

Emotional Analysis Strategy and Application of Internet Celebrity Marketing Under the Background of Big Data

Langxuan Zhang

College of Business, Shanghai University of Finance and Economics, Shanghai, 200433, China

ABSTRACT

The deepening development of the digital economy has made internet celebrity marketing the core path for brands to reach users. According to the "2025 China Internet Celebrity Marketing Solution Industry Report" by Boyan Consulting, the domestic market size will reach 680 billion yuan in 2024. However, the traditional model relies on traffic indicators such as exposure and clicks, ignores emotional appeals, and leads to pain points such as "high traffic, low conversion" and "weak user stickiness". Big data technology provides a path to crack pain points, mining users' multi scenario emotional tendencies through sentiment analysis, and helping them achieve the transformation from "traffic driven" to "emotion driven". This paper combs the integration status quo, clarifies the emotional analysis technology links (data collection, tagging, modeling), refines three core strategies, and analyzes their applications in live video delivery, short video operation, and brand reputation maintenance by using the literature research method in combination with Taobao Live emotional monitoring, Tiktok content emotional adaptation and other cases. To address the challenges of noise data accounting for 40% (supported by IDC), emotional ambiguity, and privacy compliance, optimization paths such as multimodal fusion, privacy computing, and industry standards are proposed. Research has confirmed that sentiment analysis helps internet celebrities and brands accurately capture demand, driving conversion rates to increase by 12% -20% (HypeAuditor 2024 Internet Celebrity Marketing Insights Report), providing support for their sustainable development.

KEYWORDS

Big data; Influencer marketing; Sentiment analysis; User sentiment modeling; Marketing effectiveness optimization

1. INTRODUCTION

Internet technology drives the evolution of marketing form. From TV advertising, search engine marketing to online celebrity marketing, the core logic is to "narrow the distance between brands and users". As of June 2024, there are 830 million live streaming users and 1.05 billion short video users in China (respectively cited from the "2024 Live Streaming and Short Video Industry Report" and the 54th Statistical Report of CNNIC), laying a solid foundation for internet celebrity marketing. But with the intensification of industry competition, the crux of traditional internet celebrity marketing has become apparent: some internet celebrities focus on "traffic gimmicks", and the mismatch between content and user emotions has caused resistance; Brands rely on surface data such as fan base to select partners, ignoring emotional feedback, resulting in resource depletion. Xiao Rui's (2022) research on the college student population also confirms this point - through a questionnaire survey, it was found that 38.2% of college students engage in irrational consumption due to "excessive emotional agitation and disconnection between content and demand" during live streaming, resulting in a subsequent return rate of 27.5%, directly reflecting the resource waste and user experience

problems of traditional marketing models [1]. The rise of big data has solved this dilemma. Its "4V" features (massive, high-speed, diverse, low value density) can integrate the full chain data of internet celebrity marketing, covering user browsing trajectory, interactive behavior, purchase, and after-sales data, providing abundant data sources for sentiment analysis. As a core branch of natural language processing (NLP), sentiment analysis can transform users' unstructured emotional expressions into structured labels (positive, negative, neutral), helping to decode users' "emotional passwords". Some platforms and brands have already tested the waters: Taobao Live has launched real-time emotion monitoring, using barrage emotions to adjust language; Huaxizi selects "high emotional resonance" experts for collaboration based on emotional analysis. However, industry applications are still in the "fragmented" stage, lacking system strategies and facing challenges in data quality and privacy compliance. Based on this, this article focuses on the research of "Emotional Analysis of Internet Celebrity Marketing under the Background of Big Data", aiming to clarify the technical path and practical strategies, solve the difficulty of capturing emotional needs, provide practical references, and promote the transformation of Internet Celebrity Marketing towards "precision and emotionalization".

2. THE INTEGRATION STATUS OF BIG DATA AND INTERNET CELEBRITY MARKETING

The deep penetration of big data technology is reconstructing the entire chain of influencer marketing - from influencer selection and content creation to user engagement and effectiveness analysis, all require data as the underlying support. Clarifying the current state of integration between the two is the cornerstone of designing subsequent sentiment analysis strategies.

From the perspective of the evolution of internet celebrity marketing, it has moved from "individual IP combat" to the stage of "industrialization and dataization". The early model relied on empirical decision-making, with brands selecting cooperative internet celebrities based on their "fan base" and "external image". Content creation was centered around the personal preferences of internet celebrities, and the evaluation of effectiveness only focused on surface indicators such as "sales volume" and "views", failing to accurately anchor user needs and easily leading to problems such as "conflicting tone between internet celebrities and brands" and "content being ignored by audiences". In the context of big data, this situation has greatly improved: platforms rely on user behavior data to build "user profiles", and brands can choose partners with high target user overlap based on this; During the content creation phase, data tools can analyze the emotional tendencies of recent highly interactive content, providing creative anchors for internet celebrities; The effectiveness evaluation has shifted from a "single data dimension" to a "multi-dimensional matrix", covering user stay time, interaction rate, repurchase rate, and even tracking "user emotional feedback after purchase", building a marketing closed loop.

The empowering value of big data for internet celebrity marketing is mainly reflected in three aspects. Firstly, refine the user profile. Taking Tiktok's "huge cloud map" as an example, it can aggregate users' age, region, consumption ability and "emotional content preference", helping online celebrities anchor core customer groups. As Wang J et al. (2022) found through over 340000 Weibo and WeChat posts, the interaction rate of users with "humorous" content is 27% higher than that of "promotional" content, providing empirical support for refining emotional tags [2]. Secondly, intelligent content distribution. The algorithm can dynamically adjust the distribution strategy based on user emotional feedback: if users interact frequently with "humorous" content, their exposure weight will be increased; If there is negative feedback on "exaggerated propaganda" content, reduce its exposure ratio. Thirdly, the marketing effectiveness can be traced. Brands can use big data tools to track user emotional changes in marketing activities in real time. For example, during live streaming sales, the proportion of negative emotions such as "high price" in product reviews increases, which can optimize promotional plans in a timely manner and avoid user churn.

It should be noted that there is still a problem of "emotional dimension omission" in the current integration of the two: most data tools focus on "behavioral data" and overlook "emotional data". According to a 2024 survey by iResearch Consulting, 70% of brands in internet celebrity marketing have not conducted systematic analysis of user emotional feedback, and only 30% of top brands have attempted emotional data monitoring, with limited application scenarios. This "behavior heavy, emotion light" model makes it difficult for internet celebrity marketing to deeply reach the user's mind, and also reserves practical space for sentiment analysis.

3. THE CORE TECHNOLOGY PATH OF SENTIMENT ANALYSIS IN INTERNET CELEBRITY MARKETING

Real Time Sentiment Analysis for Live Streaming Marketing: To implement influencer marketing, it is necessary to follow the entire process of "data collection emotion annotation model construction result application", clarify the technical points and paradigms of each link, ensure accurate and practical analysis, and avoid disconnection between technology and business. Data collection focuses on full scene multi-source coverage - user emotions are dispersed across multiple platforms, and a single data source is difficult to capture completely. It is divided into three categories: social interaction data (microblog comments, Tiktok bullet screen, and Xiaohongshu messages), and real-time feedback of emotions; E-commerce feedback data (such as "product reviews" and "asking everyone" on Taobao and JD.com), linked to purchasing behavior, can guide marketing more effectively; Private domain data (online celebrity fan group records, official account messages) are private and authentic, reflecting the needs of core fans. The acquisition technology focuses on "API docking" and "compliance crawler": public platforms (Tiktok, Weibo) borrow open APIs to retrieve data, such as Tiktok's "Massive Engine Open Platform" comment interface, to obtain compliance information; Private domains and other unopened API scenarios require compliant crawlers and compliance with the Personal Information Protection Law. After authorization, only emotional texts can be provided. Jabbar et al. (2019) pointed out that real-time capture of review data through open APIs on e-commerce platforms, such as the Amazon Product Advertising API, can ensure data compliance and timeliness. Similarly, the comment interface of Tiktok's "Massive Engine Open Platform" (such as /data/comments/list) can realize the millisecond level collection of live broadcast bullet screen, and dynamically adjust the content distribution strategy in combination with the emotion analysis model [3]. After collection, initial cleaning is required to screen out invalid and duplicate data. According to the "2024 Marketing Data Governance White Paper", the data efficiency after cleaning needs to reach over 80%. Emotional annotation requires transforming unstructured text into structured labels to serve as a benchmark for the model, with the core being to establish a unified paradigm to eliminate subjective errors. The commonly used dimensions for annotation are classified into three levels: positive (like, recommended), negative (poor quality, not recommended), and neutral (just received, average logistics). Some scenarios can be subdivided and adapted as needed: live streaming uses a three-level speed guarantee, and after-sales segmentation excavates the root cause of dissatisfaction. The annotation adopts "human-machine collaboration": it is manually completed by a special training team, and cross checked after annotation (reviewed by 2-3 people and judged by the leader of the disagreement group); The machine borrows the basic model for initial calibration and then manually corrects it. According to the "Application Report of Natural Language Processing Technology in Marketing Scenarios (2024)", this model can accelerate by 3-5 times and achieve an accuracy rate of over 90% after correction. Model construction needs to balance scene selection algorithms - emotional models as the core driver, live streaming requires fast response, and after-sales requires high precision. There are two types of commonly used algorithms: traditional machine learning (SVM, Naive Bayes), which has a lightweight architecture, fast training, and is suitable for live streaming. For example, Taobao Live uses SVM algorithm to split barrage emotions in 1 second; Deep learning (RNN, LSTM), capturing context and breaking ambiguity, adapting to after-sales and content optimization, such as using LSTM to analyze comments and select internet

celebrities for beauty brands. After model training, it needs to be evaluated according to the "AI Marketing Technology Application Specification (2024)", with an accuracy rate of over 85% before use; User emotions evolve over time, and the model iterates every 3-6 months to avoid bias.

4. EMOTIONAL ANALYSIS STRATEGY DESIGN FOR INTERNET CELEBRITY MARKETING UNDER THE BACKGROUND OF BIG DATA

The technological path provides "tool support" for sentiment analysis, and strategy design is the core of transforming it into "marketing effectiveness". Combining the characteristics of real-time interaction and diverse content in internet celebrity marketing, sentiment analysis strategies can be developed from two dimensions: "layered adaptation" and "real-time response".

4.1. User Sentiment Stratification and Content Adaptation Strategy

User sentiment stratification uses sentiment analysis to classify fans into three categories and target operations: According to the "2024 Internet Celebrity Marketing User Sentiment Report" by iResearch Consulting, positive emotional users account for 30% -40% of the core customer group. Exclusive coupons and member live broadcasts are used to promote repeat purchases, and top beauty anchors achieve a repeat purchase rate of over 50% for core users; Neutral emotional users account for 40% -50%, building trust through product testing and word-of-mouth collections. Beauty influencers rely on fan feedback to increase their interaction rate by more than 20%; Negative emotional users account for 10% -20%. Firstly, identify the causes of dissatisfaction (such as logistics and allergies), and use refund compensation and special content to repair emotions. Domestic beauty brands promote special policies for allergy feedback, and the conversion effect is improved by more than 10% [4].

Content emotional adaptation is divided into pre planning and post optimization: in the early stage, users' emotional tendencies can be judged by tools, such as planning healing content when users are anxious (during the epidemic, top internet celebrities promoted home funny videos with single views of tens of millions); Optimize content based on emotional feedback in the later stage, such as food influencers who, due to user feedback of "slow pace", shorten the time and add key shots, with positive emotional feedback accounting for over 80%. The core is centered on user emotions, avoiding content self talk, shifting content production from "subjective judgment" to "data-driven", and strengthening the emotional connection between users, brands, and internet celebrities.

4.2. Real Time Emotional Response Strategy

This strategy focuses on live streaming sales and is implemented in three steps: first, real-time monitoring, relying on tools to integrate bullet comments, updating the proportion of emotions every second and setting warnings (such as triggering reminders if the negative proportion exceeds 40%), such as Taobao Live providing real-time emotion data visualization function for the anchor's backend; The second is dynamic adjustment. When positive emotions climb, push limited time flash sales. When neutral, increase actual test sharing. When negative, respond immediately (negotiate discounts if there is a price dispute). If top anchors borrow real-time price adjustments, sales of a certain product can exceed 50000 units in a short period of time; The third is to review the effectiveness, correlate emotional changes with marketing actions, and optimize subsequent strategies. The core is to grasp the "emotional window period" and transform emotional advantages into marketing results.

5. THE PRACTICAL APPLICATION SCENARIOS OF SENTIMENT ANALYSIS IN INTERNET CELEBRITY MARKETING

The sentiment analysis strategy needs to be implemented in mainstream marketing scenarios such as live streaming sales, short video operations, and brand reputation maintenance in order to unleash practical value and provide a reusable practice paradigm for the industry. Live streaming sales is a scenario with outstanding conversion efficiency, but it has a high real-time threshold and limited user retention time. Sentiment analysis can solve the pain points of weak conversion and high user churn rate. During the product selection stage, deep exploration of user emotional demands can be carried out, such as top live streaming operators using sentiment analysis to screen high resonance products. The accuracy of product selection and live streaming conversion efficiency have been significantly improved; In the process of optimizing anchor language, the expression logic is optimized by comparing the emotional feedback of different language styles. The new anchor team uses this to optimize user acceptance and live streaming duration; After live streaming, follow-up visits can also be conducted based on the emotional tendencies of users' post purchase evaluations. After optimization, the user retention time and conversion efficiency in live streaming scenes are usually significantly improved. Short video operations are deeply trapped in the dilemma of content homogenization, and sentiment analysis can help anchor differentiation paths: the topic selection process focuses on current high emotional resonance themes, such as a top lifestyle influencer adjusting the focus of content through sentiment analysis, focusing on emotional topics that users pay attention to, and significantly expanding the playback volume and fan base of individual content; The interactive process relies on sentiment analysis to identify the emotional tendencies of comments and respond to user demands in a targeted manner, such as classifying comments with different emotional tendencies by automotive influencers, significantly optimizing comment response rates and user stickiness. Cross platform research shows that the interaction rate of comments is 37% higher than that of dithering due to users' preference for in-depth technical analysis (Huang Bingyi, 2024). Therefore, in Station B, the auto vertical online celebrity can give priority to respond to professional issues such as "chassis adjustment", and Tiktok focuses on intuitive topics such as "appearance design" [5] . At the same time, the content style can be iterated based on emotional feedback to effectively improve the dissemination indicators such as the playback volume and likes of short videos. In brand reputation maintenance, sentiment analysis can monitor users' emotional tendencies at all times to deal with public opinion. In response to negative public opinion, for example, when a health product recommended by an internet celebrity receives negative feedback from users, the brand uses sentiment analysis to timely capture the signs of negative emotions, quickly take measures such as product removal, popular science interpretation, and after-sales compensation, effectively curbing the spread of negative public opinion; In terms of building positive word-of-mouth, emotional analysis is used to deeply explore the core value points recognized by users. For example, a new emerging beverage brand focuses on the product characteristics that receive positive reviews from users, organizes internet celebrities to produce a collection of word-of-mouth content, and subsequently achieves growth goals in brand favorability and product sales. The core is to prevent negative public opinion in advance, strengthen positive emotional cognition, and build a brand emotional moat.

6. CHALLENGES AND OPTIMIZATION DIRECTIONS OF SENTIMENT ANALYSIS IN INTERNET CELEBRITY MARKETING

6.1. Core Challenges of Sentiment Analysis Applications

Although sentiment analysis has demonstrated practical value in internet celebrity marketing, it faces three core challenges in the actual implementation process, which directly hinder its effectiveness release. One reason is that the quality of data varies greatly, with emotional data coming from social

media platforms and e-commerce scenarios, naturally carrying a large amount of noisy data (such as meaningless comments, malicious smearing content, and internet buzzwords and puns). Such data can interfere with analysis results, and the proportion of noise is always high, especially in high attention areas such as beauty and clothing; The second is that there are barriers to recognizing emotional ambiguity, and users' emotional expression is constrained by context. The semantic interpretation of ironic expressions is more difficult, and such misjudgments are particularly significant in the context of internet celebrity marketing; The third issue is the prominent risk of privacy compliance. Some platforms and internet celebrities engage in excessive or unauthorized collection of emotional data (such as capturing user browsing trajectories and sensitive information), directly crossing the red lines of the Personal Information Protection Law and the Data Security Law, which can easily trigger user complaints and brand trust crises.

6.2. Optimization Direction of Sentiment Analysis

Anchoring the above challenges, optimization can be promoted from four aspects: firstly, introducing multimodal data fusion technology to fuse non text data such as text, images, and live voice to reduce noise interference. The multimodal sentiment analysis model launched by leading technology institutions achieves a significant leap in analysis accuracy compared to a single text model through the fusion of multiple types of data; Secondly, iteratively upgrading the context aware model and pre trained language model, analyzing the context before and after data of a single comment, deepening the semantic understanding of network buzzwords and ironic expressions. Previous studies have captured emotional turning points in text through the "context emotion" bidirectional attention mechanism (such as "product is good but poor after-sales service"), and the F1 score has increased by 12.3% in social comment analysis (Liu, Bo, et al., 2020); Some platforms currently integrate pre trained models such as BERT on this basis, further improving the accuracy of sentiment recognition for hot words such as "Juejuezi" and "YYDS" to over 90%, significantly enhancing the ability to recognize complex expressions [6]. Some platforms use emotion analysis tools optimized based on pre trained language models, which significantly improve their ability to recognize complex expressions compared to traditional models; Thirdly, implementing privacy computing technologies (such as federated learning) to achieve "data availability but invisibility". Alibaba, Tencent, and other companies have launched pilot applications in internet celebrity marketing related scenarios, and need to collaborate with the industry to develop compliance guidelines for sentiment analysis data, clarifying data collection boundaries and authorization standards; Fourthly, establish an industry standard system led by professional organizations such as the China Advertising Association, in collaboration with platforms, brands, and technology enterprises, to clarify the dimensions of sentiment classification, model evaluation standards, and core bottom lines of data compliance, and promote industry applications towards standardization and normalization.

7. CONCLUSION

This article focuses on the research of emotional analysis strategies and applications in internet celebrity marketing under the background of big data. It summarizes the current status of their integration, emotional analysis technology paths, core strategies, and practical optimization directions. The following conclusions are drawn: firstly, big data provides a data foundation and technical support for emotional analysis in internet celebrity marketing. The traditional "traffic driven" model is difficult to match users' emotional needs. Big data can integrate multiple sources of emotional data and rely on sentiment analysis to transform unstructured expressions into structured labels, resolving pain points such as "high traffic, low conversion" and "weak user stickiness". Industry practice has shown that the application of sentiment analysis in internet celebrity marketing significantly improves conversion efficiency and user retention time, confirming its practical value. Secondly, the implementation of sentiment analysis requires a dual wheel drive of "technical path+strategic design".

Technically, multi-source collection ensures comprehensive coverage, emotional annotation establishes standards, adapts algorithm models, and updates them regularly; Strategically, user sentiment stratification, content sentiment adaptation, and real-time sentiment response form a complete application loop to avoid disconnection between technology and business. Thirdly, current sentiment analysis faces challenges in data quality, sentiment ambiguity, and privacy compliance, which can be gradually resolved through multimodal fusion, context model optimization, privacy computing, and industry standards. In the future, it will move from "fragmented applications" to "full chain penetration", becoming a core marketing capability for internet celebrities and promoting the industry's transformation towards "emotionalization". Under the trend of the industry, users' demand for emotional resonance is increasing. Brands, internet celebrities, and platforms need to pay attention to emotional analysis, increase investment and strategy iteration, build an "emotional driven" model centered on user emotions, strengthen compliance awareness, and ensure sustainable development. This article can provide reference for the industry, but it does not break down industry differences. In the future, targeted strategies can be studied.

REFERENCES

- [1] Xiao Rui Research on the Impact of Live Streaming Marketing on College Students' Irrational Consumption under the Internet Celebrity Economy [D]. Kunming University of Science and Technology, 2022. DOI: 10.27200/d.cnki.gkmlu.2022.000861.
- [2] Wang J, Yang Y, Liu Q, etc. Empirical Study on User Participation in Weibo and WeChat Influencer Marketing [J]. IEEE Journal of Computational Social Systems, 2022, 10 (6): 3228-3240.
- [3] Jabbar J, Urooj I, JunSheng W, etc. Real time sentiment analysis for e-commerce applications [C]//2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC). IEEE, 2019: 391-396.
- [4] Galagra PH, Veloso A, Meira Jr W, etc. From Bias to Viewpoint: Transfer Learning Methods for Real time Sentiment Analysis [C]//Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2011: 150-158.
- [5] Huang Bingyi Emotional analysis of pan knowledge short video comments: cross platform comparative research based on Tiktok and Station B [J]. New Media Research, 2024, 10 (18): 28-35. DOI: 10.16604/j.cnki.issn2096-0360.2024.18.020.
- [6] Liu, Bo, et al. "Situational Perceived Social Media User Emotion Analysis". Tsinghua University Science and Technology 25.4 (2020): 528-541.