Capitalizing on big data and revolutionary 5G technology: Extracting and visualizing ratings and reviews of global chain hotels☆

Loveleen Gaur [a,1], Anam Afaq [a,1], Arun Solanki [c,1], Gurmeet Singh [b,1], Shavneet Sharma [b,1], N.Z. Jhanjhi [d,1], Hoang Thi My [e,1], Dac-Nhuong Le [f,g,*,1]

[a] Amity International Business School, Amity University, Noida, Uttar Pradesh, India
[b] The University of the South Pacific (USP), Faculty of Business & Economics, Suva, Fiji
[c] Gautam Buddha University, Greater Noida, Uttar Pradesh, India
[d] School of Computer Science and Engineering SCE, Taylor's University, Malaysia
[e] Faculty of Tourism, Haiphong University, Haiphong 180000, Viet Nam
[f] Institute of Research and Development, Duy Tan University, Danang 550000, Viet Nam
[g] School of Computer Science, Duy Tan University, Danang 550000, Viet Nam

ARTICLE INFO

ABSTRACT

This paper aims to use machine learning (ML) algorithm for natural language pre-processing (NLP), text mining (TM), and sentiment analysis (SA) techniques to analyze and examine 45,500 online reviews of customers of 50 global chain hotels from different online review sites. Furthermore, the paper addresses the new business value and experiences that the revolutionary 5G technology can bring to the hotel industry. The research findings revealed that the general review star rating corresponds with the opinion (sentiment) scores for the title and the full substance of the online reviews. The case study's contextual analysis also uncovered that both fulfilled and disappointed customers have a frequent inclination for five categories: food, stay, rooms, service, and staff. This study contributes both theoretically and practically to the multidisciplinary domains of computer science, information systems, and tourism and discovers hidden patterns in data using visual analytics techniques.

1. Introduction

Businesses worldwide are engaged in different web-based activities using social media platforms regardless of their size [1]. The e-commerce era has affected many industries, including hospitality and tourism [2]. Online feedback by customers plays an increasingly important role for businesses due to its fast and up to date nature. These reviews assist in comprehending customer's needs [3]. According to budding customers' reviews, the reviews are influential, trustworthy, and helpful [4]. These reviews provide a base to evaluate the customers' expectations from the service providers. Another paper predicted sentiment of the news article is positive or negative using the two popular approaches of Naïve Bayes Text Categorization, i.e., Multivariate Bernoulli Naïve Bayes Classification and Multinomial Naïve Bayes Classification [5].
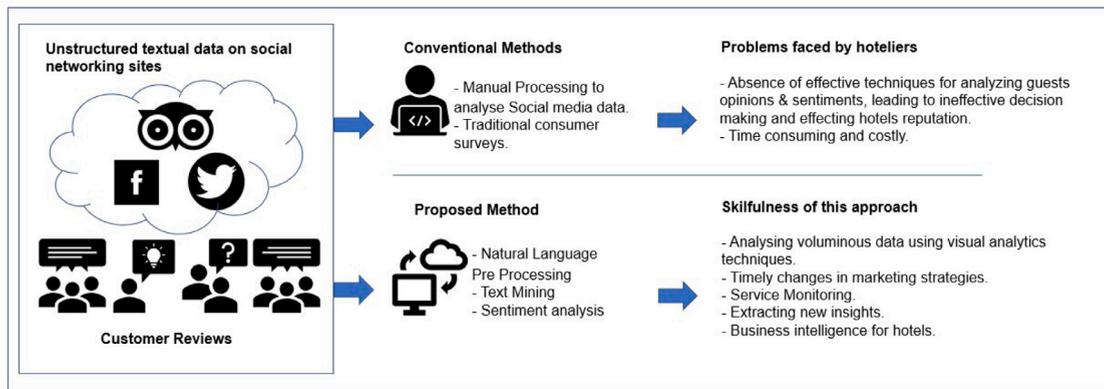

Fig. 1. Problem statement.

Moreover, businesses acknowledge the benefits of social media data as it provides them valuable and timely insights that can be useful for making meaningful decisions and improving and refining their products. Moreover, businesses are strengthening business analytics to draw out deep and meaningful information through data available on social platforms. The application of business analytics for analyzing bulk social data can be an effective way through which businesses can discover innovative ideas for providing solutions to the concerns raised by their customers, which in turn can lead to competitive advantage [3]. Machine learning for NLP and TM plays a significant role in understanding the meaning of text documents. These documents can be anything that contains the text: survey responses, online reviews, social media comments. ML's role in NLP and TM is to enhance, facilitate, and automate the implied TM functions and NLP features, making the unstructured data into usable insights. ML aims at making SA more convenient. As this paper emphasizes the importance of analyzing bulk user-generated content, it is also important to discuss how the hotel industry can further benefit if they embrace revolutionary 5G and 6G technologies for data analytics. The fifth-generation networks can help data analytics effectively in the coming years. The fast-downloading speed and low latency can help analysts clean, collect, and analyze the voluminous data in a brief time. To further step in providing better customer experience apart from getting insights from the customer reviews, it is also crucial for hotels to embrace these technological advancements and incorporate technology like 5G and 6G, which can further assist in improving user experiences at hotels. The technology is available only in few countries; however, the hotel industry needs to embrace it at a faster pace to be ahead of the competition and cater to the needs of the customers who already have 5G devices. This technology gives the potentiality to digitalize the hotel industry more rapidly and embrace new business models. The wide-scale use of 5G technology can support a high customer engagement level in the hospitality industry [6].

The paper in [7] explains the reason for the widespread adoption of social media in the hospitality domain, primarily due to the critical role of online reviews in shaping new customers' decision-making. Customer satisfaction is a crucial ingredient for hotel businesses. Many potential customers spend a lot of time reading online reviews of previous customers before deciding on their choice of hotels. These online reviews provide insights to the hotels and improve their hotel performance and stand out profitably than their competitors [8]. Additionally, these reviews are useful for hotel managers who can make timely changes to their marketing strategies to offer better services to their guests and strengthen their brand for their future growth [9]. The below-mentioned Fig. 1 depicts the problems faced by the hoteliers and the proposed approach that could help them improve service delivery.

The availability and accessibility of large volumes of data generated by users on social platforms have offered researchers and markets new opportunities. They have identified the benefits of using automated computational techniques like TM and SA to capture a massive amount of user-generated information for extracting meaningful and relevant information about the customers [10]. The old school methods of content analysis are outdated. They cannot meet the organizations' needs in analyzing a considerable amount of everyday content available on social media. To avail this opportunity of extracting and using the user-generated content for drawing inferences, the businesses must be equipped with a data analytics skillset for analyzing such massive amounts of data. [10], through their research, they found that companies exploring and accepting the opportunities offered by these analytics techniques have shown better improvement in their performance. Consequently, this paper attempts to focus on the emerging technologies that can be applied and integrated for analyzing unstructured data available on social media by using a case study.

This paper aims to analyze online reviews of English-conversing hotel guests and understand their preferred hotel attributes, guest demands, and concerns. This research uses Machine learning (ML) for Natural Language Pre-processing (NLP), Text Mining (TM), and Sentiment Analysis (SA) techniques to analyze customers' online reviews collected from different review sites. Positive customer feedback and emotions are critical for a service-based industry like tourism [11]. Researchers have resorted to the use of various instruments to measure this. Surveys have primarily been used to collect data in this regard, which has resulted in several drawbacks [11]. The use of new technologies and the availability of enormous user-generated content have enabled research to adopt SA. Such data analysis techniques have just begun to gain popularity for researchers with particular relevance to the tourism industry [12]. Studies have started to use techniques to provide valuable insights into the tourism domain. However, researchers

have acknowledged hotel reviews using SA are still scarce [13]. Literature has shown that SA has mostly performed using reviews extracted from Amazon and Twitter. Therefore, this study aims to address that gap. Furthermore, the research also highlights the areas through which the hotels can leverage the 5G and 6G technology.

The fifth generation of cellular technology can transform the way hoteliers connect with their guests. 5G goes much beyond just "faster connection speed." As it advances, 5G is anticipated to improve the very DNA of the user experience in exciting ways. The ultimate aim of hotels is to provide a positive experience to their guests, converting their one-time guests into loyal guests. If the hotels look to deploy 5G networks, it can give ample facilities to the guests. For example, they can have full access to their rooms by using their mobiles or connected tablets while on the property's wireless 5G network. Cellphones could be used for keyless entry and for enhancing the in-room experiences by pre-setting the room temperature, window shades, television options, and lights with only a touch of a button on their devices. 5G can bring technologies like machine learning and artificial intelligence to the frontage and effectively link the digital and physical worlds. The intelligent connectivity enabled by 5G can create new hotel industry cases and transform hotel operations and business activities. The combination of artificial intelligence (AI) and 5G can enable swift check-in and check-out within a fraction of seconds through facial recognition, enhancing security, and service efficiency. These technologies will help the hoteliers offer more personalized services and provide customized treatment that the clients demand. Therefore, it can be said that 5G will be a significant boost to the service industry. The incorporation of 5G technology into tourism will promote smart tourism, which is customer-centric and aims to upgrade the traveling experience [14].

This research can be of relevance to the hoteliers in many ways. Firstly, the analysis of online reviews can give the hotel managers directions in getting insights about the admired hotel attributes, issues, and expectations of customers. There is a shortage of studies in this specific niche. However, millions of foreigners visit different countries as tourists every year. The authors illustrated the approach that uses ML for NLP, TM, and SA techniques for analyzing the online content using a case study. As presented in the case study, the multifariousness of the extracted information indicates the usefulness of this approach. The visualization techniques used in the case study that showed meaningful insights can give the hotel service providers directions in terms of visualizing and capitalizing upon collected data for their hotels and their competitors to stimulate decision-making regarding strategic planning and operation. The displayed results can be used by the hotel managers for a quick service check, troubleshooting, and assessment. Secondly, it draws the hoteliers' attention towards the emerging 5G and 6G technologies that can be explored by them for providing enhanced hotel guest experience.

This study will provide useful insights and findings for hotels to better understand the needs and requirements of guests. It will help businesses understand their customer needs and requirements and use their resources effectively by knowing what improvement and changes are expected to satisfy customers better. Big Data analysis can provide new business opportunities and be a source of competitive advantage [15]. Visualization techniques like bar/pie charts, multi-Y axis graphs, word clouds, and adjacency matrix graphs have been used to depict how hotels can utilize to transform guests' reviews into valuable information. This will help managers in decision making and planning. The charts related to sentiment by rating and correlation networks used in this paper aid to the diverseness of visualization methods for carrying out online content analysis. Moreover, the study depicts the correlation between review star rating and the sentiment scores for the title and the online customers' complete review content. As the claims are short in many cases, the full content reviews provide better and more precise insights along with text analytics.

The rest of this paper is organized as follows. Section 2 details the review of TM and SA. Section 3 has its focus on a case study that uses the approach, which is a combination of ML for NLP, SA techniques, and TM for analyzing online hotel reviews. Section 4 of this paper presents the insights of the case study. The conclusion and need for future research are mentioned in Section 5.

2. Literature review

The studies mentioned above stipulate the requirement for further research about this area (see Table 1). These guiding insights were useful for our present paper. Travelers can provide new insights and knowledge for the hotel managers. Thereby, we were prompted to apply ML, TM, and SA techniques for analyzing the data. Also, studies that have attempted to investigate customer reviews in the hotel industry using survey method, econometrics, modeling [2]. However, these studies considered the number of customer reviews, hotel stars, or overall hotel rating. This study will overcome the limitations of these methods by using textual analysis and visualization techniques that allow for more detailed understanding and insights into customers.

As social media platforms have gained a lot of popularity in expressing opinions and sharing information by customers, service-based firms have become necessary to use this source to capture knowledge about their customers. Online reviews can be a game-changer for hospitality practitioners. The traditional questionnaire and interview methods cannot compete with the voluminous self-expressed real-life responses of customers available on online platforms. Moreover, the massive volume of data on the internet has made manual processing next to impossible, and hence there is a need to adopt new technical approaches. Competition in the service industry is intensive, and hotels that get timely and relevant knowledge about its guests can edge over their competitors. For example, Marriot has recognized the benefits of big data analytics for meeting the modern customer-centric world's demand. It uses big data analytics to track competitor brands and hotels. Marriot uses its revenue management approach called a dynamic pricing approach to enable the hotel chain to forecast demand and customer behavior patterns correctly [21]. However, the majority of hotel firms have not fully adopted big data analytics technology. The absence of useful techniques for analyzing customer opinions and sentiments has led to ineffective decision-making and has tampered with many companies' business reputation. In contrast, the firms which developed significant data analytics capabilities have shown a positive effect on their performance [22]. In the present study, implementing big data analytics can open the doors for effective decision making, quality service monitoring, and business intelligence for hotels.

Table 1
Social media analytics and the tourism industry.

Author	Findings
[8]	The authors adopted the TM approach for inspecting the undermining of fulfilled and dissatisfied customers by analyzing and comparing the online reviews posted by the hotel guests. Findings revealed that satisfied customers focused more on impalpable aspects of their stay at hotels like the staff or employees than dissatisfied customers.
[12]	The authors implied TM techniques in their study and stated that gamification features impact travelers when they write reviews.
[15]	The authors suggested that Bigdata plays a significant role in tourism and hospitality research and how it can enhance other digital technologies like artificial intelligence and IoT in the tourism and hospitality sectors.
[16]	Automation methods like TM and SA have gained a lot of popularity for processing vast amounts of user-generated information and pulling informative insights posted by users on social media.
[17]	They used TM for comparing three key online review forums (Yelp, Expedia, and TripAdvisor) concerning quality details associated with the online feedback regarding complete hotel occupants at Manhattan, New York City. Their findings projected a large disparity in the portrayal of the hotel sector on these channels. They also found that online opinions differ significantly in respect of their language-producing features, rating, sentiment, benefits, and association among these characteristics.
[18]	The social media data from tourists can enhance the value creation process to give way to a smart tourism destination.
[19]	The authors compiled and contextualized the application of IoT in hospitality and applied an input-process–output model, and demonstrated its link with bigdata.
[20]	The authors examined the essential features of luxury hotel service in Malaysia and analyzed bigdata in online comments generated from trip advisor. Content analysis was performed by using word frequency to identify main themes from the dataset

3. Methodology

The authors created the data set by extracting customer views from different online review sites. TripAdvisor was also one of the review sites for extracting data due to its popularity and an extensive collection of user-generated content [13,23]. This increases the likelihood of collecting sufficient and abundant data for analysis. We used web scraping to collect reviews in January 2020. This data set consists of 45,500 English reviews given by the guests for the hotels. We gathered structured systemized information (overall rating score and specific rating scores) and unstructured data for every discussion, including titles and textual review content.ML for NLP, TM, and SA techniques applied for analyzing these reviews.

This study aims to ascertain the sentiment of hotel guests on different review sites. Fig. 2 above depicts the flowchart of the methodology employed by this study and below mentioned Fig. 3 illustrates the layered system architecture of this study.

First, a program extracted guests' reviews concerning hotels and the reviews stored in a `.csv` file in excel. Web scraping and HTML parsing methods were used for retrieving the data from the review sites and converted into a format used for analysis. As mentioned, the authors created the data set; the below-mentioned Fig. 4 shows the flowchart depicting the entire process of creation of the dataset.

For creating the dataset, the hotel review sites were identified, where the reviews were available, post that the URL list of those hotels was fetched whose data was required. The next step was to pass the hotel review URLs to the RCrawler library function to scrape its content and produce structured data set and URL of the following review page link. The step was repeated till the time the next page links were available. The above process was applied to all hotel URLs to get the structured data set for all hotels.

Second, NLP was conducted for cleaning the text, like the removal of stop words. The data pre-processing step involves transforming raw data into a workable format by data cleaning. *R* libraries were viable for TM for analyzing the text for word-frequency analysis as well as for co-occurrence analysis. We constructed an $(n \times t)$ document term matrix (DTM), where n stands for number of documents and t is the number of unique terms. Each DTM column represents unique term, the (i, j)th cell illustrates the number of terms j that are present in document i. We applied sparse = 0.997 for term t to retain all terms for which $df_t > N \times (1 - 0.997)$, where N is the number of documents and df_t is document frequency for term t. Next $(\rho(statist, ?) \geq 0.05)$ is used to compute the correlations between terms.

Third, a list of possible categories from the dataset developed using *R* libraries. The new categories were identified from the sample of small data. Afterward, a program to undergo the complete set of data and position the applicable words allied to the located categories into multiple categories. For example, the Ambience category can contain related words like "overall environment," "Décor and design," "table setting and seating." Another example can be of food category that includes "foods," "breakfast," and "Menu," "taste," "starters," "desserts," "soup," "Brunch," "main course," "cuisines," "Ingredients," "presentation."

Fourth, for every category SA technique was applied for grouping the comments under various categories like positive, neutral, and negative.

Fifth, the results were examined thoroughly to identify emerging trends and the latest topics for better insights and decision-making. Varied query searches were carried out by using the search feature option of NVivo software. This search generally assists in finding associations and uncommon insights from the content.
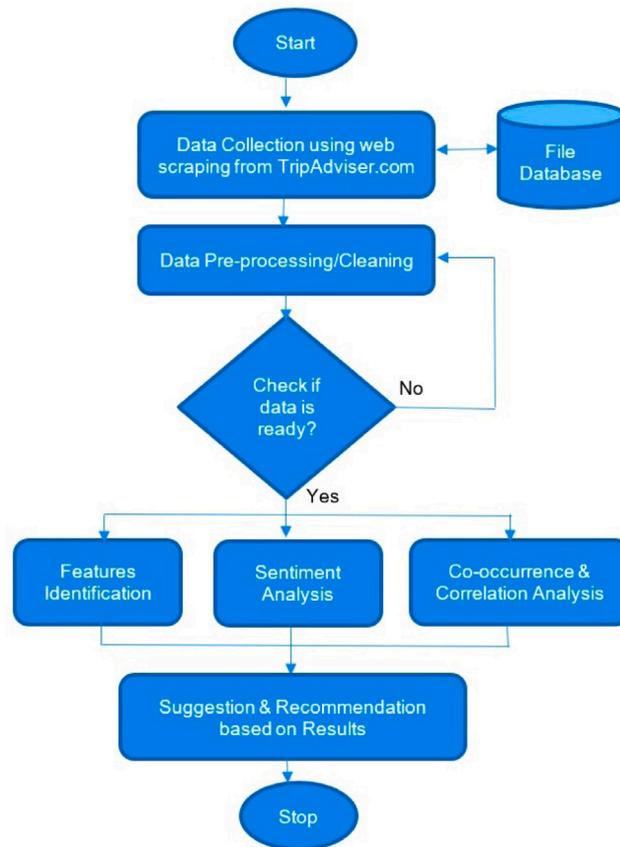
Fig. 2. The flow chart of the methodology.
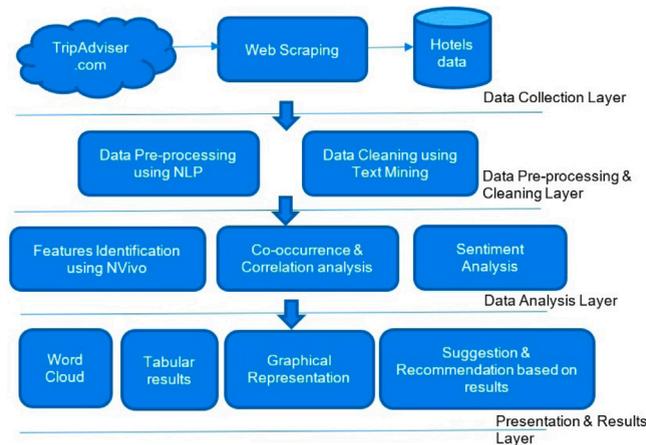
Fig. 3. Layered system architecture.

4. Results, discussions, and implications

We collected a total of 45,500 English reviews obtained from different review sites. Before applying the TM techniques, the cleaning of the dataset by removing special symbols such as "&," "!," "#," "@," and so on. A pivotal step to be followed is the processing of the dataset before analyzing it. The entire data set changed into a lower case, and common stop words were removed from each sentence. For example, "a," "the," "and," and other common stop words were taken out. Next, the DTM Document Term
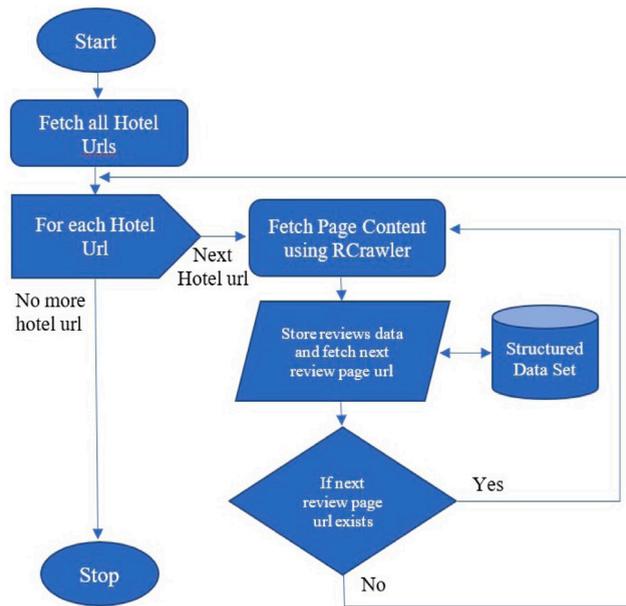
Fig. 4. Flowchart showing dataset creation.

Table 2
Data corpus statistics.

Term	#
Hotel	48083
Rooms	35401
Stay	34157
Staff	30685
Good	26754
Services	26413
Food	20356
Great	18715
Thank	11741
Restaurants	11460
Nice	11445
Breakfast	11211
Experience	11100
Excellent	10697
Like	10128
Best	9552
Amazing	8577
Place	8187
Helpful	7460
Visit	6964

Matrix was created to study this data. Open-source *R* libraries applied for evaluating the word frequency. Word frequency analysis recognized the terms that repeatedly appeared in the dataset's comments, as exhibited in Table 2.

After getting the high-frequency terms, a word co-occurrence search and analysis were conducted to identify the words that have a strong association with the high prevalence terms in *R*. Correlation checked between 0 and 1. For example, if the correlation is 1, the search request will present just those words which invariably co-occur with the search term. Therefore, it is said that the program will depict the results of the words which have a co-occurrence of not less than 80% of the high searched frequency-terms. This case study presents five higher frequency terms, and its results are depicted in Table 3 listed below. To illustrate, the word hotel repeatedly co-occurs alongside food. This portrays that generally, whenever customers post reviews about hotels, they also mention rooms and the quality of food along with the term hotel.

Similarly, the words friendly and helpful are seen co-occurring with the term staff. This analysis of co-occurrence terms as presented in Table 3 brings out the fact that few words are interconnected and should not be dealt with separately; for instance, in a hotel, customers will also look for good food and staff, and they have a weight-age and is a significant part for overall customer experience.

Table 3
Co-occurrence terms.

Search term: *hotel*		Search term: *stay*		Search term: *rooms*		Search term: *service*	
Term	*Correlation*	*Term*	*Correlation*	*Term*	*Correlation*	*Term*	*Correlation*
Rooms	0.30	Hotel	0.24	Hotel	0.30	Customer	0.17
Stay	0.24	Rooms	0.23	Bathroom	0.28	Rooms	0.16
Staff	0.19	Comfortable	0.20	Bed	0.28	Excellent	0.15
Airport	0.17	Staff	0.19	Breakfast	0.26	Food	0.11
Breakfast	0.17	Breakfast	0.15	Check in	0.26	Hotel	0.11
Restaurants	0.16	Memorable	0.14	Stay	0.23	Restaurants	0.10
Bed	0.14	Trip	0.13	Booked	0.23	Impeccable	0.10
Booked	0.14	Check-in	0.12	Floor	0.23	Quality	0.09
Bathroom	0.14	Pleasant	0.11	Clean	0.21	Lounge	0.09
Pool	0.14	View	0.10	Shower	0.21	Outstanding	0.09
Like	0.13	Member	0.20				
Check-in	0.13	Friendly	0.20				
Deals	0.12	Hotel	0.19				
Luxury	0.12	Stay	0.19				
		Helpful	0.19				
		Courteous	0.15				
		Rooms	0.14				
		Desk	0.14				
		Housekeeping	0.14				
		Attentive	0.13				
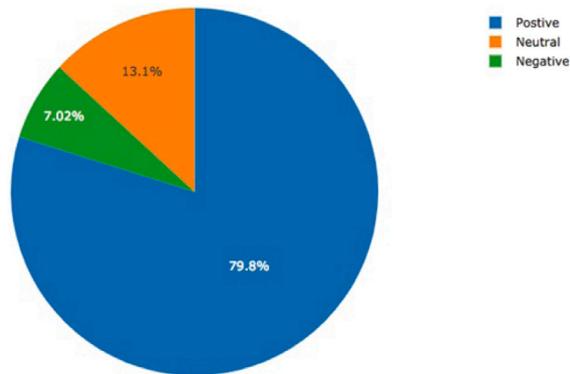
Fig. 5. Sentiment distribution.
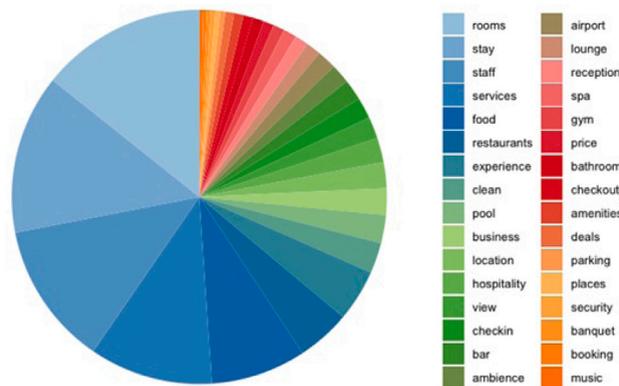
Fig. 6. Category magnitude distribution.

Based on these comments, SA conducted for all comments in the data set using *R* programming. The data set results depicted 79.8 percent comments as positive, 13.1 percent as neutral, while the reminder 7.02 percent as unfavorable, as shown in Fig. 5.

Fig. 6 presents the category magnitude distribution of the data. These have been classified into 32 categories.
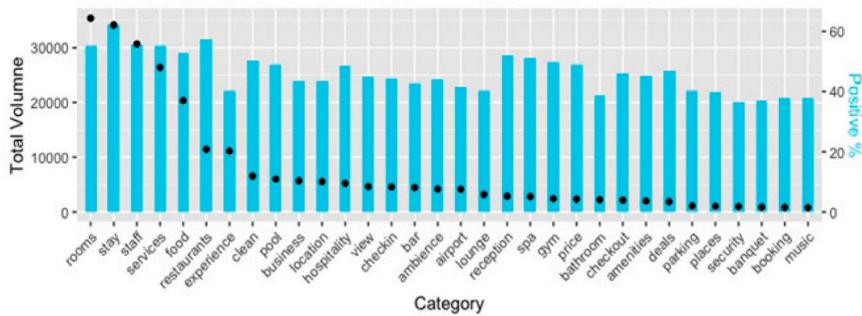
Fig. 7. The sentiment of every category.
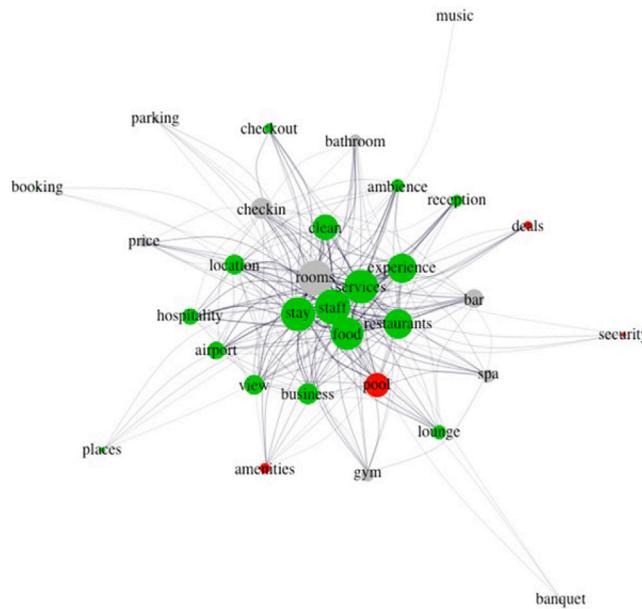
Fig. 8. Category correlation network.

Conducive to every category, the percentage of positive remarks compared. The comparability depicted a clearer image of customers' pursuits or interests, as portrayed in Fig. 7. The total reviews in every category and their overall sentiments, as shown in Fig. 7. For instance, rooms, food, staff, services are the top things for which the customers cared the most as these categories had a maximum number of reviews. For every group, positive comment vs. total grand comment is shown. For example, food received about 55 percent positive feedback from full comments, and the stay category gained around 60 percent of positive feedback.

The correlation between the categories is also presented in Fig. 7. For every type, the frequency of comments present in both groups was computed. Fig. 8 shows the category correlation network; programming language *R* is used to reconstruct the data in the form of an adjacency matrix, after that a term–term adjacency matrix is developed, here the rows and columns illustrate terms, and each entry represents the count of co-occurrences of the two words. Moreover, the graph. The adjacency function of the library "graph" of R was used to construct the graph. The colors of the correlation network portray the sentiment of the categories. The fatness of the edges bridging the two groups is dependent on the frequencies in which the two types occur together. The size of the bubble represents the volume of the category. Fig. 8 depicts the categories with a maximum number of comments. There are some complaints as well, as problems with security, amenities, and deals. However, mostly the comments are positive.

Sentiment and category interpretation helped in projecting the basic idea of hotel industries. The comments can be transformed into opinion knowledge by running the same analysis for each hotel.

Moreover, the connection within the overall review rating for each review is identified; it also includes the sentiment scores of the comment and its title and full content.

Figs. 9 and 10 present the reviews content and title's sentiment by rating. The projected results reveal a correlation between the general review star rating and the sentiment scores of the title as well as the complete content of online reviews. In comparison to the title, the full review details depict further useful results from TM; this is so because, in numerous cases, the title is minimal.
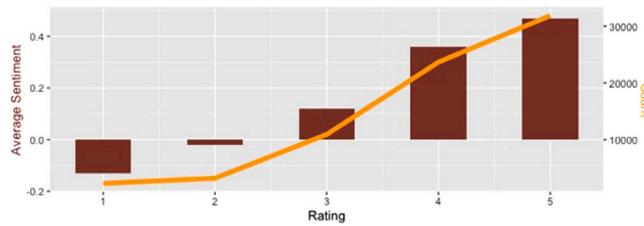
Fig. 9. Reviews content sentiment by rating.
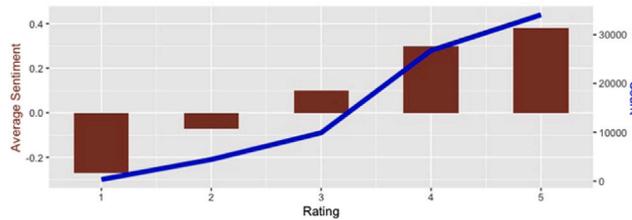
Fig. 10. Reviews titles sentiment by rating.

Fig. 11. Word cloud showing amazingly fulfilled and incredibly disappointed customers.

Table 4

Frequently quoted categories of fulfilled and disappointed customers.

Customers	Categories
Amazingly fulfilled customers	rooms, stay, staff, services, food, experience, breakfast
Incredibly disappointed customers	rooms, services, stay, staff, check-in, food

Satisfied customers are an asset for hotels, as they will come back to the same hotels if their expectations are met and spread positive word of mouth. Table 4 presents the frequently quoted categories of fulfilled and disappointed customers. These positive online reviews can be beneficial for hotels to improve revenues [24].

Moreover, [25,26] stated that the comments depicting extreme sentiments are also considered to help make significant and strategic business decisions. [27,28] found that extremely positive as well as extremely negative ratings are more decisive as compared to the neutral ratings. For this purpose, we used R for comparing and analyzing the full substance of all the reviews, which includes both amazingly fulfilled customers giving 5-star ratings and very disappointed customers providing a 1-star rating. The word clouds for astonishingly accomplished and incredibly frustrated customers are generated in Fig. 11. This figure also reveals some of the categories that are commonly found in both positive as well as negative reviews of customers.

The Web 2.0 era has witnessed an explosion of social media data. Using social networking sites, these digital consumers share their experiences about products and services with each other. There is an immediate need for businesses to adapt to new high-tech analytics and new sophisticated tools for improved business expansion and penetration. Here, the case study approach used, how companies in general and hotels, in particular, can use online reviews to find customers' perceptions and deep insights from online reviews on different social media platforms. The hotel managers can attain crucial managerial information to understand the strengths and weaknesses of their services and explore the opportunities that they can grab without wasting more time and resources. One example of this is the assessment that can be drawn from the findings of this research about the categories that attained less sentiment score and can improve those areas to avoid any negative consequences.

Different hotel attributes discussed by the customers on online forums could be explored for gaining better insights on how the products and services can be improved to satisfy customers and expand the business. Based on the findings from research, advises hotel managers to be prompt in giving responses to the hotel guests in a transparent and customized way. The managers need to continuously keep comparing customer satisfaction level after making changes based on the review results. This will bring clarity as to whether the decisions made by the managers have an impact on customer satisfaction. The word co-occurrence analysis can be an effective way for hotel managers to know how the categories are interrelated. They can focus on providing better services to the guests by understanding the link between the co-occurring terms and can pay attention to the concerned aspects.

Considering the management outlook, [8] suggested scrutinizing categories that have cropped up from online reviews. These reviews provide an excellent opportunity to understand every hotel guest's voice and help in examining the overall picture gathered through these conversations collectively. The completion stage of TM calls for the assessment of knowledge collected from the data and to present the insights with the assistance of conceptualized techniques like graphs. Fig. 6 unveils the 32 categories of hotels used by the hoteliers to handle their conversations and engagement with the guests. Fig. 7 depicts a few of the categories that receive fewer sentiment scores and should be looked upon by the hotel managers to make additional efforts to improvise these areas to meet their guests' expectations. Fig. 8 depicting the category correlation network brings out a clearer picture of the categories representing the satisfied and dissatisfied customers; this figure can help find the inherent relationship amongst categories. At times the problem of a particular category may not be tackled without understanding the problems of related categories. Therefore to solve the main problem, the hotel managers first need to fix the issues of the associated categories. The managers should look at a bigger picture through these figures and should analyze the potential opportunities they could grab by removing some of the obstacles coming in the way of serving the customers better.

Few categories are uncontrollable like location, view, and airport as once a hotel is built, it cannot change its position; similarly, the view outside the hotel cannot be drastically altered, and if the airport is far from the hotel, nothing can be done in this context. Therefore, the hotel managers need to focus on categories which are controllable and can have a practicable impact. For example, efforts can be put to improve the food quality or train the employees to be more courteous towards the guests, to improvise check-in and check-out services and to provide better deals and adequate security to the guests. The graphs generated through the presented case study like sentiment by rating and category correlation networks bring out social media analytics' potential to create a business vision. This case study discovered a connection between the overall star rating of review and sentiment scores for the title and the entire content of the reviews.

This case study's empirical findings reveal that both amazingly fulfilled and incredibly disappointed customers hold a universal inclination towards five categories, food, stay, service, room, and staff. These findings indicate to the hotel managers that the services, facilities, and issues that create satisfied customers can also become the cause for their dissatisfaction. Therefore, a lot more emphasis should be given to these categories. These dimensions can make as well as break the image of the hotels [8]. The strong performance of hotels in these categories can be a benchmark for the hotels. This can be a useful marketing tactic for inviting more guests and leading to a more competitive advantage.

Several research studies in this area indicate that the organizations using and relying on analytics have overshadowed their counterparts in terms of profits and growth [10]. SM Analytics is a hub for generating essential business knowledge and uses content to fabricate business intelligence [16]. A content analysis of big data analytics applications in healthcare cases mapped the benefits of adopting big data analytics technology [29]. E-Commerce firms that injected big data analytics (BDA) within their value chain experienced 5 to 6 percent higher returns as compared to their competitors [30]. Social media analytics possess a lot of potential and extensiveness, and in today's world of big data, SM analytics assists businesses to make informed and better decisions. It is now becoming necessary for organizations to boost their competence in gathering, keeping, and evaluating social media data. By doing so, these organizations can reap knowledge and attain practical insights for planning, forecasting, and making effective decisions. Not all enterprises are fully equipped with the capacity to analyze the collected social media data. The omnipresent nature of user-generated content has created challenges for companies. Therefore, it is a requirement for them to develop advanced analytics competence to perform better than competitors.

SM analytics is a booming versatile research subject, and it is essential to develop effective analytics techniques to manage digital customers. Graphs produced like category correlation network and sentiment by ratings can be a necessary contribution towards visualization techniques that can be used for carrying out online content analysis and discovering new revelations for business and availing the benefits of analytics. Attaining competency in using these techniques can benefit the enterprises in developing strong business analytics proficiency and equip them well in improving their current business analytics operations. The approach applied in the presented paper can be adopted by an extensive range of service domains for improving enterprise knowledge procurement capacity. However, from a futuristic point of view, it is also vital for the hoteliers and others in the service industry to keep on upgrading themselves with the upcoming technological advancements. It is also essential for the service industry to embrace future technologies in 5G and 6G networks for delivering an exceptionally amazing service delivery and enhanced guest experience. For hotel operations and management in the coming years, if 5G used along with big data and AI, it can help collaborate among devices at the front end like patrol robots and Augmented Reality (AR) headsets and data at the back end. These combinations can guarantee intelligent safety and security both indoors and outdoors.

5. Conclusion

The implementation of big data analytics surpasses other conventional methods of analysis and provides significant insights that could benefit hotels in the tourism industry to uplift service delivery. The authors have applied ML algorithm for NLP, TM and

[13] Chang Y-C, Ku C-H, Chen C-H. Social media analytics: Extracting and visualising Hilton hotel ratings and reviews from TripAdvisor. Int J Inf Manage 2019;48:263–79. http://dx.doi.org/10.1016/j.ijinfomgt.2017.11.001.

[14] Peng R, Lou Y, Kadoch M, Cheriet M. A human-guided machine learning approach for 5G smart tourism IoT. Electronics 2020;9(6):947. http://dx.doi.org/10.3390/electronics9060947.

[15] Mariani M. Big data and analytics in tourism and hospitality: a perspective article. Tour Rev 2019;75(1):299–303. http://dx.doi.org/10.1108/tr-06-2019-0259.

[16] He W, Tian X, Tao R, Zhang W, Yan G, Akula V. Application of social media analytics: a case of analysing online hotel reviews. Online Inf Rev 2017. http://dx.doi.org/10.1108/oir-07-2016-0201.

[17] Xiang Z, Du Q, Ma Y, Fan W. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. Tour Manag 2017;58:51–65. http://dx.doi.org/10.1016/j.tourman.2016.10.001.

[18] Del Vecchio P, Mele G, Ndou V, Secundo G. Creating value from social big data: Implications for smart tourism destinations. Inf Process Manage 2018;54(5):847–60. http://dx.doi.org/10.1016/j.ipm.2017.10.006.

[19] Nadkarni S, Kriechbaumer F, Rothenberger M, Christodoulidou N. The path to the hotel of things: Internet of things and big data converging in hospitality. J Hosp Tour Technol 2019;11(1):93–107. http://dx.doi.org/10.1108/jhtt-12-2018-0120.

[20] Padma P, Ahn J. Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. Int J Hosp Manag 2020;84:102318. http://dx.doi.org/10.1016/j.ijhm.2019.102318.

[21] Dallemand J. Why is marriott the big data analytics leader in hospitality? 2019, [Case]. from https://blog.datumize.com/big-data-analytics-in-hospitality-marriott-international-case-study.

[22] Ferraris A, Mazzoleni A, Devalle A, Couturier J. Big data analytics capabilities and knowledge management: impact on firm performance. Manage Decis 2019;57(8):1923–36. http://dx.doi.org/10.1108/md-07-2018-0825.

[23] Kim Sohyeon, Lee Won Seok. Network text analysis of medical tourism in newspapers using text mining: the south korea case. Tourism Management Perspectives 2019;31:332–9. http://dx.doi.org/10.1016/j.tmp.2019.05.010.

[24] Phillips Paul, Barnes Stuart, Zigan Krystin, Schegg Roland. Understanding the impact of online reviews on hotel performance: an empirical analysis. Journal of Travel Research 2017;56(2):235–49. http://dx.doi.org/10.1177/0047287516636481.

[25] Fang Bin, Ye Qiang, Kucukusta Deniz, Law Rob. Analysis of the perceived value of online tourism reviews: influence of readability and reviewer characteristics. Tourism Management 2016;52:498–506. http://dx.doi.org/10.1016/j.tourman.2015.07.018.

[26] Kumar Arun, Gupta Manoj, Le Dac-Nhuong, Aly Ayman A. Pts-papr reduction technique for 5g advanced waveforms using bfo algorithm. Intelligent Automation and Soft Computing 2021;27(3):713–22. http://dx.doi.org/10.32604/iasc.2021.015793.

[27] Park Sangwon, Nicolau Juan L. Asymmetric effects of online consumer reviews. Annals of Tourism Research 2015;50:67–83. http://dx.doi.org/10.1016/j.annals.2014.10.007.

[28] Kumar Arun, Dhanagopal R, Albreem Mahmoud A, Le Dac-Nhuong. A comprehensive study on the role of advanced technologies in 5g based smart hospital. Alexandria Engineering Journal 2021;60(6):5527–36. http://dx.doi.org/10.1016/j.aej.2021.04.016.

[29] Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organisations. Technol Forecast Soc Change 2018;126:3–13. http://dx.doi.org/10.1016/j.techfore.2015.12.019.

[30] Akter S, Wamba SF. Big data analytics in E-commerce: a systematic review and agenda for future research. Electron Mark 2016;26(2):173–94. http://dx.doi.org/10.1007/s12525-016-0219-0.

Loveleen Gaur, Ph.D. is a Professor of Management and Information Systems in Amity International Business School, at Amity University, Noida, India. She is serving in various editorial board of journals of repute. Her area of research is Artificial Intelligence and Digital Marketing , Business Intelligence and Data Mining.

Anam Afaq is a Research Scholar and pursuing her Ph.D. in Marketing Management from Amity University Noida. Her area of interest in research includes Big Data Analytics and Digital Marketing. She is a certified Digital Marketer.

Arun Solanki is Assistant Professor in the Department of Computer Science and Engineering, Gautam Buddha University, Greater Noida, India. His research interests span Expert System, Machine Learning, and Search Engines. He is the editor/reviewer of many Books with reputed publisher journals.

Gurmeet Singh is a Professor and Head of Graduate School of Business at the University of the South Pacific in Fiji. His current research interests include Marketing, Information Systems, Social Entrepreneurship, Strategy, Project Management, Business Ethics, Open Innovation and Service Operations Management.

Shavneet Sharma is Assistant Lecturer in the School of Management and Public Administration at the University of the South Pacific. He is also a Certified Fraud Examiner (CFE) from USA. His research interests are information systems, technology adoption, digital marketing, and social media marketing.

N. Z. Jhanjhi currently with Taylor's University, Malaysia. He has an intensive background of academic quality accreditation in higher education besides scientific research activities, he had worked a decade for academic accreditation, and earned ABET accreditation twice for three programs at CCSIT, KFU Saudi Arabia.

Hoang Thi My has obtained an M.Sc. in traditional culture from Hanoi Culture University, Hanoi, Vietnam. Presently, She is working as a lecturer, faculty of Tourism, Haiphong University. She is having more than 12 years of teaching/research experience. Her research interests include traditional culture, religious beliefs, tourism and hotel management.

Dac-Nhuong Le is Associate Professor, Associate Dean of Faculty of Information Technology, Haiphong University, Vietnam. He is having 15+ years of teaching/research experience with 50+ reputed publications. His research interests include evolutionary multi-objective optimization, network security and cloud computing.