Journal of Business Research 128 (2021) 391–404

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Exploring the drivers of customers' brand attitudes of online travel agency services: A text-mining based approach

Arghya Ray [a,*], Pradip Kumar Bala [b], Nripendra P. Rana [c,*]

[a] FORE School of Management, Qutub Institutional Area, New Delhi 110016, India
[b] Indian Institute of Management Ranchi, Suchana Bhawan, 5th Floor, Audrey House Campus, Meur's Road, Ranchi, Jharkhand 834008, India
[c] School of Management, University of Bradford, Richmond Road, BD7 1DP Bradford, UK

ARTICLE INFO

Keywords:
Brand attitude
Online user-generated content
Online travel agency services
Qualitative aspects
Textual data

ABSTRACT

This paper aims to explore the important qualitative aspects of online user-generated-content that reflects customers' brand-attitudes. Additionally, the qualitative aspects can help service-providers understand customers' brand-attitudes by focusing on the important aspects rather than reading the entire review, which will save both their time and effort. We have utilised a total of 10,000 reviews from TripAdvisor (an online-travel-agency provider). This study has analysed the data using statistical-technique (logistic regression), predictive-model (artificial-neural-networks) and structural-modelling technique to understand the most important aspects (i.e. sentiment, emotion or parts-of-speech) that can help to predict customers' brand-attitudes. Results show that sentiment is the most important aspect in predicting brand-attitudes. While total sentiment content and content polarity have significant positive association, negative high-arousal emotions and low-arousal emotions have significant negative association with customers' brand attitudes. However, parts-of-speech aspects have no significant impact on brand attitude. The paper concludes with implications, limitations and future research directions.

1. Introduction

The penetration of Internet and wide popularity of smartphones have seen a changing trend in online shopping behaviour (McClure & Seock, 2020). Customers usually search in the website about various services or products, read the online reviews or social media posts, and decide on whether to use a service or product (Hong & Cha, 2013; Saumya & Singh, in press; Simonson & Rosen, 2014). The reviews are usually perceived as valuable and credible by the customers (Gupta & Harris, 2010; Mayzlin, 2006). Online customer reviews (OCRs) are considered the second most reliable information source for customers after they get suggestions from their near and dear ones (Chatterjee, 2020; Grimes, 2012). Researchers (e.g. Chen & Xie, 2008; Jiménez & Mendoza, 2013) have also voiced the importance of online reviews in affecting customers' various decisions (Chatterjee, 2019; Ray, Bala, Chakraborty, & Dasgupta, in press).

The availability of large amount of data present in online reviews helps managers of various service providers to understand the views of customers about the brands (Gensler, Völckner, Egger, Fischbach, & Schoder, 2016). In addition to the informal advice that is present in the user-generated content (UGC), the main advantage of using UGC is the lack of commercial biasness (East, Hammond, & Lomax, 2008). The second advantage of using UGC is that the data is easily accessible in real time and is widely available (Berger, Sorensen, & Rasmussen, 2010; Chatterjee, 2020). Third, the OCRs help in longitudinal analysis of changes in customer opinions or views over time (Gensler et al., 2016). Because of these advantages, there has been a rise in research using UGC in various contexts such as finding review helpfulness (Chatterjee, 2020), understanding customer perspectives (Ray, Bala, & Dwivedi, in press), and predicting customer ratings (Ray, Bala, & Jain, in press).

This study utilises data from online travel agency services (OTAs). OTAs refer to the online travel solutions that sells services related to travel such as linking customers to hotel, flight, etc. (Landman, 2020). OTAs are an important area of research, since the expected global revenues from OTAs is 1,955 billion USD by 2026 (ZionMarketResearch, 2018). Although researchers have explored different factors that affect customers' intention like, online reviews (Nisar, Hajli, Prabhakar, & Dwivedi, 2019), accessibility, navigation, reliability (van Riel, Semeijn, & Pauwels, 2004), process quality (Chen & Kao, 2010), information quality, customer safety (Tsang, Lai, & Law, 2010), and hedonic

* Corresponding authors.
E-mail addresses: mailarghyaray@gmail.com (A. Ray), pkbala@iimranchi.ac.in (P.K. Bala), nrananp@gmail.com (N.P. Rana).

https://doi.org/10.1016/j.jbusres.2021.02.028
Received 24 July 2020; Received in revised form 12 February 2021; Accepted 14 February 2021
Available online 28 February 2021

motivation (Tamilmani, Rana, Nunkoo, Raghavan, & Dwivedi, in press), in the context of OTAs, there is still a lack of research in understanding customers' brand attitudes from textual data. Branding is important for various service providers since it enables them to contest on features other than price (Bailey & Ball, 2006). Although brand attitude is a crucial precursor of repurchase intention (Gwang-In, Jung-Ae, & Gi-Jin, 2013; Priester, Nayakankuppam, Fleming, & Godek, 2004) and brand equity (Park, MacInnis, Priester, Eisingerich, & Iacobucci, 2010), there has been a significant gap in research on utilising customer reviews for understanding their attitude towards different brands (Zablocki, Schlegelmilch, & Houston, 2018).

Understanding the factors that affect customers' brand attitudes from their online posts can help business owners to not only formulate proper strategies for retaining customers but also to take necessary steps to reduce service issues that may exist. However, the research on capturing customers' brand attitudes from UGC is still in its nascent stage (Kostyra, Reiner, Natter, & Klapper, 2016; Zablocki et al., 2018). Although researchers have attempted to utilise UGC for finding out robustness of a brand (Ho-Dac, Carson, & Moore, 2013), effect of brand's image (Gensler et al., 2016), customers' selection of brands (Kostyra et al., 2016), and their attitude towards brands (Zablocki et al., 2018), researchers have not utilised the different qualitative aspects of textual data to understand how these aspects can help in understanding customers' brand attitudes. Gensler et al. (2016) affirmed that although OCRs are rich in information content, earlier scholars have often ignored the qualitative aspects hidden in texts and have rather focused on the overall quantitative aspects such as valence and volume of the product or service reviews (Zablocki et al., 2018). Second, understanding the emotional aspect of textual reviews can help business owners gain a deeper insight of customer perspectives. Researchers (e.g. Siering, Deokar, & Janze, 2018) also expressed concerns about the limited work on "emotional aspect". Additionally, from an organisation point of view, it is not an easy task to detect the issues raised by customers and filter out the bad-comments (Oheix, 2018), which can depict an unsatisfied customer from a large amount of customer reviews or posts. Exploring the textual aspects to understand customers' brand attitudes can be an added advantage for handling such issues.

Earlier research on understanding customer brand attitude was either quantitative-survey based analysis (Davtyan, Cunningham, & Tashchian, 2020; Wang, Cao, & Park, 2019) or qualitative interview based studies (Mattke, Muller, & Maier, 2019). There is a gap in research on understanding how UGC affect customer brand attitude (Zablocki et al., 2018). Even the studies utilising UGC for understanding how UGC affects brands is in the nascent stage (Zablocki et al., 2018). The studies (e.g. Gensler et al., 2016; Ho-Dac et al., 2013) on understanding the impact of online reviews on brands have utilised the quantitative aspects, like, number of positive and negative reviews, degree centrality, etc. and not the qualitative aspects of textual data, like, the emotional content and content polarity. Only a handful of studies (e.g. Bao, 2017; Zablocki et al., 2018) on analysing UGC for understanding brand attitude are conceptual papers. Hence, there is a need for research on understanding the various qualitative aspects of textual content that can help in understanding users' attitude towards various brands based on the experience they share through online reviews or posts. Additionally, although Zablocki et al. (2018) have proposed the use of valence score to measure brand attitude, the other qualitative aspects that can help to capture customer perspectives, like, total sentiment content (Chatterjee, 2020), high and low arousal emotions (Chatterjee, 2020), and parts of speech tags (Lee, Trimi, & Yang, 2017) have not been explored in earlier studies on brand attitude. This research will help researchers to utilise the power of emotional and sentiment aspects of textual data in various other analyses. Therefore, the research question that drives this research work is:

RQ1: What are the important qualitative aspects of UGC (sentiment aspects, emotional aspects or parts of speech tags) that predict customers' attitude towards various brands?
RQ2: Can sentiment aspects, emotional aspects and parts of speech tags help to understand customers' attitude towards brands?

While this study has theoretical contributions like, being the first paper to explore qualitative aspects from textual data to understand reviewers' brand attitudes and utilising sentiment and emotional scores to predict brand attitude, the managerial implications are multi-fold. The use of both qualitative and quantitative aspects of online user reviews will help managers understand the important aspects in different contexts. The results of this study also highlight how sentiment and emotional aspects in textual data helps to explore users' brand attitudes. The importance of the total number of sentiment filled words in reviews along with the valence of those words can help managers predict whether the customer will continue the service or not. Additionally, since OTA services represent the hospitality segment, the qualitative aspects play a vital role in affecting customers' electronic word of mouth intentions as well.

The remaining sections of this study are structured as follows: Section 2 provides a comprehensive review of the relevant literature. Section 3 provides an overview of the theoretical background, proposes conceptual model and formulates the hypotheses. Section 4 presents the research methodology adopted by this study. Section 5 presents the results whereas Section 6 provides discussion of results along with contributions to theory, implications for practice and limitations and future research directions. Finally, Section 7 presents the conclusion of this study.

2. Literature review

In this section, we first review some aspects of social media marketing and the need for understanding customers' brand attitudes. Second, we review earlier studies on UGC. Third, we discuss research on using UGC for understanding customers' attitude towards various brands. Fourth, we explore various qualitative aspects used by earlier researchers. Finally, we discuss how text mining has been used by researchers to explore different textual aspects.

2.1. Social media marketing

Social media marketing (SMM) deals with the use of various social media (SM) platforms for conducting marketing activities (Dahnil, Marzuki, Langgat, & Fabeil, 2014). In recent years, the popularity of SM platforms such as Facebook, Twitter, etc. have seen it emerge as an important marketing tool (Hawkins & Vel, 2013). SM marketing provides a cost-effective easy communication channel (Dwivedi, Kapoor, & Chen, 2015) for (1) fast and effective delivery of information (Shilbury, Westerbeek, Quick, Funk, & Karg, 2014), (2) fast and accurate feedback from customers (Shilbury et al., 2014), (3) better customer engagement (Abreza, O'Reilly, & Reid, 2013; Nisar, Prabhakar, & Patil, 2018), and (4) targeting a wider customer base (Weinberg, 2009). Drummond, O'Toole, and McGrath (2020) have proposed various tactics and strategies, like, information sharing posts, bilateral conversation, etc. based on the different SM marketing capability layers (connect, engage, coordinate, and collaborate).

Chatterjee and Kar (2020) found factors such as perceived usefulness and perceived ease of use influence SMM in the context of small and medium enterprises (SMEs) in India. Dwivedi et al. (2020) have also noted that SMM improves customer engagement behaviour. SM platforms can also help in increasing the firms' overall value (Kim, Koh, Cha, & Lee, 2015). Grover and Kar (2020) proposed an SM Engagement model to understand customer dynamics in case of how mobile wallet firms advertise on Twitter. However, researchers are divided in the way customer decision making is affected: some focus on online reviews and comments (Duan, Gu, & Whinston, 2008) whereas some others focus on pricing and promotions (Ajorlou, Jadbabaie, & Kakhbod, 2016), and customer psychology (Mowen, Park, & Zablah, 2007). For example, Jacobson, Gruzd, and Hernández-García (2020) found that marketing comfort influences SMM. Search Engine Marketing (SEM), if not done properly, not only affects customer benefits but also destroys firms' value (Aswani, Kar, Ilavarasan, & Dwivedi, 2018).

Hence, it is equally important for providers to focus on branding since it enables them to contest on aspects other than price (Bailey & Ball, 2006). In this competitive market scenario, it is essential to create a durable brand image for developing a strong emotional bond with customers (González-Mansilla, Berenguer-Contrí, & Serra-Cantallops, 2019). In case of online services, whether it is to choose food, restaurants or hotels, a similar trend is followed, i.e., users search in the website or mobile application based on their preferred needs, read or view recommendations and ratings to help them make decisions (Hong & Cha, 2013; Simonson & Rosen, 2014). In the tourism sector, researchers have explored the impact of review helpfulness, review credibility (Filieri, Acikgoz, Ndou, & Dwivedi, in press) and social-media (Nunkoo, Gursoy, & Dwivedi, 2020) on behavioural intention. Since customers have started perceiving online reviews and recommendations as an important and credible source of information (Gupta & Harris, 2010; Mayzlin, 2006), OCRs has been found helpful in influencing customer decisions (Chen & Xie, 2008; Jiménez & Mendoza, 2013) and is also considered to be useful in measuring customer brand attitude (Zablocki et al., 2018).

2.2. User-generated content (UGC)

UGC depicts the online content developed by users for sharing with others (Tang, Fang, & Wang, 2014). This can be retrieved from blogs (Christodoulides, Michaelidou, & Argyriou, 2012, p. 1689), social media platforms (Ray & Bala, 2020a), and merchandise websites (Ray et al., in press) to name a few. The online UGC is shared in various forms such as online ratings, reviews and feedback, blogs, experience sharing, etc. (Mishra & Satish, 2016). Researchers have not only voiced the strategic importance of UGC for service providers (Chatterjee, 2019), but have also mentioned the usefulness of online reviews and recommendations in case of prospective customers (Ray & Bala, 2020c; Ray et al., in press). UGC related to a product or service can be divided into two divisions, namely the qualitative as well as the quantitative part (Sridhar & Srinivasan, 2012). While the qualitative part contains detailed description of the customer experience related to the product or service (Jiménez & Mendoza, 2013), the quantitative part contains numeric details like ratings, which describes the customers' usage experience (Kostyra et al., 2016). Online reviews and ratings serve as great information sources for both service-providers and prospective customers (Chatterjee, 2019; Ray et al., in press). Online reviews and ratings are important from the business point of view because they drive profitability, revenues, etc. (Chevalier & Mayzlin, 2006; Duan et al., 2008). Similar to online recommendations (Cascio, O'Donnel, Bayer, Tinney, & Falk, 2015), online reviews and ratings guide users on 'using' or 'avoiding' a product or service (Zablocki et al., 2018). The other advantage of OCRs is that it provides an evaluation of the product/service by the customer. OCRs usually reflect the customer experience about a product or service (Zablocki et al., 2018). Additionally, the UGC also help to explore social-identity (Reyes-Menendez, Saura, & Thomas, 2020), identify Socio-Citiezenry factors (Aladwani & Dwivedi, 2018), examine brand communities (Kamboj, Sarmah, Gupta, & Dwivedi, 2018) and understand customers' brand attitudes from the online reviews and ratings. However, only a few studies have explored the effect of UGC on brand attitudes (Kostyra et al., 2016; Zablocki et al., 2018).

2.3. Capturing brand attitude from UGC

Brand attitude reflects what customers think about a product or service. Additionally, brand attitude is an essential driver of repurchase intention (Priester et al., 2004) and brand equity (Park et al., 2010). However, research on understanding how UGC influence brand attitude is still in its nascent stage (Kostyra et al., 2016). Brand attitude contains three main dimensions, namely, affective, cognitive, and behavioural components. The affective component deals with emotions related to brand attitude; the cognitive component deals with the evaluation of various factors like prior knowledge, beliefs, etc. related to brand attitude; whereas the behavioural component deals with the impact of brand attitude on customer behaviour (Zablocki et al., 2018). Roma and Aloini (2019) have examined how UGC related to a brand varies across different social media platforms. In this study, we focus on the affective-cognitive components of brand attitude.

We try to capture how the qualitative aspects (e.g. sentiment aspect, parts of speech tags, and emotional aspect) of the reviews posted by a user about a brand can help to understand the continuance/avoidance attitude of the individual towards the brand. For example, Ho-Dac et al. (2013) explored the impact of OCRs on a brand's strength whereas Gensler et al. (2016) assessed the influence of OCRs on brand image and brand associations. Moreover, Kostyra et al. (2016) examined the effect of UGC behind the choice of various brands. In addition, Ballantine and Yeung (2015) studied the relationship between valences of OCRs and brand attitude, perceived credibility and intention to buy.

Based on our search on Google Scholar using the keyword combinations of "brand attitude" "user generated content", and "online reviews" and their plural counterparts all in the title, we found that only two researchers have focused on utilising the textual aspects of user-generated content to explore customers' brand attitudes (refer Table 1). However, both these papers have just proposed a conceptual framework and hypotheses. For example, Bao (2017) has proposed the use of richness, usefulness, and sociality of UGC to capture brand attitude whereas Zablocki et al. (2018) utilised both the qualitative and quantitative aspects of textual data to frame a conceptual model to analyse the effect of valence of OCRs on brand attitude. The authors have used factors like variance, volume, brand types, and source as moderators of the association between the valence and brand attitude. Though researchers have voiced the importance of OCRs as electronic word of mouth (eWOM) on influencing purchase intention (Floyd, Freling, Alhoqail, Cho, & Freling, 2014), and the effect of both OCRs and brand attitude on intentions and brand equity (Priester et al., 2004), the lack of studies on influence of OCRs on brand attitude is visible. In this study, we attempt to contribute to the existing work on the influence of OCRs on brand attitude.

2.4. Qualitative aspects of UGC used in earlier studies

Researchers (e.g. Chatterjee, 2020; Lee et al., 2017; Zablocki et al.,

Table 1
Review of literature on understanding brand attitude from UGC.

Authors	Variables used	Methods Used	Findings
Bao (2017)	**Independent**: Richness, Usefulness, and Sociality of UGC **Dependent:** Online brand experience, brand attitude	Conceptual Paper	The researcher has presented a theoretical model to find the effect of different variables on brand attitude.
Zablocki et al. (2018)	**Independent**: Valence of OCR **Moderator**: Volume of OCR, Brand Type, Variance of OCR, Source of review **Dependent**: Brand Attitude	Conceptual Paper	The researcher has presented a theoretical model to find the effect of different moderators on relationship between valences of online reviews on brand attitude.

2018) over the years have used various qualitative aspects of textual data to understand drivers of customer behaviour in different contexts like, review helpfulness (Chatterjee, 2020), and human–computer-interaction design (Lee, Choi, Marakas, & Singh, 2018). Rathore and Ilavarasan (2020) have examined the emotions expressed by customers based on tweets captured 'before the launch' and 'after the launch' in case of three different products. Chatterjee (2020) has used the different qualitative aspects, namely, total sentiment content, the content polarity, the different emotional aspects (low and high arousal emotions) and the review length to explore review helpfulness. Lee et al. (2018) have mentioned many other emotions such as cheerful, happy, elated, quiescent, relaxed, calm, serene, disappointed, gloomy, dejected, nervous, restless, tense and agitated. We have decided to use the arousal emotions as mentioned by Chatterjee (2020). Lee et al. (2017) used the different qualitative aspects of textual data to capture reviewers' tendency (through positive score and negative score), and reviewer characteristics (through use of various parts of speech tags, namely, the count of nouns, verbs, adjectives and adverbs used) while evaluating the effect on review helpfulness. We have not explored dominant emotions in this study because these emotions are difficult to define for an overall population. Table 2 summarised the related studies from where we derived the scale items used in this study.

2.5. Text mining approaches for exploring UGC

Text mining deals with the process of extracting knowledge from textual content. Researchers have used various text-mining approaches to explore customer perspectives from online UGC such as topic modeling (Blei, 2012), Naïve Bayes approach (Ray & Bala, 2020a, 2020b), sentiment and emotion mining (Chatterjee, 2020), NLP-SEM (Ray et al., in press), finding readability metrics (Bafna & Saini, 2021), etc. Researchers have been using text mining approaches in different contexts like, finding review helpfulness (Chatterjee, 2020), predicting customer ratings (Ray et al., in press), understanding values customers prefer (Ray et al., in press), improving disaster management processes (Ray & Bala, 2020a). Researchers have usually used sentiment analysis in contexts like analysing movie reviews (Anandarajan, Hill, & Nolan, 2019), product reviews (Vyas & Uma, 2019), etc. Another aspect of customer behaviour can be understood from the textual data by extracting the emotion scores (anger, fear, disgust, sadness, joy, surprise, anticipation, and trust) from UGC (Chatterjee, 2020). In this work, we utilise text-mining and emotion mining to derive qualitative aspects from UGC for exploring the impact of the qualitative aspects of textual data on predicting customers' brand attitudes.

3. Theoretical background and hypotheses development

In this modern era, understanding brand attitude helps the service-providers to cater to the customers in an improved manner (Kostyra et al., 2016). Limited studies have utilised user-generated content to understand customers' brand attitudes (Zablocki et al., 2018). Additionally, Zablocki et al. (2018) have voiced concerns regarding the lack of studies utilising the qualitative aspects of textual data. They utilised both the qualitative and quantitative aspects of textual data to frame a conceptual model for analysing the effect of the valence of OCRs on brand attitude. They further proposed predicting brand attitude using valence-score as the only independent variable and four moderators, namely, volume of OCRs, variance of OCRs, brand-type and the source of the review. The volume, variance, and brand-type are not relevant since this study's attempts at predicting the reviewers' brand attitudes.

Table 2
Relevant earlier works from where the qualitative aspects used in this study.

Qualitative Aspect	Authors	Our comments
Total sentiment content, content polarity, low and high arousal emotions	Chatterjee (2020)	The authors have used the qualitative aspects as well as quantitative aspects of UGC to understand drivers of review helpfulness.
Positive score, Negative score, and Parts of speech tags like count of nouns, verbs, adjectives and adverbs	Lee et al. (2017)	Researchers have used these aspects to analyse the effect on review helpfulness.

Although Zablocki et al. (2018) used valence as a single independent variable, Chatterjee (2020) used two aspects to capture the sentiment scores, namely, the total sentiment content and the content polarity. While the total sentiment content refers to the overall sentiment captured, content polarity refers to the valence score as obtained from the difference of the number of positive and negative words. This helps to capture the sentiment present in the textual data in a better way. Additionally, Chatterjee (2020) also used the emotional aspects, namely, high arousal emotions (such as anger and fear) and low arousal emotions (such as sadness) to examine review helpfulness. Moreover, Lee et al. (2017) utilised different qualitative aspects of textual data like, sentiment polarity and count of parts of speech tags (noun, verb, adjective and adverb) to capture review tendency and reviewer characteristics for understanding review helpfulness. None of the existing studies have utilised the emotional aspects and parts of speech tags to understand brand attitude. Based on these studies (e.g. Chatterjee, 2020; Lee et al., 2017; Zablocki et al., 2018), the constructs we have used in this study are as follows.

The sentiment aspect is divided into total sentiment content and content polarity (Chatterjee, 2020). The emotional aspect is divided into high and low arousal emotions (Chatterjee, 2020). The parts of speech tags used in the textual data are divided into noun counts, verb counts, adjective counts and adverb counts (Lee et al., 2017). Fig. 1 depicts the conceptual model used in this study. The constructs used for this study and the relevant sources are presented in Table 2. We have three main independent variables (i.e. sentiment aspects, emotional aspects and parts of speech aspects), and one dependent variable (brand attitude). All the constructs used for predicting customers' brand attitudes has not been used in earlier studies for understanding this construct. Although Zablocki et al. (2018) used valence score to reflect the sentiment aspects; we have divided sentiment aspects into content polarity and total sentiment content to extract more information from the textual content as proposed by Chatterjee (2020) in his work.

3.1. Relationship between sentiment aspects and brand attitude

The total sentiment content reflects whether the online customer review contains rich sentiment content (Chatterjee, 2020). Filieri (2016) have noted that reviews containing good description of content, and those having equal amount of positive and negative views to be more beneficial for the readers. Cao, Duan, and Gan (2011) have also noted that a review containing both positive and negative opinions is more helpful (Salehan & Kim, 2016) since a two-sided argument is more compelling (Crowley & Hoyer, 1994). Chatterjee (2020) stated that the total sentiment content of an online customer reviews is denoted by the overall number of positive and negative words present. Therefore, the following hypothesis can be formulated:

H1a: Higher total sentiment content present in the online customer review reflects stronger brand attitude.

Polarity of the text denotes the degree to which the OCR is either positive or negative (Geetha, Singha, & Sinha, 2017). Researchers have noted that extreme polarity reviews are perceived as untrustworthy by the customers (Filieri, 2016). Hence, keeping this aspect in mind, it indicates that the higher the polarity of the OCR, the lesser will be the perceived customer attitude towards the brand. Researchers (Kostyra et al., 2016; Zablocki et al., 2018) have noted a direct effect of valence score on brand attitude. We also propose that the total sentiment content

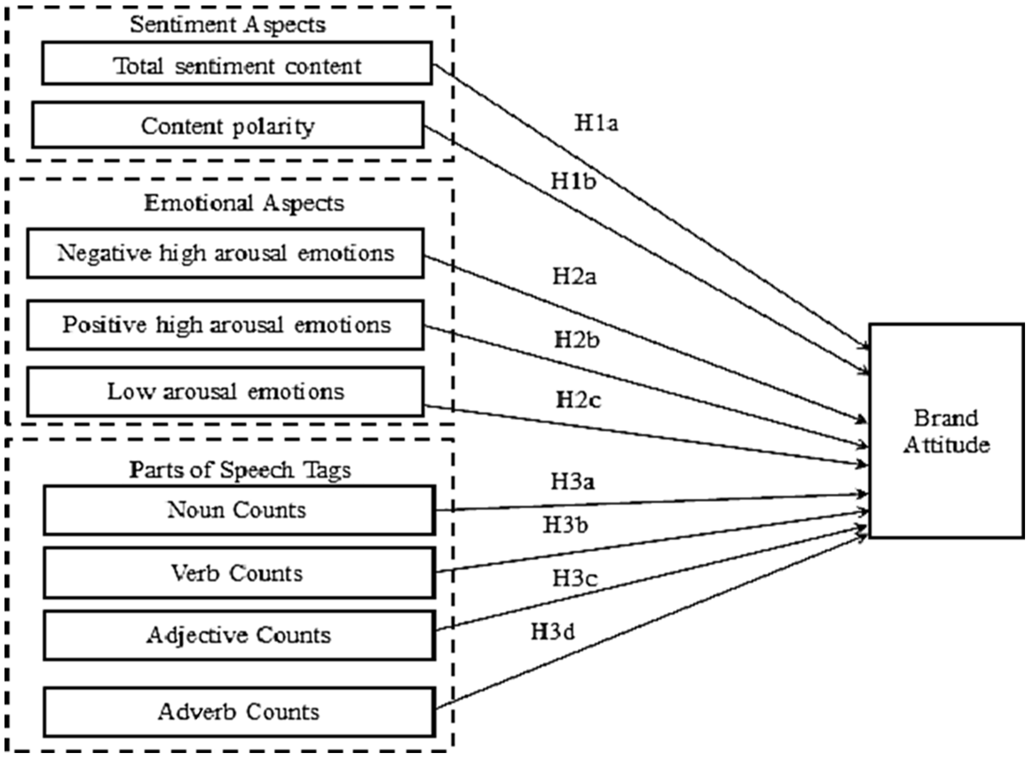


Fig. 1. Proposed conceptual model (Adapted from Zablocki et al., 2018; Chatterjee, 2019; Lee et al., 2017). [Note: The constructs used in this study had not been used earlier to understand brand attitude from textual data.]

and the extremeness of valence scores in the online textual reviews will have a negative impact on brand attitude. Thus, we propose:

H1b: Higher polarity of the online customer review reflects an unfavorable customer attitude towards the brand.

3.2. Relationship between emotional aspects and brand attitude

The emotional aspect present in textual data helps to have a deeper understanding of the textual data. In this modern era, due to increase in the number of fraudulent activities, customers perceive most reviews to be fake and the reviews with negative emotions to be more credible (Chatterjee, 2020). We have also utilised one important dimension of emotions, which is related to the level of arousal. Cavanaugh, MacInnis, and Weiss (2015) have divided emotions into various dimensions like anger and fear belong to high arousal negative emotions, sadness belongs to the low arousal negative emotion category, etc. Researchers believe that reviews denoting high arousal emotions are supposed to be less credible (Chatterjee, 2020) and are often less informative in nature because the cognitive ability of humans does not function properly under high arousal emotions (Filieri, 2016; Salehan & Kim, 2016) and hence affects brand attitude negatively.

Negative high arousal emotions reflect deep negative feelings of a user related to some experience. The negative high arousal emotions are reflected through anger and fear (Cavanaugh et al., 2015) that a user expresses. These emotions reflect dissatisfaction from a user (Cavanaugh et al., 2015). However, researchers have also noted that high arousal emotions are considered less credible (Chatterjee, 2020). We feel that strong negative emotions like anger and fear express dissatisfied customers, because they might have faced some issues because of which they may not use the service or product in future. Therefore, we propose:

H2a: Higher negative high arousal emotions reflect unfavorable brand attitude.

Positive high arousal emotions usually reflect highly satisfied customers. Researchers have usually considered trust and joy as high arousal positive emotions (Cavanaugh et al., 2015). Positive high arousal emotions help to increase the satisfaction of customers (Cavanaugh et al., 2015). Liu et al. (2017) noted that positive high arousal emotions help to increase the motivation and enjoyment in students which ultimately enhance their academic involvement. Sas and Zhang (2010) also stated that positive high arousal emotions increase motivation. Thus, in line with the above discussions, we also feel that positive high arousal emotions will reflect strong brand attitude. Thus, we hypothesise:

H2b: Greater positive high arousal emotions reflect the stronger brand attitude.

The low arousal emotions reflect subtle emotional changes in a person and are generally considered more credible (Chatterjee, 2020). These emotions are captured by the feelings sadness and anticipation (Cavanaugh et al., 2015). When a user does not express some strong emotions, they have low arousal emotions. Low arousal emotions are perceived to be neutral and take time to arouse as compared to high arousal emotions (Rasa, Leonidas, & Gintautas, 2020). Since low arousal emotions portray neutral feelings, we feel that it will not reflect a strong brand attitude. Hence we propose:

H2c: The low arousal emotions reflect moderate negative influence on perceived brand attitude.

3.3. Relationship between parts of speech aspects, review length and brand attitude

In addition to sentiment and emotional aspects of textual data, it is also essential to consider the content present in the online review (Lee et al., 2017). The writing style of a user and the choice of words they use for describing the experience reflect their attitude towards a product/

service (Lee et al., 2017). The prior studies (e.g. Saumya, Singh, Baabdullah, Rana, & Dwivedi, 2018; Singh et al., 2018) have utilised the different parts of speech aspects like noun, adjective and adverb count to understand review helpfulness and rank the reviews. The researchers have also noted that parts of speech aspects are helpful in predicting user behaviour. Although the style of writing the review reflects users' attitude towards the brand, only a limited studies have taken the parts of speech aspect into consideration. Thus, we propose:

H3: The count of (a) nouns (b) verbs (c) adjectives (d) adverbs in a review is positively associated with users' brand attitude.

Researchers (Saumya et al., 2018; Singh et al., 2018) have noted that review-length has a major influence on customers' behaviour (Schindler & Bickart, 2012). This is due to the fact that longer reviews are perceived to be rich in information content (Schwenk, 1986) and are more persuasive in nature (Tversky & Kahneman, 1974). Lee et al. (2017) found that the parts of speech tags have an influence on review length. Additionally, the whole review also reflects the overall sentiment polarity due to the presence of positive and negative words (Salehan & Kim, 2016). Thus, along with the above hypotheses, we also propose that an interaction between the total number of parts of speech tags and content polarity will reflect users' brand attitude.

H4: The interaction between content polarity and total parts of speech count will reflect users' attitude towards the brand.

For positive polarity, users are supposed to have a positive brand attitude and vice versa. Additionally, we also propose the following hypotheses:

Sentiment analysis deals with analysing the valence score (positive, negative or neutral) for understanding the writer's attitude (Turney, 2002). Researchers have used sentiment analysis in various contexts like, analysing movie reviews (Anandarajan et al., 2019), product reviews (Vyas & Uma, 2019), etc. Chatterjee (2020) used two aspects to capture the sentiment scores, namely, the total sentiment content and the content polarity. Zablocki et al. (2018) have proposed that the text valence has a direct effect on brand attitude. In line with the above discussion, we propose:

H5: Sentiment aspects are a good reflector of users' attitude towards various brands.

Apart from understanding the sentiment score of textual data, it is also important to capture the emotional content present in textual data to understand the affective-cognitive view of UGC (Chatterjee, 2020). The emotional scores (i.e. anger, fear, disgust, sadness, joy, surprise, anticipation, and trust) provide an additional dimension to analyse the users' perspectives. Researchers have used emotions in different contexts like, analysing reader's and writer's perspectives (Yang, Lin, & Chen, 2009), analysis of articles (Alsmearat, Shehab, Al-Ayyoub, Al-Shalabi, & Kanaan, 2015), etc. Chatterjee (2020) also noted that emotions are important predictors of review helpfulness. Earlier studies (e.g. Zablocki et al., 2018; Bao, 2017) have not proposed the use of emotional aspects to predict brand attitude. We feel that emotional aspects will be helpful in predicting customers' brand attitudes. Hence we propose:

H6: Emotional aspects are good reflectors of customer attitude towards various brands.

Apart from sentiment and emotional aspects, researchers (e.g. Lee et al., 2017; Singh et al., 2018) have also focused on capturing the different parts of speech word counts to understand the writing style of users and thus predict their behaviour. Prior studies (e.g. Lee et al., 2017; Singh et al., 2018) have found that parts of speech aspects are also helpful in understanding drivers of review helpfulness. Although researchers (Bao, 2017; Zablocki et al., 2018) have not proposed the use of parts of speech aspects to predict brand attitude, we feel that parts of speech aspects will be helpful in predicting customers' brand attitudes. Hence we hypothesise:

H7: The parts of speech aspects are good reflectors of customer attitude towards various brands.

4. Research methodology

4.1. Data description

We have used 10,000 reviews from the TripAdvisor (an online travel agency) dataset, which contains around 500,000 reviews (Datafiniti, 2019). Reviews having a minimum of 100 words were separated and then 10,000 reviews were selected randomly based on an algorithm. Hence, the final OTA dataset contains lengthy reviews (minimum 100 words). Using sentences with having a minimum of 100 words ensure that the sentiment and emotional aspects have been captured properly (Watson, 2020). The dataset has 6,823 unique user ids and reviews for 1,667 hotels. Although this dataset contains user reviews related to their experience of staying in a hotel, we have used each review as a separate independent entity. Hence, each review by a user is used to understand his/her brand attitude based on the learning from the whole dataset, which contains reviews given by all users about different hotels. We have used both statistical and machine learning techniques in this study. Additionally, combining each review's qualitative aspects and the respective predicted brand attitude can help to form the dataset for analysing the aspects that are most important in explaining users' brand attitudes.

4.2. Data processing

For processing textual data, researchers have utilised several techniques like, frequency analysis, topic modeling, topic analysis, etc. (Karami, Lundy, Webb, & Dwivedi, 2020; Ray & Bala, 2020c) in various contexts like, theory building (Kar & Dwivedi, 2020), SM prediction (Rousidis, Koukaras, & Tjortjis, 2020), etc. In this work, we have used text-mining (topic modeling and topic frequency) and sentiment and emotion-mining to prepare the dataset from the qualitative UGC. While topic modeling helps to generate an estimated probability of the topics and words used to represent the document (Ray et al., in press), sentiment and emotion-mining helps to understand the underlying hidden feelings engraved in the reviews (Ray et al., in press). Topic frequency helps to find the frequency of words related to a particular context (Karami et al., 2020).

In our analysis, we have considered each review as a separate document similar to what Eusebius, Parackal, and Gnoth (2016) had proposed. We have utilised the techniques mentioned by the researchers (e.g. Chatterjee, 2020; Ho-Dac et al., 2013; Zablocki et al., 2018) to formulate the dataset for our analysis. We have used the standard R-libraries "sentimentr" and "syuzhet" for calculating the total sentiment score and the various emotional scores. The package "sentimentr" is used for finding out the polarity scores (Jockers, 2017). The underlying algorithm uses a pre-defined table containing a collection of positive and negative scores to determine the overall polarity scores of each sentence. The advantage of using "sentimentr" package is that it not only checks for inversions, but also takes into account valence shifters, like, the negators, the amplifiers, the de-amplifiers, and the adversative conjunctions (Document, 2019). Negators are the words that negate the sentence meaning like, "not". Amplifiers (or intensifiers), like, "seriously", enhance the meaning of the sentence. De-amplifiers (or downtoners), like, "barely", decrease the intensity of the sentence. Adversative conjunctions trump the previous clause. For example, "The services are good but not up-to-date" (Valence shifters, 2019). For calculating the emotion scores, we have used the "syuzhet" package (Document, 2017;

Jockers, 2017), which utilises the NRC Emotion Library containing emotion lexicons (Mohammad & Turney, 2011). For capturing the data which defines brand attitude we have used the method suggested by Ho-Dac et al. (2013) in which we have tried to extract only those probabilities that contain topics like 'love', 'like', 'satisfied', 'hate', 'good', etc.

The dataset for the other variables were created based on the definitions e.g. for content polarity users value comments which have both positive and negative words rather than those which are having higher polarity. Based on the main emotions fear, anger, joy, surprise, anticipation, sadness, trust and disgust, we have utilised fear and anger to represent negative-high-arousal-emotions; trust and joy as positive-high-arousal-emotions; and sadness and anticipation as low-arousal-emotions (Cavanaugh et al., 2015). For the parts of speech aspect counts, the count of the verbs, adjectives, etc. were taken for each review. The scores received from sentiment and emotional analysis related to the variables under study, the parts of speech aspect counts were normalised in a scale of '1' to '5' so that the influence of the variables can be captured properly.

4.3. Data analysis

We have used econometric (logistic regression), structural modeling (NLP-SEM) and machine-learning (artificial neural networks (ANN)) techniques to identify the important aspects of textual data that affects brand attitude. The logistic regression helps in predictive analysis, in describing the associations between the dependent binary variable and the independent variable(s) (Statistics Solutions, 2010), and for answering questions which results in a 'yes' or 'no' answer (in our case, positive brand attitude or negative brand attitude) (Vito, Schaefer, Higgins, Marcum, & Ricketts, 2019; Yan, Radwan, & Abdel-Aty, 2005). Since the key objective of this study is to understand the important predictors of brand attitude, we have used logistic regression to analyse the data. Although there are other alternatives of logistic regression available including decision tree classifier, Poisson regression, neural networks and support vector machines (SVM) (Glen, 2019), the techniques like neural networks and SVM are mainly considered under machine learning algorithms. Additionally, the statistical technique such as Poisson regression has a tendency to provide conservative estimates for confidence intervals (Glen, 2019). The advantage of using logistic regression is that it ensures good accuracy in case of simple and linearly separable datasets (GeeksforGeeks, 2020). Additionally, it has been found that logistic regression and step-wise logistic regression perform better than even k-nearest neighbours and decision trees based on lift scores (Bichler & Kiss, 2004; Glen, 2019). The ANN is used since it helps to handle noisy data and helps to provide the normalised importance of the predictors. Although there are other alternative techniques available, like, SVM, Random Forest, etc., we have preferred ANN because ANN can be helpful for handling multi-class problems and that with the increase in data, ANNs are able to generalise better and can provide more accurate results (Prem, 2020). The NLP-SEM is used to explore the overall influence of different aspects on predicting users' attitude towards OTAs. This is a new technique proposed by Ray et al. (in press), which helps to understand user perspectives from user-generated content through structural model analysis. We wanted to find the importance of each aspect as evident from the online reviews. Hence we have used the NLP-SEM technique. We have used SPSS statistics, Python and R for analysing the datasets. Table 3 presents the sample statistics:

5. Results

5.1. Explanatory model

The model as described in the Equation (1) below is used to explain the attitude towards brands as reflected in the data:

Table 3
Descriptive statistics of the variables under study.

Variable	Mean	Standard Deviation
Total Sentiment Content	4.449	3.028
Content Polarity	0.513	0.423
Negative High Arousal Emotions	0.595	1.103
Positive High Arousal Emotions	3.872	2.989
Low Arousal Emotions	2.102	1.735
Noun Count	11.701	9.520
Verb Count	3.775	4.634
Adjective Count	4.780	3.959
Adverb Count	1.840	2.087

$$PBA = f(\beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC) \quad (1)$$

Additionally, we have also tested for an interaction between total parts of speech count and the content polarity. Hence, the Equation (2) is as given below:

$$PBA = f(\beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC + \beta_{10} \times CP \times TPOSC) \quad (2)$$

[Note: PBA = Perceived Brand Attitude; TSC = Total Sentiment Content; CP = Content Polarity; NHAE = Negative High Arousal Emotions; PHAE = Negative High Arousal Emotions; LAE = Low Arousal Emotions; NC = Noun Counts; VC = Verb Counts; Adjective Counts; ADVC = Adverb Counts; TPOSC = Total Parts of speech Count]

In both Equation (1) and Equation (2), the perceived brand attitude (denoted by PBA) is the only response variable. The other variables such as total sentiment content, content polarity, negative high arousal emotions, positive high arousal emotions, low arousal emotions, verb count, noun count, adjective count and adverb count act as predictor variables. While Equation (1) is approximated by the variables, total sentiment content, content polarity, negative high arousal emotions, positive high arousal emotions, low arousal emotions, verb count, noun count, adjective count and adverb count, it does not take into account the interaction effects. However, interaction effects are also important since the words representing total parts of speech count can also imply the overall sentiment of the review. This is represented in Equation (2). We have mainly used logistic regression to approximate the functions mentioned in Equation (1) and Equation (2). Since the function will predict whether the user has a positive or negative brand attitude, it has just two outcomes and hence logistic regression is preferred (Statistics Solutions, 2010). Thus, Equation (1) and Equation (2) can be represented as Equation (3) and Equation (4) respectively using the link function g(PBA).

$$g(PBA) = \beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC \quad (3)$$

$$g(PBA) = \beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC + \beta_{10} \times CP \times TPOSC \quad (4)$$

The link function g(PBA) is used to represent generally two criteria: (a) the probability of success (p) and the probability of failure (1-p), where, p should always be positive (since p >= 0) and must always be less than equal to 1. Thus, equations (3) and (4) can further be represented as Equation (5) and Equation (6) respectively.

$$\log(p/1 - p) = \beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC \quad (5)$$

$$\log(p/1 - p) = \beta_0 + \beta_1 \times TSC + \beta_2 \times CP + \beta_3 \times NHAE + \beta_4 \times PHAE + \beta_5 \times LAE + \beta_6 \times NC + \beta_7 \times VC + \beta_8 \times ADJC + \beta_9 \times ADVC + \beta_{10} \times CP \times TPOSC \quad (6)$$

The correlation coefficient for the variables is tested using Pearson

Correlation and presented in Table 4. In the correlation matrix, we noticed that some variables are somewhat highly correlated. Hence, we checked for the variation inflation factor (VIF) of the variables and found that there is no problem of multi-collinearity (with VIF < 5).

The binary logistic regression helps in predictive analysis and in describing the associations between the dependent binary variable and the independent variable(s) (Statistics Solutions, 2010). Logistic regression is helpful for answering questions which results in a 'yes' or 'no' answer (in our case, positive brand attitude or negative brand attitude) (Vito et al., 2019; Yan et al., 2005). Result of the logistic regression is shown in Table 5. The Hosmer and Lemeshow goodness of fit criteria was fulfilled by the significant variables. The omnibus test showed good fit for model coefficients. Additionally the Nagelkerke R-square value was 0.475.

The results of the logistic regression indicate that there is no significant association between positive high arousal emotions and parts of speech tag counts (namely, noun counts, verb counts, adverb counts, adverb counts) with customers' attitude towards various brands (see Table 5). Thus, hypotheses H2b, H3a, H3b, and H3c are not supported. The results also suggest a significant positive association between total sentiment content and content polarity with users' perceived brand attitude, and a significant negative association between negative high arousal emotions and low arousal emotions with users' perceived brand attitude. Thus, H1a, H1b, H2a, and H2c are supported. Additionally, we find that the interaction between total parts of speech tag count and content polarity has a non-significant association with users' brand attitude. Therefore, H4 is not supported.

5.2. Structural equation modelling

The measurement model was tested using PLS-SEM technique. The reliability and discriminant validity scores are presented in Table 6. The items demonstrated good reliability (composite reliability and average-variance-extracted scores > 0.5) and discriminant validity scores (Hair, Black, Babin, & Anderson, 2010; Hair, Ringle, & Sarstedt, 2013). The Cronbach's alpha score for sentiment analysis was low. The potential reason for this could be because both total sentiment content and content polarity use the same positive and negative word counts. We checked the VIF score and the results showed VIF < 5. We have decided to use both these variables as they are important for the purpose of our analysis.

The model shows good fit indices i.e. standardised root-mean-square-error (SRMR) (0.08) and normed-fit-index (NFI) score (0.797). The hypotheses results and the path coefficients are presented in Table 7 and Fig. 2 respectively.

Results of the analysis show that sentiment and emotional aspects are significant drivers of brand attitude with sentiment aspects being the more important predictor. Parts of speech aspects have a non-significant impact. This has also been found in the regression analysis. Hence, hypotheses H5 and H6 hold. However, Hypothesis H7 is rejected.

Table 5
Results of logistic regression explaining brand attitude.

	(1)	(2)	Supported?
(Intercept)	−4.725***	−4.781***	Yes
Total sentiment content	0.166**	0.166**	Yes
Content polarity	4.133***	4.194***	Yes
Negative high arousal emotions	−2.434***	−2.425***	Yes
Positive high arousal emotions	−0.052^	−0.052^	No
Low arousal emotions	−2.092***	−2.090***	Yes
Noun Counts	0.005^	0.008^	No
Verb Counts	0.040^	0.043^	No
Adjective Counts	0.054^	0.057^	No
Adverb Counts	0.006^	0.011^	No
Total Parts of speech tags × Content Polarity	–	−0.004^	No

[Note: **: p < 0.01; ***: p < 0.001; ^: Non Significant, (1) denotes the results of the model based on Equation (1) (2) denotes the findings of the model based on Equation (2).]

5.3. Predictive model

Along with the econometric and SEM techniques, we have used a machine learning technique, since machine learning technique (in this case, ANN) has the capability to analyse complex datasets and reduces the effects due to noise (Akgül, 2018). Akgül (2018) has also stated that since SEM can lead to oversimplification of the dynamic dataset, it is usually preferred to conduct an SEM-ANN analysis to verify the output of the SEM analysis and find out the order of importance. The ANN contains various neurons in multiple interconnected networks, which are mainly specified as the input layer, hidden layer and the output layer. In the ANN, every neuron has an activation function and every connection has weights assigned to them (Han, Pei, & Kamber, 2011). In this study, we have used the sigmoid activation function as proposed by earlier researchers (Fuller, Biros, & Wilson, 2009). The dataset was divided into 70:30 ratio for training and testing purpose. The model's performance was examined using the root-mean-square-error (RMSE). To minimise the sampling influence while creating the training and testing datasets, the operation was run five times and the standard deviation and RMSE values are reported. Findings from the ANN analysis are presented in Table 8.

When we considered only the important factors that emerged from the statistical and SEM-based analysis, the results of the study revealed that sentiment aspect is a more important predictor (68%) than emotional aspects (32%). This is also revealed through Fig. 3(a). The RMSE of the model was 0.4701 for the training dataset and 0.4669 for the test dataset. We also noted that the standard deviations noted for the training set were 0.0097 and 0.0063 for the testing dataset. Fig. 3(b) shows the importance of different sentiment and emotional aspects. The independent variable 'importance' score depicts how much the value predicted by the ANN varies with different values of the independent variable (Chong, 2013). We find that content polarity (42.1%) has the highest importance followed by total sentiment content (28.8%) and

Table 4
Correlation matrix for the independent variables in context of OTAs.

	TSC	CP	NHAE	PHAE	LAE	NC	VC	ADJC	ADVC
TSC	1								
CP	0.198	1							
NHAE	0.513	−0.199	1						
PHAE	0.769	0.387	0.259	1					
LAE	0.710	0.143	0.462	0.662	1				
NC	0.746	−0.015	0.500	0.517	0.539	1			
VC	0.671	−0.092	0.490	0.441	0.533	0.816	1		
ADJC	0.753	0.078	0.460	0.552	0.550	0.799	0.735	1	
ADVC	0.595	0.066	0.380	0.453	0.495	0.651	0.688	0.656	1

[Note: TSC = Total sentiment content; CP = content polarity; NHAE = Negative High Arousal Emotions; PHAE = Positive High Arousal Emotions; LAE = Low Arousal Emotions; NC = Noun Count; VC = Verb Count; ADJC = Adjective Count; ADVC = Adverb Count. Correlation is significant at 0.01 level (2-tailed)]

Table 6
Checking for consistency and accuracy measures of the scale items.

	Composite Reliability	Average Variance Extracted	Brand Attitude	Emotional Aspects	Parts of speech Aspects	Sentiment Aspects
Brand Attitude	0.782	0.642	0.801			
Emotional Aspects	0.798	0.571	0.403	0.755		
Parts of speech Aspects	0.903	0.700	0.306	0.629	0.837	
Sentiment Aspects	0.686	0.523	0.581	0.503	0.441	0.723

Table 7
Hypotheses results.

Hypothesis	β-values, p-values	Supported?
H5: Sentiment Aspects → Brand Attitude	$\beta = 0.509, p < 0.001$	Yes
H6: Emotional Aspects → Brand Attitude	$\beta = 0.159, p < 0.001$	Yes
H7: Parts-of-speech Aspects → Brand Attitude	$\beta = -0.019, p > 0.1$	No

negative high arousal emotions (25.4%) while predicting users' brand attitudes.

An interesting finding is noted in Table 8. We find that when we include the overall parts of speech aspects, namely, the nouns, verbs, adjectives and adverbs, we find that the performance of the ANN improves. Additionally, when we include the interaction effect there is a further improvement of the ANN performance. However, our earlier analysis shows that parts of speech aspects and the interaction effect have no significant association with brand attitude.

6. Discussion

Earlier studies (e.g. Bao, 2017; Zablocki et al., 2018) on utilising qualitative and quantitative aspects to understand brand attitude are conceptual papers. Additionally, Zablocki et al. (2018) need to take into account the reviews provided by other reviewers as well since it takes into account the volume aspect as well. Hence, earlier studies have not empirically examined the important qualitative or quantitative aspects that affect users' brand attitudes. This study aims to address the gap in literature on understanding the qualitative aspects of textual data that can predict the reviewers' brand attitudes.

For the purpose of this study, we have used econometric model (logistic regression), predictive model (ANN) and structural model for analysing the OTA dataset. Unlike what Zablocki et al. (2018) has proposed, the use of different techniques helps to get a better accurate view of the aspects that affect users' brand attitudes. Additionally, it is important to consider each review separately to predict the brand attitude of the user based on the different aspects present in the review. Although researchers (e.g. Chatterjee, 2020; Lee et al., 2017; Singh et al., 2018) have noted the importance of aspects like content polarity, total sentiment content, noun count, adjective count, adverb count, low arousal emotions, high arousal emotions, etc. in predicting review helpfulness, researchers (Bao, 2017; Zablocki et al., 2018) working on online textual data for understanding users' brand attitudes has only focused on valence, richness, sociality, and usefulness for predicting brand attitude.

Thus, the use of emotional aspects and parts of speech aspects help to improve the understanding of users' brand attitudes from online textual data. Results of the logistic regression revealed that total sentiment content and content polarity are positively associated with users' brand attitudes. This is in line with previous research studies (Cao et al., 2011; Filieri, 2016; Geetha et al., 2017). This means that if the reviews contain a good mix of positive and negative words, as well as if the polarity of the review is more towards neutral-positive, the users are supposed to have a positive brand attitude. Additionally, results also suggest that when the reviews of users contain words depicting fear, anger, sadness or anticipation, the users will be reluctant to use that brand. Researchers have found that high arousal emotions (Chatterjee, 2020; Salehan & Kim, 2016) and mainly negative emotions have a negative influence on brand attitude (Hwang & Mattila, 2019). We note that sentiment aspects have a greater predictive importance than emotional aspects. This means that a review containing a good balance of positive and negative words will portray a better brand attitude than the review which contains more words related to fear and anger. The results show that the use of emotional aspects in addition to sentimental aspects and parts of speech aspects will help to gain a better understanding of users' brand attitudes which was lacking in earlier studies (Bao, 2017; Zablocki et al., 2018).

Additionally, the study results also show that although the predictive model shows that inclusion of parts of speech aspects, represented by the number of nouns, verbs, adverbs and adjectives present in the reviews along with the interaction between parts of speech aspects and content polarity improve the model performance, these aspects have no significant impact on predicting customers' brand attitudes. Since, we have also represented the length of the review by using the parts of speech count, as proposed by Lee et al. (2017), the findings exhibit that the review length also has a non-significant influence on brand attitude. This is in contrast to what researchers (e.g. Chatterjee, 2020) have found in context of understanding review helpfulness. Singh et al. (2018) have also noted that variables like adjectives, nouns, verbs and set length do not have a strong effect on prospective customers. This means that positive/negative words, and words depicting fear or anger are considered more valuable than just the parts of speech aspects or length of the review. Thus, results of all the techniques show sentiment and emotional aspects as the important predictors of brand attitude and parts of speech aspects having a non-significant impact on brand attitude.

6.1. Theoretical contributions

This study has three main theoretical contributions. First, this study extends the conceptual paper of Zablocki et al. (2018). This is one of the

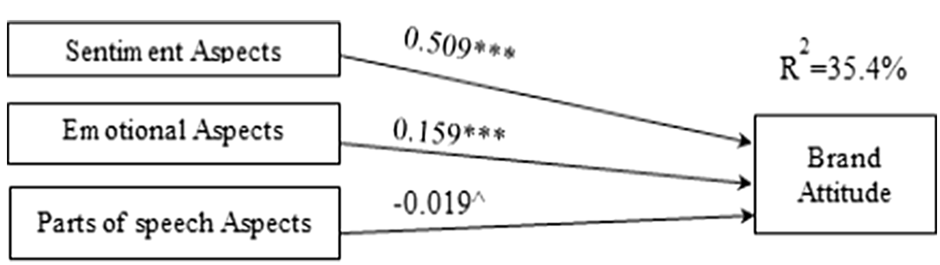
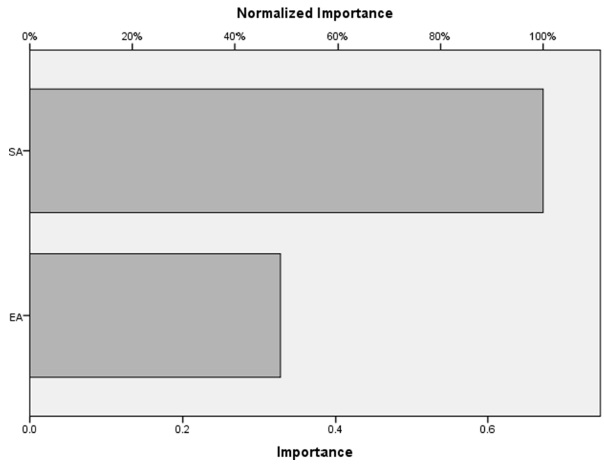


Fig. 2. Validated research model.

Table 8
Results of the ANN analysis on the textual aspects.

	Mean RMSE		Mean Standard Deviation	
	Training	Testing	Training	Testing
Sentiment and Emotional aspects only	0.4701	0.4669	0.0097	0.0063
Sentiment, Emotional and Parts of speech Aspects	0.4588	0.4597	0.0053	0.0123
Sentiment, Emotional and Parts of speech Aspects and the interaction effect	0.4140	0.4132	0.0154	0.0130



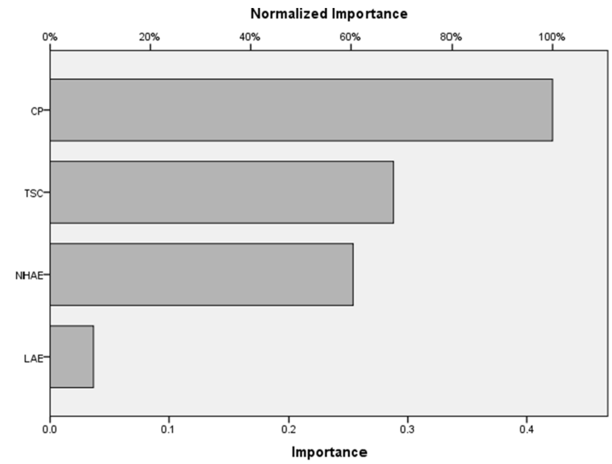


Fig. 3. Factor importance based on ANN results on (a) overall sentiment and emotional aspects; (b) the different significant sentiment and emotional aspects from UGC. [Note: CP = Content Polarity; EA = Emotional Aspects; NHAE = Negative High Arousal Emotions; LAE = Low Arousal Emotions; SA = Sentiment Aspects; TSC = Total Sentiment Content.]

few studies that have utilised the qualitative aspects of textual data to understand users' attitude towards brand and is the first study that captures the effect of various qualitative aspects of textual data on brand attitude in context of OTAs. This will help researchers to understand the customer perspectives regarding the brand from the online reviews. Since users' attitude towards brands is an important precursor of purchase intention, researchers can utilise this study to understand users' behavioural intention from textual content in various contexts (Zablocki et al., 2018).

Second, this study is unique in that it has utilised the qualitative aspects of textual content, namely, the sentiment, the emotional, and the parts of speech aspects, to understand customers' brand attitudes. This will help to make the model even more exhaustive and it contributes to the literature on brand attitudes using textual content. Earlier research works have mainly focused on the quantitative aspects of online UGC (Gensler et al., 2016; Zablocki et al., 2018). This study will also help researchers to utilize the qualitative aspects of textual data since textual aspects reveal a lot of information about customer perspectives (Chatterjee, 2019).

Third, the use of sentiment and emotional content to understand predictors of users' brand attitudes is still new (Chatterjee, 2019). Researchers have voiced their concerns about the lack of studies that have utilised emotional aspects of textual data (Siering et al., 2018). This study demonstrates the importance of emotional aspects in understanding customer perspectives better. Researchers have also used the parts of speech content of textual data to understand the predictive importance, which is not used quite often by any studies.

6.2. Implications for practice

This study has three main managerial implications. First, the use of both qualitative and quantitative aspects of online UGC will help managers understand the important aspects in different contexts. This will help them formulate better strategies. Additionally, we have found that brand attitude is an important predictor of purchase intention. Hence, understanding the users' underlying feelings as depicted by the online textual content can also help to understand their continuance intention. The findings suggest that the qualitative aspects have varying effects on understanding users' brand attitudes, such as (a) the total sentiment aspects and the content polarity are a significant driver of brand attitude. It means that, if the users' review contains good amount of positive and negative words, the user will have a positive brand attitude. Additionally, if the user has a good balance of positive and negative words as expressed in his reviews, they will have a positive brand attitude; (b) the negative high arousal emotions and the low arousal emotions both have a moderate negative impact on brand attitude. This means if the user reviews contain words that portray more anger, fear, sadness or anticipation, the user is most likely to have a negative brand attitude. This understanding will help service providers to solicit reviews properly and work for improving customers' brand attitudes. These findings will also help managers to have a quick easy understanding of the customer perspectives by just looking at the textual aspects rather than going through each and every review, which can be really time-taking.

Second, the results of this study show the importance of sentiment and emotional aspects of the textual data on exploring users' brand attitudes. The importance of the total number of sentiment-filled words in reviews along with the valence of those words can help managers predict whether the customer will be likely to continue the service or not. Additionally, this study has also revealed the negative effect of words related to anger, fear, sadness and anticipation on customers' brand attitudes. In this modern era, with hundreds of reviews getting posted every single minute about different product/services, extracting the textual aspects as discussed in the study, can help managers gain an easy overall understanding of the brand attitude of different customers. This study can help them to focus on customers whose reviews reflect a negative brand attitude. Developing appropriate strategies can help companies retain customers who are more likely to switch to a different provider. Understanding the underlying sentiments and emotions expressed by customers provides an additional advantage for organisations to cater to individual customers. Organisations can work on reducing the service gaps that exists by tracking the reviews depicting strong negative arousal emotions.

Third, since OTA services represent the hospitality segment, the qualitative aspects play a vital role in affecting customers' eWOM intentions as well. Results of this study show the sentiment and emotions expressed play a crucial role in impacting brand attitude. A good mix of positive and negative words in a review reflects a satisfied customer. Satisfied customers hence will spread positive eWOM. Augusto and Torres (2018) have stated how brand attitude and eWOM affect users' intention. However, reviews reflecting fear and anger are highly likely to predict a negative brand attitude and in turn a negative eWOM. Chatterjee (2019) earlier noted that online textual content is rich in information for both prospective customers as well as service providers. Hence service providers can aim at providing better services for getting positive reviews from customers to not only retain existing customers but also allure new prospects.

6.3. Limitations and future research directions

Like any other studies, this research is also not without limitations. First, this work has not looked into source credibility. Future researchers can utilise the study by Halliday (2016) to check the source credibility. Second, because online reviews are used to predict customer brand attitude, this study is prone to online review bias and social-proof (Bassig, 2020). It can be mitigated if the source credibility is checked (Halliday, 2016) and also by checking the use of words in the review. Additionally, text valence and social identity also helps to avoid this bias (Kusumasondjaja, Shanka, & Marchegiani, 2012). Third, this study has also not looked into the reviewers' mental state and cultural background. Future researchers can look into solving these aspects. Future researchers can also look into avenues to capture the demographics of reviewers, namely, location, gender, age, etc. The addition of these aspects will boost the generalisability of this work (Sharma, Al-Badi, Rana, & Al-Azizi, 2018). Fourth, this work is limited by the dataset used (Rana & Dwivedi, 2016). This study has utilised dataset of only one provider from OTA services. The impact of sentiment and emotional aspects on brand attitude can be different for different service-providers. Future research can focus on contextual effects and also explore how the drivers of customer attitude can vary for different providers in the same domain. Future research can also utilise the sentiment and emotional aspects in other contexts. Additionally, future studies can attempt to understand user mood from textual content.

7. Conclusion

The advancement of technological innovations and the penetration of Internet have witnessed a growth of not only various online services but also the popularity of various SM platforms. This has also increased the comments posted by users on SM platforms and merchandise websites about their experiences related to a product or service. These online reviews can not only be of strategic importance to service providers but also will be beneficial for prospective customers. The objective of this work is to highlight how the qualitative aspects of textual data can be useful to academics and managers for exploring the drivers of brand attitude from the vast amount of available user-generated content. In this study on OTAs, we have utilised 10,000 reviews from TripAdvisor (an OTA provider). We have utilised econometric technique (logistic regression), structural modeling technique (NLP-SEM) and machine-learning technique (ANN) to have an in-depth analysis of the different predictors of customers' perceived brand attitudes.

Results reveal that: (a) sentiment aspect is the most important predictor of brand attitude; (b) total sentiment content and content polarity have a significant positive association with brand attitude; (c) negative high arousal emotions and low arousal emotions have significant negative association with customers' brand attitudes; and (d) parts of speech aspect is not a significant predictor of brand attitude. This study contributes to the existing literature on understanding user brand attitude using user-generated content by (a) exploring the impact of different qualitative aspects of textual data, namely sentiment content, emotional content and parts of speech tags, (b) exploring the different emotional components that affect brand attitude namely high arousal emotions and low arousal emotions, and (c) exploring the impact of different parts of speech tags on examining brand attitude. This study can help scholars by providing an avenue for research on brand attitude using the qualitative aspects of user-generated content. This study will help service providers gain an easy overview of the different textual aspects to understand customers' brand attitudes and focus on customers whose comments reveal fear, anger, sadness or anticipation because these customers are more likely to switch to a different provider.

Acknowledgement

The infrastructural support provided by FORE School of Management, New Delhi in completing this paper is gratefully acknowledged.

References

Abreza, G., O'Reilly, N., & Reid, I. (2013). Relationship marketing and social media in sport. *International Journal of Sport Communication, 6*(2), 120–142. https://doi.org/10.1123/ijsc.6.2.120.

Ajorlou, A., Jadbabaie, A., & Kakhbod, A. (2016). Dynamic pricing in social networks: The word-of-mouth effect. *Management Science, 64*(2), 971–979. https://doi.org/10.1287/mnsc.2016.2657.

Akgül, Y. (2018). A SEM-Neural Network Approach for Predicting Antecedents of Factors Influencing Consumers' Intent to Install Mobile Applications. In F. J. Mtenzi, G. S. Oreku, D. M. Lupiana, & J. J. Yonazi (Eds.), *Mobile Technologies and Socio-Economic Development in Emerging Nations* (pp. 262–308). IGI Global.

Aladwani, A. M., & Dwivedi, Y. K. (2018). Towards a theory of SocioCitizenry: Quality anticipation, trust configuration, and approved adaptation of governmental social media. *International Journal of Information Management, 43*, 261–272. https://doi.org/10.1016/j.ijinfomgt.2018.08.009.

Alsmearat, K., Shehab, M., Al-Ayyoub, M., Al-Shalabi, R., & Kanaan, G. (2015). Emotion analysis of Arabic articles and its impact on identifying the author's gender. In *12th ACS/IEEE International Conference of Computer Systems and Applications AICCSA, Morocco*. https://doi.org/10.1109/aiccsa.2015.7507196.

Anandarajan, M., Hill, C., & Nolan, T. (2019). *Practical Text Analytics* (Vol. 2,, 193–220. https://doi.org/10.1007/978-3-319-95663-3_13.

Aswani, R., Kar, A. K., Ilavarasan, P. V., & Dwivedi, Y. K. (2018). Search engine marketing is not all gold: Insights from Twitter and SEOClerks. *International Journal of Information Management, 38*(1), 107–116. https://doi.org/10.1016/j.ijinfomgt.2017.07.005.

Augusto, M., & Torres, P. (2018). Effects of brand attitude and eWOM on consumers' willingness to pay in the banking industry: Mediating role of consumer-brand identification and brand equity. *Journal of Retailing and Consumer Services, 42*, 1–10. https://doi.org/10.1016/j.jretconser.2018.01.005.

Bafna, P. B., & Saini, J. R. (2021). On Readability Metrics of Goal Statements of Universities and Brand-Promoting Lexicons for Industries. In N. Sharma, A. Chakrabarti, V. E. Balas, & J. Martinovic (Eds.), *Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing* (pp. 63–72). Springer. https://doi.org/10.1007/978-981-15-5616-6.

Bailey, R., & Ball, S. (2006). An exploration of the meaning of hotel brand equity. *Service Industries Journal, 26*(1), 15–38. https://doi.org/10.1080/02642060500358761.

Ballantine, P. W., & Yeung, C. (2015). The effects of review valence in organic versus sponsored blog sites on perceived credibility, brand attitude, and behavioural intentions. *Marketing Intelligence & Planning, 33*(4), 508–521. https://doi.org/10.1108/MIP-03-2014-0044.

Bao, L. (2017). The Impact of User-generated Content Quality on Consumers' Brand Attitude. *Proceedings of the 8th International Conference on E-Education, E-Business, E-Management and E-Learning - IC4E '17, Kuala Lumpur, Malaysia*.

Bassig, M. (2020). Is Online Review Bias Real? Retrieved from https://www.reviewtrackers.com/blog/online-review-bias/. Accessed January 24, 2021.

Berger, J., Sorensen, A., & Rasmussen, S. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science, 29*(5), 815–827. https://doi.org/10.1287/mksc.1090.0557.

Bichler, M., & Kiss, C. (2004). A Comparison of Logistic Regression, k-Nearest Neighbor, and Decision Tree Induction for Campaign Management. In *AMCIS 2004 Proceedings, New York, USA* (pp. 1918–1925).

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84. https://doi.org/10.1145/2133806.2133826.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems, 50*(2), 511–521. https://doi.org/10.1016/j.dss.2010.11.009.

Cascio, C., O'Donnel, M., Bayer, J., Tinney, F. J., & Falk, E. B. (2015). Neural correlates of susceptibility to group opinions in online word-of-mouth recommendations.

Journal of Marketing Research, 52(4), 559–575. https://doi.org/10.1509/jmr.13.0611.

Cavanaugh, L. A., MacInnis, D. J., & Weiss, A. M. (2015). Perceptual dimensions differentiate emotions. *Cognition and Emotion, 30*(8), 1430–1445. https://doi.org/10.1080/02699931.2015.1070119.

Chatterjee, S. (2019). Explaining customer ratings and recommendations by combining qualitative and quantitative user generated contents. *Decision Support Systems, 119*, 14–22. https://doi.org/10.1016/j.dss.2019.02.008.

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management, 85*, Article 102356. https://doi.org/10.1016/j.ijhm.2019.102356.

Chatterjee, S., & Kar, A. K. (2020). Why do small and medium enterprises use social media marketing and what is the impact: Empirical insights from India. *International Journal of Information Management, 53*, Article 102103. https://doi.org/10.1016/j.ijinfomgt.2020.102103.

Chen, C.-F., & Kao, Y.-L. (2010). Relationships between process quality, outcome quality, satisfaction, and behavioural intentions for online travel agencies – evidence from Taiwan. *The Service Industries Journal, 30*(12), 2081–2092. https://doi.org/10.1080/02642060903191108.

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science, 54*(3), 477–491. https://doi.org/10.1287/mnsc.1070.0810.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345.

Chong, A. Y. L. (2013). Predicting m-commerce adoption determinants: A neural network approach. *Expert Systems with Applications, 40*(1), 523–530. https://doi.org/10.1016/j.eswa.2012.07.068.

Christodoulides, G., Michaelidou, N., & Argyriou, E. (2012). Crossnational differences in e-WOM influence. *European Journal of Marketing, 46*(11/12), 1689–1707. https://doi.org/10.1108/03090561211260040.

Crowley, A. E., & Hoyer, W. D. (1994). An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research, 20*(4), 561–574. https://doi.org/10.1086/209370.

Dahnil, M. I., Marzuki, K. M., Langgat, J., & Fabeil, N. F. (2014). Factors Influencing SMEs Adoption of Social Media Marketing. *Procedia - Social and Behavioral Sciences, 148*, 119–126. https://doi.org/10.1016/j.sbspro.2014.07.025.

Datafiniti (2019). Hotel Reviews: A list of 1,000 hotels and their online reviews. Retrieved from https://www.kaggle.com/datafiniti/hotel-reviews/. Accessed June 10, 2019.

Davtyan, D., Cunningham, I., & Tashchian, A. (2020). Effectiveness of brand placements in music videos on viewers' brand memory, brand attitude and behavioral intentions. *European Journal of Marketing, 55*(2), 420–443. https://doi.org/10.1108/EJM-08-2019-0670.

Syuzhet Document (2017). Introduction to the Syuzhet Package. Retrieved from https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html/. Accessed May 20, 2020.

Drummond, C., O'Toole, T., & McGrath, H. (2020). Digital engagement strategies and tactics in social media marketing. *European Journal of Marketing, 54*(6), 1247–1280. https://doi.org/10.1108/EJM-02-2019-0183.

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233–242. https://doi.org/10.1016/j.jretai.2008.04.005.

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., … Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168.

Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. *The Marketing Review, 15*(3), 289–309. https://doi.org/10.1362/146934715X14441363377999.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing, 25*(3), 215–224. https://doi.org/10.1016/j.ijresmar.2008.04.001.

Eusebius, S., Parackal, M., & Gnoth, J. (2016). Text mining the expression of "brand love" in user-generated social media content. In *Presentation in Marketing in a Post-Disciplinary Era ANZMAC 2016, Christchurch, New Zealand* (pp. 308–315).

Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research, 58*, 46–64. https://doi.org/10.1016/j.annals.2015.12.019.

Filieri, R., Acikgoz, F., Ndou, V., & Dwivedi, Y. K. (2020). Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention. *International Journal of Contemporary Hospitality Management*. https://doi.org/10.1108/IJCHM-05-2020-0402 (in Press).

Floyd, K., Freling, R., Alhoqail, S., Cho, H., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing, 90*(2), 217–232. https://doi.org/10.1016/j.jretai.2014.04.004.

Fuller, C. M., Biros, D. P., & Wilson, R. L. (2009). Decision support for determining veracity via linguistic-based cues. *Decision Support Systems, 46*(3), 695–703. https://doi.org/10.1016/j.dss.2008.11.001.

GeeksforGeeks (2020). Advantages and Disadvantages of Logistic Regression. Retrieved from https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/. Accessed January 20, 2021.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. *Tourism Management, 61*, 43–54. https://doi.org/10.1016/j.tourman.2016.12.022.

Gensler, S., Völckner, F., Egger, M., Fischbach, K., & Schoder, D. (2016). Listen to your customers: Insights into brand image using online consumer-generated product reviews. *International Journal of Electronic Commerce, 20*(1), 112–141. https://doi.org/10.1080/10864415.2016.1061792.

Glen, S. (2019). Alternatives to Logistic Regression. Retrieved from https://www.datasciencecentral.com/profiles/blogs/alternatives-to-logistic-regression/. Accessed January 20, 2021.

González-Mansilla, Ó., Berenguer-Contrí, G., & Serra-Cantallops, A. (2019). The impact of value co-creation on hotel brand equity and customer satisfaction. *Tourism Management, 75*, 51–65. https://doi.org/10.1016/j.tourman.2019.04.024.

Grimes, M. (2012). Nielsen: Global Consumers' Trust in 'Earned' Advertising Grows in Importance. Retrieved from https://www.nielsen.com/us/en/press-releases/2012/nielsen-global-consumers-trust-in-earned-advertising-grows/. Accessed January 20, 2021.

Grover, P., & Kar, A. K. (2020). User engagement for mobile payment service providers-introducing the social media engagement model. *Journal of Retailing and Consumer Services, 53*. https://doi.org/10.1016/j.jretconser.2018.12.002.

Gupta, P., & Harris, J. (2010). How e-WOM recommendations influence product consideration and quality of choice: A motivation to process information perspective. *Journal of Business Research, 63*(9), 1041–1049. https://doi.org/10.1016/j.jbusres.2009.01.015.

Gwang-In, B., Jung-Ae, K., & Gi-Jin, K. (2013). A Study on the Influence of Price Discount Policy in Brand Coffee Shops on Perceived Value, Brand Attitude, and Repurchase Intention. *Culinary Science and Hospitality Research, 13*(3), 274–290.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long Range Planning, 46*, 1–12.

Halliday, S. V. (2016). User-generated content about brands: Understanding its creators and consumers. *Journal of Business Research, 69*(1), 137–144. https://doi.org/10.1016/j.jbusres.2015.07.027.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques* (1st ed.). Elsevier.

Hawkins, K., & Vel, P. (2013). Attitudinal loyalty, behavioural loyalty and social media: An introspection. *The Marketing Review, 13*(2), 125–141. https://doi.org/10.1362/146934713X13699019904605.

Ho-Dac, N. N., Carson, S. J., & Moore, W. L. (2013). The effects of positive and negative online customer reviews: Do brand strength and category maturity matter? *Journal of Marketing, 77*(6), 37–53. https://doi.org/10.1509/jm.11.0011.

Hong, I. B., & Cha, H. S. (2013). The mediating role of consumer trust in an online merchant in predicting purchase intention. *International Journal of Information Management, 33*(6), 927–939. https://doi.org/10.1016/j.ijinfomgt.2013.08.007.

Hwang, Y., & Mattila, A. S. (2019). Feeling left out and losing control: The interactive effect of social exclusion and gender on brand attitude. *International Journal of Hospitality Management, 77*, 303–310. https://doi.org/10.1016/j.ijhm.2018.07.010.

IBM Watson (2020). Personality Insights. Retrieved https://cloud.ibm.com/docs/personality-insights/getting-started.html/. Accessed January 20, 2021.

Jacobson, J., Gruzd, A., & Hernández-García, Á. (2020). Social media marketing: Who is watching the watchers? *Journal of Retailing and Consumer Services, 53*. https://doi.org/10.1016/j.jretconser.2019.03.001.

Jiménez, F. R., & Mendoza, N. A. (2013). Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *Journal of Interactive Marketing, 27*(3), 226–235. https://doi.org/10.1016/j.intmar.2013.04.004.

Jockers, M.L. (2017). Syuzhet: Extract sentiment and plot arcs from text. Retrieved from https://github.com/mjockers/syuzhet/. Accessed January 20, 2021.

Kamboj, S., Sarmah, B., Gupta, S., & Dwivedi, Y. K. (2018). Examining branding co-creation in brand communities on social media: Applying the paradigm of Stimulus-Organism-Response. *International Journal of Information Management, 39*, 169–185. https://doi.org/10.1016/j.ijinfomgt.2017.12.001.

Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research – Moving away from the "What" towards the "Why". *International Journal of Information Management, 54*, Article 102205. https://doi.org/10.1016/j.ijinfomgt.2020.102205.

Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and Research: A Systematic Literature Review through Text Mining. *IEEE Access, 8*, 67698–67717. https://doi.org/10.1109/ACCESS.2020.2983656.

Kim, S., Koh, Y., Cha, J., & Lee, S. (2015). Effects of social media on firm value for US restaurant companies. *International Journal of Hospitality Management, 49*, 40–46. https://doi.org/10.1016/j.ijhm.2015.05.006.

Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2016). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing, 33*(1), 11–26. https://doi.org/10.1016/j.ijresmar.2014.12.004.

Kusumasondjaja, S., Shanka, T., & Marchegiani, C. (2012). Credibility of online reviews and initial trust. *Journal of Vacation Marketing, 18*(3), 185–195. https://doi.org/10.1177/1356766712449365.

Landman, P. (2020). OTA – Online Travel Agency. Retrieved from https://www.xotels.com/en/glossary/ota/. Accessed January 20, 2021.

Lee, K., Choi, J., Marakas, G. M., & Singh, S. N. (2018). Two Distinct Routes for Inducing Emotions in HCI design. *International Journal of Human-Computer Studies, 124*, 67–80. https://doi.org/10.1016/j.ijhcs.2018.11.012.

Lee, S.-G., Trimi, S., & Yang, C.-G. (2017). Perceived Usefulness Factors of Online Reviews: A Study of Amazon.com. *Journal of Computer Information Systems, 58*(4), 344–352. https://doi.org/10.1080/08874417.2016.1275954.

Liu, R.-D., Zhen, R., Ding, Y., Liu, Y., Wang, J., Jiang, R., & Xu, L. (2017). Teacher support and math engagement: Roles of academic self-efficacy and positive emotions. *Educational Psychology, 38*(1), 3–16. https://doi.org/10.1080/01443410.2017.1359238.

Mattke, J., Muller, L., & Maier, C. (2019). Paid, Owned and Earned Media: A Qualitative Comparative Analysis Revealing Attributes Influencing Consumers' Brand Attitude in Social Media. In *Proceedings of the 52nd Hawaii International Conference on System Sciences, Honolulu, United States*. https://doi.org/10.24251/HICSS.2019.098.

Mayzlin, D. (2006). Promotional chat on the internet. *Marketing Science, 25*(2), 155–163. https://doi.org/10.1287/mksc.1050.0137.

McClure, C., & Seock, Y.-K. (2020). The role of involvement: Investigating the effect of brand's social media pages on consumer purchase intention. *Journal of Retailing and Consumer Services, 53*, Article 101975. https://doi.org/10.1016/j.jretconser.2019.101975.

Mishra, A., & Satish, S. M. (2016). eWOM: Extant research review and future research avenues. *Journal for Decision Makers, 41*(3), 222–233. https://doi.org/10.1177/0256090916650952.

Mohammad, S. M., & Turney, P. D. (2011). Crowdsourcing a Word–Emotion Association Lexicon. National research council Canada. Retrieved from http://www.saifmohammad.com/WebDocs/Crowdsourcing-MohammadTurney-CI.pdf/. Accessed January 20, 2021.

Mowen, J. C., Park, S., & Zablah, A. (2007). Toward a theory of motivation and personality with application to word-of-mouth communications. *Journal of Business Research, 60*(6), 590–596. https://doi.org/10.1016/j.jbusres.2006.06.007.

Nisar, T. M., Hajli, N., Prabhakar, G., & Dwivedi, Y. (2019). Sharing economy and the lodging websites: Antecedents and mediators of accommodation purchase intentions. *Information Technology and People, 33*(3), 873–896. https://doi.org/10.1108/ITP-06-2018-0297.

Nisar, T. M., Prabhakar, G., & Patil, P. P. (2018). Sports clubs' use of social media to increase spectator interest. *International Journal of Information Management, 43*, 188–195. https://doi.org/10.1016/j.ijinfomgt.2018.08.003.

Nunkoo, R., Gursoy, D., & Dwivedi, Y. K. (2020). Effects of social media on residents' attitudes to tourism: Conceptual framework and research propositions. *Journal of Sustainable Tourism, 1–17*. https://doi.org/10.1080/09669582.2020.1845710.

Oheix, J. (2018). Detecting bad customer reviews with NLP. Retrieved https://towardsdatascience.com/detecting-bad-customer-reviews-with-nlp-d8b36134dc7e/. Accessed January 26, 2020.

Park, C. W., MacInnis, D. J., Priester, J., Eisingerich, A. B., & Iacobucci, D. (2010). Brand attachment and brand attitude strength: Conceptual and empirical differentiation of two critical brand equity drivers. *Journal of Marketing, 74*(6), 1–17. https://doi.org/10.1509/jmkg.74.6.1.

Prem (2020). Which is better – Random Forest vs Support Vector Machine vs Neural Network. Retrieved from https://www.iunera.com/kraken/fabric/random-forest-vs-support-vector-machine-vs-neural-network/. Accessed January 24, 2021.

Priester, J. R., Nayakankuppam, D., Fleming, M. A., & Godek, J. (2004). The A^2SC^2 model: The influence of attitudes and attitude strength on consideration and choice. *Journal of Consumer Research, 30*(4), 574–587. https://doi.org/10.1086/380290.

Rana, N. P., & Dwivedi, Y. K. (2016). Using clickers in a large business class: Examining use behavior and satisfaction. *Journal of Marketing Education, 38*(1), 47–64.

Rasa, K., Leonidas, S., & Gintautas, D. (2020). Kriging Predictor for Facial Emotion Recognition Using Numerical Proximities of Human Emotions. *Informatica, 31*(2), 249–275. https://doi.org/10.15388/20-INFOR419.

Rathore, A. K., & Ilavarasan, P. V. (2020). Pre-and post-launch emotions in new product development: Insights from twitter analytics of three products. *International Journal of Information Management, 50*, 111–127. https://doi.org/10.1016/j.ijinfomgt.2019.05.015.

Ray, A., & Bala, P. K. (2020a). Social media for improved process management in organizations during disasters. *Knowledge and Process Management, 27*(1), 63–74. https://doi.org/10.1002/kpm.1623.

Ray, A., & Bala, P. K. (2020b). Use of NLP and SEM in Determining Factors for E-Service Adoption. In Y. Akgül (Ed.), *Structural Equation Modeling Approaches to E-Service Adoption* (pp. 38–47). IGI Global. https://doi.org/10.4018/978-1-5225-8015-7.ch003.

Ray, A., & Bala, P. K. (2020c). User generated content for exploring factors affecting intention to use travel and food delivery services. *International Journal of Hospitality Management, 92*, Article 102730. https://doi.org/10.1016/j.ijhm.2020.102730.

Ray, A., Bala, P. K., Chakraborty, S., & Dasgupta, S. A. (2020). Exploring the impact of different factors on brand equity and intention to take up online courses from e-Learning platforms. *Journal of Retailing and Consumer Services*. https://doi.org/10.1016/j.jretconser.2020.102351 (in Press).

Ray, A., Bala, P. K., & Dwivedi, Y. K. (2020). Exploring values affecting e-Learning adoption from the user-generated-content: A consumption-value-theory perspective. *Journal of Strategic Marketing*. https://doi.org/10.1080/0965254X.2020.1749875 (in Press).

Ray, A., Bala, P. K., & Jain, R. (2020). Utilizing emotion scores for improving classifier performance for predicting consumers' intended ratings from social media posts. *Benchmarking: An International Journal*. https://doi.org/10.1108/BIJ-01-2020-0004 (in Press).

Reyes-Menendez, A., Saura, J. R., & Thomas, S. B. (2020). Exploring key indicators of social identity in the# MeToo era: Using discourse analysis in UGC. *International Journal of Information Management, 54*, Article 102129. https://doi.org/10.1016/j.ijinfomgt.2020.102129.

Roma, P., & Aloini, D. (2019). How does brand-related user-generated content differ across social media? Evidence reloaded. *Journal of Business Research, 96*, 322–339. https://doi.org/10.1016/j.jbusres.2018.11.055.

Rousidis, D., Koukaras, P., & Tjortjis, C. (2020). Social media prediction: A literature review. *Multimedia Tools and Applications, 79*, 6279–6311. https://doi.org/10.1007/s11042-019-08291-9.

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems, 81*, 30–40. https://doi.org/10.1016/j.dss.2015.10.006.

Sas, C., & Zhang, C. (2010). Investigating Emotions in Creative Design. In *DESIRE '10 Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design. Desire Network, Aarhus, Denmark* (pp. 138–149).

Saumya, S., & Singh, J. P. (2020). Spam review detection using LSTM autoencoder: An unsupervised approach. *Electronic Commerce Research.*. https://doi.org/10.1007/s10660-020-09413-4 (in Press).

Saumya, S., Singh, J. P., Baabdullah, A. M., Rana, N. P., & Dwivedi, Y. K. (2018). Ranking online consumer reviews. *Electronic Commerce Research and Applications, 29*, 78–89. https://doi.org/10.1016/j.elerap.2018.03.008.

Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour, 11*(3), 234–243. https://doi.org/10.1002/cb.1372.

Schwenk, C. H. (1986). Information, cognitive biases, and commitment to a course of action. *Academy of Management Reviews, 11*(2), 298–310. https://doi.org/10.5465/amr.1986.4283106.

Sentiment Document (2019). Package 'sentimentr'. Retrieved from https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf/. Accessed January 21, 2021.

Sharma, S. K., Al-Badi, A., Rana, N. P., & Al-Azizi, L. (2018). Mobile applications in government services (mG-App) from user's perspectives: A predictive modelling approach. *Government Information Quarterly, 35*(4), 557–568.

Shilbury, D., Westerbeek, H., Quick, S., Funk, D., & Karg, A. (2014). *Strategic Sport Marketing* (4th ed.). Routledge, Tayor & Francis Group.

Siering, M., Deokar, A. V., & Janze, C. (2018). Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews. *Decision Support Systems, 107*, 52–63. https://doi.org/10.1016/j.dss.2018.01.002.

Simonson, I., & Rosen, E. (2014). Absolute value: What really influences customers in the age of (nearly) perfect information. Retrieved from https://www.gsb.stanford.edu/faculty-research/books/absolute-value-what-really-influences-customers-age-nearly-perfect/. Accessed January 21, 2021.

Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Kumar, R. P. (2018). Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research, 70*, 346–355. https://doi.org/10.1016/j.jbusres.2016.08.008.

Sridhar, S., & Srinivasan, R. (2012). Social influence effects in online product ratings. *Journal of Marketing, 76*(5), 70–88. https://doi.org/10.1509/jm.10.0377.

Statistics Solutions (2010). Logistic Regression. Retrieved from https://www.statisticssolutions.com/wp-content/uploads/kalins-pdf/singles/what-is-logistic-regression.pdf/. Accessed March 3, 2020.

Tamilmani, K., Rana, N. P., Nunkoo, R., Raghavan, V., & Dwivedi, Y. K. (2020). Indian Travellers' Adoption of Airbnb Platform. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-020-10060-1 (in Press).

Tang, T., Fang, E., & Wang, F. (2014). Is neutral really neutral? The effects of neutral user-generated content on product sales. *Journal of Marketing, 78*(4), 41–58. https://doi.org/10.1509/jm.13.0301.

Tsang, N. K. F., Lai, M. T. H., & Law, R. (2010). Measuring EService Quality for Online Travel Agencies. *Journal of Travel & Tourism Marketing, 27*(3), 306–323. https://doi.org/10.1080/10548401003744743.

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Pennsylvania, USA* (pp. 417–424).

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124.

Valence shifters (2019). hash_valence_shifters: Valence Shifters. Retrieved from https://rdrr.io/cran/lexicon/man/hash_valence_shifters.html/. Accessed January 20, 2021.

van Riel, A. C. R., Semeijn, J., & Pauwels, P. F. J. (2004). Online Travel Service Quality: The Role of Pre-Transaction Services. *Total Quality Management, 15*(4), 475–493. https://doi.org/10.1080/1478336042000183578.

Vito, A. G., Schaefer, B., Higgins, G. E., Marcum, C., & Ricketts, M. (2019). Self-control, social learning theory, social bonds and binge drinking: Results from a national sample. *Journal of Substance Use, 24*(6), 655–659. https://doi.org/10.1080/14659891.2019.1642406.

Vyas, V., & Uma, V. (2019). Approaches to Sentiment Analysis on Product Reviews. In D. S. Rajput, R. S. Thakur, & S. M. Basha (Eds.), *Sentiment Analysis and Knowledge Discovery in Contemporary Business* (pp. 15–30). IGI Global.

Wang, X.-W., Cao, Y.-M., & Park, C. (2019). The relationships among community experience, community commitment, brand attitude, and purchase intention in social media. *International Journal of Information Management, 49*, 475–488. https://doi.org/10.1016/j.ijinfomgt.2019.07.018.

Weinberg, T. (2009). *The new community rules: Marketing on the social web* (1st ed.). O'Reilly Media, Inc.

Yan, X., Radwan, E., & Abdel-Aty, M. (2005). Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. *Accident Analysis & Prevention, 37*(6), 983–995. https://doi.org/10.1016/j.aap.2005.05.001.

Yang, C., Lin, K. H.-Y., & Chen, H. H. (2009). Writer Meets Reader: Emotion Analysis of Social Media from Both the Writer's and Reader's Perspectives. In *2009 IEEE/WIC/*

ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. Milan, Italy (pp. 287–290). https://doi.org/10.1109/WI-IAT.2009.50.

Zablocki, A., Schlegelmilch, B., & Houston, M. J. (2018). How valence, volume and variance of online reviews influence brand attitudes. *AMS Review, 9*, 61–77. https://doi.org/10.1007/s13162-018-0123-1.

ZionMarketResearch (2018). Global Online Travel Booking Market Will Reach USD 1,955 Billion By 2026. Retrieved from https://www.globenewswire.com/news-release/2018/12/26/1678253/0/en/Global-Online-Travel-Booking-Market-Will-Reach-USD-1-955-Billion-By-2026-Zion-Market-Research.html/. Accessed December 31, 2020.

Arghya Ray is an Assistant Professor at FORE School of Management, Delhi. He had earlier worked as an Assistant Professor in ADAMAS University, and as a Visiting Faculty at Birla Global University. He has done his PhD from IIM Ranchi in the area of Information Systems and Business Analytics and received his B.Tech. in Computer Science and Engineering from SRM University. He has also worked at Tata Consultancy Services as a System Engineer for two years. His areas of interest include Text Mining & NLP, Recommender Systems, Data mining applications and algorithms, Business Analytics, Marketing Analytics and Computer Networks.

Pradip Kumar Bala obtained his B.Tech, M.Tech and Ph.D., all three from Indian Institute of Technology Kharagpur. He worked in Tata Steel and taught in IIT, Roorkee, Xavier Institute of Management Bhubaneswar (XIMB) before joining IIM Ranchi where he is currently working as professor in the area of Information Systems & Business Analytics. His areas of expertise include text mining and NLP, recommender systems, data mining applications and algorithms. He has published more than 90 research papers in reputed journals, conference proceedings and as book chapters of reputed publishers. He is also a member of International Association of Engineers (IAENG).

Nripendra P. Rana is a Professor in Digital Marketing and the Head of International Business, Marketing and Branding at the School of Management at University of Bradford, UK. His current research interests focus primarily on adoption and diffusion of emerging ICTs, e-commerce, m-commerce, e-government and digital and social media marketing. He has published more than 200 papers in a range of leading academic journals, conference proceedings, books, etc. He has co-edited six books on digital and social-media marketing, emerging markets, adoption and diffusion of information technology and systems and supply and operations management. He has also co-edited special issues, organised tracks, mini-tracks and panels in leading conferences. He is a Chief-Editor of International Journal of Electronic Government Research and an Associate Editor of International Journal of Information Management. He also holds a Senior Fellow of the Higher Education Academy (SFHEA). He is a Visiting Professor at IIM Jammu in India.