EAI Endorsed Transactions
on Context-aware Systems and Applications

Research Article EAI.EU

Analyzing online reviews at the word level to understand customer experience

Ha Thi Thu Nguyen

Univerity of Greenwich, Ho Chi Minh City, Vietnam
FPT University, Vietnam

Abstract

INTRODUCTION: In the competitive business environment, customer experience plays a pivotal role in driving brand success. Brands that deliver exceptional customer experiences benefit from increased loyalty, advocacy, and stronger market differentiation. With the rise of digital platforms, customers frequently share post-purchase experiences online, making sentiment analysis essential for strategic marketing.
OBJECTIVES: This study aims to explore customer experience with the Trung Nguyen Legend coffee brand by analyzing user-generated content on TripAdvisor. It seeks to identify key aspects of customer feedback and measure satisfaction and loyalty levels.
METHODS: The research employs natural language processing (NLP) techniques and Python-based sentiment analysis tools. Specifically, aspect-based sentiment analysis (ABSA) is used to extract and evaluate sentiment associated with different service dimensions based on online reviews.
RESULTS: The analysis reveals that Trung Nguyen Legend achieves a Customer Satisfaction (CSAT) score exceeding 66% and a Net Promoter Score (NPS) over 34%. These results indicate a generally positive customer experience, with specific strengths and areas for improvement clearly identified.
CONCLUSION: The study demonstrates that ABSA is a cost-effective and time-efficient method for understanding customer sentiment. The findings offer valuable insights for enhancing customer experience management and inform strategic improvements for the Trung Nguyen Legend brand.

Keywords: Customer Experience, Trung Nguyen Legend, Sentiment Analysis, CSAT, NPS, Natural Language Processing

Received on 26 August 2024, accepted on 04 July 2025, published on 08 July 2025

doi: 10.4108/eetcasa.7059

[1]Corresponding author. Email: hantt194@fe.edu.vn

1. Introduction

Understanding customer experience is crucial for businesses operating in a highly competitive market [8], [12]. Customers interact with companies through various touchpoints across multiple channels. Customer experience extends beyond direct product or service interaction and includes all brand-related encounters, particularly online. Providing seamless integration between online and offline interactions ensures a consistent and high-quality experience. Over the past decade, customer experience has gained significant attention in both marketing research and practice. Business leaders now consider it a strategic factor, while marketing scholars view it as a foundation of marketing management [24]. This increased interest has led to numerous academic publications and advancements in understanding the concept.

The internet has generated a vast amount of unstructured textual data through newspaper articles, social media posts, product reviews, and more [11, [26], [33]. Management researchers are leveraging natural language processing (NLP) to extract theoretical insights from this data [23]. NLP, a key component of big data analytics, allows researchers to analyze text for theory building [10], [33]. Online customer reviews, a form of digital word-of-mouth, have become a trusted source of consumer information

[26], [31]. Over 90% of customers read reviews before making a purchase, using them to assess product quality, service, and value. Negative reviews can highlight issues such as poor service or faulty products. Analyzing these reviews enables managers to gain insights into customer sentiments, although the increasing volume of data poses challenges [28], [20].

Sentiment analysis has emerged as a powerful tool for understanding customer experience. It enables businesses to gauge customer sentiment and make informed decisions [21], [12]. Its popularity has grown across industries, governments, and academic communities. Often referred to as opinion mining, sentiment analysis helps monitor public sentiment using user-generated content [32], [9], [22].

Various studies in this domain employ sentiment analysis to enhance product development and customer insights [20]. While many focus on overall sentiment, aspect-based sentiment analysis (ABSA) provides a more nuanced view by associating sentiments with specific aspects of a product or service [2], [3]. However, few studies focus on coffee restaurants, and none provide NPS at both aspects and overall levels. This research addresses that gap by analyzing customer reviews of Trung Nguyen Legend on TripAdvisor. The study objectives include:

- Collecting TripAdvisor reviews on Trung Nguyen Legend
- Identifying aspects customers discuss
- Measuring satisfaction for each aspect and overall
- Calculating NPS at both the aspect and overall levels

Section 2 presents a literature review. Section 3 outlines the methodology. Section 4 presents the results. Section 5 includes discussion, and Section 6 concludes the paper.

2. Literature Review

Internet customer experience refers to how consumers perceive products or services they interact with through digital platforms [40]. Compared to traditional customer experiences, the online environment provides greater speed and convenience in managing customer interactions. However, it also presents unique risks; negative experiences can have immediate and wide-reaching effects. A single unfavorable interaction or a critical online review can discourage potential customers. Likewise, technical issues such as an inaccessible website may lead to a loss of customer trust and loyalty. Therefore, any digital touchpoint, including social media, websites, and other online platforms that enable customer-brand interaction, is a key part of the internet customer experience [10], [11], [24], [25], [40].

In recent years, sentiment analysis of customer-generated online content has become an important focus of research in the service industry. This method allows for efficient analysis of large amounts of data, often involving millions of records [2], [5], [6], [11], [18], [20], [28], [35]. It has become a key tool in big data analytics for assessing customer feedback across various industries [30], [38]. Although user-generated content is spread across various platforms, such as Google Maps and TripAdvisor, researchers often prefer TripAdvisor due to its robust algorithms for identifying fake reviews [24], [19]. To measure customer satisfaction in specific geographic areas, representative review data can be gathered from major cities. Additionally, sentiment analysis has been successfully applied to non-English datasets, with several studies demonstrating its effectiveness in multilingual settings [7]. This wide-ranging applicability highlights the technique's value in cross-cultural and international customer experience studies.

Aspect-Based Sentiment Analysis (ABSA) is a powerful tool for analyzing text data, as it enables researchers to classify and evaluate customer feedback by pinpointing sentiments associated with specific aspects of a product or service [2], [7]. Unlike traditional sentiment analysis, which often returns a general sentiment score, ABSA provides a more detailed view by connecting sentiments to specific aspects, such as service quality, pricing, or ambiance, allowing for more targeted managerial responses [6], [3]. With the large volume of customer reviews and support interactions produced daily, manual analysis becomes impractical, emphasizing the importance of automated ABSA methods [4], [5].

To address the challenge of extracting relevant aspects from unstructured text, a variety of computational methods have been developed. Latent Dirichlet Allocation (LDA) and other topic modeling techniques are among the most commonly used for aspect identification [9], [10]. For example, Ozyurt and Akcayol [2] introduced Sentence Segment LDA (SS-LDA), an innovative adaptation of LDA specifically designed for aspect extraction at the sentence level. Their experimental results showed that SS-LDA performs well in identifying product-related aspects, providing higher granularity and accuracy in sentiment analysis.

Nikolić et al. [4] further advanced ABSA by applying it to educational settings, specifically analyzing students' reviews of teaching staff. Using a combination of natural language processing, machine learning models, linguistic rules, and domain-specific dictionaries, their system was able to identify sentiment with high accuracy. Positive sentiment was detected with an F-measure of 0.83, while negative sentiment reached 0.94. The effectiveness of aspect identification varied based on how often it appeared in the dataset, with F-measure scores ranging from 0.49 to 0.89.

The importance of detailed sentiment analysis over general sentiment scoring has been confirmed by several studies, especially in the hospitality and food service sectors [6], [2], [13], [15]. These studies also show the growing use of artificial intelligence (AI), machine learning (ML), and NLP to improve ABSA models [26], [7], [23], [25], [29], [37], [39].

In the food and beverage (F&B) and restaurant industries, sentiment analysis has provided particularly useful insights. Algorithms have performed well when applied to datasets from platforms like Twitter and TripAdvisor [1], [34]. A notable study by Suryadi and Sabarman [9] analyzed restaurant reviews from five Indonesian tourist destinations, Borobudur, Lake Toba, Likupang, Labuan Bajo, and Mandalika, using LDA to identify key topics. They highlighted seven main themes: Experience, Taste-Value, Food & Italian Ambience, Beach-Bar, Service, and Coffee-Pastry. Sentiment scores were then calculated for each aspect, with Mandalika's restaurants showing the highest positive sentiments across all aspects. Their model achieved nearly 90% accuracy in detecting sentiment and aspect combinations, and restaurants were further grouped based on sentiment profiles using the k-means algorithm. These insights are valuable for business owners and policymakers seeking to improve service quality and customer satisfaction.

Burkov et al. [10] conducted a study based on the Theory of Planned Behavior and the cognitive-affective-conative model to examine how different satisfaction dimensions affect consumer behavioral intentions. Using structural topic modeling, they analyzed a dataset of 286,642 TripAdvisor reviews. Their findings showed that affective factors such as food quality, service quality, ambiance, and perceived fairness of costs significantly influenced customers' willingness to express positive emotions. Specifically, customers more likely to share positive feedback tended to emphasize food quality and staff performance, while those less inclined to do so more often mentioned auxiliary services.

Programming tools like Python and R have become crucial in sentiment analysis research, especially through libraries such as VADER (Valence Aware Dictionary and sEntiment Reasoner) [27], [28], [38]. For instance, a study by Pleerux and Nardkulpat [11] examined TripAdvisor reviews of restaurants in Pattaya City, Chon Buri, Thailand, covering the years 2017 to 2022. Using the VADER model, the researchers observed a significant decline in review volume during the COVID-19 pandemic and a simultaneous increase in negative sentiments. The analysis identified two main areas of customer concern: service and staff, along with food quality and taste. The authors suggest these insights can help improve service quality and better meet customer expectations.

In another study, Nguyen and Nguyen [12] proposed a method for analyzing customer reviews of Vietnamese hotel services. The study collected data from 12 hotels across six major Vietnamese cities: Hanoi, Ho Chi Minh City, Da Nang, Hue, Quy Nhon, and Nha Trang, resulting in a dataset of 20,551 pre-processed reviews. Using Python and the VADER library, along with lexical rule-based techniques, the study found that customers were most satisfied with the "place" aspect (78%). In comparison, the "room" aspect received the lowest satisfaction rating (61.3%). Based on these findings, the authors recommend that hotel managers focus on addressing low-performing service dimensions to improve overall customer satisfaction.

Although these studies show the effectiveness of various NLP and sentiment analysis techniques, none used substring generation from original reviews to boost analytical accuracy. Additionally, few addressed customer satisfaction through Net Promoter Score (NPS) at either the aspect or overall level. While machine learning and topic modeling are common tools, more detailed methods like aspect-based sentiment analysis with improved text segmentation are still largely unexplored in this area.

3. Methodology

3.1 Fundamental concepts

This study uses a variety of concepts and symbols. For clarity, the concepts are defined in detail, and the accompanying symbols and meanings are described in Table 1 below.

Table 1. The definitions and notations

Meaning	Notation	Express	Definition
Set of reviews	R	$R = \{r_1, r_2, \dots, r_s\}$	A set of reviews collected from TripAdvisor
Set of customer	G	$G = \{g_1, g_2, \dots, g_n$	A set of customers refers to guests who have experience and have written reviews on the TripAdvisor platform. Each customer, g_j has a review r_i .
Satisfaction	$Sas(g_j)$		Satisfaction of guest g_j is who wrote review , and it is measured by the value of sentiment of r_i by Python.
Net Promoter Score in overall level	$NPS_{overall}$		Net Promoter Score (NPS) in overall level is a measure used to gauge customer loyalty, satisfaction.
Net Promoter Score in aspect level	$NPS(a_i)$		Net Promoter Score (NPS) in aspect level is a measure used to gauge customer loyalty, satisfaction, and enthusiasm with each aspect.
Aspect	A	$A = \{a_1, a_2, \dots, a_m\}$	Aspects are the set of attributes and services that provided to customer, such as service, staff, and location. Aspects are nouns.

Aspect group	T	$T = \{t_1, t_2, .., t_m\}$	The words with the same meaning are used to describe attributes, these words are extracted from customer reviews.
Aspect substring	aspect s ubstring (a_i)		They are substrings that were cut from original reviews and containing aspects words a_i

3.2 The process of measuring Customer Experience

The customer experience measurement process is performed using the combined algorithms in section 3.3 and Python libraries, according to Figure 1 below.

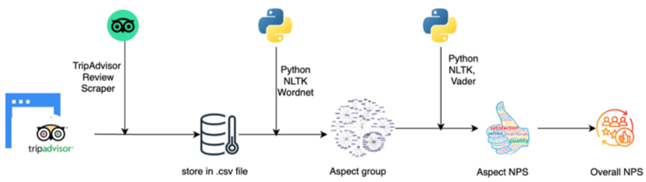


Figure 1. Measuring Customer Experience Process

First, reviews are collected from the TripAdvisor site using the TripAdvisor Review Scaper tool and stored as a .csv file. Next, algorithm (1) and Python's WordNet library are used to determine the aspect word set. To measure satisfaction by dimension, use algorithms number (2) and number (3), combining Python's NLTK and Vader libraries. Finally, the satisfaction scores of the aspects are aggregated to obtain the overall NPS and NPS for each aspect.

3.3 Measuring

Identifying the set of aspects of interest to customers based on natural language analysis.

Previous studies have identified domain-specific dimensions. For example, in the hotel industry, there are set terms such as "room," "staff," etc. However, for the coffee restaurant industry, there is currently no standard set of aspects to measure customer satisfaction. Many studies use topic modeling to determine the set of aspects. In this study, we build a dictionary of these aspects, which are nouns with high frequency and reflect the attributes of the Trung Nguyen Legend restaurant based on natural language analysis and statistics. To build this term set, follow algorithm 1, which includes the following steps:

Algorithm 1: Identifying the set of aspects

- Step 1. Extracting nouns from the review data set that includes s reviews $R = \{r_1, r_2, \ldots, r_s\}$
- Step 2. Removing nouns with the frequency < threshold ∂; obtain a set of nouns $N = \{n_1, n_2, .., n_k\}$and n_j is a noun in R.
- Step 3. Building a set of similar meanings with restaurant and coffee by $Sim(n_i, coffee)$ and $Sim(n_i, restaurant)$ with the library NLTK and WordNet of Python.
- Step 4. The set of words built in step 3 will be matched with the set of nouns N. The terms that do not appear in N will be removed. Finally, we have $T = \{t_1, t_2, .., t_m\}$. With $T \subset N$. Synonymous terms are grouped together into Aspect groups.

Figure 3 below is an example of a coffee-related dictionary created by Python's NLTK and Wordnet libraries.

```
['coffees', 'gourmet_coffee', 'Coffee', 'o_joe', 'Starbucks_coffee', 'coffee_beans', 'latté', 'cappucci
no', 'brewed_coffee', 'espresso', 'java', 'iced_coffee', 'freshly_brewed_coffee', 'coffe', 'decaf', 'fr
esh_brewed_coffee', 'Wolfgang_Puck_gourmet', 'chai_latte', 'lattés', 'latte', 'coff_ee', 'caramel_latt
e', '###.##_-##.##_Liffe', 'macchiato', 'iced_tea', 'coffee_bean', 'Dunkin_Donuts_coffee', 'coffee_sho
p', 'tea_plantations_Gamage', 'iced_lattes', 'specialty_coffees', 'café_latte', 'Starbucks', 'vanilla_l
atte', 'flavored_pasteurized_milk', 'chai_tea_latte', 'cuppa_joe', 'herbal_tea', 'decaf_coffee', 'cappu
cino', 'mocha_latte', 'caramel_macchiato', 'Illy_coffee', 'nonfat_latte', 'raspberry_mocha', 'flavored_
coffees', 'steaming_mug', 'strawberry_smoothie', 'skim_latte', 'chai_tea', 'gingerbread_latte', 'Mystic
_Monk', 'lattes', 'freshly_brewed', 'macchiatos', 'expresso', 'smoothie', 'freshly_brewed_espresso', 'e
spressos_lattes', 'cafe_con_leche', 'Arabica_coffee', 'tea', 'Pancakes_sausage_juice', 'espressos_cappu
ccinos', 'cofee', 'Caramel_Macchiato', 'Cathy_Sila', 'java_jolt', 'coffeehouse', 'Lavazza_coffee', 'Cof
fees', 'cuppa', 'espresso_cappuccino', 'iced_grande', 'gourmet_coffees', 'mango_smoothie', 'drip_coffe
e', 'Iced_Mocha', 'decaf_coffees', 'peppermint_tea', 'teas', 'pancakes_sausage_orange_juice', 'Stumptow
n_coffee', 'chocolate', 'cappuccinos', 'mocha_lattes', 'bottomless_cups', 'barristas', 'steaming_mugs',
'Cafe_Grumpy', 'mocha_frappuccino', 'cappucinos', 'decaffeinated', 'Fairtrade_coffee', 'Sparkletts_®',
'decaffeinated_coffee', 'freshly_brewed_tea', 'cappuccino_latte', 'frappucino', 'café_mocha']
```

Figure 2. Synonyms for Coffee.

Similar to Figure 3, groups are built one by one and form an aspect group. Table 2 below describes a portion of the aspect group data generated by Algorithm 1.

Table 2. Aspect group của Trung Nguyen Legend

Coffee	Service	Place	Price	Staff
Cake Mocha_latte Fruit Tea Ice Capuchino Robusta Aroma Milk Legend Café Bean	Restaurant Car park Table Wifi Music Atmosphe re	Central Floor Street Location Coner Station Airport Locals Enviroment	USD Dong Bill pricing Prices theprice cost	Staffs English staffers staffing

Calculating NPS score for aspects of Trung Nguyen Legend.

Although a customer may express overall satisfaction in a positive sentence, their reviews may still highlight certain service aspects with which they are not happy. Currently, the VADER library is unable to analyze satisfaction for each aspect. Therefore, we have developed an algorithm to segregate each service aspect into a substring and then used the VADER library to measure the satisfaction level for each aspect. To perform this task, we create substrings that are segments containing aspects according to the following algorithm 2.

Algorithm 2. Generating substrings

- Step 1: Identify the position of the aspect in the reviews
- Step 2: From the position of the aspect, shift to the left. If encounter a polarity word, create an aspect substring including the polarity word from the aspect. If encounter 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 from the aspect to the polarity word. If encounter articles or sentence endings, punctuation marks (“;”, “,”, “.”), Then stop and do not create an aspect substring.

Example 2: Given a review “The restaurant is large and the kindly staff.”

In the above review, there are 2 aspects that are detected: “restaurant” and “staff”.

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

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

After creating substrings corresponding to each aspect. Research and calculate satisfaction for each aspect as in algorithm 3 below.

Algorithm 3. Measuring the NPS score for each aspect.

- Step 1: Identify aspects in the reviews, including “coffee”, “service”, “staff”, “location”, “price”
- Step 2: Create an aspect substring
- Step 3: Gather aspect substrings into aspect groups.
- Step 4: Use the VADER library to measure satisfaction with each aspect of Trung Nguyen Legend.

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

$$Sas(a_i) = \frac{\#\ positive\ aspect\ substring\ (a_i)}{total\ of\ aspect\ substring\ (a_i)} x100\% \quad (1)$$

Where :

$\#\ positive\ aspect\ substring\ (a_i)$: Number of positive aspect substring of aspect a_i

$total\ of\ aspect\ substring\ (a_i)$: Total of aspect substring about the aspect a_i

When determining the satisfaction level with each aspect, we also calculate the dissatisfaction of each aspect by the formula (2) below

$$DisSas(a_i) = \frac{\#\ negative\ aspect\ substring\ (a_i)}{total\ of\ aspect\ substring\ (a_i)} x100\% \quad (2)$$

In which $\#\ negative\ aspect\ substring\ (a_i)$: Number of negative aspect substring of aspect a_i .

Companies use Net Promoter Score (NPS) to measure customer loyalty by assessing their likelihood to recommend products or services. Measure the NPS of each aspect according to the following formula:

$$NPS(a_i) = Sas(a_i) - DisSas(a_i) \quad (3)$$

Measuring overall satisfaction, dissatisfaction, and NPS score for Trung Nguyen Legend.

- *Overall satisfaction*

By measuring satisfaction across different dimensions, we can calculate the overall satisfaction score as the average of these values. This approach allows us to gain a more comprehensive understanding of satisfaction levels and make informed decisions based on this data.

$$Sas_{overall} = \frac{\sum_1^k Sas(a_i)}{k} \quad (4)$$

In which:

$Sas_{overall}$: The overall satisfaction with all aspects of Trung Nguyen Legend

$\sum_1^k Sas(a_i)$: The summary of satisfaction of aspect a_1 to a_k.

Example 3: Trung Nguyen’s Legend has five aspects, so k = 5. The $Sas(a_1)$=0.13; $Sas(a_2)$=0.24; $Sas(a_3)$=0.65; $Sas(a_4)$=0.34; $Sas(a_5)$=0.64.

The $Sas_{overall} = \frac{0.13+0.24+0.65+0.34+0.64}{5} = 0.4$

With the value of $Sas_{overall}$=0.4. Customers’ satisfaction with all aspects of Trung Nguyen Legend is equal to 0.4.

- *Overall dissatisfaction*

Similar to the formula for measuring overall satisfaction with Trung Nguyen Legend, the study applies the following formula to measure overall dissatisfaction:

$$DisSas_{overall} = \frac{\sum_1^m DisSas(a_i)}{m} \quad (5)$$

In which:

$DisSas_{overall}$: The overall dissatisfaction with all aspects of Trung Nguyen Legend

$\sum_1^m DisSas(a_i)$: The summary of dissatisfaction of aspect a_1 to a_m.

- *Overall NPS score*

Companies use Net Promoter Score (NPS) to measure customer loyalty by assessing their likelihood to recommend products or services. Đo lường NPS của từng khía cạnh theo công thức sau:

$$NPS_{overall} = Sas_{overall} - DisSas_{overall} \quad (6)$$

4. Results

4.1 Description of data

TripAdvisor is a widely used travel platform that offers reviews, recommendations, and information about hotels, restaurants, attractions, and other travel-related businesses. It enables users to search for accommodations, restaurants, and activities in particular destinations, read reviews and ratings from other travelers, and make reservations directly through the platform. TripAdvisor is well-known for its vast collection of reviews and ratings of hotels, restaurants, and attractions worldwide, which are all provided by users. These reviews allow travelers to share their experiences, opinions, and recommendations and rate businesses on a scale of one to five stars.TripAdvisor ensures reliable reviews by verifying authenticity, monitoring for fraud, and displaying relevant information like travel history and review date [1].

To collect data on the TripAdvisor site, we use an add-on tool on Chrome called TripAdvisor Review Scraper. Data is collected from 14 different Trung Nguyen restaurants around the world. Trung Nguyen chain stores are currently present in several countries on TripAdvisor, such as Vietnam, Singapore, Canada, and Malaysia. The restaurant with the most reviews has only nearly 500 reviews, while Starbucks has 2,510 reviews. This shows the popularity of Starbucks stores and the greater customer interest in Starbucks products than in Trung Nguyen Legend.

The main components of a customer review are the name of the reviewer, the date of the review, the free text review, an overall star rating, and sometimes a rating by some aspect of Trung Nguyen Legend. The collected data is then stored in .csv format, including the name of the reviewer, the date of the review, the content of the review, the title of the review, and the overall rating. After processing the data, we obtained a total of 520 reviews, as shown in Figure 3.

Review Id	User ID	Display Nam	User Name	User Profile	User Avatar	User Locatio	User Is Verif	Rating	Additional Ra	Review Title	Review Text	Helpful Vote	Photos	Stay Date	Created Date	Published Da	Language	Location	Location Id	URL
720568778		Jackie Holm	jackiehK3035	https://www	https://medi	Perth	No	2		disappointing	stayed for 2	3		2019-10-31	2019-10-22	2019-10-22	en	Trung Nguye	601995	https://www
646498411		carol w	511carolw	https://www	https://medi		No	4	Value: 5Serv	Very good fo	Great price f	0		2019-01-31	2019-01-17	2019-01-16	en	Trung Nguye	601995	https://www
597849146		Helvin3	Helvin3	https://www	https://medi	Auckland	No	3		Ask for a roo	This place co	0		2018-07-31	2018-07-20	2018-07-20	en	Trung Nguye	601995	https://www
581745954		Kerryn H	kerrynh2015	https://www	https://medi	Mount Dune	No	4		Great stay o	We stayed 2	0		2018-05-31	2018-05-21	2018-05-21	en	Trung Nguye	601995	https://www
580786227		anandp2011	anandp2011	https://www	https://medi	Bengaluru	No	4	Cleanliness:	Good locatio	Super friendl	0		2018-05-31	2018-05-17	2018-05-17	en	Trung Nguye	601995	https://www
540285968		None of You	noneofy94	https://www	https://medi	Canada	No	4		Exceptional	A bit wary ab	0	https://medi	2017-11-30	2017-11-11	2017-11-11	en	Trung Nguye	601995	https://www
486356079		50600ejea	50600ejea	https://www	https://medi	Suzhou	No	4		Good value	The rooms a	1		2017-05-31	2017-05-22	2017-05-22	en	Trung Nguye	601995	https://www
485589900		Franzi K	655franzik	https://www	https://medi	Germany	No	4	Rooms: 5Ser	Centrally loc	I usually stay	2		2017-04-30	2017-05-19	2017-05-19	en	Trung Nguye	601995	https://www
464024729		andy w	andyw550	https://www	https://medi	Candolim	No	4	Location: 4Cl	Clean and co	Nice, clean h	1		2017-03-31	2017-03-02	2017-03-02	en	Trung Nguye	601995	https://www
456101220		oggse	oggse	https://www	https://medi	Swieqi	No	4	Location: 5Cl	noisy but oka	I stayed in th	1	https://medi	2017-01-31	2017-01-31	2017-01-31	en	Trung Nguye	601995	https://www
452614461		Cinamoon	Cinamoon	https://www	https://medi		No	4	Cleanliness:	Clean and gr	We just stay	2		2017-01-31	2017-01-17	2017-01-17	en	Trung Nguye	601995	https://www
448145087		whoisthatgu	whoisthatgu	https://www	https://medi	Seoul	No	3	Location: 5Cl	Clean and sn	Great locatic	1		2016-11-30	2016-12-31	2016-12-31	en	Trung Nguye	601995	https://www
445519601		Rosa Ester S	r0saesters	https://www	https://medi		No	3		Terrible staff	The hotel wa	2		2016-12-31	2016-12-20	2016-12-19	en	Trung Nguye	601995	https://www
444350075		mige1988	mige1988	https://www	https://medi		No	3		Hotel was ok	We stayed h	0		2016-12-31	2016-12-14	2016-12-14	en	Trung Nguye	601995	https://www
437242855		fojane	fojane	https://www	https://medi	Seoul	No	3		Good locatio	This is a goo	0		2016-11-30	2016-11-14	2016-11-14	en	Trung Nguye	601995	https://www
435988515		loek_de_bee	loek_de_bee	https://www	https://medi		No	4		Decent place	This place is	0		2016-11-30	2016-11-09	2016-11-09	en	Trung Nguye	601995	https://www
434129266		Clanerock	Clanerock	https://www	https://medi	Sudbury	No	3	Cleanliness:	Great locatic	Great bathro	0		2016-11-30	2016-11-02	2016-11-02	en	Trung Nguye	601995	https://www
419969838		Son_H_Tran	Son_H_Tran	https://www	https://medi	Ho Chi Minh	No	4		Budget hotel	The hotel ha	0	https://medi	2016-09-30	2016-09-18	2016-09-18	en	Trung Nguye	601995	https://www
403414911		annatimmer	annatimmer	https://www	https://medi	Rotterdam	No	1		Terrible serv	The location	0		2016-08-31	2016-08-09	2016-08-09	en	Trung Nguye	601995	https://www

Figure 3. Reviews of customers on TripAdvisor

A total of 520 reviews were collected from 14 Trung Nguyen Legend chains in Vietnam, Singapore, Canada, and Malaysia. The longest review is 1008 words, the shortest is 1 word, and the average sentence length is 82.15.

When calculating the length of reviews and ratings, we see that the proportion of long sentences often appears in 1-3 star reviews. Figure 4 shows that dissatisfied customers tend to write longer sentences to complain about the quality and service of Trung Nguyen Legend.

Figure 4. Statistics on sentence length in reviews from 1 - 5 5-star ratings.

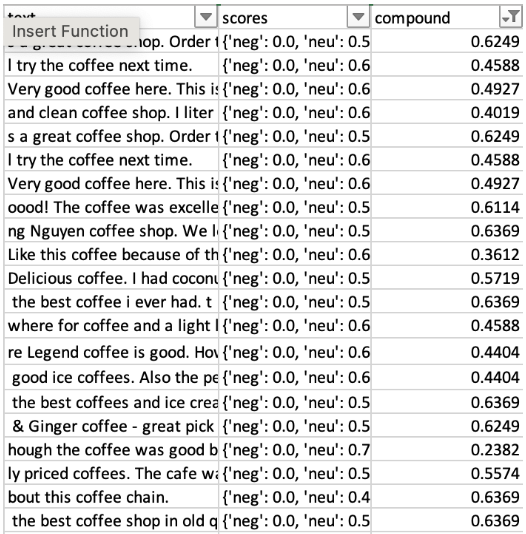
A word cloud visually represents the frequency of words within a given text or dataset. Typically, the more frequently a word appears in the text, the larger and more prominent it appears in the word cloud. It's a handy tool for quickly identifying common themes or topics within a body of text. Figure 5 below shows the word cloud from reviews.

Figure 5. Word Cloud of word frequency in reviews.

Upon analyzing Figure 5, it is evident that the words coffee, cafe, place, service, and great are the most commonly used terms in the reviews for Trung Nguyen Legend. This suggests that customers are highly satisfied with the quality of coffee and service provided by the cafe. It also indicates that the overall ambiance of the cafe is pleasant, and customers find it to be a great place to relax and enjoy a cup of coffee. Thus, based on this preliminary assessment, Trung Nguyen Legend appears to have a strong reputation for providing good coffee and service to its customers.

4.2 Measuring customer satisfaction with aspects

In Python programming, the Vader library provides tools to measure the score of sentiment in sentences. This score is to parse positive, negative, and neutral scores and compound scores. Figure 6 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



text	scores	compound
[illegible]op. Order t	{'neg': 0.0, 'neu': 0.5	0.6249
I try the coffee next time.	{'neg': 0.0, 'neu': 0.6	0.4588
Very good coffee here. This is	{'neg': 0.0, 'neu': 0.6	0.4927
and clean coffee shop. I liter	{'neg': 0.0, 'neu': 0.6	0.4019
s a great coffee shop. Order t	{'neg': 0.0, 'neu': 0.5	0.6249
I try the coffee next time.	{'neg': 0.0, 'neu': 0.6	0.4588
Very good coffee here. This is	{'neg': 0.0, 'neu': 0.6	0.4927
oood! The coffee was excelle	{'neg': 0.0, 'neu': 0.5	0.6114
ng Nguyen coffee shop. We l	{'neg': 0.0, 'neu': 0.5	0.6369
Like this coffee because of th	{'neg': 0.0, 'neu': 0.6	0.3612
Delicious coffee. I had cocont	{'neg': 0.0, 'neu': 0.5	0.5719
the best coffee i ever had. t	{'neg': 0.0, 'neu': 0.5	0.6369
where for coffee and a light l	{'neg': 0.0, 'neu': 0.6	0.4588
re Legend coffee is good. Hov	{'neg': 0.0, 'neu': 0.6	0.4404
good ice coffees. Also the pe	{'neg': 0.0, 'neu': 0.6	0.4404
the best coffees and ice crea	{'neg': 0.0, 'neu': 0.5	0.6369
& Ginger coffee - great pick	{'neg': 0.0, 'neu': 0.5	0.6249
hough the coffee was good b	{'neg': 0.0, 'neu': 0.7	0.2382
ly priced coffees. The cafe wa	{'neg': 0.0, 'neu': 0.5	0.5574
bout this coffee chain.	{'neg': 0.0, 'neu': 0.4	0.6369
the best coffee shop in old q	{'neg': 0.0, 'neu': 0.5	0.6369

Figure 6. Scores of reviews

The substring will be identified by the algorithms (2). Then, we used Python to calculate their scores. Figure 5 below shows the score of each substring for each aspect. We use formulas (1), (2), and (3) to calculate the satisfaction score for each aspect. Figure 7 shows the word cloud of positive reviews and negative reviews.

Figure 7. Word Cloud for positive and negative reviews.

Based on the compound score, we classify it into three categories: satisfied, dissatisfied, and neutral. After filtering, some reviews do not mention this aspect, and some are neutral. Specifically, the statistics of 520 reviews of the Trung Nguyen Legend store chain are as follows:

- 52 reviews did not include coffee
- 127 reviews were dissatisfied
- 341 reviews were satisfied

There are no neutral reviews for the coffee aspect. Next, proceed with other aspects such as Service, Location, Price, and Staff. The number of reviews for each class is displayed in Table 3 below.

Table 3. Classify customers' reviews into satisfaction, dissatisfaction, and neutral for each aspect.

Aspect	# positive substrings	# negative substrings	# substrings was not contained this aspect	# neutral substring	# substrings mentioned this apspect
Coffee	341	127	52	0	468
Service	220	145	143	12	377
Location	277	149	87	7	433
Price	144	62	314	0	206
Staff	282	129	86	23	434

From the data in Table 3, it is possible to calculate customer satisfaction or dissatisfaction according to each aspect by formula (1) and (2). After excluding reviews that do not contain the aspect and reviews with neutral values, the study only retained reviews that contained aspects of interest to measure. Figure 8 below shows customer satisfaction and dissatisfaction with Trung Nguyen Legend in each specific aspect.

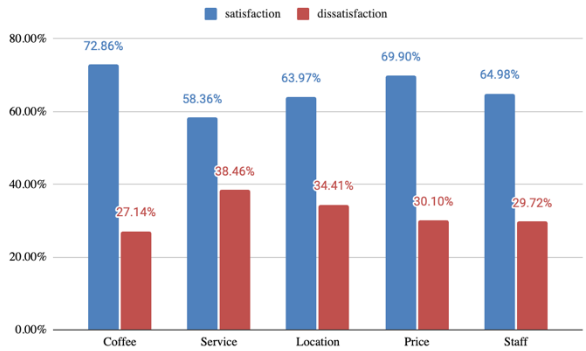


Figure 8. The ratio of satisfaction and dissatisfaction for each aspect.

The highest satisfaction is in the aspect of coffee, with a score of 72.86%, followed by the aspect of price, which has a value of 69.90%. The lowest is the service aspect, which only reaches 58.36%. The service aspect also has the highest dissatisfaction score, reaching 38.46%, while the coffee aspect has only 27.14% of dissatisfied customers.

4.3 Measuring the NPS score of customers with Trung Nguyen Legend

The study conducted an analysis of the NPS (Net Promoter Score) at both the aspect and overall level. To obtain the value of the NPS score at the aspect level, the formula (2) was applied. The formula measured the NPS score for each aspect under consideration. The NPS score for each aspect is presented in Figure 6 below, which provides an insightful overview of the performance of different aspects. The study results provide an in-depth analysis of customer feedback and can help identify areas that require improvement to enhance overall customer satisfaction.

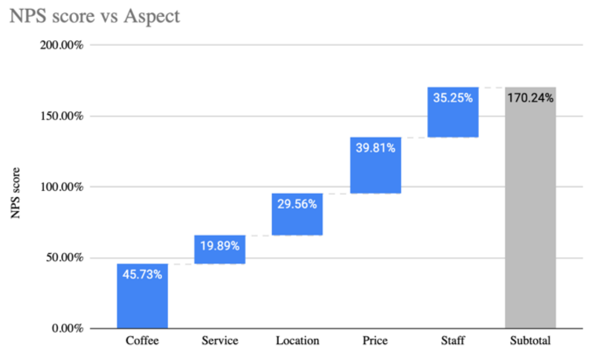


Figure 9. The NPS score for each aspect of the Trung Nguyen Legend.

The NPS score for each aspect represents customer satisfaction and loyalty scores for each aspect. There are five dimensions measured in this study. The NPS score for the coffee aspect reached the highest at 45.73%, followed by price at 39.81%. Service achieved the lowest value with an NPS score of 19.89%. NPS scores below 30% include Service and Location. This shows that these two aspects are rated lowest as needing improvement in the Trung Nguyen Legend chain. Then, we measure the NPS score for all aspects of Trung Nguyen Legend. The study took the overall average value as formula (3). Figure 7. Below are the results for the overall NPS score.

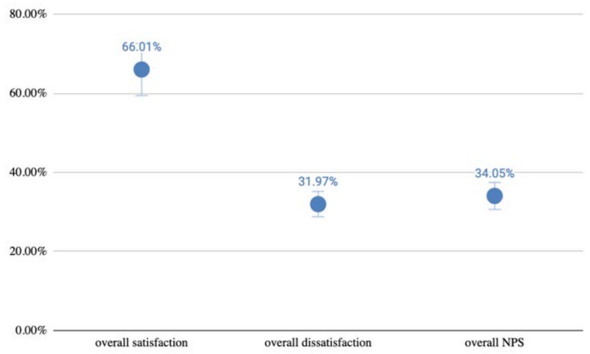


Figure 10. The overall satisfaction, dissatisfaction, and NPS score of Trung Nguyen Legend.

Based on the data presented in Figure 10, the overall satisfaction rate is 66.01%. This means that 66.01% of customers are satisfied with Trung Nguyen Legend. On the other hand, the overall dissatisfaction rate is 31.97%, which means that 31.97% of customers are not satisfied with Trung Nguyen Legend. The NPS score is 34.05%, which indicates that Trung Nguyen Legend is doing great and has more happy customers than unhappy ones.

5. Discussion

The success and longevity of businesses in the Food and Beverage (F&B) industry greatly depend on the quality of customer experience. Trung Nguyen Legend offers its customers an enjoyable and engaging retail experience at its coffee shops and retail outlets [17], [18]. The stores' ambiance, decor, and layout are designed to create a welcoming and comfortable atmosphere where customers can unwind and savor their coffee [14], [16]. Trung Nguyen Legend celebrates Vietnam's rich coffee culture and heritage by incorporating traditional brewing techniques and methods into its products [42]. The brand's commitment to preserving and honoring Vietnamese coffee traditions resonates with customers who value authenticity and cultural significance. When penetrating domestic and international markets, unique meanings are also brought to Trung Nguyen Legend's mission and vision [41].

While collecting data for research, the research team noticed that the number of stores on the TripAdvisor platform is not much, but it is large on the Google Maps platform. Similar to the Cong coffee brand. This shows that brand presence on Google Maps is given priority over TripAdvisor. Meanwhile, Starbucks is fully visible on both Google Maps and TripAdvisor platforms. According to statistics, the number of Starbucks stores is currently about 38,587 worldwide, of which 20,656 are located outside North America.

Trung Nguyen Legend recently opened its second coffee store in Shanghai and the first store in the US, and the upcoming goal is to open 1,000 coffee stores in China. Currently, five Vietnamese F&B chains are opening stores abroad: Cong Ca Phe, Phuc Long, King Coffee, Highlands Coffee, and E-Coffee. Cong Coffee has 16 stores in Korea and Canada. Phuc Long, a tea and coffee chain, has opened its first store in the US. The competition between coffee shop chains not only domestically but also internationally. In the study, it was clear that Trung Nguyen Legend's current points that need improvement include: Regarding satisfaction in each aspect, coffee and related beverages have the highest satisfaction, reaching nearly 80%. This is followed by price and staff, determined to be reasonable prices and friendly staff. The service part includes accompanying services and space, which seems unreasonable because the layout does not create a good experience for customers in this aspect. The atmosphere, as well as cigarette smoke, also affects the customer's bad experience.

Similarly, the NPS score also shows the highest customer satisfaction for the coffee aspect. Next is price and staff. The service aspect achieved the lowest NPS score.

In the overall satisfaction state, overall process satisfaction reached 66.01%, and dissatisfaction reached 31.97%

(~40%). The overall satisfaction value is 34% when measuring the overall NPS score. These are some not-so-high scores. Therefore, Trung Nguyen Legend still needs to improve, especially in terms of service quality.

This study shows that it is possible to understand customer experience in many ways just by analyzing customer reviews on the Internet platform. Good customer feedback is an effective word-of-mouth marketing method. However, it has been famous for many years in the international market, products under the Trung Nguyen brand mainly export coffee. The focus on chain stores creating a legendary Vietnamese coffee experience for customers has just begun to develop abroad in the past few years in the context of a fiercely competitive retail market. This requires creativity in the style and quality of services and products to promote sustainable development and engage customers.

6. Conclusion

This study presents a method for understanding customer experience at the Trung Nguyen Legend store chain, based on customer reviews collected from the TripAdvisor platform and analyzed using natural language processing. Python and natural language processing libraries and algorithms are combined to measure customer satisfaction scores by each dimension and overall NPS score.

The research results underscore the potential benefits of improving service quality for Trung Nguyen Legend in a domestic and international competitive market. This improvement can create a good customer experience space, fostering customer engagement through other added values. Moreover, the brand's presence on various internet platforms can significantly increase customer touchpoints, further enhancing the Trung Nguyen Legend brand.

To continue our future direction and measure competition in the coffee market in Vietnam. In future studies, we will compare customer experiences with those of famous brands in Vietnam. From there, we will evaluate the competition in the Vietnamese coffee market and the opportunities for the brands in Vietnam.

References

[1] E. (olivia) Park, W.-H. Kim, and J. Kwon, “Understanding the relationship between green restaurant certification programs and a green restaurant image: the case of TripAdvisor reviews,” *Kybernetes*, vol. 50, no. 6, pp. 1689–1703, 2021.

[2] B. Ozyurt and M. A. Akcayol, “A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA,” *Expert Syst. Appl.*, vol. 168, no. 114231, p. 114231, 2021.

[3] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, “Aspect-based sentiment analysis: A survey of deep learning methods,” *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 6, pp. 1358–1375, 2020.

[4] N. Nikolić, O. Grljević, and A. Kovačević, “Aspect-based sentiment analysis of reviews in the domain of higher education,” *Electron. Libr.*, vol. 38, no. 1, pp. 44–64, 2020.

[5] G. Zhao, Y. Luo, Q. Chen, and X. Qian, “Aspect-based sentiment analysis via multitask learning for online reviews,” *Knowl. Based Syst.*, vol. 264, no. 110326, p. 110326, 2023.

[6] M. M. Trușcă and F. Frasincar, “Survey on aspect detection for aspect-based sentiment analysis,” *Artif. Intell. Rev.*, vol. 56, no. 5, pp. 3797–3846, 2023.

[7] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, “Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions,” *Heliyon*, vol. 10, no. 9, p. e27863, 2024.

[8] A. Mawadati, W. Ustyannie, A. Hindarto Wibowo, and R. Adelina Simanjuntak, “Analysis of Yogyakarta coffee shop visitor reviews to increase customer satisfaction using sentiment analysis,” *KnE Soc. Sci.*, 2024.

[9] D. Suryadi and J. T. Sabarman, “Analyzing restaurants in tourism destinations through online reviews using topic modeling and sentiment analysis,” in *2023 IEEE 9th Information Technology International Seminar (ITIS)*, 2023.

[10] I. Burkov, A. Gorgadze, and I. Trabskaia, “Satisfaction dimensions influencing consumers’ behavioral intentions through structural topic modeling analysis of restaurant reviews,” *Consumer Behavior in Tourism and Hospitality*, vol. 18, no. 2, pp. 200–214, 2023.

[11] N. Pleerux and A. Nardkulpat, “Sentiment analysis of restaurant customer satisfaction during COVID-19 pandemic in Pattaya, Thailand,” *Heliyon*, vol. 9, no. 11, p. e22193, 2023.

[12] H. T. T. Nguyen and T. X. Nguyen, “Understanding customer experience with Vietnamese hotels by analyzing online reviews,” *Humanit. Soc. Sci. Commun.*, vol. 10, no. 1, 2023.

[13] O. T. K. Vu *et al.*, “Coffee culture unravelled: exploring the coffee shop experience model in the Vietnamese context,” *Tourism Recreation Research*, pp. 1–19, 2023.

[14] T. Nguyen, D. T. Tuyen, N. N. Trung, and T. Cuong, “Analysing costs and margins of smallholder farmers in the coffee value chain: M4P approach,” *International Journal of Economics and Business Research*, vol. 26, no. 3, pp. 325–353, 2023.

[15] T. D. Vo, L. Yang, and M. Tran, “Determinants influencing Vietnam coffee exports,” *Cogent Business & Management*, vol. 11, no. 1, 2024.

[16] O. T. K. Vu *et al.*, “Coffee and gastronomy: a potential ‘marriage’? The case of Vietnam,” *Int. J. Contemp. Hosp. Manag.*, 2022.

[17] T. T. H. Nguyen, “Factors Influencing the Performance of the Coffee Supply Chain-The Case Study of Vietnam,” in *World Conference on Information*

Systems for Business Management, Singapore; Singapore: Springer Nature, 2023, pp. 167–177.

[18] L. Ngoc *et al.*, "Machine Learning for Agricultural Price Prediction: A Case of Coffee Commodity in Vietnam Market," in *2023 IEEE/ACIS 8th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, IEEE, 2023, pp. 38–41.

[19] J. F. Sánchez-Rada and C. A. Iglesias, "Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison," *Inf. Fusion*, vol. 52, pp. 344–356, 2019.

[20] N. Hajli, M. Tajvidi, A. Gbadamosi, and W. Nadeem, "Understanding market agility for new product success with big data analytics," *Ind. Mark. Manag.*, vol. 86, pp. 135–143, 2020.

[21] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," *Int. J. Res. Mark.*, vol. 40, no. 1, pp. 75–87, 2023.

[22] I. Chakraborty, M. Kim, and K. Sudhir, "Attribute sentiment scoring with online text reviews: Accounting for language structure and missing attributes," *J. Mark. Res.*, vol. 59, no. 3, pp. 600–622, 2022.

[23] M. Z. Asghar *et al.*, "Senti-eSystem: A sentiment-based eSystem-using hybridized fuzzy and deep neural network for measuring customer satisfaction," *Softw. Pract. Exp.*, vol. 51, no. 3, pp. 571–594, 2021.

[24] J. Park, "Framework for sentiment-driven evaluation of customer satisfaction with cosmetics brands," *IEEE Access*, vol. 8, pp. 98526–98538, 2020.

[25] M. N. Imtiaz and M. K. Ben Islam, "Identifying significance of product features on customer satisfaction recognizing public sentiment polarity: Analysis of smart phone industry using machine-learning approaches," *Appl. Artif. Intell.*, vol. 34, no. 11, pp. 832–848, 2020.

[26] S. Yi and X. Liu, "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review," *Complex Intell. Syst.*, vol. 6, no. 3, pp. 621–634, 2020.

[27] R. Annamalai, S. A. Rasool, Deena, K. Venkatraman, and Soundaram, "Sentiment analysis using VADER: Unveiling customer sentiment and predicting buying behavior in the airline industry," in *2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)*, 2024.

[28] K. Barik and S. Misra, "Analysis of customer reviews with an improved VADER lexicon classifier," *J. Big Data*, vol. 11, no. 1, 2024.

[29] V. Vought, R. Vought, A. S. Lee, I. Zhou, M. Garneni, and S. A. Greenstein, "Application of sentiment and word frequency analysis of physician review sites to evaluate refractive surgery care," *Adv. Ophthalmol. Pract. Res.*, vol. 4, no. 2, pp. 78–83, 2024.

[30] H. Ahn and E. Park, "The impact of consumers' sustainable electronic-word-of-mouth in purchasing sustainable mobility: An analysis from online review comments of e-commerce," *Res. Transp. Bus. Manag.*, vol. 52, no. 101086, p. 101086, 2024.

[31] G. Agag, B. Ali Durrani, Z. Hassan Abdelmoety, M. Mostafa Daher, and R. Eid, "Understanding the link between net promoter score and e-WOM behaviour on social media: The role of national culture," *J. Bus. Res.*, vol. 170, no. 114303, p. 114303, 2024.

[32] L. J. Callarisa-Fiol, M. Á. Moliner-Tena, R. Rodríguez-Artola, and J. Sánchez-García, "Entrepreneurship innovation using social robots in tourism: a social listening study," *Rev. Manag. Sci.*, vol. 17, no. 8, pp. 2945–2971, 2023.

[33] P. Balakrishnan, C. Kroiss, T. Keskes, and B. Friedrich, "Perception and use of reversible contraceptive methods in Germany: A social listening analysis," *Womens. Health (Lond. Engl.)*, vol. 19, p. 174550572211473, 2023.

[34] E. Zarrabeitia-Bilbao, M. Jaca-Madariaga, R. M. Rio-Belver, and I. Álvarez-Meaza, "Nuclear energy: Twitter data mining for social listening analysis," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, p. 29, 2023.

[35] R. Sivakumar and C. Ganeshkumar, "Enhancement of touchpoint interactions on customer experience management in the banking sector," *Int. J. Electron. Cust. Relatsh. Manage.*, vol. 13, no. 4, p. 314, 2022.

[36] H. Daou, "Sentiment of the public: the role of social media in revealing important events," *Online Inf. Rev.*, vol. 45, no. 1, pp. 157–173, 2020.

[37] H. N. Thi Thu, T. D. Thanh, T. N. Hai, and V. H. Ngoc, "Building Vietnamese topic modeling based on core terms and applying in text classification," in *2015 Fifth International Conference on Communication Systems and Network Technologies*, 2015.

[38] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, 2014.

[39] M. Wang and F. Hu, "The application of NLTK library for Python natural language processing in corpus research," *Theory Pr. Lang. Stud.*, vol. 11, no. 9, pp. 1041–1049, 2021.

[40] M. Gómez-Suárez and M. Veloso, "Designing Facebook publications focused on hotel customer experience: How to improve brand attitude and booking intention," in *Brand, Label, and Product Intelligence*, Cham: Springer International Publishing, 2022, pp. 247–258.

[41] M. Intelligence, "Vietnam coffee market - share, report - industry size." .

[42] O. Schulte, J. Mumber, and T. T. Nguyen, "Agricultural commercialisation, asset growth and poverty in rural Vietnam," *Australian Journal of Agricultural and Resource Economics*, vol. 67, no. 3, pp. 388–416, 2023.

.