International Journal of Asian Business and Information Management
Volume 15 • Issue 1

Classifying Different Levels of Customer Satisfaction With Vietnamese Hotel Services by Analyzing Customer Feedback

Ha Thi Thu Nguyen, Vietnam FPT University, Vietnam
https://orcid.org/0000-0003-4414-6944

Hung Nguyen Manh, Vietnam Thuongmai University, Vietnam*
Thoa Bui Thi Kim, Vietnam Thuongmai University, Vietnam

ABSTRACT

The development of online booking systems has created information platforms for sharing customers when choosing a destination. Mining this information helps to understand the customer's experience and measure customer satisfaction with hotel services. Recent studies used this approach with machine learning or language models to mine the data generated by customers on the internet. However, this approach still has some limits when wanting to understand more customer insight.
This article uses linguistics rules to measure customer satisfaction by combining aspects and polarity words. In the first step, the dataset with 21,196 reviews on seven main cities in Vietnam was collected from TripAdvisor. Next, the study developed a series of formulas to measure customer satisfaction with Vietnamese hotel service aspects based on inferential statistics and linguistic rules. Python's VADER library was used to measure overall customer satisfaction for Vietnamese hotels. In the final step, by language analysis, the authors calculate and grade the satisfaction score with hotel aspects from 1 to 5. Moreover, the study discovered the negative aspects of positive reviews, while previous studies were rarely mentioned.

KEYWORDS

customer satisfaction, data analysis, hotel service, linguistic rule, online review, TripAdvisor

1. CLASSIFYING SATISFACTION WITH VIETNAMESE HOTEL SERVICES BY ANALYZING CUSTOMER REVIEWS

Customer experience has become a vital focus for researchers and business leaders in the last two decades. This is particularly true for the hotel service industry, where customer experience is one of the most crucial factors for survival (Mohamed, 2021; Lee et al., 2019; Paulose & Shakeel, 2022; Kim & Kim, 2022). Understanding customer experience is essential for hotels to thrive in the market. The trend of online tourism causes the number of tourists booking traditional tours to decrease rapidly,

DOI: 10.4018/IJABIM.335855
*Corresponding Author

forcing businesses to find new directions to catch up with market demand (Özen & Özgül Katlav, 2023). Travelers also tend to use omni-channel. Before deciding where to go, they usually collect and search for information about the destination. The source of online reviews from customers on online booking sites is a channel that attracts a lot of customers' consideration (Quan Xiao et al., 2021; Alrawadieh & Law, 2019). Over the past decade, hotels have faced numerous challenges in terms of customer reviews (Vo et al., 2022). According to researchers and companies, customer reviews play a critical role in determining the service quality of a hotel and whether it meets international standards (Thu et al., 2020; Mohamed, 2021; Adi, 2022). Visitors often rely on reviews and feedback from previous guests to gather detailed information about the services, additional amenities, and the destination's appeal (Breda et al., 2020). As a result, online booking websites for tours, hotels, and restaurants, such as Agoda and TripAdvisor, have gained global popularity (Gómez-Suárez & Veloso, 2022; Chalupa & Petricek, 2022; Breda et al., 2020). Nowadays, most tourism companies are also part of this supply chain, leading to an ever-expanding tourism industry and increasing international competition (Hu et al., 2019; Breda et al., 2020).

The competition in the global hotel industry is fiercer due to the rise of online booking systems (Chalupa & Petricek, 2022; Gómez-Suárez & Veloso, 2022). Therefore, customer satisfaction is crucial in ensuring customer loyalty and generating positive word-of-mouth, increasing the hotel's reputation (Vo et al., 2022; Paulose & Shakeel, 2022). The hotel industry is vast, and as a result, numerous studies measure customer satisfaction in this area (Zhao et al., 2019; Adi, 2022; Cherdouh et al., 2022; Kim & Kim, 2022; Quan Xiao et al., 2021; Alrawadieh & Law, 2019). Hotels often have relied on questionnaire surveys to gauge customer satisfaction. These surveys are typically conducted using Linkert's measuring scale. However, this approach has a limited sample size, with only a few hundred responses obtained. Additionally, this method provides little information, as it only depends on the content by designers. To address these limitations, modern techniques such as analyzing online reviews posted by customers on booking sites have been adopted (Li et al., 2019; Li et al., 2020). When mining a large number of reviews, hotel managers can gain a better understanding of customer preferences and feedback compared to traditional questionnaires.

Vietnam is a country with high tourism potential with many attractive destinations. In recent years, Vietnam has continuously been involved in the group of destinations with the highest growth in the world. The average growth rate is from 50% to 75%. The number of projects in Vietnam relating to international and regional hotel brands is expected to double in the next three years, from 127 to 261 projects by 2025 (Bui et al., 2022). The development of international investment projects shows that the hotel market in Vietnam will explode in the coming years (Bui et al., 2022). The global and domestic competition will become more intense. However, it is necessary to evaluate the customer's experience with Vietnamese hotel services to assess whether the current hotel quality meets the requirements of tourists or not. Are customers delighted with the quality of hotel services in Vietnam? If not, what aspects are customers unsatisfied with hotel services?

In the past five years, studies on data analysis of customer reviews have become increasingly popular (Alaei et al., 2019; Barnes et al., 2020; Chalupa & Petricek, 2022; Chen et al., 2019; Li et al., 2019; Leal et al., 2019). Customer review data is always available and usable for real-time mining (Ba et al., 2019). These reviews carry the opinions and feelings of customers after experiencing hotel services. This way, we can see customer satisfaction or dissatisfaction after a service experience (Ba et al., 2019; Chen et al., 2019). However, when this data is too big, it is not easy to analyze by conventional means but rather by techniques of extensive data analysis. Machine learning models, artificial intelligence, or natural language processing solve this problem (Ba et al., 2019; Özen & Özgül Katlav, 2023; Akhtar et al., 2017). These results are meaningful for the researchers and are expected by many hotel managers for expanding their business and building marketing campaigns (Zhang & Niu, 2024). To the best of our knowledge, previous studies have rarely used linguistic rules to exploit in detail the unsatisfied aspects of customer satisfaction ratings (Alaei et al., 2019; Barnes et al., 2020; Chalupa & Petricek, 2022; Chen et al., 2019). Very few studies have focused on analyzing data to group satisfaction according to five levels: very satisfied, satisfied, neutral, dissatisfied, and very dissatisfied. Moreover, no studies focus on developing a series of formulas for measuring different levels of customer satisfaction.

This study will contribute a process and formulas to calculate different levels of customer satisfaction. This study used the linguistics rule and inferential statistics to build customer satisfaction measurement formulas. We use the Harvy Web tool to collect representative 3-5 star hotel data from seven major tourist cities of Vietnam, including Hanoi, Ho Chi Minh City, Quy Nhon, Nha Trang, Hue, Da Nang, and Sapa. These are large or famous tourist cities with many visitors and a high demand for hotel stays. Next, the Python VADER Library analyzes customer satisfaction with 21,196 reviews. We measure customer satisfaction with Vietnamese hotel services on three levels:

(1). The overall customer satisfaction with Vietnamese hotel service
(2). The satisfaction and dissatisfaction with hotel aspects, including room, location, staff, services, and restaurant
(3). Calculate and grade scores from 1-5 scale for hotel aspects by analyzing language deeply

The next part of the paper is structured as follows: Section 2 presents research-related issues, concepts, and definitions of data analysis and literature review. The research method is presented in section 3. Section 4 presents the research results. Section 5 are discussions and conclusions.

2. LITERATURE REVIEW

2.1. Big Data in the Hospitality Industry

The use of big data has become increasingly popular in various industries, ranging from product development to machine learning, fraud detection, and customer security. The hotel industry also uses big data analytics to gain valuable business insights. When analyzed and utilized effectively, big data presents numerous opportunities for hotel managers to make informed decisions and improve their operations (Chen et al., 2019; Li et al., 2019; Liu et al., 2017; Lee et al., 2019; Zhang & Niu, 2024).

The hotel industry is a data-rich industry that collects large volumes of different types of data (UNWTO et al., 2016). But data remains an underutilized and undervalued asset for most hotel managers. Most people understand customer loyalty. However, only a few know that data analytics will enhance customer insight and a more detailed understanding of customer needs and preferences and identify new opportunities to attract new customers (Padma & Ahn, 2020; Zhao et al., 2019). Data analytics in the hospitality industry is often used to segment guests by booking trends, behaviors, and other factors to reveal their psychology, emotions, and emerging travel trends. It is paramount for hotel managers to be able to understand guest preferences (food, needs, and room type) and purchasing behavior (frequency, length of stay, time of year) to increase satisfaction and brand loyalty (Adi, 2022; Quan Xiao et al., 2021).

Modern online booking service platforms have become a significant source of big data for the tourism industry, providing an accessible way to gather information (Breda et al., 2020). The data collected from hotel managers alone runs into millions of records, while customer reviews also provide a significant source of unstructured data, similar to a social network. Analyzing this data can help hotel managers and customers (Godnov & Redek, 2018) better understand the hospitality industry and make informed decisions. The typical customer behavior when choosing a hotel is to read reviews from previous experiences and book at a discounted price (Breda et al., 2020; Wu et al., 2022). These online reviews will help customers make suitable decisions. The benefit of these reviews for hotel managers is getting the customer’s feelings and psychology about their services for better customer care (Quan Xiao et al., 2021). In addition, the availability of data helps managers save on operation costs (Godnov & Redek, 2018; Chalupa & Petricek, 2022). Managers can also use customer reviews as a word-of-mouth marketing strategy online. This helps promote the image and reputation of the hotel better, attracting more tourists to the hotel (Breda et al., 2020; Zarezadeh et al., 2022; Zhang& Niu, 2024).

Big data analytics is a cost-effective means of extracting useful information, as per Li et al. (2019). It is much quicker and easier than conducting surveys or interviews (Li et al., 2019). Companies favor this technique to analyze customer behavior since it is created by customers objectively (Paulose & Shakeel, 2022; Sampetua Hariandja & Vincent, 2022; Barnes et al., 2020; Gómez-Suárez & Veloso, 2022). Analyzing customer reviews is essential for hotels to understand their strengths and weaknesses. However, manual data mining can be overwhelming, mainly when a hotel receives thousands of records. Data analysis methods now involve applying data mining models, artificial intelligence, and computer statistics to increase efficiency and accuracy (Hu et al., 2019; Leal et al., 2019). Using these techniques, hotels can make sense of survey data containing various issues. For example, the satisfaction of customers will be measured from reviews that have become popular in existing studies. The authors proposed more methods for customer insight by knowing the overall level of satisfaction or aspect level (Liu et al., 2017; Padma & Ahn, 2020; Sampetua Hariandja & Vincent, 2022). Another task is knowing about hotel quality by rating customers. These reviews are calculated and rated from 1-5 score. It can match the quality of the hotel from 1 to 5 stars (Leal et al., 2019).

Moreover, the hotel data analysis can also develop automatic destination recommendation systems from customer reviews to support decision-making. This function can help customers choose a destination more quickly and effectively (Cui et al., 2022). Another example is customer segmentation from review data (Hananto et al., 2022). In this task, the customer with the same opinion will be ordered into a group. The separating groups will create an effective marketing campaign.

2.2. VADER Library for Text Mining

Text mining is the process of analyzing and extracting information from text. It is a vital part of text analysis in data mining. Text mining extracts information from patterns, trends, or ordinal orders using statistical rules or learning. The process breaks down a complex issue into minor problems such as text categorization, classification, clustering, concept/entity extraction, sentiment analysis, document summarization, and entity relation modeling. (Ristova & Cvetanka, 2020). Companies and organizations are researching text mining to analyze customer sentiment. The objective is to provide personalized customer care by using sentiment analysis to identify the sensitivity behind the customers' words. This analysis can yield valuable insights into customers' behavior and preferences, which can help hotel owners make better management decisions (Ristova & Cvetanka, 2020). Customer reviews often contain sentiment adjectives that reveal their experiences. By analyzing these adjectives that indicate emotional poles, like positive or negative, one can gain a deeper understanding of customers' experiences.

Currently, there are many tools or software to mine textual data. However, Python is still a top tool for performing text-mining tasks. Python provides hundreds of different libraries, but several are used for text. For example, NLTK is a popular analysis library; text can be split word by word or sentence by sentence, filter stop words in the text, label word type, and calculate word frequency (Hutto & Gilbert, 2022). Except for NLTK, Python also provides VADER (Valence-aware dictionary and sentiment reasoner). VADER is a rule-based and vocabulary-based sentiment analysis tool. It is suitable to the sentiment analyzed on social media. It uses a list of lexical features (e.g., words) labeled as positive or negative according to their semantic orientation to calculate text sentiment. VADER returns the probability of a given input sentence being positive, negative, or neutral. VADER is optimized for social media data and can yield good results with data from Twitter, Facebook, and more. The results of Hutto & Gilbert (2022) prove that the polarity and probabilities of words are pos, neg neu, and compound words, which are better than current tools and libraries.

2.3. Knowing the Customer Satisfaction Aspect by Analyzing Online Reviews

Online booking sites gather reviews from customers, which are in the form of unstructured data. While this data type can be read and understood, it must be aggregated to extract valuable information and insights. These reviews often reflect customers' opinions and emotions about the service they received during their stay at the hotel (Alaei et al., 2019; Barnes et al., 2020). Hotel managers often want to gain insights into customer psychology by analyzing reviews. However, manually diagnosing extensive data can take time and effort. To address this challenge, automated tools are used to mine and analyze information and summarize the content (Chalupa & Petricek, 2022). This helps managers identify whether customers are satisfied or dissatisfied with the hotel's services, features, or facilities. This process is commonly referred to as aspect-based extraction in customer sentiment analysis.

Example 1. Let's consider a customer review of a hotel: "The service was great, but the food wasn't that good." The following set of words assembles the review: ["The," "service," "was," "great," "but," "the," "food," "wasn't," "that," "good"]

This customer review identified two important hotel attributes: "service" and "food," both nouns. The customer's opinion is expressed as satisfaction or dissatisfaction in a specific aspect; it has positive or negative connotations. Adjectives and accompanying adverbs describe the customer's opinion (positive or negative). For example, "great" and "good" are two adjectives, and "not" is an adverb. In this sentence, the customer's opinion expresses satisfaction with the service and dissatisfaction with the aspect "food."

Therefore, to extract attributes and aspects of customer review data about hotel service quality, we need to find keywords representing their perception of hotel attributes. With hotel services mentioned in customer reviews, these attributes and aspects are nouns, and if many customers mention them, their frequency will be high (Thu et al., 2020).

2.4. Data Analysis for Measuring Customer Satisfaction

Internet customer experience refers to customers' perception of goods or services that are utilized and described on the internet or online environment (Gómez-Suárez & Veloso, 2022). Compared to the previous concept of customer experience in a traditional environment, customer experiences on the internet for business touchpoints are faster. While managing customer experiences is more convenient due to quick access, it can become more complicated if the experience is terrible and tricky to control. Research shows that customers tend to switch to a competitor after encountering a single negative experience with a brand (Hu et al., 2019). Even reading a negative review can leave a bad impression on a customer, resulting in them leaving the brand. Additionally, if a customer cannot access a business's website, it can lead to an immediate loss of satisfaction. This highlights the importance of internet customer experience, which refers to the various online channels such as social media, websites, and other media that businesses can utilize to connect with customers and improve their overall experience with the brand.

One of the recent research directions on hotel experience is text mining of customer-generated content on the internet (Mohamed, 2021). This technique has been shown to have powerful applications, with data sets that can contain millions of records but can still be easily mined (Liu et al., 2017). Text mining is the primary tool for analyzing customer reviews in the hotel industry (Hutto & Gilbert, 2022; Ristova & Cvetanka, 2020). Customers share their feedback not only on booking sites like Booking, Hotels, and TripAdvisor but also on many other platforms (Breda et al., 2020). However, TripAdvisor's data is commonly used for research because its algorithms can detect fraudulent hotel ratings and fake reviews (Leal et al., 2019). To measure customer satisfaction in a country or region, researchers often select a representative number of major cities from that country. Li et al. (2019) provide an excellent example of this approach, having studied five major Chinese cities, including Sanya, Beijing, Guangzhou, Shanghai Hai, and Hangzhou. Additionally, some studies have collected online reviews from platforms in languages other than English. Text mining is often used to analyze these non-English reviews. It also brings significant results when using texting to understand customer experiences.

In previous studies, identifying topics was the core issue when analyzing customer reviews (Ba et al., 2019; Özen & Özgül Katlav, 2023). It is also known as the problem of entity extraction. It is identifying aspects of hotel service such as "room" or identifying positive words "good" or "terrible." The studies use related techniques such as LDA (Godnov & Redek, 2018) to identify topics to perform the extraction task. They also found that topic words are often the nouns mentioned in the reviews. Therefore, researchers extracted frequently used words representing important aspects of the customer experience (Ristova & Cvetanka, 2020). These words are considered for hotel service topics (aspects) "restaurant," "breakfast," "food," "pool," "reservation," "buffet," "lounge," "bar," "check-in," "floor," "view," "cleanliness," "bed," "bathroom," and "staff."

Many software tools can mine textual data and measure customer satisfaction through sentiment analysis. R and Python are popular choices for research due to their efficiency when working with big data and their ability to integrate open machine-learning libraries (Li et al., 2019; Godnov & Redek, 2018). Python recently became more popular because its libraries support text mining (Zhao et al., 2019). The NLTK library and TextBlob support customer satisfaction measurement at the overall level of services (Zhao et al., 2019). Aylien software and R also identified positive, negative, and neutral values in the analysis (Godnov & Redek, 2018). For Chinese text mining, researchers usually use the ROST CM6.0 tool (Dong et al., 2014). Accordingly, there is a wide selection of tools and programming languages for measuring customer satisfaction. VADER is a highly accurate library for emotional analysis of customer reviews (Hutto & Gilbert, 2022).

Previous studies have identified several issues when measuring customer satisfaction with hotel services. Most methods measure overall satisfaction (Godnov & Redek, 2018) and satisfaction with each aspect of hotel service (Hu et al., 2019). Most studies include aspect extraction and statistics to analyze customer needs. Examples of such studies include Ba et al. (2019) and Özen & Özgül Katlav (2023). Among the statistical methods used, frequency analysis is the most commonly applied to identify specific services that customers are most interested in, such as Location (21.7%), Room (21.4%), Service (14.3%), Dining & Food (8.1%), Value (4.8%), and Facility Availabilities (3.8%) (Dong et al., 2014). In a study conducted by Hu et al. in 2019, it was found that customer complaints often contain negative words related to three main topics: Facilities (such as noise, air, heat, and sound), Service (including issues with smoking, reservations with Priceline, and room type preferences), and Location (factors like distance, proximity to the subway, and being centrally located). These complaints can contribute to customer dissatisfaction with the services provided. In each service aspect, the needs of customers are also very different. For example, in the aspect of "room," " customers mainly mention size, furniture, amenities, location, bathroom, cleanliness, and design (Alrawadieh & Law, 2019). Besides the statistical frequency method, numerous studies have also discovered associations between aspects and polarity words to measure customer satisfaction in each aspect based on inferential statistics (Hu et al., 2019). In addition, some studies have implemented quantitative methods to test the relationship between hotel service dimensions (Godnov & Redek, 2018; Alrawadieh & Law, 2019). Using a regression model, they identified aspects affecting overall customer satisfaction and tested the hypothesis. Similar to the regression model proposed by Li (2019). Recent studies rank overall through numerous variables, such as cleanliness, price, room, service, location, etc. Their studies analyze customer satisfaction for some hotel attributes; they classify domestic and foreign customers. Hotel quality is also classified into different levels; customers often demand better standard hotels (Blomberg-Nygard & Anderson, 2016; Adi, 2022). Similarly, other studies have noted that long reviews have a relationship with lower customer satisfaction; a higher level of diversity and sentiment polarity of textual review leads to higher overall customer satisfaction (Zhao et al., 2019).

Existing studies have focused on analyzing hotel customer experience data that customers have generated on booking sites or online travel platforms (Cui et al., 2022; Hananto et al., 2022; Ristova & Cvetanka, 2020). By utilizing the availability of such data, businesses can uncover valuable insights into customer sentiments, emotions, and perspectives, leading to new business opportunities. By understanding their customers' experiences, companies can improve their service quality and develop better customer care plans, attracting more tourists to their hotels (Chalupa & Petricek, 2022; Bian et al., 2022).

The studies mentioned above have utilized various text-mining approaches. However, no studies have used lexical rules in their research or mentioned how to combine nouns and polar words (adjectives). In addition, the studies have described the method applied but have not yet explained the development of standard formulas and measurement methods. Moreover, they must provide a detailed explanation of the procedure required to address the issue productively. Some recent studies have only examined a relatively small number of samples (Godnov & Redek, 2018; Alrawadieh & Law, 2019); therefore, the results are typical. They only deal with the problem of measuring aspects in satisfied sentences without combining aspect and polarity words (Godnov & Redek, 2018; Li et al., 2020; Alrawadieh & Law, 2019). Some studies have used regression models to determine the impact of independent variables on customer satisfaction (Zhao et al., 2019). Long sentences negatively affect overall ranking (Zhao et al., 2019).

3. METHODOLOGY

3.1. Fundamental of Concepts

In this study, we have provided some definitions integral to understanding the topic.

Definition 1. Set of customer

A set of customers is a set of guests who have used hotel services in Vietnam and are reviewed on the TripAdvisor system. Each customer has a review r_i . So, it is expressed by:

$$\mathcal{G} = \left\{g_1,\ g_2,\ldots,g_m\right\} \tag{1}$$

Definition 2. Satisfaction

Satisfaction of guest g_j is who wrote review r_i and it is measured by the value of sentiment of r_i with a combination function λ.

$$Sas\left(g_j\right) = \lambda.\left(neg\left(r_i\right) + pos\left(r_i\right) + neu\left(r_i\right)\right) \tag{2}$$

Definition 3. Aspect

Aspects are the set of attributes and services provided to guests during their stay at the hotel, such as rooms, staff, and hotel location. Aspects are nouns.

$$\mathrm{A} = \left\{a_1,\ a_2,\ldots,a_m\right\} \tag{3}$$

With a_k is the aspect of the hotel.

Definition 4. Aspect group

The words with the same meaning describe hotel attributes extracted from customer reviews.

Example: Given two reviews

S1: The hotel is *located* in central Hanoi
S2: *Location* of the hotel is in central Hanoi

In the two sentences above, "*located*" and "*location*" have similar meanings.

This study uses aspect groups inherited from the research (Thu et al., 2020) with similar content terms. There are eight representatives of groups including: "location," "room," "check-in," "services," "staff," "meal," "surrounding," and "value". However, in our study, we only used some aspects including: "location," "check-in," "room," "services," and "food" (meal).

Definition 5. Aspect satisfaction

Customers are satisfied with each aspect of the service provided by the hotel and are denoted by $Sas\left(a_k\right)$ With a_k is the aspect of hotel.

Definition 6. Polarity words set

It is a set of adjectives and adverbs extracted from customer reviews. These polarity words are intended to describe the customer's sentiment toward the hotel aspects or the state of each hotel aspect.

$$\mathrm{P} = \{p_i \mid p_i \text{ is adjective or adverb}\} \quad (4)$$

Definition 7. Satisfaction scale for hotel aspect

The satisfaction scale for the hotel aspect is scored on a scale of 1 to 5. For example: 1= very dissatisfied, 2= dissatisfied, 3= neutral, 4= satisfied, 5= very satisfied. This scale is determined by the degree of polarity of the polarity words.

Example 2: Give a table with a list of polarity words. For each satisfaction level, there will be a set of polarity words representing.

The polarity level of satisfaction depends on the expression level of words such as "very" accompanied by descriptive adjectives that will increase the level of expression (called gradable adjective/adverb). For example, "good" indicates satisfied, but "very good" indicates very satisfied (Table 1). In addition, words with a negative meaning will have the opposite meaning of the original polarity word. For example: not, no,…

Definition 8. Aspect substring

They are substrings containing aspects. These strings are generated by the following:

- Step 1: Identify the position of the aspect in the reviews.
- Step 2: From the position of the aspect shift to the left, if we encounter a polarity word, create an aspect substring including the polarity word from the aspect. If we meet the articles or sentence endings, punctuation marks (";",".") then stop and do not create aspect substring.
- Step 3: In case the left substring cannot be created, continue from the position of the aspect to the right; if it meets the polarity word, create the aspect substring, including from the aspect to the polarity word. If encounter articles or sentence endings, punctuation marks then stop and do not create aspect substring.

Table 1. An example of the satisfaction levels with polarity words

No	*Polarity words*	*Level of satisfaction*
1	Exellent, very good, wonderful, gorgeous, stunning,	very satisfied
2	Good, delicious, beautiful, great, majestic, exquisite, good-looking, delightful, cute, easy going, friendly, funny, kind, polite, out going, lovely, nice	satisfied
3	Fair, acceptable, dependable	neutral
4	Bad, terrible, horrible, uncomfortable, awful, bad, diabolical, mediocre, careless, cold, impolite, small, dejected, chilly, heavy, unfriendly	Dissatisfied
5	Very terrible, damnably, worst	very dissatisfied

Table 2. List of some gradable adjectives/adverbs

No	Word	Ord	Word
1	Very	7	Slightly
2	Really	8	A little
3	So	9	Relatively
4	Quite	10	extremely
5	A bit	11	much
6	A lot	12	fairly

Example: Given a review, "The room is large and the kindly staff."

In the above review, two aspects are detected: "room" and "staff."

For aspect "room": From the position of aspect "room," the left aspect substring cannot be created by encountering the article "the," so switching to the right. Have an aspect substring = "room is large and."

For aspect "staff": From the position of the staff, create the left aspect substring = "kindly staff."

In this study, we measure customer satisfaction but at a deeper level and different from previous works. We have three main tasks performed in this study: First, we use Python to measure overall customer satisfaction. Next, we use linguistic rules to split into shorter and group substrings by aspect. The goal is to explore the satisfaction of each aspect in more detail, uncovering the unsatisfactory aspects of the overall satisfaction ratings. This issue has yet to be mentioned in the previous studies. Finally, we discovered a list of common polarity words in this dataset. Grouping polarity words according to each level of satisfaction: very dissatisfied, dissatisfied, neutral, satisfied, and very satisfied to measure customer satisfaction by satisfaction level. It is a new finding based on linguistics rule and inferential statistics.

3.2. Measuring Customer Satisfaction With Hotel Services

Python language offers the VADER library, which can easily measure the sentiment score of reviews. We have used this library to calculate the overall customer satisfaction with hotel services. The overall satisfaction score represents the percentage of satisfied customers with hotel services in Vietnam, and we use the following formula to calculate it:

$$Sas_{overall} = \frac{\#\,positive\ \ reviews}{total\ \ of\ \ reviews} x100\% \tag{5}$$

In which:

- $\#\,positive reviews$: Number of positive reviews
- $total\,of\,reviews$: Total number of reviews

Example: The dataset includes 20,000 customer reviews about hotel services. In wich, there are 16,000 positive reviews. Overall customer satisfaction with hotel services = $\frac{16,000}{20,000} x100\% = 80\%$. It means that 80% customers are satisfied with the quality of service provided by the hotel.

3.3. Measuring Customer Satisfaction With Hotel Aspects

In a positive sentence that expresses overall satisfaction, there may be service aspects that the customer is not satisfied with. The VADER library currently cannot analyze the satisfaction for each hotel aspect. Therefore, we used linguistics rules to separate each service aspect in the substring, then used the VADER library to measure satisfaction for aspects. To measure the hotel aspects, we use the group of aspects as mentioned in definition 4.

We follow these steps:

Step 1: Identify aspects in the reviews including "food," "room," "staff," "location," "service"
Step 2: Create aspect substring
Step 3: Gather aspect substrings into aspect groups.
Step 4: Use the VADER library to measure satisfaction with each aspect of hotel service

Measure the rate of customer satisfaction according to the service aspects extracted in the previous step according to the following formula:

$$Sas\left(a_i\right) = \frac{\#\ positive\,aspect\,substring\left(a_i\right)}{total\,of\,aspect\,substring\left(a_i\right)} x100\% \tag{6}$$

Where:

- $\#\ positive\,aspect\,substring\left(a_i\right)$: Number of positive aspect substring of aspect a_i
- $total\,of\,aspect\,substring\left(a_i\right)$: Total number of aspect substring about aspect a_i

3.4. Determining the Satisfaction Level With Each Hotel Aspect

The level of customer satisfaction with each service aspect is expressed on a scale from 1-5 as defined in definition 7. To perform this task, inherit the results from the previous result and perform the pairing of aspects and polarity words as follows:

- Step 1: Gather groups of aspect substrings according to aspects
- Step 2: Extract pairs of polarity words and aspects and remove unrelated words. Each word pair is with the frequency of occurrence.
- Step 3: Calculate the score and determine the satisfaction level of each aspect according to the formula

$$Level\_Sas\left(a_i\right) = \frac{1*\theta_1\left(a_i\right) + 2*\theta_2\left(a_i\right) + 3*\theta_3\left(a_i\right) + 4*\theta_4\left(a_i\right) + 5*\theta_5\left(a_i\right)}{n} \tag{7}$$

In which:

$-Level\_Sas\left(a_i\right)$: Level of satisfaction with aspect a_i

$-\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$: The number of occurrences of polarity word and aspect pairs is equivalent to the score 1, 2, 3, 4, 5.

n : Total number of side-by-side occurrences of polarity word and aspect

Example: Suppose that aspect "staff" has a number of pairings with polarity words in Table 3.

Table 3. A combination of aspect word and polarity word with their occurrences

Aspect	*Polarity word*	*# occurence*
Staff	Kindly, kind	23
Staff	Great	12
Staff	Very kind	5
Staff	Careless	20

Scoring scale for kind/kindly, great=4; careless=2; very kind=5.

$$Level\_Sas(staff) = \frac{1*0+2*20+3*0+4*35+5*5}{60} = 3.4$$

Thus, the aspect "staff" has a customer satisfaction score of 3.4

4. RESULTS

4.1. Description of Data

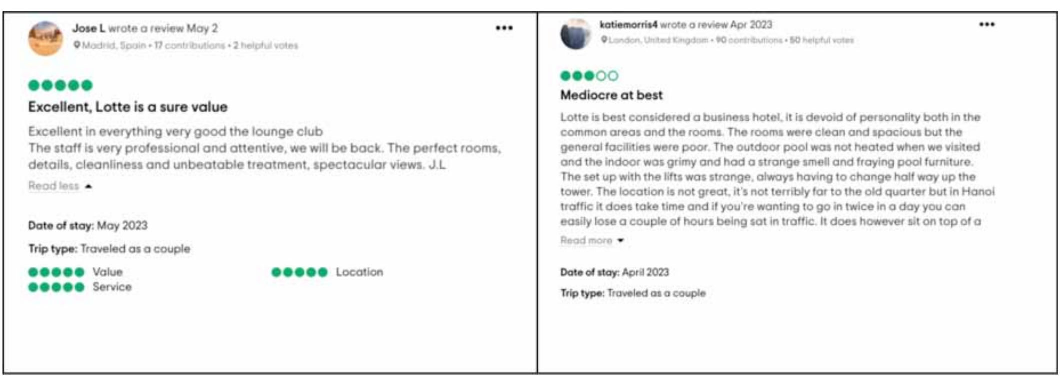
The data used in this study was collected from the TripAdvisor site. TripAdvisor is a platform that provides reliable travel advice from real travelers. The TripAdvisor website is considered the largest travel community in the world, with more than 60 million monthly visits, 44 million market members, and over 100 million reviews from real travelers. Many countries use TripAdvisor, now available in 49 markets in 28 different languages.

To collect data on the TripAdvisor site, we use the Harvy tool. Data is collected from 14 different hotels in seven cities of Vietnam, including Hanoi, Hochiminh, Danang, Quynhon, Nhatrang, Hue, and Sapa. These are major cities or tourist cities with many tourists and represent regions in Vietnam, including the North, the Central, and the South.

The structure of a customer review includes the following key information: reviewer name, review date, free text review, overall star rating, and rating by some aspect of hotel service, as shown in Figure 1. below.

Next, we store the data in .csv format with the following information: name of the reviewer, date of review, the content of the review, the title of the review, and overall rating. Aspect ratings sometimes can not be found in reviews of guests, so it is not necessary to retrieve them. Then, we used Python

Figure 1. Reviews of customers on TripAdvisor (Source: TripAdvisor)



to preprocess the data; after processing, it has 21,596 reviews, including 2,687,646 words in total, and a vocabulary size of 34,452. Figure 2 below depicts the data that Python ran.

From 21,596 reviews collected from 3-5 star hotels in Vietnam, the longest review is 3075 words; the shortest is one word, and the average sentence length is 110.39. Table 4 is the descriptive statistics on the length of customer reviews.

4.2. Measuring Customer Satisfaction With Hotel Services

The VADER library provides tools for measuring sentence sentiment scores in Python programming. This score is to parse positive, negative, and neutral scores and compound scores. Figure 3 below shows the detailed value of each assessment with the scores included:

- Neg (negative): Negative score
- Neu (Neutral): Score represents neutral
- Pos (Positive): The score represents positivity
- Compound: Compound score

Based on the compound score, we classify it into three categories: satisfied, dissatisfied, and neutral. The number of reviews for the satisfaction class is 19,756; dissatisfaction is 1,756; and neutral

Figure 2. Description of data that were collected from the TripAdvisor site

```
                                                      text
0      My husband and I traveled here for business an...
1      We booked the full board package with 3 meals ...
2      Nice location and rooms are spacious with sea ...
3      Yes I would recommend Avani Quy Nhon, check by...
4      My family and I had stayed in Avani Quy Nhon i...
...                                                  ...
21591   The staff is amazingly helpful and friendly. ...
21592                                Th  ch á nice emplyee
21593  á Location is in the center of Hanoi. Close to...
21594                 They provided me a wonderful room.
21595   kaylee from sales really helped us with the t...

[21596 rows x 4 columns]
```

Table 4. Descriptive statistics on the length of customer reviews

description_lengths	
Mean	110.3910272
Standard Error	0.669627048
Median	91
Mode	42
Standard Deviation	95.99518888
Minimum	1
Maximum	3075
Sum	2268646
Count	21956

Figure 3. Scores of reviews

text	description_lengths	scores
My husband	172	{'neg': 0.033, 'neu': 0.77, 'pos': 0.197, 'compound': 0.9885}
We booked	115	{'neg': 0.017, 'neu': 0.716, 'pos': 0.268, 'compound': 0.9895}
Nice locatior	60	{'neg': 0.0, 'neu': 0.888, 'pos': 0.112, 'compound': 0.6055}
Yes I would r	72	{'neg': 0.0, 'neu': 0.71, 'pos': 0.29, 'compound': 0.9769}
My family	203	{'neg': 0.09, 'neu': 0.835, 'pos': 0.075, 'compound': -0.3153}
We spent 8 c	70	{'neg': 0.0, 'neu': 0.881, 'pos': 0.119, 'compound': 0.8435}
I got good	44	{'neg': 0.035, 'neu': 0.815, 'pos': 0.15, 'compound': 0.674}
This is by	167	{'neg': 0.021, 'neu': 0.822, 'pos': 0.157, 'compound': 0.979}
Our family st	138	{'neg': 0.117, 'neu': 0.803, 'pos': 0.081, 'compound': -0.6662}
A beautiful	108	{'neg': 0.0, 'neu': 0.779, 'pos': 0.221, 'compound': 0.9763}
The hotel is ı	41	{'neg': 0.076, 'neu': 0.597, 'pos': 0.327, 'compound': 0.9304}
The hotel is ı	41	{'neg': 0.076, 'neu': 0.597, 'pos': 0.327, 'compound': 0.9304}
Unfortunatel	40	{'neg': 0.117, 'neu': 0.742, 'pos': 0.142, 'compound': 0.2926}
This private ı	193	{'neg': 0.0, 'neu': 0.731, 'pos': 0.269, 'compound': 0.9967}
This hidden	147	{'neg': 0.0, 'neu': 0.788, 'pos': 0.212, 'compound': 0.9889}
What a	37	{'neg': 0.0, 'neu': 0.496, 'pos': 0.504, 'compound': 0.9808}
Beautiful roc	125	{'neg': 0.0, 'neu': 0.718, 'pos': 0.282, 'compound': 0.9918}
Slow	133	{'neg': 0.077, 'neu': 0.822, 'pos': 0.1, 'compound': 0.6697}
I stayed with	71	{'neg': 0.0, 'neu': 0.784, 'pos': 0.216, 'compound': 0.9612}
Celebrated	129	{'neg': 0.014, 'neu': 0.672, 'pos': 0.314, 'compound': 0.9942}
Great service	38	{'neg': 0.061, 'neu': 0.543, 'pos': 0.396, 'compound': 0.9455}
Have a confi	37	{'neg': 0.0, 'neu': 0.78, 'pos': 0.22, 'compound': 0.8622}
Convenient	35	{'neg': 0.0, 'neu': 0.613, 'pos': 0.387, 'compound': 0.9538}
I visited	110	{'neg': 0.025, 'neu': 0.76, 'pos': 0.215, 'compound': 0.968}
I stayed for	114	{'neg': 0.026, 'neu': 0.759, 'pos': 0.215, 'compound': 0.9769}
We had a	130	{'neg': 0.028, 'neu': 0.643, 'pos': 0.329, 'compound': 0.9943}
I had a wond	75	{'neg': 0.041, 'neu': 0.736, 'pos': 0.223, 'compound': 0.9305}
Excellent exc	49	{'neg': 0.0, 'neu': 0.564, 'pos': 0.436, 'compound': 0.9863}

is 84. From the data in Table 3, we calculate customers' satisfaction and dissatisfaction rates. 91.48% are satisfied, 8.13% are dissatisfied, and 0.39% do not clearly express attitudes. The measurement results are shown in Figure 4.

With a rate of 91.48% of customers satisfied with the 3-5 star hotels in Vietnam, it shows that the overall service quality of 3-5 star hotels has met customers' requirements. With the data that has the overall rating according to customer reviews, we found that with compound score values of 0.9 or higher, equivalent to an overall rating of 5-star level, similar to an overall rating of 4 stars, compound values score ranges from 0.7 to 0.9. This can determine the overall rating when, in many cases, customers write reviews without grade stars.

4.3. Measuring the Satisfaction of the Customer by Service Aspect

The customer satisfaction rate with hotel services is over 90%. However, unsatisfied aspects are still found in customer satisfaction reviews. This shows that there are still problems with Vietnam hotel services. So, it is necessary to study each aspect of the hotel in more detail. Figure 5 describes the common words when extracting the aspect substring of the aspect "room."

Figure 4. Satisfaction and dissatisfaction rates with 3-5 star hotels in Vietnam

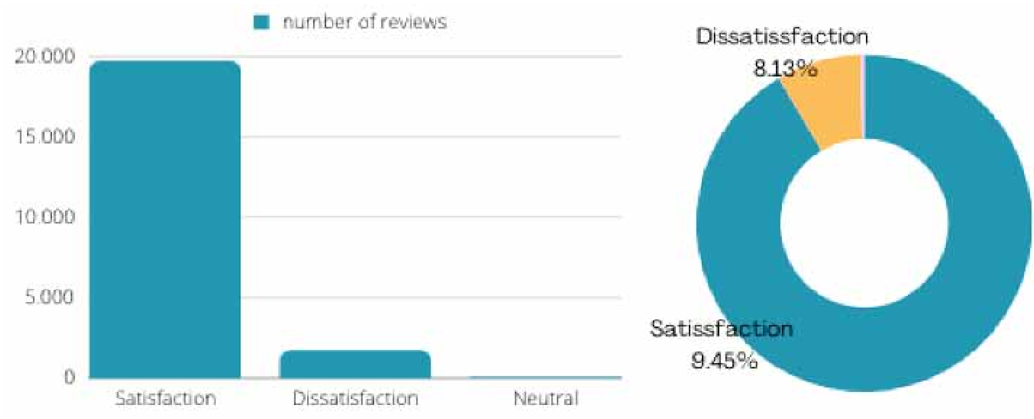
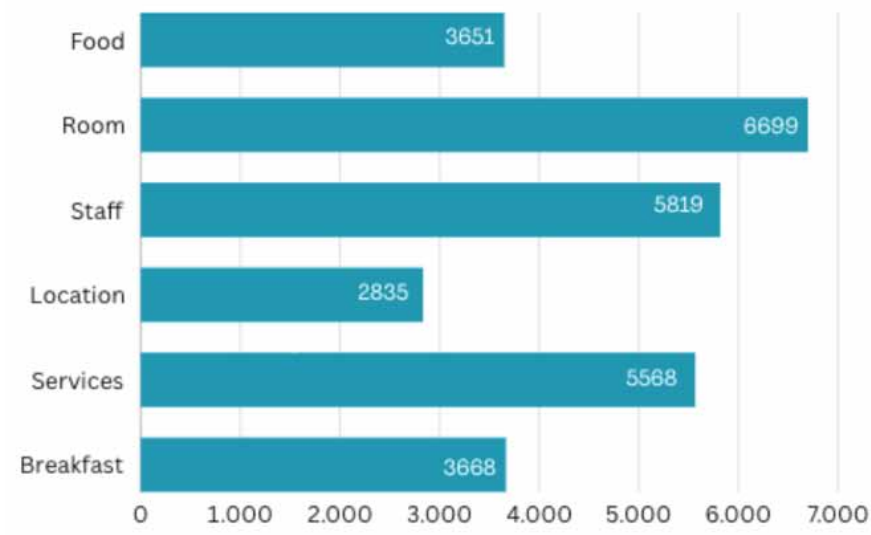


Figure 5. WordCloud of the most frequent polarity words for aspect "room"

We have extracted aspect substrings of aspects: room, staff, food, service, and location. After extracting the aspect substrings, the total number of substrings appearing in 21,596 reviews is listed in Figure 6. This shows that, despite over 20,000 reviews, not all reviews cover all aspects of hotel service. Therefore, the number of aspect substrings extracted from reviews is much smaller than the number of reviews collected.

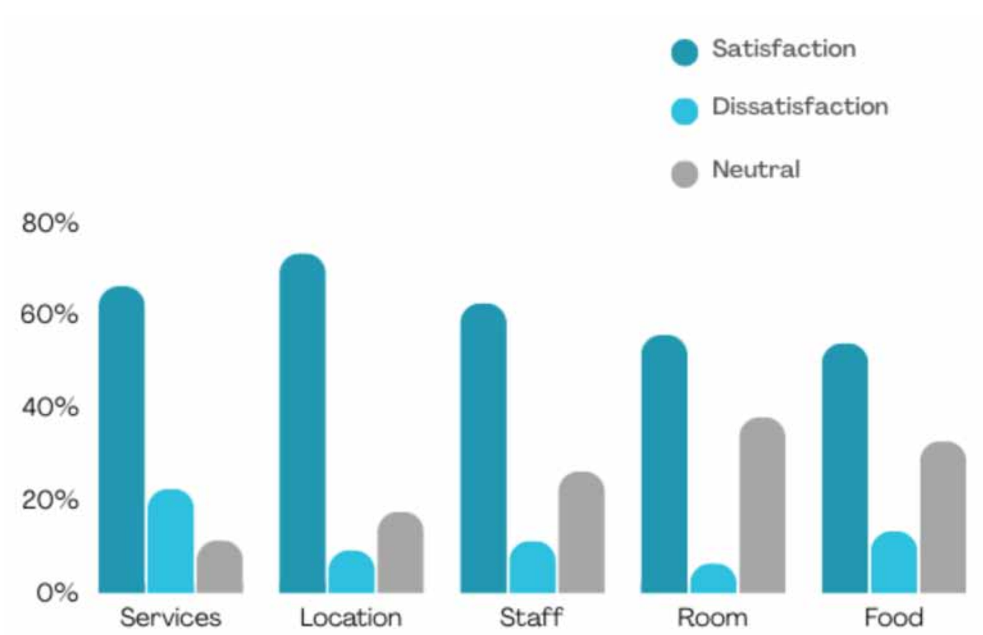
Figure 6. Substring for each aspect



The aspect substrings measure customer satisfaction by using the VADER library. The satisfaction scale is divided into satisfaction, dissatisfaction, and neutral. Figure 7 below is the result of measuring satisfaction according to hotel aspects: services, location, staff, room, and food.

As the results of Figure 7 show, the aspect "location" has the highest satisfaction: 73.24%. The other aspects only reach an average of nearly 60%. Aspect "food" achieved the lowest satisfaction of 53.88%, with no opinion up to 32.78%. This can be easily explained due to the eating habits of

Figure 7. Measuring satisfaction for each aspect by calculating the aspect substring score



each region, so the taste can be difficult to match. On the other hand, in Vietnam, the food service is very developed, and it is easy to find restaurants outside the hotel, so the rate of customers with no comments reached 32.78%. Aspect "room" is the service with the second lowest satisfaction, but also has a high percentage of customers with unclear opinions. Finally, the "service" aspect has the worst review of service quality. This issue also needs to be taken care of and improved to increase the quality of the "services" aspect.

4.4. Measuring the Satisfaction Level With Each Hotel Aspect

In section 3.3, the percentage of customers satisfied with each hotel aspect can be seen. However, customer satisfaction levels have not been measured. An in-depth analysis of the word category provides a more detailed level of satisfaction. At this point, polarity words are found in customer reviews. The frequency of polarity words appearing a lot with a high rate shows that the customer satisfaction shown on the reviews through positive polarity words is very clear. Based on the aspect string obtained with each aspect in the above step, we perform statistics of polarity words. Table 5 below lists some of the polarity words in reviews with their frequency.

We have separated into two word-cloud negative and positive of the aspect "room" for illustrating the frequency of polarity words appearing in each group when combined with the aspect "room." From WordCloud it can be seen that: In the group of positive polarity words for the aspect "room," there will be many positive polarity words: excellent, nice, comfortable, clean, huge, well, good, lovely... Conversely, in the group of negative polarity words combined with aspect "room" include the polarity words: tired, noisy, lower, bad, nothing, trouble, … (Figure 8).

We continue to detail the table by combining aspects "room," "staff," "services," "location," and "food" with polarity words. Table 6 lists some combinations and their frequency when they appear together in aspect substrings. Splitting into aspect substrings makes searching faster and more accurate across all reviews.

Polarity words are assigned a satisfaction score of 1-5 when combined with aspects. We calculate the level of satisfaction based on the combination of polarity words. The formula for measuring satisfaction level by each aspect is presented in section 2.4. Satisfaction results on a scale of 1-5 are shown in Figure 9 below.

Figure 9 shows that the highest customer satisfaction score based on customer reviews analysis is the aspect "location," with a score of =0.475. Next is the aspect "staff" with a score of = 0.467, and the aspect "services" with a score of = 0.442. There is a difference between the satisfaction concerning the aspect and the evaluation scores of the two aspects "staff" and "service." The reason may be that the customer satisfaction rate is high because many people are satisfied at the same level. However, points 4 and 5 are satisfied when polarized according to the scale. If the number of

Table 5. List of polarity words in reviews

Words	*Frequency*	*Words*	*Frequency*
good	0.374	best	0.115
great	0.318	beautiful	0.100
nice	0.213	amazing	0.095
excellent	0.162	old	0.083
friendly	0.161	Small	0.064
clean	0.148	uncomfortable	0.042
helpful	0.145	little	0.009
comfortable	0.128	terrific	0.003

Figure 8. WordCloud with positive and negative of aspect "room"

customers rated 5 points more, the average score of each aspect will be higher. From the results, star level for each aspect as Figure 10 below.

With the calculated scores of aspects, we transform them into different levels of star rating like Figure 10. The staff is rated 4.7 stars, the room is 4.4 stars, the location is 4.8, the services are 4.4 stars, and the food is 3.85 stars.

5. DISCUSION AND CONCLUSION

The customer-centric model has become a strategy in modern business administration, helping companies increase customer satisfaction and customer loyalty, thereby increasing revenue and profit (Luturlean & Anggadwita, 2016; Lee et al., 2019; Rahimian et al., 2021). In the strongly digitalized century, Vietnam's hotel industry has joined the global value chain through online hotel service platforms. Customer experience with hotel services in the context of developing technology is also multidimensional, emotional, and complex (Sampetua Hariandja & Vincent, 2022). The competition is increasing, and just one bad experience will make customers leave and go to competitors (Hu et al., 2019).

Using text-mining techniques, we have tried to show a part of hotel management in Vietnam by understanding customer experiences, feelings, needs, and aspirations. It helps hotel managers understand customers and form more explicit customer portraits through text-mining techniques and deep language analysis. We have divided into three levels of understanding customer experience, from overall to detailed levels, using only analytical techniques and linguistic rules. As follows:

(1). Overall satisfaction with hotel services: The customer satisfaction rate for 3-5 star hotels in seven main cities of Vietnam is 91.48%.
(2). Satisfied with the hotel aspect: We have divided it into five aspects including "services," "staff," "location," "room," and "food." Among these five aspects, the highest satisfaction rate is the "location" aspect, reaching just over 73%.
(3). Difference levels of customer satisfaction with hotel aspect: Expressed by rating each service on a scale and measuring from 1-5 stars.

Table 6. List of some combination of polarity words and aspects

aspect	polarity	# occurrance	with very, extremely, really…	with not, no, don't, doesn't…	with little, quite, …	very + not, extremely, ..
room	beautiful	223	6	32	3	1
	nice	184	93	6	5	
	well	256	32	14	2	
	large	323	34	2	7	
	comfortable	371	99	2		
	great	300		3	9	3
	excellent	197			3	
	clean	749	98	9		1
	wonderful	56			4	
	huge	88	1	17		
	dirty	7		4	4	
	wet	9				
	good	428	34		1	
	small	75	5		12	
	terrific	6				
	ok	73				3
staff	friendly	613	3	6		5
	helpfull	902	5	1		
	incredibly	49	45	5		
	attentive	210	2	2		
	nice	288	3	7		
	enthusiast	13	6	2		
	kind	94	7	12		2
	absolutely	27	2			
	pleasant	91	7			
	lovely	143	8			
	happy	23	2			1
	warm	31	1	2		
	good	315	7	26		7
	polite	23	11	3		
	professional	67	31	6		6
	genuinely	7				
	terrific	7	2			
	cheerful	1				

This research has brought a different approach when quantifying and measuring customer satisfaction using text data analysis. By analyzing at a deep level, research can discover that regarding overall satisfaction level, up to 91.48% of customers are satisfied with Vietnamese hotel services.

Figure 9. The level of customer satisfaction with each hotel aspect

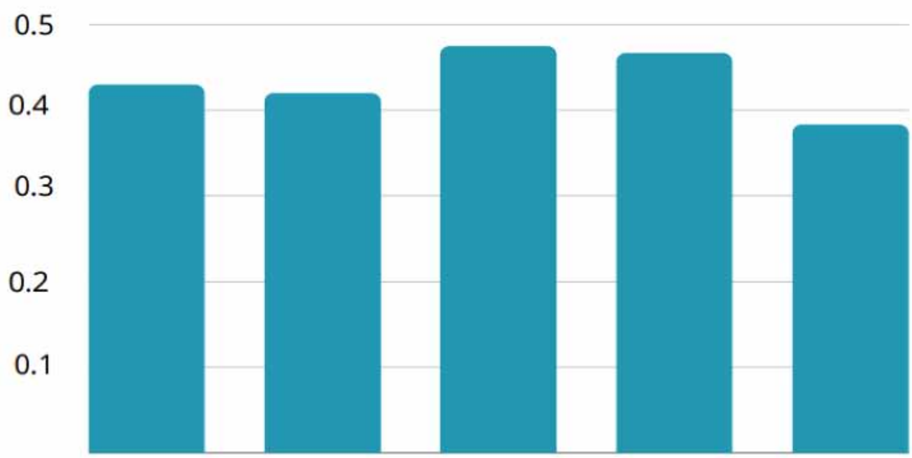


Figure 10. Star rating for each aspect by customers

However, when analyzing the language at a deeper level, one finds that there are many aspects that are unsatisfactory in the overall assessment. In previous studies, the "room" or "staff" aspect often scored highest. This is also easy to explain because the data retrieved is data from large cities and tourist cities near the Vietnamese sea. Therefore, tourists are quite satisfied when choosing hotel locations in Vietnam.

Previous studies have focused on measuring customer satisfaction, however, there are very few studies addressing the level of customer satisfaction. This is a new contribution of this study, which can be further extended to other studies in the future.

REFERENCES

Adi, A. (2022). Does the perceived quality of applications affect customer's trust and satisfaction in online food delivery services. *Jurnal Ekonomi dan Bisnis (EK dan BI), 5*(1), 122-135.

Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based sentiment oriented summarization of hotel reviews. *Procedia Computer Science*, *115*, 563–571. doi:10.1016/j.procs.2017.09.115

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on big data. *Journal of Travel Research*, *58*(2), 175–191. doi:10.1177/0047287517747753

Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. *International Journal of Culture, Tourism and Hospitality Research*, *13*(1), 84–97. doi:10.1108/IJCTHR-08-2018-0104

Ba, H., Nguyen, S., Thang, P. V. M., Le, T., & Huynh, N. (2019). Investigating the Relationships among Sentiment, Hotel Aspects, and Customer's Home Country from Online Reviews: A Machine Learning Approach. SSRN *Electronic Journal*. doi:10.2139/ssrn.3357751

Barnes, S. J., Mattsson, J., Sørensen, F., & Jensen, J. F. (2020). Measuring employee-tourist encounter experience value: A big data analytics approach. *Expert Systems with Applications*, *154*(113450), 113450. doi:10.1016/j.eswa.2020.113450

Bian, Y., Ye, R., Zhang, J., & Yan, X. (2022). Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis. *Computers & Industrial Engineering*, *172*, 108648. doi:10.1016/j.cie.2022.108648

Blomberg-Nygard, A., & Anderson, C. K.United Nations World Tourism Organization (UNWTO). (2016). United nations world tourism organization study on online guest reviews and hotel classification systems: An integrated approach. *Service Science*, *8*(2), 139–151. doi:10.1287/serv.2016.0139

Breda, Z., Costa, R., Dinis, G., & Martins, A. A. (2020). ewow of guests regarding their hotel experience: Sentiment analysis of TripAdvisor reviews. In Handbook of research on social media applications for the tourism and hospitality sector (pp. 295-308). IGI Global.

Bui, H. T., Phi, G. T., Pham, L. H., Do, H. H., Le, A., & Nghiem-Phu, B. (2022). *Vietnam Tourism: Policies and Practice*. Cabi.

Chalupa, S., & Petricek, M. (2022). Understanding customer's online booking intentions using hotel big data analysis. *Journal of Vacation Marketing*, *135676672211221*. doi:10.1177/13567667221122107

Chen, M.-C., Hsiao, Y.-H., Chang, K.-C., & Lin, M.-K. (2019). Applying big data analytics to support Kansei engineering for hotel service development. *Data Technologies and Applications*, *53*(1), 33–57. doi:10.1108/DTA-05-2018-0048

Cherdouh, S., Kherri, A., Abbaci, A., & Kebir, S. (2022). Using Sentiment Analysis of Online Hotel Reviews To Explore the Effect of Information and Communication Technologies on Hotel Guest Satisfaction. *Journal of Tourismology*, *8*(1), 49–67. doi:10.26650/jot.2022.8.1.1038566

Cui, C., Wei, M., Che, L., Wu, S., & Wang, E. (2022). Hotel recommendation algorithms based on online reviews and probabilistic linguistic term sets. *Expert Systems with Applications*, *210*, 118503. doi:10.1016/j.eswa.2022.118503

Dong, J., Li, H., & Zhang, X. (2014). Classification of customer satisfaction attributes: An application of online hotel review analysis. In *IFIP Advances in information and communication technology* (pp. 238–250). Springer Berlin Heidelberg.

Godnov, U., & Redek, T. (2018). Good food, clean rooms and friendly staff: Implications of user-generated content for Slovenian skiing, sea and spa hotels' management. *Management*, *23*(1), 29–57. doi:10.30924/mjcmi/2018.23.1.29

Gómez-Suárez, M., & Veloso, M. (2022). Designing Facebook publications focused on hotel customer experience: How to improve brand attitude and booking intention. In *Brand, label, and product intelligence* (pp. 247–258). Springer International Publishing. doi:10.1007/978-3-030-95809-1_12

Hananto, V. R., Serdült, U., & Kryssanov, V. (2022). A text segmentation approach for automated annotation of online customer reviews, based on topic modeling. *Applied Sciences (Basel, Switzerland)*, *12*(7), 3412. doi:10.3390/app12073412

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management*, *72*, 417–426. doi:10.1016/j.tourman.2019.01.002

Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 216-225). IEEE. doi:10.1609/icwsm.v8i1.14550

Kim, Y.-J., & Kim, H.-S. (2022). The impact of hotel customer experience on customer satisfaction through online reviews. *Sustainability (Basel)*, *14*(2), 848. doi:10.3390/su14020848

Leal, F., Malheiro, B., & Burguillo, J. C. (2019). Analysis and prediction of hotel ratings from crowdsourced data. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *9*(2), e1296. doi:10.1002/widm.1296

Lee, M., Lee, S., & Koh, Y. (2019). Multisensory experience for enhancing hotel guest experience: Empirical evidence from big data analytics. *International Journal of Contemporary Hospitality Management*, *31*(11), 4313–4337. doi:10.1108/IJCHM-03-2018-0263

Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, *32*(5), 1713–1735. doi:10.1108/IJCHM-06-2019-0581

Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). A review of text corpus-based tourism big data mining. *Applied Sciences (Basel, Switzerland)*, *9*(16), 3300. doi:10.3390/app9163300

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism Management*, *59*, 554–563. doi:10.1016/j.tourman.2016.08.012

Luturlean, B. S., & Anggadwita, G. (2016). A framework for conceptualizing customer experiences management in the hotel industry. *Proceedings of the 3rd International Seminar and Conference on Learning Organization (Isclo-15)*. IEEE. doi:10.2991/isclo-15.2016.25

Maglovska, C. R. (2020). What do hotel guests really want? An analysis of online reviews using text mining. *Menadzment u Hotelijerstvu i Turizmu*, *8*(1), 37–48. doi:10.5937/menhottur2001037R

Mohamed, E. S. A. (2021). The impact of customer experience and relationship quality on corporate reputation in the hotel sector. *International Journal of Customer Relationship Marketing and Management*, *12*(2), 53–79. doi:10.4018/IJCRMM.2021040104

Özen, İ. A., & Özgül Katlav, E. (2023). Aspect-based sentiment analysis on online customer reviews: A case study of technology-supported hotels. *Journal of Hospitality and Tourism Technology. Journal of Hospitality and Tourism Technology*, *14*(2), 102–120. doi:10.1108/JHTT-12-2020-0319

Padma, P., & Ahn, J. (2020). Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. *International Journal of Hospitality Management*, *84*, 102318. doi:10.1016/j.ijhm.2019.102318

Paulose, D., & Shakeel, A. (2022). Perceived experience, perceived value and customer satisfaction as antecedents to loyalty among hotel guests. *Journal of Quality Assurance in Hospitality & Tourism*, *23*(2), 447–481. doi:10.1080/1528008X.2021.1884930

Quan Xiao, Q. X., Quan Xiao, S. L., Shun Li, X. Z., & Xing Zhang, M. Z. (2021). Understanding consumer's satisfaction to online hotel review systems: Quality perception and usability. *Wangji Wanglu Jishu Xuekan*, *22*(6), 1299–1311. doi:10.53106/160792642021112206009

Rahimian, S., ShamiZanjani, M., Manian, A., & Esfidani, M. R. (2021). A framework of customer experience management for hotel industry. *International Journal of Contemporary Hospitality Management*, *33*(5), 1413–1436. doi:10.1108/IJCHM-06-2020-0522

Sampetua Hariandja, E., & Vincent, F. (2022). Linking customer experience, satisfaction, and loyalty to brand power and performance in international hotels. *Innovative Marketing*, *18*(3), 59–71. doi:10.21511/im.18(3).2022.06

Thu, H. N. T., Minh, T. T., Binh, G. N., & Ngoc, V. H. (2020). Determining criteria for evaluating quality of Vietnamese hotel through guests online reviews. *International Journal of Business Information Systems*, *1*(1), 1. doi:10.1504/IJBIS.2020.10036288

Vo, N. T., Hung, V. V., Tuckova, Z., Pham, N. T., & Nguyen, L. H. L. (2022). Guest online review: An extraordinary focus on hotel users' satisfaction, engagement, and loyalty. *Journal of Quality Assurance in Hospitality & Tourism*, *23*(4), 913–944. doi:10.1080/1528008X.2021.1920550

Wu, J., Liu, C., Wu, Y., Cao, M., & Liu, Y. (2022). A novel hotel selection decision support model based on the online reviews from opinion leaders by best worst method. *International Journal of Computational Intelligence Systems*, *15*(1), 19. doi:10.1007/s44196-022-00073-w

Zarezadeh, Z. Z., Rastegar, R., & Xiang, Z. (2022). Big data analytics and hotel guest experience: A critical analysis of the literature. *International Journal of Contemporary Hospitality Management*, *34*(6), 2320–2336. doi:10.1108/IJCHM-10-2021-1293

Zhang, D., & Niu, B. (2024). Leveraging online reviews for hotel demand forecasting: A deep learning approach. *Information Processing & Management*, *61*(1), 103527. doi:10.1016/j.ipm.2023.103527

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111–121. doi:10.1016/j.ijhm.2018.03.017

Ha Nguyen Thi Thu is a lecturer researcher from Business Management Department at FPT University. She received a Master's in IT at Guilin University of Electronic Technology in 2006 and is now a Doctor. Her major is in E-commerce and Data analysis. She has more than 20 international papers in this field. She is interested in AI, big data, and how to apply it to Business management.

Hung Nguyen Manh is a lecturer in the Faculty of Human Resource Management at Thuongmai University. He graduated with a master's degree in 2012 and a doctorate in Economic Management from Thuongmai University in 2019. His major research are in Human resource management, Human Development, Human resource economics and Human Capital ; Skills ; Occupational Choice ; Labor Productivity.

Thoa Bui Thi Kim is a lecturer in the Faculty of Human Resource Management at Thuongmai University. She graduated with a master's degree in 2016. Her major research are in Human resource management, Human Capital investment, OSH.