Journal of Destination Marketing & Management 20 (2021) 100570

Contents lists available at ScienceDirect

Journal of Destination Marketing & Management

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

What drives the helpfulness of online reviews? A deep learning study of sentiment analysis, pictorial content and reviewer expertise for mature destinations

Enrique Bigne[a,*], Carla Ruiz[a], Antonio Cuenca[a], Carmen Perez[a], Aitor Garcia[b]

[a] University of Valencia, Faculty of Economics - Department of Marketing, Av. Naranjos S/n, 46022, Valencia, Spain
[b] Vicomtech, Parque Científico y Tecnológico de Gipuzkoa, Paseo Mikeletegi 57, 20009, Donostia / San Sebastián, Spain

A R T I C L E I N F O

Keywords:
Perceived helpfulness
Dual-processing theory
User-generated content
Sentiment analysis
Deep learning
Mature destinations

A B S T R A C T

Tourist destinations are increasingly affected by travel-related information shared through social media. Drawing on dual-process theories on how individuals process information, this study examines the role of central and peripheral information processing routes in the formation of consumers' perceptions of the helpfulness of online reviews of mature destinations. We carried out a two-step process to address the perceived helpfulness of user-generated content, a sentiment analysis using advanced machine-learning techniques (deep learning), and a regression analysis. The database was 2023 comments posted on TripAdvisor about two iconic Venetian cultural attractions, St. Mark's Square (an open, free attraction) and the Doge's Palace (which charges an entry fee). Using deep-learning techniques, with logistic regression, we first identified which factors influenced whether a review received a "helpful" vote. Second, we selected those reviews which received at least one helpful vote to identify, through linear regression, the significant determinants of TripAdvisor users' voting behaviour. The results showed that reviewer expertise is influential in both free and paid-for attractions, although the impact of central cues (sentiment polarity, subjectivity, pictorial content) differs for both attractions. Our study suggests that managers should look beyond individual ratings and focus on the sentiment analysis of online reviews, which are shown to be based on the nature of the attraction (free vs. paid-for).

1. Introduction

In the analysis of big data generated by tourists through online reviews posted on tourism platforms (for a review, see Li, Xu, Tang, Wang, & Li, 2018), a key question is: how helpful are these reviews for other tourists in their decision-making? The helpfulness of online comments for consumer decision-making adds a valuable new feature to the four big data Vs (volume, velocity, variety, validity), as recently highlighted in the destination management organization (DMO) literature (Hlee, Lee, Koo, & Chung, 2020).

Helpful user-generated content (UGC) posted on social media platforms plays a key role in reducing information overload and driving brand choice for mature tourism destinations (Bigné, Ruiz, & Currás, 2019; González-Rodríguez, Martínez-Torres, & Toral, 2016), so a better understanding of this customer-driven influence is required (Jiménez-Barreto & Campo-Martínez, 2018; Ratchford, 2020). For example, St. Mark's Square and the Doge's Palace in Venice are famous attractions, but potential travellers may doubt whether their reputations make them worth visiting. There are more than 3000 reviews about St. Mark's Square on TripAdvisor. Therefore, the strategic question for mature destinations is what constitutes a helpful review?

Sentiment polarity analysis is based on techniques that assign polarities, positive, neutral, and negative, to opinions, and obtains valuable information from them through text analysis (Ma, Cheng, & Hsiao, 2018). Sentiments expressed about tourist attractions in reviews provides tacit context-specific explanations of the reviewer's feelings, experiences and emotions, which go beyond numeric ratings. Recent studies (Cao, Duan, & Gan, 2011; Liu & Park, 2015) have indicated that the qualitative aspects of reviews are more influential determinants of travel-review helpfulness than ratings. This research focuses on content (text/pictures); we determine the helpfulness of online review content through deep-learning (DL) data analysis methods applied to a mature

* Corresponding author.
E-mail addresses: enrique.bigne@uv.es (E. Bigne), carla.ruiz@uv.es (C. Ruiz), antonio.cuenca@uv.es (A. Cuenca), carmen.perez-cabanero@uv.es (C. Perez), agarciap@vicomtech.org (A. Garcia).

https://doi.org/10.1016/j.jdmm.2021.100570
Received 10 November 2019; Received in revised form 1 December 2020; Accepted 22 February 2021
Available online 4 March 2021

tourism destination. ML techniques are becoming popular in tourism industry online data mining (Bigné, William, & Soria, 2020). DL describes ML approaches that stack computational layers of depth (as opposed to more "plain", single-layer, ML approaches). DL has recently emerged as a powerful technique for analysing digital content based on natural language processing (NLP), becoming a promising area for tourism research (Li et al., 2018). A crucial element of any tourism destination – its attractions – has been very little studied using ML. Attractions are the core products of tourism destinations and require deeper involvement and higher investments in time and in money than do amenities (Simeon, Buonincontri, Cinquegrani, & Martone, 2017). The present study analyses the perceived helpfulness of UGC about tourism attractions in a two-step process, a sentiment polarity analysis using deep learning, and a regression analysis that identifies the drivers of perceived helpfulness.

Recent research on the influence of UGC on decision-making has demonstrated the significant influence of online reviews on hotel and restaurant bookings (Bigné, Ruiz, & Currás-Pérez, 2019), but paid little attention to tourist attractions (Fang, Ye, Kucukusta, & Law, 2016; Hlee et al., 2020). Consumer electronic word of mouth (eWOM) adoption in the selection of tourist attractions differs between hotels and restaurants, for two main reasons. First, some attractions are free, whereas others require an entry fee. A research question arises: does the perceived helpfulness of online reviews vary based on whether an attraction is free or paid-for? Second, when comparing attractions to other tourism services, consumers may face smaller loss if they choose a bad restaurant or hotel (meal/accommodation cost) compared to the choice of a disappointing attraction (entire travel plan). Hence, exploring the effect of online reviews on tourists' decision-making processes as regards attractions is important and, similarly, assessing perceived helpfulness is critical in tourists' choices of mature destinations.

Dual-process theories (Eagly & Chaiken, 1993) posit that consumers process information through two routes, central (i.e., an analysis of all relevant pieces of information) and peripheral (i.e., decision-making by assessing whatever information is available). Previous studies have assessed the role of these two routes in consumers' willingness to purchase. However, few studies have explored the effect of these processing routes on the perceived helpfulness of online reviews. To address this gap, this study uses dual processing models to classify the factors that influence the helpfulness of online reviews of tourism attractions into peripheral cues for heuristic information processing, and central cues for systematic information processing. This study has three goals: (i) to assess whether sentiment polarity, the subjectivity of the review, pictorial content and the reviewer expertise have an impact on the helpfulness of UGC about mature destinations; (ii) to identify which review components have the highest impact on the helpful UGC about mature destinations; (iii) to analyse whether the perceived helpfulness of online reviews varies based on price (free or paid-for).

This research adds to the sentiment analysis research literature by combining ML techniques applied to NLP, and DL, to study UGC about tourist attractions. The study provides insights into UGC sentiment and analyses its polarity on the basis of tourist online comments about specific features of tourist attractions. The study applies NLP to extract semantic characteristics from the reviews, and linear regression models to examine which semantic characteristics determine: (i) if a review is perceived as helpful or not (i.e., helpful reviews have received at least one "helpful" vote on the TripAdvisor page); (ii) the number of votes helpful reviews receive (i.e., the number of "helpful" votes for the review shown in the TripAdvisor page). The study is organised as follows. In section 2 we review the recent NLP and sentiment analysis literature. Then, the impact of heuristic information cues, sentiment polarity and the objectivity/subjectivity of UGC on the helpfulness of online reviews is addressed. Section 3 explains the methodology and section 4 sets out the results of our analyses. The last section discusses conclusions, implications and future research lines.

2. Literature review and hypotheses development

2.1. Sentiment analysis and natural language processing

Cambria and White (2014) proposed that NLP employs computational models to process natural language through learning cognitive activities of the human brain. NLP tasks include information extraction, information retrieval, text summarisation, question answering, topic modelling and, more recently, sentiment polarity analysis (Chang, Ku, & Chen, 2020).

Sun, Luo, and Chen (2017) reviewed the general NLP techniques required for text pre-processing, deep-learning approaches to opinion mining and the relations between information fusion and opinion mining. More recently, advanced ML methods based on DL algorithms have shown high accuracy in determining the polarity of online reviews (Chang et al., 2020). One of the main advantages of applying DL to NLP is that it is independent of expert knowledge and lexical resources (Rojas-Barahona, 2016). Thelwall, Buckley, Paltoglou, Cai, and Kappas (2010) developed an algorithm (Sentitrengh) for MySpace comments that measured word sentiment polarity strength using human-annotated scores, combined with the sentiment polarity in the reviewers' comment history. Aldayel and Azmi (2016) used both lexical-based and support vector machine (SVM) classifiers to perform hybrid sentiment polarity analysis, obtaining over 80% accuracy. Amplayo and Song (2017), showed that online reviews could be effectively summarised by using a fine-grained sentiment polarity extraction model for short texts that combines NLP and ML techniques, taking a three-level classification approach (first-level: NLP techniques; second-level: support vector machine (SVM); third level: combination of the other levels using a feed-forward neural network).

Recent studies have shown that good results can be achieved in tourism research by using a combination of advanced ML models and NLP in sentiment analysis. Marrese-Taylor, Velásquez, and Bravo-Marquez (2014) applied an extension of traditional NLP-based rules for both subjective and sentiment polarity classification of tourism product reviews. Kim and Park (2017) applied a DL approach based on Stanford sentiment analysis to analyse the sentiment polarity of reviews of Paris; they noted that tasks are performed better when suitable algorithms are applied. Al-Smadi, Qawasmeh, Al-Ayyoub, Jararweh, and Gupta (2018) showed that deep recurrent neural network approaches (RNN) outperform other ML methods (SVMs) in polarity identification. Chang et al. (2020) demonstrated that combining DL and NLP tools can assist hotels in decision-making by prioritising which reviews need responses.

2.2. User-generated content and mature destinations

According to Buhalis (2000, p.97) "tourism destinations are amalgams of tourism products, offering an integrated experience to consumers". From the demand-oriented viewpoint, tourism destinations are in a state of continuous change. The concept "destination life cycle" (Butler, 1980) has been used to explain the successive stages destinations go through over time: (i) Exploration (a small number of adventurous visitors); (ii) Involvement (increased visitor numbers and improved facilities); (iii) Development (the destination becomes well-known/popular); (iv) Consolidation (tourist arrivals slow down); (v) Stagnation (tourism peak, repeat visitors); (vi) Decline (destination no longer fashionable, mainly visited by excursionists).

Mature destinations enjoy a consistent flow of guests, adjusted for seasonal variations; however, they may be unable to differentiate themselves from similar vacation options (Kozak & Martin, 2012). In this context, tourist preferences are drivers that influence destination trends. Butler (1980) argued that, in maturity, the original cultural atmosphere, that made the destination unique, diminishes as non-local actors (e.g. excursionists) come to dominate the consumption of goods and services. In Venice, as in other heritage mature destinations, strong tourism pressure has trapped the city in a vicious circle that is eroding the quality of its tourism attractions (Ganzaroli, De Noni, & Van Baalen). UGC may help destinations arrest this trend by providing tourists with aggregated, up-to-date information on the quality of their tourism attractions. This is particularly important for mature tourism destinations, which seek to avoid the stagnation phase by taking appropriate management decisions (Bernini & Cagnone, 2014).

In an attempt to address the increasing competition and market saturation of tourism destinations, modern destination management approaches suggest that DMOs use tourist information and business intelligence to improve the positioning, and attraction, of mature destinations (Van der Zee, Bertocchi, & Vanneste, 2020). Marine-Roig and Clavé (2015) showed how tourism managers in Barcelona use UGC to better position the city as a tourist destination and enhance its business intelligence. Fuchs, Höpken, and Lexhagen's (2014) model can help mature destinations use big data mining to learn more about real-time tourist behaviour. Phillips, Zigan, Silva, and Schegg (2015) analysed 59,688 online reviews to improve Swiss hotel performance. Recent research (van der Zee et al., 2018; Ganzaroli, De Noni, & Van Baalen, 2017) suggests that UGC about mature heritage cities may have a clustered review pattern, with many frequently reviewed core attractions creating a vicious circle. Thus, helpful UGC is a valuable information source for tourists. The next sections discuss how consumers process UGC and the impact of review cues and reviewer characteristics on the perceived helpfulness of UGC about mature destinations.

2.3. Information processing of UGC and helpfulness of online reviews

Two of the most prominent models in the dual-process literature are the elaboration likelihood model-ELM (Petty & Cacioppo, 1981) and the heuristic–systematic model-HSM (Chaiken, 1980). Both have been applied to study online reviews. The HSM proposes that, in systematic processing, message recipients carefully examine all information items to assess their importance and relevance to tasks before making a final decision. In heuristic processing, message recipients use just a few informational cues, such as simple decision rules, to reach a conclusion. The ELM distinguishes between the central route and the peripheral route. The central route involves the message recipient in a high level of cognitive effort to process the argumentation; the peripheral route persuades through associations the individual makes with positive or negative non-content cues in the stimulus (Petty & Cacioppo, 1981).

Dual-processing theories propose that consumers faced with large amounts of complex information, and who are limited by their cognitive abilities to process it in a limited timeframe, often reduce their cognitive effort by resorting to simplifying strategies and heuristics to arrive at a decision (Chaiken, 1980). Information that requires less effort to process, and is easily assimilated, such as details about a reviewer's expertise, may be used to simplify the consideration set. Thereafter, more strenuous mental strategies, such as analysing the sentiment polarity expressed in the reviews, or the subjectivity of the review, may be used to arrive at a final decision (selecting the tourism attraction). The extant literature suggests that information embedded in the text of online reviews is centrally processed, whereas information associated with text attributes is peripherally processed (Cheng & Ho, 2015; Hong, Xu, Wang, & Fan, 2017; Tsai, Chen, Hu, & Chen, 2020). Based on dual-processing models, this study defines the reviewer's expertise as a peripheral cue, and factors related to review content (subjectivity, pictures, sentiment polarity) as central cues for evaluating online reviews. Dual-processing models suggest that both routes might influence the perceived helpfulness of online reviews.

2.3.1. Peripheral information cues of user-generated content

The peripheral cues in online reviews are non-content cues used in heuristic information processing. Consumers use reviewer characteristics as peripheral information cues when assessing the helpfulness of online reviews (Cheng & Ho, 2015; Hong et al., 2017; Tsai et al., 2020). Some researchers have argued that online information posted by experienced travellers is regarded as more credible than information posted by novice users (e.g. Fang et al., 2016). The credibility given to a source reflects the extent to which (s)he has previously made valuable contributions to the online travel community. In addition, the administrators of virtual communities rank their members based on their contributions. In online communities, a user's activity level has a positive effect on the credibility of his/her reviews (Baek, Ahn, & Choi, 2012; Tsai et al., 2020). TripAdvisor awards badges to highly ranked reviewers so that consumers will be aware of their reputations. In this study we use reviewers' expert badges and experience on TripAdvisor as a proxy for reviewer expertise.

The ELM proposes that consumers are more likely to use peripheral cues to judge a reviewer's credibility when assessing the overall helpfulness of a review (Cheng & Ho, 2015), such as reviewer expertise (Park & Nicolau, 2015). A recent meta-analysis by Hong et al. (2017) showed that reviewer expertise positively influences the perceived helpfulness of reviews. Thus, we hypothesise that:

H1a. Reviewer expertise positively influences the perceived helpfulness of UGC about mature destinations.

H1b. Reviewer expertise positively influences the number of helpful votes given to UGC about mature destinations.

2.3.2. Central cues of user-generated content

Previous research has suggested that content cues, such as the proportion of positive/negative words in a review, pictorial content and emotional tone (subjective/objective) are used as central cues in online reviews (Baek et al., 2012; Cheng & Ho, 2015).

2.3.2.1. Pictorial content. Cultural attractions are regarded as experience goods. Consumers prefer to view images of attractions in reviews because they provide additional information. Following Cheng and Ho (2015), this study uses the number of images in an online review as a central cue for evaluating its helpfulness. Pictorial content enhances the diagnosticity of online reviews as it communicates emotion, increases vividness and increases the user's empathy with the reviewer (Manganari & Dimara, 2017). Pictorial content increases vividness because it is more salient than text and, thus, attracts more attention. Therefore:

H2a. Pictorial content embedded in UGC positively influences the perceived helpfulness of UGC about mature destinations.

H2b. Pictorial content embedded in UGC positively influences the number of helpful votes given to UGC about mature destinations.

2.3.2.2. Sentiment polarity. The influence of sentiment polarity (positive/negative versus neutral) on perceived helpfulness is debated. Some studies have suggested that favourable/unfavourable reviews have higher perceived helpfulness than neutral reviews. Cao et al. (2011) demonstrated that reviews for software programs with extreme opinions are considered more helpful than those with neutral opinions. Extreme reviews are considered more diagnostic and less ambiguous than mixed reviews and, thus, more helpful for making judgments (Bhandari & Rodgers, 2018). In the travel and tourism industry, scholars have found that consumers vote restaurant reviews with extreme ratings more helpful than reviews with moderate ratings (Park & Nicolau, 2015); however, other studies have shown that travellers regard some extreme reviews as untrustworthy, and thus unhelpful (Filieri, 2016).

Mudambi and Schuff (2010) argued that reviews with higher word counts are perceived as more useful because readers treat word count as representing greater depth of information comprehensiveness and usefulness. Baek et al. (2012) showed that the percentage of negative words in a review impacts positively on its perceived helpfulness. In consequence, we conclude that consumers centrally process negative and positive reviews differently and, thus, the perceived helpfulness of reviews varies. The study posits that reviews containing many positive (favourable) or negative (unfavourable) words are perceived as being more helpful than those with neutral opinions. This stems from research on information processing and uncertainty reduction theory, which posits that individuals look for uncertainty-reducing information (Forman, Ghose, & Wiesenfeld, 2008). Favourable and unfavourable reviews help consumers reduce uncertainty by confirming or eliminating purchase options. Therefore:

H3a. The number of negative words in a review positively influences the perceived helpfulness of UGC about mature destinations.

H3b. The number of negative words in a review positively influences the number of helpful votes given to UGC about mature destinations.

H4a. The number of positive words in a review positively influences the perceived helpfulness of UGC about mature destinations.

H4b. The number of positive words in a review positively influences the number of helpful votes given to UGC about mature destinations.

Online reviews are informational cues that facilitate customers' evaluations of tourism destinations. Persuasion theory proposes that consumers find specific arguments more persuasive than vague arguments (Sparks, Perkins, & Buckley, 2013). Bigné et al. (2019) showed that the level of detail in a message plays a powerful role in the persuasion process. That is, highly detailed reviews alleviate the customer's uncertainty about tourism attractions and provide confidence in the decision-making process. We extend this reasoning to posit that UGC containing extreme information (positive or negative) about the specific features of a tourism attraction, because it is more persuasive and diagnostic, will be perceived as more helpful than neutral reviews.

H5a. The number of positive words in a review about a specific feature of a tourism service positively influences the perceived helpfulness of UGC about mature destinations.

H5b. The number of positive words in a review about a specific feature of a tourism service positively influences the number of helpful votes given to UGC about mature destinations.

H6a. The number of negative words in a review about a specific feature of a tourism service positively influences the perceived helpfulness of UGC about mature destinations.

H6b. The number of negative words in a review about a specific feature of a tourism service positively influences the number of helpful votes given to UGC about mature destinations.

2.3.2.3. Subjectivity. Sentiment analysis is based on the premise that information provided in a review is either subjective (i.e., opinionated) or objective (i.e., factual). Reviewers give "subjective opinions" based on personal opinions, feelings, beliefs and judgments about products/services and destinations, and "objective statements" based on facts, evidence and measurable observations, or a mixture of both (Feldman, 2013). Therefore, objective reviews tend to reflect cognitive behaviour while subjective reviews reflect affective behaviour.

The positive effect of subjectivity on the perceived helpfulness of reviews is grounded in dual-processing models and schema theory. If a tourism destination is emergent, or its attractions are not well-known, tourists will first build their schema with objective comments. This is unnecessary with mature destinations with well-known attractions. Subjective reviews such as "worst museum I ever visited", put less cognitive load on consumers' information processing capabilities than objective comments, which raises the likelihood of the review being perceived as helpful (He et al., 2017https://onlinelibrary.wiley.com/doi/full/10.1002/mar.20934). Consumers who post subjective reviews are more emotional and tend to generate more extreme and, thus, diagnostic and helpful evaluations (Zhao, Xu, & Wang, 2019). Ghose and Ipeirotis (2011) argued that reviews with more subjective words are more likely to be helpful. Forman et al. (2008) found that the subjectivity of reviews impacted positively on the perceived helpfulness of reviews of electronic goods. Therefore, we posit:

H7a. Subjective reviews about mature destinations are perceived as more helpful than objective reviews.

H7b. Subjective reviews about mature destinations will receive a higher number of helpful votes than objective reviews.

3. Method

3.1. Sample description

Venice is among the 50 top worldwide ranked cities, with 5.5 million visitors annually (Euromonitor International, 2019), and is a UNESCO World Heritage Site. St Marks's Square and the Doge's Palace were selected based on: (i) their overall ratings of 4.5/5 based on UGC comments, and their TripAdvisor certificate of excellence awards; (ii) both are cultural attractions; (iii) consumers face similar information overload when processing their online comments; (iv) they are very close; (v) one is free to visit (St. Mark's Square), and the other charges 20 euros (Doge's Palace). St. Mark's Square is an open, free attraction and is visited by almost all tourists who go to Venice. The Doge's Palace is the most visited Venetian museum. To ensure review homogeneity, 3469 original English-language comments about the attractions were selected. We analysed the full texts of the reviews, the individual ratings, the headings, the pictures, the number of "helpful" votes, reviewer expertise (years on TripAdvisor) and their number of badges. From the initial database we selected all reviews (2,023) of St. Mark's Square (1,206) and the Doge's Palace (817). Removing missing values gave us a final sample of 1199 comments about St. Mark's Square, and 812 for the Doge's Palace. Of the sample of some 88.7% comments about St. Mark's Square contained positive words, and 19.5% positive words about related activities; 41.2% comments included negative words. Only 13% of the reviews included picture(s). Some 30.3% of the comments about the Doge's Palace included negative words, and 16.7% positive words. Some 56.3% of the comments were classified as objective, and 43.7% as subjective.

The reviewers' profiles are shown in Table 1. Similar numbers of males (49.6%) and females (50.4%) posted comments on St. Mark's Square, while more females (59.44%) than males (40.6%) posted about the Doge's Palace. The age distribution is similar. They had been using TripAdvisor for a similar number of years (around 8, on average), and had been awarded a similar number of badges (around 38, on average). The means of the number of reviews posted and the number of cities the reviewers had visited are also similar, both having high standard

Table 1
Reviewers' profiles.

St. Mark's Square				
Demographics	%	Reviewer experience	Mean	S.D
Male	49.6%	Years on TripAdvisor	8.20	2.74
Female	50.4%	Badges	38.48	35.71
<34	25.0%	Helpful votes obtained	2.07	2.02
35–49	30.7%	Number of reviews	138.87	237.58
50–64	32.8%	Number of cities visited	136	175.90
>65	11.5%			
Doge's Palace				
Demographics	%	Reviewer experience	Mean	S.D.
Male	40.6%	Years on TripAdvisor	8.30	2.87
Female	59.44%	Badges	37.49	38.18
<34	22.7%	Helpful votes obtained	3.32	6.08
35–49	32.5%	Number of reviews	133.71	258.26
50–64	35.1%	Number of cities visited	112.94	140.94
>65	9.7%			

Source: Own design

deviations.

3.2. Method

To test the proposed hypotheses two main analyses were carried out. First, we applied logistic regression to identify the factors that cause TripAdvisor reviews to receive helpful votes. Logistic regression transforms the number of helpfulness votes reviews receive into nominal data, such as "unhelpful" or "helpful" (dichotomous variable), based on whether the number of votes exceeds a benchmark cut-off value (0). Logistic regression has previously been used in sentiment analysis research (Cao et al., 2011). The independent variables relate to the reviewers' profiles (number of years on TripAdvisor, number of badges), pictorial content of the reviews (number of pictures posted) and sentiment polarity-based variables (number of positive/negative words, number of positive/negative words related to attraction-linked activities, number of positive/negative words about environmental conditions, number of positive/negative words about price, number of positive/negative words about recommending the attraction, and its classification as an objective or subjective review).

We next analysed the factors that explain the perceived helpfulness of online reviews by measuring the number of "helpful" votes they received on TripAdvisor. Many studies have adopted linear regression to predict review helpfulness scores (Forman et al., 2008; Kim & Park, 2017; Mudambi & Schuff, 2010). From the original database we selected the comments about St. Mark's Square (177) and the Doge's Palace (273) which received at least one helpfulness vote. We then applied a linear regression analysis using the stepwise method. After deleting the outliers identified by the case wise diagnostics, the final sample was 172 reviews for St. Mark's Square and 268 for the Doge's Palace. Of the 172 comments about St. Mark's Square, 60.5% received only one helpful vote, and 19.8% contained at least one picture. Positive words featured in 92.4% of the comments, 15.1% contained positive words about the environment and 7.6% contained negative words about activities. Of the 268 comments about the Doge's Palace, 55.2% received only one helpful vote, and only 12.7% contained at least one photo. Positive words featured in 88.4% of the comments, and 1.9% included positive words, in particular, about price. Only 8.2% contained negative recommendations, and 0.4% included negative words about price. Of the comments, 48.9% were classified as objective, and 51.1% as subjective.

3.3. Sentiment analysis using machine learning and deep learning

We carried out an automatic sentiment analysis using deep learning applied to NLP. Deep learning is an ML technique that exploits multiple layers of non-linear information processing and feature extraction and transformation (Do, Prasad, Maag, & Alsadoon, 2018).

In this study we used software specifically developed for the hospitality industry by Vicomtech[1]. This is based on free open-source tools available in OpeNER (https://www.opener-project.eu/project/), a NLP platform (García-Pablos, Cuadros, & Linaza, 2016), and DL open-source tools. The sentiment polarity analysis process involves several steps. First, the data (text) were extracted and cleaned. The languages used were automatically detected and unwanted languages were filtered out to avoid noise. The texts were then segmented into sentences and tokens which, broadly speaking, are words and punctuation marks. Following this, part-of-speech tagging (PoS tagging) and lemmatisation processes were carried out. The PoS-tagging consists of determining the corresponding morphosyntactic category for a word, given its context (e.g., in the sentence *"the town is beautiful"*, "*town*" is a noun, "*beautiful*" is an adjective, etc.). Lemmatisation consists of obtaining the canonical form of a word, as it would appear in a dictionary (e.g., the words "*ate*" and "*eaten*" would share the same lemma, "*eat*"). Both PoS tagging and lemmatisation help reduce the target vocabulary and simplify later processes.

Next, the polarity of each sentence was calculated using the well-known Bing Liu's (2010) English polarity lexicon[2]. As this lexicon is domain independent, some manual revision was undertaken to discard irrelevant words. In this step, negation is considered as a polarity shifter, which means that words in the scope of a negation particle represent the opposite polarity (e.g., "good" is positive but, in the scope of the negation "not", as in "not good", it becomes negative). The overall sentiment polarity of sentences was calculated using the ratio of positive (or negative) words detected in them, after taking into account negation.

A further step applied a similar dictionary-based approach to classify each piece of text into a set of predefined categories or topics. We focused on 4 categories of the tourist experience: environmental conditions, activities, recommendations and price (see Table 2). The environmental category captures words related to how visitors cognitively and emotionally perceive the physical characteristics and the surroundings of a cultural attraction (Alnawas & Hemsley-Brown, 2019). The activities category captures words linked to the act of consumption (interacting with the cultural attraction) and is linked to the behavioural dimension of consumer experience (Brun, Rajaobelina, Ricard, & Berthiaume, 2017). Recommendations capture advice about what to do, or not to do, when visiting the cultural attraction (Park & Lee, 2008); it is part of the social component of the tourism experience. Price category captures words related to consumers' perceptions of the money they spent on the attraction and the surrounding hospitality services (i.e., restaurants); it is related to the cognitive component of customer experience. We preferred manually curated resources over automatically generated (e.g., using topic-modelling techniques, such as latent Dirichlet allocation [LDA], Osmani, Mohasefi, & Gharehchopogh, 2020), because user-generated content, being informal text, is prone to generate a lot of noise and the resulting categories are difficult to control and require additional curation.

The objective/subjective classification of the texts was addressed as a binary classification problem powered by a deep neural network. The

Table 2
Categories of the tourist experience.

Environmental conditions	Overcrowded, Motion, Hustle, Busy, Wet, People everywhere, Smelly, Crowded, Footpaths, Waterways, Dirty, Chaotic
Price	Reasonable, Overpriced, Expensive, Value for money, Cheap
Activities	Walk, Footage, Merchants, Guide, Explore, Breathe, Glass-making, Glass, Traffic, Gondola ride, Sing, Photo, Aperitives, Parade, Palaces, Sailing, Views, Mansions, Race, Banks, Watch, Shopping, Tour, Shops, Ticket, Burano, Churches, Gondola, Routes, Film festival, Gelateria, View, Restaurants, Sea, Boat size, Sunset, Bridges, Restaurant, Murano, Boat tour, Sunrise, Vaporetto, Boat
Recommendations (what to do)	App, Wine, Seat near exit, Day, Night, Toilets, Old buildings, Alleys, Selfie stick, Sculpture, Dinner, Taxi, No coffees, Not by foot, Validation, Afternoon, Validate ticket, Gondola trip, Gondola, Private taxi boat, Spring, Foot, Shop, Route 1, Well dressed, Early, Sit, Travel card, Walk, Group, Explore, Guide, September, Rick Stevens guide, No cruise, Evening, No tourist traps, Not October, No high tide, Free tour, Rialto Market, Day pass, Coffee, Photos, Move, Watch, Private gondola, Tour, Stroll around, Dusk, Not summer, Shared Gondola, Lunch, Taxi, Murano, Boat side, High tide, Tour, Vaporetto, Day, Share

Source: Own design

[1] http://www.vicomtech.org.

[2] https://www.cs.uic.edu/~liub/FBS/sentiment polarity-analysis.html.

first step was to generate in-domain word embeddings (Rudkowsky et al., 2018). Word embeddings are dense vectors of continuous numeric values that represent and capture semantic and lexical information. By in-domain, we mean that these dense vectors are generated using texts from the target domain (online reviews) to enhance their semantic representation of the domain concepts. These word embeddings are fed into a recurrent neural network, a long short-term memory (LSTM), which uses all the words of a given document (i.e., online review) in the order they appear in the text to build a dense representation of the whole document. This document representation is the numeric vector obtained from the last state of the recurrent neural network after feeding in all the words, and it serves as the input into a fully connected feed-forward layer that performs a linear transformation. The outcome is that the document representation is converted into two values which, after passing through a softmax layer, describe a probability distribution over two possible classes (i.e., being objective vs. being subjective).

Neural networks are powerful tools for creating models for a variety of tasks, including text classification. For neural networks the text needs to be converted into dense vectors of continuous values. We used the well-known word2vec algorithm (Mikolov, Chen, Corrado, & Dean, 2013) to create a word-embedding model which was then applied directly to the UGC. Thus we obtained domain-adapted semantic representations of the words which, in turn, helped us gather domain-aware semantic representations of the text documents (i.e., online reviews). The word2vec algorithm is an unsupervised language model, so it does not need manually labelled examples, just pieces of natural text. This results in a vocabulary of words, each with a corresponding vector of continuous numerical values, that implicitly encode information about the words they represent. We used a word-embedding size of 200, that is, the resulting vectors for each word had 200 dimensions. When we had a semantic representation in the form of a dense vector for each word in the vocabulary, we used them to feed the recurrent neural network.

We used recurrent neural network architecture with LSTMs (Hochreiter & Schmidhuber, 1997) as the recurrent unit. LSTMs keep a hidden state after each processing time-step, so that each time a word is fed into the network, the output for the next step is based both on the new input plus the information retained in the hidden state. This allows the LSTM units to have "memory" that helps them keep track of long-range dependencies among words in a piece of text. We used a hidden unit size of 100, and a truncated backpropagation through time (BPTT), of a maximum of 300 steps, which is enough for the length of most of the comments analysed.

The words in a given comment are fed one by one recurrently into the LSTM units. The hidden state of the last step serves as the semantic representation of the whole document, and is input into a fully connected layer of dimensions 100×2. Here, 100 is the size of the LSTM hidden layer, while 2 is the number of classes for the classification task (objective vs. subjective). The raw output of this layer, a vector of size 2, is passed through a softmax layer that normalises the probabilities for each of the two possible outcomes. Fig. 1 depicts the described neural network architecture.

For training purposes two researchers manually labelled a set of customer comments as being subjective or objective. From these manually labelled comments, the algorithm automatically balanced both classes to avoid biasing the classifier towards the more frequently used class. From the resulting balanced set of labelled comments, 80% were used as training data, while the remaining 20% were retained as evaluation data to assess if the classifier keeps learning and improving its predictions on unseen data after each training epoch. In our experiments more than the 70% of the test comments (not used for the training) were correctly classified as objective/subjective based on the gold annotations, which is reasonably high considering the subjective nature of the task. When the training was completed, the resulting model was stored and used to predict the most probable class of new unseen examples that used the outputs of the last layer to predict the likelihood of the text being objective or subjective.

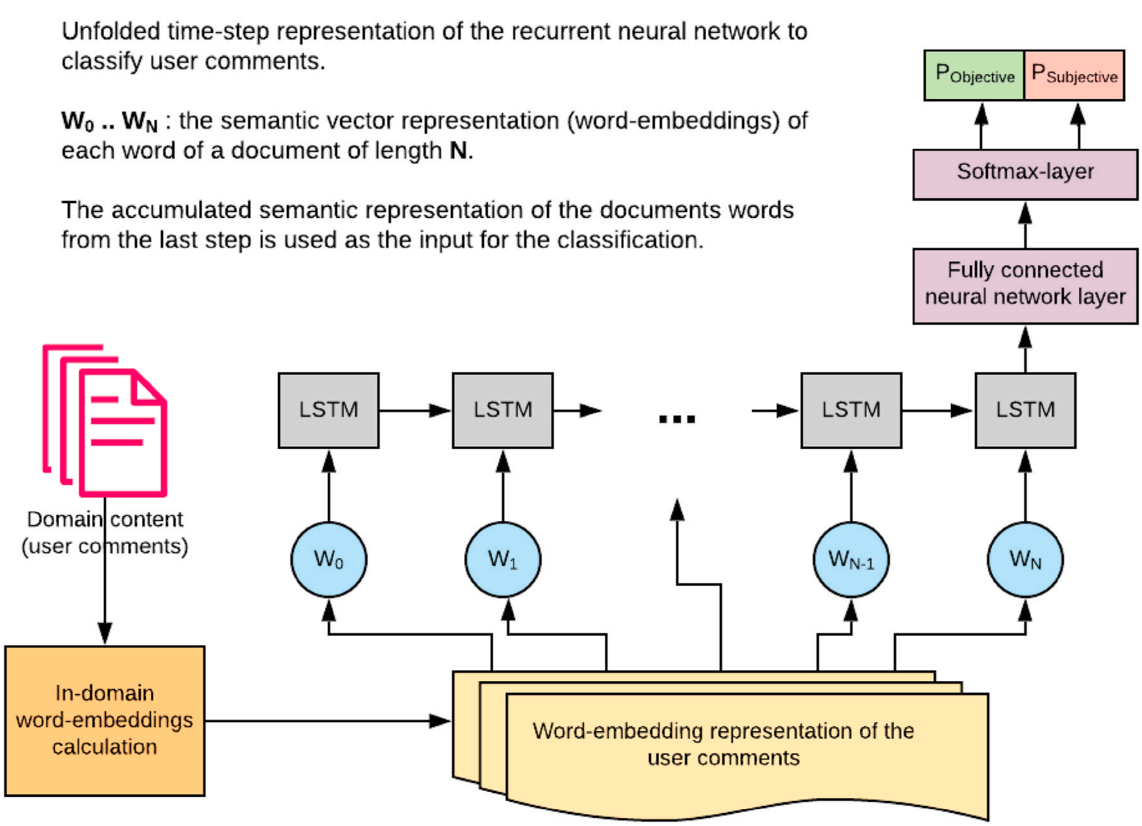

Fig. 1. Diagram of the recurrent neural network that classifies user comments as being objective or subjective.

4. Results

As shown in Table 3, the variables that most influence whether or not a review about St. Mark's Square received a helpful vote were the number of photos embedded in the comment (B = 0.195; H2a supported), the number of positive words (B = 0.160; H4a supported), the number of negative words (B = 0.178; H3a supported), and the number of years the reviewer has been on TripAdvisor (B = 0.145; H1a supported). Reviews with positive sentiment polarity about activities are less likely to receive a helpful vote (B = −0.146; H6a rejected). Unlike St. Mark's Square, which is a free attraction where price perceptions are related to visits to nearby museums, restaurants and souvenir shops, the Doge's Palace is a paid-for attraction that costs 20 euros (2019) to enter. The variables that most influence whether or not a review about the Doge's Palace receives a helpful vote are the subjectivity of the message (B = 0.422; H7a supported), the number of negative words (B = 0.367; H3a supported) and the reviewer's expertise (B = 0.168; H1a supported). Reviews with positive recommendations have a negative effect on the probability of receiving a helpful vote (B = −0.147; H6a rejected for recommendation category).

The results of the linear regression model showed an R^2 of 0.327 for St. Mark's Square ($F = 16.101$, $p = 0.000$) and 0.206 for the Doge's Palace ($F = 13.574$, $p = 0.000$); the Durbin-Watson statistics are 1.151 and 1.047, respectively. As they are in the range 1–3, they are acceptable (Field, 2009). Further underlying assumptions of the linear regression analysis were also tested, such as the normality, linearity and homogeneity of variances with normal P–P plots and scatterplots. The absence of multicollinearity is confirmed as tolerance is close to 1, and the variable inflation factor (VIF) is less than 3 (see Table 4). The linear regression analysis provided different coefficient significance results for the two attractions (Table 4).

The results at Table 4 suggest that reviewer expertise is the only variable that influences the perceived helpfulness of the reviews of both types of attraction. The greater the reviewer's number of years on TripAdvisor, the greater the number of helpful votes the reviews receive. Indeed, in the case of the Doge's Palace, this is the most influential factor. Online review sentiment polarity influences helpfulness for both tourist attractions. The main positive influencing factor on the number of helpful votes received by reviews about St. Mark's Square was negative sentiment polarity about activities (B = 0.482; H5b supported), followed by the reviewer's expertise (B = 0.193; H1b supported), the number of positive words in the review (B = 0.162; H4b supported), and the number of photos embedded in the review (B = 0.157; H2b supported). In contrast, positive sentiment polarity about the environmental conditions of the attraction impacted negatively on the helpfulness votes the reviews received (H6b rejected). The variables with the greatest influence on the number of helpful votes received by reviews of the Doge's Palace are: the reviewer's expertise (B = 0.336; H1b supported), positive sentiment polarity about price (B = 0.226; H6b supported for price category), recommendations of what not to do (negative recommendations) (B = 0.124; H5b supported for recommendation category) and the subjectivity of the message (B = 0.129; H7b supported). Users' experiences expressed in subjective messages are important for readers, as are personal recommendations about what not to do. Price seems to be an important factor for the palace, as the reviews with positive words about price impacted positively (B = 0.226; H6b supported for price category), while reviews with negative words about price impacted negatively (B = −0.150; H5b not supported for price category). Pictorial content is not significant. Consumers perceive purchase risk when they buy a ticket to visit the Doge's Palace, so they process not only pictures but also information cues, such as sentiment polarity about price, and recommendations made by expert reviewers, to make their final purchase decision.

Table 4
Linear regression results for St. Mark's Square and the Doge's Palace.

	Standard. Beta coeff.	*t*-value	Sig.	Tolerance	VIF
St. Mark's Square					
Constant		0.487	0.627		
Negative words (activities)	0.482	7.262	0.000	0.805	1.087
Positive words (environ. Conditions)	−0.216	−3.103	0.002	0.920	1.192
Years using TripAdvisor (expertise)	0.193	2.946	0.004	0.839	1.062
Positive words in the review	0.162	2.281	0.024	0.942	1.242
Number of photos	0.157	2.446	0.015	0.980	1.021
$R^2 = 0.327$. Adjusted $R^2 = 0.306$					
Durbin-Watson = 1.151					
Doge's Palace					
Constant		−3.775	0.000		
Years using TripAdvisor (expertise)	0.336	6.045	0.000	0.979	1.021
Positive words (price)	0.226	3.564	0.000	0.754	1.326
Negative words (price)	−0.150	−2.372	0.018	0.760	1.316
Negative recommendation words	0.124	2.217	0.027	0.972	1.028
Objective vs subjective review	0.129	2.328	0.021	0.993	1.007
$R^2 = 0.206$. Adjusted $R^2 = 0.191$					
Durbin-Watson = 1.047					

Source: Own design

5. Conclusions and discussion

DMOs should study online reviews patterns both because helpful reviews reflect the tourist behaviour of the active UGC users, and

Table 3
Logistic regression results for St. Mark's Square and the Doge's Palace.

	St. Mark's Square			Doge's Palace		
	Beta	Wald	Sig	Beta	Wald	Sig
Number of photos	0.195	6.155	0.013			[a]
Positive words	0.160	15.134	0.000			[a]
Positive words (activities)	−0.146	4.336	0.037			[a]
Years using TripAdvisor	0.145	23.963	0.000	0.168	36.611	0.000
Negative words	0.178	7.127	0.008	0.367	17.974	0.000
Objective vs subjective review			[a]	0.422	7.250	0.007
Positive recommendation words			[a]	−0.147	4.083	0.043
Constant	−3.507	157.844	0.000	−2.706	69.544	0.000
	Chi square = 64.862, p = 0.000			Chi square = 66.462, p = 0.000		
	Cox & Snell R Square = 0.053			Cox & Snell R Square = 0.079 Nagelkerke R Square = 0.109		
	Nagelkerke R Square = 0.094					

[a] = non significant.
Source: Own design

because they influence the behaviour of the far bigger group of passive social media users who use UGC to make travel choices. Ganzaroli et al. (2017) suggested that if DMOs analyse the UGC they might alter tourist behaviour and break the vicious circle faced by mature destinations. Opinion-mining tools allow DMOs to analyse large amounts of UGC and, therefore, to gain control over what is being said, to tackle negative opinions and, in general, to make smarter decisions. This study makes two methodological contributions: (i) combining sentiment polarity analysis techniques and regression analysis, it explains the perceived helpfulness of online reviews for free and paid-for attractions; (ii) it identifies, using DL, which type of content (subjective/objective) of UGC reviews consumers consider more helpful.

One goal of this study is to assess the impact of UGC information cues on the helpfulness of UGC for mature destinations. The study, to the best of the authors' knowledge, is the first to consider both qualitative review components (pictorial content, reviewer expertise, subjectivity/objectivity of the content) and the sentiment about the features of tourism attractions in the assessment of readers' perceptions of the helpfulness of reviews. The expertise of the reviewer, and the number of negative words in the review, positively influence the perceived helpfulness of UGC for both Venetian attractions. The influence of reviewer expertise on perceived helpfulness of UGC shows that trust in online settings is driven by the expertise of peer recommenders (Smith, Menon, & Sivakumar, 2005). This result supports González-Rodriguez et al.'s (2016) findings about the influence of reviewer expertise on the perceived helpfulness of UGC about mature destinations. However, not all previous studies support these findings. Filieri, Alguezaui, and McLeay (2015) found that source credibility and user experience do not influence consumer trust in UGC. Our findings, aligned with the first research stream, contribute to this debate. They also support previous research that posited that negative UGC is perceived by consumers as more diagnostic and, therefore, more helpful than positive UGC (Baek et al., 2012; Bhandari & Rodgers, 2018).

Another goal was to analyse the impact of the attributes of consumer experiences on the perceived helpfulness of UGC based on the pricing of attractions (free vs. paid-for). Subjective reviews are the more influential driver of whether paid-for attractions receive helpful votes. In the case of free, open attractions, the most influential factor is the number of photos embedded in reviews. Focusing on helpful reviews, in the case of the open attraction St. Mark's Square, the number of negative words about activities is the main driver for the number of helpful votes it receives. The number of positive words about environmental conditions impacts negatively on the number of helpful votes received. Schema theory suggests that consumers progressively gain knowledge about products and attributes. Positive UGC about the environmental conditions in St. Mark's Square perhaps makes a review seem less helpful because its readers have already stored Venice's environmental conditions in their schema, so additional judgments do not add much value, while they have to spend time and effort reading them. Most of the words included in the environmental category refer to the square as being dirty, smelly, overcrowded and/or chaotic; readers probably process this information as a peripheral cue, and do not consider it to be important. Sentiment about price is an important driver of the number of helpful votes received by reviews about the Doge's Palace (a goal-oriented attraction). Any positive piece of information about the price, or value-for-money, of museums is processed through the central route, and is perceived as being helpful for deciding whether to visit them. As Venice is perceived as an expensive city, sharing negative sentiments about the price of paid-for attractions doesn't provide added value to other tourists. Venice is a well-known destination, therefore, negative recommendations (what not to do) influence the number of helpful votes received for paid-for attractions, but positive recommendations (what to do), are not perceived as helpful. This finding supports previous research on diagnoscity and vividness (Baek et al., 2012; Baek et al., 2012). Given the overwhelming number of positive online consumer reviews about tourist attractions, negative reviews are expected to be visually more salient (vivid) and credible (diagnostic) and, consequently, attract greater attention than positive reviews. The Doge's Palace is a closed, paid-for attraction, so environmental conditions are processed as peripheral cues; thus, they are not drivers of the tourist experience. Subjective comments positively influence the number of helpful votes received by Dodge's Palace. However, results might be different for other types of attraction, such as non-iconic/non-famous attractions, where previous schemas would not be valid because the tourists lack prior knowledge. In those cases, the objectivity of the review might be more influential in building the tourists' schemas for the attractions.

5.1. Theoretical contributions

The main contributions of this study is that it identifies: (i) which of the review components linked to consumers' experiences of cultural attractions in mature destinations impact on the perceived helpfulness of online reviews; (ii) whether the perceived helpfulness of the online reviews varies as a consequence of the price base of the attraction, that is, if it is free or paid-for. A further important contribution is, unlike previous research, this study differentiates between positive and negative sentiment about features of tourism attractions as distinct variables. This takes forward the debate on the application of dual-processing models and positivity and negativity bias in tourism services.

Another important contribution of this study is that it analyses the specific impact of information cues on the perceived helpfulness of UGC, based on whether the attraction is free or paid-for. In extending compensatory model theory (for details, see Hoyer, 1984) to the context of UGC in tourism services, we posit that consumers make trade-offs between the attraction's features as discussed in UGC, by a type of linear compensation, when choosing a brand (tourist attraction). Some attributes are not taken into account in non-compensatory decision--making processes. This suggests that, because consumers make limited cognitive effort, only attributes they consider important are taken into account when they make a decision.

The influence of reviewer expertise on the perceived helpfulness of UGC supports the solid framework in the literature on consumer trust and source credibility (Zhang, Zhang, & Yan, 2016). Thus, trust in the expertise and credibility of the reviewer (source) are key drivers of the perceived helpfulness of online reviews. This study also highlights the importance of capturing consumers' visual attention, which should be regarded as a primary objective in marketing communications, given the information overload in today's marketplace. Specifically, we analysed the relative impact of pictorial content and sentiment polarity on the perceived helpfulness of online reviews. When the attraction involves an economic cost and, therefore, a higher perceived purchase risk (paid-for attractions), pictorial content is not a determinant of the helpfulness of the UGC. Pictorial content helps consumers reduce the number of attractions they consider, and sentiment polarity is used to make the final choice of a tourism attraction.

5.2. Managerial implications

The sentiment analysis performed in this study can guide DMOs in their implementation of specific eWOM marketing strategies to better respond to customer's information needs. For example, DMOs should prioritise strategies to deal with negative comments about activities in open attractions. This also holds true for paid-for attractions, that is, the more negative words included in reviews, the more helpful votes the reviews obtained. We recommend that DMOs and tourism companies keep track of helpful reviews, use opinion-mining tools to process the vast amount of UGC that appears on social media, and fix the problems revealed in helpful reviews before they influence the decisions of potential customers.

Heuristic information cues have different impacts on free and paid-for tourist attractions. We suggest that DMOs and tourism services companies prominently display reviews on their websites that contain the information that tourists have said they find helpful for the particular destination being promoted, for example, pictures for open attractions. For both free and paid-for attractions, we recommend managers prominently display reviews posted by high-source credibility authors; web features/tools might be implemented to allow these reviews to be more easily found by consumers. Reviewer experience, measured by the number of reviews they have posted, might be used by DMOs to filter the high number of opinions to obtain a reduced set of opinions with which to work when drawing conclusions about customers' needs and preferences for destinations.

DMOs should also pay special attention to reviews about free attractions, which contain a large number of positive words, to reviews with negative words about activities at the attraction, and to reviews with pictorial content. Second, we recommend that managers of museums that charge for entry display on their opening web pages, first, subjective reviews and, second, clear price information. We also suggest that for paid-for tourism attractions, DMOs focus on price (sales promotions) and how to improve subjective recommendations about what not to do.

5.3. Limitations and future research lines

This study focused on the analysis of the content of online reviews. Sentiment analysis of the headings of reviews, individual ratings and review length could also explain the perceived helpfulness of online reviews. This paper carried out the sentiment analysis on reviews posted originally only in English. In future research we intend to analyse the review sentiment of attractions posted both in Spanish and English, as reviews posted in these languages represent more than 85% of the UGC posted on attractions on TripAdvisor (TripAdvisor, 2020).

This research combines opinion-mining tools with regression analysis. Future research could use other DL methods to obtain results beyond the regression analysis. This study uses ML and deep-learning techniques, but the field is progressing very fast and we have merely scratched the surface. More advanced approaches allow the processing of multilingual data, unsupervised semantic clustering, and training of more robust models with less manually labelled data. Combining DL with big data opens new avenues for tourism research. Artificial neural networks have the ability to deal with non-linear relationships, missing data and outliers, and offer great predictive accuracy (Bigné, William, & Soria-Olivas, 2020). Furthermore, other approaches based on artificial intelligence should be explored.

We focused on the analysis of helpful reviews. Aside from argument quality, there may be many other reasons why a specific review does not receive a helpful vote (e.g., recently posted reviews, too long, format). Further research using the characteristics of reviews with many helpful votes and data-mining techniques may aid in the development of predictive models that can determine the helpfulness of reviews that do not receive helpfulness votes.

This study achieved novel results about the influence of words linked to four specific categories of the tourism experience: attraction-linked activities, environmental conditions, price and recommendations. Further research might use sentiment analysis to examine other categories, such as weather, facilities close to the attraction and accessibility; this could help identify what specific features of a cultural tourist attraction will be perceived as helpful by tourists, and what will not. This study investigated only two popular cultural attractions with a high number of online reviews; more areas and types of attraction need to be examined. Our results can be only generalised to well-known cultural attractions in mature destinations. However, this study raises interesting questions for the design and management of open, free, and paid-for cultural tourist attractions in mature destinations.

Declaration of competing interest

None.

Acknowledgments

The authors gratefully acknowledge the financial support of Spanish Ministry of Science and Innovation, Spain project PID 2019-111195RB-I00.

References

Al-Smadi, M., Qawasmeh, O., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2018). Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *Journal of Computational Science, 27*, 386–393.

Aldayel, H., & Azmi, A. (2016). Arabic tweets sentiment analysis–a hybrid scheme. *Journal of Information Science, 42*(6), 782–797.

Alnawas, I., & Hemsley-Brown, J. (2019). Examining the key dimensions of customer experience quality in the hotel industry. *Journal of Hospitality Marketing & Management, 28*(7), 833–861.

Amplayo, R., & Song, M. (2017). An adaptable fine-grained sentiment analysis for summarization of multiple short online reviews. *Data & Knowledge Engineering, 110*, 54–67.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce, 17*(2), 99–126.

Bernini, C., & Cagnone, S. (2014). Analysing tourist satisfaction at a mature and multi-product destination. *Current Issues in Tourism, 17*(1), 1–20.

Bhandari, M., & Rodgers, S. (2018). What does the brand say? Effects of brand feedback to negative eWOM on brand trust and purchase intentions. *International Journal of Advertising, 37*(1), 125–141.

Bigné, E., Ruiz, C., & Currás-Pérez, R. (2019). Destination appeal through digitalized comments. *Journal of Business Research, 101*, 447–453.

Bigné, E., William, E., & Soria-Olivas, E. (2020). Similarity and consistency in hotel online ratings across platforms. *Journal of Travel Research, 59*(4), 742–758.

Brun, I., Rajaobelina, L., Ricard, L., & Berthiaume, B. (2017). Impact of customer experience on loyalty: A multichannel examination. *Service Industries Journal, 37*(5–6), 317–340.

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management, 21*(1), 97–116.

Butler, R. W. (1980). The concept of a tourist area cycle of evolution: Implications for management of resources. *Canadian Geographer/Le Géographe canadien, 24*(1), 5–12.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine, 9*(2), 48–57.

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.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752–766.

Chang, Y., Ku, C., & Chen, C. (2020). Using deep learning and visual analytics to explore hotel reviews and responses. *Tourism Management, 80*, 104–129.

Cheng, Y., & Ho, H. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research, 68*(4), 883–887.

Do, H., Prasad, P., Maag, A., & Alsadoon, A. (2018). Deep learning for aspect-based sentiment analysis: A comparative review. *Expert Systems with Applications, 118*, 272–299.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth: Harcourt Brace Jovanovich College Publishers.

Euromonitor International. (2019). *Top 100 city destinations. 2019 edition*. Available at: https://go.euromonitor.com/white-paper-travel-2019-100-cities.html#download-link. (Accessed 25 April 2020).

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

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM, 56*(4), 82–89.

Field, A. (2009). *Discovering statistics using SPSS: And sex and drugs and rock 'n' roll* (3rd ed.). London: Sage.

Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research, 58*, 46–64.

Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management, 51*, 174–185.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291–313.

Fuchs, M., Höpken, W., & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations–A case from Sweden. *Journal of Destination Marketing & Management, 3*(4), 198–209.

Ganzaroli, A., De Noni, I., & Van Baalen, P. (2017). Vicious advice: Analyzing the impact of TripAdvisor on the quality of restaurants as part of the cultural heritage of Venice. *Tourism Management, 61*, 501–510.

García-Pablos, A., Cuadros, M., & Linaza, M. T. (2016). Automatic analysis of textual hotel reviews. *Information Technology & Tourism, 16*(1), 45–69.

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering, 23*(10), 1498–1512.

González-Rodríguez, M., Martínez-Torres, R., & Toral, S. (2016). Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management, 28*(11), 2609–2627.
He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: A case of analyzing online hotel reviews. *Online Information Review, 41*(7), 921–935.
Hlee, S., Lee, H., Koo, C., & Chung, N. (2020). Will the relevance of review language and destination attractions be helpful? A data-driven approach. *Journal of Vacation Marketing*. https://doi.org/10.1177/1356766720950356 (in press).
Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.
Hong, H., Xu, D., Wang, G., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, 102*, 1–11.
Hoyer, W. (1984). An examination of consumer decision making for a common repeat purchase product. *Journal of Consumer Research, 11*(3), 822–829.
Jiménez-Barreto, J., & Campo-Martínez, S. (2018). Destination website quality, users' attitudes and the willingness to participate in online co-creation experiences. *European Journal of Management and Business Economics, 27*(1), 26–41.
Kim, D., & Park, B. (2017). The moderating role of context in the effects of choice attributes on hotel choice: A discrete choice experiment. *Tourism Management, 63*, 439–451.
Kozak, M., & Martin, D. (2012). Tourism life cycle and sustainability analysis: Profit-focused strategies for mature destinations. *Tourism Management, 33*(1), 188–194.
Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of natural language processing, 2*(2010), 627–666.
Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management, 47*, 140–151.
Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management, 68*, 301–323.
Ma, E., Cheng, M., & Hsiao, A. (2018). Sentiment analysis–a review and agenda for future research in hospitality contexts. *International Journal of Contemporary Hospitality Management, 30*(11), 3287–3308.
Manganari, E., & Dimara, E. (2017). Enhancing the impact of online hotel reviews through the use of emoticons. *Behaviour & Information Technology, 36*(7), 674–686.
Marine-Roig, E., & Clavé, S. (2015). Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing & Management, 4*(3), 162–172.
Marrese-Taylor, E., Velásquez, J. D., & Bravo-Marquez, F. (2014). A novel deterministic approach for aspect-based opinion mining in tourism products reviews. *Expert Systems with Applications, 41*(17), 7764–7775.
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv preprint arXiv:1301.3781.
Mudambi, S., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*(1), 185–200.
Osmani, A., Mohasefi, J. B., & Gharehchopogh, F. S. (2020). Enriched latent dirichlet allocation for sentiment analysis. *Expert Systems*, Article e12527.
Park, D., & Lee, J. (2008). eWOM overload and its effect on consumer behavioral intention depending on consumer involvement. *Electronic Commerce Research and Applications, 7*(4), 386–398.
Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research, 50*, 67–83.
Petty, R., & Cacioppo, J. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: William C. Brown.
Phillips, P., Zigan, K., Silva, M., & Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network analysis. *Tourism Management, 50*, 130–141.
Ratchford, B. (2020). The history of academic research in marketing and its implications for the future. *Spanish Journal of Marketing - ESIC, 24*(1), 3–36.
Rojas-Barahona, L. (2016). Deep learning for sentiment analysis. *Language and Linguistics Compass, 10*(12), 701–719.
Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, Š., & Sedlmair, M. (2018). More than bags of words: Sentiment analysis with word embeddings. *Communication Methods and Measures, 12*(2–3), 140–157.
Simeon, M., Buonincontri, P., Cinquegrani, F., & Martone, A. (2017). Exploring tourists' cultural experiences in Naples through online reviews. *Journal of Hospitality and Tourism Technology, 8*(2), 220–238.
Smith, D., Menon, S., & Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing, 19*(3), 15–37.
Sparks, B., Perkins, H., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management, 39*, 1–9.
Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. *Information Fusion, 36*, 10–25.
Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology, 61*(12), 2544–2558.
Tripadvisor. (2020). https://tripadvisor.mediaroom.com/es-about-us.
Tsai, C., Chen, K., Hu, Y., & Chen, W. (2020). Improving text summarization of online hotel reviews with review helpfulness and sentiment. *Tourism Management, 80*, 104122.
Van der Zee, E., Bertocchi, D., & Vanneste, D. (2020). Distribution of tourists within urban heritage destinations: A hot spot/cold spot analysis of TripAdvisor data as support for destination management. *Current Issues in Tourism, 23*(2), 175–196.
Zhang, Z., Zhang, Z., & Yang, Y. (2016). The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tourism Management, 55*, 15–24.
Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management, 76*, 111–121.