



Universitat de Lleida

Document downloaded from:

<http://hdl.handle.net/10459.1/64698>

The final publication is available at:

<https://doi.org/10.1016/j.ijhm.2017.10.016>

Copyright

cc-by-nc-nd, (c) Elsevier, 2018



Està subjecte a una llicència de [Reconeixement-NoComercial-SenseObraDerivada 4.0 de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Modelling a grading scheme for P2P accommodation: Stars for Airbnb

Eva Martin-Fuentes*^a, Cesar Fernandez^b, Carles Mateu^b and Estela Marine-Roig^a
University of Lleida

^a Department of Business Administration.

E-mail: eva@aegern.udl.cat estela.marine@aegern.udl.cat

^b INSPIRES Research Institute.

E-mail: cesar@diei.udl.cat carlesm@diei.udl.cat

University of Lleida. C/ Jaume II, 73, 25001 LLEIDA, Spain.

Telephone number: + 34 973 703216

*corresponding author

Abstract

This study aims, firstly, to determine whether hotel categories worldwide can be inferred from features that are not taken into account by the institutions in charge of assigning such categories and, if so, to create a model to classify the properties offered by P2P accommodation platforms, similar to grading scheme categories for hotels, thus preventing opportunistic behaviours of information asymmetry and information overload. The characteristics of 33,000 hotels around the world and 18,000,000 reviews from Booking.com were collected automatically and, using the Support Vector Machine classification technique, we trained a model to assign a category to a given hotel. The results suggest that a hotel classification can usually be inferred by different criteria (number of reviews, price, score, and users' wish lists) that have nothing to do with the official criteria. Moreover, room prices are the most important feature for predicting the hotel category, followed by cleanliness and location.

Keywords: Airbnb; hotel classification system; Support Vector Machine; big data; peer-to-peer accommodation platform

1. Introduction

The sharing economy is defined as “an economic system based on sharing underused assets or services, for free or for a fee, directly from individuals” (Botsman and Rogers, 2011), and it is significantly changing consumption patterns (Byers et al., 2013). In tourism, the sharing economy is not a new phenomenon because this peer-to-peer (P2P) exchange has existed for a long time with, for example, the typical advertisement “For rent” hanging from beach apartments, with the owners directly offering short-term rentals or short stays to others, or with individuals waiting for backpackers to arrive at bus stations to offer them a room in their home to get extra income.

With the advance of the Internet, the tourist accommodation sector is experiencing a revolution (Cheng, 2016), with businesses such as Couchsurfing, HomeExchange, Airbnb, HomeAway or Roomorama acting as intermediaries to facilitate contact between host and guest in a simple, convenient and fast way, allowing hosts to earn extra income (Sigala, 2015).

One of the detected barriers to using P2P accommodation platforms is the lack of trust (Tussyadiah and Pesonen, 2016), so overcoming this barrier is a challenge for these platforms and for people who offer their properties for use by others. This lack of trust is related to information asymmetry, which is generated in any market. This theory, developed by Akerlof (1970) in “The Market for Lemons” explains that the seller (i.e., host) knows exactly the true state of the service offered (apartment, room, studio) and the purchaser (guest) does not know it and does not trust in it. Thus, poor services drive out good quality services from the market, leading to an adverse selection problem that ends up negatively affecting those who offer quality services but are drawn down by those who do not provide good service. There are different ways to avoid the adverse effects of information asymmetry such as transmitting credible information. An example of this is when sellers offer post-sales warranties, since only those sellers who are sure of their products would offer them (Stiglitz, 2002). In this sense, the more information available about their services and the more accurate it is, the more people will be willing to use such services (Harford, 2010).

Moreover, with the huge amount of information generated on the Internet for a single item, e.g., thousands of reviews for a single company or destination, an additional problem of information overload may occur, where users find it impossible to sift out useful or high-quality information or to read all opinions (Marine-Roig, 2017). As a consequence, they become overwhelmed. This issue is also a barrier to P2P consumption as it makes decision-making more difficult.

Thus, and given that hotel classification compensates for information asymmetry (Martin-Fuentes, 2016; Nicolau and Sellers, 2010; Öğüt and Onur Taş, 2012), it can help to reduce the problem of information overload. The aim of this study is to predict a hotel category by taking into consideration certain user-generated content (UGC) parameters and other factors in order to create a model to classify the properties offered by P2P accommodation platforms, similar to grading scheme categories of hotels, driven by the need to provide users with certain guarantees for such accommodation services, thereby allowing them to trust in them and preventing opportunistic behaviours of information asymmetry. This model is applied to Airbnb, the leading platform in the P2P accommodation sector, based on information extracted from 18,000,000 reviews on Booking.com written by guests staying at any of 33,000 hotels in outstanding international destinations.

In order to establish a model to classify accommodation on sharing economy platforms, the Support Vector Machine classification technique developed by Vapnik (1995) will be used. The technique is explained in detail in the existing literature, and although its application has been proven in fields such as medicine, engineering, biology, marketing and others, it has not been widely used in the field of tourism (Akin, 2015; Zheng and Ye, 2009), despite the good results reported.

2. Literature review

This section reviews the collaborative economy with special emphasis on the accommodation sector. The importance of hotel classification in order to avoid information asymmetry and information overload is also reviewed.

2.1. Sharing economy

The sharing economy is a phenomenon that can be considered a consequence of the global financial crisis that began in the late 2000s (Buczynski, 2013). It has exploded in recent years thanks to the information and communication technologies (ICTs) that have enabled purchasers and sellers to get in touch with each other directly and conveniently.

Collaborative consumption, or the sharing economy, promotes the use of goods and services without having ownership of them. In the case of property, ownership is increasingly being replaced by use (Rifkin, 2000). Also called the peer-to-peer (P2P) economy, in collaborative consumption, individuals participate in sharing activities by renting, lending, trading, bartering or swapping goods, services, transportation solutions, space or money (Möhlmann, 2015).

In tourism, P2P platforms have experienced tremendous growth. This applies not only to platforms related to the accommodation sector, but also to those related to the catering, transport and tour-guide sectors (Cheng, 2016). The factors that have led to an increase in the use of these new forms of accommodation are economic, because they are potentially cheaper for travellers than other kinds of accommodation (Guttentag, 2015; Tussyadiah and Pesonen, 2016), social, especially because they allow travellers to be in touch with the local community, and others such as authenticity and sustainability (Botsman and Rogers, 2011; Sigala, 2015; Tussyadiah and Pesonen, 2016), because excessive consumption and unnecessary purchases of products that subsequently will not be used can be avoided by sharing goods (Bulchand Gidumal and Melián González, 2016).

Tussyadiah and Pesonen (2016), in an exploratory study with American and Finnish travellers, found that, among Americans, trust is a barrier to using P2P accommodation (not only trust in hosts but also in technology and transaction safety) and conclude that a significant challenge for P2P accommodation companies is the need to create a mechanism of trust among customers, for example, by including reputation scoring or other consumer protection measures such as safe and secure transaction systems.

In this respect, as Ert et al. (2016) claim, P2P product platforms involve economic risks only, while sharing a home involves additional risks. Moreover, “risks are higher for transactions involving products whose attributes can be evaluated only after purchase and use” (Ba and Pavlou, 2002: 12). Therefore, sharing economy platforms base the way they operate and also their trust system on P2P communication through UGC (Tussyadiah and Zach, 2017). Barriers to creating and consuming UGC have been lowered dramatically (Ayeh et al., 2013). Personal thoughts and opinions posted by users are easily accessible to the global community (Dellarocas, 2003) and potentially affect travellers’ decisions in terms of creating ideas and reducing alternatives (Barreda and Bilgihan, 2013). Indeed, many studies have demonstrated the influence that UGC in general and online travel reviews in particular have on travel-related decisions through the electronic Word-of-Mouth (eWOM) effect (Schuckert et al., 2015). Social media is a particularly powerful and credible source of information among users, and especially among digital natives.

Consumers' opinions have been found to generate more confidence than communications from a company (Gretzel and Yoo, 2008; Vermeulen and Seegers, 2009). Although some users are afraid of biased information and false comments (Blomberg-Nygaard and Anderson, 2016; Hensel and Deis, 2010), the reality is that most users trust social media reviews (Pirolli, 2016), and this is demonstrated by their travel-related behaviour, searching for online advice or information before making reservations (Blomberg-Nygaard and Anderson, 2016; Kim et al., 2011). Thus, online opinions are essential not only for a sharing economy service, but also for the traditional hotel sector (Guttentag, 2015), and should be included in future classification systems to be consistent with customer needs (Blomberg-Nygaard and Anderson, 2016). Moreover, online reviews are useful for promoting properties –especially the less-known ones– (Vermeulen and Seegers, 2009) and for taming the possible adverse effects of asymmetric information (Ba and Pavlou, 2002; Park and Nicolau, 2015).

2.2 Information asymmetry, information overload and star-rating classification system

In a market where one of the parties involved in a buying/selling transaction does not have the same information as the other about a product or service, so-called information asymmetry occurs, which could cause the market to fail (Akerlof, 1970).

There are different mechanisms to avoid opportunistic behaviours of information asymmetry, such as guarantees of certain claims that only those sellers who are confident in the quality of their products would offer, or certification by external auditors to ensure the quality of the product or service (Stiglitz, 2002).

Hotel customers rely on recommendations by friends and family to solve their informational disadvantage because tourism services cannot be tried or tested before purchase (Fernández-Barcala et al., 2010). To some extent, this has been replaced by the role of the travel agent, who acts as an intermediary in a market characterised by such asymmetry (Clerides et al., 2005; Jeacle and Carter, 2011).

Information asymmetry in the hospitality industry can also be countered using other elements such as price, customer review ratings, number of recommendations and average display rank (Cezar and Ögüt, 2016; Martin-Fuentes, 2016; Neirotti et al., 2016; Ögüt and Onur Taş, 2012). Moreover, star-rating classification systems established by third party institutions serve as a tool to mitigate asymmetric information (Martin-Fuentes 2016; Nicolau and Sellers 2010; Núñez-Serrano et al., 2014) and provide guidelines for reducing the hotel booking risk (Neirotti et al., 2016).

In addition to the problem of information asymmetry, sharing economy establishments may face the problem of information overload caused by UGC, which is key to the way they operate and also to the trust system. In travel and hospitality, online travel reviews have increased exponentially and there is usually a huge number of reviews available for the same product or service (De Ascaniis and Gretzel, 2012). However, increased amounts of information can be both a blessing and a curse (O'Connor, 2010). Although UGC information provides users with unbiased, unsolicited and cost-effective data on products and services, information overload may actually prevent consumers from getting a comprehensive idea or high-quality information and, moreover, can complicate the decision-making process (Fang et al., 2016; Marine-Roig, 2017).

Therefore, given that it is impossible for users to read all reviews about a product or service, it is crucial for them to have simplified, uniform and comparable indicators available. In this respect, simplified integrative classification systems, easily understood by all, such as accommodation star-rating levels or simplified indicators, could also help users overcome the information overload.

However, no grading scheme categories similar to hotel classifications exist in the case of P2P accommodation. Hotel classification systems are established using various standards set by governments or by independent organisations. These systems are universally recognised, and the most common method for classifying hotels is to rank them from 1 to 5 stars, although the requirements for assigning the stars differ depending on the institutions responsible for doing so, which can be split into official and non-official (Zhan-Qing and Liu, 1993).

The star-rating classification mechanism is the most common customer segmentation pattern in the hotel industry (Dioko et al., 2013); hotel quality can be inferred from the number of stars (Fang et al., 2016), the highest hotel categories can be considered as an indicator of high quality (Abrate et al., 2011), and it plays a general role in the selection of hotels (Callan 1998; Núñez-Serrano et al., 2014). Furthermore, although the star-rating classification systems are different all over the world, it has been proven that there is a relationship between star-rating classification and satisfaction measured from the point of view of scores assigned by users (Martin-Fuentes et al., 2016).

However, the current hotel star-rating classification system presents some weaknesses. The hotel classification system does not follow the same pattern all over the world as each country has its own criteria. At the European level, attempts to launch a process of harmonisation of different regulations have nevertheless been made (Arcarons i Simon, 2008).

There is an initiative by hotel associations from some European countries, sponsored by the Hotrec Association (Hotels, Restaurants & Cafes in Europe), that is trying to implement a scoring system to enable the unification of criteria for the allocation of stars in different countries (Hotrec, 2015), but it is not an easy task because, even within an individual country, there are different systems in place. This is the case for Spain, which has 17 different classification systems, one for each of the autonomous governments that have the power to regulate in this field.

Moreover, the current hotel classification system does not take into account guests' opinions in the form of UGC. Such increasingly popular UGC, which has now become central to accommodation bookings (Blomberg-Nygaard and Anderson, 2016), would provide a quality check on the amenities and characteristics required for the classification system. Thus, Blomberg-Nygaard and Anderson (2016) suggest that future hotel classification systems should be refined by integrating online reviews into them. The idea of integrating online travel reviews in P2P accommodation classification systems is even stronger because UGC production and user opinions are at their very core. This is so because P2P platforms are online "engagement platforms", whose operation and success is based on value co-creation, information exchange and the production of UGC, among various economic actors in a service ecosystem (Breidbach and Brodie, 2017).

3. Research aim and methodology

This study aims, firstly, to determine whether a hotel classification system can be inferred in general from criteria and standards that are not considered by the rules and regulations of public and non-public organisations in charge of assigning hotel categories and, if so, to use this model to design a classification system for P2P accommodation platforms in order to avoid a market of “lemons” and information overload and to use it on platforms like Airbnb, which are based on user interaction and value co-creation.

In order to contribute to the P2P accommodation literature and, by so doing, to provide additional insights, this study aims to answer this research question: Is it possible to design a classification system for P2P accommodation platforms similar to the hotel classification system in order to avoid a market of “lemons” and information overload on P2P platforms like Airbnb?

As seen in the literature review, the problems of a lack of trust arising from asymmetric information can cause businesses and even markets to collapse, so it is necessary to provide a mechanism that offers its users certain guarantees. Similarly, to address the problems arising from UGC information overload, indicator classification systems that are both comprehensive and simple are needed.

The conclusions by (Ert et al., 2016: 72) state that “the strong need for trust in sharing economy platforms leads consumers to use any information available to them” and the literature review confirms that, on the one hand, hotel star-ratings counter the adverse effects of asymmetric information and, on the other, the hotel classification system is internationally recognised (despite being applied differently in each country); it is a comprehensive system that can help simplify tourists’ decision-making. Moreover, within a context of the ever-increasing popularity of UGC, integrated classifications should combine hotel classification systems focusing on objective features and on guest reviews – providing information about how service-related elements are perceived – since both are complementary (Blomberg-Nygard and Anderson, 2016). In fact, “both consumers and the industry are interested in seeing a closer fit between the two” so that offerings are presented in keeping with consumer needs (Blomberg-Nygard and Anderson, 2016: 3). Therefore, and responding to these needs, the merging of hotel-like classifications with UGC could provide users with a holistic integrative classification system capable of preventing online information overload.

So, on that basis, we try to build a model to classify P2P accommodation that increases customer trust, integrates several elements and is easy to understand worldwide. In this sense, we propose building a simple model for P2P accommodation platforms that consumers can trust, similar to the hotel classification system. The proposed classification model is applied to the leading P2P accommodation platform Airbnb, though it could be used for any other P2P platform.

3.1 Case study: Airbnb

The proposed classification model is applied to the P2P accommodation platform Airbnb. Among P2P accommodation services, Airbnb, founded in 2008, is an example of a leading company that offers properties in 194 countries worldwide (Airbnb, 2016a). Airbnb is a company that bases its business model on putting individuals who have a space for rent in contact with other individuals who want to rent it for a period of time in exchange for money, all of which is done via the Internet. Airbnb is the best-known P2P accommodation community marketplace platform, with more than 34,000 cities, 60 million guests and 2 million listings (Airbnb, 2016a). It can be considered a sharing economy engagement platform based on user interaction and value co-creation (Breidbach and Brodie, 2017). Airbnb connects travellers with local hosts that provide a space, which can be entire properties, castles, rooms, beds, sofas, airbeds or any kind of accommodation. The way it works is that guests contact hosts through the Airbnb platform to confirm availability and to get more information through the messaging system. Airbnb charge the guests and holds their money until 24 hours after the check in, to allow guest and host to confirm that everything is in order, and later on Airbnb transfers the money to the host.

As previously seen, trust management in P2P platforms like Airbnb is a critical element of their business, which the company knows and announces with the slogan “Trust is what makes it work”. Airbnb uses a reputation mechanism that allows hosts and travellers to write reviews about each other; a review is based on a previous stay, and is done within 14 days after that stay. Moreover, guests and hosts can scan a government ID to verify their online profiles, and there is a secure messaging system and a host guarantee to cover any possible damage to the property (Airbnb, 2016b).

In addition, Airbnb uses a quality certification called ‘Superhost’ that serves to prevent opportunistic behaviours of information asymmetry and information overload. However, the percentage of properties that have the ‘Superhost’ badge on Airbnb is very limited (e.g., a mere 2.9% in Hong Kong) (Liang et al., 2017).

This reputation system, in which both the service supplier and the service demander can give their opinions about each other, has also been applied by other collaborative economy companies such as Couchsurfing or BlaBlaCar and can make people think twice before posting a bad review because of fear of revenge (Ert et al., 2016). To avoid it, Airbnb does not publish reviews until both parties have given their opinions or until the deadline has expired to do so.

3.2 Data collection

To create a classification model applicable to P2P accommodation platforms similar to the hotel star-rating classification system, in this study, data is taken from Booking.com because it is a prominent example of an online accommodation-booking website. It has 895,589 properties in 224 countries and deals with over 1 million room-night reservations per day (Booking.com, 2016). On Booking.com, travellers can compare prices and customer reviews (Neirotti, 2016) and it is a popular online source for hotel information (Sun et al., 2015) that draws the attention of researchers.

Moreover, Booking.com and Airbnb have some similarities, apart from being leaders in accommodation reservations; both websites allow customers to give their opinions and have similar systems for collecting them. Thus, both websites only allow users to leave

a review if someone has booked accommodation through the respective site and actually stays at the reviewed property, so all are “verified reviews from real people”, as Booking.com claims.

As mentioned above, the way both websites collect reviews is similar; after a stay, the user receives an e-mail to rate the property and to post a review, the time to rate the property is limited (for Booking.com it is 28 days and for Airbnb it is 14 days) to avoid out-dated opinions.

On Booking.com, users rate 6 items (value, cleanliness, location, services, comfort and staff) and, on Airbnb, users also rate 6 items (value, cleanliness, location, check in, communication and accuracy), the common variables (value, cleanliness, and location) were taken into account in this study.

The selection of these variables and not others to create a star-rating system for P2P accommodation is also justified by the preferences of sharing economy guests compared to those of hotel guests. As emphasized by Tussyadiah and Zach (2017), cleanliness and location are fundamental elements for both types of guest. However, P2P accommodation guests place much greater value on location and social interactions with hosts than hotel guests do. For the latter, convenience and room features are more important (Belarmino et al., 2017; Tussyadiah and Zach, 2017). Earlier studies have matched or compared hotels and P2P accommodation based on price, among other aspects (Belarmino et al., 2017). Value for money was found to be a fundamental driver of booking P2P accommodation, as well as a basic element of satisfaction therewith (Belarmino et al., 2017), since emphasis was placed on this type of accommodation providing better value than hotels did (Harrington, 2015). As a result, cleanliness, location and price should be taken as the core indicators of P2P accommodation analyses.

In April 2016, we automatically gathered the hotel data from Booking.com; the number of reviews, the general scoring, the score for the 6 items, the hotel name, the city, the country, the number of users that saved the hotel in their wish list, the hotel category and the price (ranked from 1 to 5 on Booking.com). The hotels in the world’s 443 top destinations according to TripAdvisor were classified into four regions: Europe (EUR), America (AME), Asia and Pacific (ASP), and the Middle East and Africa (MEA) as suggested by Banerjee and Chua (2016), as shown in Table 1.

The data was collected using an automatically controlled web browser (developed in Python) that simulates user navigation (clicks and selections).

We filtered the results by “Property type”, selecting “Hotels” and “Review score”, with the rated by “All reviewers” option.

Table 1. Data collection by regions

Region	Countries	Destinations	Hotels	Reviews
EUR	17	168	14,395	11,097,703
AME	23	122	7,022	3,285,925
ASP	16	109	10,448	3,559,306
MEA	9	44	1,179	767,947
Total	65	443	33,044	18,710,881

3.3. Support Vector Machine

Support Vector Machines (SVMs) are a useful technique for binary data classification. By finding hyperplanes that separate n-dimensional data, they learn to separate data into two classes (Vapnik, 1995), turning the problem into a set of linear equations. A dataset is often not linearly separable, so the concept of kernel is introduced in SVMs. Different types of kernels may be devised, but the common idea is to cast the original data into a higher dimension dataset that may be separated. Some of the most successful kernels are Radial Basis Function (RBF) kernels, particularly when the number of features is not large, as in our case. RBF kernels are exponential functions defined by a single parameter, namely the exponent constant, being preferable to other types of kernels with a high number of parameters such as polynomials.

In our experiment, we use LIBSVM (CC01a), which is an open source implementation of SVMs written in C code.

When dealing with multiclass data (it should be noted that in our experiment the instances belong to five classes according to the hotel star classification), LIBSVM implements multiclass classification using one-against-all methods (Hsu and Lin, 2002).

Having k classes $\binom{k}{2}$, binary classifiers are constructed. Then, each point to be predicted is classified according to each of the binary $\binom{k}{2}$ classifiers, giving one vote to the class (or classes) to which it has been assigned. Finally, the point is designated to the class with a higher number of votes received.

The decision to use SVM classifying techniques instead of more traditional techniques like an ordered logistic regression was due to some of the core advantages of SVM and because the results obtained with the ordered logistic model were worse, as shown in Table 2.

SVMs enable non-linear data to be easily classified, while providing a more robust model, thanks to the maximisation of the support vector margins (the distance to the separation hyperplane of the support values). It also provides a simple unique solution to the classification problem (whereas deep neural networks can offer multiple solutions to the same problem). Furthermore, it is computationally efficient and widely available (many statistical packages provide SVM implementations), so research reproducibility is guaranteed because it is not dependent on using the same code as that used in deep neural networks.

Because of these features, SVMs are widely used in pattern recognition problems: face and speech recognition, face detection and image recognition, financial classification, medical analytics, etc.

In short, SVMs are a supervised Machine Learning technique that creates a model from sample data (training set); using that model, its validity is then checked on testing data (validation set) to compute a predefined performance measure, called accuracy, which is the ratio of correctly classified instances.

4. Results

A classification task usually involves separating data into training and testing sets. Each instance in the training set contains one “target value” (i.e., the class labels) and several “attributes” (i.e., the features or observed variables). The goal of SVM is to produce a model (based on the training data) that predicts the target values of the test data given only the test data attributes (Hsu et al., 2003).

In our case, the number of classes is 5 when the 5 star-rating classification is used, or 4 when grouped categories are used (i.e. [1,2], [2-3], [3-4] and [4-5]). Regardless of the number of classes, the number of features employed is 6 (1: Cleanliness, 2: Value, 3: Location, 4: Reviews, 5: ListSaved, 6: Price), which are the 6 common features on Booking.com (from where we collected the data) and on Airbnb.

In order to avoid creating large datasets that make SVM computation unfeasible, raw data was split for the training and the testing phases into 10 sets for each region, and the results show the average over the 10 datasets. Furthermore, we have tried to balance the datasets, when possible, with the same number of instances for each class. For the training phase, in EUR we used 300 instances of 1-star and 900 of 2- to 5-star hotels for each of the 10 sets. In AME, we used 40 instances of 1-star and 500 of 2- to 5-star hotels; in ASP, 290 instances of 1-star and 1,000 of 2- to 5-star hotels; and, finally, in MEA, 20 instances of 1-star, 50 of 2-star and 200 of 3- to 5-star hotels.

The results show a high level of accuracy except for the 1-star hotel category in EUR, AME and ASP, and 2-star category in MEA, as shown in Table 2.

Table 2. Results for 5 star-rating classification

Region	Category	Test phase instances	Accuracy SVM	Ratio SVM	Ratio Logit
EUR	1	1,000	12	0.01	0.00
	2	1,000	729	0.73	0.77
	3	1,000	452	0.45	0.28
	4	1,000	371	0.37	0.39
	5	1,000	837	0.84	0.84
AME	1	500	0	0.00	0.00
	2	1,000	686	0.69	0.75
	3	1,000	442	0.44	0.26
	4	1,000	256	0.26	0.33
	5	1,000	876	0.88	0.87
ASP	1	1,000	28	0.03	0.00
	2	1,000	741	0.74	0.79
	3	1,000	363	0.36	0.19
	4	1,000	314	0.31	0.43
	5	1,000	841	0.84	0.82
MEA	1	200	74	0.37	0.00
	2	200	2	0.01	0.00
	3	200	151	0.76	0.15
	4	200	119	0.60	0.08
	5	200	153	0.77	0.17

Grouping the hotels into four categories as budget accommodation (1 and 2-star hotels), mid-low range accommodation (2- and 3-star hotels), mid-high range accommodation (3- and 4-star hotels), and superior accommodation (4- and 5-star hotels), the results improve because there is one category less. Worthy of note is that the training phase was done with the 5 categories coinciding with the 5 star-rating classification, but the results are explained in 4 categories, i.e., the categories in the training phase have not been mixed.

Table 3. Results grouped into 4 categories

Region	Category	Test phase Instances	Accuracy SVM	Ratio SVM	Ratio Logit
EUR	Budget	2,000	1,623	0.81	0.90
	Mid-low range	2,000	1,716	0.86	0.92
	Mid-high range	2,000	1,196	0.60	0.54
	Superior	2,000	1,724	0.86	0.80
AME	Budget	1,500	1,113	0.74	0.46
	Mid-low range	2,000	1,642	0.82	0.89
	Mid-high range	2,000	1,095	0.55	0.52
	Superior	2,000	1,658	0.83	0.78
ASP	Budget	2,000	1,660	0.83	0.94
	Mid-low range	2,000	1,657	0.83	0.90
	Mid-high range	2,000	1,092	0.83	0.50
	Superior	2,000	1,686	0.84	0.80
MEA	Budget	400	110	0.28	0.03
	Mid-low range	400	308	0.77	0.17
	Mid-high range	400	346	0.87	0.19
	Superior	400	355	0.89	0.14

Although in some categories of some regions the ordered logistic model predicts a bit better the category than SVM, the SVM results are more regular worldwide and, in general, show a better accuracy.

The SVM results show that around 80% predicts the accommodation category (divided into budget, mid-low range, mid-high range, and superior category) that a property could have by just knowing the price, the number of reviews, the ratings awarded by past users for location, cleanliness and value, and the number of people that have saved the property in their wish list.

The greatest accuracy of the SVM model is in the superior and the mid-high range categories in MEA. The accuracy in ASP is very high in all categories. In EUR and AME, the lowest accuracy is in the mid-high range category (3- to 4-star) with 60% and 55%, respectively. The worst prediction is found in MEA in the budget category, probably due to the fact that the parameters do not capture the difference in this range, as shown in Table 3.

In order to determine the importance of the features when classifying, LIBSVM proposes a method based on F-scores (Chen and Lin, 2006). For a given feature, its F-score is computed as the ratio of the discrimination between the positive and negative

sets, over the particular value of the feature within each of the two sets. The larger the F-score is, the more likely this feature will be more discriminative, it being useful as a feature selection criterion. Even though the F-score does not reveal mutual information among features, it is a simple and efficient method.

Table 4. F-scores

Features	EUR	AME	ASP	MEA
Cleanliness	0.44	0.63	0.35	0.63
Value	0.05	0.12	0.03	0.19
Location	0.14	0.35	0.10	0.37
Listsaved	0.02	0.02	0.06	0.04
Reviews	0.06	0.01	0.07	0.08
Price	1.35	1.41	1.45	1.23

Table 4 shows that price is the most important feature when classifying, followed by cleanliness, and in AME and MEA, the location is also important; the other factors serve to explain the model but the significance is considerably lower. In this respect, the classification model based on the features of price, cleanliness and location is also well-suited to P2P accommodation platforms, which are engagement platforms intrinsically bound to UGC. Price is a key issue for P2P guests' choice and satisfaction (Wang and Nicolau, 2017), and cleanliness and location are among the most frequently mentioned topics – and in a similar order – in both P2P and hotel guests' UGC (Belarmino et al., 2017).

An ordered logistic regression was also performed to see if better results could be obtained, but the accuracy was generally worse when compared to the SVM results; the worst results were in MEA, where a maximum accuracy of 17% in 5-star hotels was obtained because there were fewer instances to predict the category than in the other regions.

To check the validity of the model in Airbnb properties, a small random sample of two cities was downloaded manually. The results show that properties were distributed between all categories (Table 5) and not solely among the top of the classification, something that might have been expected given that Airbnb properties' ratings tend to have a high positive skew and that it is rare to find a property with a lower than 3.5 rating (Zervas et al., 2015).

Table 5. Percentage of properties in Airbnb by four categories

Category	Subsample Airbnb
Budget	29.73%
Mid-low range	21.62%
Mid-high range	27.03%
Superior	21.62%

5. Discussion

Although there are differences among countries and even among regions within the same country, and while there are some systems that group hotel categories into fewer than 5 levels (e.g. Malta from 2- to 5-star), we can affirm that the hotel classification

system can be inferred in general through other criteria and standards that have nothing to do with the official criteria assigned by regulations of public and non-public organisations.

Attempts to launch a process of harmonisation of the different regulations (Arcarons i Simon et al., 2008) to enable the unification of criteria for allocating stars in different countries (Hotrec, 2015) could be made by taking into account the User-Generated Content translated into ratings along with other factors such as price or score, because the results of this study show that it can be predicted with great accuracy worldwide, thereby avoiding criteria that could become out-dated with the passage of time (Torres et al., 2014). Such allocation of stars could also help users by reducing the problem of UGC information overload and by making P2P accommodation classification more comprehensive and simple.

The greatest accuracy is in 5-star hotels worldwide because it “is the only category that has a certain uniformity from an international point of view” (Minazzi 2010: 80). The accuracy of the remaining categories is fairly high, except the 1-star category in EUR, AME, and ASP, and the 2-star category in MEA. This may be because customers of 1- and 2-star hotels use ratings systems less often than those staying in 3- to 5-star hotels in the United Kingdom (Callan, 1995). It may also be because user satisfaction coincides with the hotel category in nine European countries, except for 1- and 2-star hotels, where there are no significant differences in seven countries (Austria, Germany, Greece, Italy, Poland, Portugal and Spain) (in press).

With the results grouped into 4 categories, the lowest accuracy is in the mid-high range category (3- to 4-star) in EUR and AME because hotel supply in that category varies from country to country (Minazzi, 2010).

The results confirm that price is the most important characteristic for inferring the hotel category, which is consistent with the idea that the hotel classification system is a regular forecaster of prices (Israeli, 2002), and that room prices have a very strong direct linear relationship with hotel categories (Martin-Fuentes, 2016). In this respect, it is worth mentioning that price could also be considered a good predictor in the case of P2P accommodation, as market rules would prevent the system from being tricked through prices (e.g., if a price does not correspond to what is being offered, customers would not buy it or would create negative reviews about it, resulting in reputation loss). This prevention would be stronger in the case of P2P platforms as their success is attributed to their wide range of prices (Wang and Nicolau, 2017) suited to different offerings.

Although charging higher prices might seem to be a way of achieving a higher rating, this is something that would not happen, at least in the short run, because, Airbnb or any other P2P accommodation platform should only classify a property if there are at least 5 reviews, which is what Booking.com does.

The next important feature for inferring the star category is cleanliness, which is also consistent with other studies that confirm cleanliness is a factor that strongly affects travellers' choices (Atkinson, 1988) and is a relevant feature to the hotel users' satisfaction (Barreda and Bilgihan, 2013; Choi and Chu, 2001), although it is not a determinant of hotel room price (Zhang et al., 2011). This variable had previously been

identified as fundamental to P2P accommodation travellers' choices and satisfaction, as expressed in their online UGC (Belarmino et al., 2017; Tussyadiah and Zach, 2017).

Location, especially in AME and MEA, is also important for predicting the star category, which is a variable that has a significant effect on price (Saló et al., 2014; de Oliveira Santos 2016) and is a feature that cannot be changed or improved once the property has been built (Xie et al., 2014). This variable has been identified as key for P2P accommodation guests through UGC analysis (Belarmino et al., 2017; Tussyadiah and Zach, 2017).

The application of the model to a sample of Airbnb properties demonstrates the validity and applicability of the model for P2P accommodation platforms. Remarkably, this model brings together most of the factors identified by Varma et al., (2016) that Airbnb users and non-users employ when choosing accommodation, such as price, location, past experiences and reputation, and classify them into categories that are understandable worldwide.

In this respect, trust can generate positive outcomes by reducing transaction risks (Ba and Pavlou, 2002). It is vital for organisations like Airbnb, which is vulnerable to bad press (Guttentag, 2015), and it is the main concern for Consumer-Generated Media organisations (Fileri et al., 2015) or for websites where UGC creates more confidence than communications from a company do (Gretzel and Yoo, 2008; Vermeulen and Seegers, 2009). Consequently, our model can give consumers of Airbnb or any other P2P accommodation platform additional confidence because the grading scheme categories in the hospitality sector are useful when it comes to mitigating asymmetric information (Martin-Fuentes 2016; Nicolau and Sellers 2010; Núñez-Serrano et al., 2014).

In addition, this model helps to provide customers with holistic information, which might make them feel they can rely more on Airbnb, and presents a comprehensive classification that may help users deal with the huge amounts of information available online. In a synthetic way, the aforementioned classification combines a well-known hotel-like classification system with UGC, and it is especially useful for users' faced with information fragmentation and overload. Moreover, a strength of the model is that it is partially based on UGC, along with other relevant factors, which consumers find highly trustworthy and is therefore consistent with the philosophy of P2P platforms as online engagement platforms, based on user interaction, UGC exchange and value co-creation.

Airbnb is trying to put measures in place to encourage service providers to become engaged in quality certifications like the 'Superhost' badge, which might seem as though it is offering a new classification system. However, such certifications have limited scope because a very low percentage of properties has the 'Superhost' badge (Liang et al., 2017), thus restricting the number of options available to users. The distinctive trait of our model is that all properties with at least five reviews will have a star assigned to them.

Unlike Airbnb ratings, which are mostly positive, this model classifies properties into all categories, which provides the user with more information because high star-ratings

alone “are likely not informative enough for users to make informed consumer choice” (Bridges and Vásquez, 2016: 16).

Finally, this model will give more visibility to Airbnb for non-users who seem to be unaware of the existence of this alternative (Varma et al., 2016). It may also be used as a tool for comparing different types of accommodation in a given destination because most travellers use Airbnb to find alternatives to hotels. It is worth noting that only 2.3% of respondents indicated that without Airbnb they would not have taken the trip (Guttentag and Smith, 2017).

6. Conclusions

The rapid growth of the sharing economy, especially in the tourism sector, can be hampered by trust issues, as most business relationships will be customer-to-customer (C2C). To provide a measure of trust, several methods are being used: customer reviews, reputation-based systems, sharing on social media and so on. But, for those situations where prospective customers or others within their social networks do not have previous experience, a classification system is required to avoid opportunistic behaviours due to asymmetric information. Moreover, an integrative and synthetic classification is necessary, on the one hand, to prevent the information overload that tourists experience online and, on the other, to assist tourists in their decision-making processes. Such systems are in place for hotels, namely the star-rating system, but they are not implementable for sharing economy sites, as they would require too many resources and too much effort to work: inspections, bureaucracy, etc. Our system would overcome these limitations because it is not resource hungry; it only requires affordable computational resources. Besides, as the proposed system is partly based on UGC, it would also be useful for ensuring that hosts put every effort into offering a faultless service to get the best online reviews from users, thereby managing to reach the highest category.

The proposed system is not only useful for classifying P2P accommodation and overcoming trust issues in P2P platforms, but also represents an improvement for the entire hotel classification system. The current official system for assigning hotel categories creates many differences within the same category since stars do not mean the same worldwide. Many hotel category classifications work with a system of points that lets hotels organise themselves as they wish, meeting certain minimum requirements, but the same category has many differences. Besides, a classification system that includes guests’ reviews is a proposal by the United Nations World Tourism Organization (Blomberg-Nygaard and Anderson 2016: 26), principally because consumers use reviews and hotel classifications in different ways: “classification systems help filter hotels, whereas guest reviews provide a means to help select from a smaller set of acceptable options. These similar yet distinct uses indicate a continued need for both hotel classification and guest reviews.” Our proposed system goes a step further since it combines reviews with a new classification system for hotels and P2P accommodation platforms in order to help simplify users’ decision-making.

In this respect, our model would therefore provide several advantages:

1. Matching users’ point of view: Experts deciding which criteria should be applied to officially allocate hotel categories should know that the ideal classification

system is the one adapted to the users' needs, which takes into account that satisfaction might be obtained through eWOM.

2. Converging different systems: To standardise and bring in line the criteria used in different countries and regions in order to help users understand what each hotel category stand for.
3. Validating the official classification system: Audits would not need to be carried out to check that the criteria applied are being met, thus eliminating bureaucracy.

In short, the implementation of this model would achieve a reorientation of hotel classification systems to improve them, thereby bringing different countries, regions and systems in line and matching the systems to the users' opinions. This model contributes to the literature by determining which elements are more significant for inferring international hotel categories, bearing in mind that the system of hotel classification is not unified (each country and region applies its own regulations). Furthermore, as these elements are also available in P2P accommodation platforms, they can be used to create a model of accommodation categories equivalent to the hotel grading scheme, thus achieving a unified, comprehensive and real-time accommodation classification model based on online data.

6.1. Practical implications and limitations

This work could be used by adapting the classification model to other P2P platforms, in addition to Airbnb, such as Couchsurfing and home exchanges, or even to tourist attractions, based on data from travel opinion websites like TripAdvisor. This model could even be used for managing and consolidating online reputation through websites like Traity, with information coming from more than one platform (i.e., a property that is on HomeAway, Airbnb and Couchsurfing) and by combining this information. The application of this model implies in practice that accommodation categories are constantly updated through computerised systems based on available online information, representing a time- and cost-saving option.

The application of the proposed model has one disadvantage; properties that have only just been listed and do not have any reviews by previous users. They would start without any prior classification, but this is the same as currently happens when a user has to rely on a host that does not have a reference because he has not yet received an online travel review. On the positive side, once hosts achieve a category within the proposed model, it will prevent them from quitting their online identity if they have negative reviews (Ba and Pavlou, 2002), since the cost of getting a new online identity would involve the loss of the qualifying category. Such a cost might therefore be higher than trying to reverse the situation by improving the service offered to guests. Moreover, even if price (as set by the host) is part of the classification system, market rules and the inclusion of users' reviews would prevent hosts from setting higher prices to achieve higher categories than would otherwise correspond the characteristics of their offerings. Similarly, to prevent the system from being manipulated through the use of price change strategies to achieve either a higher or a lower category, a moving price average for a given number of days or months in the past could be used in the model, or, even, instead of assigning categories on a particular date, it could be done online instantly, i.e., every time a user visits the accommodation information. By doing so, the price reflected would be the real price and any manipulation of the system could be mitigated.

Like any other research, the research presented here is not exempt from limitations. The proposal is made using data obtained from Booking.com, which, despite being a very popular website, is not the only one in existence. The results could be different if data from other websites like Ctrip, HolidayCheck, or TripAdvisor were used. Indeed, replicating the study using other data is a challenge for further research.

Finally, this research has been done using a large volume of data downloaded automatically from Booking.com and has explored the capacity of big data in the hospitality field, as suggested by Xiang et al. (2015).

Acknowledgements

This work was partially funded by the Spanish Ministry of the Economy and Competitiveness: research project TIN2015-71799-C2-2-P and ENE2015-64117-C5-1-R.

This research article has received a grant for its linguistic revision from the Language Institute of the University of Lleida (2017 call).

7. References

- Abrate, G., Capriello, A., Fraquelli, G., 2011. When quality signals talk: Evidence from the Turin hotel industry. *Tour. Manag.* 32, 912–921. doi:10.1016/j.tourman.2010.08.006
- Airbnb, 2016a. Airbnb [WWW Document]. URL <http://www.airbnb.com> (accessed 11.9.16).
- Airbnb, 2016b. Airbnb [WWW Document]. URL <https://www.airbnb.co.uk/help/article/4/how-does-airbnb-help-build-trust-between-hosts-and-guests> (accessed 12.20.16).
- Akerlof, G.A., 1970. The Market for“ Lemons”: Quality Uncertainty and the Market Mechanism. *Q. J. Econ.* 84, 488–500.
- Akin, M., 2015. A novel approach to model selection in tourism demand modeling. *Tour. Manag.* 48, 64–72. doi:10.1016/j.tourman.2014.11.004
- Arcarons i Simon, R., Goitia Serra, V., González Aznar, N., 2008. La clasificación hotelera en la Unión Europea: Un mercado poco común. *Pap. Tur.* 49–67.
- Atkinson, A., 1988. Answering the eternal question: what does the customer want? *Cornell Hotel Restaur. Adm. Q.* 29, 12–13.
- Ayeh, J.K., Au, N., Law, R., 2013. “Do We Believe in TripAdvisor?” Examining Credibility Perceptions and Online Travelers’ Attitude toward Using User-Generated Content. *J. Travel Res.* 52, 437–452. doi:10.1177/0047287512475217
- Ba, S.L., Pavlou, P.A., 2002. Evidence of the effects of trust building technology in electronic markets: Price, premiums and buyer behaviors. *MIS Q.* 26, 243–268. doi:10.2307/4132332
- Banerjee, S., Chua, A.Y.K., 2016. In search of patterns among travellers’ hotel ratings in TripAdvisor. *Tour. Manag.* 53, 125–131. doi:10.1016/j.tourman.2015.09.020
- Barreda, A., Bilgihan, A., 2013. An analysis of user-generated content for hotel experiences. *J. Hosp. Tour. Technol.* 4, 263–280. doi:10.1108/JHTT-01-2013-0001
- Belarmino, A., Whalen, E., Koh, Y., Bowen, J.T., 2017. Comparing guests’ key attributes of peer-to-peer accommodations and hotels: mixed-methods approach. *Curr. Issues Tour.* 1–7. doi:10.1080/13683500.2017.1293623
- Blomberg-Nygaard, A., Anderson, C.K., 2016. United Nations World Tourism

- Organization study on online guest reviews and hotel classification systems: An integrated approach. *Serv. Sci.* 8, 139–151. doi:10.1287/serv.2016.0139
- Booking.com, 2016. About Booking.com [WWW Document]. URL <http://www.booking.com/content/about.en-gb.html> (accessed 9.23.16).
- Botsman, R., Rogers, R., 2011. *What's mine is yours: how collaborative consumption is changing the way we live*. Collins London.
- Breidbach, C.F., Brodie, R.J., 2017. Engagement platforms in the sharing economy: conceptual foundations and research directions. *J. Serv. Theory Pract.* 27.
- Bridges, J., Vásquez, C., 2016. If nearly all Airbnb reviews are positive, does that make them meaningless? *Curr. Issues Tour.* 1–19. doi:10.1080/13683500.2016.1267113
- Buczynski, B., 2013. *Sharing is good: How to save money, time and resources through collaborative consumption*. New Society Publishers.
- Bulchand-Gidumal, J., Melián-González, S., 2016. *Una guía para entender la economía colaborativa*.
- Byers, J., Proserpio, D., Zervas, G., 2013. The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *SSRN Electron. J.* (June 9, 2). doi:10.2139/ssrn.2366898
- Callan, R.J., 1998. Attributional Analysis of Customers' Hotel Selection Criteria by U.K. Grading Scheme Categories. *J. Travel Res.* 36, 20–34. doi:10.1177/004728759803600303
- Callan, R.J., 1995. Hotel classification and grading schemes, a paradigm of utilisation and user characteristics. *Int. J. Hosp. Manag.* 14, 271–283.
- Cezar, A., Ögüt, H., 2016. Analyzing conversion rates in online hotel booking. *Int. J. Contemp. Hosp. Manag.* 28, 286–304. doi:10.1108/IJCHM-05-2014-0249
- Chen, Y.-W., Lin, C.-J., 2006. Combining SVMs with various feature selection strategies, in: *Feature Extraction*. Springer, pp. 315–324.
- Cheng, M., 2016. Sharing economy: A review and agenda for future research. *Int. J. Hosp. Manag.* 57, 60–70. doi:10.1016/j.ijhm.2016.06.003
- Choi, T.Y., Chu, R., 2001. Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. *Int. J. Hosp. Manag.* 20, 277–297. doi:10.1016/S0278-4319(01)00006-8
- Clerides, S., Nearchou, P., Pashardes, P., 2005. Intermediaries as bundlers, traders and quality assessors: the case of UK tour operators.
- De Ascaniis, S., Gretzel, U., 2012. What's in a review title?, in: Fuchs, M., Ricci, F., Cantoni, L. (Eds.), *Information and Communication Technologies in Tourism*. Springer, New York, NY, pp. 460–470.
- de Oliveira Santos, G.E., 2016. Worldwide hedonic prices of subjective characteristics of hostels. *Tour. Manag.* 52, 451–454. doi:10.1016/j.tourman.2015.07.001
- Dellarocas, C., 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Manage. Sci.* 49, 1407–1424.
- Dioko, L.A.N., So, S.-I. (Amy), Harrill, R., 2013. Hotel category switching behavior—Evidence of mobility, stasis or loyalty. *Int. J. Hosp. Manag.* 34, 234–244. doi:10.1016/j.ijhm.2013.04.002
- Ert, E., Fleischer, A., Magen, N., 2016. Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tour. Manag.* 55, 62–73. doi:10.1016/j.tourman.2016.01.013
- Fang, B., Ye, Q., Kucukusta, D., Law, R., 2016. Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tour. Manag.* 52, 498–506. doi:10.1016/j.tourman.2015.07.018
- Fernández-Barcala, M., González-Díaz, M., Prieto-Rodríguez, J., 2010. Hotel quality

- appraisal on the Internet: a market for lemons? *Tour. Econ.* 16, 345–360.
doi:10.5367/000000010791305635
- 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. *Tour. Manag.* 51, 174–185.
doi:10.1016/j.tourman.2015.05.007
- Gretzel, U., Yoo, K.H., 2008. Use and Impact of Online Travel Reviews, in: O'Connor, P., Höpken, W., Gretzel, U. (Eds.), *Information and Communication Technologies in Tourism*. Springer Vienna, Vienna, pp. 35–46. doi:10.1007/978-3-211-77280-5_4
- Guttentag, D., 2015. Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Curr. Issues Tour.* 18, 1192–1217.
doi:10.1080/13683500.2013.827159
- Guttentag, D.A., Smith, S.L.J., 2017. Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. *Int. J. Hosp. Manag.* 64, 1–10. doi:10.1016/j.ijhm.2017.02.003
- Harford, T., 2010. *The undercover economist*. Hachette UK.
- Harrington, R., 2015. Vacation rentals: Commercial activity butting heads with CC&Rs. *Calif. West. Law Rev.* 51, 187–224.
- Hensel, K., Deis, M.H., 2010. Using social media to increase advertising and improve marketing. *Entrep. Exec.* 15, 87–98.
- Hotrec, 2015. Hotrec [WWW Document]. Hotel. Union polished criteria its Hotel. Categ. URL <http://www.hotrec.eu/newsroom/press-releases-1714/hotelstars-union-polished-the-criteria-for-its-hotelstars-categories.aspx>
- Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2003. A practical guide to support vector classification.
- Hsu, C.-W., Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks* 13, 415–425.
- Israeli, A.A., 2002. Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. *Int. J. Hosp. Manag.* 21, 405–424.
doi:10.1016/S0278-4319(02)00037-3
- Jeacle, I., Carter, C., 2011. In TripAdvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organ. Soc.* 36, 293–309.
doi:10.1016/j.aos.2011.04.002
- Kim, M.-J., Chung, N., Lee, C.-K., 2011. The effect of perceived trust on electronic commerce: Shopping online for tourism products and services in South Korea. *Tour. Manag.* 32, 256–265. doi:10.1016/j.tourman.2010.01.011
- Liang, S., Schuckert, M., Law, R., Chen, C.-C., 2017. Be a “Superhost”: The importance of badge systems for peer-to-peer rental accommodations. *Tour. Manag.* 60, 454–465.
- Marine-Roig, E., 2017. Online travel reviews: A massive paratextual analysis, in: Xiang, Z., Fesenmaier, D.R. (Eds.), *Analytics in Smart Tourism Design: Concepts and Methods*. Springer, Heidelberg, Germany, pp. 179–202. doi:10.1007/978-3-319-44263-1_11
- Martin-Fuentes, E., 2016. Are guests of the same opinion as the hotel star-rate classification system? *J. Hosp. Tour. Manag.* 29, 126–134.
doi:10.1016/j.jhtm.2016.06.006
- Martin-Fuentes, E., Fernandez, C., Mateu, C., 2016. Las estrellas muestran el camino: coincidencia entre satisfacción y categoría hotelera, in: *Universidad de Malaga (Ed.), ICT & Tourism International Conference TURITEC*. Malaga.

- Minazzi, R., 2010. Hotel classification systems: a comparison of international case studies. *Acta Univ. Danubius. Economica* 6, 64–86.
- Möhlmann, M., 2015. Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *J. Consum. Behav.* 14, 193–207. doi:10.1002/cb.1512
- Neirotti, P., Raguseo, E., Paolucci, E., 2016. Are customers' reviews creating value in the hospitality industry? Exploring the moderating effects of market positioning. *Int. J. Inf. Manage.* 36, 1133–1143. doi:10.1016/j.ijinfomgt.2016.02.010
- Nicolau, J.L., Sellers, R., 2010. The quality of quality awards: Diminishing information asymmetries in a hotel chain. *J. Bus. Res.* 63, 832–839. doi:10.1016/j.jbusres.2009.06.009
- Núñez-Serrano, J.A., Turrión, J., Velázquez, F.J., 2014. Are stars a good indicator of hotel quality? Assymmetric information and regulatory heterogeneity in Spain. *Tour. Manag.* 42, 77–87. doi:10.1016/j.tourman.2013.10.004
- O'Connor, P., 2010. Managing a hotel's image on TripAdvisor. *J. Hosp. Mark. Manag.* 19, 754–772. doi:10.1080/19368623.2010.508007
- Ögüt, H., Onur Taş, B.K., 2012. The influence of internet customer reviews on the online sales and prices in hotel industry. *Serv. Ind. J.* 32, 197–214. doi:10.1080/02642069.2010.529436
- Park, S., Nicolau, J.L., 2015. Asymmetric effects of online consumer reviews. *Ann. Tour. Res.* 50, 67–83. doi:10.1016/j.annals.2014.10.007
- Pirolli, B., 2016. Travel information online: navigating correspondents, consensus, and conversation. *Curr. Issues Tour.* 1–7. doi:10.1080/13683500.2016.1273883
- Rifkin, J., 2000. *The age of access: The new culture of hypercapitalism. Where All Life is a Paid.* Exp. Tarcher, New York 33, 40–51.
- Saló, A., Garriga, A., Rigall-I-Torrent, R., Vila, M., Fluvià, M., 2014. Do implicit prices for hotels and second homes show differences in tourists' valuation for public attributes for each type of accommodation facility? *Int. J. Hosp. Manag.* 36, 120–129. doi:10.1016/j.ijhm.2013.08.011
- Schuckert, M., Liu, X., Law, R., 2015. Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *J. Travel Tour. Mark.* 32, 608–621. doi:10.1080/10548408.2014.933154
- Sigala, M., 2015. Collaborative commerce in tourism: implications for research and industry. *Curr. Issues Tour.* 1–10. doi:10.1080/13683500.2014.982522
- Stiglitz, J.E., 2002. Information and the Change in the Paradigm in Economics. *Am. Econ. Rev.* 92, 460–501.
- Sun, S., Fong, L.H.N., Law, R., Luk, C., 2015. An Investigation of Gen-Y's Online Hotel Information Search: The Case of Hong Kong. *Asia Pacific J. Tour. Res.* 1–14. doi:10.1080/10941665.2015.1062405
- Torres, E.N., Adler, H., Behnke, C., 2014. Stars, diamonds, and other shiny things: The use of expert and consumer feedback in the hotel industry. *J. Hosp. Tour. Manag.* 21, 34–43. doi:10.1016/j.jhtm.2014.04.001
- Tussyadiah, I.P., Pesonen, J., 2016. Impacts of Peer-to-Peer Accommodation Use on Travel Patterns. *J. Travel Res.* 55, 1022–1040.
- Tussyadiah, I.P., Pesonen, J., 2016. Drivers and barriers of peer-to-peer accommodation stay – an exploratory study with American and Finnish travellers. *Curr. Issues Tour.* 1–18. doi:10.1080/13683500.2016.1141180
- Tussyadiah, I.P., Zach, F., 2017. Identifying salient attributes of peer-to-peer accommodation experience. *J. Travel Tour. Mark.* 34, 636–652. doi:10.1080/10548408.2016.1209153

- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media. (2nd ed. reprinted by Springer-Verlag 2000, 1995 New York, Inc.).
- Varma, A., Jukic, N., Pestek, A., Shultz, C.J., Nestorov, S., 2016. Airbnb: Exciting innovation or passing fad? *Tour. Manag. Perspect.* 20, 228–237.
doi:10.1016/j.tmp.2016.09.002
- Vermeulen, I.E., Seegers, D., 2009. Tried and tested: The impact of online hotel reviews on consumer consideration. *Tour. Manag.* 30, 123–127.
doi:10.1016/j.tourman.2008.04.008
- Wang, D., Nicolau, J.L., 2017. Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *Int. J. Hosp. Manag.* 62, 120–131. doi:10.1016/j.ijhm.2016.12.007
- Xiang, Z., Schwartz, Z., Gerdes, J.H., Uysal, M., 2015. What can big data and text analytics tell us about hotel guest experience and satisfaction? *Int. J. Hosp. Manag.* 44, 120–130.
- Xie, K.L., Zhang, Z., Zhang, Z., 2014. The business value of online consumer reviews and management response to hotel performance. *Int. J. Hosp. Manag.* 43, 1–12.
doi:10.1016/j.ijhm.2014.07.007
- Zervas, G., Proserpio, D., Byers, J., 2015. A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. *Where Every Stay is Above Aver.* (January 23, 2015).
- Zhan-Qing, L., Liu, J.C., 1993. Assessment of the hotel rating system in China. *Tour. Manag.* 14, 440–452.
- Zhang, Z., Ye, Q., Law, R., 2011. Determinants of hotel room price. *Int. J. Contemp. Hosp. Manag.* 23, 972–981. doi:10.1108/09596111111167551
- Zheng, W., Ye, Q., 2009. Sentiment classification of Chinese traveler reviews by support vector machine algorithm, in: *Intelligent Information Technology Application, 2009. IITA 2009. Third International Symposium on. IEEE*, pp. 335–338.