International Journal of Hospitality Management 71 (2018) 120–131

Contents lists available at ScienceDirect

International Journal of Hospitality Management

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

Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning

Yufeng Ma[a], Zheng Xiang[b,*], Qianzhou Du[c], Weiguo Fan[d]

[a] Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA
[b] Department of Hospitality and Tourism Management, Pamplin College of Business, Virginia Tech, Blacksburg, VA 24061, USA
[c] Department of Business Information Technology, Pamplin College of Business, Virginia Tech, Blacksburg, VA 24061, USA
[d] Center for Business Intelligence & Analytics, Department of Accounting and Information Systems, Pamplin College of Business, Virginia Tech, Blacksburg, VA 24061, USA

ARTICLE INFO

Keywords:
Online reviews
Hotel
User-provided photos
Computer vision
Deep learning
Social media analytics

ABSTRACT

Online reviews have been extensively studied in the hospitality and tourism literature. However, while user-provided photos embedded in online reviews accumulate in large quantities, their informational value has not been well understood likely due to technical challenges. The goal of this study is to introduce deep learning for computer vision to understand information value of online hotel reviews. Using a dataset collected from two social media sites, we compared deep learning models with other machine learning techniques to examine the effect of user-provided photos on review helpfulness. Findings show that deep learning models were more useful in predicting review helpfulness than other models. While user-provided photos alone did not have the same impact as review texts, combining review texts and user-provided photos produced the highest performance. Implications for the applications of deep learning technologies in hospitality and tourism research, as well as limitations and directions for future research, are discussed.

1. Introduction

With the widespread adoption of the smartphone and other hand-held devices in everyday life as well as the availability of multitudes of social media platforms, online photo sharing has become a common social activity that facilitates people's exchange of experiences and opinions. Particularly in hospitality and tourism, photography plays an integral role in the travel experience, and photo sharing is an important activity in documenting, reliving, and sharing that experience (Chalfen, 1979; Garrod, 2008; MacKay and Fesenmaier, 1997; Markwell, 1997). While photo sharing may serve many purposes, user-provided photos on the Internet are becoming increasingly important in the context of product evaluation (Konijn et al., 2016; Vu et al., 2015). For example, many online review sites such as TripAdvisor and Yelp allow users to post photos along with product reviews, which in tandem may generate higher impact on consumers' perception of the travel products during the decision making process (e.g., Lo et al., 2011). As such, understanding the effects of user-provided photos can be of great interest to both research and practice in hospitality and tourism.

Recent literature studying online consumer reviews in hospitality and tourism has focused on understanding a range of research questions including product attributes, guest satisfaction, product defects as well as the impact of consumer reviews on business performance (e.g., Schuckert et al., 2015; Xiang et al., 2015). Also, there is a growing interest in applying the so-called analytics approach to collect, analyze, summarize and interpret online review data in order to extract useful patterns and insights pertaining to managerial problems (Xiang et al., 2017). However, likely due to technical challenges in collecting and analyzing visual data in large quantities, the impacts of user-provided photos embedded in online reviews have not yet been examined and well understood. Therefore, the primary goal of this study is to introduce deep learning techniques for image processing recently developed in computer science to assess the information value of user-provided photos embedded in online hotel reviews. As the old saying goes, "is a picture worth a thousand words?" Specifically, we designed an analytics exercise by applying deep learning techniques to examine the extent to which user-provided photos influence the perceived helpfulness of online hotel reviews. As such, this study evaluates the utilities of deep learning techniques within a social media setting and discusses its implications for research and practice in hospitality and tourism.

* Corresponding author.
E-mail addresses: yufengma@vt.edu (Y. Ma), philxz@vt.edu (Z. Xiang), qiand12@vt.edu (Q. Du), wfan@vt.edu (W. Fan).

https://doi.org/10.1016/j.ijhm.2017.12.008
Received 8 April 2017; Received in revised form 1 December 2017; Accepted 15 December 2017

2. Research background

2.1. Analytics of online reviews

Online consumer reviews have been widely considered a rich information source that represents consumer experiences and evaluation of products (Chevalier and Mayzlin, 2006). Within the context of hospitality and tourism, online reviews is an important type of user-generated content due to the fact that the products are experiential, and consumers tend to rely on the electronic word-of-mouth for decision making purposes (Sparks et al., 2013; Xiang et al., 2015). By taking advantage of the abundance of data on the Internet, recent literature has attempted to understand a number of research problems such as travel motivation (e.g., Wu and Pearce, 2014), impact of online reviews on hotel performance (e.g., Zhang et al., 2016; Ye et al., 2009), and opinions and sentiments reflected in the experiences of hospitality products (e.g., Crotts et al., 2009; Levy et al., 2013; Xiang et al., 2015). Although these studies may differ from each other in terms of study focus and specific methodology, they can be generally defined as analytics research, which primarily relies on applying analytical tools and frameworks to collect, analyze, and interpret social media data to extract useful patterns and insights (Fan and Gordon, 2014). In this stream of literature, online reviews are generally conceptualized as a bundle of information components that reflect the website's business model, technological affordances, and user segments with the goal to communicate with and even persuade other online consumers (e.g., Filieri et al., 2015; Sparks et al., 2013). As such, research on online reviews has been focused upon how, and the extent to which, they influence consumers' perceptions, attitudes and behaviors in response to these user-generated contents.

There has been a variety of conceptual framework focusing on different characteristics of online reviews (e.g., Filieri et al., 2015; Hong et al., 2017). Based upon the recent literature, Xiang et al. (2017) deconstruct online reviews into several basic components including linguistic features, semantic features, sentiment, and peripheral cues such as the source (i.e., information about the reviewer profile) and user-provided product photos. Linguistic features refer to characteristics related to the review textual content such as appropriate amount of information, readability, timeliness, relevancy, and completeness, etc., which can be used to measure the quality of information when one aims to make an argument. Semantic features can include words, topics and semantic relationships between linguistic entities that represent meanings, which can be extracted from the textual content as linguistic tokens, their latent dimensions (e.g., topics) and network structures. Sentiment, which represents the valence of an opinion, i.e., a measure along the spectrum from positive to negative, is an important feature widely used in understanding the emotional effect of text (Pang and Lee, 2008). While the textual contents and their characteristics carry the core message, other components may serve as the peripheral cues in the communication process (Petty and Cacioppo, 1986). For example, reviewer profile, which may include the reviewer's id, gender (if shown), photo, tenure with the website, etc., can be used to measure reviewer credibility or trustworthiness (e.g., Filieri et al., 2015; Park et al., 2014; Sparks et al., 2013). These characteristics have been included in analytics research to explain and predict the impacts of online reviews, particularly, on related variables such as product rating and review helpfulness score, with the latter often seen as a direct measure of the information value of online reviews.

In the past several years, our understanding of the impacts of online product reviews seems to have been improved, to a great extent, due to technical advancements in our capacity to process and analyze new forms of data in increasingly large quantities. Taking textual review data for instance, our analytical methods have evolved from simply counting word frequencies based upon relatively small samples (e.g., Crotts et al., 2009; Stringam and Gerdes, 2010), to using factor analysis to identify underlying dimensions based on word frequency (e.g., Xiang et al., 2015), to more sophisticated machine learning tools such as sentiment analysis and topic modeling to extract deep meanings from large quantities of texts (e.g., Levy et al., 2013; Xiang et al., 2017; Ye et al., 2009). As social media websites continue to evolve and user-generated contents become more diverse and richer in terms of both substance and format, our ability to understand managerial problems, old and new, will likely be defined by technical tools to process, analyze, and interpret these new data.

Visual contents such as photos and videos, perhaps, represent the next boundary for us to explore with cutting-edge technological tools. Photos can provide rich and powerful messages in communicating various aspects and holistic nature of a tourism product (Garrod, 2008; Haywood, 1990; MacKay and Couldwell, 2004). This is presumably true within the context of social media and particularly online product reviews. User-provided photos provide visual cues for showing specific product information from one's actual experience; besides, they also serve the purpose of verifying that the reviewer has physically visited or stayed at a place (e.g., a hotel) and thus engendering trust (Ert et al., 2016). Although there is growing interest in social media analytics research using user-provided photos as data to describe travel patterns (e.g., Girardin et al., 2008; Vu et al., 2015), how these photos contribute to the communicative effect of online reviews has not been empirically examined. For example, compared to review texts and other user-provided cues, how do user-provided photos influence users' perception of the quality of online reviews? In order to gain a deeper understanding of the information value of online reviews in hospitality and tourism, it is important to identify and develop new approaches to process and interpret these user-provided product photos. As such, deep learning techniques recently developed in natural language processing and, especially, computer image processing appears to be an ideal tool for many of the problems related to user-generated contents on the Internet.

2.2. Deep learning for natural language processing and computer image processing

Conventional computer programs are hard-coded by humans with specific instructions on the tasks intended to execute and, as such, they are inflexible and may not be suitable to handle ever changing, complex problems in the real world. Therefore, the concept of machine learning is to give a computer program a large number of "experiences" (i.e., training data) and a basic principle for learning in order to identify patterns, associations, and insights from the data. Machine learning is widely used in pattern recognition, classification and prediction in social media analytics (Fan and Gordon, 2014; Lazer et al., 2009). Conventional machine learning techniques such as regressions, support vector machines, and decision trees, etc., mainly rely on manually designed features (which can be understood as independent variables in regression-based analysis) that can be included in the algorithms in order to produce best performance for a classifier or regression model as in supervised learning. However, these types of feature extraction and engineering not only require plenty of specific domain knowledge but also are limited in terms of their scope of coverage, as we humans cannot include all the possible and necessary attributes needed for a specific learning task.

To solve this problem, neural networks come to the fore with the idea that computer programs may learn features hierarchically by mimicking the way a biological brain solves problems with large clusters of biological neurons (Jain et al., 1996). As it would be difficult for researchers or even domain experts to abstract effective data directly from scratch, which are suitable for a task, neural network programs may rely on multiple hierarchies of transformations. Typically, a neural network can help extract simple features or concepts first through lower layers, and then concepts that are more complicated can be represented based upon these simpler ones. Intuitively, we may also interpret each neuron within a network as doing one nonlinear mapping like logistic