horticulturae

Article

Study of Zhejiang Tangerine E-Commerce Reviews Based on Natural Language Processing

Leiming Yuan [1,*], Haoyang Liu [1], Fangfang Fu [2], Yimin Liu [1], Xiaoyu Zuo [1] and Limin Li [1,*]

1 College of Electrical & Electronic Engineering, Wenzhou University, Wenzhou 325035, China
2 Wenzhou Evening News, Wenzhou 352000, China
* Correspondence: yuan@wzu.edu.cn (L.Y.); lilimin@wzu.edu.cn (L.L.)

Abstract: In recent years, the global economy has experienced significant shifts, leading to a trend of consumption downgrading. Amid economic pressures and uncertainties, consumers are increasingly turning to cost-effective shopping methods. The COVID-19 lockdowns further accelerated the growth of e-commerce platforms, presenting both opportunities and challenges for sales. Electronic commerce has played a crucial role in enhancing the sales of agricultural products with regional characteristics in China, thereby opening new channels for farmers. This article utilizes tangerines, particularly popular in Zhejiang Province, as a case study to explore e-commerce reviews and assist merchants in delivering more satisfactory products. The analysis of tangerine reviews revealed that customers primarily focused on the taste, service, quality, and price. By applying the latent Dirichlet allocation (LDA) topic model, comments were categorized into four themes: 'quality', 'service', 'price', and 'flavor', with key terms identified for each theme. Through sentiment analysis using SnowNLP and bidirectional encoder representations from transformers (BERT), it was found that online shoppers generally expressed positive sentiment toward tangerines. However, there was also some negative feedback. These findings are of paramount importance for businesses aiming to meet consumer demands. The study acknowledges certain limitations including the reliability of data mining and the accuracy of Chinese corpus analysis. Future research could benefit from employing more precise language models to enhance the analysis, ultimately improving the consumer shopping experience and aiding businesses in service improvement.

Keywords: tangerine; online comments; text mining; sentiment analysis; latent Dirichlet allocation (LDA)

check for updates

Academic Editors: Christian Fischer and Francesco Sottile

Received: 23 December 2024
Revised: 14 January 2025
Accepted: 24 January 2025
Published: 1 February 2025

Citation: Yuan, L.; Liu, H.; Fu, F.; Liu, Y.; Zuo, X.; Li, L. Study of Zhejiang Tangerine E-Commerce Reviews Based on Natural Language Processing. *Horticulturae* **2025**, *11*, 151. https://doi.org/10.3390/horticulturae11020151

1. Introduction

In recent years, the global economic landscape has experienced significant transformations, resulting in a discernible decline in the quality of consumer goods purchased. Confronted with economic pressures and uncertainty, consumers have increasingly gravitated toward more cost-effective shopping methods. Additionally, measures implemented to curb the spread of the COVID-19 pandemic have catalyzed the rapid growth of e-commerce platforms. The propensity for online shopping has become increasingly evident, as illustrated by the 2023 Statistical Report on the Development of the Internet in China. The report indicates that by 2023, China's Internet user base had surged to 1.079 billion, with a penetration rate of 76.4%. Moreover, the number of individuals engaging in online shopping had risen to 840 million.

Due to the convenience of shopping online and the availability of more affordable goods, an increasing number of merchants have joined e-commerce platforms. To expand their market reach, many agricultural products including those with regional characteristics have been transferred online [1]. Among these, mandarin tangerines, a popular fruit produced in Taizhou City, Zhejiang Province, China, are marketed globally through e-commerce annually. Since customers cannot physically inspect the items before purchasing, they largely depend on e-commerce review sections to guide their decision-making [2,3]. Consequently, it is essential to develop a rating system for mandarin tangerines by meticulously scrutinizing and evaluating the comments on e-commerce platforms. Such a system would assist merchants in delivering products that align with customer expectations [4]. Therefore, to enhance the provision of satisfactory products, we aimed to establish a comprehensive review system for tangerines through the analysis of e-commerce reviews. Today's e-commerce platforms facilitate interaction and transactions between consumers and merchants, fostering a close connection between products and online commerce. These platforms enable merchants to display and sell a diverse array of products, while consumers can conveniently browse and make purchases through the platform. By refining our understanding of consumer feedback, we can assist businesses in better meeting the evolving demands of their clientele.

The relationship between products and e-commerce is multifaceted and can be explained from various perspectives. Firstly, sales channels: e-commerce platforms have emerged as crucial sales channels, offering a solution to the geographical and temporal constraints faced by traditional brick-and-mortar stores [5]. They enable products to reach and be accessible to a broader consumer base. Next, there is product display and promotion: these platforms provide a virtual space for merchants to exhibit and promote their offerings. By employing images, detailed descriptions, and videos, merchants can effectively highlight the unique features and benefits of their products, thereby capturing consumer interest and stimulating purchases [6]. Thirdly, transactions and payments: e-commerce platforms facilitate secure and user-friendly transaction processes. They simplify the act of purchasing for consumers, who can select products, choose payment options, and receive prompt and secure delivery services. Finally, reviews and feedback: products on e-commerce platforms are often accompanied by consumer reviews and feedback. After making a purchase, consumers can review the products, sharing their usage experiences and opinions [7]. These reviews and feedback serve as valuable references for other consumers, aiding them in making informed purchasing decisions [8,9].User online reviews are a visual representation of the customer's voice, and through the study of user online reviews, can better assist merchants in improving customer satisfaction. Currently, some researchers have studied and analyzed the user online ratings, using digital forms to intuitively and clearly display the degree of customer satisfaction with the store and the product [10,11]. However, ratings alone cannot fully respond to the store's strengths and weaknesses. Online reviews, which are open-ended feedback written by users after using the goods, can provide a comprehensive response to their ideas for the merchant and other users to consider. Therefore, this paper will take the online reviews of Honey Tangerine E-commerce as the source of information, carrying out data analysis and processing.

In the domain of text mining for attitude analysis, several works have laid the groundwork for understanding consumer behavior and satisfaction. Guo et al. [12] utilized LDA thematic modeling to delve into customer satisfaction with merchants, uncovering a number of key factors that were not explicitly captured by traditional questionnaire. Xu et al. [13] focused on mining textual information to establish a correlation between the type of hotels and the purposes of the customers. They employed regression analysis to explore various issues, revealing significant insights into how different hotel attributes catered to the diverse needs of their clientele. This study underscored the value of textual data in predicting consumer preferences and tailoring services accordingly. Srinivas and

Rajendran [14] delved into the realm of educational services by obtaining influence factors of student satisfaction with their schools through data mining [15]. They further employed sentiment analysis to comprehend the profound impact of these factors on student satisfaction. This sequential use of data mining and sentiment analysis provided a comprehensive view of the dynamics at play, emphasizing the importance of considering both quantitative and qualitative aspects of consumer feedback. These studies collectively demonstrate the potency of analyzing online product review data in unearthing strategies to enhance merchandising and bolster the competitiveness of businesses. By tapping into the wealth of information contained within customer reviews, companies can gain a competitive edge by refining their offerings and services to meet the evolving demands of the market [16].

In this paper, we focused on two mainstream e-commerce platforms in China as the source of our research dataset, conducting an in-depth analysis of online review information for mandarin tangerine products listed on these platforms. By leveraging advanced text sentiment analysis techniques, we intended to mine these reviews and identify the primary factors that influenced the customer purchasing decisions. Our objective was to furnish merchants with actionable insights to enhance their products and services, thereby elevating the overall user satisfaction. Specifically, through methodical data collection and analysis, we aimed to uncover the elements that consumers prioritized when purchasing honey tangerine products including but not limited to product quality, taste, packaging, logistics efficiency, and customer service. Utilizing natural language processing technology, we conducted sentiment analysis on the review texts to discern positive and negative sentiment expressions, gaining an understanding of the consumers' genuine feedback and potential needs. Meanwhile, this research was not only dedicated to assist merchants in providing better services, but also to enrich the shopping experience of customers. By considering the reviews of previous buyers and comprehending their insights, new buyers can minimize trial-and-error costs, swiftly identify quality products that meet their requirements, and avoid the disappointment associated with purchasing unsatisfactory products. We are confident that both merchants and customers can benefit from this systematic analysis, optimizing the shopping process to increase sales and customer loyalty for merchants, while customers enjoy a more gratifying shopping experience [17,18].

In summary, the goal of this paper was to analyze online reviews of citrus products on major e-commerce platforms and establish the scientific evaluation topic. This will help businesses provide services that customers need the most, thereby enhancing consumer satisfaction and shopping experience. It hopes to provide actionable guidance for e-commerce merchants specializing in mandarin tangerines and serve as a reference for analyzing the e-commerce reviews of other similar products.

2. Research Content and Methodology

This study primarily employed Python-based web scraping programs (Version 3.9) to collect online review data from the e-commerce platform in China, aiming to capture a diverse range of consumer feedback. The web scraping process was carefully designed to target product reviews related to tangerines, using custom-built scripts that interacted with the platform's APIs and HTML structures to extract data such as product names, ratings, review texts, and timestamps. These raw data were then subjected to a thorough preprocessing pipeline to clean and organize them including tasks such as removing irrelevant or duplicate entries, handling missing values, and standardizing the text by converting it into a uniform format (e.g., lowercase conversion, removing special characters, and stop-words) [19,20]. Once the data had been cleaned, latent Dirichlet allocation (LDA) topic modeling was applied to extract the underlying topics from the reviews. LDA is particularly suited for this task as it assumes that each document (review) is a mixture of topics, with each topic characterized by a set of keywords, thus identifying patterns in large datasets. By fitting the LDA model to the preprocessed data, this study was able to extract a set of topics and the associated keywords that represented the main themes in the reviews. To gain deeper insights into the sentiment expressed by consumers within these topics, natural language processing (NLP) techniques were then employed [21]. Specifically, sentiment analysis was performed on the topic keywords and review texts to assess the overall sentiment (positive, negative, or neutral) expressed by consumers in relation to each identified topic. Finally, the main results of the analysis, which include the identification of key topics and their corresponding sentiment trends, were derived from this process. The main workflow is illustrated in Figure 1, providing a clear visual representation of the steps involved, from data collection to sentiment analysis, ensuring transparency and reproducibility of the research methodology.

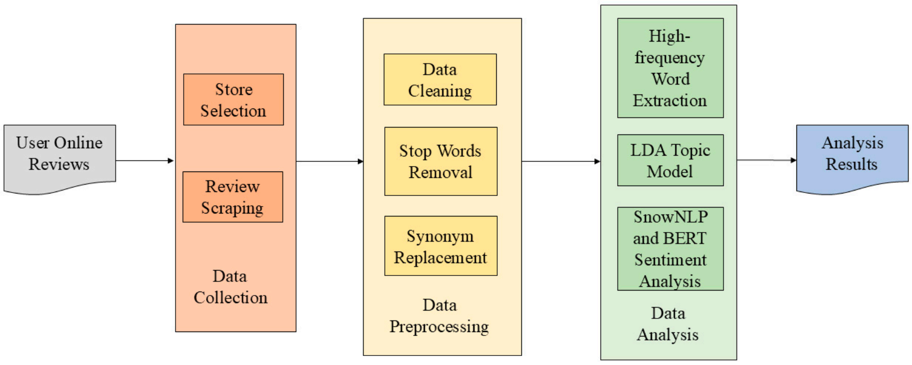


Figure 1. Process of analyzing the tangerine reviews.

2.1. Data Sources and Scraping

Considering the seasonal nature of tangerines, which are typically in peak availability during the fall season, this study scheduled the commencement of data scraping to occur after the period in which consumers have had the opportunity to make their post-purchase reviews. Consequently, data collection was initiated around December to ensure that the reviews reflected the seasonal peak of tangerine availability and consumer feedback. This research zeroed in on tangerine reviews from two of China's leading e-commerce platforms, JD.com and Taobao.com. The decision to select only Taobao and JD.com as the two e-commerce platforms was made after multiple experiments and discussions. In the context of everyday online shopping, most consumers tend to use Taobao, JD.com, and TEMU, as these three platforms dominate the majority of online shopping traffic. However, TEMU primarily targets mobile users, making data scraping more challenging. The remaining platforms either have anti-scraping measures in place or offer insufficient data. Therefore, we ultimately chose Taobao and JD.com for data collection. To ensure a focused and representative sample, the reviews were sorted based on sales volume, and the study selected approximately the first 1000 reviews from the top ten stores with the highest sales for scraping. This approach aimed to capture a broad spectrum of consumer opinions while focusing on the most popular and relevant products within the market.

The utilization of Python for web scraping allowed for the efficient collection of a vast dataset from e-commerce platforms, which served as the foundation for our analysis. By pressing F12 on a keyboard to enter developer mode on the review page, the request URL was obtained along with the contents of the cookies, which were decoded to extract the review information, and then saved into a text file.

After obtaining the desired data, we needed to perform some filtering operations to remove fake and irrelevant comments in order to facilitate the subsequent research. Below are the three methods that were used [22]:

Length Filtering

Comments that are too short may lack meaningful content. For example, comments consisting of just one or two words (e.g., "good", "bad") often do not provide sufficient details. A minimum length threshold can be set, such as requiring comments to be at least 10 characters long, to filter out superficial or careless remarks.

Comments that are too long may also be problematic. Some merchants may intentionally generate fake reviews by filling them with irrelevant content such as repetitive sentences or long, unrelated stories. A reasonable maximum length, such as 500 words, can be established, and comments exceeding this length can be subjected to manual inspection or further analysis to determine whether they are suspicious.

Keyword Filtering

Identifying and filtering out comments containing advertisement-related keywords. For instance, words like "professional review booster" or "paid review", which are clearly associated with fake reviews, can be flagged. Additionally, marketing terms unrelated to the product, such as "follow the shop for discounts" or "other great products in our store", can also be used as filters. Comments filled with excessive emojis without substantial content should also be considered for filtering. For example, comments like "😊😊😊" or "👍👍👍" signs repeated multiple times may be intended to inflate review numbers rather than provide genuine feedback.

Text Similarity Filtering

Calculating the similarity between comments. A large number of nearly identical comments could indicate fake reviews. Text similarity algorithms, such as cosine similarity, can be used. If the cosine similarity between two comments exceeds a certain threshold (e.g., 0.9), these comments should be further investigated to determine whether they come from real users. For instance, if multiple comments on the same product contain nearly identical content, such as "This product is of great quality, I'm very satisfied, it looks beautiful and is worth buying", they may be flagged as suspicious due to their high similarity.

2.2. Text Data Preprocessing

Ensuring the integrity and accuracy of online review data is essential for the reliability of subsequent analyses. The preprocess details a systematic approach to data cleaning and preprocessing to enhance the validity of textual data analysis in our work including text cleaning, tokenization, pause word removal, synonym replacement, text normalization, outlier detection, sentiment analysis, data annotation, and integration. Through these steps, the online review data were refined to enhance cleanliness and organization, establishing a robust foundation for advanced data analysis and knowledge discovery.

In e-commerce review analysis, age differences among consumers significantly impact their evaluation styles. Younger generations (such as Gen Z and Millennials) tend to provide concise and direct reviews. Influenced by cognitive biases (such as overconfidence and the Dunning–Kruger effect), they may overestimate their judgment abilities and simplify complex product evaluations. In contrast, older consumers (such as Gen X and Baby Boomers) tend to give more detailed and comprehensive reviews, covering aspects like user experience and product features. This is related to their higher sensitivity to risk and accumulated experience [23].

When analyzing reviews from these different age groups, it is important to consider the impact of cognitive biases. The brief evaluations from younger consumers may not fully reflect the true value of a product, while older consumers' detailed reviews may sometimes rely too heavily on traditional experiences, which could affect their acceptance of new products. To ensure an accurate analysis, it is crucial to integrate the generational cognitive characteristics and evaluate the reviews comprehensively.

This study used the TF-IDF weighting method to balance the impact of long and short reviews in the LDA model. TF-IDF weighs the words in the reviews based on their frequency and importance in the corpus, ensuring that longer reviews, which contain more information, do not lose their contribution due to common word frequencies. Longer reviews typically include more information and diverse topics, and TF-IDF effectively highlights their thematic contributions in the model, while short reviews, with less information, have lower weight. This approach helps the LDA model better uncover underlying topics and subtle emotional differences.

2.3. Natural Language Process

Natural language processing (NLP) is a field that lies at the intersection of linguistics, computer science, and artificial intelligence, dedicated to facilitating interaction between human language and computer systems. It involves the development of algorithms and techniques that allow computers to analyze, understand, and generate human language in a way that is both meaningful and practical. NLP technologies aim to automate the understanding of natural language, enabling computers to extract valuable information from text and interact effectively with humans. Techniques such as syntactic analysis, semantic interpretation, and pragmatic understanding are employed to tackle the intricacies of language, which is often nuanced and context-dependent [24].

Applications of NLP are diverse and impactful across various sectors. In information retrieval, NLP powers advanced search engines capable of interpreting natural language queries and retrieving relevant information from vast textual databases [25,26]. This capability enhances the user experience by providing accurate and contextually appropriate search results. Machine translation is another significant application where NLP shines, enabling the automatic translation of text from one language to another. This facilitates cross-language communication and understanding, crucial in globalized environments where language diversity is prevalent. NLP also plays a crucial role in text classification, categorizing documents based on their content, and sentiment analysis, determining the emotional tone of text, which is valuable in areas such as social media monitoring and customer feedback analysis. Question answering systems leverage NLP to process and comprehend user queries, retrieving precise answers from large datasets. Automatic summarization techniques condense extensive documents into concise summaries, aiding in information retrieval and knowledge management. Named entity recognition (NER) is pivotal in identifying and categorizing entities such as the names of people, organizations, and locations within the text, facilitating tasks like information extraction and database population. As NLP continues to advance, its applications will likely expand, offering even more ways to process and understand human language, and thereby enhancing the capabilities of various technologies and systems that rely on natural language interaction.

This study employed two tools from NLP technology, SnowNLP and bidirectional encoder representations from transformers (BERT), for data analysis. SnowNLP is a lightweight natural language processing library specifically designed for Chinese text, offering features such as word segmentation, sentiment analysis, and keyword extraction, making it particularly suitable for the quick analysis of short texts. BERT is a pre-trained language model based on the transformer architecture, capable of understanding the deep meanings of text through bidirectional context, and is well-suited for handling complex Chinese corpus tasks such as text classification and sentiment analysis [21].

2.4. LDA Subject Modeling

Latent Dirichlet allocation (LDA) is a commonly used algorithm for topic modeling, designed to uncover the latent thematic structures within textual data. As an unsupervised learning method, LDA is adept at identifying underlying themes without the need for pre-labeled data [27]. It operates on a three-tiered Bayesian model that emulates the process of document generation, assuming that each document is a mixture of various topics, and each topic follows a specific distribution of words within a vocabulary. During the training stage, LDA takes a corpus of documents and a predefined number of topics to identify the latent topics and their respective distributions. This process does not necessitate manually annotated datasets, allowing for the discovery of topics in an unbiased manner. The obtained topic distribution and vocabulary distribution can be used as a linear classifier or processed further. LDA is a very important topic modeling method used in NLP fields such as text classification. Several applications can be implemented using LDA models such as text topic modeling, text classification, and information retrieval [28,29]. With LDA modeling, the topic structure in a document set can be discovered to understand the correlation between documents and the distribution of topics. This is important for text mining, information organization, and text understanding.

When applying the LDA model, both model training and inference are required. Model training aims to discern the distribution of topics across the document set and the distribution of words within each topic. Inference, on the other hand, seeks to determine the topic distribution for new documents. To achieve these objectives, methods such as variational inference or Gibbs sampling are commonly employed. LDA topic modeling is a powerful tool to help us understand the topic structure and associative relationships in text data. It has a wide range of applications in the fields of text mining, information retrieval, and text understanding. Further research and improvement can further enhance the performance and applicability of the LDA model.

Firstly, assume that all online reviews from the tangerine e-commerce platform constituted a dataset D, where each review is composed of N words, represented as $w = (w_1, w_2, \ldots, w_n)$. Utilizing the LDA model, we hypothesized the presence of K topics within the entire dataset. The generation of a single review commences with the creation of a topic distribution specific to that review. Following this, a set of words is generated, corresponding to the identified topics according to their distributions. For each word generation, a topic is randomly selected based on the review's topic distribution, and then a word is chosen at random based on the selected topic's word distribution. This process is iterated until all of the reviews are fully composed [30].

The probabilistic framework of the topic extraction model is illustrated in Figure 2. Initially, N words are extracted, which obeys the distribution of $N \sim Poinsson(\varepsilon)$. Subsequently, the topic distribution θ across the online reviews is formatted from $Dirichlet(\alpha)$. For each of N words w_n, its corresponding topic Z_n is first identified, where $Z_n \sim Multinomial(\theta)$. Based on $p(w_n|Z_n, \beta)$, the corresponding word Wn is generated. Here, α serves as the *Dirichlet* prior parameter for the multinomial distribution of topics under any single comment, β serves as the Dirichlet prior parameter for the multinomial distribution of characteristic words under this topic, and θ represents the overall topic distribution across the comments. K denotes the total number of topics; D denotes the total number of comments; N_d denotes the total number of words in the d-th comment. $W_{d,n}$ represents the n-th word in the d-th comment; θ_d and φ_k respectively denote the topic distribution in the $d-th$ comment and the characteristic word distribution under the $k-th$ topic; $Z_{d,n}$ denotes the topic of the $n-th$ word in the $d-th$ comment [31].

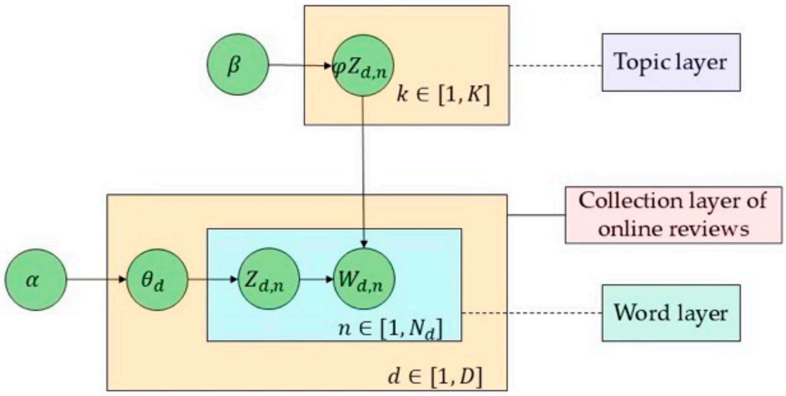


Figure 2. LDA topic model probability graph.

3. Analysis of E-Commerce Mandarin Reviews

3.1. Word Frequency Extraction

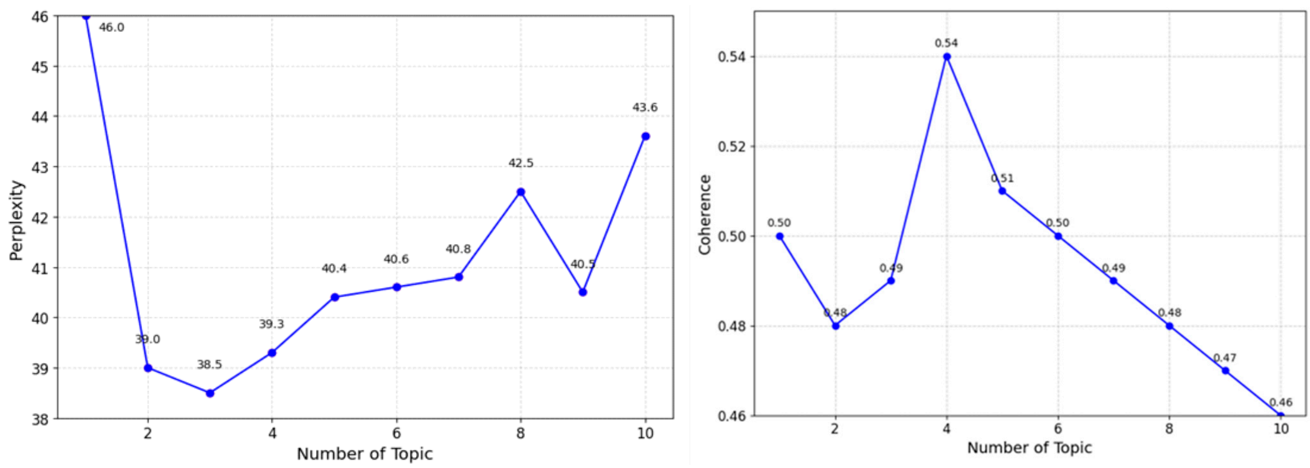
Initially, over twenty thousand comments were collected using a Python crawler program. This initial dataset included substantial irrelevant information, such as repetitive phrases and filler words, necessitating a textual data-cleaning process. The raw text data was first imported into an Excel spreadsheet for preliminary processing, where duplicate comments were identified and eliminated. Following this, the cleaned dataset was imported back into Python for further refinement using the *jieba* dictionary toolbox for data cleaning, yielding 18,896 valid comments. To visualize the distribution of high-frequency words, a word cloud was generated using a word cloud map production tool. In the word cloud representation, the size of a word correlates with its frequency of occurrence, where the larger the word, the more frequently it appears. As depicted in Figure 3, there were several prominent high-frequency words. Some words were used to describe the taste such as "delicious", "taste", "sour", "sweet", etc. Other words described the honey tangerines themselves including "even", "juicy", "fresh", etc. These high-frequency words reflect different perspectives from which the consumers evaluated the purchased products. Taking a closer examination of the low-frequency words, some words were generally meaningless in the context of shopping experience, fruit quality, or online shopping service, and may have been arbitrarily pasted or typed by shoppers. It is apparent that customer evaluations of tangerines predominantly focused on flavor. These meaningless low-frequency words, for which no corresponding English translations can be found, should be excluded from the analysis. The process of data cleaning and word cloud generation not only refined the dataset for more accurate analysis, but also provided a visual representation of consumer sentiments and priorities. This approach allows for a more focused examination of the factors that influence customer satisfaction and can provide strategies for quality improvement and service enhancement in e-commerce. In addition, the Chinese word cloud is provided in Supplementary Materials Figure S1, and the Chinese–English glossary compiled for the high-frequency words is shown in Supplementary Materials Table S1.

3.2. LDA Topic Model

In this section, we delve into the exploration of relationships among the corresponding words, utilizing a large language model to indicate potential connections. The LDA topic model was employed to uncover the hidden relationships within the textual data. This method leverages the LDA module within the Sklearn package to reveal the underlying patterns and associations that may not be immediately apparent. It focused on retaining positive and negative reviews within the text data, which corresponded to positive and negative sentiments expressed by consumers. This binary classification allowed for a deeper analysis of customer feedback, which can be crucial for understanding the dynamics of customer satisfaction and dissatisfaction, focusing on retaining positive and negative reviews within the text data that correspond to the positive and negative sentiments expressed by consumers.

Figure 3. Word Cloud for the high-frequency comments in the English version.

Perplexity is a commonly used evaluation metric in LDA topic modeling, and is primarily used to assess how well the model fits the data. The concept of perplexity originates from information theory in probabilistic models, reflecting the model's 'uncertainty' or 'level of confusion' regarding the test data. Lower perplexity values indicate better clustering performance, suggesting that the model is more successful in capturing the underlying structure of the data. This is visually represented in Figure 4A, which provides a graphical representation of the perplexity scores for different numbers of topics, allowing us to identify the point at which the model achieves the most optimal balance between complexity and coherence [32,33].



(**A**) Topic perplexity (**B**) Topic coherence

Figure 4. Result of top by latent Dirichlet allocation (LDA).

Coherence is another commonly used evaluation metric in LDA topic modeling, primarily used to assess the coherence or interpretability of topics. Unlike perplexity, which measures how well the model fits the data, coherence focuses on the semantic relatedness of the words within a topic. Higher coherence values indicate that the words in a topic are semantically related and easier to interpret and understand. Figure 4B illustrates the coherence scores for different numbers of topics, allowing us to identify the point at which the model achieves the optimal balance between topic coherence and complexity. Compared with the perplexity graph (Figure 4A), coherence emphasizes the semantic quality of the model.

From Figure 4, it was observed that the LDA topic perplexity value peaked when the number of topics was set to 1. As the number of topics increased, the perplexity value dropped significantly, indicating improved model performance. The perplexity reached its nadir when the number of topics was set to 3, suggesting that this was the point at which the LDA model provided the most coherent and distinct clustering of the data. Therefore, three topics were initially considered as a potential optimal choice. However, when combining this with the coherence values observed in Figure 4B, which showed greater stability and coherence as the number of topics increased, we decided to adjust the number of topics to four. This decision was made after balancing the perplexity results with the consistency considerations. The selection of the optimal number of topics is crucial for the effectiveness of the LDA model, as it balances the trade-off between model simplicity and the ability to capture the complexity of the data. With four topics, the model is able to group the textual data into distinct thematic clusters that are both meaningful and relevant to the analysis of consumer reviews.

The keywords within these topics, as presented in Table 1, serve as representative markers for each cluster. These words are indicative of the central themes or topics that emerge from the data and can be used to label and interpret each topic. For instance, topics might be centered around specific aspects of the product experience such as taste, quality, packaging, price, or customer service. The keywords help to distill the essence of each topic, providing a clear and concise way to understand and communicate the key points of consumer feedback. By focusing on these four topics, the analysis can provide targeted insights into the specific areas that are highly relevant to consumers. This information can be invaluable for merchants looking to refine their offerings and enhance the customer experience. It also lays the groundwork for further qualitative analysis or more in-depth exploration of consumer sentiment and behavior.

Table 1. LDA topic high-frequency word list.

Topic Categories	Topic Strength	Topic Keywords
Quality	0.29	Fresh, Even, Plump, Juiciness, Thin-skinned, Rotten, Size
Service	0.15	Logistics, Friendly, Packaging, Fast, Self-operated, attitude, Speed
Taste	0.40	Sweet, Delicious, Sour, Juicy, Bitter, Not-tasty, Tasteless
Price	0.16	Cheap, Worthwhile, Cost-effective, Cost performance

Based on the insights gleaned from the LDA topic model, it can provide a more detailed analysis and interpretation of each topic as shown in Figure 5. In this figure, it can be seen that there are four centers corresponding to each topic, where each contains numerous high-frequency keywords. Some key works are unique to a single topic, and some are shared across other topics, indicating the interrelatedness of different consumer concerns. Topic 1 was characterized by words that are intimately associated with the quality of tangerines such as 'fresh', 'even', 'thin-skinned', 'moisture', and 'sweet'. These words reflect the attributes of tangerines that consumers prioritize, particularly regarding their appearance and intrinsic quality. The word 'fresh' indicates that consumers place a high premium on the freshness and recently harvested state of the fruit, while 'even' suggests that they expect consistency in size, color, and other characteristics. Descriptors like 'thin-skinned' and 'moisture' underscore the importance of texture and juiciness, which are pivotal in the fruit selection process. Consequently, Topic 1 is aptly categorized as 'Quality', signifying that consumers are highly concerned with the physical attributes and quality standards of the product.

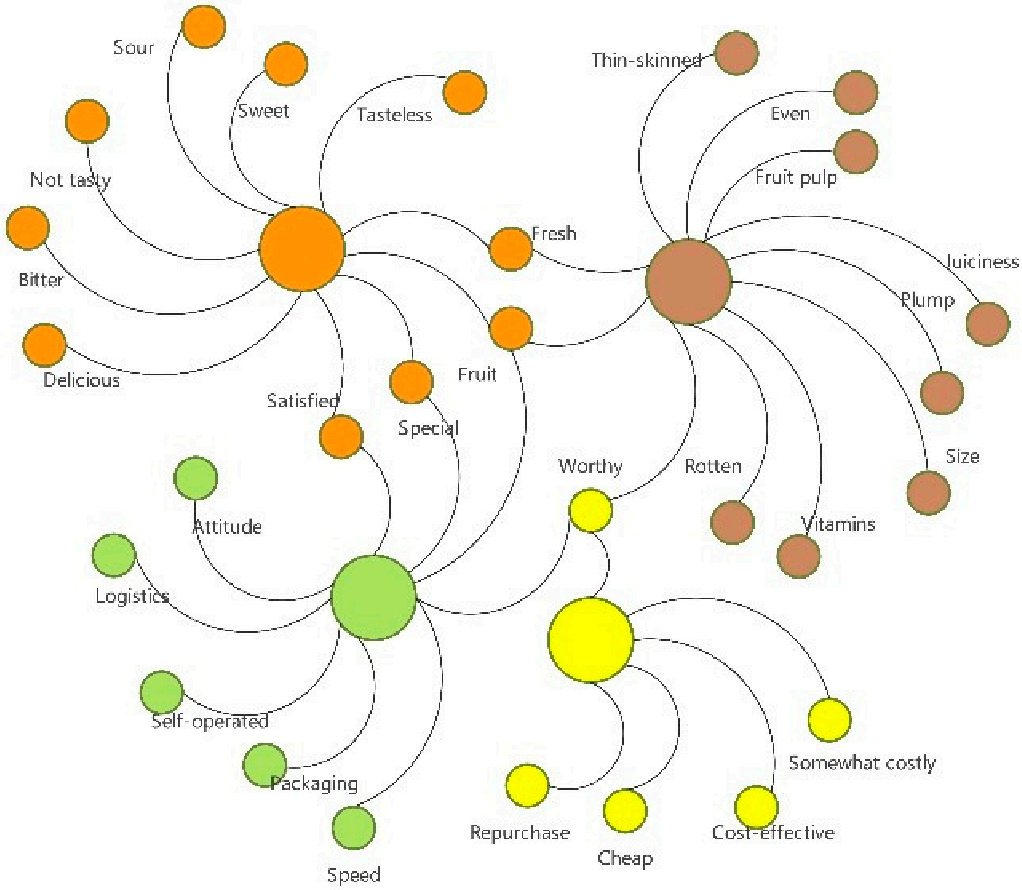


Figure 5. Topic relationship graph of LDA.

Topic 2 includes the words primarily related to the quality of services rendered by e-commerce platforms such as 'logistics', 'customer service', 'packaging', and 'speed'. This shows that consumers are not only focused on the product itself, but also place significant importance on their overall shopping experience. Elements such as delivery expedience, customer service responsiveness, and suitable packaging are integral to shaping customer satisfaction. Notably, terms like 'packaging' and 'speed' have a direct bearing on the consumer's holistic experience upon receiving the product. Thus, Topic 2 was designated as 'Service,' underscoring the significant impact of e-commerce service performance, both pre- and post-sale, on the customer reviews. Topic 3 concentrated primarily on the flavor and taste of tangerines, mirroring the consumers' emphasis on their sensory experience. This topic likely features terms such as 'sweetness', 'acidity', and 'fragrance', which articulate the consumers' specific expectations for the flavor profile of tangerines. Therefore, Topic 3 was labeled as 'Taste', highlighting the importance that consumers ascribe to their tasting experience when consuming the fruit.

By observing the keywords of the final topic, such as "price", "cheap", "worthy", and "somewhat costly", it is evident that the last topic was related to price, reflecting the consumer responses regarding pricing. Therefore, Topic 4 was labeled as "Price".

Overall, these four topics—Quality, Service, Price, and Taste—together formed the main concerns that consumers focused on when reviewing tangerines, reflecting their comprehensive expectations during both the purchasing and consumption journey. The extraction of these topics not only helps businesses understand consumer preferences and expectations, but also provides clear directions for enhancing product quality and service standards.

3.3. SnowNLP Sentiment Analysis

In this section, the SnowNLP sentiment analyzer was utilized in Python to calculate the sentiment scores for the extracted comments [24]. This tool utilizes an ontology library of Chinese emotional vocabulary, a valuable resource compiled and annotated by the teaching and research team under the guidance of Professor Lin Hongfei at the Information Retrieval Laboratory, Dalian University of Technology. This ontology provides a multi-faceted description of Chinese words or phrases, encompassing parts of speech, emotional category, emotional intensity, and polarity. Detailed emotion Chinese–English words can be found in Supplementary Materials Table S2.

With this comprehensive emotion lexicon, a sentiment analysis was conducted on the cleaned comments. Comments were scored based on the sentimental intensity of the words they contained. A score above 10 was considered to reflect a 'positive' attitude, while a score below 0 was categorized as 'negative'. Scores falling between these thresholds were classified as 'neutral'. The pie plot in Figure 6 visually represents the distribution of sentiment across the comments.

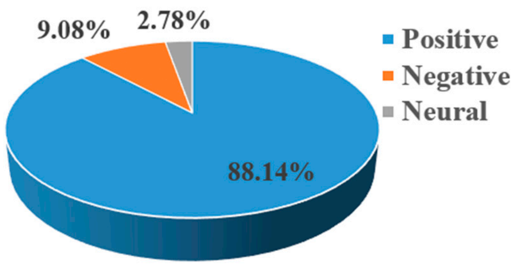
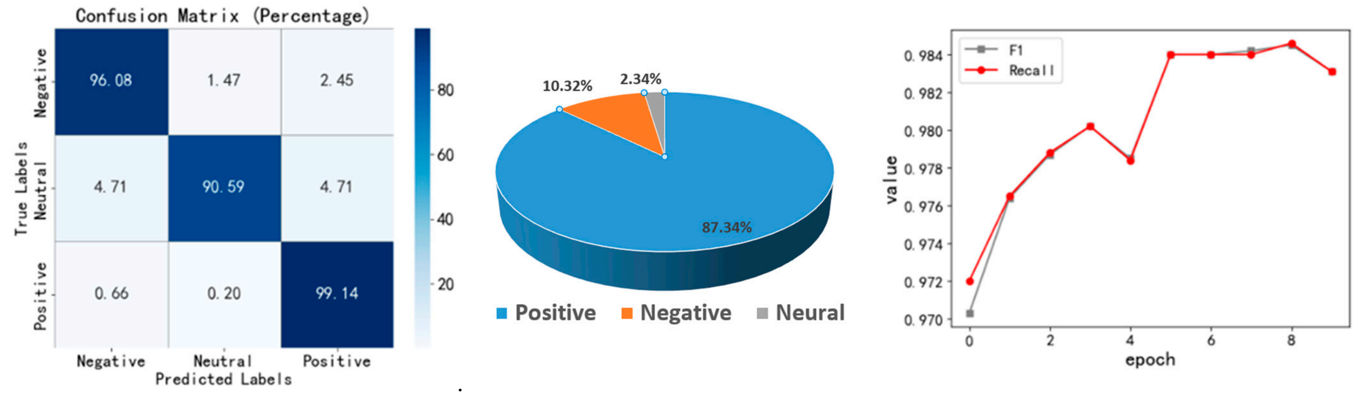


Figure 6. Sentiment analysis proportion chart classified by sentiment scores.

As depicted in Figure 6, positive vocabulary constituted 88.14% of all of the analyzed comments, with negative vocabulary accounting for approximately 9.08%. These data suggest that the general perception of tangerines among users is highly favorable, indicating broad consumer appreciation for the product. These findings underscore the overall positive sentiment in the consumer reviews of tangerines, reflecting satisfaction with the product's quality, taste, and shopping experience. The high proportion of positive comments indicates that the majority of consumers have had favorable interactions with the product and service offerings related to tangerines.

3.4. Bert Sentiment Classification

To illustrate the suitability of SnowNLP for short reviews, particularly in the context of e-commerce, we conducted text classification experiments using the f model. For this purpose, a subset of the reviews was manually annotated to form a dataset for the classification task. The BERT model was then trained and tested on this dataset, and its performance was evaluated using the confusion matrix, recall, and F1-score, with the results presented in Figure 7. The confusion matrix (Figure 7A) provides a visual representation of the model's performance in a three-class classification task, with the accuracy of each class indicated on the diagonal. Specifically, the accuracy for the positive class was 99.14%, 90.59% for the neutral class, and 96.8% for the negative class. These high accuracy rates demonstrate the BERT model's effectiveness in distinguishing from different sentiment categories.



(**A**) Bert confusion matrix (**B**) Bert sentiment classification chart (**C**) Bert F1-score and recall

Figure 7. Result of Bert sentiment classification.

Recall, a critical metric for evaluating a model's ability to identify relevant instances, reflects the proportion of actual instances that were correctly identified. The experimental results showed a steady increase in recall from 0.97 to 0.984, further validating the model's enhanced recognition capabilities across different categories. The F1-score (shown in Figure 7C), which balances precision and recall, also showed a continuous increase as the recall improved, indicating the model's consistent performance in the classification task. The experimental outcomes suggest that SnowNLP performed comparably to the BERT model on e-commerce review datasets. Moreover, SnowNLP demonstrated greater efficiency and resource-friendliness when processing short texts, thereby confirming its value in such scenarios. This makes SnowNLP a viable and practical choice for sentiment analysis in e-commerce settings, where short and concise reviews are prevalent.

The positive sentiment expressed in the reviews can primarily be attributed to three key factors. Firstly, the taste of the excellent flavor of the tangerines surpassed the expectations of most users, with the sweetness being particularly satisfying. This suggests that the quality of tangerines offered a quality that aligned well with the consumers' desires for fresh and delicious fruit, which is perhaps why tangerines can enjoy the title of local characteristic fruit. Secondly, the attentive service provided by the sellers was highly praised by many users. Finally, regarding the price, although a few users mentioned that the price of the e-commerce mandarin oranges was slightly high, the majority of consumers believed that these mandarins offered an excellent balance between price and quality, providing great value for money. Aspects such as expedited shipping, careful packaging, and responsive after-sales support contributed significantly to improving consumer satisfaction with the product. Despite the predominantly positive feedback, there were certain areas that generated negative sentiments. Some tangerines were reported to have spoiled during storage and transport, which greatly marred the shopping experience for those users who received them. Additionally, some users received underripe or esthetically unpleasing fruit [34]. The small size and unappealing appearance, coupled with the sour or bitter taste due to insufficient ripeness, did not meet the expectations of these consumers.

In conclusion, while there was some noted negative feedback, the overall consumer evaluation of the tangerines was highly positive. The areas of flavor and service stood out as particularly strong points where the tangerines garnered significant approval. It is recommended that the sellers continue to uphold the high quality of their products and services to maintain consumer satisfaction. Addressing the issues related to spoilage during transit and ensuring that only ripe and visually appealing fruit is shipped could further enhance the positive sentiment and loyalty of consumers.

4. Conclusions and Recommendations

This article provides a comprehensive analysis of online reviews for tangerines from major domestic e-commerce platforms in China, revealing three key themes that are of paramount concern to consumers: 'taste', 'service', 'price', and 'quality'.

Regarding **taste** item, businesses must meticulously monitor the flavor of their tangerines, and a challenge is exacerbated by the natural variability in fruit flavors due to factors such as ripeness, sweetness, and acidity. To enhance and maintain a consistent taste, sellers are advised to select superior varieties and harvest at the optimal time. Scientific cultivation practices that control environmental factors like temperature and humidity, along with proper water and fertilizer management, significantly influence the final taste of the fruit. Post-harvest, appropriate packaging is crucial for preserving the tangerines' quality and flavor, ensuring they reach the consumers in prime condition. By vigilantly tracking customer feedback and making data-driven adjustments, businesses can fine-tune their processes to improve taste and bolster customer satisfaction [35].

In terms of **service** item, robust customer service and after-sales support are highlighted as essential components of consumer experience. Businesses should ensure that customer service representatives are courteous and professional, addressing any legitimate concerns or requests for returns or exchanges in a timely manner. Efficient and empathetic service can significantly enhance customer satisfaction with the overall service.

Regarding **quality** item, the packaging of the product is underscored as impactful on the freshness and condition of the fruit upon delivery. While packaging is vital to prevent damage during transit, excessive packaging can drive up costs and negatively affect consumer perception. Businesses are encouraged to adopt functional and minimalistic packaging that balances protection with sustainability [36].

Price item, is considered a key factor influencing consumer purchasing decisions. While setting competitive prices is crucial to attract buyers, excessively low prices may reduce the perceived quality of the fruit and impact profitability. E-commerce businesses are encouraged to find a balance between affordability and value, offering competitive pricing that reflects the quality and freshness of the fruit while ensuring sustainable profit margins. Finally, this study provides valuable insights for Zhejiang mandarin orange e-commerce businesses, enabling them to better cater to consumer needs and enhance customer satisfaction. By focusing on the four identified themes and prioritizing them according to their significance, these businesses can strategically refine their sales approaches and optimize their service delivery to meet consumer expectations more effectively.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae11020151/s1, Figure S1: Word Cloud for the high-frequency comments in the Chinese version; Table S1: The translation of the high-frequency words; Table S2: Sentiment Classification; Table S3: Example of Sentiment Lexicon Ontology Format.

Author Contributions: Conceptualization, F.F.; Data curation, Y.L.; Formal analysis, H.L.; Funding acquisition, L.Y.; Investigation, F.F., X.Z. and L.L.; Methodology, L.Y.; Project administration, L.Y.; Resources, F.F. and L.L.; Software, H.L. and L.L.; Supervision, L.L.; Validation, H.L., Y.L. and X.Z.; Writing—original draft, H.L. and Y.L.; Writing—review and editing, L.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Wenzhou Science and Technology Specialist Project (X2023011), the Science and Technology Plan Project of Wenzhou Municipality (N20220015), and the Zhejiang Province College Students' Science and Technology Innovation Activity Plan (2024R429A031).

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Marrese-Taylor, E.; Velásquez, J.D.; Bravo-Marquez, F.; Matsuo, Y. Identifying Customer Preferences about Tourism Products Using an Aspect-Based Opinion Mining Approach. *Procedia Comput. Sci.* **2013**, *22*, 182–191. [CrossRef]
2. Chen, L.; Wang, F. Preference-Based Clustering Reviews for Augmenting e-Commerce Recommendation. *Knowl.-Based Syst.* **2013**, *50*, 44–59. [CrossRef]
3. Agarwal, A.; Hosanagar, K.; Smith, M.D. Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets. *J. Mark. Res.* **2011**, *48*, 1057–1073. [CrossRef]
4. Kang, D.; Park, Y. Review-Based Measurement of Customer Satisfaction in Mobile Service: Sentiment Analysis and VIKOR Approach. *Expert Syst. Appl.* **2014**, *41*, 1041–1050. [CrossRef]
5. Chen, Q.; Sheng, R.; Wang, P.; Ouyang, Q.; Wang, A.; Ali, S.; Zareef, M.; Hassan, M.d.M. Ultra-Sensitive Detection of Malathion Residues Using FRET-Based Upconversion Fluorescence Sensor in Food. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *241*, 118654. [CrossRef] [PubMed]
6. Gao, J.; Peng, P.; Lu, F.; Claramunt, C.; Qiu, P.; Xu, Y. Mining tourist preferences and decision support via tourism-oriented knowledge graph. *Inf. Process. Manag.* **2024**, *61*, 103523. [CrossRef]
7. Zhang, W.; Wang, S.; Wang, Q. KSAP: An Approach to Bug Report Assignment Using KNN Search and Heterogeneous Proximity. *Inf. Softw. Technol.* **2016**, *70*, 68–84. [CrossRef]
8. Huang, Z.; Li, S.X. Co-Op Advertising Models in Manufacturer–Retailer Supply Chains: A Game Theory Approach. *Eur. J. Oper. Res.* **2001**, *135*, 527–544. [CrossRef]
9. Baxendale, S.; Macdonald, E.K.; Wilson, H.N. The Impact of Different Touchpoints on Brand Consideration. *J. Retail.* **2015**, *91*, 235–253. [CrossRef]
10. Gavilan, D.; Avello, M.; Martinez-Navarro, G. The Influence of Online Ratings and Reviews on Hotel Booking Consideration. *Tour. Manag.* **2018**, *66*, 53–61. [CrossRef]
11. Fang, H.; Jiang, G.; Li, D. Sentiment Analysis Based on Chinese BERT and Fused Deep Neural Networks for Sentence-Level Chinese e-Commerce Product Reviews. *Syst. Sci. Control Eng.* **2022**, *10*, 802–810. [CrossRef]
12. Guo, Y.; Barnes, S.J.; Jia, Q. Mining Meaning from Online Ratings and Reviews: Tourist Satisfaction Analysis Using Latent Dirichlet Allocation. *Tour. Manag.* **2017**, *59*, 467–483. [CrossRef]
13. Xu, X.; Wang, X.; Li, Y.; Haghighi, M. Business Intelligence in Online Customer Textual Reviews: Understanding Consumer Perceptions and Influential Factors. *Int. J. Inf. Manag.* **2017**, *37*, 673–683. [CrossRef]
14. Srinivas, S.; Rajendran, S. Topic-Based Knowledge Mining of Online Student Reviews for Strategic Planning in Universities. *Comput. Ind. Eng.* **2019**, *128*, 974–984. [CrossRef]
15. Dai, Y. The Impact of Data Analytic on Marketing Models in the Digital Economy—The Case of Online Education Institutions. In Proceedings of the 6th International Conference on Economic Management and Green Development (Part4), San Francisco, CA, USA, 6–12 August 2022; Eliwise Academy: Oxford, UK, 2022; p. 9.
16. Sun, T.; Sun, T.; Tan, L. Deciphering Public Sentiment on iPhone 14:A BERT-Based Analysis of Twitter Discourse. In Proceedings of the 2nd International Conference on Interdisciplinary Humanities and Communication Studies (Part2), Stanford, CA, USA, 18 December 2023; Northeastern University: Boston, MA, USA, 2023; p. 9.
17. Ma, D.; Dong, J.; Lee, C.C. Influence of perceived risk on consumers' intention and behavior in cross-border e-commerce transactions: A case study of the Tmall Global platform. *Int. J. Inf. Manag.* **2025**, *81*, 102854. [CrossRef]
18. Wang, B.J.; Fritstiono, E.H. The Relationship of Consumer Trust and Willingness to Buy: The Mediating Role of Consumers' Attitude in Indonesian E-Commerce. In Proceedings of the 2015 SSR International Conference on Social Sciences and Information (SSR-SSI 2015 V10), Tokyo, Japan, 29–30 November 2015; Singapore Management and Sports Science Institute: Singapore, 2015; p. 6.
19. Hobbs, L.; Shanoyan, A.; Wiley, Z.Z.; Aldrich, G. Analysis of Pet-food Customer Postpurchase Experience Using Online Customer Reviews: Implications for Product and Marketing Strategies. *Agribusiness* **2024**, *40*, 885–907. [CrossRef]
20. Guo, X.; Wang, Y.; Tao, J.; Guan, H. Identifying Unique Attributes of Tourist Attractions: An Analysis of Online Reviews. *Curr. Issues Tour.* **2024**, *27*, 479–497. [CrossRef]
21. Liu, Y.; Lu, J.; Yang, J.; Mao, F. China Sentiment Analysis for E-Commerce Product Reviews by Deep Learning Model of Bert-BiGRU-Softmax. *Math. Biosci. Eng.* **2020**, *17*, 7819–7837. [CrossRef]
22. Guangwen, M.; Ran, W.; Shufang, W. A review of China's overseas economic and trade cooperation zones along the Belt and Road:Progress and prospects. *J. Geogr. Sci.* **2023**, *33*, 1505–1526.
23. Da Silva, S.; Paraboni, A.; Matsushita, R. Adapting the National Financial Capability Test to Address Generational Differences in Cognitive Biases. *Int. J. Financ. Stud.* **2024**, *12*, 124. [CrossRef]

24. Zhou, B.; Zhu, Y.; Mao, X. Sentiment Analysis on Power Rationing Micro Blog Comments Based on SnowNLP-SVM-LDA Model. In Proceedings of the 2022 International Conference on Software, Data Processing and Information Technology (SDPIT 2022), Shenyang, China, 14–15 May 2022; Wuhan Zhicheng Times Cultural Development Co., Ltd.: Wuhan, China, 2022; p. 7.
25. He, H.; Ma, C.; Ye, S.; Tang, W.; Zhou, Y.; Yu, Z.; Yi, J.; Hou, L.; Hou, M. Low Resource Chinese Geological Text Named Entity Recognition Based on Prompt Learning. *J. Earth Sci.* **2024**, *35*, 1035–1043. [CrossRef]
26. Hou, Q.; Li, N.; Chao, Y.; Li, S.; Zhang, L. Design and regulation of the surface and interfacial behavior of protein molecules. *Chin. J. Chem. Eng.* **2020**, *28*, 2837–2847. [CrossRef]
27. Tirunillai, S.; Tellis, G.J. Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *J. Mark. Res.* **2014**, *51*, 463–479. [CrossRef]
28. Balla, S.J.; Xie, Z. Online Consultation and the Institutionalization of Transparency and Participation in Chinese Policymaking. *China Q.* **2021**, *246*, 473–496. [CrossRef]
29. Shibata, N.; Kajikawa, Y.; Takeda, Y.; Matsushima, K. Detecting Emerging Research Fronts Based on Topological Measures in Citation Networks of Scientific Publications. *Technovation* **2008**, *28*, 758–775. [CrossRef]
30. Korea Trade Research Association; Chen, X.; Kim, H.-M. Consumer Feedback Analysis Using LDA Approach in Cross-Border E-Commerce. *J. Korea Trade* **2024**, *28*, 77–102. [CrossRef]
31. Han, Y.-Q.; Jeong, W.-J.; Oh, G.-S.; Oh, S.; Whangbo, T.-K. Using LDA Topic Modeling to Understand Regrowth Factors of the Chinese Gaming Industry in the COVID-19 Era: Current Situation, Future and Predicament. *JWE* **2023**, *22*, 433–464. [CrossRef]
32. Bo, H.; Jiaji, J.; Huan, C.; Yimin, Z.; Jin, L.; Zhicai, S. Online Latent Dirichlet Allocation Model Based on Sentiment Polarity Time Series. *Wuhan Univ. J. Nat. Sci.* **2021**, *26*, 464–472.
33. Li, Y. Short Text Classification Improved by Feature Space Extension. In Proceedings of the 2019 the 5th International Conference on Electrical Engineering, Control and Robotics (EECR 2019), Guangzhou, China, 12–14 January 2019; Sichuan Electronics Society: Chengdu, China, 2019; p. 8.
34. Xu, Q.; Bai, J.; Ma, L.; Li, Z.; Tan, B.; Sun, L.; Cai, J. Identification of Multiple Symptoms of Huanglongbing by Electronic Nose Based on the Variability of Volatile Organic Compounds. *Ann. Appl. Biol.* **2023**, *183*, 181–195. [CrossRef]
35. Zhu, A.; Ali, S.; Jiao, T.; Wang, Z.; Xu, Y.; Ouyang, Q.; Chen, Q. Facile Synthesis of Fluorescence-SERS Dual-Probe Nanocomposites for Ultrasensitive Detection of Sulfur-Containing Gases in Water and Beer Samples. *Food Chem.* **2023**, *420*, 136095. [CrossRef]
36. Zareef, M.; Arslan, M.; Mehedi Hassan, M.d.; Ali, S.; Ouyang, Q.; Li, H.; Wu, X.; Muhammad Hashim, M.; Javaria, S.; Chen, Q. Application of Benchtop NIR Spectroscopy Coupled with Multivariate Analysis for Rapid Prediction of Antioxidant Properties of Walnut (Juglans Regia). *Food Chem.* **2021**, *359*, 129928. [CrossRef] [PubMed]