Word Cloud Analysis of Live Shopping User Comments Based on Python Web Crawling Technology

Boya Wang

School of Economics and Trade, Hao Jing College of Shaanxi University of Science and Technology, Xi'an, Shaanxi, 712046, China
miaboya@qq.com

Abstract: In recent years, live streaming e-commerce has experienced rapid development, with the number of live streaming platforms and their audience sizes growing exponentially. Online shopping and live streaming have greatly altered people's purchasing behavior and strengthened their willingness to buy. However, concurrently, the term "consumerism trap" has gained prominence, suggesting that consumers may fall into traps while participating in live shopping, prompting reflections on the relationship between live shopping and consumerism. This paper attempts to conduct a word cloud analysis using text information obtained through Python web scraping technology from Weibo. The goal is to remind consumers to make rational purchases and protect their legitimate rights, and to provide businesses with insights to create a favorable environment for live streaming e-commerce, influencing consumer purchasing decisions and promoting rational buying.

1. Research Background and Literature Review

The live shopping trend on internet platforms has been rapidly gaining momentum, with 2019 being dubbed the "year of online live streaming." Weibo, Douyin, Kuaishou, Taobao, JD.com, and even the emerging Pinduoduo have all enhanced their live streaming features, making live streaming an unstoppable trend. According to iiMedia Research, the number of online live streaming users in China is expected to reach 660 million by 2022. With the development of internet platforms and vertical applications, the live streaming industry is expected to witness diverse applications, leading to an increase in potential user base. Among various applications, live e-commerce remains the main profit model in the online live streaming industry, with China's live e-commerce market reaching 961 billion yuan in 2020. "Live shopping" is a new shopping method that emerged after 2016, integrating e-commerce, live streaming, and social interaction. It unlocks a new shopping scene where users can place orders while watching videos. While watching shopping live streams, users can click on product links on the screen to access the product details page, opening up a new entry point for shopping. Compared to traditional online shopping, which mainly relies on text and images to describe product information, live shopping uses video streaming to showcase product information to users, providing a more immersive shopping experience through real-time interaction between hosts and users. As a result, many users have gradually developed a

preference for purchasing products in live streams, and live shopping has entered a stage of rapid growth.However, amidst the fervor of live shopping development, questions arise: Are we ensnared by consumerism traps? Have we purchased goods that we don't actually need, thereby falling into consumerism traps? Consumerism is defined as "a value orientation that advocates and pursues excessive possession and consumption as a way to satisfy self and life goals, and the behavioral practices under this value orientation [1]." Thus, consumerism is not merely a value orientation or a behavioral practice but a combination of both.

The Theory of Planned Behavior, proposed by American social psychologist Ajzen in 1985 and further developed in 1991, has been widely applied to explain and predict individual behavior. According to this theory, all factors considered by individuals before actual action influence behavioral intention through attitudes, subjective norms, and perceived behavioral control [2-4]. Thus, consumers' attitudes and perceptions toward live shopping will affect their purchasing behavior. Whether in traditional or online markets, consumers exhibit herd behavior, whereby individuals follow the actions of others [5]. Herd behavior refers to the social phenomenon of individuals acting according to the behavior of others, which is more pronounced in online markets than in traditional markets [6-7].

Under the rapid development of various online information dissemination platforms such as Sina Weibo, Zhihu, Twitter, etc., more people are discussing current hot topics, expressing their positions and attitudes, and commenting on sensitive social issues, leading to a dramatic increase in data volume on social media platforms. More and more users tend to express their opinions, attitudes, views, and emotions on public platforms. Although the vast amount of data seems trivial and irregular, it holds immense research value. Live shopping platforms offer resources like "limited-time promotions." Zhang Xiaoyu (2020) believes that in e-commerce, online users often experience "information loss" during online shopping, lacking sufficient knowledge and time to make rational decisions [8]. Yin Meng et al. (2019) argue that consumers are more likely to follow others' purchasing behavior while shopping online, leading to herd behaviour [9]. Wang Lei (2021) suggests that with the rapid development of internet technology, advertising videos with bullet comments can increase consumers' attention to products. The positivity or negativity of bullet comments affects consumers' perception of value and risk, thus influencing their purchasing intentions [10].

Therefore, this paper attempts to integrate the relevant text information of live shopping through Python web scraping technology and word cloud analysis, aiming to explore current consumer perceptions of live streaming e-commerce and provide a basis for online shopping and marketing for businesses.

2. Word Cloud Analysis Based on Python Web Scraping Technology

This paper selects Weibo posts as the research objects for text mining, using data scraping techniques to capture relevant text information, analyze user sentiment, and provide feedback to businesses and the live streaming industry. Using Python, 2000 pieces of text data were scraped, and 20 examples are listed in Table 1.

Table 1: Examples of Text Data

Author Code	Posting Time	Post Content
A1	Mar-23-16:21:03	Starting today to break out of the consumerism trap.
A2	Mar-23-15:16:46	Teacher: This is a consumer trap! This is serving capital! Advertising sets standards and disciplines consumers! Me: Stop cursing.
A3	Mar-23-15:37:02	I fell into the consumer trap, deceived by the outer packaging and cheap price. This

		coffee is more bitter than my life, and it has a Chinese medicine smell. So annoyed!
A4	Mar-23-15:40:30	In this era where traffic is king, whoever has traffic embraces success. ... Short videos have become people's first choice for entertainment, socializing, shopping, etc. In this situation, ordinary people need to be vigilant against the consumer traps on the internet.
A5	Mar-23-20:50:11	One thing I admire about her is that she hasn't been kidnapped by luxury goods. Luxury goods always come with an unavoidable sense of superiority, weaving a consumerism trap for the masses... Countless consumers are kidnapped by it, buying items priced for their logos. Real consumption isn't being realized...
A6	Mar-22-12:54:23	In today's entanglement with consumerism, every manufacturer meticulously calculates consumer preferences and packages them... Desperately creating consumer traps 'exclusive' to women...
A7	Mar-22-22:53:22	Not being able to shop in malls feels like nothing's happening, and not being able to shop online makes me feel bad. After replying to deliveries, the first thing is probably retaliatory consumption. When will things go back to normal?
A8	Mar-22-22:27:51	This month I've spent a lot less because there's nowhere to shop even if I wanted to. Can't go out to buy, and online shopping isn't delivering.
A9	Mar-22-10:32:45	I complained a lot about the inconvenience of not being able to shop online. But now that it's been a while since I've done any online shopping, it's made me rethink the meaning of consumption. Some things really aren't necessary to buy.
A10	Mar-22-07:45:32	Haven't done any online shopping for over half a month now. It feels really good and comfortable not to have that impulsive desire to spend money anymore! Also, I've noticed that I'm eating less for dinner these days. I hardly have any appetite, and I feel full after just a few bites. I wonder if I'm going to lose weight.
A11	Mar-22-00:03:28	Day two of not shopping online, just checking in.
A12	Mar-22-22:39:20	Just a few days into this and I'm already going crazy. I want to drink tea from DaDaoGuMing, have some McDonald's, receive a delivery, indulge in some crazy online shopping spree. Six days at home and there's basically nowhere to spend money. I'm so miserable.
A13	Mar-21-17:50:01	I've noticed that lately my spending habits revolve around eating, online shopping, and foot massages. Is this normal? Why do I feel like I'm wasting away?
A14	Mar-20-22:06:55	I want to shop online. I want to spend money. I want to buy new clothes. It's so hard.
A15	Mar-20-11:12:39	The suspension of delivery services is very conducive to consumption downgrading—because there is simply no way to shop online, as deliveries are not being made.
A16	Mar-19-19:35:09	These days, it's rare to see anyone practicing filial piety like the 'Pinduoduo filial child'. These rogue online shopping platforms squeeze merchants, engage in false advertising to consumers, and deliberately recruit new users for cash rewards. Apart from the associated interest chains, who among ordinary people would come out to whitewash the stench of these rogue Pinduoduo-like behaviors?
A17	Mar-19-09:58:23	Never browse online shopping at night; it's all impulsive buying. Waking up in the morning, I deeply regret it.
A18	Mar-18-19:01:37	I'm really fed up. I can't control myself from buying all sorts of things even though I'm short of money. This behavior is despicable. Starting from today, I won't shop online anymore except for essential items.
A19	Mar-22-14:24:40	Except for some daily expenses and regular consumption, if we don't consume new things, we'll be less likely to be deceived, right? Increasing vigilance is also an important aspect.
A20	Mar-22-12:17:00	"Bought eight skirts... Beware of consumption traps."

Table 2: Word Frequency Analysis

Rank	World	Frequency
1	Trap	276
2	Shopping	255
3	Consumption	190
4	Network	50
5	Onlineshopping	49
6	Customer	48
7	Vigilant	38
8	Link	37

The crawler technology preprocesses the generated text data to generate the analytical text, and

then performs deduplication and noise reduction, word segmentation. The crawler generates the final document and performs a frequency analysis, selecting the top 8 words, as shown in Table 2.

At the same time, the crawler technology summarizes the text information, carries on the visualization operation, and generates the text cloud. The original word cloud analysis diagram is shown in Figure 1. And the word cloud diagram of the top eight words by frequency is shown in Figure 2.

Figure 1: Word Cloud Analysis Diagram

Figure 2: Word Cloud Analysis Diagram of top 8 worlds

3. Results and Discussions

The paper focuses on the topic of live shopping and conducts a word cloud analysis based on Weibo microblog data obtained through Python web scraping technology. It reveals that among the evaluations and sentiments expressed by consumers on the Weibo platform regarding live shopping, "trap," "shopping," "consumption," "network," "customer," "vigilance," and "link" are the top eight high-frequency words. This also indicates that in the rapid development of livestream shopping today, consumers gradually realize the consumerism trap inherent in livestream shopping through multiple purchases, leading to feelings of regret after purchase. Consequently, this reduces their subsequent purchases, thereby influencing consumer purchasing decisions.In consumer society, businesses have reached unprecedented levels of emphasis and propaganda regarding consumption, using various techniques to create desires and needs in consumers, driving them to seek more goods and engage in unconscious shopping behavior. The word frequency analysis and word cloud analysis show the continuous appearance of advertisements and promotion traps in online shopping platforms. These traps make it difficult for consumers to perceive the harm to their own interests.

Seckill, free shipping, buy one get one free, trial use, and discounts are frequently employed tactics on online shopping platforms. However, many online shoppers still find it difficult to resist the temptation of these traps, willingly falling into them. The reasons behind this lie in the satisfaction of diverse expectations among the online shopping community, especially for college students who rely on parental allowances or part-time income. Promotions signify "getting a good deal" and reflect personal interests, leading consumers into the consumerism trap jointly constructed by businesses and platforms. The lack of norms in the live streaming e-commerce industry and disorderly competition inevitably lead to behaviors such as "excessive marketing," "selling counterfeit goods," "exaggerating product utility," and "false advertising." Under the influence of consumers' lack of judgment, mainstream modes, and blind, impulsive consumption concepts, impulsive buying behavior easily arises. Purchasing goods that do not match the promotion or are counterfeit on live shopping platforms can lead to post-purchase regret experiences for consumers and dissatisfaction with the live shopping method and platforms, reducing their willingness to buy again. The lack of regulation in live streaming e-commerce platforms, coupled with unhealthy consumer shopping habits, fosters the "consumerism trap" in live streaming e-commerce platforms.

Therefore, livestream shopping is both a trend and an inevitable product of the internet age, but the creation of a favorable livestream shopping environment relies on the efforts of both consumers and enterprises. For consumers, it is essential to discern their needs and actual circumstances during livestreams, actively engage in pre-purchase reflection, allocate more time for purchases, identify the compatibility between livestream shopping products and their own needs, reduce impulsive purchasing behaviors and post-purchase regret caused by the fundamental characteristics of livestream shopping, and approach their purchasing decisions rationally. Consumers should enhance their self-control ability and acquire reasonable consumption knowledge when engaging in live shopping. They should develop healthy and rational consumption concepts and behaviors, avoiding becoming obsessed with live shopping. Overcoming impulsive consumption and avoiding falling into the "consumerism trap" are essential. For enterprises, on the one hand, they can gather insights into consumers' purchasing decisions through various channels and then market their products in a reasonable manner, leveraging the advantages of livestream shopping platforms to promote products and expand distribution channels, rather than solely relying on impulsive purchasing to drive consumer purchases. On the other hand, in order to safeguard the legitimate rights and interests of consumers, it is hoped that live streaming platforms and live streaming enterprises can promptly regulate their own behavior, present product information more truthfully, and provide higher-quality goods. In addition, relevant authorities should increase supervision of online live streaming platforms, collectively creating a more healthy and genuine environment for online live streaming shopping through fostering fair competition.

Finally, the limitations of this study include, firstly, although a certain number of user comments were selected, only a portion of comments from the Weibo platform were obtained using web crawling technology. In reality, consumers express their opinions on various platforms, and while Weibo is representative, it may not cover all perspectives comprehensively. Secondly, this study did not further segment the user groups, thus failing to delve deeper into the discussion of customer segments. Lastly, this study only conducted data collection and analysis through word cloud analysis, without further exploration and discussion of the collected data. Future research could delve deeper into this aspect.

References

[1] Lu J R. "Consumerism in China: Manifestation, Hazards, and Management". *Journal of Hubei University of Economics*, 2005, (04):118-123.

[2] Azjen I. "The theory of planned behavior", *Organizational Behavior and Human Decision Processes. Journal of*

Lsure Research, 1991, 50(2): 176 -211.

[3] Han Y, Xue Y. "Development and Application of the Theory of Planned Behavior." Psychological Monthly, 2019, 14(07): 195-196.

[4] Xuan Y, and Sun J H. "From Rationality to Impulsivity: The Model of Influencing Factors of Live Shopping User Behavior and Its Development and Changes During the COVID-19 Pandemic". Journal of Communication University of China (Natural Sciences Edition), 2020, 27(06): 25-35.

[5] Li X, Wu L. Herding and social media word-of-mouth: evidence from Groupon. MIS Quarterly, 2018, 42(4): 1331-1351.

[6] Ye Q, Cheng Z, Fang B. Learning from other buyers: the effect of purchase history records in online marketplaces. Decision Support Systems, 2013, 56(1).

[7] Bikhchandani S, Hirshleifer D, Welch I. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 1992, 100(5).

[8] Zhang X Y. "Research on Herd Behavior of Online Shopping Users Based on Information Association." Jiangxi University of Finance and Economics, 2020.

[9] Yin M, Li Q, Liu Y. "Herd Effect in Online Shopping Festivals and Its Impact on Impulse Buying—Based on Social Influence Theory and Herd Effect Empiricism." China Business and Economics, 2019, 33(08):99-107.

[10] Wang L. "The Influence of Online Shopping Attitude on User Purchase Intention Based on Different Bullet Screen Contents." Hebei University, 2021.