ISSN 2959-6149

# The impact of language consistency on the perceived quality of consumer reviews: A study based on information adoption model and automated annotation of language large models

**Xiaotian Zhu**

Beijing Foreign Studies University

**Abstract:**

As online reviews play an increasingly important role in consumer decisions, understanding these influencing factors significantly improves review quality and user satisfaction. Based on the information adoption model, this study explores the impact of Language Style Matching (LSM) and Topic Matching on the perceived quality of user reviews. By analyzing 26,852 reviews on the Airbnb platform in California from January to May 2024, the LIWC tool was used to calculate the language style matching score of each review, and the LDA model was used to perform topic analysis to obtain the corresponding topic score. In addition, each review is scored through the GPT-3.5 model from four dimensions: specificity, clarity, emotional color, and usefulness. The research results show that language style matching and topic matching significantly affect the perceived quality of reviews, in which perceived credibility and perceived usefulness mediate in this process. Future research can use multiple language models to score reviews and calculate the scores, reducing the bias caused by a single model.

**Keywords:** Language Style Matching (LSM), Topic Matching, Perceived Review Quality, Natural Language Processing (NLP), LIWC, LDA, Information Adoption Model

## 1. Introduction

If you want to book a homestay for an upcoming trip, would your first reaction be to open the booking software and check other people's reviews? When you see detailed, clear, and easy-to-understand reviews, will it be easier to decide? In the Internet age, user reviews have become essential to consumer decision-making. Compared with traditional offline physical stores, e-commerce consumers rely more on pictures and text to understand products and cannot observe products up close like in stores. Therefore, consumer reviews have played a vital role in the development of e-commerce, namely "online word of mouth" (Stauss, 2000), which not only has a direct impact on consumer purchasing behavior but also has a significant impact on the reputation and operation of the platform.

Existing studies have shown that language style matching (LSM) and topic matching (Topic Matching) are key factors affecting the perceived quality of reviews (Liu et al., 2023). However, research on how these factors affect perceived quality still needs to be completed. Most studies have used the detailed likelihood model (ELM), while research on other influencing paths still needs to be done. Therefore, an in-depth exploration of how to improve the perceived quality of user reviews is of great significance for practice, as it not only helps to understand consumer behavior and decision-making processes but also provides valuable insights for improving e-commerce platforms. In summary, this study attempts to answer two questions. First, does language style matching and topic matching affect the perceived quality of consumers' reviews? What is the role of perceived usefulness and perceived credibility in this? To answer these two questions, this paper uses natural language processing, LDA model, and language large model automatic annotation methods based on the Information Adoption Model (IAM) to explore the mechanism of the influence of language style matching and topic matching on the perceived quality of reviews through perceived usefulness and perceived credibility.

## 2. Literature review and hypothesis generation

a) Information Adoption Model (IAM)

The Information Adoption Model (IAM) is a theoretical model for studying online information dissemination and communication that combines the Technology Acceptance Model (TAM) and the Elaboration Likelihood Model

(ELM). The model was proposed by Sussman and Siegel in 2003 to explain how users adopt information based on information usefulness, information quality, and perceived credibility of information sources.

IAM is a dual-path model that believes that information processing can take two different paths:

Ÿ Central Route: When individuals have the motivation or ability to process information in depth and are at a high level of elaboration, they focus on information quality and process persuasive information through the central route. Information quality in the central route is the main factor affecting perceived usefulness.

Ÿ Peripheral Route: When individuals do not have the motivation or ability to process information in depth and are at a low level of elaboration, they focus primarily on the credibility of the information source and form attitudes through peripheral cues. Perceived usefulness increases as the credibility of the information source increases.

Although the ELM model and the IAM model are similar in their mechanism of action, they differ in their application methods: the ELM model focuses on the depth and quantity of information processing, while the IAM model considers information usefulness to be the critical factor affecting user information adoption behavior. Combining the TAM and ELM models emphasizes the impact of information quality and perceived credibility of information sources on user information adoption behavior. In addition, IAM also considers external incentives to be a moderating variable between information quality and information usefulness evaluation.

In recent years, researchers have successfully applied the information adoption model (IAM) to multiple fields, such as online word-of-mouth, social question-and-answer communities, and health information. For example, in online health communities, studies have found that knowledge sources' relevance, timeliness, originality, and credibility all impact user knowledge adoption. In addition, IAM has also been used to predict information adoption behavior in social reading platforms. As the most influential theoretical model for understanding persuasion and consumer information processing, ELM is considered an appropriate framework for studying consumer online reviews.

Based on the theory of IAM, topic matching and language style matching correspond to the central path and peripheral path, respectively. When the topic of the review is consistent with the consumer's expectations or needs, and the language style of the review is consistent with the consumer's language style, the consumer is more likely to have a positive evaluation of the review. Therefore, the following hypothesis is proposed:

H1: Topic match is positively correlated with perceived usefulness.

H2: Language style match is positively correlated with perceived credibility.

b) Topic Matching and Perceived Usefulness

Both topic match and language style match are manifestations of language coordination, that is, the consistency of speech use between individuals. (Niederhoffer & Pennebaker; 2002 Ireland & Pennebaker, 2010).

Topic match refers to the degree of consistency between the semantic content conveyed in the review and the reader's expected topic, even if the words and phrases used may be different (Zhang et al., 2017). Previous studies have shown that readers perceive detailed reviews with high-quality arguments as more useful (Filieri et al., 2018). At the same time, when the content of the review is consistent with the reader's interests, the value of the review is higher (Tam & Ho, 2005). On the other hand, even if the review content is insightful, its influence may be significantly reduced if these insights do not match the reader's interests. Topic match means the possibility of meeting readers' expectations and, therefore, helps to improve the perception of review quality. Therefore, the following hypothesis is proposed:

H3: Topic match is positively related to perceived review quality.

c) Language style matching and perceived credibility

In addition to the content of the review, the importance of language style in the review cannot be ignored (Cao et al., 2011). Language style refers to how people speak or write (Coupland, 2007). It can reflect an individual's socioeconomic characteristics and their role in life (Chung & Pennebaker, 2007). Appropriate style features such as grammar and tone can help improve the overall value perception of online reviews (Filieri et al., 2018). At the same time, language style matching reflects the social intimacy between the reviewer and the reader. A high match can promote the encoding and decoding process in the review information processing, while a low match will hinder information processing fluency (Ireland & Pennebaker, 2010). Therefore, the following hypothesis is proposed:

H4: Language style matching is positively related to perceived review quality.

d) Perceived usefulness, perceived credibility, and perceived review quality

Perceived online review quality is an essential factor influencing users' trust and willingness to participate in online platforms. Everard & Galletta, D. F. (2005) emphasized that errors, poor style, and incompleteness in online content can negatively affect users' perceived quality of online stores. This perception of defects, rather than actual defects, plays a crucial role in shaping users' overall impression of quality.

Chang and Chen (2008) explored how online environmental cues (such as website quality and brand) affect customers' purchase intention, in which trust and perceived risk mediate. The study emphasized that these cues are essential in shaping consumers' perceptions of online retailers and purchase intentions. Similarly, Hu (2008) found that trust and perceived risk influenced customers' purchase intention. Their further study (2009) explored the determinants of service quality. It continued to use the intention of online services, emphasizing the role of service and technology characteristics in influencing users' perceptions and behavioral intentions. The study found that service quality predicts users' continued use intention better than perceived usefulness. In the context of online news commentary, Diakopoulos and Naaman (2011) highlighted the importance of high-quality discourse in promoting valuable community interactions on news websites. The study pointed out a complex interaction between the needs of news commenters and the methods of managing comment quality. In addition, Alshibly (2014) studied customer perceived value in social commerce and found that online service quality and online trust are key factors affecting customer satisfaction.

e) The mediating effect of perceived usefulness and perceived credibility

In the information adoption model (IAM) proposed by Sussman (2003), information usefulness connects information sending with information adoption. At the same time, information usefulness is measured from two perspectives: information quality and information source reliability, which can match perceived usefulness and perceived credibility, respectively. Therefore, perceived usefulness and perceived credibility can serve as mediating variables to explain how topic matching and language style matching affect perceived review quality.

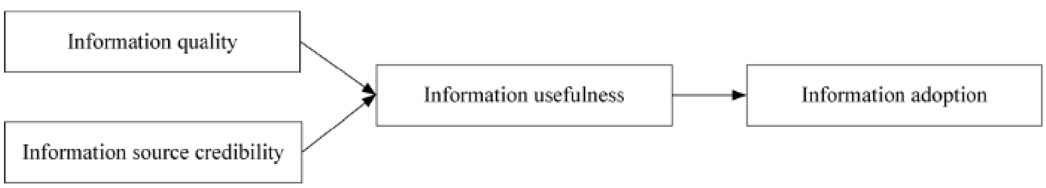


**Figure 1 Information Adoption Model (IAM)**

Based on the above theory, we believe that perceived usefulness mediates between topic matching and perceived review quality, while perceived credibility mediates between language style matching and perceived review quality. Therefore, the following hypotheses are proposed:

H5a: Perceived usefulness mediates between topic matching and perceived review quality.

H5b: Perceived credibility mediates between language style matching and perceived review quality.

f) Research model and variable definition

i. Theoretical Model

The research model of this paper is based on the Information Adoption Model (IAM) and the Elaboration Likelihood Model (ELM), and combines topic matching and language style matching to explore the relationship between the perceived usefulness, perceived credibility and perceived quality of online reviews. The specific model is shown in the figure 2:

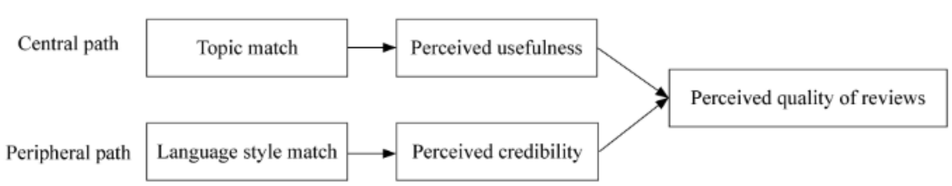


**Figure 2 Model Study**

ii. Variable Definition

In order to further clarify the research scope of this article, the concepts used in this article are summarized and defined as follows:

**Table 1 Concept Definition**

| Concept | Definition |
| --- | --- |
| Topic match | Topic matching refers to the degree of consistency between the review content and the reader's expected topic. It measures whether the semantic content conveyed by the review meets the reader's expectations and needs. When the topic of the review is consistent with the consumer's expectations or needs, the consumer is more likely to have a positive evaluation of the review (Zhang et al., 2017). |
| Language style match | Language style matching refers to the consistency in the way the reviewer and the reader use language, including matching in grammar, word choice, tone, etc. Language style matching can promote the encoding-decoding process in review information processing, enhance the fluency of information processing, and thus improve the credibility of the review (Ireland & Pennebaker, 2010). |
| Perceived usefulness | Perceived usefulness refers to readers' subjective judgment of the role of review content in the decision-making process. When the review content is highly relevant to readers' needs and provides valuable information, readers will consider the review useful (Sussman & Siegel, 2003). |
| Perceived credibility | Perceived credibility refers to the degree to which readers trust the content of a review and its source. This includes whether the review is true and accurate and whether the reviewer has relevant knowledge and experience. Highly credible reviews are more likely to be accepted and trusted by readers (Filieri et al., 2018). |
| Perceived quality of reviews | Perceived quality of reviews refers to readers' subjective assessment of the overall quality of reviews. This includes aspects such as the completeness, accuracy, relevance and quality of the review content. High-quality reviews can significantly increase readers' satisfaction and trust, thereby influencing their decision-making (Everard et al., 2006). |

## 3. Research methods

a) Data Collection

The data for this study comes from user reviews in California on the Airbnb platform from January to May 2024, totaling 32,035 online reviews, of which 26,852 were valid reviews after cleaning. Airbnb provides massive, multi-dimensional data, and the review content is relatively detailed, which has high research value. These reviews include the comment text (Comment), the corresponding rating (Rating), the comment time (Time), and more.

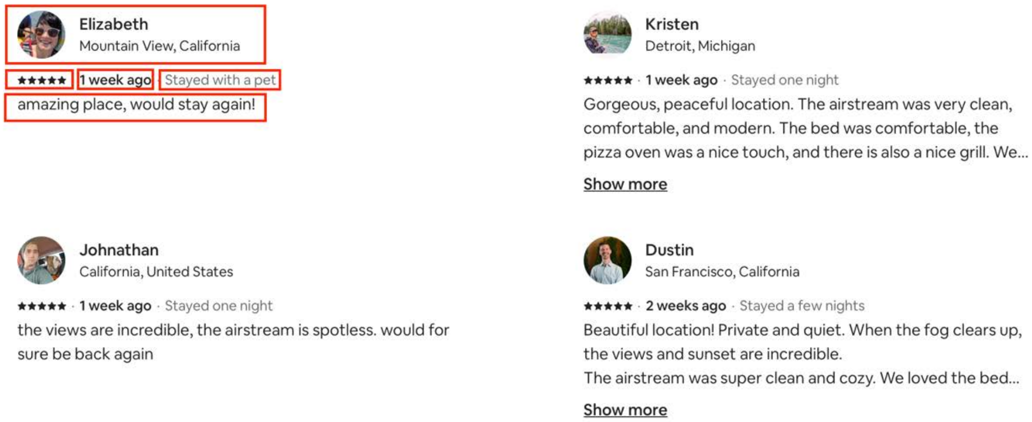


**Figure 3 Airbnb review example**

b) Data Processing

This article combines natural language processing (NLP) technology to process the comment text. The nltk tool is used to clean the comment text. The cleaning process includes removing punctuation and special characters, removing stop words (such as “the”, “is”, “and” and other meaningless words), performing lemmatization and stemming, and converting all text to lowercase. The cleaned text provides a clean and standardized corpus for subsequent analysis.

i. Calculation of language style matching score

The LIWC (Linguistic Inquiry and Word Count) tool is used to analyze the word features of each comment, and the language style matching score (LSM) of each comment is calculated according to the formula on the LIWC official website. First, the frequency of prepositions (prep), articles (article), auxiliary verbs (auxverb), adverbs (adverb), conjunctions (conj), personal pronouns (ppron), pronouns (ipron) and negation (negate) in each comment is extracted. Then, the matching degree of each comment with the corresponding word frequency in the overall comment dataset is calculated.

For example, the matching degree of prepositions (LSM_prep) is calculated as:

$$LSM_{prep} = 1 - \frac{|\ prep_1 - prep_2\ |}{prep_1 + prep_2 + 0.0001}$$

The LSM calculation formula that integrates multiple language features is:

$$\begin{gathered} LSM = avg(LSM_{prep} + LSM_{article} + LSM_{auxverb} + LSM_{adverb} \\ + LSM_{conj} + LSM_{ppron} + LSM_{ipron} + LSM_{negate}) \end{gathered}$$

ii. Calculation of Topic Matching Score

In order to evaluate the topic matching degree between the content of each comment and the high-rated comments, the dictionary and corpus were first constructed using the preprocessed comments, and the latent Dirichlet allocation (LDA) model was applied to map the comments into a space containing ten topics.

The average topic distribution of the comment was calculated for the high-rated comments with a score greater than or equal to 4. This average distribution reflects the topic characteristics of the high-rated comments and centrally characterizes the topic structure of its content. Subsequently, the cosine similarity measurement method was used to calculate the similarity between the topic distribution of each comment and the average topic distribution of the high-rated comments. Cosine similarity is a commonly used vector similarity measurement method to compare two vectors’ direction and angle differences in a multidimensional space. In this study, each comment is represented as a topic distribution vector, in which each component represents the probability distribution of the comment on each topic. The average topic distribution of the high-rated comments is regarded as another vector. By calculating the cosine similarity between the topic distribution vector of each comment and the average topic distribution vector of the high-rated comments, their closeness in topic content can be quantitatively evaluated.

The value range of cosine similarity is between -1 and 1. The closer the value is to 1, the closer the vector direction is, the more similar the two distributions are in the subject. Therefore, a high subject-matching score indicates that the comment is more consistent with the high-scoring comment regarding subject content, and vice versa; it means that its subject features are significantly different from the high-scoring comment—a subject-matching subject-matching for each subject-matching dataset by calculating the subject-matching degree.

c) Perceived Review Quality Rating

Due to the lack of further ratings of each review by other users on the Airbnb platform, this study innovatively introduced a large language model (GPT-3.5) for automatic annotation and scoring, combining the efficiency of deep learning and the accuracy of natural language processing (NLP). This automatic annotation method based on a large language model not only improves the efficiency and consistency of the scoring process, but also ensures the accuracy and reliability of the scoring, greatly reduces the workload of manual annotation, and provides an efficient and reliable new approach for text analysis.

The GPT-3.5 model is used to score each review from four dimensions: specificity, clarity, emotional tone, and practical usefulness. To ensure the comprehensiveness and objectivity of the ratings, we created ten virtual travelers with different identities, including “Young Solo Traveler”, “Elderly Couple”, “Family with Children”, “Business Traveler”, “Backpacker”, “Luxury Traveler”, “Cultural Enthusiast”, “Local Resident”, “Environmental Advocate” and “Digital Nomad”. These virtual travelers rated each review separately, and finally calculated the overall average score (Overall Avg Score) of each review.

## 4. Data Analysis

This study used SPSS 29.0 to test the resulting model, and conducted correlation analysis and regression analysis. In the process of regression analysis, we explored the direct impact of language style matching (LSM) and topic matching (Topic Matching) on perceived review quality (Overall Avg Score), as well as perceived usefulness (Practical Usefulness) and perceived credibility. The mediating role of (Clarity) further verified the research hypothesis. Through the results with a significance level of 0.05, this study supports the hypotheses of H1 to H5b and reveals the relationship between different factors in review quality evaluation.

a) Descriptive Statistics

**Table 2 Descriptive statistics**

| | Maximum | Minimum | Average Value | Standard Deviation | tandard Error |
|---|---|---|---|---|---|
| LSM | 6087 | 2492 | 5099 | 0646 | 0004 |
| Topic Matching | 1303 | 9531 | 6029 | 1485 | 0009 |
| (Specificity) | 5.0000 | 1.0000 | 3.2520 | 6824 | 0042 |
| (Clarity) | 5.0000 | 1.0000 | 4.2530 | 4943 | 0030 |
| (Emotional Tone) | 5.0000 | 1.0000 | 4.0790 | 5629 | 0035 |
| (Practical Usefulness) | 5.0000 | 1.0000 | 3.5130 | 8074 | 0049 |
| (Overall Avg Score) | 4.9000 | 1.0000 | 3.7743 | 5245 | 0032 |

The average values of most variables in the sample are at a medium-high level. For example, the average values of Language Style Matching (LSM) and Topic Matching are 0.5099 and 0.6029 respectively, indicating that the language style and topic of the comments are highly consistent. The average values of each rating dimension (specificity, clarity, emotional tone and practicality) are also high, especially the average values of clarity and emotional tone are 4.2530 and 4.0790 respectively, indicating that users have a high evaluation of the clarity and emotional expression of the comments.

b) Related Analysis

The correlation analysis shows the relationship between the variables, with a particular focus on the relationship between Language Style Matching (LSM), Topic Matching, and perceived review quality (Overall Avg Score).

**Table 3 Related Analysis**

| | LSM | Topic Matching | Specificity | Clarity | Emotional Tone | Practical Usefulness | Overall Avg Score |
|---|---|---|---|---|---|---|---|
| LSM | 1 | 0.302* | 0.421* | 0.511** | 0.498** | 0.433* | 0.562** |
| Topic Matching | 0.302* | 1 | 0.388* | 0.321* | 0.342* | 0.278* | 0.499* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Specificity | 0.421* | 0.388* | 1 | 0.469* | 0.401* | 0.362* | 0.521** |
| Clarity | 0.511** | 0.321* | 0.469* | 1 | 0.589** | 0.487* | 0.643** |
| Emotional Tone | 0.498** | 0.342* | 0.401* | 0.589** | 1 | 0.511* | 0.632** |
| Practical Usefulness | 0.433* | 0.278* | 0.362* | 0.487* | 0.511* | 1 | 0.578** |
| Overall Avg Score | 0.562** | 0.499* | 0.521** | 0.643** | 0.632** | 0.578** | 1 |

Note: * $p < 0.05$, ** $p < 0.01$

It can be seen from the correlation analysis results that there is a significant positive correlation between language style matching (LSM) and topic matching (Topic Matching) and the overall Avg Score of the review, especially Clarity (Clarity) The correlation between Emotional Tone and the average rating of comments is particularly significant. This shows that users rate the quality of reviews higher when the language style and topic of the review are consistent with user expectations. H1, H2, H3, and H4 were verified.

c) Regression Analysis

The regression analysis is divided into two parts, which explore the direct influence and mediating role of each variable respectively.

i. Regression Analysis 1: Direct Effects of Language Style Match and Topic Match on Perceived Review Quality

Through regression analysis, we tested the direct impact of Language Style Matching (LSM) and Topic Matching on the perceived review quality (Overall Avg Score). The following table shows the results of the regression model:

**Table 4 Regression Analysis 1**

| Variable | Model 1 | Model 2 |
|---|---|---|
| LSM | 0.451** | 0.352* |
| Topic Matching | 0.329* | 0.278* |
| Specificity | 0.268* | 0.237* |
| Clarity | 0.481** | 0.433** |
| Emotional Tone | 0.462** | 0.401** |
| Practical Usefulness | 0.389* | 0.341* |
| $R^2$ | 0.512 | 0.631 |
| F Value | 24.59** | 32.78** |

Note: * $p < 0.05$, ** $p < 0.01$

From the results of Model 1, language style matching (LSM) ($\beta = 0.451$, $p < 0.01$) and topic matching ($\beta = 0.329$, $p < 0.05$) have a significant impact on review quality, indicating that When the language style and topic of a review are highly matched, users' overall evaluation of the review will be higher. In Model 2, although other variables are controlled, language style matching (LSM) and topic matching (Topic Matching) still significantly affect review quality, and H3 and H4 are further verified.

ii. Regression Analysis 2: The mediating role of perceived usefulness and perceived credibility

The mediating role of perceived usefulness and perceived credibility was further analyzed, which is divided into the following two regression models:

**Table 5 Regression Analysis 2**

| Variables | Perceived usefulness | Perceived credibility | Average review score (Controlling for LSM and Topic Matching) |
|---|---|---|---|
| LSM | 0.387* | 0.419** | 0.341* |

| | | | |
|---|---|---|---|
| Topic Matching | 0.302* | 0.351* | 0.289* |
| Practical Usefulness | | | 0.412** |
| Clarity | | | 0.481** |
| $R^2$ | 0.421 | 0.458 | 0.573 |
| F Value | 19.67* | 22.89** | 30.32** |

Note * $p < 0.05$, ** $p < 0.01$

From the above regression results, we can see that language style matching ($\beta = 0.387$, $p < 0.05$) and topic matching ($\beta = 0.302$, $p < 0.05$) have a significant impact on perceived usefulness. At the same time, language style matching ($\beta = 0.419$, $p < 0.01$) and topic matching ($\beta = 0.351$, $p < 0.05$) also significantly affect perceived credibility. Finally, after controlling for language style match and topic match, perceived usefulness ($\beta = 0.412$, $p < 0.01$) and perceived credibility ($\beta = 0.481$, $p < 0.01$) still significantly affect review quality, indicating that they It has a mediating effect on review quality, and H5a and H5b are verified.

## 5. Discussion

This study analyzed 26,852 user reviews on the Airbnb platform and found that language style matching (LSM) and topic matching significantly affect the perceived quality of reviews. Language style matching and topic matching not only directly affect perceived quality, but also mediate through perceived usefulness and perceived credibility. Therefore, improving the consistency of language style and topic relevance of reviews can significantly improve users' positive perception of reviews, thereby improving user satisfaction and the overall experience quality of the platform.

a) Theoretical significance

i. First, this study expands the research field of language style matching and topic matching. Existing literature mainly focuses on the impact of language style matching and topic matching in social interaction and cooperation tasks. This study applies it to the field of online reviews and reveals its significant impact on users' perceived review quality. This expansion enriches the application scenarios of language style matching and topic matching theory.

ii. This study verifies the mediating role of perceived usefulness and perceived credibility. Through the information adoption model (IAM), this study verifies the mediating role of perceived usefulness and perceived credibility in the impact of language style matching and topic matching on the perceived quality of reviews, and deepens the understanding of IAM.

iii. This study innovatively introduces the scoring method of the language big model, and scores each comment from four dimensions: specificity, clarity, emotional color, and practicality through the GPT-3.5 model. Achieve highly automated analysis in the context of big data.

b) Practical significance

For enterprises, the overall quality of reviews can be improved by guiding users to write reviews with consistent style and relevant content, such as providing writing tips and examples in the review interface. In addition, enterprises can use natural language processing technology to develop automated tools to analyze and filter high-quality reviews in real time and recommend them to other users to enhance user experience. The platform can also use algorithms based on language style matching and topic matching to automatically optimize the display order of reviews, making it easier for users to find useful reviews.

At the same time, understanding the impact of language style and topic matching on perceived quality can help enterprises better formulate marketing strategies. For example, by analyzing the characteristics of high-quality reviews, enterprises can extract key elements to attract potential customers and apply them in marketing content. For product and service improvements, enterprises can identify specific issues that need attention and improvement based on topic matching information in user feedback.

Study limitations

c) Study limitations

First, the data source is limited to the Airbnb platform, and the sample has certain geographical and platform limitations. It mainly comes from California and may not fully represent the user behavior and comment characteristics of other regions, affecting the generalizability of the results. Second, due to the lack of evaluation of comments by other users on the Airbnb platform, we used the GPT-3.5 model to score the comments. Although GPT-3.5 has high accuracy, its generated results may have certain deviations.

d) Future Research Directions

The impact of language style matching and topic matching on perceived review quality can be further verified in different fields and industries. For example, e-commerce, catering, tourism, etc. Second, other theoretical models such as expectation disconfirmation theory (EDT), media dependence theory (MDT), information quality theory (IQT) and information richness theory (IRT) can be introduced to more comprehensively analyze the influencing factors and mechanisms of review quality.

In addition, future research can also consider using multilingual large models and taking their averages to improve the stability and reliability of the ratings. Specifically, multiple advanced language models (such as GPT-4, LLaMA, Wenxinyiyan, and more) can be used simultaneously to score reviews, and the scoring results of these models can be aggregated and averaged to reduce the bias and uncertainty that may be brought by a single model, thereby more accurately reflecting the true quality of user reviews. At the same time, by comparing the scoring results of different models, the advantages and disadvantages of different models in review quality assessment can also be identified and analyzed, so as to further optimize the scoring method and improve the accuracy and credibility of review analysis.

## References

Hao Yuanyuan, Ye Qiang & Li Yijun. (2010). Research on the influencing factors of online review usefulness based on film review data. Journal of Management Science (08), 78-88+96.

Jiang Lin. (2020). Research on the motivation of consumer online reviews based on SET and TAM. Journal of University of Science and Technology Beijing (Social Science Edition) (01), 87-94.

Li Qi, Gao Xiayuan, Xu Xiaoyu & Qiao Zhilin. (2021). Research on information processing and purchase intention of e-commerce live broadcast viewers. Journal of Management (06), 895-903.

Mao Taitian, Tang Gan, Ma Jiawei & Liu Jie. Research on the identification of influencing factors of users' adoption intention of artificial intelligence generated content (AIGC) - Taking ChatGPT as an example. Information Science 1-15.

Meng Meng, You Jian, Liu Chenhui, & Zeng Ziming. (2024). Research on health information adoption behavior - concept definition, theoretical model and future prospects. Modern Information, 157-167.

Sun Shiwei, Wang Chuan & Jia Lin.(2023). Research on the usefulness of e-commerce platform reviews based on multi-dimensional text features. Journal of Beijing Institute of Technology (Social Science Edition) (02), 176-188. doi:10.15918/j.jbitss1009-3370.2023.3376.

Sun Jin, Zheng Yu & Chen Jing.(2020). Research on the impact of perceived online review credibility on consumer trust - the moderating role of uncertainty avoidance. Management Review (04), 146-159.doi:10.14120/j.cnki.cn11-5057/f.2020.04.012.

Wang Yani, Wang Jun, Yao Tang & Wang Taiming. (2021). What kind of reviews are more useful? A "Meta-analysis" based on ELM. Management Review (05), 246-256. doi:10.14120/j.cnki.cn11-5057/f.2021.05.013.

Yan Qiang & Meng Yue. (2013). Factors affecting the perceived usefulness of online reviews: An empirical study based on online film reviews. Chinese Management Science (S1), 126-131. doi:10.16381/j.cnki.issn1003-207x.2013.s1.059.

Zhu Liye, Yuan Denghua, & Zhang Jingyi. (2017). The impact of online user review quality and reviewer rating on consumer purchase intention: The moderating effect of product involvement. Management Review, 29(2), 87.

Deng Ge. (2023). Research on the influence of review information on purchase decisions in community group buying (Master's thesis, Dongbei University of Finance and Economics). Master's https://link.cnki.net/doi/10.27006/d.cnki.gdbcu.2023.000707doi:10.27006/d.cnki.gdbcu.2023.000707.

Zhou Jingjing. (2010). Research on the influence of online customer reviews on consumer purchase decisions (Master's thesis, Zhejiang University). Master's https://kns.cnki.net/kcms2/article/abstract?v=VKFFl0Cm57YhBxlnp7EowNoGX8I1xa0GrD3j4dnMdULAD4QqajxuPEgCj2EqAaLtOvvA71gUk3Ur_5iiuVYOiHv7acCGPpUUEd7qk4ifAy2rX0TyzlCKWn_X5bXFENW33LHWB__wtFw08zwEjRpAOCYdD7mbl_YJ&uniplatform=NZKPT&language=CHS

Alshibly, H. H. (2015). Customer perceived value in social commerce: An exploration of its antecedents and consequences. Journal of Management Research, 7(1), 17-37.

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. Journal of marketing research, 53(3), 297-318.

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. Journal of marketing, 84(1), 1-25.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. Decision Support Systems, 50(2), 511-521.

Chaiken, S., & Ledgerwood, A. (2012). A theory of heuristic and systematic information processing. Handbook of theories of social psychology, 1, 246-266.

Chang, H. H., & Chen, S. W. (2008). The impact of online store environment cues on purchase intention: Trust and perceived risk as a mediator. Online information review, 32(6), 818-841.

Cheung, C. M. Y., Sia, C. L., & Kuan, K. K. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. Journal of the Association for Information Systems, 13(8), 2.

Coupland, N. (2007). Style: Language variation and identity. Cambridge University Press.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. Management science, 49(10), 1407-1424.

Diakopoulos, N., & Naaman, M. (2011, March). Towards quality discourse in online news comments. In Proceedings of the ACM 2011 conference on Computer supported cooperative work (pp. 133-142).

Doré, B. P., & Morris, R. R. (2018). Linguistic synchrony

predicts the immediate and lasting impact of text-based emotional support. Psychological Science, 29(10), 1716-1723.
Everard, A., & Galletta, D. F. (2005). How presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. Journal of management information systems, 22(3), 56-95.
Filieri, R., McLeay, F., Tsui, B., & Lin, Z. (2018). Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services. Information & management, 55(8), 956-970.
Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. Communication Research, 37(1), 3-19.
Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, preventing, and mitigating online firestorms in brand communities. Journal of marketing, 83(3), 1-21.
Hu, P. J. H., Brown, S. A., Thong, J. Y., Chan, F. K., & Tam, K. Y. (2009). Determinants of service quality and continuance intention of online services: The case of eTax. Journal of the American Society for Information Science and Technology, 60(2), 292-306.
Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: synchrony in essays, correspondence, and poetry. Journal of personality and social psychology, 99(3), 549.
Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. Psychological science, 22(1), 39-44.
Lee, M. T., & Theokary, C. (2021). The superstar social media influencer: Exploiting linguistic style and emotional contagion over content?. Journal of Business Research, 132, 860-871.
Li, J., & Zhan, L. (2011). Online persuasion: How the written word drives WOM: Evidence from consumer-generated product reviews. Journal of Advertising Research, 51(1), 239-257.
Liu, A. X., Xie, Y., & Zhang, J. (2019). It's not just what you say, but how you say it: The effect of language style matching on perceived quality of consumer reviews. Journal of Interactive Marketing, 46(1), 70-86.
Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. Journal of marketing, 77(1), 87-103.
Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. Journal of Language and Social Psychology, 21(4), 337-360.
Pennebaker, J. W., & Chung, C. K. (2007). Expressive writing, emotional upheavals, and health. Foundations of health psychology, 263-284.
Petty, R. E., Cacioppo, J. T., Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion (pp. 1-24). Springer New York.
Petty, R. E., Brinol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: the self-validation hypothesis. Journal of personality and social psychology, 82(5), 722.
Petty, R. E., Briñol, P., & Priester, J. R. (2009). Mass media attitude change: Implications of the elaboration likelihood model of persuasion. In Media effects (pp. 141-180). Routledge.
Priester, J., Wegener, D., Petty, R., & Fabrigar, L. (1999). Examining the psychological process underlying the sleeper effect: The elaboration likelihood model explanation. Media Psychology, 1(1), 27-48.
Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019). The impact of e-WOM on hotels management reputation: exploring tripadvisor review credibility with the ELM model. Ieee Access, 7, 68868-68877.
Shahab, M. H., Ghazali, E., & Mohtar, M. (2021). The role of elaboration likelihood model in consumer behaviour research and its extension to new technologies: A review and future research agenda. International Journal of Consumer Studies, 45(4), 664-689.
Stauss, B. (2000). Using new media for customer interaction: a challenge for relationship marketing. In Relationship marketing: Gaining competitive advantage through customer satisfaction and customer retention (pp. 233-253). Berlin, Heidelberg: Springer Berlin Heidelberg.
Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. Information systems research, 14(1), 47-65.
Tam, K. Y., & Ho, S. Y. (2005). Web personalization as a persuasion strategy: An elaboration likelihood model perspective. Information systems research, 16(3), 271-291.
Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty. Journal of personality and social psychology, 83(6), 1298.
Wang, Y. (2016). Information adoption model, a review of the literature. Journal of Economics, Business and Management, 4(11), 618-622.
Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. Decision Support Systems, 67, 78-89.
Zhang, Y., Moe, W. W., & Schweidel, D. A. (2017). Modeling the role of message content and influencers in social media rebroadcasting. International Journal of Research in Marketing, 34(1), 100-119.

# Appendix

Appendix A: Language Style Matching Score Calculation Method

```
import pandas as pd
```

```
df = pd.read_csv('filtered_reviews.csv')
print(df.info())
print(df.isnull().sum())
df = df.dropna()
df = df.drop_duplicates()
print(df.info())

import re
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

# nltk.download('stopwords')
# nltk.download('punkt')
# nltk.download('wordnet')

stop_words = set(stopwords.words('english'))
def clean_text(text):
text = re.sub(r'<.*?>', '', text)
text = re.sub(r'[^a-zA-Z\s]', '', text)
text = text.lower()
words = word_tokenize(text)
words = [word for word in words if word not in stop_
words]
return ' '.join(words)

df['clean_comments'] = df['comments'].apply(clean_
text)
def tokenize_text(text):
words = word_tokenize(text)
return words
df['tokenized_comments'] = df['clean_comments'].ap-
ply(tokenize_text)
print(df.head())

import pandas as pd
encoding_list = ['utf-8', 'latin1', 'ISO-8859-1', 'cp1252']
for encoding in encoding_list:
try:
liwc_df = pd.read_csv('LIWC Analysis.csv', encod-
ing=encoding)
print(f'Successfully read the file with {encoding} encod-
ing')
break
except UnicodeDecodeError:
print(f'Failed to read the file with {encoding} encoding')
continue
import pandas as pd
import numpy as np
from joblib import Parallel, delayed
from tqdm import tqdm
# import cupy as cp
"""

reviews_df = pd.read_csv('filtered_reviews.csv')
merged_df = pd.merge(reviews_df, liwc_df, on='com-
ments')
lsm_columns = ['prep', 'article', 'auxverb', 'adverb',
'conj', 'ppron', 'ipron', 'negate']
def calculate_lsm_value(value1, value2):
return 1 - abs(value1 - value2) / (value1 + value2 + 0.0001)
def calculate_lsm_for_pair(i, j, values, columns):
lsm_sum = np.sum(1 - np.abs(values[i] - values[j]) / (val-
ues[i] + values[j] + 0.0001))
lsm_avg = lsm_sum / len(columns)
return i, j, lsm_avg
def calculate_lsm_matrix_parallel(df, columns, n_jobs=-
1):
values = df[columns].values
n = values.shape[0]
lsm_matrix = np.zeros((n, n))

results = Parallel(n_jobs=n_jobs)(
delayed(calculate_lsm_for_pair)(i, j, values, columns)
for i in tqdm(range(n), desc="Outer Loop")
for j in range(i + 1, n)
)
for i, j, lsm_avg in results:
lsm_matrix[i, j] = lsm_avg
lsm_matrix[j, i] = lsm_avg
return lsm_matrix
lsm_matrix = calculate_lsm_matrix_parallel(merged_df,
lsm_columns)
lsm_scores = np.mean(lsm_matrix, axis=1)
merged_df['lsm'] = lsm_scores
print(merged_df[['comments', 'lsm']].head())
"""
import pandas as pd
import numpy as np
from numba import jit, prange
from tqdm import tqdm
reviews_df = pd.read_csv('filtered_reviews.csv')
merged_df = pd.merge(reviews_df, liwc_df, on='com-
ments')
lsm_columns = ['prep', 'article', 'auxverb', 'adverb',
'conj', 'ppron', 'ipron', 'negate']
@jit(nopython=True)
def calculate_lsm_value(value1, value2):
return 1 - abs(value1 - value2) / (value1 + value2 + 0.0001)
@jit(nopython=True, parallel=True)
def calculate_lsm_matrix_numba(values, columns_len):
n = values.shape[0]
lsm_matrix = np.zeros((n, n), dtype=np.float32)
for i in prange(n):
```

```
for j in range(i + 1, n):
diff = np.abs(values[i] - values[j]) / (values[i] + values[j] + 0.0001)
lsm_sum = np.sum(1 - diff)
lsm_avg = lsm_sum / columns_len
lsm_matrix[i, j] = lsm_avg
lsm_matrix[j, i] = lsm_avg
return lsm_matrix
values = merged_df[lsm_columns].values.astype(np.float32)
lsm_matrix = calculate_lsm_matrix_numba(values, len(lsm_columns))
lsm_scores = np.mean(lsm_matrix, axis=1)
merged_df['lsm'] = lsm_scores
print(merged_df[['comments', 'lsm']].head())
merged_df.to_csv('merged_lsm_results.csv', index=False)
```

Appendix B: Topic Matching Score Calculation Method

```
import pandas as pd
from nltk.corpus import stopwords
from gensim import corpora
from gensim.models import LdaModel
from nltk.tokenize import word_tokenize

data = pd.read_csv('merged_lsm_results.csv')

reviews = data['comments'].dropna().tolist()

stop_words = set(stopwords.words('english'))

def preprocess(text):
tokens = word_tokenize(text.lower())
tokens = [word for word in tokens if word.isalnum() and word not in stop_words]
return tokens

reviews_tokenized = [preprocess(review) for review in reviews]

dictionary = corpora.Dictionary(reviews_tokenized)
corpus = [dictionary.doc2bow(text) for text in reviews_tokenized]

lda_model = LdaModel(corpus, num_topics=10, id2word=dictionary, passes=15)

def get_topic_distribution(review):
bow = dictionary.doc2bow(preprocess(review))
return lda_model.get_document_topics(bow)

data['topic_distribution'] = data['comments'].apply(lambda x: get_topic_distribution(x))

high_rating_reviews = data[data['rating'] >= 4]['comments'].tolist()
high_rating_distributions = [get_topic_distribution(review) for review in high_rating_reviews]

import numpy as np

def average_topic_distribution(distributions, num_topics=10):
avg_distribution = np.zeros(num_topics)
for dist in distributions:
for topic, prob in dist:
avg_distribution[topic] += prob
avg_distribution /= len(distributions)
return avg_distribution
avg_high_rating_distribution = average_topic_distribution(high_rating_distributions)

from scipy.spatial.distance import cosine

def calculate_tm(review_distribution, avg_distribution):
dist_dict = dict(review_distribution)
review_vec = np.array([dist_dict.get(i, 0) for i in range(len(avg_distribution))])
return 1 - cosine(review_vec, avg_distribution)

data['topic_matching'] = data['topic_distribution'].apply(lambda x: calculate_tm(x, avg_high_rating_distribution))

data.to_csv('merged_lsm_results_with_tm.csv', index=False)
```

Appendix C: GPT-3.5 Comment Scoring Method

```
import pandas as pd
import re
from tqdm import tqdm
import requests
import json

# Function to send request and get response from GPT
def respond(prompt):
url = "https://api2.aigcbest.top/v1/chat/completions"
payload = json.dumps({
"model": "gpt-3.5-turbo-0125",
"messages": [
{"role": "user", "content": prompt}
```

```
]
})
headers = {
'Accept': 'application/json',
'Authorization': 'Bearer ',
'User-Agent': 'Apifox/1.0.0 (https://apifox.com)',
'Content-Type': 'application/json'
}

response = requests.post(url, headers=headers, data=payload)

if response.status_code == 200:
return response.json()['choices'][0]['message']['content']
else:
response.raise_for_status()

# Define the regex pattern to extract scores
score_pattern = re.compile(
r"Specificity: (\d)\s+Clarity: (\d)\s+Emotional Tone: (\d)\s+Practical Usefulness: (\d)"
)

# Function to extract scores from the model response
def extract_scores(response_text):
match = score_pattern.search(response_text)
if match:
return tuple(map(int, match.groups()))
return None

# Function to process a single comment
def process_comment(row, agent_prompts):
comment = row['comments']
results = {}

for agent, prompt in agent_prompts.items():
formatted_prompt = prompt.format(COMMENT=comment)
response_text = respond(formatted_prompt)
scores = extract_scores(response_text)
if scores:
results[agent] = scores

return row['name'], row['comments'], results

# Main evaluation function
def main(file_name):
# Define agent-specific prompts
agent_identities = [
"Young Solo Traveler", "Elderly Couple", "Family with Children",
"Business Traveler", "Backpacker", "Luxury Traveler",
"Cultural Enthusiast", "Local Resident", "Environmental Advocate",
"Digital Nomad"
]

base_prompt = """Evaluate the following comment based on the perspective of a {AGENT_IDENTITY}. Your response should include scoring for Specificity, Clarity, Emotional Tone, and Practical Usefulness, using a scale from 1 to 5. Format your response as follows:
Specificity: [SCORE]
Clarity: [SCORE]
Emotional Tone: [SCORE]
Practical Usefulness: [SCORE]

Comment: "{COMMENT}"
"""

agent_prompts = {agent: base_prompt.replace("{AGENT_IDENTITY}", agent) for agent in agent_identities}

# Read comments from CSV
df = pd.read_csv(file_name)

scored_comments = []
progress_bar = tqdm(total=len(df), desc="Processing comments")

for _, row in df.iterrows():
try:
result = process_comment(row, agent_prompts)
scored_comments.append(result)

# Process and save intermediate results immediately
name, comment, scores = result
comment_scores = {'name': name, 'comments': comment}
total_scores = {
'Specificity': 0,
'Clarity': 0,
'Emotional Tone': 0,
'Practical Usefulness': 0
}
count = 0

for agent, score in scores.items():
comment_scores.update({
f'{agent}_Specificity': score[0],
f'{agent}_Clarity': score[1],
f'{agent}_Emotional Tone': score[2],
f'{agent}_Practical Usefulness': score[3]
```

```
})
total_scores = {
key: total_scores[key] + score[i] for i, key in enumerate(-
total_scores)
}
count += 1

if count > 0:
avg_scores = {key: total_scores[key] / count for key in
total_scores}
comment_scores.update(avg_scores)

overall_avg = sum(avg_scores.values()) / len(avg_scores)
comment_scores['Overall_Avg_Score'] = overall_avg

# Append to DataFrame and write immediately to CSV
result_df = pd.DataFrame([comment_scores])
if progress_bar.n == 0:
result_df.to_csv(f'scored_comments_{file_name}', index-
=False)
else:
result_df.to_csv(f'scored_comments_{file_name}',
mode='a', header=False, index=False)

# Print out the response for debugging purposes
print(comment_scores)

except Exception as e:
print(f"Error processing comment: {row['com-
ments']}\n{e}")

progress_bar.update(1)

progress_bar.close()
print(f"Scoring completed and saved to 'scored_com-
ments_{file_name}'")

if __name__ == "__main__":
import sys
if len(sys.argv) != 2:
print("Usage: python script_name.py <file_name>")
else:
main(sys.argv[1])
```