İSTANBUL UNIVERSITY PRESS

Journal of Tourismology

RESEARCH ARTICLE

Journal of Tourismology, 8(1), 49-67

DOI: 10.26650/jot.2022.8.1.1038566
http://jt.istanbul.edu.en/

Submitted: 20.12.2021
Revision Requested: 13.03.2022
Last Revision Received: 17.03.2022
Accepted: 04.04.2022
Published Online: 30.05.2022

Using Sentiment Analysis of Online Hotel Reviews To Explore the Effect of Information and Communication Technologies on Hotel Guest Satisfaction

Salma Cherdouh[1], Abdenacer Kherri[2], Ayoub Abbaci[3], Salim Kebir[4]

Abstract

Online hotel reviews are a rich source of information regarding drivers of customer satisfaction or dissatisfaction with hotel services. In this paper, we study the impact of information and communication technologies (ICT) on the satisfaction of customers of 144 Algerian hotels through the analysis of 11310 online user reviews. The methodology adopted is based on the analysis of the sentiments expressed in the user reviews from one of the most used travel platforms, TripAdvisor. The results indicate that sentiment towards ICT contributes to satisfaction but not as much as sentiment towards other non-ICT services. Furthermore, we identified individually which ICTs contribute to satisfaction. We found that ICTs related to booking, comfort and entertainment were the ones that significantly contributed to satisfaction, while ICT related to safety and security did not. We believe that this study contributes to the literature because it uses innovative natural language processing techniques as opposed to traditional questionnaire-based approaches. Moreover, this study was conducted in a developing country where ICT adoption differs from that of developed countries.

Keywords

Information and communication technology, Sentiment analysis, Online hotel reviews, Hotel guest satisfaction, Algerian hotels

1 **Correspondence to:** Salma Cherdouh (PhD. Student), Laboratoire RMTQ, Faculté des Sciences Economiques, Commerciales et des Sciences de Gestion, Université de Bejaia, 06000 Bejaia, Algérie. E-mail : salma.cherdouh@univ-bejaia.dz ORCID: 0000-0002-6215-1045

2 Abdenacer Kherri (Asst. Lecturer), . Ecole des Hautes Etudes Commerciales, Koléa, Algeria. E-mail: a.kherri@hec.dz ORCID: 0000-0003-2923-5922

3 Ayoub Abbaci (Asst. Lecturer), Ecole Nationale Polytechnique, Algiers, Algeria. E-mail: ayoub.abbaci@univ-bejaia.dz ORCID: 0000-0001-5067-7579

4 Salim Kebir (Asst. Lecturer), Ecole Supérieure de Technologies Industrielles d'Annaba, Algeria. E-mail: s.kebir@esti-annaba.dz ORCID: 0000-0002-7101-6523

To cite this article: Cherdouh, S., Kherri, A., Abbaci, A., & Kebir, S. (2022). Using sentiment analysis of online hotel reviews to explore the effect of information and communication technologies on hotel guest satisfaction. *Journal of Tourismology, 8*(1), 49-67. https://doi.org/10.26650/jot.2022.8.1.1038566

Introduction

Nowadays, the use of web-based travel platforms has become ingrained in the habits of travelers and tourists before, during and after their trips. Their use is not limited to online booking of airline tickets and hotel rooms but extends to other activities such as sharing travel experiences in which tourists express their adventures and misadventures, down to the smallest details about what they liked and disliked about their destination (He et al., 2017). This intensive exchange of information and user experiences has paved the way for the emergence of what is generally known as user-generated content and more specifically in the hotel industry as online hotel reviews.

Thanks to their open structure, online hotel reviews allow customers to describe and report exhaustively on their user experience and their perception of the services they have been offered (Zhao et al., 2019). As a result, online hotel reviews are considered a reliable source of information that can be used by other potential customers to support their decision to search, choose, and finally decide to book a hotel (Fang et al., 2016; Nusair et al., 2013; Chua & Banerjee, 2016). In addition to writing a textual review, all leading web-based travel platforms such as TripAdvisor, Booking and Expedia allow users to give an overall rating of their stay using a scale generally ranging from 1 (very dissatisfied) to 5 (very satisfied). This scale reflects the degree of customer satisfaction (Bulchand-Gidumal et al., 2011; Geetha et al., 2017) and thus represents a crucial interest for hotel companies as it determines future customer behaviors such as willingness to return and intention to recommend the hotel. For this reason, identifying what makes customers happy and what irritates them has long been the subject of numerous research studies (Padma & Ahn, 2020; Davras & Caber, 2019; Nunkoo et al., 2020; Ruan, 2020). Indeed, understanding customers' expectations and preferences is a critical step for hotel managers in order to improve and adapt their marketing strategy to meet the needs of their guests (Ahani et al., 2019).

Throughout the last few years, with the success of artificial intelligence and machine learning algorithms in other fields, the exploitation of knowledge extracted from online hotel reviews has surpassed the traditional use of questionnaires in satisfaction surveys. Indeed, many studies relying on text mining, natural language processing and sentiment analysis have been used on online hotel reviews in order to deduce customers' preferences and identify the drivers of their satisfaction (Berezina et al., 2016; Geetha et al., 2017; Zhao et al., 2019; Padma & Ahn, 2020; Shin et al., 2021; Ahani et al., 2019).

Information and communication technologies (ICTs) are among the services that have received particular attention from research studies regarding their effect on hotels' guest satisfaction. In fact, many studies have investigated and demonstrated that the presence of high-quality ICT equipment is a determining factor of satisfaction (Sirirak et al., 2011; Velazquez et al., 2015; Melian-Gonzalez & Bulchand-Gidumal, 2016; Chevers & Spencer, 2017; Moliner-Velazquez et al., 2019). However, the aforementioned studies were primarily based on surveys to investigate the effect of ICTs on satisfaction (Khoo-Lattimore et al., 2019). Our study focuses on the evaluation of the impact of ICTs on the overall satisfaction of hotel guests through the textual analysis of online hotel reviews. For this purpose, we rely on innovative and well-founded natural language processing techniques to evaluate the sentiment of sentences and to study the effect of this sentiment on the rating given by the user in his review. Therefore, the aim of this study is to answer the following research questions:

RQ1 *(a) Does customer review sentiment towards ICT components have a positive impact on the rating? (b) Is this impact greater than that of customer review sentiment towards other non-ICT services?*

RQ2 *Which of these ICT components, if any, contribute the most to the review rating?*

The remainder of this paper is organized as follows. Section 2 provides a literature review on ICT adoption in hotels, hotel guest satisfaction and sentiment analysis of online hotel reviews. Section 3 presents our methodology, with a focus on the measuring instruments and the adopted data analysis technique. Section 4 presents and discusses the major findings. Section 5 concludes the paper and provides future perspectives, as well as the main theoretical and managerial implications, and addresses the potential limitations of our study.

Literature Review

Relationship Between ICT Adoption and Hotel Guest Satisfaction

ICTs can be defined as any technology that is used to "*create, capture, manipulate, communicate, exchange, present, and use information in its various forms*" (Ryssel et al., 2004). Several studies found that ICTs in hotels contribute positively to customer satisfaction (Cobanoglu et al., 2011; Sirirak et al., 2011; Velazquez et al., 2015; Moliner-Velazquez et al., 2019; Chevers & Spencer, 2017). This can be explained among other things by the fact that customers have become more sophisticated, educated, and technologically experienced and pay more attention to personalized services (Kim & Ham, 2006).

Cobanoglu et al. (2011) show that not all ICT amenities impact guest satisfaction equally. Indeed, the authors pointed out that business essentials for travelers, such as express check-in/check-out, in-room telephone, in-room alarm clock, and universal battery chargers, were found to have a more positive impact than other ICT amenities on guest's overall satisfaction. Moreover, the study showed that the variety and type of ICT used are important factors that influence hotel selection and return intention. Furthermore, Sirirak et al. (2011) studied the relationship between ICT adoption and satisfaction by measuring ICT adoption using three factors: availability, integration, and intensity of usage. They found that ICT usage intensity has a significant positive relationship with customer satisfaction. More recently, Chevers & Spencer (2017)'s research on Jamaican hotels found a significant relationship between ICT adoption and hotel guest satisfaction. The authors classify ICTs according to hotel operational domains and claim that in-room ICTs have the most positive effect on guests' satisfaction. More specifically, Bulchand-Gidumal et al. (2011) demonstrated that offering free Wi-Fi can constitute a source of revenue for hotels because Wi-Fi contributes to enhancing customer satisfaction.

Despite the numerous research studies that have investigated the relationship between ICT and hotel customer satisfaction, it can be noted that the majority of them have mainly targeted hotel customers in developed countries. However, it has been argued in the literature that the level of ICT adoption in developed countries differs from that in developing countries (Ezzaouia & Bulchand-Gidumal, 2020). Furthermore, it has been shown that one of the main barriers to ICT adoption in developing countries is the lack of technological infrastructure, the poor quality of the Internet network (Zaied, 2012) and the high cost of acquiring software for the digitalization of business processes (Alrawadieh et al., 2021). In addition, the works of Ahmad et al. (2015) and Adebanjo et al. (2016) assert that the hesitation in adopting ICT in the hospitality industry in developing countries comes from the lack of awareness of the potential that the technologies can offer as well as the lack of financial resources to implement them. In response to these circumstances, there have been calls within the research community to study the impact of ICTs in the hospitality sector in developing countries (Mihalic & Buhalis, 2013).

Identifying Drivers of Hotel Customer Satisfaction Through Sentiment Analysis of Online Hotel Reviews

Online hotel reviews have grown in importance as a means of understanding customer perception as well as forecasting future behavior, particularly in the tourism and hospitality industries (Padma & Ahn, 2020). Indeed, they are considered by several academics as a meaningful emerging alternative to traditional approaches using face-to-face interviews or surveys in which it is difficult to capture and accurately assess customer satisfaction (Berezina et al., 2016; Gao et al., 2018). However, with the advent of big data and the exponential increase in the volume of user-generated content, it is inconceivable to process hotel reviews in their raw state. For these reasons, more and more researchers in the hospitality and tourism field have relied on innovative techniques from the field of natural language processing (NLP) such as opinion/sentiment analysis, language detection and topic classification.

Lu & Stepchenkova (2012) proposed a method for evaluating ecotourism experiences from online hotel reviews extracted from TripAdvisor. The authors identified 26 attributes that influence ecotourists' satisfaction with their ecolodge stays and aggregated them into seven categories: ecolodge settings, room, nature, service, food, location and value for money. The study carried out by Li et al. (2013) found that transportation convenience, food and beverage management, convenience to tourist destinations and value for money were important and excellent factors for assessing the satisfaction of those customers who are booking both luxury and budget hotels. Both studies from Xiang et al. (2015) and Zhang et al. (2016) rely on topic classification techniques to deconstruct hotel guest experience and examine its association with overall guest satisfaction. The work of Geetha et al. (2017) found consistency between customer ratings and actual customer sentiment polarity across hotels belonging to the two categories of premium and budget. Liu et al. (2017) combined language detection and sentiment analysis techniques on a collection of 412,784 online hotel reviews extracted from TripAdvisor with the aim of offering new insights into the determinants of hotel customer satisfaction depending on their native language. More recently, Zhao et al. (2019) investigated technical attributes of online textual reviews and found that a higher level of subjectivity and readability and a longer length of textual review lead to lower overall customer satisfaction, and a higher level of diversity and sentiment polarity of textual reviews lead to higher overall customer satisfaction.

Prior studies in the area of sentiment analysis of online hotel reviews suffer from two main shortcomings. First, it should be noted that the majority of studies focus on a single natural language processing library to analyze review sentiment. This poses a significant threat to their internal validity, as some libraries may give more accurate results than others. Second, past studies can be characterized as blind in that they do not focus on a specific service offered by the hotel, but seek to exhaustively discover which services contribute to guest satisfaction. In our study, we sought to fill this research gap by performing sentiment analysis using several different sentiment analysis libraries to mitigate the threat to the internal validity of our findings. On the other hand, we focus only on one category of services offered by hotels, namely ICT.

Methodology

Data Collection

For this study, we used TripAdvisor as our primary source of information. Founded in 2000, TripAdvisor is one of the largest and most popular platforms containing travel-related reviews. The number of online reviews on TripAdvisor is constantly growing and has exceeded one billion by the end of 2021 according to forecasts (Statista, 2021). TripAdvisor allows users to post, comment, and share trip suggestions, as well as rate hotels, restaurants, and destinations. An online hotel review on TripAdvisor contains various information such as the title of the review, the body of the review, the date of publication, the name of the hotel, the number of stars of the hotel, the city of the hotel, the country of the hotel as well as the overall rating that the user gave to the hotel.

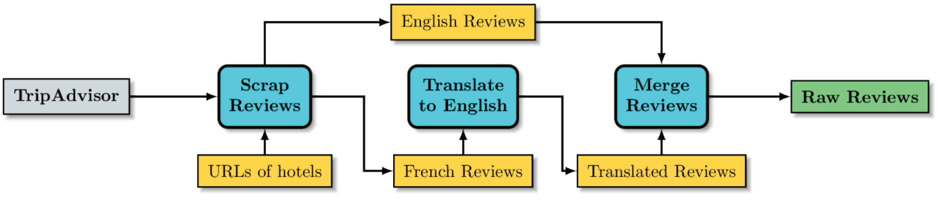

Figure 1. Data collection process.

As illustrated in Figure 1, the data collection process in our study goes through three steps. First, we developed a Python program to extract online hotel reviews from TripAdvisor. This program accepts as input a list of hotel URLs and produces as output a file containing all collected reviews. We manually collected the URLs of the first 144 Algerian hotels present on TripAdvisor sorted in the decreasing order of their value. After collecting all the hotel reviews of the hotels in question, we decided to keep only the reviews written in French and English because for all the other languages (Arabic, Italian, Chinese, etc.), the number of reviews was extremely small and to integrate them in our analysis would risk decreasing the reliability of our analysis. Since the majority of sentiment analysis libraries have been developed to handle texts written in English, we decided to translate all reviews written in French into English. The translation was done using a Python program that calls Google's t5 (Raffel et al., 2020) which is one of the most downloaded models for text translation hosted on Hugging Face Model Hub (HuggingFace, 2021a). The latter is a repository that hosts state-of-the-art machine learning models dedicated to natural language processing created by top AI researchers and web giants such as Google, Facebook and Microsoft (Wolf et al., 2020). Once the translation is complete, all reviews are collected and stored in a single JavaScript Object Notation (JSON) file containing relevant information about each review such as title, text, reason for stay, URL of the review, etc. JSON is a standard open file format that uses human-readable text to store and transmit data consisting of the following attributes: value pairs and arrays. A total of 11957 online hotel reviews were collected. Figure 2 illustrates the structure of a single review.

```
{
  "hotel_name": "Sheraton Annaba Hotel",
  "hotel_location": "Annaba",
  "hotel_country": "Algeria",
  "hotel_stars": "5 Star",
  "rating": 5,
  "title": "Great Stay for business",
  "text": "I really enjoyed my stay at the Sheraton Annaba with
      great service, friendly staff and the manager Mehdi, has been
       very helpful in handling our issues.  The building itself is
       fabulous and prices are within range of hotels of this
      standards in the country.  I would recommend this hotel to
      business travelers as it's well located, safe and includes
      high speed wifi access.",
  "date": "2017-08-27",
  "reason_of_stay": "BUSINESS",
  "url": "https://www.tripadvisor.com/ShowUserReviews-g1071600-
      d12063561-r517969346-Sheraton_Annaba_Hotel-
      Annaba_Annaba_Province.html"
}
```

Figure 2. Overview of the structure of a review.

Data Cleaning and Sample Profile

In order to perform a statistical analysis of the data and to eliminate outliers, we have kept, from among the 11957 user reviews, only the 11310 that concern 3, 4 and 5-star hotels. Upscale hotels were selected because they are more likely to invest in technology. Table 1 describes the profile of the hotel reviews we collected.

Table 1
Sample Profile

Characteristics	Frequency	Percentage (%)	Characteristics	Frequency	Percentage (%)
Hotel Category			*User Rating*		
3 Star	3618	31.99%	★☆☆☆☆	936	8.28%
4 Star	2923	25.84%	★★☆☆☆	940	8.31%
5 Star	4769	42.17%	★★★☆☆	1815	16.05%
Reason of Stay			★★★★☆	3248	28.72%
Business	6098	53.92%	★★★★★	4371	38.65%
Family	1822	16.11%	*Hotel Region*		
Couples	1346	11.90%	East	2244	19.84%
Solo	713	6.30%	West	2979	26.34%
Friends	644	5.69%	Center	6017	53.20%
Not available	687	6.07%	South	70	0.62%

Among the 11310 user reviews, almost half (42.17 percent) were written for 5-star hotels, followed by 31.99 percent written for 3-star hotels and 25.84 percent for 4-star hotels. Regarding the rating given by the users, 38.65 percent gave a rating of 5, 28.72 percent gave a rating of 4, 16.05 percent gave a rating of 3, 8.31 percent gave a rating of 2 and 8.28 percent gave a rating of 1. Finally, more than half (53.92 percent) of the users stayed for business purposes, while the remaining users stayed for other reasons.

Data Preprocessing

Since our research focuses on the analysis of the sentiment towards ICTs and on the study of its effect on the user rating, the data we have collected needs to be preprocessed to enable us to conduct the necessary statistical analyses. We describe in the following the different variables we calculated for each review *R*.

- $rating(R)$ denotes the rating of the review *R*. Its value is an integer within the range $[1, 5]$ where 1 indicates very unsatisfied and 5 indicates very satisfied.
- $sentences(R) = \{s_1, s_2, \ldots, s_n\}$ denotes the set of sentences of the review *R*. We used the sentence tokenizer provided by the NLTK Python library (Bird et al., 2009) to divide the text of a review into several individual sentences.
- $sentiment_{lib}(s_i)$ denotes the sentiment score of the sentence s_i measured using the library lib. Its value is a float within the range $[-1.0, 1{,}0]$ where −1.0 indicates a negative sentiment and 1.0 indicates a positive sentiment. In order to mitigate the threat to the internal validity of our approach, we used four different sentiment analysis libraries, namely TextBlob (Loria, 2018), Vader (Hutto & Gilbert, 2014), Flair (Akbik et al., 2019) and Transformers (Wolf et al., 2020). TextBlob and Vader are both lexicon-based Python sentiment analysis libraries, i.e., for these two libraries the sentiment of a given text is an aggregate of weights assigned to the words in that text. For example, the words "*good*", "*great*" and "*happy*" have a positive weight, whereas the words "*horrible*", "*difficult*" and "*unhappy*" have a negative weight. Flair and Transformers are both machine learning-based Python libraries for sentiment analysis. That is to say that they both use supervised learning models trained on large text corpora. Machine learning-based sentiment analysis libraries generally offer better accuracy than lexicon-based libraries because they act not on the text itself but on a tree-like representation of the text that captures the intensity of the links between words. However, due to the computational and memory requirements to implement the supervised learning models, machine learning-based libraries require much more time to run compared to lexicon-based libraries.
- $sentiment_{lib}(R) = \sum_{s_i \in sentences(R)} \frac{sentiment_{lib}(s_i)}{|sentences(R)|}$ denotes the sentiment score of the review *R* measured by calculating the arithmetic mean of the sentiment score of all its sentences.
- $label(s_i)$ denotes a label associated with the sentence s_i. To associate a label with a sentence, we used a Python program that calls Facebook's bartlarge-mnli (Lewis et al., 2019) which is one of the most downloaded models for zero-shot classification hosted on Hugging Face Model Hub (HuggingFace,

2021b). Zero-shot classification is a challenging machine learning and natural language processing problem that aims to automatically associate an appropriate label with a piece of text among several labels (Yin et al., 2019). Figure 3 shows an example of the use of zero-shot classification.

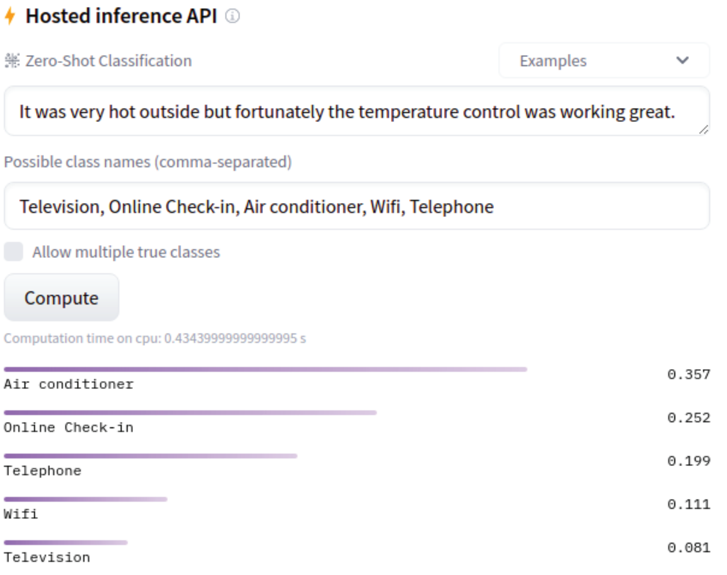

Figure 3. Zero-shot classification (screenshot taken from huggingface.co).

As illustrated in Figure 3, even though the words "*air*" and "*conditioner*" are not explicitly used in the sentence, "*Air conditioner*" has actually been found as the best candidate label among the other labels. In our study, we consider the following labels for ICT services: "*Electronic Safe Box*", "*Television*", "*Online Check-in*", "*Air conditioner*", "*Wi-Fi*", "*Email*", "*USB Charger*", "*Telephone*", "*Social Network*", "*Electronic Lock*", "*Web Site*" and "*ATM* ". For all other non-ICT services, we used the "*non-ICT services*" label.

Accuracy of Used Sentiment Analysis Libraries

Since the sentiment analysis libraries we used are the measuring instruments of our experiment, we found it useful to evaluate their degree of accuracy. To assess the accuracy, Pearson correlation analysis was applied to identify the correlation between $sentiment_{lib}(R)$ and $rating(R)$ for the 11310 reviews we collected. Table 2 presents the correlation test results.

Table 2
Results of Pearson Correlation Test

		$rating(R)$
$sentiment_{TextBlob}(R)$	Pearson's r	0.634***
	p-value	0.000
$sentiment_{Vader}(R)$	Pearson's r	0.680***
	p-value	0.000
$sentiment_{Flair}(R)$	Pearson's r	0.771***
	p-value	0.000
$sentiment_{Transformers}(R)$	Pearson's r	0.802***
	p-value	0.000

Notes: ***p<0.001

Table 2 shows a significant positive correlation between $sentiment_{lib}(R)$ and $rating(R)$ for the four libraries with r = 0.634, p < 0.001 for TextBlob, r = 0.680, p < 0.001 for Vader, r = 0.771, p < 0.001 for Flair and r = 0.802, p < 0.001 for Transformers. This indicates that the four libraries are highly accurate in analyzing the sentiment of hotel reviews. Moreover, it is worth noting the superiority of machine learning-based libraries compared to lexicon-based ones in terms of accuracy.

Findings and discussion

To address ***RQ1***, a multiple regression analysis was carried out using sentiment with hotel ICTs and sentiment with other non-ICT factors as the independent variables and rating(R) as the dependent variable. We measured sentiment with hotel ICTs and sentiment with other non-ICT factors using the following equations.

$$sentiment_{lib}^{ICT}(R) = \sum_{s_i \in sentences(R)}^{label(s_i) \in ICT} \frac{sentiment(s_i)}{|\{s_i / \, s_i \in sentences(R) \land label(s_i) \in ICT\}|} \quad (1)$$

$$sentiment_{lib}^{nonICT}(R) = \sum_{s_i \in sentences(R)}^{label(s_i) \notin ICT} \frac{sentiment(s_i)}{|\{s_i / \, s_i \in sentences(R) \land label(s_i) \notin ICT\}|} \quad (2)$$

In equation (1), $sentiment_{lib}^{ICT}(R)$ denotes the sentiment score towards ICTs of the review *R* measured using the library lib. It is obtained by calculating the arithmetic mean of the sentiment score of ICT-labeled sentences in the review *R*. In a complementary way, in equation (2), $sentiment_{lib}^{nonICT}(R)$ denotes the sentiment score towards other non-ICT services of the review *R* measured using the library lib. It is obtained by calculating the arithmetic mean of the sentiment score of sentences whose label is not an ICT in the review *R*. Table 3 summarizes the results of the multiple regression analysis.

Table 3

Regression Results Of Sentiment Score Towards ICTs And Sentiment Score Towards non-ICT Services On $rating(R)$

	$rating(R)$							
	TextBlob		**Vader**		**Flair**		**Transformers**	
	β	$p-value$	β	$p-value$	β	$p-value$	β	$p-value$
$sentiment_{lib}^{ICT}(R)$	1.31	0.000***	0.84	0.000***	0.43	0.000***	0.81	0.000***
$sentiment_{lib}^{nonICT}(R)$	2.80	0.000***	2.18	0.000***	1.06	0.000***	1.81	0.000***
R	0.602		0.679		0.724		0.771	
R^2	0.363		0.461		0.524		0.594	
$Adjusted\ R^2$	0.362		0.460		0.524		0.594	
$F-Statistics$	606		911		1173		1558	
$p-value$	0.000***		0.000***		0.000***		0.000***	

Notes: ***p<0.001

It can be seen that the variation of the rating is explained by the two independent variables using the four sentiment analysis libraries with an $adjusted\ R^2$ of 36.2% (p-value < 0.001) for TextBlob, 46% (p-value < 0.001) for Vader, 52.4% (p-value < 0.001) for Flair and 59.4% (p-value < 0.001) for Transformers. Furthermore, the results of the multiple regression analysis indicate that for all sentiment analysis libraries, both sentiment score of ICT-labeled sentences and sentiment score of sentences whose label is not an ICT have a significant relationship with overall guest satisfaction (p-value < 0.001). Moreover, with respect to β coefficients, it can be noticed that, using the four sentiment analysis libraries, the sentiment score of the sentences that were not labeled ICT had a contribution towards rating more than twice than that of ICT labeled sentences.

In summary, the results show that the sentiment towards ICTs has a significant impact on the user rating and consequently on their satisfaction. However, this impact is not as important when compared to other non-ICT services such as staff service, comfort, location, restaurant and room quality. This can be explained by two reasons related to the sample profile. Firstly, more than two-thirds of the online hotel reviews we collected were from premium category hotels (4 and 5-stars). The remaining reviews came from 3-star hotels. Premium category hotels are known to be impeccable in terms of the ICT facilities they offer. As a result, reviews of this category of hotels do not focus mainly on ICT facilities and usually mention other intangible services such as the politeness or helpfulness of staff and the cleanliness of rooms (Sirirak et al., 2011; Xu & Li, 2016). Second, more than half of the hotel reviews we collected were written by business guests. Guests belonging to this segment are usually older than leisure guests and therefore their online reviews are more sensitive to comfort, quietness of room and food. This is consistent with the findings of the study by Xu et al. (2019a).

To answer ***RQ2***, we conducted a linear regression analysis using sentiment with each single hotel ICT as the independent variable and $rating(R)$ as the dependent variable. We measured sentiment with each single hotel ICT using $sentiment_{lib}^{ICT}(R)$ defined in equation (1) by each time replacing ICT with a single hotel ICT among the following: "*Electronic Safe Box*", "*Television*", "*Online Check-in*", "*Air conditioner*", "*Wi-Fi*", "*Email*", "*USB Charger*", "*Telephone*", "*Social Network*", "*Electronic Lock*", "*Web Site*" and "*ATM*". Table 4 summarizes the results of the linear regressions we conducted. Each row of this table contains the regression result for a given ICT and indicates the value of $adjusted\ R^2$, β coefficients and p-values obtained using the four sentiment analysis libraries.

Table 4

Regression results of sentiment score for each ICT on $rating(R)$

	$rating(R)$							
	TextBlob		**Vader**		**Flair**		**Transformers**	
$sentiment_{lib}^{ICT}(R)$	$Adj. R^2$	β	$Adj. R^2$	β	$Adj. R^2$	β	$Adj. R^2$	β
Electronic Safe Box	0.603	-13.79	-0.092	0.57	0.251	0.50	0.437	0.87
Television	0.186***	1.99	0.251***	1.12	0.313***	0.72	0.368***	1.36
Online Check-in	0.180***	2.13	0.284***	1.33	0.331***	0.83	0.473***	1.69
Air Conditioner	0.176***	2.06	0.280***	1.25	0.287***	0.84	0.391***	1.16
Wifi	0.150***	1.33	0.180***	0.91	0.242***	0.59	0.334***	1.20
Email	0.150***	2.46	0.244***	1.46	0.303***	0.97	0.484***	2.03
USB Charger	0.103	2.17	0.200	1.22	0.465**	1.03	0.212	1.45
Telephone	0.064***	1.46	0.152***	1.05	0.158***	0.71	0.182***	1.14
Social Network	0.049	1.31	0.094	0.97	0.000	0.31	0.055	0.85
Electronic Lock	0.013	1.24	-0.038	0.35	0.060	0.46	0.133	1.11
Website	-0.257	7.91	-0.144	-1.99	0.319	-39.30	-0.160	2.00
ATM	-0.004	0.36	0.038*	0.63	0.008	0.23	0.107**	0.99

Notes: *p<0.005, **p<0.01, ***p<0.001

At first glance, with regard to the values of $adjusted\ R^2$ and β coefficients, it can be noted that the contribution of each ICT component varies from one library to another. Nevertheless, it can be highlighted that regardless of the used sentiment analysis library, the variation of the rating given by the customer is significantly explained by the following six ICT components: "*Television*", "*Online Check-in*", "*Air conditioner*", "*Wi-Fi*", "*Email*" and "*Telephone*".

However, we can observe that there are two exceptions: "*ATM*" and "*USB Charger*". Indeed, "*ATM*" has been identified by Vader and Transformers as an ICT component that contributes to rating. And "*USB Charger*" has been identified only by Flair as the ICT component that surprisingly contributes the most to rating ($adjusted\ R^2$ = 0.465, β = 1.03, p-value< 0.01).

In conclusion, we can say that among the twelve ICTs we initially considered, the sentiment towards only six of them, namely "*Television*", "*Online Check-in*", "*Air conditioner*", "*Wi-Fi*", "*Email*" and "*Telephone*", was found to have a significant impact on the rating. This is supported by the use of four different sentiment analysis libraries. In addition, sentiment towards two other ICTs namely "*ATM*" and "*USB Charger*" was found to contribute to the rating. But this was not supported by all sentiment analysis libraries.

In order to better understand and discuss the reasons why the ICTs mentioned above contribute positively to the rating, we thought it would be useful to visualize them according to their operational domain in the hotel as is commonly done in the literature (Siguaw et al., 2000; Sigala, 2003; Ham et al., 2005; Beldona & Cobanoglu, 2007; Sirirak et al., 2011; Chevers & Spencer, 2017). The ICTs we have chosen for this study belong to two domains, namely in-room and front office. However, some ICTs such as "*Wi-Fi*", "*Telephone*", and "*USB charger*" do not belong exclusively to one of the operational domains but can straddle both.

Figure 5 illustrates the distribution of ICTs and highlights those that contribute significantly to the rating.

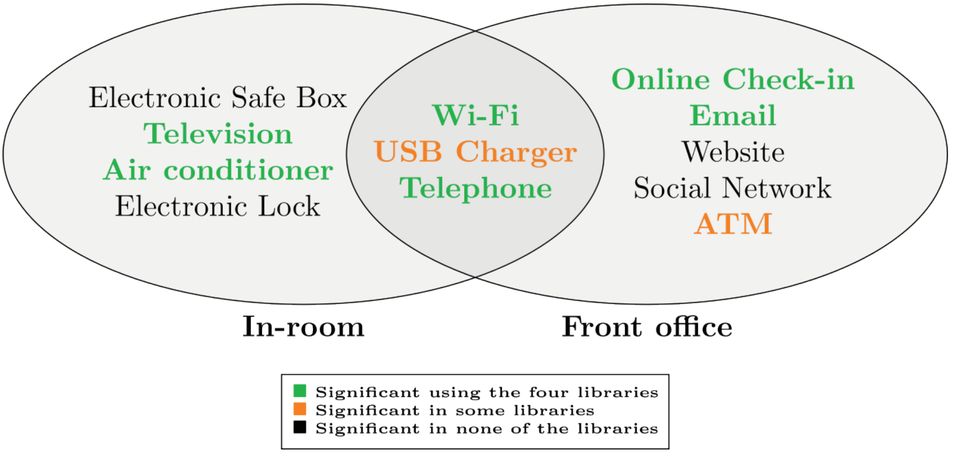

Figure 5. ICTs contribution by operational domains.

We can observe from Figure 5 that among the ICTs belonging to the front office, only those related to booking, namely "*Online Check-in*", "*Email*" and "*Telephone*", contribute significantly to the rating according to the four sentiment analysis libraries used. On the other hand, "*ATM*" was also found to contribute to the rating but only in two of the four sentiment analysis libraries. Although it is difficult to extrapolate on the reasons that may have led to these results, we suppose that even though online payment has been recently adopted in Algeria, its implementation is progressing slowly. This forces customers to book their rooms through traditional means such as phone, email or cash payment using money withdrawn from ATMs available in the lobby.

Regarding in-room ICTs, it was surprising to see that those related to safety and security ("*Electronic Safe Box*" and "*Electronic Lock*") did not contribute significantly to the rating. Again, this can be explained by reasons related to the sample profile. Indeed, premium category hotels are known to be blameless in terms of security and safety. We suppose that for this reason ICTs related to safety and security were not the central topics of the majority of hotel reviews. In fact, only the ICTs related to comfort and entertainment such as "*Air Conditioner*", "*Television*", "*Wi-Fi*" and "*Telephone*" were found to contribute positively to the rating. Moreover, "*USB charger*" which is an ICT available in both rooms and lobbies was found to contribute significantly to the rating by only one of the four sentiment analysis libraries. This can be explained by the fact that the majority of online hotel reviews were written by business guests. They are likely to bring with them smart devices such as laptops, tablets and smartphones that need to be charged at least once or twice a day.

Conclusion, Implications and Limitations

This study contributes to the tourism and hospitality literature by investigating the impact of information and communication technologies on hotel customer satisfaction in a developing country, Algeria.

Our approach is based on the use of innovative sentiment analysis techniques from natural language processing to extract knowledge from online hotel reviews collected on TripAdvisor, one of the most used travel platforms. Unlike past studies proposed in the literature, we used four different sentiment analysis libraries. Two of these libraries are lexicon-based, and the other two are machine learning-based. Our choice to use multiple sentiment analysis libraries was motivated by the goal of mitigating the threat to the internal validity of the study. Although the four sentiment analysis libraries we used differed in terms of accuracy and precision, the results obtained were consistent with respect to the research questions we initially posed. Indeed, the results we obtained indicate that information and communication technologies have a positive impact on the rating given by the user and therefore contribute to customer satisfaction. Nevertheless, this impact is two times smaller than that of other non-ICT services. Furthermore, we also identified which ICTs contribute to the rating and found that ICTs related to room booking, comfort and entertainment were the ones that contributed significantly to the users' rating. It was also surprising to find that ICTs related to safety and security did not contribute significantly to the rating.

Theoretical Implications

In terms of theoretical implications, our study mainly contributes to the existing body of knowledge on the assessment of the impact of ICT adoption in the hotel sector in developing countries.

To the best of our knowledge, our study is one of the few to focus on the relationship between satisfaction with hotel ICTs and overall guest satisfaction in Algerian hotels. We hope that it can provide a solid foundation for future research on ICTs and their impact in developing countries in general and in Algeria in particular. Furthermore, the approach we followed and the way we orchestrated the different natural language processing libraries to process the data from online hotel reviews can be generalized and serve as a framework that may be useful for academics and researchers conducting further research regarding the evaluation of other factors that may contribute to the rating. In addition, researchers can use our approach in other emerging contexts in developing countries, such as rental of homes and apartments, private rooms, and other properties, as well as to assess the impact of ICTs on other customer segments.

Managerial Implications

The results of this research have managerial implications for hotel managers as well as for policymakers.

Due to their open structure, online hotel reviews are a reliable source that can reflect the drivers of satisfaction and dissatisfaction of hotel guests (Ye et al., 2014). However, as the number of online hotel reviews increases, it becomes more difficult for hotel managers to identify trends that can explain customer satisfaction (Calheiros et al., 2017). The approach adopted by our study addresses this problem by using innovative natural language processing techniques.

Therefore, the findings of this research can provide very useful insights to hotel managers in Algeria about the preferences of their guests. In addition, they can help managers to better understand the expectations of guests in terms of ICTs and their operational domains. Furthermore, the findings obtained can also be exploited by hotel managers of other developing countries that have the same tourism potential as Algeria. For example, the results indicate that ICTs related to booking such as telephone and email contribute significantly to customer satisfaction, especially in a context where electronic payment is not yet offered by the majority of hotels. This brings us to the managerial implications for policymakers. Indeed, the Algerian government should make the widespread adoption of electronic payment in all economic sectors, and particularly in the tourism sector, a priority.

Limitations and Future Research

Despite the interesting findings of our study, we can note some limitations related mainly to the profile of the sample as well as the measurement instruments we used.

Algeria is a country where the majority of tourist attractions and hotels are concentrated in the north of the country and operate in a context steeped in cultural and sometimes religious traditions specific to this country. Therefore, the results should be used with caution, taking into account the contextual aspects of the country. In addition, due to the low number of online hotel reviews in budget hotels, the profile of the sample we studied came from 3, 4 and 5-star hotels only and it would be inappropriate to conclude that these results are also valid for other hotel categories or other types of accommodation. Furthermore, it is important to point out that more than half of the online hotel reviews collected were written by business customers whose needs differ from those of customers in other segments (Zhang et al., 2019; Kim et al., 2020). Moreover, future studies need to be conducted to include other ICT components, such as teleworking technologies that became almost the norm during the COVID-19 pandemic. Finally, the textual nature of online hotel reviews as well as the languages in which they are expressed can be a limit to the findings because on the one hand sentiment analysis remains a research-intensive field where it is still a challenge for the machine to understand some of the subtleties related to natural language such as irony, humor and sarcasm. Also, it is important to note that many online hotel reviews are not well written and contain grammatical or punctuation errors.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The authors have no conflict of interest to declare.
Grant Support: The authors declared that this study has received no financial support.
Author Contributions: Conception/Design of study: S.C., A.K., A.A., S.K.; Data Acquisition: S.C., A.K., A.A., S.K.; Data Analysis/Interpretation: S.C., A.K., A.A., S.K.; Drafting Manuscript: S.C., A.K., A.A., S.K.; Critical Revision of Manuscript: S.C., A.K., A.A., S.K.; Final Approval and Accountability: S.C., A.K., A.A., S.K.

References

Adebanjo, D., Teh, P.-L., & Ahmed, P. K. (2016). The impact of external pressure and sustainable management practices on manufacturing performance and environmental outcomes. International Journal of Operations & Production Management, 36 , 995–1013.

Ahani, A., Nilashi, M., Yadegaridehkordi, E., Sanzogni, L., Tarik, A. R., Knox, K., Samad, S., & Ibrahim, O. (2019). Revealing customers' satisfaction and preferences through online review analysis: The case of canary islands hotels. Journal of Retailing and Consumer Services, 51, 331–343.

Ahmad, S. Z., Bakar, A. R. A., Faziharudean, T. M., & Zaki, K. A. M. (2015). An empirical study of factors affecting e-commerce adoption among small- and medium-sized enterprises in a developing country: Evidence from Malaysia. Information Technology for Development, 21, 555–572.

Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., & Vollgraf, R. (2019). Flair: An easy-to-use framework for state-of-the-art NLP. In NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations) (pp. 54

Alrawadieh, Z., Alrawadieh, Z., & Cetin, G. (2021). Digital transformation and revenue management: Evidence from the hotel industry. Tourism Economics, 27(2), 328–345.

Beldona, S., & Cobanoglu, C. (2007). Importance-performance analysis of guest technologies in the lodging industry. Cornell Hotel and Restaurant Administration Quarterly, 48, 299–312.

Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: Text mining of online hotel reviews. Journal of Hospitality Marketing & Management, 25, 1–24.

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. "O'Reilly Media, Inc.".

Bulchand-Gidumal, J., Meli´an-Gonz´alez, S., & L´opez-Valc´arcel, B. G. (2011). Improving hotel ratings by offering free wi-fi. Journal of Hospitality and Tourism Technology, 2, 235–245.

Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. Journal of Hospitality Marketing & Management, 26, 675–693.

Chevers, D. A., & Spencer, A. J. (2017). Customer satisfaction in Jamaican hotels through the use of information and communication technology. Worldwide Hospitality and Tourism Themes, 9, 70–85. –59).

Chua, A. Y., & Banerjee, S. (2016). Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. Computers in Human Behavior, 54, 547–554.

Cobanoglu, C., Berezina, K., Kasavana, M. L., & Erdem, M. (2011). The impact of technology amenities on hotel guest overall satisfaction. Journal of Quality Assurance in Hospitality & Tourism, 12, 272 – 288.

Davras, O., & Caber, M. (2019). Analysis of hotel services by their symmetric and asymmetric effects on overall customer satisfaction: A comparison of market segments. International Journal of Hospitality Management, 81, 83–93.

Ezzaouia, I., & Bulchand-Gidumal, J. (2020). Factors influencing the adoption of information technology in the hotel industry. an analysis in a developing country. Tourism Management Perspectives, 34, 100675.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. Tourism Management, 52, 498–506.

Gao, B., Li, X., Liu, S., & Fang, D. (2018). How power distance affects online hotel ratings: The positive moderating roles of hotel chain and reviewers' travel experience. Tourism Management, 65, 176–186.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels - an empirical analysis. Tourism Management, 61, 43–54.

Ham, S., Gon Kim, W., & Jeong, S. (2005). Effect of information technology on performance in upscale hotels. International Journal of Hospitality Management, 24, 281–294.

He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: a case of analyzing online hotel reviews. Online Inf. Rev., 41, 921–935.

HuggingFace (2021a). Translation Models - Hugging Face huggingface.co. https://huggingface.co/models?pipeline_tag=translation&sort=downloads. [Online; accessed 21-November-2021].

HuggingFace (2021b). Zero-shot Classification Models - Hugging Face huggingface.co. https://huggingface.co/models?pipeline_tag=zero-shot-classification&sort=downloads. [Online; accessed 21-November-2021].

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In ICWSM.

Khoo-Lattimore, C., Mura, P., & Yung, R. (2019). The time has come: a systematic literature review of mixed methods research in tourism. Current Issues in Tourism, 22, 1531–1550.

Kim, D., Hong, S., Park, B.-J., & Kim, I. (2020). Understanding heterogeneous preferences of hotel choice attributes: Do customer segments matter? Journal of Hospitality and Tourism Management, 45, 330–337.

Kim, W. G., & Ham, S. (2006). The impact of information technology implementation on service quality in the hotel industry. Information Technology in Hospitality, 4, 143–151.

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). BART: denoising sequence-to-sequence 585 pre-training for natural language generation, translation, and comprehension. CoRR, abs/1910.13461.

Li, H., Ye, Q., & Law, R. (2013). Determinants of customer satisfaction in the hotel industry: An application of online review analysis. Asia Pacific Journal of Tourism Research, 18, 784–802.

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

Loria, S. (2018). textblob documentation. Release 0.15, 2, 269.

Lu, W., & Stepchenkova, S. (2012). Ecotourism experiences reported online: Classification of satisfaction attributes. Tourism Management, 33, 702–712.

Melian-Gonzalez, S., & Bulchand-Gidumal, J. (2016). A model that connects information technology and hotel performance. Tourism Management, 53, 30–37.

Mihalic, T., & Buhalis, D. (2013). ICT as a new competitive advantage factor - case of small transitional hotel sector. Economic and business review, 15, 33–56.

Moliner-Velazquez, B., Fuentes-Blasco, M. and Gil-Saura, I. (2019), "The role of ICT, eWOM and guest characteristics in loyalty", Journal of Hospitality and Tourism Technology, Vol. 10 No. 2, pp. 153-168.

Nunkoo, R., Teeroovengadum, V., Ringle, C. M., & Sunnassee, V. (2020). Service quality and customer satisfaction: The moderating effects of hotel star rating. International Journal of Hospitality Management, 91, 102414.

Nusair, K. K., Bilgihan, A., & Okumus, F. (2013). The role of online social network travel websites in creating social interaction for gen y travelers. International Journal of Tourism Research, 15, 458–472.

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

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21, 1–67.

Ruan, Y. (2020). Perceived host-guest sociability similarity and participants' satisfaction: Perspectives of airbnb guests and hosts. Journal of Hospitality and Tourism Management, 45, 419–428.

Ryssel, R., Ritter, T., & Gemünden, H. G. (2004). The impact of information technology deployment on trust, commitment and value creation in business relationships. Journal of Business & Industrial Marketing, 19, 197–207.

Shin, S., Du, Q., Ma, Y., Fan, W., & Xiang, Z. (2021). Moderating effects of rating on text and helpfulness in online hotel reviews: an analytical approach. Journal of Hospitality Marketing & Management, 30, 159–177.

Sigala, M. (2003). The information and communication technologies productivity impact on the UK hotel sector. International Journal of Operations & Production Management, 23, 1224–1245.

Siguaw, J. A., Enz, C. A., & Namasivayam, K. (2000). Adoption of information technology in U.S. hotels: Strategically driven objectives. Journal of Travel Research, 39, 192–201.

Sirirak, S., Islam, N., & Khang, D. B. (2011). Does ICT adoption enhance hotel performance? Journal of Hospitality and Tourism Technology, 2, 34–49.

Statista (2021). Tripadvisor: number of reviews 2020 — Statista. https://www.statista.com/statistics/684862/tripadvisor-number-of-reviews/. [Online; accessed 21-November-2021].

Velazquez, B. M., Blasco, M. F., & Saura, I. G. (2015). ICT adoption in hotels and electronic word-of-mouth. Academia-revista Latinoamericana De Administracion, 28, 227–250.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. L., Gugger, S., Drame, M., Lhoest, Q., & Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (pp. 38–45). Online: Association for Computational Linguistics.

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? International Journal of Hospitality Management, 44, 120–130.

Xu, X., & Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. International Journal of Hospitality Management, 55, 57–69.

Xu, X., Li, Y., & Lu, A. C. C. (2019a). A comparative study of the determinants of business and leisure travellers' satisfaction and dissatisfaction. International Journal of Services and Operations Management, 33, 87–112.

Xu, X., Liu, W., & Gursoy, D. (2019b). The impacts of service failure and recovery efforts on airline customers' emotions and satisfaction. Journal of Travel Research, 58, 1034–1051.

Ye, Q., Li, H., Wang, Z., & Law, R. (2014). The influence of hotel price on perceived service quality and value in e-tourism: An empirical investigation based on online traveler reviews. Journal of Hospitality & Tourism Research, 38, 23–39.

Yin, W., Hay, J., & Roth, D. (2019). Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach. CoRR, abs/1909.00161.

Zaied, A. N. H. (2012). Barriers to e-commerce adoption in Egyptian SMEs. International Journal of Information Engineering and Electronic Business, 4, 9–18.

Zhang, T., Seo, S., & Ahn, J. A. (2019). Why hotel guests go mobile? examining motives of business and leisure travelers. Journal of Hospitality Marketing & Management, 28, 621–644.

Zhang, X., Yu, Y., Li, H., & Lin, Z. (2016). Sentimental interplay between structured and unstructured user-generated contents: An empirical study on online hotel reviews. Online Inf. Rev., 40, 119–145.

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