration with like counts are non-linear. For sentiment intensity, we found a significant U-shaped effect, confirming H1. This suggests that content titling strategies may succeed at the extremes: very low intensity (objective, professional) and very high intensity (emotionally provocative) titles both perform well, while moderately emotional titles are associated with the lowest engagement.

For video duration, we found a significant, though diminishing, positive effect. Longer videos tend to get more likes, but the returns diminish, indicating a potential ceiling effect. This, combined with the "like efficiency" analysis, presents a strategic choice for creators between maximizing total engagement with longer content and producing efficient, easily digestible shorter content.

Second, and central to this study, is the confirmation of a significant moderation effect (H2). Video duration alters the relationship between sentiment intensity and engagement. The J-N analysis revealed that for the majority of videos (up to ~20 minutes), the dominant effect of increasing sentiment intensity is negative. This implies that for a typical-length video, a more neutral, fact-driven title is a safer strategy than a moderately emotional one. The interaction's positive coefficient indicates that as videos get longer, this negative effect is attenuated, possibly because a longer video provides the necessary space to justify a highly emotional title's claim.

Finally, our advanced analyses revealed the heterogeneity of these effects. Quantile regression showed that the nature of the interaction is not consistent across the popularity spectrum. Furthermore, the SHAP analysis corroborated the findings from the OLS regression, confirming that sentiment intensity and duration are the two most critical features and that they interact in meaningful ways.

5.2 Theoretical and Practical Implications

Theoretically, this study provides strong empirical evidence against the adequacy of simple linear models for studying user engagement, highlighting the necessity of testing for curvilinear and interactive effects. It also refines our understanding of sentiment's role by focusing on ‘intensity’(arousal) rather than just ‘valence’, revealing a U-shaped pattern that suggests a duality of successful communication strategies: one rooted in objectivity and information (low arousal), and another in emotion and provocation (high arousal).

Practically, the implications for content creators are significant. The key takeaway is that there is no one-size-fits-all strategy. First, titling is context-dependent: the choice between a neutral or an emotional title should be made in conjunction with video length. Second, creators must balance total engagement with efficiency, choosing a video length that aligns with their specific goals. Third, for titling, the least effective strategy appears to be moderate emotionality; titles should be either clearly informative or clearly emotional.

5.3 Limitations and Future Research

This study has several limitations. First, the final regression model explained a relatively small portion of the variance in like counts (Adjusted $R^2 \approx .033$), indicating that other variables play a much larger role. Future research should incorporate creator-level variables such as subscriber count. Second, the data is correlational, precluding causal claims; experimental designs would be necessary to establish causality. Third, our findings are specific to the automotive niche on Bilibili and may not be generalizable to other genres or platforms. Finally, our operationalization of sentiment intensity, while an improvement on simple valence, could be further refined by classifying discrete emotions.

6 CONCLUSION

This study demonstrates that the relationship between video content characteristics and user engagement is far more complex than linear models suggest. For automotive videos on Bilibili, the path to engagement is not a straight line but a landscape of shifting curves. Our findings pinpoint a significant U-shaped relationship for title sentiment intensity, revealing that both highly objective (low-intensity) and highly emotional (high-intensity) titles significantly outperform those with moderate emotionality. Similarly, while longer videos tend to accumulate more likes, they do so at a diminishing rate and at a steep cost to engagement efficiency. The central contribution of this research is the identification of a significant interaction effect: the power of an emotional appeal is fundamentally conditional on the temporal commitment a video asks of its audience. Our analysis reveals that for the vast majority of videos in this niche (under approximately 20 minutes), the initial effect of increasing a title's emotional intensity is, counterintuitively, negative. This complex interplay is vividly illustrated by our finding that a high-intensity title paired with a long video (>20 minutes) can yield over five times more likes (averaging ≈65,910) than the exact same titling strategy applied to a short video (<3 minutes), which averaged only ≈12,698 likes. By employing a diverse computational toolkit, from OLS regression to SHAP analysis on an XGBoost model, this research provides a robust, nuanced understanding of these dynamics, offering both theoretical depth for communication scholars and concrete, data-informed heuristics for digital content creators navigating the ever-evolving attention economy.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The authors are grateful for the finance support from the Ministry of Education Industry-Academia Collaborative Education Program (No.231104575131914), General Program of National Social Science Fund of China (No. 22BXW040) and Shanghai Jiao Tong University Participation in Research Program (No. T200PRP42015).

REFERENCES

[1] Lin Xiao, Xiaofeng Li, Jian Mou. Exploring user engagement behavior with short-form video advertising on short-form video platforms: a visual-audio perspective. Internet Research, 2024, 34(7): 2325-2350.

[2] Stappen L, Baird A, Lienhart M, et al. An Estimation of Online Video User Engagement From Features of Time- and Value-Continuous, Dimensional Emotions. Frontiers in Computer Science, 2022(4): 773154.

[3] Chen Z, Yang DL. Governing Generation Z in China: Bilibili, bidirectional mediation, and technological politics. The Information Society, 2023, 39(2): 82-94.

[4] Pan X. Motivations for game stream spectatorship: A content analysis of Danmaku on Bilibili. Global Media and China, 2023, 8(2): 190-212.

[5] Zeng D. Danmu subtitling as a self-regulative practice: a descriptive discourse analysis of Bilibili danmu subtitles from ethical perspectives. Frontiers in Psychology, 2025(16): 1577260.

[6] Sun C, Zhou D, Yang T. Sponsorship disclosure and consumer engagement: Evidence from Bilibili video platform. Journal of Digital Economy, 2023(2): 81-96.

[7] Violot C, Elmas T, Bilogrevic I, et al. Shorts vs. Regular Videos on YouTube: A Comparative Analysis of User Engagement and Content Creation Trends. Proceedings of the ACM Web Science Conference 2024. Association for Computing Machinery, New York, NY, USA. 2024, 213-223. DOI: https://doi.org/10.1145/3614419.3644023.

[8] Cui G, Chung S Y-H, Peng L, et al. Clicks for money: Predicting video views through a sentiment analysis of titles and thumbnails. Journal of Business Research, 2024(183): 114849.

[9] Chou YC. How much is too much? The nonlinear link between emotional arousal and review helpfulness. Decision Support Systems, 2023(175): 114035.

[10] Nisa MUN, Mahmood D, Ahmed G, et al. Optimizing Prediction of YouTube Video Popularity Using XGBoost. Electronics, 2021, 10(23): 2962.

[11] Boujena O, Ulrich I, Manthiou A, et al. Customer engagement and performance in social media: A managerial perspective. Electronic Markets, 2021, 31(4): 389-405.

[12] Cheng Z, Li Y. Like, Comment, and Share on TikTok: Exploring the Effect of Sentiment and Second-Person View on the User Engagement with TikTok News Videos. Social Science Computer Review, 2024, 42(3): 567-589.

[13] Tandoc EJ Jr, Vos T. The framing of video news narratives: A content analysis of online news broadcasts. Journalism & Mass Communication Quarterly, 2019, 96(4): 1045–1068.

[14] Savvas Zannettou, Olivia Nemes-Nemeth, Franziska Roesner, et al. Analyzing User Engagement with TikTok's Short Format Video Recommendations using Data Donations. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA. 2024, 1–16. DOI: https://doi.org/10.1145/3613904.3642433.

[15] Chuqi Wang, Zhi Li. Unraveling the relationship between audience engagement and audiovisual characteristics of automotive green advertising on Chinese YouTube. PLoS ONE, 2024, 19(4): e0299496.

[16] Anu Alamäki, Choong-Ki Rhee, Jari Suomala, et al. Creating effective visuals for destination marketing videos: Scenery vs people. Journal of Hospitality and Tourism Technology, 2023, 14(3): 456-475.

[17] Wondwesen Tafesse. YouTube marketing: How marketers' video optimization practices influence video views. Journal of Business Research, 2021(127): 252-263.

[18] Matthis Boer, Gonneke W. J. M. Stevens, Catrin Finkenauer. The complex association between social media use intensity and depressive symptoms: A longitudinal study. Computers in Human Behavior, 2022(127): 107197.

[19] Jinyoung Jinnie Yoo, Heejin Kim, Sungchul Choi, et al. Expanding knowledge on emotional dynamics and viewer engagement: The role of travel influencers on YouTube. Journal of Innovation & Knowledge, 2024, 9(4): 100550.