Chapter 6

Using Data Science to Predict Hotel Booking Cancellations

Nuno António
ISCTE Instituto Universitário de Lisboa, Portugal

Ana de Almeida
ISCTE Instituto Universitário de Lisboa, Portugal & Centro de Informática e Sistemas da Universidade de Coimbra, Portugal

Luis M. M. Nunes
ISCTE Instituto Universitário de Lisboa, Portugal & Instituto de Telecomunicações, Portugal

ABSTRACT

Booking cancellations in the hospitality industry not only generate revenue loss and affect pricing and inventory allocation decisions, but they also, in overbooking situations, have the potential to affect the hotel’s online social reputation. By employing data sets from four resort hotels and addressing this issue as a classification problem in the scope of data science, the authors demonstrate that it is possible to build models for predicting booking cancellations with accuracy results in excess of 90%. This research also demonstrates that despite what was alleged by Morales and Wang (2010), it is possible to predict with high accuracy whether a booking will be canceled. Results allow hotel managers to act on bookings with high cancellation probability and contain the associated revenue losses, produce better net demand forecasts, improve overbooking/cancellation policies, and have more assertive pricing and inventory allocation strategies.

INTRODUCTION

Bookings represent a contract between a customer and a service provider (Talluri & Van Ryzin, 2004). This contract gives the customer the right to use the service in the future at a settled price, usually with an option to cancel the contract prior to the service provision. In the case of the hospitality industry, this option to cancel the booking puts the risk on the hotel. The hotel has to guarantee rooms to the customers who honor their bookings, but at the same time it has to bear with the opportunity cost of
vacant capacity when a customer cancels a booking or does not show up (Talluri & Van Ryzin, 2004). Cancellation rates vary from hotel to hotel. For the purpose of this chapter, no-shows will be treated as cancellations, even though there are some differences between them. A cancellation occurs when the customer terminates the contract prior to his or her arrival. A no-show occurs when the customer does not inform the hotel and fails to check in.

Canceled bookings can represent up to 20% of the total bookings (Morales & Wang, 2010). However, in airport/roadside hotels this number can rise as high as 60% (Liu, 2004). These cancellations can have a substantial affect on revenue, not only because of the revenue loss they represent themselves, but also because of the effect they can have on pricing and inventory allocation decisions (Morales & Wang, 2010).

To compensate for the potential revenue losses caused by cancellations, hotels often sell above their capacity (overbooking) (Ivanov & Zhechev, 2012; Mehrotra & Ruttley, 2006; Morales & Wang, 2010). However, overbooking can also generate costs (Hayes & Miller, 2011; Mehrotra & Ruttley, 2006): reallocation of customers to alternative hotels, cash compensations, or social reputation. Thus, classifying hotel bookings with high cancellation probability is relevant to enable hotels to act on those bookings to prevent or mitigate their cancellation effects. At the same time, this prediction facilitates an easier identification of cancellation patterns, hence allowing a better understanding of net demand and a better definition of overbooking/cancellation policies.

Using uncensored data from four hotel Property Management Systems (PMS) that represent this tendency for hotels to have increasingly higher booking cancellations rates (illustrated in Figure 1), the authors aim to demonstrate how data science can be applied in the scope of hotel revenue management to:

1. Identify features from PMS databases with predictive strength regarding a booking cancellation probability.
2. Build a model that could predict bookings with a high cancellation probability.
3. Understand if one prediction model fits all hotels or if a specific model should be built for each hotel.

BACKGROUND

Booking cancellations in the hospitality industry have a well-known body of knowledge in the scope of revenue management, particularly how the definition of a hotel’s overbooking and cancellation policies can affect its revenue and inventory allocation (Hayes & Miller, 2011; Ivanov, 2014; Talluri & Van Ryzin, 2004). Revenue management is defined as “the application of information systems and pricing strategies to allocate the right capacity to the right customer at the right price at the right time” (Kimes & Wirtz, 2003, p. 125). Originally developed in 1966 by the aviation industry (Chiang, Chen, & Xu, 2007), revenue management was gradually adopted by other services industries, such as hotels, rental cars, golf courses, and casinos (Chiang et al., 2007; Kimes & Wirtz, 2003). For the hospitality industry (rooms division), the revenue management definition was adapted to “making the right room available for the right guest and the right price at the right time via the right distribution channel” (Mehrotra & Ruttley, 2006, p. 2). It requires constant monitoring of supply and demand, historical data analysis, forecasts, implementation of strategic and tactical restrictions/controls by market segments and distribution channels, detection of patterns, and tendencies and anomalies, among other processes that require data from multiple sources (Hayes & Miller, 2011; Ivanov & Zhechev, 2012; Serra, 2013).
Related Content

The Role of Self-Confidence in the Criteria of Aesthetic Labour Recruitment
www.igi-global.com/article/the-role-of-self-confidence-in-the-criteria-of-aesthetic-labour-recruitment/218906?camid=4v1a

Inertia Stages and Tourists' Behavior: Moderator Effects of Zone of Tolerance, Switching Barriers and External Opportunities
Zhiyong Li, Rui Cui, Li Li, Yingli Hu and Ruwan Ranasinghe (2018). *International Journal of Tourism and Hospitality Management in the Digital Age* (pp. 1-17).
www.igi-global.com/article/inertia-stages-and-tourists-behavior/201103?camid=4v1a

Sustainable Entrepreneurship Development in the Accommodation Sector: The Case of the “Albergo Diffuso” (Diffused Hotel) in Italy
www.igi-global.com/article/sustainable-entrepreneurship-development-in-the-accommodation-sector/189741?camid=4v1a

Impact and Role of Social Media in Tourism Sector
www.igi-global.com/chapter/impact-and-role-of-social-media-in-tourism-sector/165076?camid=4v1a