Sentiment drivers of hotel customers: a hybrid approach using unstructured data from online reviews

Ree Chan Ho
Taylor's University, Subang Jaya, Malaysia
Madusha Sandamali Withanage
University of Nottingham–Malaysia Campus, Semenyih, Malaysia, and
Kok Wei Khong
Taylor's University, Subang Jaya, Malaysia

Received 23 September 2019
Revised 26 February 2020
14 May 2020
26 September 2020
Accepted 29 September 2020

Abstract

Purpose – With the growth of social media and online communications, consumers are becoming more informed about hotels' services than ever before. They are writing online review to share their experiences, as well as reading online review before making a hotel reservation. Hotel customers considered it as reliable source and it influences customers' hotel selection. Most of these reviews reside in unstructured format, scattered across in the Internet and inherently unorganized. The purpose of this study was to use predictive text analytics to identify sentiment drivers from unstructured online reviews.

Design/methodology/approach – The research used sentiment classifications to analyze customers' reviews on hotels from TripAdvisor. In total, 9,286 written reviews by hotel customers were scrapped from 442 hotels in Malaysia. A detailed text analytic was conducted and was followed by a development of a theoretical framework based on the hybrid approach. AMOS was used to analyze the relationship between customer sentiments and overall review rating.

Findings – With the use of Structural Equation Modeling (SEM) and clustering technique, a list of sentiment drivers was detected, i.e. location, room, service, sleep, value for money and cleanliness. Among these variables, service quality and room facilities emerged as the most influential factors. Sentiment drivers obtained in this study provided the insights to hotel operators to improve the hotel conditions.

Research limitations/implications – Although this study extended the existing literature on sentiment analysis by providing valuable insights to hoteliers, it is not without its limitations. For instance, online hotel reviews collected for this study were limited to one specific online review platform. Despite the large sample size to support and justify the findings, the generalizability power was restricted. Thus, future research should also consider and expand to other type of online review channels. Therefore, a need to examine these data reside various social media applications, i.e. Facebook, Instagram and YouTube.

Practical implications – This study highlights the significance of hybrid predictive model in analyzing the unstructured hotel reviews. Based on the hybrid predictive model we developed, six sentiment drivers emerged from the data analysis, i.e. location, service quality, value for money, sleep quality, room design and cleanliness. This consideration is critical due to the ever-increasing unstructured data resides in the online space. This explores the possibility of applying data analytic technique in a more efficient manner to obtain customer insights for hotel managerial consideration.

Originality/value – This study analyzed customer sentiments toward the hotel in Malaysia with the use of predictive text analytics technique. The main contribution was the list of sentiment drivers and the insights needed to improve the hotel conditions in Malaysia. In addition, the findings demonstrated motivating findings from different methodological perspective and provided hoteliers with the recommendation for improved review ratings.

Keywords Online review, Unstructured data, Text analytics, Hotel sentiment drivers, Hybrid approach, Predictive analytics

Paper type Research paper

1. Introduction

With the growth of social media and online communications, consumers are becoming more informed about products and brands ever before (Schivinski and Dabrowski, 2016; Ho and

Asia-Pacific Journal of Business Administration

1757-4323
DOI 10.1108/APJBA-09-2019-0192

Rajandram, 2016). Due to the Internet's highly pervasive nature, hotel customers are inclined to use online and social media to share travel information and experiences (Magno *et al.*, 2018; Ho and Amin, 2019). Viglia *et al.* (2016) confirmed that e-word of mouth is a trusted source of accessing reliable and useful information for hotel selection. Furthermore, customers treat online reviews as important compared to personal references or recommendations (Ladhari and Michaud, 2015). Online reviews were more reliable than information provided by hotels' communication channels, i.e. Website and social media accounts. Online customer reviews are progressively influencing the choice of hotel (Vermeulen and Seegers, 2009, Noone and Mcguire, 2016; Gavilan *et al.*, 2018). Hence, customer reviews are referred and taken into consideration before the final booking decision via third-party hotel booking sites, online travel agencies and other booking intermediaries or hotels' sites directly (Geetha *et al.*, 2017; De Pelsmacker *et al.*, 2018; Park *et al.*, 2018).

Both structured and unstructured are varied in nature and the ways they were organized. Structured data from traditional data techniques are stored in easily accessible databases. Most online reviews are unstructured generated from a myriad of business channels (Liu *et al.*, 2016; Singh *et al.*, 2017; Park *et al.*, 2018). These unstructured contents are progressively gained from social media and smart devices (Ho and Rajadurai, 2020); hence, resulted in ambiguity and uniformity of the outcomes gained (Zhang *et al.*, 2016). Therefore, they need to be managed differently. Since the majority of information is presented in the unstructured format, businesses must convert it for decision making.

Massive quantity of unstructured online reviews are shared and scattered across social media, forums and, other platforms, hence this warrants the need to organize them properly. It takes a long time in converting them to be useful for hotel management. A need to analyze this source of data is critical to understand customer sentiment toward the hotels. Data mining converts unstructured information into structured datasets that can be analyzed to reveal trends, patterns and insights (Chen *et al.*, 2015; Salehan *et al.*, 2016). It can process substantial amounts of data faster (Hung, 2012; Lazard *et al.*, 2015). Furthermore, data mining reveals hidden connections between information and uncovers business insights previously not possible with traditional tools (Torii *et al.*, 2016; Kang *et al.*, 2018). The advancement of data mining techniques has been applied in many studies, i.e. health issues (Torii *et al.*, 2016), new product development (Xu *et al.*, 2016) and e-commerce logistics (Wu and Lin, 2018). Similarly, the data analytics approach was used in the hotel industry. These studies investigated online reviews and their relationship with hotel financial performance (Xie *et al.*, 2017), guest experience (Xiang *et al.*, 2015) and service quality (He *et al.*, 2018).

Existing studies have examined several rewards gained from the sentiments of the online review. Sentiments from customer review promote brand and the sale performance (Öğüt and Onur, 2012; Lee *et al.*, 2017; Xia *et al.*, 2019). Combining with other data at firms' hands, customer sentiment analysis allows firms to learn more about customer preference and consumer behavior (Li *et al.*, 2019). This enables better marketing intelligence to compete with other brands. In this context, scholars also recognized that customers' sentiments play a significant role in predicting the booking reservation for hotels (Chong *et al.*, 2016; Zhao and Wang, 2019).

The potential to gain more insights with the use of text mining in analyzing the massive amount of online reviews could provide higher predictive power. It is observed that several research works on online review were emphasized on the hotel satisfaction after the patronage (Min *et al.*, 2015; De Pelsmacker *et al.*, 2018; Yang *et al.*, 2018). These researches were not focused on motivational factors relating to the intended selection of the future hotel. Hence, a need to develop a model for hotel sentiments based on the unstructured data scraped from the online travel websites. The purpose of this study was to develop a novel data-driven approach to identify sentiment drivers of online hotel reviews in Malaysia. With that, the following two research questions were developed. What is the sentiment of factors that influence the customer to produce the online review? Second, what are the inter-relationships among the identified sentiment drivers and their impacts on hotel selection?

The major contribution of this study is to identify the drivers derived from online customer reviews leading to the hotel booking. First, we discovered the sentiment drivers in the selection of hotels with the use of text mining techniques. These new sentiment drivers explain the consumer choice in the extensive unstructured datasets and provide higher generalizability. Second, the empirical findings from our study deduce the need to emphasize on online customer reviews that were ignored in prior studies. Third, this study identified the key motivators based on customers' perceptions toward their behavioral intention in choosing the future hotel patronage. This study would provide useful insights and recommendations to hoteliers by examining the sentiments collected. This consideration sheds light on the ever-increasing online customer reviews in the hotel and hospitality sector.

2. Literature review

2.1 Hotel online review

Customers are flooded with information about products and brands before purchase due to the growth of social media and online communications (Ho, 2020, pp. 1–19). Online customer review is regarded as the popular source of product and brand information for customers (King *et al.*, 2014). The evaluation of products generated by customers based on personal usage or purchase experience (Liu and Park, 2015). Hotel customers are writing their experiences and reading online reviews before making a hotel reservation. It is common for customers to read online reviews before clicking on brands' sites. It has attracted the attention of the hotel and hospitality industry and no dearth of research works on the impact of hotel customer reviews on their choice of hotels (Herrero *et al.*, 2015; Flavián *et al.*, 2016). Furthermore, Gil-Soto *et al.* (2019) confirmed that e-WOM is a trusted source of accessing reliable and useful information for hotel selection. Hence, customers treat online reviews as important to personal references or recommendations. Since online reviews are distributed across the Internet widely, consumers are facing the challenge of bombarded with incomplete and distorted hotel reviews.

2.2 Hotel customer sentiment

Customer sentiments are customers' feelings communicated via the online review (Geetha and Sinha, 2017). It can either be positive, negative, or neutral sentiment. The sentiments of customers provide the information ques about the specific features of the product reviewed (Liu and Park, 2015). Therefore, sentiment drivers are recurring variables that motivate the consumer to produce online reviews (Xie *et al.*, 2016; Park, 2019).

Hoteliers were urged to be aware of the impact of sentiment from reviews on hotel customers' choice of hotel (Gavilan *et al.*, 2018). They must measure the reliability of the reviews posted by the consumer (Noone and Mcguire, 2016). Lo and Yao (2019) asserted that the positive sentiments have direct linkage to attract the consumers to visit their websites and high chance to convert them into sales. Hence, hotel operators must have an accurate and full picture of the sentiments vented by customers via various customer review platforms.

Previous studies have identified the role of hotel customer sentiments in improving the hotel sale performance (Kim, 2013; Uddin, 2015; Lee *et al.*, 2017). Positive sentiments over the hotel facilities improved the customer rating and increase the sale for the hotel (Salehan and Kim, 2016; Geetha *et al.*, 2017). For example, customers assess the quality of services by evaluating their perception of the service delivery against the service they received (Uddin, 2015). Also, Dube and Renaghan (2000) and Tsai *et al.* (2011) had further found that friendliness of the staff also contributes to hotel attributes considered by guests.

Scholars also recognized that customers' sentiments play a significant role in predicting hotel booking reservations (Ogut and Onur, 2012; Zhou *et al.*, 2014; Gavilan *et al.*, 2018). Kim (2013) revealed that promptness and courteousness are the two key sentiment attributes when selecting a hotel. However, Lau *et al.* (2005)'s research had shown that the Malaysian hoteliers failed to meet the expectations of the customers in terms of quality of service. Also, prior hotel studies have found that customers' sentiment over the location is important in the choice of hotels (Baruca and Civre, 2012; Dubé and Renaghan, 2000). This is consistent with Abdullah and Haan (2012)'s study conducted on hotels in Malaysia, revealed that 66.4% of the hotel customers strongly agreed that location matter the most for their selection of hotels. Therefore, location is a crucial factor leading to positive customer sentiments.

Hotel customers' sentiments have been the precursor of hotel customer retention (Liang *et al.*, 2015; Chong *et al.*, 2016; Park *et al.*, 2018). In this context, Bilgihan *et al.* (2015) validated customers' intention to book the same hotel depend on their previous experiences. Gu and Ryan (2008) revealed that cleanliness, comfortable mattresses and towel quality were among the key attributes rated by the hotel guests. Moreover, Ahmad *et al.* (2008) study on the Malaysian hotels had confirmed that the cleanliness of the room was an important factor for hotel guests when choosing a particular hotel. Also, Ahmad *et al.* (2008) further identified that the room layout, quietness and the surrounding of the room also contribute to the choice of a hotel. However, these factors depend on the demographic profiles of the customers. Specifically, Ruys and Wei (1998) revealed that the comfortability of the bed mattress was important for senior travelers.

Customer sentiments have been identified as one main criterion in developing relevant marketing intelligence (Li *et al.*, 2019). For example, Zhao *et al.* (2019) highlighted that customer sentiment was a major source in predicting hotel customer satisfaction. Consumers' perceived service value determines consumers' willingness to pay for a particular service. In general, price plays a crucial role in the value and quality of the products (Mattila and O'Neill, 2003). Similarly, Dubé and Renaghan (2000) revealed that customer sentiment on value for money when choosing a hotel. This is consistent with Atkinson (1988)'s assertion on customer perceived value which affects their decisions. A study conducted on 3-star hotels in Malaysia verified the direct relationship between customers' perceived value and customer satisfaction (Amirreza *et al.*, 2013).

3. Research approach

The main purpose of this study was to collect unstructured dataset from online forums. Therefore, a hybrid approach combining both inductive and deductive reasoning was appropriate. This approach permitted us to draw the knowledge from the text corpus in the customer reviews. Complex reasoning on large dataset were conducted in many online and social media studies (Zeng *et al.*, 2010; Barbieri *et al.*, 2010). We collected unstructured data in hotel reviews from 442 hotels located across 12 states in Malaysia, except Perlis. The key objective was to use sentimental classification to analyze customers' opinions about the hotel industry in Malaysia.

4. Text mining and data analytics techniques

Specifically, we used a descriptive mining method to identify the sentiment drivers in the Malaysian hotel industry. Based on the works of Chen *et al.* (2015) and Torii *et al.* (2016), the data of this study were acquired from the online forums of TripAdvisor. We selected TripAdvisor, a renowned and popular online customer review portal due to its impact on consumer choice in hospitality studies (Xiang *et al.*, 2017). A total of 9,286 hotel reviews and comments were extracted with the use of Octoparse, a cloud-based web crawler software program. Following that, several software tools were used, i.e. SAS Enterprise Miner to conduct the text analytics, Microsoft Power BI to compute the sentiment scores and SPSS AMOS to perform the structured equation modeling.

In particular, we used SAS Enterprise Miner to perform text analytics. The entire process was carried out in the following steps: data parsing, followed by data filtering and finally clustered the data into groups. These computational expansive steps were needed to expose the terms relating to the hotel reviews. Text parsing node tool was used to decompose textual data and to generate a quantitative representation that was suitable for data mining purposes. Hence, we were able to exclude the stop-word in the data, i.e. prepositions, pronouns and auxiliary verbs during the text parsing process. Text filtering was conducted to reduce the total number of parsed terms that were analyzed by the Text Miner module. By doing so, we were able to maintain the related terms that were displayed by the frequency of occurrence in the dataset. Therefore, the measurement quality of the data was improved with the high signal to noise ratio (Kim *et al.*, 2017). Following that, text cluster node was used to perform the cluster analysis on the dataset. This was used to identify the sentiment drivers of online hotel reviews from the scrapped data. The clusters detected were organized based on the frequency of the data collection. Each cluster was assigned with a randomly generated Cluster ID. Table 1 displays a summary of the clusters extracted.

Subsequently, SPSS AMOS was used to analyze the data extracted from the reviews. The customer sentiments scores were calculated by the Microsoft Power BI. We used SEM in this research as it could examine the complex and multi-faceted behavioral concepts derived from this study (Petrescu, 2013). Also, it could estimate a series of separate, but interdependent, multiple regression equations simultaneously in a structural model. Therefore, SEM was useful to estimate the underlying relationship among the three parameters in this study, i.e. sentiment drivers, online hotel review ratings and customer sentiments.

5. Data analysis and findings

SAS Text Miner software was based on top-down approach and clusters were identified from clusters detected earlier. Out of the 14 clusters, only 7 clusters were found to be significant. Text analysis was conducted on these 7 clusters to identify the sentiment drivers (Table 2). For example, terms encountered in clusters such as “bathroom”, “water” and “shower”, revealed that bathroom facilities were important to customers. Among these sentiment drivers, the term “bed” was recurring in cluster 6 and lead to the new element of “sleep quality”. Based on the text analysis, sentiment drivers of online hotel reviews were identified, i.e. location (distance from the main mode of transportations and city venues), value for money, room quality and facilities (room design, bathroom facilities, help desk, etc.), sleep quality (bed), hotel facilities (pool, car parking, etc.), cleanliness and service quality.

Following that, SPSS AMOS was used to conduct the SEM analysis to test the relationships among sentiment drivers, online hotel review ratings and customer sentiments, which formed the conceptual framework of this study. As mentioned earlier, the SEM module could estimate the relationships among the variables after taking account of the errors ($e1$ and $e2$) as shown in Figure 1. It rectified these errors when establishing the causal relationship among the variables. Also, the squared multiple correlations for review (0.691) and sentiment (0.264) obtained from AMOS, illustrate that 69% of the review ratings explained 26% of the customer sentiments. These results confirmed and established the causal relationship among the sentiment drivers, online hotel review ratings and customer sentiments.

Table 3 shows the variances among the customer review and the list of sentiment drivers. Online review differs mostly with service quality (1.427) followed by value for money (1.226), sleep quality (1.214), room facilities and quality (1.133) and location (0.918).

Table 4 presents the result of the regression analysis conducted. The highest coefficient recorded was service, followed by room, value for money, sleep and cleanliness, and location has the minimum impact on the user behavior.

Table 1. Summary of clusters extracted for online hotel reviews in Malaysia

Cluster-ID	Cluster description	Frequency	Percentage
1		5,572	24%
4	+ price + pool parking spacious + service + place + good comfortable + friendly + food + area + breakfast	2,981	13%
2	+ bed + walk + great + night + clean + water + time + small stayed helpful + location + room	2,591	11%
3	+ bed + great + walk + staff + night + time stayed helpful rooms + bathroom + stay + small	1,846	8%
7	+ pool + price + service + family + well + breakfast spacious + bathroom + place rooms + good + room	1,572	7%
9	+ pool + service + family + price + breakfast spacious + check + room + well + good + bathroom + place	1,572	7%
11	parking walking + distance "walking distance" + area + food nearby + location located + price comfortable + friendly	1,409	6%
6	+ bed + shower + night + time + check + back + bathroom stayed + water + stay + room + family	1,091	5%
15	+ hotel + good + clean + staff + friendly + nice + location + night + price + great comfortable money	892	4%
10	+ walk + station kl located + bus + city + great + location helpful + clean + hotel + staff	755	3%
5	+ mall shopping + distance nearby "walking distance" walking located + clean Wi-Fi + location + walk + food	745	3%
13	parking + car + park + tv + free located + shower walking + distance spacious "walking distance" + small	705	3%
16	+ food walking + distance "walking distance" + restaurant helpful + friendly + staff + place + service + location + area	704	3%
14	+ apartment + bedroom + desk + front + water + floor + family + service + well + bathroom + food + pool	680	3%

Table 2. Sentiment drivers of online hotel reviews in Malaysia

Hierarchy level	Cluster-ID	Parent	Cluster description	Frequency	Sentiment driver
4	6	3	+ bed + shower + night + time + check + back + bathroom stayed + water + stay + room + family	1,091	Room quality and facility, sleep quality
4	15	3	+ hotel + good + clean + staff + friendly + nice + location + night + price + great comfortable money	892	Staff service, cleanliness and value for money
4	10	7	+ walk + station kl located + bus + city + great + location helpful + clean + hotel + staff	755	Location
4	5	11	+ mall shopping + distance nearby "walking distance" walking located + clean wifi + location + walk + food	745	Location
4	13	11	parking + car + park + tv + free located + shower walking + distance spacious "walking distance" + small	705	Location and hotel facilities
5	16	9	+ food walking + distance "walking distance" + restaurant helpful + friendly + staff + place + service + location + area	704	Location and staff service
5	14	9	+ apartment + bedroom + desk + front + water + floor + family + service + well + bathroom + food + pool	680	Facilities and design of the apartment

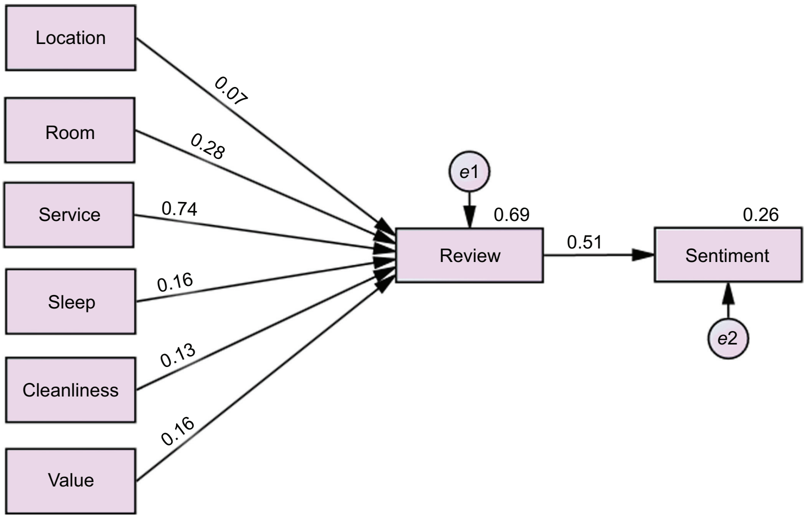

Figure 1.
List of sentiment drivers

Table 3.
Variances of sentiment drivers obtained from SPSS AMOS

Path	Estimate	S.E.	C.R.	*p*
Location	0.918	0.028	33.302	0.000*
Room	1.133	0.034	33.816	0.000*
Service	1.427	0.027	53.539	0.000*
Sleep	1.214	0.037	32.738	0.000*
Value	1.226	0.036	33.764	0.000*
Cleanliness	1.173	0.035	33.726	0.000*
*e*1	0.300	0.009	34.211	0.000*
*e*2	0.098	0.001	68.136	0.000*

Table 4.
Unstandardized estimates of sentiment drivers obtained from SPSS AMOS

Path	Estimate	S.E.	C.R.	*p*-value
Location → Review	0.076	0.014	5.611	0.000*
Room → Review	0.261	0.011	24.837	0.000*
Service → Review	0.607	0.007	92.950	0.000*
Value → Review	0.140	0.012	12.146	0.000*
Sleep → Review	0.142	0.011	12.758	0.000*
Cleanliness → Review	0.114	0.012	9.880	0.000*
Review → Sentiment	0.190	0.003	57.735	0.000*

Note(s): * Denotes regression weights were significant at 0.05

Regression analysis was conducted with SPSS AMOS and the estimate, critical ratio and standard error for the sentiment drivers were obtained. Table 5 shows that on average, the customer review ratings were above 3 for all sentiment drivers measured during this study. This showed that the customers were satisfied with the sentiment drivers in general. Specifically, location achieved the highest score and indicated that most customers were satisfied with it. The result also revealed that customers were least satisfied with service quality. This is consistent with the findings of Lau *et al.* (2005) which asserted that Malaysian hotels have failed to satisfy customer expectations in terms of service quality.

6. Conclusion

Based on the hybrid predictive model we developed, six sentiment drivers emerged from the data analysis, i.e. location, service quality, value for money, sleep quality, room design and cleanliness. This study revealed that customers pay particular attention to location convenience. For example, customers have used terms such as "KL located", "bus", "city" and "station" alongside with "walking distance" to describe the location that indicates that customers are interested in finding hotels that are located near transportation modes (i.e. bus and train stations) and near the city for the ease of commuting. Hence, location convenience is a significant factor in the selection of hotels and consistent with past studies (Dubé and Renaghan, 2000).

Secondly, our text analysis revealed that customers pay attention to staff attitude when serving and greeting guests. Customers use terms such as "friendly" and "nice" to describe staff attitude. Service quality was explained in terms of promptness, friendliness and courteousness (Kim, 2013). This finding confirmed that service quality plays a key role in customers' selection of a hotel (Sampaio *et al.*, 2019; Lee and Cheng, 2018). Furthermore, this study revealed that customers use terms such as "price" and "money" often when describing hotel experience. This indicates that customers evaluate hotel experience in terms of "value for money". This is in line with the findings of many scholars in determining the importance of customers' value toward the hotels (Nasution and Mavondo, 2008; Lockyer, 2005).

The term "bed" was a recurring term in some of the clusters, which indicates that the comfortability of the bed is an important attribute to customers as it affects their sleep quality. This result corroborated with (Ruys and Wei, 1998) findings on the bed mattress that served as an important attribute to customer satisfaction. This further indicated that sleep quality is linked directly to hotel customer satisfaction. Also, this study showed that cleanliness is an important consideration for hotel customers and shared the same view with past researchers (Dubé and Renaghan, 2000; Ahmad *et al.*, 2008).

Finally, room design and facilities obtained from clusters examined in this study contained terms such as "bathroom", "water" and "shower" which revealed that bathroom facilities are important to customers. This corroborates with (Ahmad *et al.*, 2008)'s study that revealed room facilities are a crucial criterion in affecting customers' selection of a hotel in Malaysia.

6.1 Theoretical contribution

The importance of unstructured online review was validated as the source of sentiment drivers in influencing customers' choice of hotel in this study. It has three main contributions. First, a list of sentiment drivers was identified as the antecedent for online hotel review. The findings validated that location, service quality, room facilities and quality, cleanliness and value for money are the main sentiment drivers affecting customers' selection of a hotel. Also, sleep quality was a discovery as it has a significant impact on review ratings and sentiments. This indicated that customers pay attention to the comfortability of the bed and the quality of sleep during their stay in the hotel. Earlier research works supported the inclusion of this new

	Estimate	S.E.	C.R.	*p*-value
Room	3.607	0.021	170.788	0.000
Service	3.697	0.015	251.263	0.000
Sleep	3.763	0.023	161.172	0.000
Cleanliness	3.752	0.022	167.239	0.000
Value	3.743	0.023	164.527	0.000
Location	3.948	0.020	194.955	0.000

Table 5. Means of sentiment drivers obtained from SPSS AMOS

element. For example, Ruys and Wei (1998) revealed that the comfortability of the bed mattresses is an important factor for senior travelers. This shed new light for hoteliers to provide the right services to meet the customers' requirements.

Second, the empirical findings from our study deduce the inter-relationship among customer sentiment, online review ratings and hotel selection, which was overlooked in prior studies. The study validated a hybrid predictive model in explaining the implication of customer sentiments for online review. The review ratings act as the mediator in linking the drivers to the customer satisfaction. Finally, this study emphasizes the usefulness of a hybrid predictive model in analyzing the unstructured hotel reviews. This consideration is critical due to the ever-increasing unstructured data resides in the online space. This explores the significance of applying data analytic techniques to obtain customer insights more efficiently.

6.2 Managerial contribution

This study provides valuable insights to hoteliers. The findings indicated that service quality has the highest impact on customer behavior intention. Hence, hotel owners should focus on improving the customer service. Efforts such as training needed to prepare the staff to be friendly, courteous, attentive and ready to provide prompt service (Atkinson, 1988; Buttle and Bok, 1996; Knutson, 1988). This indicated a need to hire staff members who are empathetic and helpful to customers (Barsky and Labagh, 1992). In terms of room quality and facilities, cleanliness of the room and amenities offered could enhance customer satisfaction (Dubé and Renaghan, 2000). Therefore, hotel owners continuously ensure that customers are comfortable with the level of cleanliness in the hotel. Also, the convenience of the hotel location when travelers selecting a hotel. Thus, hoteliers must provide easy access to bus stations, train stations, shopping centers and city centers. One key takeaway is the introduction of sleep quality from this study. The comfort of the bed mattress is crucial in this context and ensuring that night entertainment spots are out of the proximity of the hotel.

6.3 Limitations and future research

Although this study extended the existing literature on sentiment analysis by providing valuable insights to hoteliers, it is not without its limitations. For instance, online hotel reviews collected for this study were limited to one specific online review platform, i.e. TripAdvisor. Despite the large sample size to support and justify the findings, the generalizability power was restricted. Thus, future research should also consider and expand to other types of online review websites and also, the need to examine these unstructured data which reside in various social media applications, i.e. Facebook, Instagram and YouTube. Future works could be replicated in other countries to determine whether the sentiment drivers differ from one country to another. Other dimensions such as switching cost and customer involvement as proposed by Bügel *et al.* (2011)can be considered as a moderator in the hotel customer's view toward hotel selection. Also, other important metrics for measuring hotel performance within the online reviews should be investigated too, i.e. hotel customer loyalty and re-patronage intention (Kandampully and Suhartanto, 2000; Tsai, 2017).

References

Abdullah, A.A. and Haan, M.H. (2012), "Internal success factor of hotel occupancy rate", *International Journal of Business and Social Science*, Vol. 3, pp. 199-218.

Ahmad, M.S., Ariffin, A.A.M. and Ahmad, A. (2008), "Evaluating the services of Malaysian hotels from the leisure traveler's prespactive", *Jurnal Pengurusan (UKM Journal of Management)*, Vol. 27, pp. 69-83.

Amirreza, F., Mohammad, S.Z. and Gilani, M. (2013), "Customer satisfaction in hospitality industry: Middle East tourists at 3star hotels in Malaysia", *Research Journal of Applied Sciences, Engineering and Technology*, Vol. 5, pp. 4329-4335.

Atkinson, A. (1988), "Answering the eternal question: what does the customer want", *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 29, pp. 12-14.

Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Huang, Y., Tresp, V., Rettinger, A. and Wermser, H. (2010), "Deductive and inductive stream reasoning for semantic social media analytics", *IEEE Intelligent Systems*, Vol. 25, pp. 32-41.

Barsky, J.D. and Labagh, R. (1992), "A strategy for customer satisfaction", *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 33, pp. 32-40.

Baruca, P.Z. and Civre, Z. (2012), "How do guests choose a hotel", *Academica Turistica*, Vol. 5, pp. 75-84.

Bilgihan, A., Nusair, K., Okumus, F. and Cobanoglu, C. (2015), "Applying flow theory to booking experiences: an integrated model in an online service context", *Information and Management*, Vol. 52 No. 6, pp. 668-678.

Bügel, M.S., Verhoef, P.C. and Buunk, A.P. (2011), "Customer intimacy and commitment to relationships with firms in five different sectors: preliminary evidence", *Journal of Retailing and Consumer Services*, Vol. 18, pp. 247-258.

Buttle, F. and Bok, B. (1996), "Hotel marketing strategy and the theory of reasoned action", *International Journal of Contemporary Hospitality Management*, Vol. 8, pp. 5-10.

Chen, A.T., Zhu, S.-H. and Conway, M. (2015), "What online communities can tell us about electronic cigarettes and hookah use: a study using text mining and visualization techniques", *Journal of Medical Internet Research*, Vol. 17, e220.

Chong, A., LI, B., Ngai, E., Ch'ng, E. and Lee, F. (2016), "Predicting online product sales via online reviews, sentiments, and promotion strategies: a big data architecture and neural network approach", *International Journal of Operations and Production Management*, Vol. 36 No. 4, pp. 358-383, doi: 10.1108/IJOPM-03-2015-0151.

De Pelsmacker, P., Van Tilburg, S. and Holthof, C. (2018), "Digital marketing strategies, online reviews and hotel performance", *International Journal of Hospitality Management*, Vol. 72, pp. 47-55.

Dubé, L. and Renaghan, L.M. (2000), "Creating visible customer value: how customers view best-practice champions", *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 41, pp. 62-72.

Flavián, C., Gurrea, R. and Orús, C. (2016), "Choice confidence in the webrooming purchase process: the impact of online positive reviews and the motivation to touch", *Journal of Consumer Behaviour*, Vol. 15, pp. 459-476.

Gavilan, D., Avello, M. and Martinez-Navarro, G. (2018), "The influence of online ratings and reviews on hotel booking consideration", *Tourism Management*, Vol. 66, pp. 53-61.

Geetha, M., Singha, P. and Sinha, S. (2017), "Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis", *Tourism Management*, Vol. 61, pp. 43-54.

Gil-Soto, E., Armas-Cruz, Y., Morini-Marrero, S. and Ramos-Henríquez, J.M. (2019), "Hotel guests' perceptions of environmental friendly practices in social media", *International Journal of Hospitality Management*, Vol. 78, pp. 59-67.

Gu, H. and Ryan, C. (2008), "Chinese clientele at Chinese hotels—preferences and satisfaction", *International Journal of Hospitality Management*, Vol. 27, pp. 337-345.

He, W., Tian, X., Hung, A., Akula, V. and Zhang, W. (2018), "Measuring and comparing service quality metrics through social media analytics: a case study", *Information Systems and E-Business Management*, Vol. 16, pp. 579-600.

Herrero, Á., San Martín, H. and Hernández, J.M. (2015), "How online search behavior is influenced by user-generated content on review websites and hotel interactive websites", *International Journal of Contemporary Hospitality Management*, Vol. 27, pp. 1573-1597.

Ho, R.C. and Amin, M. (2019), "What drives the adoption of smart travel planning apps? The relationship between experiential consumption and mobile app acceptance", *KnE Social Sciences*, Vol. 1, pp. 22-41.

Ho, R.C. and Rajadurai, K.G. (2020), "Live streaming meets online shopping in the connected world: interactive social video in online marketplace", *Strategies and Tools for Managing Connected Consumers*, IGI Global, pp. 130-142.

Ho, R.C. and Rajandram, K.V. (2016), "The influence of social media data on online purchase: a study on relative advantage of social commerce", *Encyclopedia of E-Commerce Development, Implementation, and Management*, IGI Global, pp. 2039-2050.

Ho, R.C. (2020), *Strategies and Tools for Managing Connected Consumer*, IGI Global, Hershey, doi: 10.4018/978-1-5225-9697-4.

Hung, J.L. (2012), "Trends of e-learning research from 2000 to 2008: use of text mining and bibliometrics", *British Journal of Educational Technology*, Vol. 43 No. 1, pp. 5-16.

Kandampully, J. and Suhartanto, D. (2000), "Customer loyalty in the hotel industry: the role of customer satisfaction and image", *International Journal of Contemporary Hospitality Management*, Vol. 12, pp. 346-351.

Kang, M., Ahn, J. and Lee, K. (2018), "Opinion mining using ensemble text hidden Markov models for text classification", *Expert Systems with Applications*, Vol. 94, pp. 218-227.

Kim, A., Miano, T., Chew, R., Eggers, M. and Nonnemaker, J. (2017), "Classification of Twitter users who tweet about e-cigarettes", *JMIR public health and surveillance*, Vol. 3 No. 3, e63, doi: 10.2196/publichealth.8060.

Kim, M.R. (2013), "The relationship between leisure traveler's hotel attribute satisfaction and overall satisfaction", *Hospitality Review*, Vol. 31 No. 1, Article 12, available at: https://digitalcommons.fiu.edu/hospitalityreview/vol31/iss1/12.

King, R.A., Racherla, P. and Bush, V.D. (2014), "What we know and don't know about online word-of-mouth: a review and synthesis of the literature", *Journal of Interactive Marketing*, Vol. 28, pp. 167-183.

Knutson, B.J. (1988), "Frequent travelers: making them happy and bringing them back", *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 29, pp. 82-87.

Ladhari, R. and Michaud, M. (2015), "eWOM effects on hotel booking intentions, attitudes, trust, and website perceptions", *International Journal of Hospitality Management*, Vol. 46, pp. 36-45.

Lau, P.M., Akbar, A.K. and Fie, D.Y.G. (2005), "Service quality: a study of the luxury hotels in Malaysia", *Journal of American Academy of Business*, Vol. 7, pp. 46-55.

Lazard, A.J., Scheinfeld, E., Bernhardt, J.M., Wilcox, G.B. and Suran, M. (2015), "Detecting themes of public concern: a text mining analysis of the Centers for Disease Control and Prevention's Ebola live Twitter chat", *American Journal of Infection Control*, Vol. 43 No. 10, pp. 1109-1111.

Lee, W.-H. and Cheng, C.-C. (2018), "Less is more: a new insight for measuring service quality of green hotels", *International Journal of Hospitality Management*, Vol. 68, pp. 32-40.

Lee, J.H., Jung, S.H. and Park, J. (2017), "The role of entropy of review text sentiments on online WOM and movie box office sales", *Electronic Commerce Research and Applications*, Vol. 22, pp. 42-52.

Li, X., Wu, C. and Mai, F. (2019), "The effect of online reviews on product sales: a joint sentiment-topic analysis", *Information and Management*, Vol. 56 No. 2, pp. 172-184.

Liang, T.P., Li, X., Yang, C.T. and Wang, M. (2015), "What in consumer reviews affects the sales of mobile apps: a multifacet sentiment analysis approach", *International Journal of Electronic Commerce*, Vol. 20 No. 2, pp. 236-260.

Liu, Z. and Park, S. (2015), "What makes a useful online review? Implication for travel product websites", *Tourism Management*, Vol. 47, pp. 140-151.

Liu, X., Singh, P.V. and Srinivasan, K. (2016), "A structured analysis of unstructured big data by leveraging cloud computing", *Marketing Science*, Vol. 35 No. 3, pp. 363-388.

Lo, A.S. and Yao, S.S. (2019), "What makes hotel online reviews credible? An investigation of the roles of reviewer expertise, review rating consistency and review valence", *International Journal of Contemporary Hospitality Management*, Vol. 31, pp. 41-60.

Lockyer, T. (2005), "The perceived importance of price as one hotel selection dimension", *Tourism Management*, Vol. 26, pp. 529-537.

Magno, F., Cassia, F. and Bruni, A. (2018), "'Please write a (great) online review for my hotel!' Guests' reactions to solicited reviews", *Journal of Vacation Marketing*, Vol. 24, pp. 148-158.

Mattila, A.S. and O'neill, J.W. (2003), "Relationships between hotel room pricing, occupancy, and guest satisfaction: a longitudinal case of a midscale hotel in the United States", *Journal of Hospitality and Tourism Research*, Vol. 27, pp. 328-341.

Min, H., Lim, Y. and Magnini, V.P. (2015), "Factors affecting customer satisfaction in responses to negative online hotel reviews: the impact of empathy, paraphrasing, and speed", *Cornell Hospitality Quarterly*, Vol. 56, pp. 223-231.

Nasution, H.N. and Mavondo, F.T. (2008), "Customer value in the hotel industry: what managers believe they deliver and what customer experience", *International Journal of Hospitality Management*, Vol. 27, pp. 204-213.

Noone, B.M. and Mcguire, K.A. (2016), "Impact of attitudinal loyalty on the frequent unmanaged business traveler's use of price and consumer reviews in hotel choice", *Journal of Revenue and Pricing Management*, Vol. 15, pp. 20-36.

Öğüt, H. and Onur Taş, B.K. (2012), "The influence of internet customer reviews on the online sales and prices in hotel industry", *Service Industries Journal*, Vol. 32 No. 2, pp. 197-214.

Park, E., Kang, J., Choi, D. and Han, J. (2018), "Understanding customers' hotel revisiting behaviour: a sentiment analysis of online feedback reviews", *Current Issues in Tourism*, Vol. 23 No. 5, pp. 605-611.

Park, E. (2019), "Motivations for customer revisit behavior in online review comments: analyzing the role of user experience using big data approaches", *Journal of Retailing and Consumer Services*, Vol. 51, pp. 14-18.

Petrescu, M. (2013), "Marketing research using single-item indicators in structural equation models", *Journal of Marketing Analytics*, Vol. 1, pp. 99-117.

Ruys, H. and Wei, S. (1998), "Accommodation needs of mature Australian travellers", *Australian Journal of Hospitality Management*, Vol. 5, pp. 51-60.

Salehan, M. and Kim, D.J. (2016), "Predicting the performance of online consumer reviews: a sentiment mining approach to big data analytics", *Decision Support Systems*, Vol. 81, pp. 30-40.

Sampaio, C.A., Hernández-Mogollón, J.M. and Rodrigues, R.G. (2019), "Assessing the relationship between market orientation and business performance in the hotel industry–the mediating role of service quality", *Journal of Knowledge Management*, Vol. 23, pp. 644-663.

Schivinski, B. and Dabrowski, D. (2016), "The effect of social media communication on consumer perceptions of brands", *Journal of Marketing Communications*, Vol. 22, pp. 189-214.

Singh, J.P., Irani, S., Rana, N.P., Dwivedi, Y.K., Saumya, S. and Roy, P.K. (2017), "Predicting the "helpfulness" of online consumer reviews", *Journal of Business Research*, Vol. 70, pp. 346-355.

Torii, M., Tilak, S.S., Doan, S., Zisook, D.S. and Fan, J.-W. (2016), "Mining health-related issues in consumer product reviews by using scalable text analytics", *Biomedical Informatics Insights*, Vol. 8 No. S1, pp. 1-11, doi: 10.4137/BII.S37791.

Tsai, H., Yeung, S. and Yim, P.H. (2011), "Hotel selection criteria used by mainland Chinese and foreign individual travelers to Hong Kong", *International Journal of Hospitality and Tourism Administration*, Vol. 12, pp. 252-267.

Tsai, S.-P. (2017), "Driving holistic innovation to heighten hotel customer loyalty", *Current Issues in Tourism*, Vol. 20, pp. 1604-1619.

Uddin, M.M. (2015), "Assessing the quality of hospitality services: a study on hotels in chittagong", *European Journal of Business and Management*, Vol. 7, pp. 18-29.

Vermeulen, I.E. and Seegers, D. (2009), "Tried and tested: the impact of online hotel reviews on consumer consideration", *Tourism Management*, Vol. 30, pp. 123-127.

Viglia, G., Minazzi, R. and Buhalis, D. (2016), "The influence of e-word-of-mouth on hotel occupancy rate", *International Journal of Contemporary Hospitality Management*, Vol. 28, pp. 2035-2051.

Wu, P.-J. and Lin, K.-C. (2018), "Unstructured big data analytics for retrieving e-commerce logistics knowledge", *Telematics and Informatics*, Vol. 35, pp. 237-244.

Xia, H., Yang, Y., Pan, X., Zhang, Z. and An, W. (2019), "Sentiment analysis for online reviews using conditional random fields and support vector machines", *Electronic Commerce Research*, Vol. 13, May, pp. 1-18, doi: 10.1007/s10660-019-09354-7.

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

Xiang, Z., DU, Q., Ma, Y. and Fan, W. (2017), "A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism", *Tourism Management*, Vol. 58, pp. 51-65.

Xie, K.L., Chen, C. and Wu, S. (2016), "Online consumer review factors affecting offline hotel popularity: evidence from tripadvisor", *Journal of Travel and Tourism Marketing*, Vol. 33 No. 2, pp. 211-223.

Xie, K.L., So, K.K.F. and Wang, W. (2017), "Joint effects of management responses and online reviews on hotel financial performance: a data-analytics approach", *International Journal of Hospitality Management*, Vol. 62, pp. 101-110.

Xu, Z., Frankwick, G.L. and Ramirez, E. (2016), "Effects of big data analytics and traditional marketing analytics on new product success: a knowledge fusion perspective", *Journal of Business Research*, Vol. 69, pp. 1562-1566.

Yang, Y., Mao, Z. and Tang, J. (2018), "Understanding guest satisfaction with urban hotel location", *Journal of Travel Research*, Vol. 57, pp. 243-259.

Zeng, D., Chen, H., Lusch, R. and Li, S.-H. (2010), "Social media analytics and intelligence", *IEEE Intelligent Systems*, Vol. 25, pp. 13-16.

Zhang, X., Yu, Y., Li, H. and Lin, Z. (2016), "Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews", *Online Information Review*, Vol. 40 No. 1, pp. 119-145.

Zhao, Y., Xu, X. and Wang, M. (2019), "Predicting overall customer satisfaction: big data evidence from hotel online textual reviews", *International Journal of Hospitality Management*, Vol. 76, pp. 111-121.

Zhou, L., Ye, S., Pearce, P.L. and Wu, M.Y. (2014), "Refreshing hotel satisfaction studies by reconfiguring customer review data", *International Journal of Hospitality Management*, Vol. 38, pp. 1-10.

Corresponding author
Ree Chan Ho can be contacted at: reechanho@gmail.com