Table of Contents:

Journal of Environmental Management and Tourism is an interdisciplinary research journal, aimed to publish articles and original research papers that should contribute to the development of both experimental and theoretical nature in the field of Environmental Management and Tourism Sciences.

Journal will publish original research and seeks to cover a wide range of topics regarding environmental management and engineering, environmental management and health, environmental chemistry, environmental protection technologies (water, air, soil), pollution reduction at source and waste minimization, energy and environment, modeling, simulation and optimization for environmental protection; environmental biotechnology, environmental education and sustainable development, environmental strategies and policies, etc. This topic may include the fields indicated above, but are not limited to these.

Authors are encouraged to submit high quality, original works that discuss the latest developments in environmental management research and application with the certain scope to share experiences and research findings and to stimulate more ideas and useful insights regarding current best-practices and future directions in environmental management.

Journal of Environmental Management and Tourism is indexed in SCOPUS, RePEC, CEEOL, ProQuest, EBSCO and Cabell Directory databases.

All the papers will be first considered by the Editors for general relevance, originality and significance. If accepted for review, papers will then be subject to double blind peer review.

Deadline for submission:	31st July 2022
Expected publication date:	September 2022
Website:	https://journals.aserspublishing.eu/jemt
E-mail:	jemt@aserspublishing.eu

To prepare your paper for submission, please see full author guidelines in the following file: JEMT_Full_Paper_Template.docx, then send it via email at jemt@aserspublishing.eu.

Tourist Attraction Satisfaction Factors from Online Reviews. A Case Study of Tourist Attractions in Thailand

Vimolboon CHERAPANUKORN
College of Arts, Media and Technology
Chiang Mai University, Thailand
vimolboon.c@cmu.ac.th

Prompong SUGUNNASIL
College of Arts, Media and Technology
Chiang Mai University, Thailand
prompong.sugunnasil@cmu.ac.th

Abstract:

In order to survive and gain competitive advantages in the post COVID-19 pandemic, tourism destinations should plan their business strategy by focusing on customer expectation. A number of research have studied tourist satisfaction, particularly, hotels and transports, however there is limited investigation with tourist attractions which have different prominence from other tourism service providers. The purpose of this study was to identify the tourists' satisfaction components for tourist attractions by adopting an opinion mining technique and using the zero-shot text classification method. The total of 40,000 online tourists' reviews from 40 tourist attractions in Thailand, that were posted up thought TripAdivisor.com between 2010 and 2021, were analyzed. The research findings reveal six components of tourist attraction satisfaction (TATSAT) model that includes 1) ambiance, 2) hospitality, 3) price, 4) accessibility, 5) cleanliness, and 6) security. All attributes of TATSAT model are generated from tourists' point of view and was analyzed by the focus group discussion with five tourism experts from both academics and practitioners. This model expands the idea of HOLSAT and SERVQUAL by focusing on the tourist attraction business sector. The results can serve academics and practitioners in the research and improvement of tourist satisfaction to maximize competitive advantages for tourist attraction sector in the future.

Keywords: tourist attraction; tourist satisfaction; zero-shot classification; opinion mining; online review; Thailand.

JEL Classification: L83; M30; Z30; Z31.

Introduction

Tourism contributes to the economies of many countries in both direct and indirect ways, and it has become one of the world's most important economic activities. Tourist attractions have been noticed as a key role to the successful tourism development of many countries (Pearce 1998). In early 2020, the outbreak of the Coronavirus pandemic has strongly impacted the tourism industry, resulting in the largest world economic crisis in the last 10 years (UNWTO, 2021). Tourist attractions have been challenged to improve their service quality attract all travelers for the post COVID-19 situation due to the new normal travelling behavior (Lew *et al.* 2020). Moreover, tourism marketers have been concerned with increasing tourist satisfaction levels as focusing on attracting new customers alone is insufficient, an increasing customer retention is also importance. Also, the major challenges

for a tourist attraction are fostering the implementation of innovative and successful products and developing the ability to provide all visitors a memorable experience, which can lead to customer satisfaction and loyalty (Baggio and Caporarello 2005).

Most tourists, nowadays, employ social media platforms as an important channel for sharing their previous trips experiences. To the best knowledge of the authors, there is limited empirical study to date has examined the tourist satisfaction factor for tourism attraction sector, especially towards the online tourist comments. Therefore, this research aims to identify the main tourists' satisfaction components for tourist attractions from online reviews by adopting an opinion mining with zero-shot text classification technique and using Thailand, one of the world's most visited destinations, as a case study.

1. Literature Review

A tourism attraction, Pearce (1998) is defined as a named place with a specific feature that is the consider of tourist attention and management; includes natural features, cultural and history features and commercial man-made purposed features. Baggio (2011) outlined that it is a place of interest that tourists visit, typically due to its inherent or exhibited natural or cultural value, historical status, or built beauty, which offer leisure, adventure, and amusement services. Tourist attractions are generally individual sites or single units that are small, while destinations are larger areas that include a variety of individual attractions together with the support services required by visitors (Swarbrooke and Page 2012). Tourist Attraction is one of the important elements of tourism sector which motivate and encourage people to visit the destination (Crouch and Ritchie 1999) and develop the successfulness of tourism destination (Hu and Wall 2005). Ritchie *et al.* (2003) divided the attractions into five categories based on the main proposed of their activities; natural, cultural, events, recreation and entertainment.

1.1. Tourist Satisfaction and Tourist Attraction Attributes

In tourism sector, customer satisfaction is referred to a consumer's attitude and emotional response to the entire intangible service at the stage of post-purchasing, which occur when visitors experience the service and compare the encounter with their personal expectations (Li *et al.* 2013). Tourists visit a destination for consuming and experiencing the services that are offered, to create great memories and to interact with the service providers or the resident hosts (McDowell 2010). Also, most tourists will form their attitudes toward a destination by comparing their actual experiences with their expectations based on their past experiences, relatives, friends, marketing promotions, and companies' promises (McDowell 2010), which leads to their satisfaction or dissatisfaction. Chenini and Touaiti (2018) explained that tourist satisfaction is a significant focus in terms of maintaining the tourism industry as it affects individuals' perceptions when selecting where to go and what to purchase.

Furthermore, there are different satisfaction factors that have been identified in studies on service sectors that include tourism businesses. SERVQUAL model by Parasuraman *et al.* (1985) is a distinctive model for studying the customer expectations and perceptions of a service (Kozak & Rimmington, 2000). SERVQUAL includes five constructs that are believed to represent the quality of a service: reliability, assurance, tangibles, empathy, and responsiveness (Parasuraman *et al.* 1985). Also, this service quality measurement model mainly focuses on service skills and the quality of the staff within an organization. Hence, this instrument is not enough to measure the level of satisfaction in the tourism sector (Ceylan and Ozcelik 2016). Also, Tribe and Snaith (1998) presented the holiday satisfaction model or HOLSAT, which they used to evaluate holiday satisfaction at the famous resort in Varadero, Cuba by expanding the ideas from SERVQUAL. Ceylan and Ozcelik (2016) grouped the components of HOLSAT theoretical model into 7 constructs that includes the constructs of accommodation, ambiance, physical facilities, restaurants, bars, shops and nightlife, transfers, heritages and culture, responsiveness, empathy and tangible. According to Alegre and Garau (2010), HOLSAT is proposed from tourists' points of view by focusing on tourists' expression regarding their satisfaction and dissatisfaction and evaluating both positive and negative attributes, which conveys favorable and unfavorable impressions of tourists' experiences of a destination.

The determinations of satisfaction for tourism destinations have been debated among tourism literatures because it is determined by a variety of factors (Bernini and Cagnone 2014). Most of the previous scholars focus on the destination attributes as antecedents of overall tourist satisfaction (Shahrivar 2012). Pizam *et al.* (1978) mentioned seven factors of tourist satisfaction with a national destination, which are derived as beach, cost, hospitality, restaurant and facilities, accommodation, environment and commercialization. According to Shahrivar (2012), the destination attributes, that influence on tourist satisfaction, are divided into 8 attributes; 1) natural factors, 2) cultural factors, 3) shopping, 4) accessibility, 5) infrastructure, 6) reception, 7) services, and 8) cost and prices. While, there are other related scholars evaluate destination attributes that are considered relevant to

tourist satisfaction as a quality of accommodation/restaurant, price of the services, signpost/information about the destination, traffic congestion/ease to access, Personal safety, culture/historical/natural attractions, cleanliness of destination, hospitality of staffs/local hosts, parking, entertainments/activities and infrastructure (Kozak and Rimmington 2000; McDowell 2010; Marin and Taberner 2008; Eusébio and Vieira 2013; Bernini and Cagnone, 2014; Della Corte *et al.* 2015; Ceyland & Ozcelik, 2016). However, Albayrak and Caber (2016) noted that most previous tourist satisfaction models aim to measure and analysis tourists' satisfaction at a tourism destination, as being the locations (cities, region, countries).

1.2. Opinion Classification for the Tourism Sector

An opinion represents the individual's judgments, attitudes, beliefs, experience and evaluations about a topic, subject and item (Khan *et al.* 2014). Customer opinion is very useful for the organization because it creates a huge impact on human's decision-making process which can be collected from verities sources, such as face-to-face interviewing, printed documents, television and Internet (Khan *et al.* 2018). With the development of Internet, it allows all users to create, explore, and participate in information through the online platform, which develops a new electronic word of mouth. According to Verma and Yadav (2001), eWOM from customers' online reviews shows a powerful of User Generated Content (UCG), and this has been recognized as one of the strongest channels for forming both positive and negative consumer attitudes toward products and services. Many online reviews in a social science and business researches apply the online content and customers' opinion review for evaluating and proposing business strategy (Cherapanukorn and Charoenkwan 2017).

The term of Opinion mining is a text mining technical exercise that is related to big data analysis, which focus on mining knowledge and information from user-generated content mainly for extracting, analyzing and evaluating and rating users' opinion (Khan *et al.* 2014). Commonly, there are three supervised machine learning algorithms techniques including; Naive Bayes, a support vector machine (SVM), and the character-based N-gram model, that have mostly been applied to classify the sentiment of online reviews (Schuckert *et al.* 2015). Research on opinion mining is challenged with some limitations. Firstly, Zhang *et al.* (2011) claimed that a precision in text mining is still limited due to the sentiment classification divides textual contents into positive and negative comments is only around 70% accurate. They also explained that the technique's ability to generate and group the factual meaning of reviews is still not intelligent enough. Additionally, the text mining process only returns a collection of keywords, which is not enough for providing any clear analyzed results (Zhang *et al.* 2011). Secondly, tourism is a global business sector, and customers' opinion are expressed in various linguistic structures based on their nationality. Therefore, it is difficult for opinion mining programs to analysis different languages (Khan *et al.* 2018). Thirdly, there is a lack of applications and motivations.

However, the purposes of opinion mining in nowadays aims to explore customers' e-complaints, satisfaction level and experiences from customer feedback (Lee and Hu 2005) and to mine word patterns using the semantic clustering of consumer opinions from a large amount of qualitative data retrieved from online platforms (Capriello *et al.* 2013). Opinion mining techniques and sentiment classification have been widely applied for successfully segmenting and targeting customers across various business sectors. A related study that applied data mining to improve service quality was shown in the research by Emel *et al.* (2007). The results showed that the better a manager of a destination marketing organization (DMO) understands traveler profiles and traveling patterns, the better he/she can market his/her destination. Sahadev and Islam (2004) also provide an effective approach to discovering and understanding patterns in customer behavior in order to help businesses make decisions regarding their customers' feedback. Additionally, Lui *et al.* (2020) adopted data mining techniques to propose conceptual intelligent recommendations for tourist attractions based on the time restrictions and cost that related to travel.

Opinion mining studies have been interested in Zero-shot learning (ZSL) in recent years. ZSL is capable of transferring knowledge from seen classes to unseen classes (Ye *et al.* 2020). Extensive research has applied ZSL to image classification tasks (Del Chiaro *et al.* 2019; Ye and Guo 2017). Most researches in ZSL have been studied widely in image classification (Ye *et al.* 2020). For zero-short text classification, it has been practiced by embedding text and labels into a joint space while sharing word representations between the text and labels (Sappadla *et al.* 2016). Even though more ZSL approaches to text classification have been developed recently, studies that focus on ZSL in tourism and hospitality sectors are still limited. Furthermore, most of the initial research projects were focused on accommodation sector and that only a few studies have looked at tourist destinations. Description and justification of the research methods used. Normally, the methods will be selected

from known and proven examples. In special cases the development of a method may be a key part of the research, but then this will have been described in Introduction section and reviewed in first one.

2. Methodology

In this section, the working strategy involves employing natural language processing (NLP) to analyses customers' sentiments toward Thailand's different tourist attraction sites, covering various categories and locations. The first step of this research requires the extraction of sentimental information from a tourism-related website. Next, the processes of analyzation and classification keywords based on their meaning are confirmed by the experts' focus group discussion. Then, the sentimental data are pre-processed using several techniques to discover the keywords that reflect the sentimental landscape of Thailand's tourism. Finally, the comments are classified based on the keywords and analyzed using the zero-shot classification technique.

2.1. Opinion Data Collection

This study collected data from tourists' online comments that have been posted through social media platform regarding Thai Tourist attractions, as a case for studying. A report from World Tourism Organization (UNWTO) in 2019, Thailand was listed in the top 10 of the world's most visited destinations. TripAdvisor.com is the data repository used, and the four well-known cities that are popular with tourists in Thailand (Bangkok, Chiang Mai, Chonburi, and Phuket) comprise the area of interest. The top 10 most popular attractions of each city, that were rated by TripAdvisor.com on 2nd February 2021, were selected to be the sample group, hence there are 40 attractions in total, including temples, natural attractions (national parks, beaches, botanic garden), villages, shopping malls, amusement theme parks, museums. For each attraction, 500-1,000 online comments are extracted, there are 40,000 comments in total, and all of them are written in the English language. The data were downloaded in February 2021 by using Beautyfulsoup4 which is a Python module as a web scraping, saved then converted into the excel file for analyzing. The extracted data includes the rating awarded to the site, the title of the comment, and the main text of the comment. All online comments collected were posted on TripAdvisor.com from January 2010 to January 2021.

2.2. Important Word Identification and Analysis

Vocabulary knowledge is a basic element required for understanding a language. Despite the fact that tourists from different countries use the website, English is still the most popular language. However, the English used in the tourism industry is different from general English 9Lam, 2007). As a consequence, a list of keywords must first be extracted from the dataset. This research uses two techniques to discover this word list: word frequency and word range. The Python WordNet lemmatize is applied using a natural language toolkit (Nltk) and through elimination of common words or stop words from the comments. Moreover, only nouns are adjectives are studied in this work, and the pre-processing process also converts all the words into lowercase. The Nltk is a popular python library for processing natural language. The aim of WordNet lemmatization is to morphologically change words to their general base forms using WordNet, a large well-known lexical semantic database. Unlike a word stemming process, lemmatization takes the context into consideration. This research uses the part of speech (POS) to assist the lemmatization process. The overall design is displayed in Figure 1.

Figure 1. Overall process of important word identification

1. remove stop words → 2. Apply lemmatization → 3. Analyze the structure of the comments

Source: (Own illustration)

The first technique uses the word frequency. A frequency analysis generates the number of occurrences of a word. The word frequency highlights the importance of the word itself and provides information about tourism. In other words, it provides information on the appearance and disappearance of certain words. Lindquist (2018) suggested that frequency is one of the most important criteria when studying language. However, data collection can be subjective. The resulting corpus can be biased toward certain possibilities, rather than representing the whole domain (Schmitt and Schmitt 2014). Therefore, frequency analysis alone is not sufficient. Thus, the most frequently used words are listed and grouped to form the proposed tourist attraction satisfaction

factors based on the previous related literature, both nouns and adjectives are collected from the dataset (Keng and Yu 2011; Puangyaem 2013). Then, the group of keywords in each factor is confirmed by the focus group discussion. The focus group session was conducted in March 2021 by inviting a total of five experts, three academics from the area of tourism and hospitality research and two practitioners from the tourism industry. After, this study performs opinion analysis using the zero-shot technique in order to classify the sentiments of the comments for each factor. Additionally, the research framework is illustrated in Figure 2.

Figure 2. Research framework
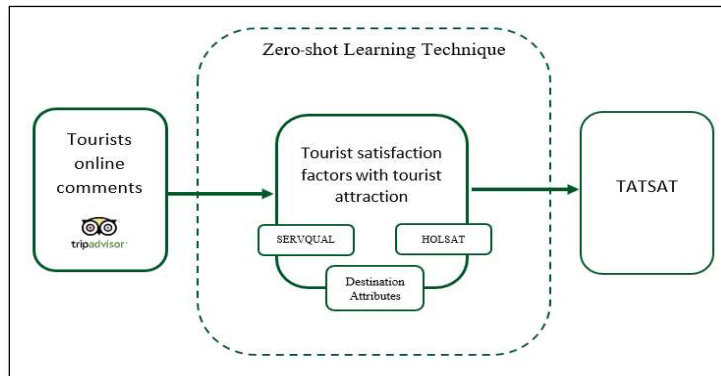
Source: (Own illustration)

3. Results and Analysis

The results are based on a total of 40,000 online comments from various popular attractions in the famous 40 tourist attractions in Thailand. Table 1 displays examples of words and their frequencies extracted from the comments. Only the first ten words with the highest frequencies are shown. Moreover, 27,501 unique nouns and 1,670 unique adjectives are identified from all the comments. However, common words, like he, she, they, is, am, are, a, an, and the, are removed from the data.

Table 1. The top 10 most frequently used words when referring to tourist attractions in Thailand

	Word	Frequency		Word	Frequency
1	place	19,767	6	gorgeous	12,687
2	beautiful	18,344	7	best	10,870
3	beauty	15,575	8	beach	10,170
4	great	14,824	9	temple	9,591
5	nice	14,227	10	price	8,667

Table 2. Examples of co-occurring nouns and adjectives.

Noun			Adjective		
beauty	place	food	gorgeous	beautiful	peaceful
money	temple	safety	terrible	cheap	historic
beach	hospitality	fee	calm	safer	crowded
price	market	access	young	attractive	clear
peace	service	water	harmful	memorable	aesthetic
employee	value	favor	personal	modern	comfortable
convenience	faith	time	professional	funny	aware
admire	charge	decoration	clean	safe	unsafe
joy	romance	faith	pricy	overpriced	costly
antique	people	photo	great	nice	best
pleasure	temple	police	polite	rude	near
staff	green	dream	dirty	unhealthy	photogenic
atmosphere	communication	Buddha	underwhelmed	walkable	favorite

Furthermore, this study uses lexical co-occurrence patterns analysis. Unlike frequency analysis, which focuses on a single corpus, word co-occurrence analysis focuses on the mutual existence of certain words of the

same domain in different collections. Additionally, this co-occurrence denotes the probability of the occurrence of words across different datasets. This concept can apply to both grammar and words. This approach can reduce the bias from the frequency analysis. In this research, words are selected based on the criteria that they must exist in at least ten attractions. As a result, the total keywords are identified as 4,647 nouns and 428 adjectives. Table 2 displays examples of co-occurring nouns, and adjectives from the dataset.

The keywords extracted can be classified into six preliminary categories according to their semantic similarities: ambiance, cleanliness, security, accessibility, price, and hospitality. Then, the focus group discussion was conducted with the experts in order to analyze and confirm the items of each factor in the tourist attraction satisfaction framework and define words for each category based on the list of co-occurring nouns and adjectives. The proposed satisfaction factors of tourist attractions are presented in Table 3 along with examples of keywords relating to each factor.

Table 3. Six dimensions of tourist attraction satisfaction (TATSAT) model and examples of keywords for each factor.

Dimension	Example keywords
Ambiance	beauty, beautiful, gorgeous, calm, peaceful, attractive, peace, atmosphere, decoration, romance, green, charming, decoration, terrible, memorable, comfortable, quiet, noisy, photogenic, cream.
Hospitality	staff, employee, people, nice, rude, honest, polite, professional, personal, communicate, service, hospitality, respond.
Price	price, cheap, pricy, overpriced, costly, expensive, inexpensive, affordable, value, money, charge fee, worth.
Accessibility	crowded, traffic, access, accessibility, near, far, time, transportation, comfortable, convenience, sign.
Cleanliness	clean, clear, cleanliness, healthy, unhealthy, health, hygiene, hygienic, dirty, waste.
Security	safe, safer, unsafe, safety, security, crowded, dangerous, danger, light, walkable, police, dark, aware.

The results present that the ambiance factor refers to keywords that are used to express the beauty of the attractions, overview of infrastructures/ facilities and the atmosphere or type of decoration in the locations. The hospitality factor includes keywords that reference the service quality and professional skills of staff at the attraction sites. It also refers to the hospitality of local people or residences in the surrounding areas. With regard to the price factor, it contains keywords relating to money and expenses or the entrance fees of the sites. It also includes keywords that express what results in value for money in terms of an attraction. The accessibility factor embodies keywords related to the ease of accessing each attraction, as well as the difficulty or simplicity of visiting or entrancing a place. In terms of the cleanliness factor, it covers keywords that refer to the hygiene of the attraction. Lastly, the security factor includes all keywords concerning safety or dangerous aspects of the sites.

Furthermore, the findings of all dimensions of the proposed model are supported and increased the previous tourist satisfaction models by considering narrow down to the context of tourist attraction business sector. In Table 4, it reveals the discussion results of each tourist attraction satisfaction factor by evaluating and analyzing with the previous tourist satisfaction models; SERVQUAL and HOLSAT. There are four components of TATSAT are matched with the definition of both two tourist satisfaction models includes the construct of ambiance, staffs' hospitality, price and accessibility. However, other two TATSAT components that are revealed from the results (cleanliness and security groups), are expanded from the previous models.

Table 4. Comparing dimensions of tourist satisfaction models with the proposed constructs of TATSAT.

Tourist satisfaction models	TATSAT					
	Ambiance	Hospitality	Price	Accessibility	Cleanliness	Security
SERVQUAL						
Tangibles	√	-	-	-	-	-
Reliability	√	-	√	√	-	-
Assurance	-	√	-	-	-	-
Empathy	-	√	-	-	-	-
Responsiveness	-	√	-	-	-	-
HOLSAT						
Accommodation	-	-	-	-	-	-
Physical assets	√	-	-	-	-	-

Tourist satisfaction models	TATSAT					
	Ambiance	Hospitality	Price	Accessibility	Cleanliness	Security
Social life	-	-	-	-	-	-
Ambiance	√	-	-	-	-	-
Heritage and Culture	√	-	-	-	-	-
Transportation	-	-	-	√	-	-
Ability to solve problems	-	√	-	-	-	-

Figure 3 displays a pie chart of the relative proportions of the six categories of tourist attraction satisfaction. It indicates that slightly over half of the online tourists' comments collected (20,568 comments or 51.4%) mentioned the ambiance of a place as their first classification ranking. The second classification ranking, with 6,378 comments (15.9%), is the hospitality factor, followed by the factors of price, accessibility, and cleanliness with 5,493 comments (13.7%), 4,302 comments (10.8%), and 2,588 comments (6.5%), respectively. The findings show that the smallest class is the security factor with 671 comments and 1.7% of the total data collected.

Figure 3. Frequency of each component
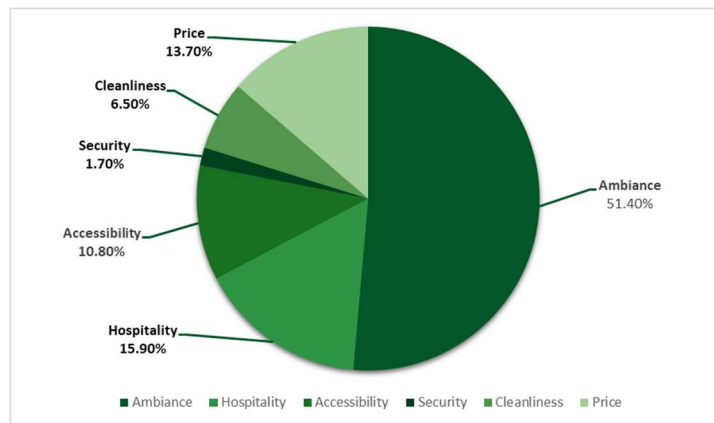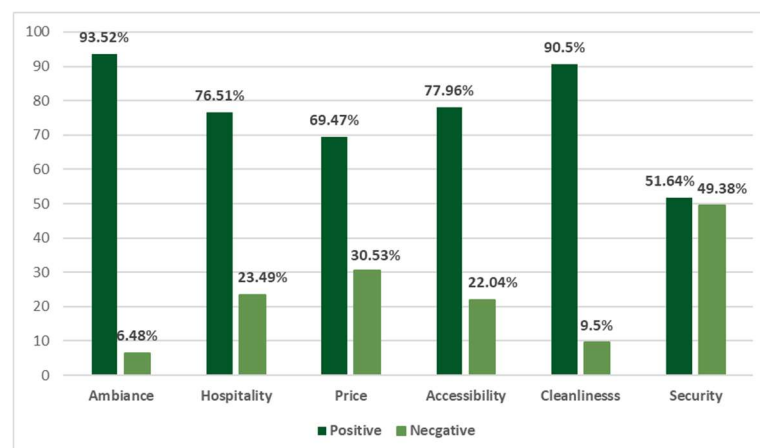

Additionally, this study analyses the sentimental data classification by using the zero-shot technique for the tourist attractions in Thailand as a case study. Given the keyword categories from the previous section, this section classifies the comments into six categories using the zero-shot technique.

Figure 4. Frequency of each component

From 40,000 online tourists' comments, there are 33,974 positive comments (85%) with tourists' rating as 3-5 stars and 6,026 comments (15%) are a negative review with 1-2 stars on tourists' rating. Since the zero-shot technique generates a list of probabilities for each comment, the comment is labelled using the keyword with the highest probability. The results of the zero-shot classification technique for each component are illustrated the frequency of comments for each sentimental class in each province is presented in Figure 4.

From the total of 40,000 online comments, the factor that has the highest number of classified comments is the ambiance factor. This indicates that most tourists pay attention mainly to the ambiance of the attraction site when compared to the other satisfaction factors. Moreover, this component has the highest of positive comments (93.52%) and the lowest negative comments from tourists (6.48%). Besides, the findings present that the proportion of positive comments of all components of TATSAT are greater than the negative one. Therefore, it can assume that the majority of tourists who visited Thai's tourist attractions were happy and have a positive attitude toward these attractive places. However, the security factor has almost equal results between positive and negative feedbacks. Comparing with other five factors, the security has highest percentage of negative comments with 48.36% while others have less than 30%.

4. Discussion

After analyzing the data of 40,000 online tourist comments from 40 tourist attractions in Thailand, the research presents 6 main factors of tourists' satisfaction with tourist attractions mentioned. The proposed of tourist attraction satisfaction (TATSAT) model were confirmed by both academics and practitioner experts about the attributes of each component. As illustrating in Table 4, the proposed model advocates and expands the previous literature related tourist satisfaction by constructing the attributes for a tourist attraction business sector from the tourists' point of view. This framework expands the concepts by proposing that the factors of tourists' satisfaction with tourist attractions that generates from tourists' point of view. The conceptual model includes 1) ambiance, 2) hospitality, 3) price, 4) accessibility, 5) cleanliness, and 6) security, as shown in Figure 6. Even though, the concept of TATSAT is agreed with some dimensions of HOLSAT and SERVQUAL in term of the beautiful attractive ambiance of the place, hospitable and service quality of staffs, price and accessibility to the area (Parasuraman *et al.* 1985; Tribe and Snaith 1998), TATSAT model focus on mainly for only the tourist attraction business sector. Furthermore, the results reveal other two important dimensions, which are cleanliness and security factors that impact towards overall tourist satisfaction with tourist attractions.

Figure 4. The constructs of tourists' attractions satisfaction (TATSAT) model
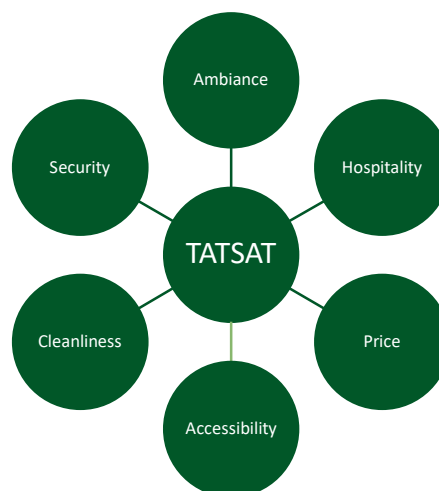
Furthermore, Definitions and descriptions for each component of the proposed model are presented as follow;

Ambiance refers to the pleasant atmosphere or beautiful decoration of a location (Kozak and Rimmington 2000; Bernini and Cagnone 2014), which is covered an amusement and quality of the activities at the site (Della Corte *et al.* 2015). This component also includes both tangible appearances and intangible or invisible forces (Ceyland and Ozcelik 2016) that impacts to tourists' feeling.

Hospitality covers the service-mind and service quality of staffs (Albayrak and Caber 2016) which also refers to the staffs' responsiveness on their duty, empathy to tourists and ability to solve the problems (Pizam *et al.* 1978; Tribe and Snaith 1998). It also includes staffs' communication skill and knowledge in their expertise area (Kozak and Rimmington 2000). Moreover, a friendliness and hospitable of local people (McDowell 2010; Della Corte *et al.* 2015) is included in this component.

Price is a perception of value and cost those tourists need to handle when visit the attraction (Pizam *et al.* 1978; Marin and Taberner 2008; Shahrivar 2012). It is covered the perception regarding overall value for money after the visiting (Kozak and Rimmington 2000; Della Corte *et al.* 2015), which lead to the reliability of the attraction (Parasuraman *et al.* 1985).

Accessibility refers to an ease to access to the attraction area (Marin and Taberner 2008; Albayrak and Caber 2016) and covers a process and quality of transportation and traffic flow to the attraction (McDowell 2010; Bernini and Cagnone 2014; Della Corte *et al.* 2015). It also includes the crowd of visitors and the lead time for the service process at the site.

Cleanliness is referred to the overall hygiene and sanitation of the facility (Marin and Taberner 2008; Della Corte *et al.* 2015; Albayrak and Caber 2016), which also may lead to a clear and clean environment and infrastructure of the attraction (Kozak and Rimmington 2000).

Security relates to a feeling of safety and security when tourists visit the site (Kozak and Rimmington 2000; Bernini and Cagnone 2014). It also leads to all safety and dangerous perception aspects that visitors perceive during their visiting (Della Corte *et al.* 2015).

After applying TATSAT to analyze the 40,000 online tourist reviews of the tourist attractions in Thailand, the reviews were mainly mentioned about the ambiance factor, following by hospitality factor, price factor, accessibility factor, cleanliness factor and security factor. Additionally, the sentimental classification reveals that the majority tourists satisfy with the performance and services of tourist attractions in Thailand because the greater amount of positive online feedbacks were posted through the social media platform compares to the negative feedbacks in all components of the model. Besides, it might be seen that most tourists tend to share or review their positive experience or satisfied expectation through online platform rather than complain about the dissatisfied experience.

Conclusion

This study aims to analyze and identify the important factors of tourists' satisfaction with tourist attractions from online tourists' reviews by applying the zero-shot learning (ZSL) technique and sentimental classification. The research outcomes contribute a comprehensive understanding for both scholars and practitioners who are interested in developing greater customer satisfaction in the context of the tourist attraction sector. TATSAT model takes HOLSAT, SERVQUAL and destination attributes from previous related scholars as a base to develop all component. In order to increase the tourist satisfaction levels and competitive advantages of tourist attraction, not only entrepreneurs and marketers, but also the national policy makers or tourism regulator should consider the proposed factors of the tourist attractions' satisfaction model. These six components are the main factors affecting overall tourist satisfaction with tourist attractions; include 1) ambiance, 2) hospitality, 3) price, 4) accessibility, 5) cleanliness, and 60 securities. As an exploratory study conducted with Thailand's attractions, this proposed model provides a holistic picture and examine the main components of tourist satisfaction factors with tourist attraction by analyzing online tourist reviews. Due to limited research focusing mainly on tourist attraction sector, this could offer an opportunity for further research in applying the model in other research areas.

For a practitioner sector, the Tourism Authority of Thailand or the tourist attractions' entrepreneur, should concern not only the ambiance of the attraction, hospitable of staffs, pricing and accessibility to the place, but also attention more on the hygienic and security at the attraction in order to reduce the negative reviews and improve the overall tourist satisfaction level. Although the collected data are mainly online comments that were posted before the COVID-19 pandemic, many online reviews mentioned about the cleanliness and security of the attraction sites. This also proves that, in order to increase a competitive advantage and overall tourist satisfaction level, the personal safety and hygiene components are other two important aspects which tourist attractions should consider. Especially, the security factor which was complained negatively more than others construct of TATSAT. Moreover, during the COVID-19 pandemic, people across the world have changed their behaviors and are more concerned about the new normal and social distancing. This new normal lifestyle also influences travel behavior, and service providers must take hygienic and safety into account in order to fulfill tourists' satisfaction in the post-COVID-19 era. Countries, like Thailand, where tourism businesses are main contributors to the economy, should prepare to reopen post-COVID-19.

In spite of the research limitations, a greater number and equal of positive and negative comments should be studied to increase the results' trustworthiness. Also, the relationship between comment posting dates and tourists' satisfaction levels should be considered in further research. Nonetheless, the proposed model should be applied with other research techniques to confirm the reliability and validity of the findings.

References

[1] Albayrak, T. and Caber, M. 2016. Destination attribute effects on rock climbing tourist satisfaction: An Asymmetric Impact-Performance Analysis. *Tourism Geographies,* 18(3): 280-296. DOI:10.1080/14616688.2016.1172663

[2] Alegre, J. and Garau, J. 2010. Tourist satisfaction and dissatisfaction. *Annals of tourism research,* 37(1): 52-73. DOI: 10.1016/j.annals.2009.07.001

[3] Baggio, R. and Caporarello, L. 2005. Decision support systems in a tourism destination: literature survey and model building. In U: Proceedings itAIS-2nd Conference of the Italian chapter of AIS (Association for Information Systems). Verona, Italy.

[4] Bernini, C. and Cagnone, S. 2014. Analysing tourist satisfaction at a mature and multi-product destination. *Current Issues in Tourism,* 17(1): 1-20. DOI: 10.1080/13683500.2012.702737

[5] Capriello, A., Mason, P.R., Davis, B. and Crotts, J. C. 2013. Farm tourism experiences in travel reviews: A cross-comparison of three alternative methods for data analysis. *Journal of Business Research,* 66(6): 778-785. DOI: 10.1016/j.jbusres.2011.09.018

[6] Ceylan, C. and Ozcelik, A. B. 2016. A circular approach to SERVQUAL and HOLSAT: An implementation suggestion. *Journal of Hotel & Business Management,* 5(1): 1-10. DOI: 10.4172/2169-0286.1000125

[7] Chenini, A., and Touaiti, M.2018. Building destination loyalty using tourist satisfaction and destination image: A holistic conceptual framework. *Journal of Tourism, Heritage & Services Marketing,* 4(2): 37-43. DOI:10.5281/zenodo.1490491

[8] Cherapanukorn, V. and Charoenkwan, P. 2017. Word cloud of online hotel reviews in Chiang Mai for customer satisfaction analysis. In *2017 International Conference on Digital Arts, Media and Technology (ICDAMT):* 146-151. IEEE. DOI: 10.1109/ICDAMT.2017.7904952

[9] Crouch, G. I., and Ritchie, B.JR. 1999. Tourism, competitiveness, and societal prosperity. *Journal of business research,* 44(3): 137-152. DOI: https://doi.org/10.1016/S0148-2963(97)00196-3

[10] Del Chiaro, R., Bagdanov, A. D. and Del Bimbo, A. 2019. Webly-supervised zero-shot learning for artwork instance recognition. *Pattern Recognition Letters,* 128: 420-426. DOI:10.1016/j.patrec.2019.09.027

[11] Della Corte, V., Sciarelli, M., Cascella, C. and Del Gaudio, G. 2015. Customer satisfaction in tourist destination: The case of tourism offers in the city of Naples. *Journal of Investment and Management,* 4(1-1): 39-50. DOI: 10.11648/j.jim.s.2015040101.16

[12] Emel, G.G., Taşkin, Ç. and Akat, Ö. 2007. Profiling a domestic tourism market by means of association rule mining. *Anatolia,* 18(2): 334-342. DOI:10.1080/13032917.2007.9687209

[13] Eusébio, C. and Vieira, A.L. 2013. Destination attributes' evaluation, satisfaction and behavioural intentions: A structural modelling approach. *International Journal of Tourism Research,* 15(1): 66-80. DOI:10.1002/jtr.877

[14] Hu, Wei, and Wall, G.2005. Environmental management, environmental image and the competitive tourist attraction. *Journal of sustainable tourism,* 13(6): 617-635. DOI: 10.1080/09669580508668584

[15] Kang, N. and Yu, Q. 2011. Corpus-based Stylistic Analysis of Tourism English. *Journal of Language Teaching & Research,* 2(1): 129-136. DOI: 10.4304/jltr.2.1.129-136

[16] Khan, Khairullah, Baharum Baharudin, Aurnagzeb Khan, and Ashraf Ullah. 2014. "Mining opinion components from unstructured reviews: A review. *Journal of King Saud University-Computer and Information Sciences* 26(3): 258-275. DOI: https://doi.org/10.1016/j.jksuci.2014.03.009

[17] Khan, S. S., Khan, M., Ran, Q. and Naseem, R. 2018. Challenges in opinion mining, comprehensive review. *A Science and Technology Journal,* 33(11): 123-135. Available at: https://www.researchgate.net/publication/328980053

[18] Kozak, M. and Rimmington, M. 2000. Tourist satisfaction with Mallorca, Spain, as an off-season holiday destination. *Journal of travel research,* 38(3): 260-269. DOI: 10.1177/004728750003800308

[19] Lam, P. YW. 2007. A corpus-driven lexico-grammatical analysis of English tourism industry texts and the study of its pedagogic implications in English for Specific Purposes. In *Corpora in the foreign language classroom*: 71-89. Brill Rodopi. DOI: https://doi.org/10.1163/9789401203906_006

[20] Lee, C.C., and Hu, C. 2005. Analyzing Hotel customers' E-complaints from an internet complaint forum. *Journal of Travel & Tourism Marketing* 17(2-3): 167-181. DOI: 10.1300/J073v17n02_13

[21] Lew, A.A. *et al.* 2020. Visions of travel and tourism after the global COVID-19 transformation of 2020. *Tourism Geographies* 22(3): 455-466. DOI:10.1080/14616688.2020.1770326

[22] Li, H., Ye, Q. and Law, R. 2013. Determinants of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pacific Journal of Tourism Research* 18(7): 784-802. DOI:10.1080/10941665.2012.708351

[23] Lindquist, H. 2018. *Corpus linguistics and the description of English.* Edinburgh University Press.

[24] Liu, X., Mehraliyev, F., Liu, C. and Schuckert, M. 2020. The roles of social media in tourists' choices of travel components. *Tourist Studies,* 20(1): 27-48. DOI: 10.1177/1468797619873107

[25] Marin, J.A. and Jaume Garau Taberner. 2008. Satisfaction and dissatisfaction with destination attributes: Influence on overall satisfaction and the intention to return. *Retrieved December* 18: 2011. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.612.6418&rep=rep1&type=pdf

[26] McDowall, S. 2010. International tourist satisfaction and destination loyalty: Bangkok, Thailand. *Asia Pacific Journal of Tourism Research* 15(1): 21-42. DOI: 10.1080/10941660903510040

[27] Parasuraman, A., Zeithaml, V.A. and Berry, L.L. 1985. A conceptual model of service quality and its implications for future research. *Journal of marketing*, 49(4): 41-50. DOI: 10.2307/1251430

[28] Pearce, P.1998. Marketing and management trends in tourist attractions. *Asia Pacific Journal of Tourism Research* 3(1): 1-8. DOI: https://doi.org/10.1080/10941669908722002

[29] Pizam, A., Neumann, Y. and Reichel, A. 1978. Dimentions of tourist satisfaction with a destination area. *Annals of tourism Research* 5(3): 314-322. DOI: https://doi.org/10.1016/0160-7383(78)90115-9

[30] Puangyaem, P. 2013. *A Corpus Based Study of the Use of Adjectives in Tourist Websites.* Language Institute, Thammasat University.

[31] Ritchie, J.R., Goeldner, C. R. and McIntosh, R. W. 2003. Tourism: principles, practices, philosophies. John Wiley & Son (New Jersey).

[32] Sahadev, S. and Islam, N. 2004. Exploring the determinants of e-commerce usage in the hotel industry in Thailand: An empirical study. *Information Technology & Tourism*, 7(3-4): 171-180. DOI:10.3727/109830505774297201

[33] Sappadla, P.V., Nam, J., Mencía, E.L. and Fürnkranz, J. 2016. Using semantic similarity for multi-label zero-shot classification of text documents. In ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges, Belgium, April, 27-29, 2016.

[34] Schmitt, N. and Schmitt, D. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching,* 47(4): 484-503. DOI: https://doi.org/10.1017/S0261444812000018

[35] Schuckert, M., Liu, X. and Law, R. 2015. Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing,* 32(5): 608-621. DOI:10.1080/10548408.2014.933154

[36] Shahrivar, R.B. 2012. Factors that influence tourist satisfaction. *Journal of Travel and Tourism Research; Special Issue Destination Management-2012 (Online)* 12(1): 61. Available at: https://www.researchgate.net/publication/284969853_Factors_that_influence_tourist_satisfaction

[37] Swarbrooke, J. and Page, S. J. 2012. *Development and management of visitor attractions.* Routledge.

[38] Tribe, J. and Snaith, T.1998. From SERVQUAL to HOLSAT: holiday satisfaction in Varadero, Cuba. *Tourism management,* 19(1): 25-34. DOI: https://doi.org/10.1016/S0261-5177(97)00094-0

[39] Verma, S. and Yadav, N. 2021. Past, present, and future of electronic word of mouth (EWOM). *Journal of Interactive Marketing,* 53: 111-128. DOI: https://doi.org/10.1016/j.intmar.2020.07.001

[40] Ye, M. and Guo, Y. 2017. Zero-shot classification with discriminative semantic representation learning.In *Proceedings of the IEEE conference on computer vision and pattern recognition:* 7140-7148. IEEE. DOI:10.1109/CVPR.2017.542

[41] Ye, Z. *et al.* 2020. Zero-shot text classification via reinforced self-training. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics:* 3014-3024. IEEE. DOI: 10.18653/v1/2020.acl-main.272

[42] Zhang, Z. *et al.* 2011. Sentiment classification of Internet restaurant reviews written in Cantonese. *Expert Systems with Applications,* 38(6): 7674-7682. DOI: 10.1016/j.eswa.2010.12.147

[43] UNWTO, 2019. International Tourism Highlights: 2019 Edition. Available at: https://www.e-unwto.org/doi/pdf/10.18111/9789284421152

[44] UNWTO, 2021. UNWTO Tourism Recovery Tracker. Available at: https://www.unwto.org/unwto-tourism-recovery-tracker

ASERS