



Discussion paper

Profiling satisfied and dissatisfied hotel visitors using publicly available data from a booking platform



Elena B. Pokryshevskaya, Evgeny A. Antipov*

National Research University Higher School of Economics Kantemirovskaya, St. 3, Saint-Petersburg, 194100, Russia

ARTICLE INFO

Article history:

Received 17 February 2017

Received in revised form 14 June 2017

Accepted 19 July 2017

Keywords:

Hotel

Online reviews

Online ratings

Satisfaction

Predictive analytics

ABSTRACT

We develop a set of models for predicting hotel visitor satisfaction and the probability of complaints about various service aspects. Our empirical analysis is based on 3630 reviews from one of the Dubai hotels. We identify profiles of visitors who are likely to be dissatisfied with the hotel service and need special attention, as well as of visitors, who are likely to be satisfied with the service and, therefore, do not require extra attention. The predictions are based on observable characteristics of visitors, thus making it possible for hotel managers to apply the models in their everyday work. Using content analysis we also reveal specific problems that different groups of visitors encountered and infer which of the problems has the highest impact on the overall satisfaction with the hotel.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

According to a number of studies user-generated content on the Internet has a higher credibility than traditional sources of tourist information (Akehurst, 2009; Gretzel and Yoo, 2008). Many researchers have also shown that online reviews are important drivers of hotel bookings (Blal and Sturman, 2014; Ye et al., 2011, 2009), as well as booking intentions and trust (Sparks and Browning, 2011).

Despite the proven impact of online reviews on hotel business performance, the literature on explaining hotel ratings and the valence of reviews is still rather scarce. Antipov and Pokryshevskaya (2014), used hotel-level data from South Cyprus to find percentage contributions of various service attributes to the propensity of visitors to recommend the hotel. Their study concentrated on exploring differences between hotels, not between customers, and made use of average numeric ratings of various service attributes, which, unlike textual reviews, are rarely available on hotel booking websites and usually provide less detailed information about the specific problems encountered by clients. Stringam and Gerdes (2010), on the contrary, utilized textual reviews to explore word choices of guests scoring hotels at lower ratings versus higher ratings. Similarly, Zhou et al. (2014) identified 17 attributes that were then classified into 4 classes based on the type of their influence on the overall satisfaction: satisfiers (work

in a solely positive direction), dissatisfiers (work in an exclusively negative direction), bidirectional forces (influence either positive or negative outcomes), and neutrals (no marked influence). The study suggested that a focused reclassification of online review data can deliver powerful customer feedback messages for both researchers and hotel managers. Both Zhou et al. (2014) and Stringam and Gerdes (2010) used bivariate types of analysis without making an attempt to isolate the impact of each attribute from all others using multivariate techniques. Del Chiappa and Dall'Aglia (2012) applied content analysis of Booking.com hotel reviews to study the overall ratings given by travelers and the words used in their comments, in order to identify the most frequent complaints, but did not do any multivariate modeling of either satisfaction ratings or the incidence of certain problems. Bulchand-Gidumal et al. (2011) provided some evidence that it is possible to improve hotel ratings by 8% by offering free Wi-Fi. Recent research results indicate that satisfied customers who are willing to recommend a hotel to others refer to intangible aspects of their hotel stay, such as staff members, more often than unsatisfied customers. On the other hand, dissatisfied customers more often mention the tangible attributes of the hotel stay, such as furnishing and finances (Berezina et al., 2016).

Some of the previous studies applied various text mining techniques for automated extraction of tangible and intangible aspects of customer experience from reviews and studying their impact on the overall customer satisfaction or willingness to recommend the hotel (Berezina et al., 2016). Xu and Li (2016) used latent semantic analysis (LSA) to avoid manual coding of reviews. Their results indicate that the antecedents of customer satisfaction and dissatisfaction vary across various types of hotels. Xiang et al. (2015)

* Corresponding author.

E-mail address: enantipov@hse.ru (E.A. Antipov).

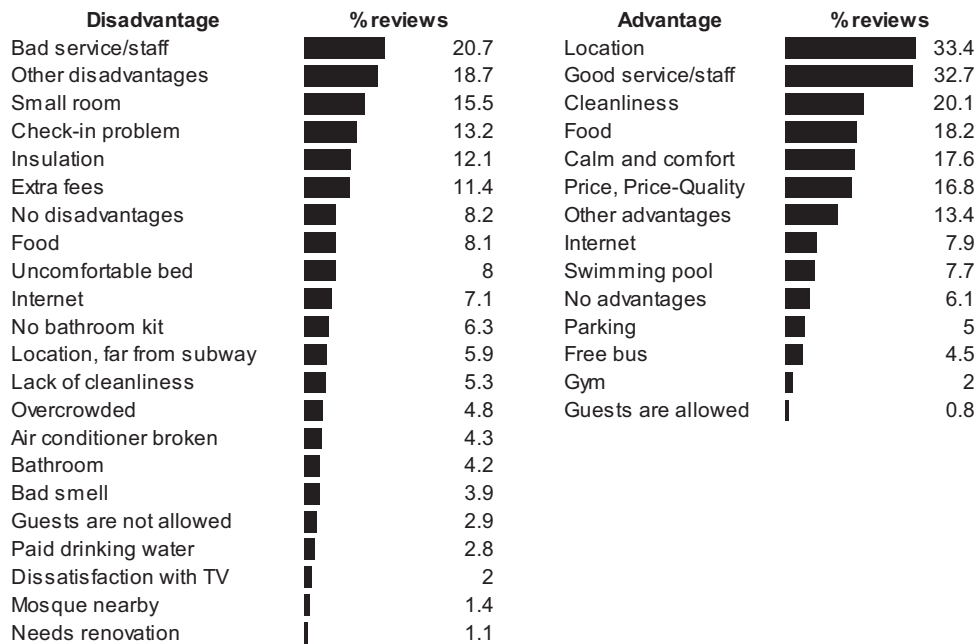


Fig. 1. Most common hotel advantages and disadvantages mentioned in customer reviews (sample size: 1994 observations).

grouped 34 words that are practically important and commonly occur in reviews into 6 components using factor analysis and showed their significance in explaining differences in average customer ratings of hotels. All these studies show the potential of using natural language processing for identifying customer satisfaction drivers. However, a major methodological drawback of the above-mentioned studies which most of their authors mention is insufficient disambiguation and semantic valence detection during the coding process.

Despite growing interest in big data and the analysis of user-generated content in the hospitality industry (Wan and Law, 2017; Xiang et al., 2015), there has still been no literature on predictive modeling in customer complaint management, whereas such modeling could be useful for designing some sort of early warning systems for predicting problems that a particular guest is likely to complain about. That is why in our study we focus on predictive modeling of customer satisfaction based on visitor-specific characteristics. More specifically, we suggest a methodology for identifying profiles of visitors who are likely to be dissatisfied with the hotel service and need special attention, as well as of visitors, who are likely to be satisfied with the service and, therefore, do not require extra attention. Using content analysis we also reveal demographic groups of visitors that complain about certain problems more often than other groups of customers. While it is widely accepted that hotels should analyze online reviews on a regular basis in order to identify the strengths and weaknesses of their property and competing ones (Berezina et al., 2016), profiling proposed in our paper can allow hotel management to predict a customer's level of satisfaction with the hotel and make an educated guess of what may cause his or her dissatisfaction and, as a result, be proactive in preventing at least some of these problems before the customer complains about them.

Our approach to the analysis of publicly available reviews allows answering 3 questions:

1. Which customer characteristics are predictors of overall satisfaction and which demographic groups have the highest/lowest average satisfaction?
2. Which customer characteristics are predictors of dissatisfaction with problems commonly faced by hotel visitors and which

demographic groups have the highest/lowest probability of complaining about each of the problem?

3. Which disadvantages/advantages mentioned in customer reviews have the strongest impact on the overall satisfaction with the hotel?

Besides making conclusions based on secondary data collected from Booking.com we also provide recommendations for developing a systematic internal customer satisfaction survey that will allow collecting data on a greater number of observable characteristics and therefore create more detailed profiles and predict customer satisfaction more accurately.

2. Data and methods

We have used data on 3630 reviews left at Booking.com in August 28, 2014–December 14, 2015 by people who stayed at City-max Hotel Bur Dubai – one of the largest hotels in Dubai. The advantages of the data source compared to other websites with hotel reviews are as follows:

1. Only a customer who has booked through Booking.com and stayed at the property in question can write a review. This ensures that all reviews come from real guests.
2. To keep the rating score and review content relevant, Booking.com archives reviews older than 24 months.
3. In each review advantages and disadvantages are clearly separated, which simplifies text processing for further quantitative data analysis.

Each reviewer answers a number of questions that allows profiling visitors based on their personal characteristics and the descriptors of their trip. Based on these profiles a hotel manager can predict what level of satisfaction and what kind of dissatisfiers are typical of the segment of visitors to which the person who has arrived belongs.

The following variables were used as dependent variables:

1. Rating (0–10)

Download English Version:

<https://daneshyari.com/en/article/5108196>

Download Persian Version:

<https://daneshyari.com/article/5108196>

[Daneshyari.com](https://daneshyari.com)