ICCK

ARTICLE

Consumer Satisfaction Based on Online Reviews of Fresh Food E-commerce Platforms: Take YuanShengXian Store As an Example

Jiayu Pan [1,*]

[1] School of Digital Economic and Management, Software Engineering Institute of Guangzhou, Guangzhou 510000, China

Abstract

Online reviews are a vital tool for both businesses and customers to understand each other's requirements and to learn about the features and quality of products. This paper uses word frequency statistics and semantic network analysis to analyze the factors affecting consumer satisfaction into five categories: service, quality, logistics, packaging, and price. The research object is text reviews of various products from YuanShengXian's flagship store. The analysis yields a consumer sentiment evaluation for each influencing factor. The theoretical framework of the new e-commerce satisfaction assessment is further enhanced and improved by this study, which also serves as a guide for the enterprise's development plan and improvement strategy.

Keywords: Yuanshengxian Store, Fresh Food E-commerce, Online reviews, Consumer satisfaction.

1 Introduction

With the Internet's quick expansion, particularly during the COVID-19 epidemic, the stay-at-home lifestyle has further fueled the online consumption pattern, leading to an explosive growth in the number of orders on fresh produce e-commerce platforms. According to the "Research Report on the Operation Big Data and Development Prospects of China's Fresh E-commerce in 2022," the domestic transaction scale of the fresh e-commerce market reached 465.81 billion yuan in 2021, an increase of 27.92% over 2020. By 2022, the transaction scale of fresh e-commerce will be 560.14 billion yuan, a year-on-year increase of 20.25%. It is expected that the scale of China's fresh market will reach 642.49 billion yuan in 2023 [1].

Since 2005, fresh food e-commerce has been called the "last blue ocean" of e-commerce, attracting the attention of many capitalists and entrepreneurs. Although fresh food e-commerce is developing steadily, most enterprises have not yet reached profitability or cash flow peaks, and the business development is very unbalanced. In order to realize further development, fresh food e-commerce enterprises need to understand consumer evaluation of fresh food e-commerce in order to adjust their business strategies.

2 Literature Review

2.1 Online Review

Owing to the growing number of e-commerce platforms and the normalization of online purchasing, research on online reviews has become reasonably advanced among domestic researchers. The research topics of domestic scholars on online reviews mainly revolve around helping consumers make purchasing decisions based on user perspective and conducting

sentiment analysis on online review texts, such as Sun (2012) found through in-depth research that if online product reviews can change consumer purchasing decisions, they are considered useful reviews [2]. This laid the foundation for research on the usefulness of online reviews in various fields. According to the relevant search results for the keyword "online reviews" in the China National Knowledge Infrastructure (CNKI), only 29 related papers on online reviews of fresh e-commerce were published between 2019 and 2022. Most scholars believe that the sentiment of online reviews for fresh e-commerce is an intuitive feedback on consumer satisfaction and a reference for consumer decision-making. Therefore, research on reviews based on different models has emerged, such as Huang Lei (2017) using an emotional dictionary to visually analyze the text [3]; Guo et al. (2020) using an optimized LDA model to deeply explore the semantic relationships in the text [4]; additionally, Zheng and Wang (2017) used the features of a semantic dictionary to collect the emotional words of online reviews, assigning them emotional scores to determine the overall emotional tendency [5].

Research on online comments abroad preceded that in our country. The main themes were the establishment of machine learning models for sentiment analysis or text mining, and in-depth study of big data analysis methods for online comments. The early research subjects were mostly online comments from financial magazines or investment websites. In the 1990s, foreign scholars had already conducted sentiment analysis using different machine learning models. For example, Maharani et al.(2022) based their sentiment mining on sentiment dictionaries and used decision trees and rule learning models to achieve high accuracy and stability [6]. Hatzivassiloglou et al.(1997) proposed a four-step supervised learning algorithm, based on the semantic bias of adjectives through linguistic connectors, to infer the semantic orientation of other words associated with known words through connectors [7]. In addition, many foreign scholars used online comments to study the impact of corporate activities, such as Trusov et al.(2009), whose study showed that compared to other marketing methods, online comments are more persuasive and bring more loyal customers to businesses [8].

2.2 Consumer satisfaction

Domestic research on consumer satisfaction in China is divided into six major themes: analysis, model establishment, enhancement, application, evaluation,

and investigation. Since 2002, domestic scholars have carried out research on customer satisfaction through various methods, and have promoted the China Customer Satisfaction Index (CCSI) model [9]. Based on Oliver's expectancy-disconfirmation model and comprehensive customer satisfaction theory, Chen Fanjiao has established a model of fresh e-commerce consumer satisfaction applied in marketing, performance appraisal, and other fields. Li and Cheng (2013) applied the satisfaction model to vehicle scheduling problems [10]. The evaluation of consumer satisfaction models is beneficial for optimizing and improving their evaluation system. Wang and Zeng (2004) first proposed using the entropy method to study customer satisfaction evaluation [11].

In terms of relevant research abroad, Sweden was the first country to establish a national consumer satisfaction index (SCSB model), which not only helps businesses calculate future economic returns and make informed investment decisions, but also measure the quality of economic output. The United States made modifications to the SCSB model and established the ACSI model. The European Customer Satisfaction Index (ECSI model) drew inspiration from the American satisfaction model and was completed with support from European quality organizations, among others [12]. After the establishment of consumer satisfaction index models in various countries, compared to China, foreign countries have applied customer satisfaction to the fields of finance and economy. For example, research based on consumer satisfaction models has explored the relationship between stock market economic returns and consumer satisfaction, as well as factors influencing the growth rate of individual consumer expenditures. For instance, a study at the University of Michigan in the United States found a significant positive correlation between the value of the consumer satisfaction model and a company's financial asset return rate and long-term financial performance [13]. The Danish postal system used satisfaction models to measure the company's perceived satisfaction values for postal services and user communication, with the perceived values having the greatest impact on consumer loyalty in the network and soft drink industries, among others [14].

3 Methodology

The representative products' online reviews from the official flagship store "Yuan Shengxian Flagship Store"

Table 1. Statistical table of high-frequency words (top 20 data).

No.	word	frequency	No.	word	frequency
1	Delicious	2552	11	Live studio	259
2	Flavor	1195	12	Affordable	183
3	Texture	924	13	Quality	154
4	Freshness	879	14	Customer service	153
5	Packaging	602	15	Dispatch	147
6	Repurchase	576	16	Deserve	142
7	Logistics	376	17	Glutinous	135
8	Evaluation	342	18	Viya(Sales anchor)	127
9	Express delivery	285	19	Cheaply	126
10	size	274	20	First	126

of Freshippo, the number one brand in the fresh e-commerce industry in recent years, were selected for research. Using web crawling software, a large amount of online review data was obtained and processed using Excel for deduplication and deletion of short sentences. ROST CM6 software was used for satisfaction analysis of the effective data. The calculation logic of the satisfaction of fresh e-commerce consumers is as follows: first, the word frequency analysis is used to determine the influencing factors of consumer satisfaction, and then the weighted analysis is used to rank the importance of the influencing factors. Semantic network analysis and inductive summarization-feature word extraction are used to classify the influencing factors into main themes for consumer satisfaction. This forms the primary indicators of consumer satisfaction influencing factors. Finally, sentiment analysis can reveal the specific expectations of fresh e-commerce consumers for the influencing factors of satisfaction.

3.1 Data acquisition and pre-processing

The tool chosen for data acquisition is the web crawler tool - Octopus Collector. This software, using the Python language, simulates user access to the comment page, sets up looped paging and looped data extraction, and copies text comments to achieve rapid capture. By using the local collection function of the Octopus Collector, 10,508 pieces of Tmall product review content can be obtained. The collected data also includes user names, review images, follow-up review content, and other information.

After completing the data collection, the first step is to deduplicate the text. Text deduplication can be achieved using the "Remove Duplicates" function in Excel. Through filtering, a total of 291 duplicate entries were identified, leaving 10,217 valid data entries after removal of duplicates.

Next is the removal of short comments. Using the LEN function in Excel to calculate the character count of the comment data, and then applying the filtering function to delete comments with less than 10 characters. This resulted in a selection of 709 invalid data entries, leaving a total of 9,508 valid entries. Finally, comment segmentation is carried out using the ROST CM software for subsequent satisfaction analysis.

3.2 Analysis of factors influencing consumer satisfaction

3.2.1 Word frequency analysis

This article conducts a frequency analysis on 9508 valid processed data and presents the top 40 mentioned words in Table 1.

### 3.2.2 Weighting analysis of feature words

From the word frequency analysis, we can find some feature words that consumers care about, but it can't fully reflect the importance percentage of these features that consumers are more concerned about for fresh food e-commerce. Therefore, further feature word weighting analysis should also be carried out. Using the TF-IDF method to calculate the weight of online review feature words of six kinds of commodities, the final calculation results are shown in Table 2. In addition, considering that there may be deviations in the features that consumers are concerned about in each commodity category, the weight analysis of various types of feature words of the three major categories is conducted for comparison.

In order to more intuitively show the degree of consumer attention to each feature word, the TF-IDF data of each type of feature word of the three categories are made into a radar chart, which shows that the

Table 2. YuanShengXian store online review TF-IDF value table.

YuanShengXian store online review TF-IDF value			
Delicious	63800	Evaluation	8550
Flavor	29875	Express Delivery	7125
Texture	23100	Price	7075
Packaging	15050	Quality	3850
Freshness	12700	Customer service	3825
Logistics	9400	Affordable	3150

three major categories of fruits, vegetables, and meat and poultry consumers are most concerned about the features are the quality of the product, such as good taste, flavor, and freshness, followed by packaging and logistics; consumers of fruits and vegetables have a lower degree of attention to the price aspect of the goods, while meat and poultry goods are required for the price; service and after-sale service are the factors that are relatively unconcerned about by the consumers.
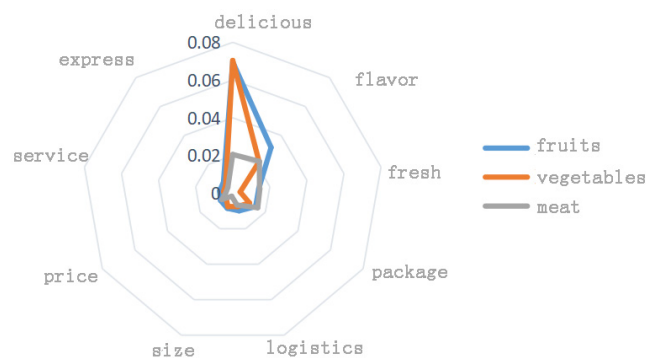
Figure 1. Radar charts for each feature word in the three categories.

Overall, the ranking of consumers' attention to various factors of Freshippo's fresh goods are delicious > flavor > package > fresh > logistics > price > service.

3.3 Consumer Satisfaction Thematic Categories

3.3.1 Semantic network analysis

From Figure 2, it can be seen that the three characteristic words "tasty," "mouthfeel," and "packaging" are key nodes, with the closest and most associations with other characteristic words.

Characteristic words related to "tasty" include "taste," "mouthfeel," "fresh," "repeated purchase," "sweet," "repurchase," "soft and glutinous," and other words, reflecting that consumers' satisfaction with the product comes from its sweet, soft and glutinous taste, good mouthfeel, and that they will make repeated purchases

of tasty products.
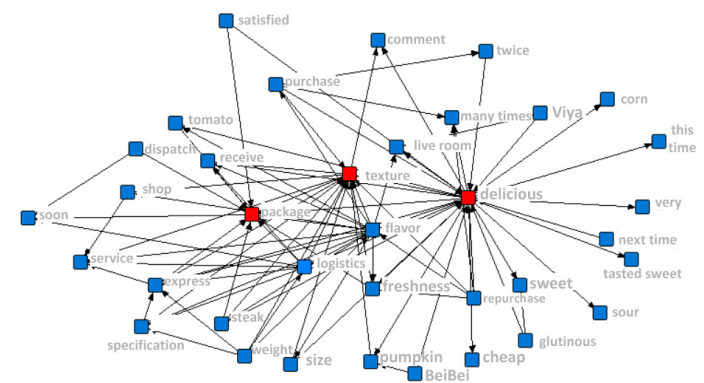
Figure 2. Semantic network structure diagram.

Characteristic words related to "mouthfeel" include "fresh," "tasty," "logistics," "service," "express delivery," showing that good mouthfeel requires the freshness of the product to be maintained, so it is necessary to consider the speed of the logistics, and when the product's logistics is fast, consumers naturally think that the product's service is good.

Characteristic words related to "packaging" include "logistics," "fresh," "taste," "express delivery," showing that the quality of packaging will affect the freshness of the product, and at the same time, in the process of logistics, the maximum integrity of the goods is ensured.

In order to measure consumer satisfaction more intuitively and clearly, this paper carries out feature word analysis on the tightness and relevance of the interconnection of feature words in the semantic network analysis, and constructs a feature word dictionary through the data, which can be finally analyzed from Table 3 to conclude that, in all the effective consumers' online review data, the influencing factors that the consumers pay attention to are related to the aspects of product quality, logistics, packaging, product service, and product price, and the ones that are closely related to these aspects are as follows: these features.

3.3.2 Sentiment analysis

In order to conduct an in-depth analysis of fresh food e-commerce consumers' satisfaction situation in the five dimensions of quality, packaging, logistics, service, and price, a study was launched using sentiment analysis, resulting in the results in Table 4:

The proportion of consumers' positive sentiment towards fresh food e-commerce is 69.72%; the proportion of neutrality is 18.77%; and the proportion

Table 3. Influencing factors of consumer satisfaction.

Factor	Feature word
Quality	Delicious, flavor, texture (of food), fresh, quality……
Packaging	Tight, wreck, aesthetically pleasing, design……
Logistics	Dispatch, fast, speed……
Serve	Customer service, attitude, bad, soon……
Price	Live room, affordable, cheap, repurchase……

Figure 3. Word Cloud of Negative Sentiments.

Figure 4. Word Cloud of positive Sentiments.

of negativity is 11.51%, which indicates that the satisfaction is generally good, but the negative sentiment also occupies a portion of the weight, and therefore should be emphasized.

Segmentation as well as word frequency analysis processing was performed again in the negative sentiment online review data, resulting in the results of Figure 3, which shows that consumers are dissatisfied with fresh food e-commerce in the following areas:

First, product quality. Most of the problems reflected in the bad taste, bad texture, or the product is not fresh or bad; Second, Packaging logistics. Consumers reflect the problem is that logistics and transportation problems lead to packaging damage or extrusion, and further damage to fresh products together. Third, logistics speed. Slow speed is the main factor of consumer negative emotions, and word frequency statistics on the platform to choose which logistics company is also a consumer concern. Fourth, the impact of the corresponding small for the evaluation of after-sales customer service, generally produces poor reviews because of the attitude of customer service slack, as well as the product of the after-sales protection is not perfect, such as commodity damage refund problem.

The results of positive emotions are shown in Figure 4. The influencing factors are primarily product quality, followed by product pricing and satisfaction with the activities. Some consumers generated positive emotions due to the discounts offered in the live broadcast rooms. As for packaging, the word cloud shows that exquisite and reliable packaging satisfies consumers, while quick delivery and convenience are the primary considerations for logistics. The term "repurchase" in comments often accompanies positive reviews.

4 Recommendations

4.1 Improve product quality

Based on data research, it can be deduced that taste, flavor, and freshness are the aspects of product quality that matter most to customers. Fresh e-commerce businesses must so tightly regulate the product selection process. Strict packing, insurance, and uniform fresh product processing ought to be done under dynamic, real-time monitoring. To accommodate changing consumer expectations and make purchases easier for customers, product categories may be horizontally enlarged and product richness can be raised simultaneously.

In addition to high quality and freshness, consumers also pay attention to local specialty products when it comes to fresh products. Just like the "Guangxi Oranges" in the YuanShengXian flagship store, having unique regional characteristics implies the product's taste and quality advantages [15].

4.2 Optimize packaging level

From the factors influencing consumer satisfaction in this article, it can be observed that the level of packaging is also an important topic for consumers. Keywords such as "secure" and "damage" are often

Table 4. Textual Sentiment Analysis.

Emotional segmentation	Analysis result		Average level		Medium level		High level	
positive	5028	69.72%	3109	43.11%	1339	18.57%	580	8.04%
neutral	1354	18.77%						
negative	830	11.51%	700	9.71%	109	1.51%	5	0.07%

associated with packaging, and in the semantic network, there are also connections to "taste" and "freshness." From this data, it can be seen that there is a strong correlation between packaging and logistics transportation, and it can affect the quality of the product. In order to prevent damage to fresh products, it is necessary to enhance the level of packaging.

Firstly, it is possible to improve the distribution equipment. In addition to strengthening, thickening, and tightly sealing the packaging, temperature regulation and protection can be set on the packaging, for example, cold fresh products need ice packs to maintain temperature. In addition, upgrading equipment such as delivery bags and foam boxes can make the packaging more refined. Secondly, differentiated packaging for products can be implemented. In terms of packaging, more attention should be paid to optimization in detail. For example, in terms of materials, biodegradable packaging can be used to reduce resource waste; in terms of design, company-specific and cultural elements can be added to create more desire for purchase.

4.3 Improve logistics efficiency and safety

The short shelf life and perishability of fresh products have posed challenges to the development of e-commerce sales. In emotional analysis of consumer satisfaction, consumers show high tolerance towards logistics. Some neutral comments mention issues during transportation, but these are resolved through after-sales service. However, consumers still have high expectations for logistics, especially regarding speed, which is a source of dissatisfaction.

To address speed issues, integrating multiple logistics methods and transportation modes can promote the integration of fresh e-commerce with traditional channels. The O2O model is rapidly developing in the e-commerce industry, and the combination of online and offline channels rapidly enhances consumer perceptions of product reliability. In the future, if fresh e-commerce expands and enters more fresh markets in various cities for regional sales, it will significantly reduce transportation workload and operational costs

in the logistics area.

The speed and stability of logistics transportation significantly affect the quality of fresh products, hence ensuring product integrity in logistics packaging and equipment is crucial for safety [16]. Strengthening cold chain logistics construction and perfecting it is essential. Communication with the government is necessary to seek policy support and to enhance infrastructure, such as expanding rural production roads, constructing high-quality agricultural parks, or obtaining greater investment in cold chain equipment.

4.4 Guarantee after-sales service

In sentiment analysis, consumers' expectations for after-sales service revolve around the attitude of customer service representatives and the adequacy of post-sales guarantees for products, including the availability of product return, exchange, and refund services. Companies can enhance after-sales service guarantees by refining consumer feedback mechanisms and improving after-sales customer service through training.

References

[1] Aimei Consulting. 2022 china fresh food e-commerce operation big data and development prospect research report. [CrossRef]

[2] Su, X. (2012). Research on factors influencing the usefulness of b2c online reviews–taking amazon.com as an example. *Hubei: Zhongnan University for Nationalities*.

[3] Lei, H. (2017). Extraction of product attributes and visual representation of their emotions based on product reviews of e-commerce websites. *Beijing University of Posts and Telecommunications*.

[4] Guoxianda, N., Gao, H., & Yang, X. (2020). Research on online review topic mining based on gaussian lda. *Journal of Intelligence*, 39(6), 630–639.

[5] Zheng, L., & Wang, H. (2017). Emotional polarity and intensity analysis of online reviews based on emotional ontology:taking cell phone as an example. *Journal of Management Engineering*, 31(2), 47–54.

[6] Maharani, W., Widyantoro, D. H., & Khodra, M. L. (2022). Dynamic aspect-based rating system

and visualization. *User Modeling and User-Adapted Interaction*, 32(1-2), 1–24. [CrossRef]

[7] Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 174–181). Madrid, Spain: Morgan Kaufmann Publishers / ACL. [CrossRef]

[8] Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of Marketing*, 73(5), 90–102. [CrossRef]

[9] Zhang, W. (2015). A review of domestic research on customer satisfaction-analysis based on china knowledge network literature from 2000-2014. *Research on Business Economy*(31), 66–68.

[10] Li, B., & Cheng, L. (2013). Study on vehicle scheduling problem based on customer satisfaction. *Logistics Technology*, 32(23), 253–254+272.

[11] Wang, Y., & Zeng, F. (2004). Customer satisfaction measurement based on entropy value method. *Business Research*(22), 11–13.

[12] Pei, F., Tang, W., & Kuitong, H. (2006). A review of customer satisfaction research and application. *World Standardization and Quality Management*(10), 4–7.

[13] Xu, Z., & Wei, B. (2001). Empirical analysis of customer satisfaction index in the united states. *World Standardization and Quality Management*(10), 16–17.

[14] Hill, N., & Alexander, J. (2017). *The handbook of customer satisfaction and loyalty measurement*. Routledge.

[15] Wang, X., Wang, N., Song, W., Wang, M., & Zhang, M. (2022). Research on factors influencing consumer satisfaction and countermeasures in fresh food e-commerce. *Investment and Entrepreneurship*, 33(18), 124–126.

[16] Liu, G. (2022). The status quo and trend of fresh food e-commerce development in the context of new retail. *China Market*(18), 169–171.

Jiayu Pan is graduated from the Department of Electronic Commerce, Software Engineering Institute of Guangzhou. Her graduation thesis, which focused on data analysis of online reviews for fresh e-commerce, was awarded the first prize for outstanding graduation thesis at the school.