Tourism Management 80 (2020) 104129

Contents lists available at ScienceDirect

Tourism Management

journal homepage: http://www.elsevier.com/locate/tourman

Using deep learning and visual analytics to explore hotel reviews and responses

Yung-Chun Chang[b,c,d], Chih-Hao Ku[a,*], Chien-Hung Chen[b,e]

[a] Monte Ahuja College of Business, Cleveland State University, Cleveland, USA
[b] Graduate Institute of Data Science, Taipei Medical University, Taipei, Taiwan
[c] Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei, Taiwan
[d] Pervasive AI Research Labs, Ministry of Science and Technology, Taiwan
[e] Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Keywords:
Deep learning
Convolutional neural network
Natural language processing
Visual analytics
Hospitality
Tourism

ABSTRACT

This study aims to use computational linguistics, visual analytics, and deep learning techniques to analyze hotel reviews and responses collected on TripAdvisor and to identify response strategies. To this end, we collected and analyzed 113,685 hotel reviews and responses and their semantic and syntactic relations. We are among the first to use visual analytics and deep learning-based natural language processing to empirically identify managerial responses. The empirical results indicate that our proposed multi-feature fusion, convolutional neural network model can make different types of data complement each other, thereby outperforming the comparisons. The visualization results can also be used to improve the performance of the proposed model and provide insights into response strategies, which further shows the theoretical and technical contributions of this study.

1. Introduction

Today, an increasing number of travelers read online consumer reviews (OCRs) to plan their trips and to make purchasing decisions (De Pelsmacker, van Tilburg, & Holthof, 2018; Hernández-Ortega, 2018). This dynamic growth of online reviews has intensified the need for analyzing OCRs because these reviews contain consumer perspectives that may relate to perceived credibility (Casaló, Flavián, Guinalíu, & Ekinci, 2015a), corporate reputation (Baka, 2016), and consumer intentions to book hotels (Ladhari & Michaud, 2015).

From the business perspective, OCRs and persuasive communications are vehicles to build credibility and influence user decisions (Fan & Gordon, 2014; Zhang et al., 2016). Therefore, the hospitality and tourism industry is ideally suited for the application of social media analytics (Fan & Gordon, 2014). Lu and Stepchenkova (2015) investigated 122 journal articles and conference proceedings in the last decade and gained insights into how user-generated content has been applied in the tourism and hospitality industry. Most related studies collect a few hundred to a few hundred thousand guest reviews using analytical methods such as content analysis, text mining, machine learning, regression analysis, econometric modelling, or combinations of these techniques (Xiang et al., 2017). As for data sources, TripAdvisor is the most frequently used because it is considered a 'premier' sampling field and also the largest travel-related review site in the world (Banerjee & Chua, 2016; Pearce & Wu, 2018; Xiang et al., 2017).

In hospitality and tourism, social media analytics is a growing area (Fan & Gordon, 2014) because consumer reviews reflect experiences on services and have been studied to gain a better understanding of research problems (Schuckert et al., 2015). Most investigations have relied mostly on statistical- and survey-based studies or experiments (Lee & Cranage, 2012; Li et al., 2017; Wei et al., 2013a) due to data unavailability (Xie et al., 2016). To date, the intricate relation between online reviews and responses has not been well studied, despite the considerable effort made on studying the effect of volume and valence (De Pelsmacker et al., 2018), travel motivation (Pearce & Wu, 2018), customer satisfaction, opinions and sentiments (Xiang et al., 2015), online reviews and hotel business performance (Xie et al., 2014), and perceived helpfulness of online reviews (Schuckert et al., 2015). This is unfortunate given that online reviews represent customers' opinions, satisfaction, and attitudes toward a hotel, and effective responses are likely to enhance the perceived hotel quality (De Pelsmacker et al., 2018; Torres et al., 2015) and boost business.

* Corresponding author. Department of Information Systems, Cleveland State University, Monte Ahuja College of Business, 2121 Euclid Avenue, Cleveland, OH, 4411-2214, USA.
E-mail addresses: changyc@tmu.edu.tw (Y.-C. Chang), c.ku17@csuohio.edu (C.-H. Ku), m946106007@tmu.edu.tw (C.-H. Chen).

https://doi.org/10.1016/j.tourman.2020.104129
Received 18 September 2019; Received in revised form 15 April 2020; Accepted 17 April 2020
Available online 5 May 2020
0261-5177/

旅游管理80（2020）104129

内容列表可在 ScienceDirect

旅游管理

期刊主页: http://www.elsevier.com/locate/tourman

利用深度学习和视觉分析探索酒店评论和回应

容俊昌b,c,d,智浩谷a,*,陈建鸿b,e

[a] 美国克利夫兰克利夫兰州立大学蒙特·阿胡亚商学院
[b] 台北医科大学数据科学研究所，台湾台北
[c] 台北医科大学医院临床大数据研究中心，台湾台北
[d] 台湾科技部普及人工智能研究实验室
[e] 国立台湾大学网络与多媒体研究所，台湾台北

文章信息

关键词:
深度学习
卷积神经网络 自然语言处理
视觉分析
酒店旅游

摘要

本研究旨在使用计算语言学、视觉分析和深度学习技术来分析TripAdvisor上收集的酒店评论和回复，并确定回复策略。为此，我们收集并分析了113685篇酒店评论和回复，以及它们的语义和句法关系。我们是最早使用视觉分析和基于深度学习的自然语言处理来经验性地确定管理层反应的公司之一。实验结果表明，我们提出的多特征融合卷积神经网络模型可以使不同类型的数据相互补充，从而优于比较。可视化结果还可用于改进所提出模型的性能，并提供对响应策略的见解，这进一步显示了本研究的理论和技术贡献。

1. 简介

如今，越来越多的旅行者阅读在线消费者评论（OCR）来规划旅行并做出购买决定（德·佩尔斯马克、范·蒂尔堡和霍尔索夫，2018；赫纳）�恩德斯·奥尔特加，2018年）。在线评论的这种动态增长增强了分析OCR的需求，因为这些评论包含可能与感知可信度有关的消费者观点（Casalo）�, 弗拉维亚�n、 Guinalíu和Ekinci, 2015a），企业声誉（Baka，2016），以及消费者预订酒店的意愿（Ladhari和Michaud, 2015）。

从商业角度来看，OCR和说服性沟通是建立可信度和影响用户决策的工具（Fan & Gordon, 2014；Zhang等人，2016）。因此，酒店业和旅游业非常适合社交媒体分析的应用（Fan & Gordon, 2014）。Lu和Stepchenkova（2015）调查了过去十年中的122篇期刊文章和会议记录，并深入了解了用户生成的内容如何应用于旅游和酒店业。大多数相关研究使用分析方法，如内容分析、文本挖掘、机器学习、回归分析、计量经济学建模或这些方法的组合，收集几百到几十万条客户评论技术（Xiang等人，2017年）。至于数据来源，TripAdvisor是最常用的，因为它被认为是“首要”抽样领域，也是世界上最大的旅游相关审查网站（Banerjee & Chua, 2016；Pearce & Wu, 2018；Xiang等人，2017）。

在好客性和旅游业中，社交媒体分析是一个不断增长的领域（Fan & Gordon, 2014），因为消费者评论反映了服务的体验，并已被研究以更好地理解研究问题（Schuckert et al., 2015）。由于数据不可用，大多数调查大多依赖于基于统计和调查的研究或实验（Lee & Cranage, 2012；Li等人，2017；Wei等人，2013a）（谢等人，2016）。迄今为止，在线评论和回复之间的复杂关系尚未得到很好的研究，尽管在研究数量和价格的影响（De Pelsmacker等人，2018年）、旅行动机（Pearce & Wu, 2018年）、客户满意度、意见和情绪（Xiang等人，2015年）方面做出了相当大的努力，在线评论和酒店经营业绩（谢等人，2014年），以及在线评论的有用性（Schuckert等人，2015年）。这是不幸的，因为在线评论代表了客户对酒店的意见、满意度和态度，而有效的回应可能会提高感知的酒店质量（De Pelsmacker et al., 2018；Torres et al., 2015）并促进业务。

*通讯作者。美国俄亥俄州克利夫兰欧几里德大道2121号蒙特阿胡亚商学院克利夫兰州立大学信息系统系，邮编4411-2214。
电子邮件地址：changyc@tmu.edu.tw（Y-c.Chang），C。ku17@csuohio.edu（C.-H. Ku），m946106007@tmu.edu.tw （C.-H.陈）。

https://doi.org/10.1016/j.tourman.2020.104129
2019年9月18日收到；于2020年4月15日收到修订版表格；2020年4月17日接受
可于2020年5月5日在线获取
0261-5177/

Online review platforms enable managers to resolve customers' complaints and engage potential customers by responding publicly to online reviews (Lui et al., 2018; Park & Allen, 2013). Despite the potential value of responding to online reviews, little research has investigated response management strategies in the hotel industry (Lui et al., 2018; Schuckert et al., 2015). We investigated the question by collecting online reviews and hotel responses from 43 London hotels. Next, we analyzed current practices such as proactive and non-proactive hotel responses (Li et al., 2017) to positive and negative reviews. We extended our previous study and analyses (Ku et al., 2019) to transform online reviews and hotel responses on the travel website into actionable insights and knowledge. The aim of this study is to use the advanced techniques of deep learning, visual analytics, and natural language processing (NLP) in order to analyze hotel reviews and responses, identify response strategies, and offer strategic recommendations to hotel practitioners. Hospitality practitioners can discover hidden online review patterns and consumer booking behavior, which may in turn offer insights into exploring and implementing drivers to take necessary actions to improve service quality and customer satisfaction.

This paper is structured as follows. Section 2 starts with a brief literature review on online consumer reviews and ratings, hotel responses, NLP and deep learning, and previous research on online reviews, typically on TripAdvisor. Section 3 draws on our initial visual analytics with 43 representative hotels to develop deep learning models. To achieve this, our analytical framework includes web scraping, data preprocessing, visual analytics, and deep learning-based NLP. We then describe our experiments with our proposed model and compare it with existing algorithms. Finally, we discuss the main findings, decision-making implications, limitations, and future research directions.

2. Related work

2.1. OCRs and ratings

Smart tourism technologies such as online tourism applications, online travel agents, social media, and mobile applications play an increasingly important role in travel planning (Huang, Goo, Nam, & Yoo, 2017; Xiang et al., 2015). Travelers often reply on OCRs, which is a form of electronic word-of-mouth (eWOM), to make purchase decisions (Filieri, 2016). Huang et al. (2017) further point out that the travel decision-making process is not sequential but involves four iterative phases: idea formation, information searching, evaluating alternatives, and booking. The overall process aims to reduce uncertainty (Fang, Ye, Kucukusta, & Law, 2016) and potential risks associated with the purchasing (Sparks et al., 2016); or, in other words, tourists usually read online ratings and reviews to plan their trip (Mauri & Minazzi, 2013).

The proliferation of OCRs and ratings are likely to influence consumers' purchase decision (Hudson & Thal, 2013; Mauri & Minazzi, 2013), re-visit intentions (Mauri & Minazzi, 2013; Zhang & Mao, 2012), search behaviors (King et al., 2014; Serra Cantallops & Salvi, 2014), online and product sales (Öğüt & Taş, 2012; Ye et al., 2009, 2011), and even intention to book a hotel (Zhao et al., 2015) as well as their attitudes toward the hotel (Vermeulen & Seegers, 2009). A higher review rating can increase online hotel sales (Öğüt & Taş, 2012) and boost online bookings (Ye et al., 2011). A similar result was found by Noone and McGuire (2013) by examining the relation between online reviews and online hotel booking. The ratings represent customers' satisfaction towards the hotel they stay in (Liu et al., 2015), and with TripAdvisor, customer reviews can include ratings on six aspects: value, location, sleep quality, rooms, cleanliness, and service. Such rating information is valuable because it contains the valence of OCRs. One study shows that a 10% increase in customer ratings can boost hotel room sales up to 5% (Ye et al., 2011). A similar result was found by Öğüt and Taş (2012) who reported that a 1% increase in customer ratings can raise sales per room up to 2.68% and 2.62% in Paris and London, respectively.

The valence of OCRs can be classified into positive and negative forms. Based on social cognition theory (Pan & Chiou, 2011), negative information is perceived to be more influential than positive information (Dickinger, 2011; Xiang et al., 2015). Casaló, Flavián, Guinalíu, and Ekinci (2015b) conducted survey studies with 46 participants to better understand the perceived usefulness of OCRs and suggested that negative online reviews were more useful than positive reviews, typically for high risk-averse travelers. However, positive reviews were likely to influence a consumer's decision making (Xiang et al., 2015) and increased customer revisit intention (Zhang & Mao, 2012). A higher volume of positive reviews may also lead to better hotel ratings and hotel performance (Gu & Ye, 2014; Sparks et al., 2016).

Ratings and rankings have created challenges and opportunities for service providers because improving online ratings may lead to more online bookings (Liu et al., 2015; Ye et al., 2011). Therefore, it would be useful to know how hotel managers respond to positive and negative reviews today.

In addition to OCR ratings and valences, it is equally important to understand different traveler types. Based on the reviews on TripAdvisor.com, the traveler types can be classified into business, couple, family, friend, and solo. Banerjee and Chua (2016) collected 39,747 ratings and found that travelers' rating patterns differed between independent and chain hotels based on travelers' profiles and geographical regions (i.e., the Americas, Asia, Europe, and Africa).

2.2. Hotel responses

The number of responses to online reviews has increased on travel review websites (Schuckert et al., 2018). Contemporary studies (Levy et al., 2013; Melo et al., 2017) point out that hotel managers should establish a digital marketing plan to actively manage online presence. It is therefore extremely important to proactively analyze online reviews and have a clear response strategy.

Gu and Ye (2014), for example, found that unsatisfied customers' future satisfaction increases when they receive responses from hotel managers. Some empirical research reports that manager responses can have a negative impact on customers' purchasing intention (Mauri & Minazzi, 2013), so it is imperative that hotel managers take appropriate corrective actions to mitigate negative outcomes. Timely responses to negative reviews can build a hotel's reputation and customer loyalty, which can elevate the hotel's ratings (Liu et al., 2015).

Park and Allen's (2013) case studies on 34 4- and 5-star hotels revealed neither pattern nor standard in how the selected hotels responded to online reviews. They found the overall response rate ranged from 0 to 64.9% and the response rate of negative reviews ranged from 0 to 45.8%. Liu et al. (2015) used correlation analysis to analyze the response rate of 187 Hong Kong hotels from all classes (from 1- to 5-star). They found that high-class hotels tended to adopt response management, but the response rate showed no significant difference between different classes of hotels. To test the factors that may influence the helpfulness of online reviews, Kwok and Xie (2016) collected 56,284 reviews and 10,797 hotel responses from 1405 hotels on TripAdvisor and tested the data using a linear regression model. They discovered that ratings, the number of words, reviewer's gender, reviewers' experience, visited cities and hotel responses may influence customers' perceived usefulness of online reviews.

Managing online reputation is an effective approach to improve consumer satisfaction (De Pelsmacker et al., 2018; Kim et al., 2015) and it is less expensive than enhancing facilities directly (Schuckert et al., 2018). However, seeking effective approaches to manage eWOM, especially if its negative, is a widely recognized challenge for hospitality management (Sparks et al., 2016). Hotel managers should go beyond the reading of consumer postings, and focus on responding to consumer reviews in a timely, proactive, and consistent manner (Kwok & Xie, 2016). Doing so opens new opportunities to engage consumers and communicate with potential customers. The perceived and proactive responses by the manager can be part of a hotel's customer relationship

在线评论平台使管理者能够通过公开回应在线评论来解决客户投诉并吸引潜在客户（Lui等人，2018年；Park & Allen, 2013年）。尽管回复在线评论具有潜在价值，但很少有研究调查酒店行业的回复管理策略（Lui等人，2018年；Schuckert等人，2015年）。我们通过收集43家伦敦酒店的在线评论和酒店回复来调查这个问题。接下来，我们分析了酒店对正面和负面评论的积极和非积极回应（Li等人，2017年）等当前做法。我们扩展了之前的研究和分析（Ku等人，2019年），将旅游网站上的在线评论和酒店回复转化为可操作的见解和知识。本研究的目的是利用深度学习、视觉分析和自然语言处理（NLP）等先进技术，分析酒店评论和回应，确定回应策略，并向酒店从业者提供策略建议。酒店从业人员可以发现隐藏的在线评论模式和消费者预订行为，这反过来可能为探索和实施驱动因素以采取必要的行动来提高服务质量和客户满意度提供见解。

本文的结构如下。第2节从一个简短的例子开始
关于在线消费者评论和评级、酒店回应、NLP和深度学习的文献综述，以及之前关于在线评论的研究，通常是在TripAdvisor上。第3部分利用我们对43家具有代表性的酒店的初步视觉分析，开发深度学习模型。为了实现这一点，我们的分析框架包括网页抓取、数据预处理、可视化分析和基于深度学习的NLP。然后，我们用我们提出的模型描述了我们的实验，并将其与现有算法进行了比较。最后，我们讨论了主要发现、决策含义、局限性和未来的研究方向。

2. 相关工作

2.1. OCR和评级

在线旅游应用、在线旅行社、社交媒体和移动应用等智能旅游技术在旅游规划中发挥着越来越重要的作用（Huang、Goo、Nam和Yoo, 2017；Xiang等人，2015）。旅行者通常会在OCR上回复，这是电子口碑（eWOM）的一种形式，以做出购买决定（Filieri，2016）。Huang等人（2017）进一步指出，旅行决策过程不是连续的，而是涉及四个迭代阶段：想法形成、信息搜索、评估备选方案和预订。整个过程旨在减少不确定性（Fang、Ye、Kucukusta和Law, 2016年）和与采购相关的潜在风险（Sparks等人，2016年）；或者，换句话说，游客通常通过阅读在线评分和评论来计划他们的旅行（Mauri & Minazzi, 2013）。

OCR和评级的激增可能会影响con-
萨默斯的购买决定（哈德逊和塔尔，2013年；莫里和米纳齐，2013年）、再次访问意向（莫里和米纳齐，2013年；张和毛，2012年）、搜索行为（金等人，2014年；塞拉·坎塔洛普和萨尔维，2014年），

[illegible]
甚至预订酒店的意愿（赵等人，2015年）以及他们对酒店的态度（Vermeulen & Seegers, 2009年）。更高的评论评级可以增加在线酒店销售额（o欧元g）�üt & Tas,, 2012）和boost
在线预订（Ye等人，2011年）。Noone和McGuire（2013）通过研究在线评论和在线酒店预订之间的关系，发现了类似的结果。这些评分代表了客户对所住酒店的满意度（Liu等人，2015年），通过TripAdvisor,客户评论可以包括六个方面的评分：价值、位置、睡眠质量、房间、清洁度和服务。这样的评级信息很有价值，因为它包含OCR的价。一项研究表明
客户评级提高10%可以将酒店客房销售额提高5%（Ye等人，2011年）。O€ g发现了类似的结果�吕特和塔斯（2012）世卫组织
报道称，在巴黎和伦敦，客户评级提高1%可以使每间客房的销售额分别提高2.68%和2.62%。

OCR的价态可分为正价态和负价态

形式。根据社会认知理论（Pan & Chiou, 2011），人们认为负面信息比正面信息更有影响力（Dickinger，2011；Xiang等人，2015）。卡萨洛�, 弗拉维亚�n、 Guinalíu和Ekinci（2015b）对46名参与者进行了调查研究，以更好地理解OCR的感知有用性，并建议负面在线评论比正面评论更有用，通常对高风险厌恶的旅行者来说。然而，正面评价可能会影响消费者的决策（Xiang等人，2015年），并增加客户的重访意愿（Zhang & Mao, 2012年）。更高数量的正面评论也可能导致更好的酒店评级和酒店绩效（Gu & Ye, 2014；Sparks等人，2016）。

评级和排名为服务提供商带来了挑战和机遇，因为提高在线评级可能会带来更多的在线预订（Liu等人，2015年；Ye等人，2011年）。因此，了解酒店管理者对今天正面和负面评论的反应会很有用。

除了OCR评级和价格，了解不同的旅行者类型同样重要。基于TripAdvisor上的评论。com，旅行者类型可以分为商务、夫妻、家庭、朋友和独行。Banerjee and Chua（2016）收集了39747个等级，发现独立酒店和连锁酒店的旅客等级模式根据旅客的个人资料和地理区域（即美洲、亚洲、欧洲和非洲）有所不同。

2.2. 酒店回应

旅游评论网站上对在线评论的回复数量有所增加（Schuckert等人，2018年）。当代研究（Levy等人，2013年；Melo等人，2017年）指出，酒店经理应制定数字营销计划，积极管理在线服务。因此，积极分析在线评论并制定明确的应对策略至关重要。

例如，Gu和Ye（2014）发现，当不满意的顾客收到酒店经理的回复时，他们未来的满意度会增加。一些实证研究报告称，管理者的反应可能会对客户的购买意愿产生负面影响（Mauri & Minazzi, 2013），因此酒店管理者必须采取适当的纠正措施来缓解负面后果。及时回应负面评论可以建立酒店的声誉和客户忠诚度，从而提高酒店的评级（Liu等人，2015年）。

Park和Allen（2013）对34家四星级和五星级酒店的案例研究显示，所选酒店对在线评论的反应既没有模式，也没有标准。他们发现总体回复率在0到64.9%之间，负面评论的回复率在0到45.8%之间。刘等人（2015）采用相关分析法分析了187年香港饭店各阶层（1～5星级）的反应率。他们发现，高级酒店倾向于采用响应管理，但不同级别酒店的响应率没有显著差异。为了测试可能影响在线评论有用性的因素，郭和谢（2016）在TripAdvisor上收集了1405家酒店的56284条评论和10797条酒店回复，并使用线性回归模型对数据进行了测试。他们发现，评分、字数、评论者的性别、评论者的经历、访问过的城市和酒店的回复可能会影响顾客对在线评论的感知有用性。

管理在线声誉是提高声誉的有效途径
消费者满意度（De Pelsmacker et al., 2018；Kim et al., 2015），而且它比直接改善设施成本更低（Schuckert et al., 2018）。然而，寻求有效的方法来管理伊渥姆，尤其是在其负面影响的情况下，是酒店管理的一个公认挑战（Sparks等人，2016）。酒店经理应超越阅读消费者帖子的范畴，专注于及时、主动、一致地回应消费者评论（郭&谢，2016）。这样做为吸引消费者和与潜在客户沟通提供了新的机会。经理所感知到的、积极主动的回应可能是酒店客户关系的一部分

management strategy (Liu et al., 2015). For example, hotel managers may invite guests back in the future and try to win back customers who post negative reviews. With the increase in online reviews, hospitality businesses can shift from the role of passive listening to one of proactive engagement by managing responses (Gu & Ye, 2014). However, little is known about the current practice by hotel managers and their strategies to respond to positive and negative reviews. In this study, we not only identify proactive and non-proactive responses, but also provide recommended strategies to hospitality practitioners.

2.3. NLP and deep learning

With a plethora of digital data increasingly becoming available for analysis, machine learning and NLP techniques have received much interest in many research fields (Xiang et al., 2017). NLP is a subfield of artificial intelligence and employs computational models to process natural language by learning cognitive activities of human brains (Cambria & White, 2014). Popular NLP tasks include information extraction, information retrieval, text summarization, question answering, topic modelling, and more recently, sentiment analysis and opinion mining. Today most NLP techniques rely on keywords, word co-occurrence, and frequencies of words, and syntactic information of text (Cambria & White, 2014). This poses a major challenge to analyze user-generated content because semantic relationships between words are often ignored (Zhang et al., 2015).

To extract sentiments or opinions of online reviews, machine learning techniques can be used. Opinion extraction identifies opinion holders from the given text, whereas sentiment analysis assigns a polarity such as positive, neutral, and negative to the extracted subjectivity. Extracting user opinions requires several NLP steps such as tokenization, word segmentation, part-of-speech (POS) tagging, stemming, and stop word removal. NLP toolkits, such as NLTK (https://www.nltk.org), OpenNLP (https://opennlp.apache.org) and Stanford's CoreNLP (https://stanfordnlp.github.io/CoreNLP/) are widely used in research projects (Sun et al., 2017) for text preprocessing. 'Furthermore, machine learning models, such as Naïve Bayes, maximum entropy, and support vector machine (SVM), are often trained to determine the polarities of online reviews (Dragoni, Federici, & Rexha, 2018). For instance, Parkhe and Biswas (2016) and Manek et al. (2017) utilized Naïve Bayes and SVM to learn a classifier for sentiment analysis of movie reviews. However, the requirement of annotated training data and domain-specific lexicons creates a major challenge for cross-domain and cross-lingual situations (Sun et al., 2017).

More recently, deep learning algorithms have shown promise with high accuracy, such as convolutional neural networks (CNN) and recurrent neural networks (RNN) that are both machine learning methods based on learning data from multiple deep layers of modules. In tourism, for instance, Kim et al. (2017) applied a deep learning approach based on the Stanford sentiment analysis to analyze the sentiments of 19,835 online reviews of Paris from www.virtualtourist.com. They noticed that different tasks perform better with higher scores when suitable algorithms were applied: CNN, for example, is good at extracting significant n-gram features to generate "informative latent semantic representation" in NLP classification tasks (Poria et al., 2016; Young et al., 2017), while RNN is effective at processing data in a sequential format. The most important strength of RNN is that it can integrate previous information into the current neural state. Therefore, in order to prevent loading too much previous information, RNN is usually chosen to deal with tasks with short texts (Lee & Dernoncourt, 2016). For example, CNN has been used for sarcasm detection based on sentiment analysis in Spanish (Chaturvedi, Cambria, & Vilares, 2016), while Al-Smadi, Qawasmeh, Al-Ayyoub, Jararweh, and Gupta (2018) used RNN, SVM, and NLP resources to conduct aspect-based opinion mining on Arabic hotel reviews, and their approach outperformed other machine learning methods.

The advantage of applying deep learning to NLP is that it is independent from expert knowledge and lexical resources (Rojas-Barahona, 2016). The Deep neural network architecture of NLP consists of (1) an input layer, (2) multiple hidden layers, and (3) an output layer (Huang, Li, Yu, Deng, & Gong, 2013). Since the input layer requires numerical data, text data firstly need to be transformed to vectors for representation. In practice, a pre-trained word embedding (*input*) such as Word2-Vec by Tomáš Mikolov at Google (Mikolov et al., 2013) and GloVe (Pennington et al., 2014) by Stanford University can be used to represent words encoded as dense numerical vectors in an n-dimensional space (Rojas-Barahona, 2016). This word embedding process also captures the syntactic and semantic information from a text (Chen, Xu, He, & Wang, 2017). For example, Zhang et al. (2015) conducted a sentiment classification based on Word2Vec and the SVM^{perf} package with more than 100,000 Chinese comments on clothing products; the experiment results reached 90% classification accuracy. The *hidden layers* were then constructed to exploit the complex compositional nonlinear functions through higher and lower layers which are composed of a great number of neurons. Each neuron receives inputs $x_1, x_2 ...,$ and is multiplied by the associated weights, and then, the activation function a (Eq. (1)) combines the weighted input data and bias to generate a single output. According to the data's property, there are different activation functions that researchers could use: the *Sigmoid* function (Eq. (2)), *ReLU* function (Eq. (3)), and Tanh function (Eq. (4)) range from 0 to 1, 0 to a, and -1 to 1, respectively. Acquired from the output of the last hidden layer z, the output layer can predict the probabilities of each class using the *softmax* function. As a result, the highest probability class would be the predicted class (Huang et al., 2013).

$$o = s(a) = s\left(\sum_i w_i x_i + b\right) \tag{1}$$

$$sigmoid\ \ (a) = \frac{1}{1+e^{-a}} \tag{2}$$

$$ReLU\ \ (a) = \max(0, a) \tag{3}$$

$$\tanh(a) = \frac{e^{2a}-1}{e^{2a}+1} \tag{4}$$

$$softmax\ (z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}} \tag{5}$$

CNN has been used in many successful NLP tasks (Do, Prasad, Maag, & Alsadoon, 2019). The typical architecture of CNN is assembled with four layers: (1) an input layer, which represents numerical word vectors with *n* vector dimensions and *m* length of a sentence; (2) a convolutional layer, which uses filters to generate new features with an activation function; (3) a max-pooling layer, which selects the maximum value in response to a filter size; and (4) an output layer, which yields the most probable class under a fully-connected layer using the *softmax* function. In addition, a drop-out layer could be added among the CNN processes to avoid over-fitting.

2.4. Previous tourism and hospitality research on TripAdvisor and deep learning

TripAdvisor has become the largest travel-related review platform in the world and is a major data source for researchers to conduct social media analytics in hospitality and tourism (Xiang et al., 2017). Since the goal of this study is to analyze hotel reviews and responses on TripAdvisor, we conducted a search for relevant research studying online reviews and responses. We used 'TripAdvisor' as a keyword to search three leading tourism and hospitality journals with an A^{+} ranking:

管理战略（Liu等人，2015年）。例如，酒店经理
将来可能会邀请客人回来，并试图赢回那些发表负面评论的客户。随着在线评论的增加，酒店业可以从被动倾听的角色转变为通过管理回应主动参与的角色（Gu & Ye, 2014）。然而，人们对酒店管理者目前的做法及其应对正面和负面评论的策略知之甚少。在这项研究中，我们不仅确定了主动和非主动的反应，而且还为酒店从业人员提供了建议的策略。

2.3. NLP与深度学习

随着越来越多的数字数据可用于分析，机器学习和NLP技术在许多研究领域受到了广泛关注（Xiang等人，2017）。NLP是人工智能的一个子领域，通过学习人脑的认知活动，利用计算模型来处理自然语言（Cambria & White, 2014）。流行的NLP任务包括信息提取、信息检索、文本摘要、问答、主题建模，以及最近的情感分析和观点挖掘。如今，大多数NLP技术依赖于关键词、单词共现、单词频率以及文本的句法信息（Cambria & White, 2014）。这对分析用户生成的内容构成了重大挑战，因为单词之间的语义关系往往被忽略（Zhang等人，2015）。

为了从在线评论中提取情感或观点，可以使用机器学习技术。意见提取识别给定文本中的意见持有者，而情绪分析为提取的主观性指定了一个极性，如积极、中立和消极。提取用户意见需要几个NLP步骤，如标记化、分词、词性标记、词干分析和停止词删除。NLP工具包，如NLTK (https://www.

nltk。org），OpenNLP (https://opennlp.apache.org)斯坦福大学的Cor- eNLP (https://stanfordnlp.github.io/CoreNLP/)广泛应用于文本预处理的研究项目（Sun等人，2017年）此外
机器学习模型，如朴素贝叶斯、最大熵和支持向量机（SVM），通常通过训练来确定目标-
在线评论的多样性（Dragoni、Federici和Rexha, 2018）。对于

例如，Parkhe and Biswas（2016）和Manek et al.（2017）利用朴素贝叶斯和支持向量机学习用于电影情感分析的分类器
评论。然而，对带注释的训练数据和特定领域词汇的需求给跨领域和跨领域的学习带来了重大挑战
跨语言情况（Sun等人，2017年）。

最近，深度学习算法显示出高精度的前景，例如卷积神经网络（CNN）和递归神经网络（RNN），这两种机器学习方法都是基于多个深层模块的学习数据。例如，在旅游业方面，Kim等人（2017年）采用了基于斯坦福情感分析的深度学习方法，分析了www.virtualtourist网站上19835篇巴黎在线评论的情感。通用域名格式。他们注意到，当应用合适的算法时，不同的任务表现更好，分数更高：例如，CNN擅长提取显著的n-gram特征，以在NLP分类任务中生成"信息性潜在语义表征"（Poria等人，2016年；Young等人，2017年），而RNN在以顺序格式处理数据方面是有效的。RNN最重要的优点是它可以将以前的信息整合到当前的神经状态中。因此，为了防止加载太多以前的信息，通常选择RNN来处理短文本任务（Lee & Dernoncourt, 2016）。例如，CNN已被用于基于西班牙语情绪分析的讽刺检测（Chaturvedi、Cambria和Vilares, 2016），而al Smadi、Qawasmeh、al Ayyoub、Jararweh和Gupta（2018）使用RNN、SVM和NLP资源对阿拉伯语酒店评论进行基于方面的意见挖掘，其方法优于其他方法

机器学习方法。

将深度学习应用于NLP的优势在于它独立于专家知识和词汇资源（Rojas Barahona，2016）。NLP的深层神经网络结构包括（1）输入层，（2）多个隐藏层，（3）输出层（Huang，Li, Yu, Deng, & Gong, 2013）。由于输入层需要数字数据，因此首先需要将文本数据转换为向量进行表示。在实践中，预先训练好的单词嵌入（输入），比如Toma的Word2-Vec□□谷歌的s Mikolov（Mikolov等人，2013年）和斯坦福大学的GloVe（Pennington等人，2014年）可以用来表示n维空间中编码为密集数字向量的单词（Rojas Barahona，2016年）。这个单词嵌入过程还可以从文本中获取句法和语义信息（Chen、Xu、He和Wang, 2017). 例如，Zhang等人（2015）基于Word2Vec和SVM^{perf}软件包进行了情绪分类，对服装产品发表了超过10万条中文评论；实验结果
分类准确率达到90%。然后构造隐藏层，通过由大量神经元组成的高层和下层来利用复杂的组合非线性函数。每个神经元接收输入$x_1, x_2 ...,$并乘以相关权重，然后，激活函数a（等式（1））将加权输入数据和偏差结合起来，生成单个输出。根据数据的性质，研究人员可以使用不同的激活函数：Sigmoid函数（等式（2））、ReLU函数（等式（3））和Tanh函数（等式（4））的范围分别为0到1、0到a和1到1。从最后一个隐藏层z的输出中获取，输出层可以使用softmax函数预测每个类别的概率。因此，最高概率等级将是预测等级（Huang等人，2013）。

o ¼ sðaÞ ¼ sð $\sum_i$ $w_i x_i$ þ bÞ (1)

乙状结肠 ðaÞ¼ 1 / 1 þ e a (2)

ReLU ðaÞ ¼ maxð0; aÞ (3)

tanh ðaÞ¼ e^{2a} 1 / e^{2a} þ 1 (4)

softmax ðzÞ$_i$ ¼ e^{z_i} / P K k¼1 e^{z_k} (5)

CNN已被用于许多成功的NLP任务（Do、Prasad、Maag和Alsadoon, 2019年）。CNN的典型结构由四层组成：（1）输入层，它表示n个向量维度和m个句子长度的数字词向量；（2）卷积层，使用过滤器生成具有激活功能的新特征；（3）最大池层，根据过滤器大小选择最大值；以及（4）输出层，其使用softmax函数在全连接层下产生最可能的类。此外，可以在CNN过程中添加一个退出层，以避免过度拟合。

2.4. 之前关于TripAdvisor和深度学习的旅游和酒店研究

TripAdvisor已成为世界上最大的旅游相关审查平台，是研究人员在酒店业和旅游业中进行社交媒体分析的主要数据源（Xiang等人，2017年）。由于本研究的目的是分析Tri- pAdvisor上的酒店评论和回复，我们搜索了研究在线评论和回复的相关研究。我们使用"TripAdvisor"作为关键字进行搜索
三本排名为þ的领先旅游和酒店杂志:

Tourism Management, *Journal of Travel Research*, and *Annals of Tourism* based on the Australian Business Deans Council (ABDC) journal quality list.[1] We searched each publisher's database directly and included forthcoming papers. The final search was completed on October 21st, 2019. As shown in Fig. 1, *Tourism Management* had 133 articles using the keyword 'TripAdvisor, followed by the *Journal of Travel Research* (68 articles) and *Annals of Tourism Research* (46). Among the 133 papers in *Tourism Management*, 7 articles were either in-press or published in 2020, and so are not listed in the following figure since 2020 was not finished at the time of writing this paper. However, we still included these articles into our examination.

Next, we manually examined the data collection, data analysis, and research methods from the 247 collected articles. We excluded studies and research notes which mentioned 'TripAdvisor' only and did not use, collect, or analyze online review data. After filtering irrelevant articles, a total of 66 studies were kept for further analyses. The key results are presented in Appendices A and B, with the former listing studies related to hotel reviews on TripAdvisor and the latter listing studies related to online reviews such as attractions, restaurants, and flight reviews from single or multiple data sources. Factors such as reviewer profile, travel types, and aspect ratings are available on TripAdvisor, but not on all online platforms, so these factors are not included in Appendix B.

Among the 66 studies, only five articles (Zhang et al. (2020), Lui et al. (2018), Li et al. (2017), Sparks et al. (2016), and Baka (2016)) studied hotel reviews and responses together. Further, most studies collected and analyzed numerical data such as hotel ratings, prices, number of words, and hotel stars. Nuance factors such as reviewer profile (e.g., reviewer activities and contributions), traveler types (e.g., solo, business, family, couple, and friend), and aspect ratings (e.g., value, location, sleep quality, rooms, cleanliness, and service) have not been well-studied.

Visualizations can bridge the gap between human analysts and machines, typically for analyzing big data. From the collected articles, we observed that most studies used static bar and line charts (Wang et al., 2020) to present research results. An interactive visualization can be used to explore data and identify hidden patterns, with the output of interactive visualizations being used to fine-tune the performance of machine learning. Rose and Willis (2018), for instance, collected 9030 Twitter images and visualized the tweeted images related to smart cities. They explored the patterns and colors of images and tried to understand the features of smart cities. For tourism research, Kirilenko et al. (2019) used geographical analysis to discover the distributions of tourist reviews in Florida and to gain a better understanding of tourist clustering.

For data analyses and research methods, most studies have used statistical (Li et al., 2017), logistic (Gao, Li, Liu, & Fang, 2018), regression (Zhang & Cole, 2016), and econometric models (Yang & Mao, 2019). Other popular methods include a survey (Kim & Stepchenkova, 2015; Sparks et al., 2016), content (Su & Teng, 2018), and geographical analysis (Kirilenko et al., 2019). More recently, sentiment analysis (Kirilenko et al., 2018), text mining (Zhang et al., 2020), and NLP (Stamolampros et al., 2019) have become popular research methods. Kirilenko et al. (2018) have investigated sentiment analysis used in tourism research, which can be classified into lexicon-based and machine learning methods, e.g., Naïve Bayes, SVM, and k-nearest neighbors (k-NN). Most studies applied NLP and text mining techniques to conduct topic modelling, concept analysis, and sentiment analysis. This is understandable because software packages such as Leximancer[2] can be applied directly, and sentiment lexicons such as SocialSent[3] can be downloaded. More advanced text mining research, on the other hand, requires a deeper knowledge of programming, model construction, and performance experiments.

We also extended our reviews to the literature tables in the collected articles, e.g., 34 articles from Hu, Teichert, et al. (2019) and Hu, Zhang, et al. (2019), p. 20 articles from Yang et al. (2018), p. 12 articles by Li et al. (2017), and 22 articles by Xiang et al. (2017) and Guo, Barnes, and Jia (2017), and found similar results. For example, Hu, Teichert, et al. (2019) and Hu, Zhang, et al. (2019) reviewed 34 related to customer loyalty, satisfaction, and re-visit intentions and found only three studies using text mining analysis. The rest of the studies used surveys, interviews, and statistical analyses, which is consistent with the findings of Vu et al. (2019) as well as our own. Both Hu, Teichert, et al. (2019) and Hu, Zhang, et al. (2019) and Vu et al. (2019) further recommended that other sophisticated text mining techniques can be used to cover a broader range of hotel attributes.

Finally, we searched the articles related to 'deep learning' from *Tourism Management,* the *Journal of Travel Research,* and *Annals of Tourism Research* and 11 articles were returned. Among these 11 articles, 2 are review papers and 7 are irrelevant to 'deep learning' but just mentioned 'deep learning.' Both deep learning-associated articles were published in 2019: Law et al. (2019) used a deep learning approach to forecast tourism demand, while Zhang et al. (2019) analyzed photos to discover tourist behavior and perceptions using deep learning techniques. However, none of these studies used NLP techniques. Further search on Google Scholar was conducted using keywords such as online reviews, TripAdvisor, machine learning, and deep learning. We found that machine learning and deep learning techniques were mostly used in sentiment analysis and opinion mining (Sun et al., 2017; Valdivia et al., 2017, 2019), computer vision and image processing (Giglio, Pantano, Bilotta, & Melewar, 2020, 2019; Ma et al., 2018), and medical text analysis (Dreisbach, Koleck, Bourne, & Bakken, 2019; Wu et al., 2020). In other words, the use of deep learning-based NLP and visualization for business strategies and decision making is still in its infancy. Alaei, Becken, and Stantic (2019), Cheng, Fu, Sun, Bilgihan, and Okumus (2019) and Li et al. (2018) suggested that using more advanced techniques such as deep learning can help tourism research gain deeper insights from different aspects of tourism data.

Few studies, therefore, have shed light on deep learning and NLP techniques on tourism and hospitality research, especially on hotel response strategies. Consequently, this paper employs the three major techniques of deep learning, NLP, and visual analytics with nuance factors such as reviewer profile, aspect ratings, sentiment, and temporal factors to offer strategic insights to hotel practitioners.

3. Methodology

What is missing thus far in the existing literature on the application of social media analytics, particularly for hospitality management, is an integrated framework to gain insights into customer reviews and hotel responses. Thus, we developed an analytical framework (see Fig. 2) that includes the components of data selection, web scraping, data preprocessing, visual analytics, and deep learning-based NLP, which correspond to each section and subsection in this study.

3.1. Data selection

According to the Mastercard Global Destination Cities Index 2018 (Julia, 2018), London has been one of the most popular cities for international travelers among the top 162 destination cities. Based on the visitor volume and average spending in 2017, London was selected in this study because it is an English-speaking city and we aim to analyze English reviews. Table 1 shows that London had 19.83 million visitors in 2017 and the average length of stay was 5.8 nights with an average spending of $153 per day.

There were approximately 1088 hotels in London based on the TripAdvisor search results for 2018. According to Brand Finance, Hilton was the most valuable hotel brand ($6330 million) in 2018 (Richard,

[1] ABDC Journal Quality List, https://abdc.edu.au/research/abdc-journal-list/.

[2] Leximancer, https://info.leximancer.com/.

[3] SocialSent, https://nlp.stanford.edu/projects/socialsent/.

《旅游管理》、《旅游研究杂志》和《旅游年鉴》
基于澳大利亚商业院长理事会（ABDC）期刊质量列表。[1]我们直接搜索了每个出版商的数据库，包括即将发表的论文。最终搜索于10月21日完成，

2019. 如图1所示，旅游管理有133篇使用关键词“TripAdvisor”的文章，其次是《旅游研究杂志》（68篇）和《旅游研究年鉴》（46篇）。在133篇关于旅游管理的论文中，有7篇是在2020年出版或发表的，因此未在下图中列出，因为在撰写本文时，2020年尚未完成。然而，我们仍然将这些文章纳入了我们的考试。

接下来，我们手动检查了247篇收集的文章中的数据收集、数据分析和研究方法。我们排除了只提到“Tripadvisor”的研究和研究笔记，没有使用、收集或分析在线评论数据。在筛选了不相关的文章后，共保留了66 项研究以供进一步分析。主要结果见附录A和附录B,前者与TripAdvisor上的酒店评论相关，后者与在线评论相关，如景点、餐厅和来自单个或多个数据源的航班评论。TripAdvisor上提供了评论者简介、旅行类型和方面评级等因素，但并非所有在线平台都提供了这些因素，因此这些因素不包括在附录B中。

在66项研究中，只有五篇文章（Zhang et al.（2020）、Lui et al.（2018）、Li et al.（2017）、Sparks et al.（2016）和Baka（2016））一起研究了酒店评论和回应。此外，大多数研究都收集和分析了数字数据，如酒店评级、价格、字数和酒店星级。细致入微的因素，如评论者简介（例如，评论者活动和贡献）、旅行者类型（例如，单人、商务、家庭、夫妇和朋友）和方面评级（例如，价值、位置、睡眠质量、房间、清洁度和服务）没有得到很好的研究。

可视化可以弥合人类分析师和机器之间的鸿沟，通常用于分析大数据。从收集的文章中，我们观察到大多数研究使用静态条形图和折线图（Wang等人，2020年）来展示研究结果。交互式可视化可用于探索数据和识别隐藏模式，交互式可视化的输出可用于微调机器学习的性能。例如，Rose和Willis（2018）收集了9030张推特图像，并可视化了与智能城市相关的推特图像。他们探索了图像的模式和颜色，并试图了解智慧城市的特征。在旅游研究方面，KirilenKo等人（2019年）利用地理分析发现了佛罗里达州旅游评论的分布情况，并更好地理解了旅游聚集。对于数据分析和研究方法，大多数研究都使用了统计学（Li等人，2017年）、logistic（高、李、刘和方，2018年）、回归（张和科尔，2016年）和计量经济学模型（杨和毛，2019年）。其他流行的方法包括调查（Kim & Stepchenkova, 2015；Sparks等人，2016）、内容（Su & Teng, 2018）和地理分析（Kirilenko等人，2019）。最近，情绪分析（Kirilenko et al., 2018）、文本挖掘（Zhang et al., 2020）和NLP（Stamolampros et al., 2019）已成为流行的研究方法。Kirilenko等人（2018年）研究了旅游研究中使用的情绪分析，可分为基于词典的方法和机器学习方法，如朴素贝叶斯、支持向量机和k近邻（k-NN）。大多数研究应用NLP和文本挖掘技术来进行主题建模、概念分析和情绪分析。这是可以理解的，因为Leximancer[2]等软件包可以直接应用，可以下载SocialSent[3]等情感词汇。另一方面，更高级的文本挖掘研究需要对编程、模型构建和应用有更深入的了解性能实验。

我们还将我们的评论扩展到了收集的文章中的文献表，例如胡、泰切特等（2019）和胡、张等（2019）的34篇文章，杨等（2018）的20篇文章，李等（2017）的12篇文章，以及项等（2017）和郭、巴恩斯和贾（2017）的22篇文章，并发现了类似的结果。例如，Hu, Teichert等人（2019年）和Hu, Zhang等人（2019年）审查了34项与客户忠诚度、满意度和重访意向相关的研究，发现只有三项研究使用了文本挖掘分析。其余的研究使用了调查、访谈和统计分析，这与Vu等人（2019年）以及我们自己的研究结果一致。Hu, Teichert等人（2019年）和Hu, Zhang等人（2019年）以及Vu等人（2019年）进一步建议，可以使用其他复杂的文本挖掘技术来覆盖更广泛的酒店属性。

最后，我们从《旅游管理》、《旅游研究杂志》和《旅游研究年鉴》中搜索了与“深度学习”相关的文章，并返回了11篇文章。在这11篇文章中，有2篇是评论文章，7篇与“深度学习”无关，只是提到了“深度学习”这两篇与深度学习相关的文章发表于2019年：Law等人（2019年）使用深度学习方法预测旅游需求，而Zhang等人（2019年）使用深度学习技术分析照片，以发现游客的行为和感知。然而，这些研究都没有使用NLP技术。使用在线评论、TripAdvisor、机器学习和深度学习等关键词对Google Scholar进行了进一步搜索。我们发现机器学习和深度学习技术主要用于情绪分析和观点挖掘（Sun等人，2017；Valdivia等人，2017, 2019）、计算机视觉和图像处理（Giglio、Pantano、Bilotta和Melewar, 2020, 2019；Ma等人，2018），和医学文本分析（Dreisbach、Koleck、Bourne和Bakken, 2019年；Wu等人，2020年）。换句话说，基于深度学习的NLP和可视化在商业战略和决策中的应用仍处于初级阶段。阿莱，
Becken和Stantic（2019）、Cheng、Fu、Sun、Bilgihan和Okumus（2019）以及Li等人（2018）认为，使用更先进的技术，如深度学习，可以帮助旅游研究从旅游数据的不同方面获得更深入的见解。

因此，很少有研究阐明旅游和酒店研究中的深度学习和NLP技术，尤其是酒店应对策略。因此，本文采用了深度学习、NLP和视觉分析三种主要技术，并结合细微差别因素，如评论员简介、方面评分、情绪和时间因素，为酒店从业者提供战略见解。

3. 方法论

到目前为止，关于社交媒体分析应用的现有文献，尤其是对于酒店管理，缺少的是一个综合框架，可以深入了解客户评论和酒店回应。因此，我们开发了一个分析框架（见图2），其中包括数据选择、网页抓取、数据预处理、视觉分析和基于深度学习的NLP的组件，这些组件对应于本研究中的每个部分和小节。

3.1. 数据选择

根据2018年万事达卡全球目的地城市指数（Julia，2018），伦敦已成为全球162个目的地城市中最受国际游客欢迎的城市之一。根据2017年的游客量和平均消费，本研究选择了伦敦，因为它是一个讲英语的城市，我们旨在分析英语评论。表1显示，2017年伦敦有1983万游客，平均住宿时间为5.8晚，平均每天消费153美元。

根据2018. 的Tri- pAdvisor搜索结果，伦敦约有1088. 酒店。根据brand Finance的数据，希尔顿是2018. 最有价值的酒店品牌（63. 3. 美元）（Richard，

[1] ABDC期刊质量列表， https://abdc.edu.au/research/abdc-journal-list/
.

[2] 莱克西曼瑟， https://info.leximancer.com/.

[3] SocialSent, https://nlp.stanford.edu/projects/socialsent/.

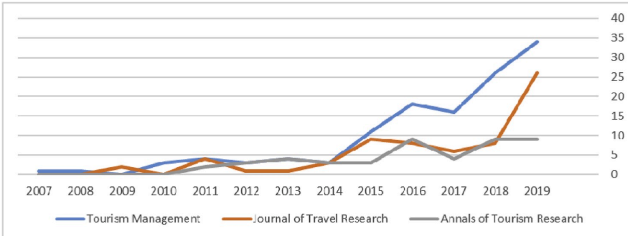

Fig. 1. The number of articles containing the keyword 'TripAdvisor' from three leading tourism and hospitality journals.

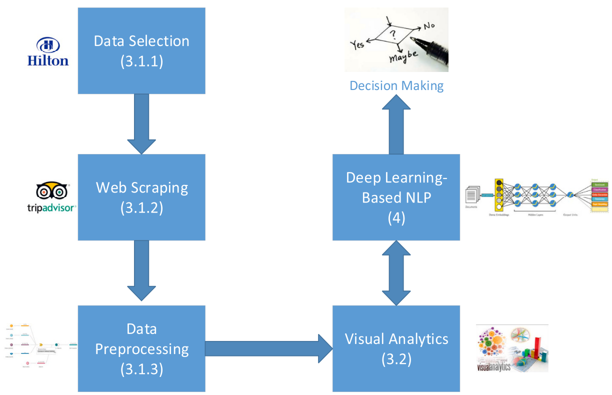

Fig. 2. The general analytical framework for hotel review & response analysis.

Table 1
Top 6 destination cities around the world.

	2017 Overnight visitors	2018 Growth forecast	Average length of stay	2017 Overnight visitor spending (USD)	Average spending per day (USD)
Bangkok	20.05 million	9.6%	4.7 nights	$16.36 billion	$173
London	**19.83 million**	**3.0%**	**5.8 nights**	**$17.45 billion**	**$153**
Paris	17.44 million	2.9%	2.5 nights	$13.05 billion	$301
Dubai	15.79 million	5.5%	3.5 nights	$29.70 billion	$537
Singapore	13.91 million	4.0%	4.3 nights	$17.02 billion	$286
New York	13.13 million	4.1%	8.3 nights	$16.10 billion	$147

2018), and for this reason, Hilton-associated hotels were selected. We first explored hotel information on Hilton Destination Travel Guides (Hilton Travel, 2018) (see Fig. 3). Next, we selected 43 Hiltons hotels operating in 2017 within 25 miles of London, based on the Hilton official website (Hilton London, 2017).

3.2. Data collection & web scraping

TripAdvisor is one of the fastest-growing travel sites, offering appropriately 702 million reviews with an average of 490 million monthly unique visitors (TripAdvisor, Inc. Earnings Press Release, 2018). Although we conducted initial data analyses with 40 hotels in our previous study (Ku et al., 2019), this study is different because we include the complete dataset with 43 hotels with new data collections, enhanced algorithms, and we conduct comprehensive visual analytics. For this purpose, we adopted C# with the Selenium package[4] to develop a crawler to retrieve or scrape hotel reviews and responses from the TripAdvisor web pages. Selenium is a browser automation tool, which supports multiple languages such as Python, Java, and C#. The collected data includes three important aspects: reviewer profile, hotel information, and hotel reviews and responses.

Regarding the reviewer profile, we gathered username, hometown and TripCollective[5] information related to the reviewer's activities. A reviewer receives points from TripCollective (a contributor program on TripAdvisor) based on types of activities such as writing a review (100 points) and posting photos (30 points), with the level of contribution

[4] Selenium.WebDriver, https://www.nuget.org/packages/Selenium.WebDriver.
[5] TripCollective, https://www.tripadvisor.com/TripCollective.

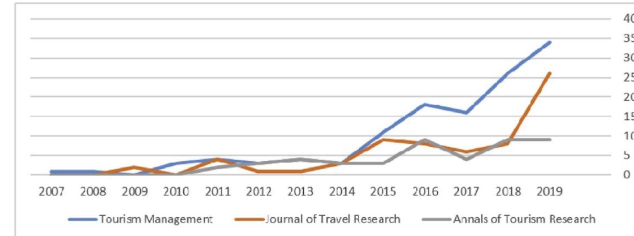

图1. 三大主要旅游和酒店杂志中包含关键词"TripAdvisor"的文章数量。

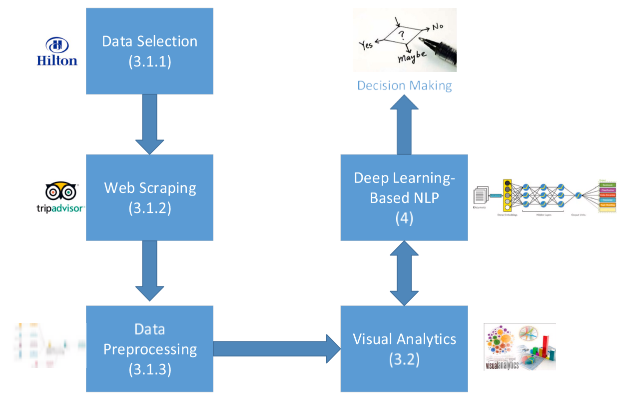

图2. 酒店审查和响应分析的一般分析框架。

表1
全球六大旅游目的地城市。

	2017 在夜间 参观者	2018 增长 预测	平均值 长度 停留	2017 在夜间 来访者 支出 (USD)	平均值 支出 每天 (USD)
曼谷	20.05 百万	9.6%	4.7晚	$16.36 十亿	$173
伦敦	**19.83 百万**	**3.0%**	**5.8 夜晚**	**$17.45 十亿**	**$153**
巴黎	17.44 百万	2.9%	2.5晚	$13.05 十亿	$301
迪拜	15.79 百万	5.5%	3.5晚	$29.70 十亿	$537
新加坡	13.91 百万	4.0%	4.3晚	$17.02 十亿	$286
纽约	13.13 百万	4.1%	8.3晚	$16.10 十亿	$147

2018年），因此选择了希尔顿联合酒店。我们首先探讨了希尔顿目的地旅游指南（希尔顿旅游，2018年）中的酒店信息（见图3）。接下来，我们选择了43家希尔顿酒店
根据希尔顿官方数据，2017年在伦敦25英里范围内运营
网站（伦敦希尔顿酒店，2017年）。

3.2. 数据收集和网络抓取

TripAdvisor是增长最快的旅游网站之一，提供约7.02亿条评论，平均每月有4.9亿独立访客（TripAdvisor，Inc.收益新闻稿，2018年）。虽然我们在之前的研究（Ku等人，2019年）中对40家酒店进行了初步数据分析，但这项研究有所不同，因为我们包含了43家酒店的完整数据集以及新的数据收集，
增强算法，我们进行全面的视觉分析。为此，我们采用了C# with the Selenium package[4]来开发一个爬虫程序，以从网站上检索或抓取酒店评论和回复
TripAdvisor网页。Selenium是一种浏览器自动化工具，支持多种语言，如Python、Java和C#。收集的数据包括三个重要方面：评审员简介、酒店信息以及酒店评审和回复。

关于审核人档案，我们收集了与审核人活动相关的用户名、家乡和TripCollective[5]信息。评论者从TripCollective（一个网站上的贡献者项目）获得分数
TripAdvisor）基于活动类型，如撰写评论（100分）和发布照片（30分），以及贡献水平

[4] 硒。网络驱动程序，https://www.nuget.org/packages/Selenium.WebDr艾弗。
[5] 集体的，https://www.tripadvisor.com/TripCollective.

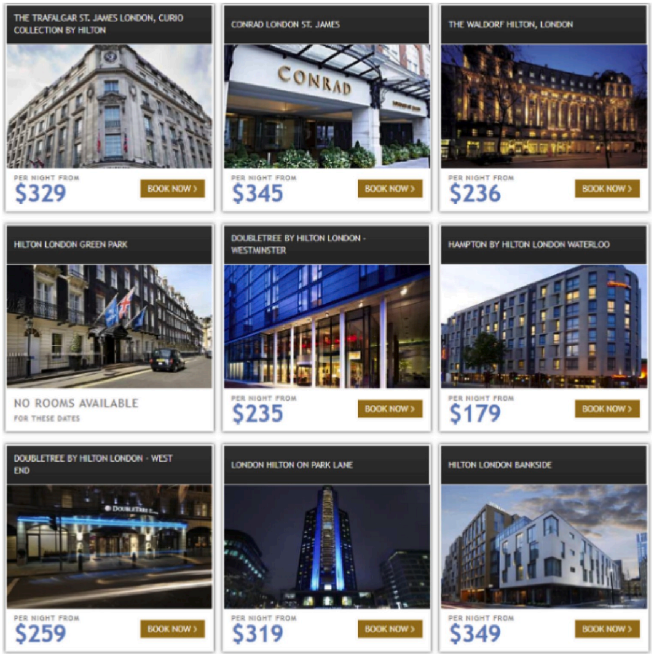

Fig. 3. A sample of hotel page from Hilton destination travel guides.

ranging from 1 (300 points) to 6 (10,000 points). Moreover, the hotel information was collected from three sections, as shown in Fig. 4, on each hotel page. For each hotel, we collected the hotel name, hotel class (star), the number of excellent, good, average, poor, and terrible reviews, an average of a price range, hotel address, location, amenities, type of rooms, hyperlink to the hotel website, and hotel descriptions. For each hotel review, shown in Fig. 5 with the highlighted sections, we collected the review title, review content, manager response, overall rating, aspect rating, traveler type, review date, response date, stay date, and reviewer information. The data were collected from the earliest available date (January 2010) for the selected hotels to the date that data analyses were conducted (October 2018). Table 2 lists an overall review of our data collection. In this research, we have collected data two times in two different years. We finished the initial data collection for the first part of the dataset in early July 2017. In August 2018, we have spent two months extending the dataset and completed the collection by the middle of October 2018. A final total of 113,685 reviews were collected. Among these, 86,907 reviews contained hotel responses, resulting in an overall 76.45% response rate. An average of the overall rating was 3.99 for all collected reviews.

3.3. Data preprocessing

Massive datasets pose immense challenges to data cleaning because manually editing datasets is impractical and ineffective (Franke et al., 2016). The collected raw data are semi-structured and contains noise information. Data preprocessing or wrangling is therefore a fundamental step to transform data into the right form for subsequent learning steps through data cleaning, extraction, transformation, and fusion (Zhou et al., 2017). We used Tableau Prep (King, 2018) to combine, shape and clean the data for data analysis. We first joined the two datasets of hotel data and review data, and then extracted keywords and values (e.g., traveler types and rating scores) out of the raw data. One major challenge we encountered was approximate date information (see Fig. 6). The date shown in the figure is an approximate date, such as 2 days and 2 weeks ago, rather than a precise date. To examine this issue, we first retrieved the date information of the same post several weeks later and found "2 days ago" can be longer than 2 days. As a result, we excluded the data records with approximate dates when we conducted the data analysis and updated the previous dataset.

Our initial data analysis shows that the Trafalgar Hotel had deleted hotel reviews with more than 1 year. To obtain more data from this hotel, we therefore collected the data twice, once in July 2017 and again September 2018. Next, we used the overall rating values to classify data into positive (a 4–5 rating value) and negative (a 1–3 rating value) reviews. The same classification approach was used by Park and Allen (2013) and Proserpio and Zervas (2017). A binary dimension 'Response or Not' was used to indicate whether a review contained a hotel response. Finally, we used the Google Map Geocoding API Get Started | Geocoding API, 2018 to automatically transform each hotel address into a pair of longitude and latitude values to show each hotel precisely on a map.

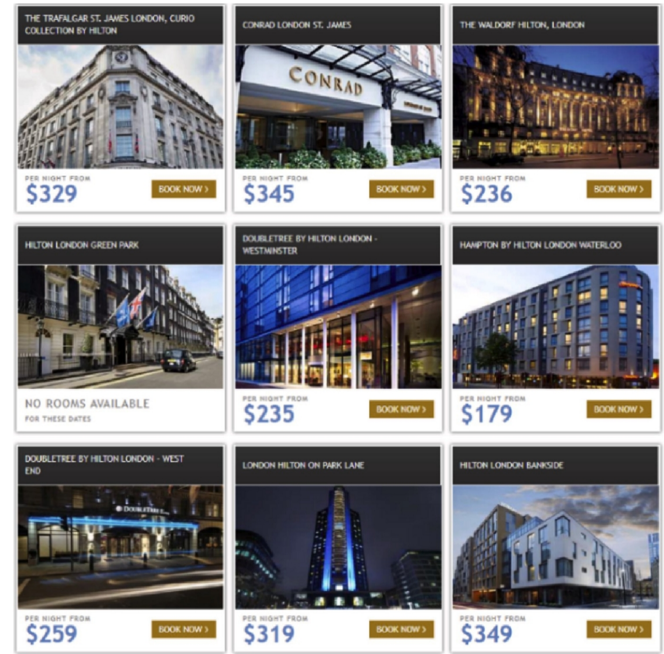

图3. 希尔顿目的地旅游指南中的酒店页面示例。

从1.（300. ）到6.（10000. ）。此外，酒店信息从每个酒店页面的三个部分收集，如图4. 示。对于每家酒店，我们收集了酒店名称、酒店等级（星级）、优秀、优秀、一般、差和糟糕评论的数量、平均价格范围、酒店地址、位置、设施、房间类型、酒店网站的超链接以及酒店描述。对于图5. 突出显示的每个酒店评论，我们收集了评论标题、评论内容、经理回复、总体评分、方面评分、旅行者类型、评论日期、回复日期、入住日期和评论人信息。这些数据是从所选酒店的最早可用日期（2010. 1. ）到进行数据分析的日期（2018. 10. ）收集的。表2. 出了我们数据收集的总体回顾。在这项研究中，我们在两个不同的年份收集了两次数据。2017. 7. 初，我们完成了数据集第一部分的初始数据收集。2018. 8. ，我们花了两个月的时间扩展数据集，并在2018. 10. 中旬完成了收集。最终共收集了113685. 评论。其中，86907. 评论包含酒店回复，总体回复率为76. 45. 。所有收集到的评论的总体评分平均为3. 99. 。

3. 3. 数据预处理

海量数据集给数据清理带来了巨大的挑战，因为手动编辑数据集是不切实际且无效的（Franke et al., 2016）。收集的原始数据是半结构化的，包含噪声信息。因此，数据预处理或争论是一个基本问题通过数据清理、提取、转换和融合，将数据转换为正确的形式，用于后续学习步骤（Zhou等人，2017年）。我们使用Tableau Prep（King，2018）来组合、塑造和清理数据，以进行数据分析。我们首先将酒店数据和审查数据这两个数据集连接起来，然后从原始数据中提取关键词和值（例如，旅行者类型和评级分数）。我们遇到的一个主要挑战是大致的日期信息（见图6）。图中显示的日期是一个大致的日期，例如2天前和2周前，而不是一个精确的日期。为了研究这个问题，我们首先在几周后检索了同一帖子的日期信息，发现"2天前"可以超过2天。因此，在进行数据分析和更新之前的数据集时，我们排除了具有大致日期的数据记录。

我们最初的数据分析显示，特拉法加酒店在一年多的时间里删除了酒店评论。为了从这家酒店获得更多数据，我们收集了两次数据，一次是2017. 7. ，另一次是2018. 9. 。接下来，我们使用总体评级值将数据分为正面（4. 5. 级值）和负面（1. 3. 级值）评论。Park和Allen（2013. ）以及Proserpio和Zervas（2017. ）采用了相同的分类方法。一个二元维度"回复与否"被用来表示评论是否包含酒店回复。最后，我们使用谷歌地图地理编码API Get Started | Geocoding API, 2018. 动将每个酒店地址转换为一对经纬度值，以便在地图上精确显示每个酒店。

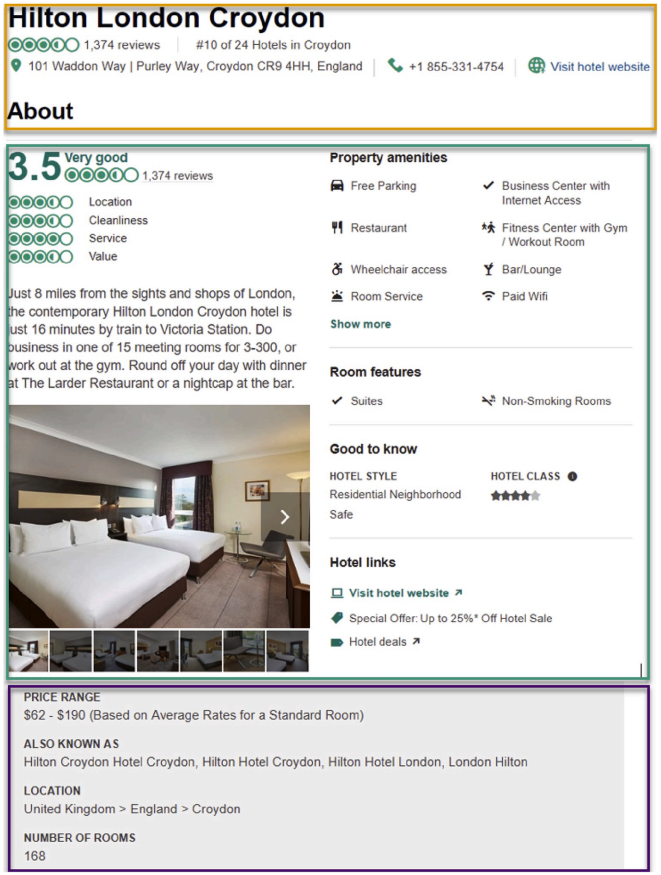

Fig. 4. A sample of hotel page from TripAdvisor – the Hilton London Croydon.

3.2. *Visual analytics*

An interactive visualization can bridge the gap between human and computational analysis (Hu, 2018), and can be used for data cleaning and data preprocessing (Puts et al., 2015). For example, Zhao et al. (2014) presented an interactive visual analysis system combined with machine learning techniques to detect anomalous tweets. In our study, we used interactive visualizations for an exploratory data analysis and for finding the patterns in the data. We used the visualization tool Tableau (Tableau, 2018) for visual analytics because of its popularity and flexibility.

3.2.1. *Spatial analysis*

Fig. 7 lists the selected hotels in this study. An interactive, street-level map was first developed, which enabled us to interactively explore the 43 Hilton-affiliated hotels interactively. The red-blue diverging colors (4 stepped colors) are used to indicate higher (blue) and lower (red) hotel ratings. A circle represents each hotel, where the bigger the size of the circle the greater the number of online reviews. Each hotel is labeled a star rating number from 3.5 to 5, and can be selected to reveal detailed information such as the aspect ratings, average minimum and maximum price, and the number of reviews of each hotel. The drop-down box (a filter) enables investigators to select a year from 2012 to 2018 to explore the rating changes for each hotel. The purpose of this spatial analysis is to observe the relation between hotel locations and other factors such as hotel price, aspect ratings, and overall ratings. The initial analysis indicates there is no evident relation found on this interactive map.

3.1.2. *Response rate analysis*

Responding to guest reviews and encouraging guests to post reviews are common management strategies for to stimulate online

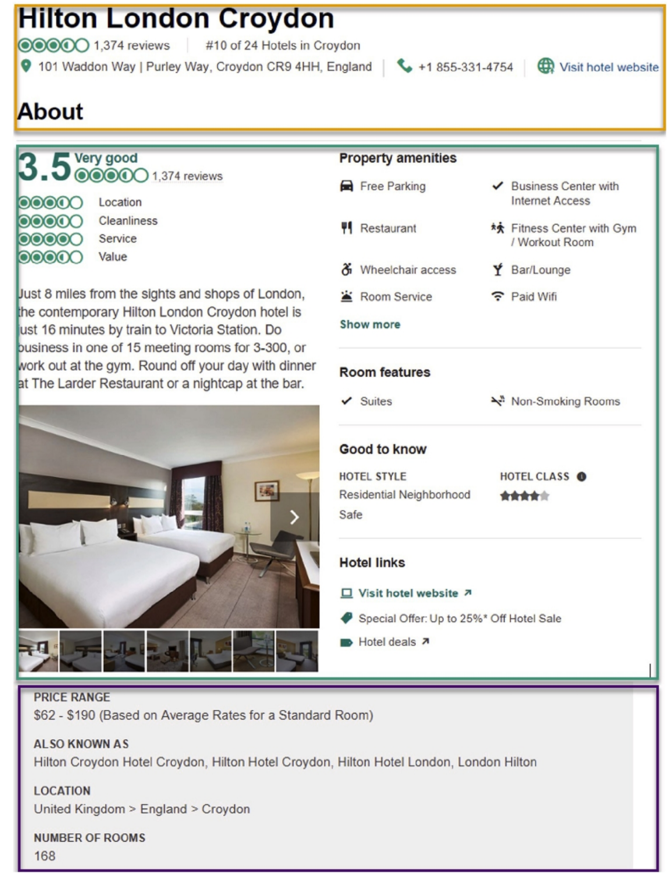

图4. TripAdvisor的酒店页面样本——希尔顿伦敦克罗伊顿酒店。

3.2. 视觉分析

交互式可视化可以弥合人类分析和计算分析之间的鸿沟（Hu，2018），并可用于数据清理和数据预处理（Puts等人，2015）。例如，Zhao等人（2014年）提出了一种交互式视觉分析系统，结合机器学习技术来检测异常推文。在我们的研究中，我们使用交互式可视化进行探索性数据分析，并在数据中找到模式。我们使用可视化工具Tableau（Tableau，2018）进行可视化分析，因为它的流行性和灵活性。

3.2.1. 空间分析

图7列出了本研究中选定的酒店。一个互动的街道地图首先被开发出来，这使我们能够以互动的方式探索希尔顿附属的43家酒店。红蓝发散颜色（4种阶梯色）用于表示较高（蓝色）和较低（红色）的酒店评级。一个圆圈代表每个酒店，其中圈子越大，在线评论的数量就越多。每家酒店都标有一个从3.5到5的星级编号，可以选择该编号来显示详细信息，例如方面评级、平均最低和最高价格，以及每家酒店的评论次数。下拉框（过滤器）使调查人员能够选择2012年至2018年的一年，以探索每家酒店的评级变化。此空间分析的目的是观察酒店位置与其他因素（如酒店价格、方面评级和总体评级）之间的关系。初步分析表明，在这张互动地图上没有发现明显的联系。

3. 1. 2. 响应率分析

回应客人的评论并鼓励客人发表评论是促进网上购物的常见管理策略略

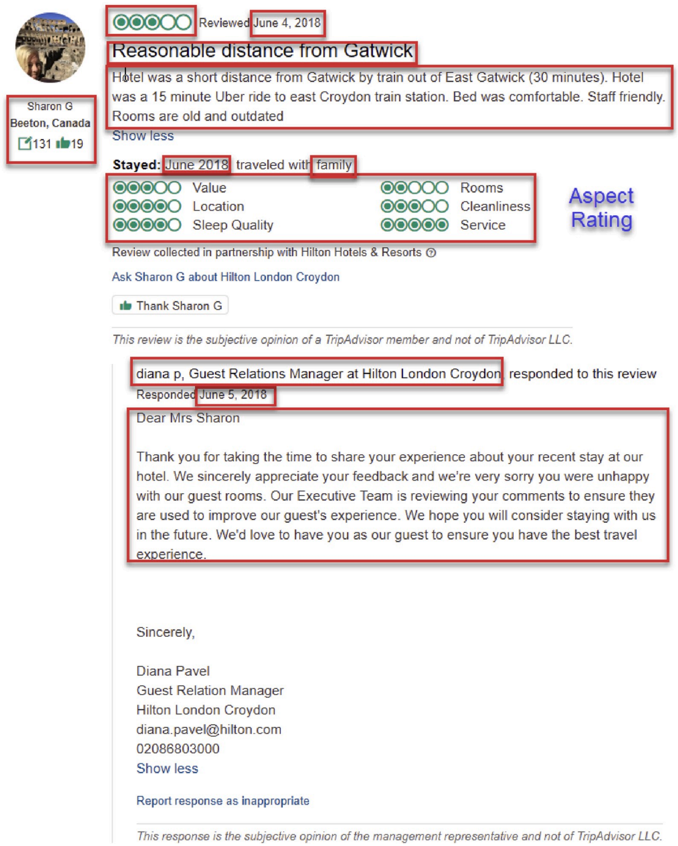

Fig. 5. A sample of hotel review from TripAdvisor – the Hilton London Croydon.

Table 2
Distribution of the sample.

Hotel star	Number of hotels	Number of reviews	Average overall rating
3	1	2423	4.39
3.5	2	4239	4.12
4	37	96,028	3.93
5	4	10,995	4.39
Total	43	113,685	3.99

conversations (De Pelsmacker et al., 2018). Torres et al. (2015) further point out that placing greater value on guest reviews is more likely to improve the perceived hotel quality. An online review can be positive or negative based on the overall rating and the associated aspect ratings, and a hotel manager may or may not respond to the review. Therefore, it would be enlightening to investigate the proportion of responses and non-responses for positive and negative reviews based on the review sentiments. To gain a holistic view of the response rate of each hotel, Fig. 8 displays a partial visualization for the percentage of reviews with hotel responses (in blue) and without (in red) for both positive and negative sentiments.

Overall, three hotels, Double Tree by Hilton Woking with 710 reviews, Hilton London Green Park with 2246 reviews, and Hampton by Hilton London Waterloo with 2423 reviews have the highest non-response rates, which are greater than 78% for positive and 60% for negative reviews. This visualization also reveals that the three 5-star hotels have the relatively higher response rates of 75%–98% than the three 3- and 3.5-star hotels with the response rates of 11%–60% for positive and negative reviews. The 37 4-star hotels show a diversity in the results with high, average, and low response rates.

Based on this visual exploration, we can classify hotels into three categories: negative-response preference, positive-response preference, and neutral preference. If the difference of response rate between positive and negative reviews is less than 10%, the hotel is classified into

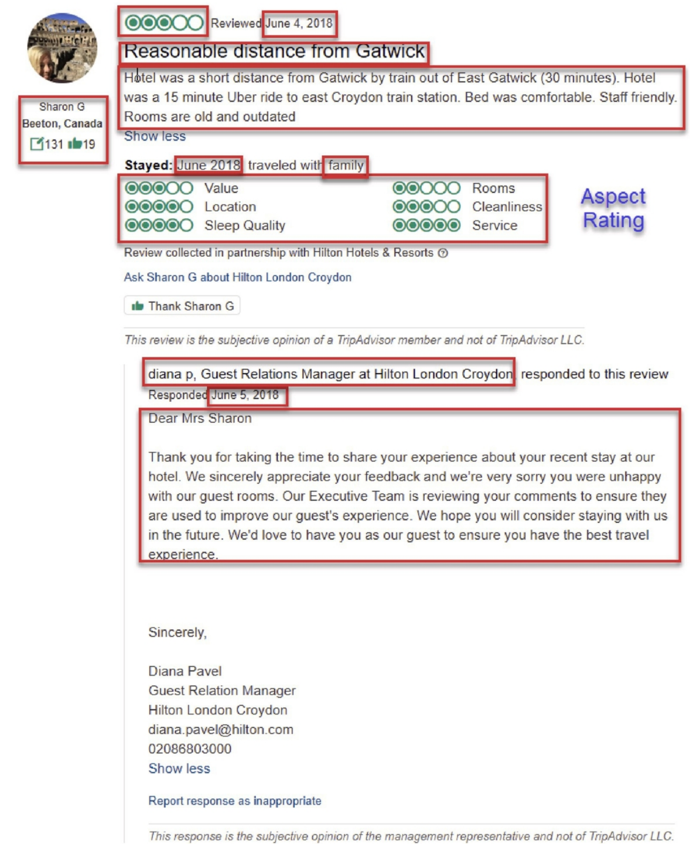

图5. TripAdvisor提供的酒店评论样本——伦敦克罗伊顿希尔顿酒店。

表2
样本的分布。

明星酒店	酒店数量	审查次数	平均总评分
3	1	2423	4.39
3.5	2	4239	4.12
4	37	96,028	3.93
5	4	10,995	4.39
总计	43	113,685	3.99

对话（De Pelsmacker等人，2018年）。Torres等人（2015年）进一步指出，更重视客人评论更有可能改善酒店质量。根据整体评级和相关方面评级，在线评估可以是正面的，也可以是负面的，酒店经理可能会也可能不会对评估做出回应。因此，在回顾的基础上调查正面和负面评论的回复和未回复比例将是有启发性的

感情。为了全面了解每家酒店的回复率，
图8显示了酒店回复（蓝色）和未回复（红色）的正面和负面评论百分比的部分可视化。

总的来说，三家酒店的无回复率最高，分别是希尔顿沃金的双树酒店（Double Tree by Hilton Woking）710条评论、希尔顿伦敦绿色公园酒店（Hilton London Green Park）2246条评论和希尔顿伦敦滑铁卢酒店（Hilton London Waterloo）的汉普顿酒店（Hampton by Hilton Waterloo）2423条评论，正面评论超过78%，负面评论超过。这一可视化还揭示了三颗五星
酒店的回复率相对较高，为75%–98%，而三星级和三星级五星级酒店的正面和负面评论的回复率为11%–60%。37家四星级酒店展现出多样化的服务风格
结果包括高、中、低应答率。

基于这种视觉探索，我们可以将酒店分为三类：消极反应偏好、积极反应偏好和中性偏好。如果正面评论和负面评论的回复率之差小于10%,则酒店被分为

Fig. 6. A sample hotel review with ambiguous dates.

neutral preference; otherwise, it is classified into negative- or positive-response preference. As shown in Fig. 9, 29 hotels have a neutral preference to respond to positive and negative reviews, while 11 and three hotels prefer to respond to negative and positive reviews, respectively.

Our analysis findings are similar to those of Park and Allen (2013) who discovered no pattern in how the hotel managers responded to online reviews for 34 hotels. Additionally, they argued that the hotels did not have a clear strategy to respond to online reviews even among hotels with the same brand, which is consistent with our observation. Although previous studies have shown that consumers are influenced by negative reviews when making purchase decisions (Berger, Sorensen, & Rasmussen, 2010; Sen & Lerman, 2007), our analysis reveals that 67% of hotels have a neutral preference to responding online reviews. That is, most hotel managers appear to put an equal amount of effort responding to both positive and negative reviews.

3.1.3. *Treemap analysis*

We collected hotel rating and review data from 2012 to 2018. A treemap visualization was then used to display the intricate relations between rating, review, and time among all 43 hotels. Each nested rectangle represents a hotel, which is then tiled with smaller rectangles representing each hotel in a specific year. The red-blue diverging color and size dimensions are correlated with the higher (deeper-blue) and lower (deeper-red) ratings and the number of reviews, respectively. In this example (see Fig. 10), we can see that the Hilton London Euston has a relatively lower rating value 1.96 in 2015 with fifty-three reviews. However, applying the same overall rating to compare all 43 hotels will be difficult for a hotel manager to make self-improvement. To address this issue, a single hotel can be selected, and the year dimension is then divided into months. In this example (see Fig. 11) the Hilton London Green Park is selected, and a deeper-red color indicates a relatively lower rating 2.9 with 30 reviews and 21 reviews in July and August 2017, respectively. By repeating this analysis with days instead of months, the hotel manager can efficiently identify the most positive and negative reviews and respond to the critical reviews first.

3.1.4. *Box-and-whisker analysis*

The temporal dimension of sentiments and ratings is often overlooked. In this analysis, we used box-and-whisker plots to reveal the relations between multiple dimensions, including the temporal, overall rating, traveler type, and the number of reviews for all hotels. In Fig. 12, each box represents the values between the first and third quartiles, with the second quartile indicating the median values. The whiskers, two lines outside of the box, represent the lowest and highest observations. Each circle denotes an average overall rating in a month and quarter (Q1 – Q4), and a larger circle size means more reviews; for colors, blue, orange, red, and green are used for the four quarters in a year. In this figure, a lower rating usually occurs in the third and fourth quarters of the year, and in general, business travelers tend to rate lower, while couple travelers tend to rate higher. Banerjee and Chua (2016) also reported a similar finding. Lower ratings occur mostly in Q3 (red circles) and Q4 (green circles) for all known traveler types. Outliers are the circles outside of the plot and are excluded in this visualization. By repeating this analysis, we can identify each type of traveler's behavior and preference in a specific month and quarter of the year.

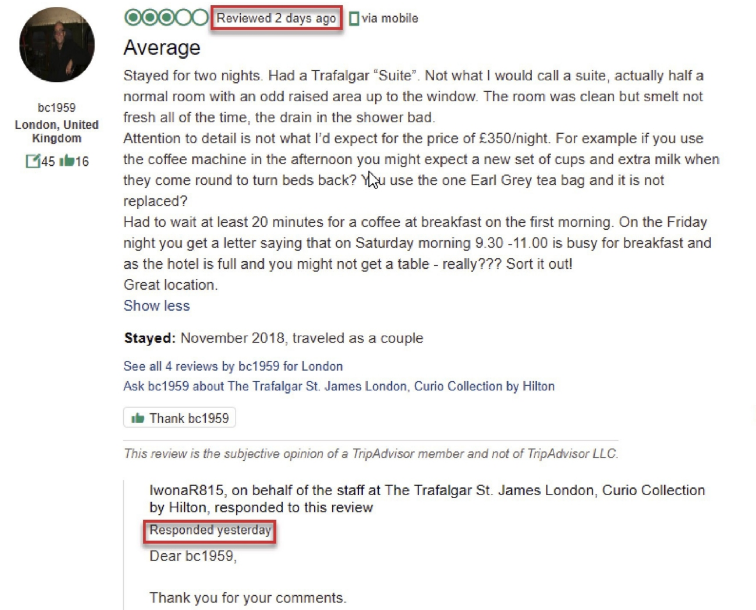

图6. 日期不明确的酒店评论样本。

中性偏好；否则，它被分为消极反应偏好和积极反应偏好。如图9所示，29家酒店对正面评价和负面评价的反应是中性的，而11家和3家酒店分别对负面评价和正面评价的反应是中性的。我们的分析结果与Park和Allen（2013）的分析结果类似，他们没有发现酒店经理对34家酒店在线评论的反应模式。此外，他们认为，即使是在同一品牌的酒店中，酒店也没有明确的策略来回应在线评论，这与我们的观察结果是一致的。尽管之前的研究表明，消费者在做出购买决定时会受到负面评论的影响（Berger、Sorensen和Rasmussen, 2010；Sen和Lerman, 2007），但我们的分析显示，67%的酒店对回复在线评论持中立的偏好。也就是说，大多数酒店经理似乎都付出了同样的努力无论是正面的还是负面的评论。

3.1.3. *树形图分析*

我们收集了2012. 至2018. 的酒店评级和审查数据。然后，使用树状图可视化显示了所有43. 酒店的评级、评论和时间之间的复杂关系。每个嵌套矩形代表一家酒店，然后用较小的矩形平铺，代表特定年份的每家酒店。红蓝色差和大小维度分别与较高（深蓝色）和较低（深红色）评分以及评论数量相关。在这个例子中（见图10. ，我们可以看到希尔顿伦敦尤斯顿酒店2015. 的评级值相对较低，为1. 96. 共有53. 评论。然而，对一位酒店经理来说，使用相同的总体评级来比较所有43. 酒店将很难自我提升。为了解决这个问题，可以选择一家酒店，然后选择年份维度分为几个月。在本例中（见图11），选择了希尔顿伦敦绿色公园，较深的红色表示相对较低的评级为2.9，2017年7月和8月分别有30次评论和21次评论。通过用几天而不是几个月来重复这一分析，酒店经理可以有效地识别最积极和最消极的评论，并首先回应批评评论。

3.1.4. *盒须分析*

情绪和评级的时间维度往往被忽视。在这项分析中，我们使用方框图和胡须图来揭示多个维度之间的关系，包括时间、整体评级、旅行者类型和所有酒店的评论数量。在图12中，每个框表示第一个四分位数和第三个四分位数之间的值，第二个四分位数表示中值。胡须在盒子外有两条线，代表最低和最高的观察值。每个圆圈表示一个月和一个季度（第一季度）的平均总体评级
–第4季度），更大的圆圈意味着更多的评论；对于颜色，蓝色、橙色、红色和绿色用于一年中的四个季度。在这个数字中，较低的评级通常出现在一年的第三季度和第四季度，一般来说，商务旅行者的评级较低，而情侣旅行者的评级较高。Banerjee和Chua（2016）也报告了类似的发现。对于所有已知的旅行者类型，较低的评级大多出现在第三季度（红圈）和第四季度（绿圈）。异常值是绘图外的圆，在该可视化中被排除。通过重复这一分析，我们可以在一年中的特定月份和季度确定每种类型的旅行者的行为和偏好。

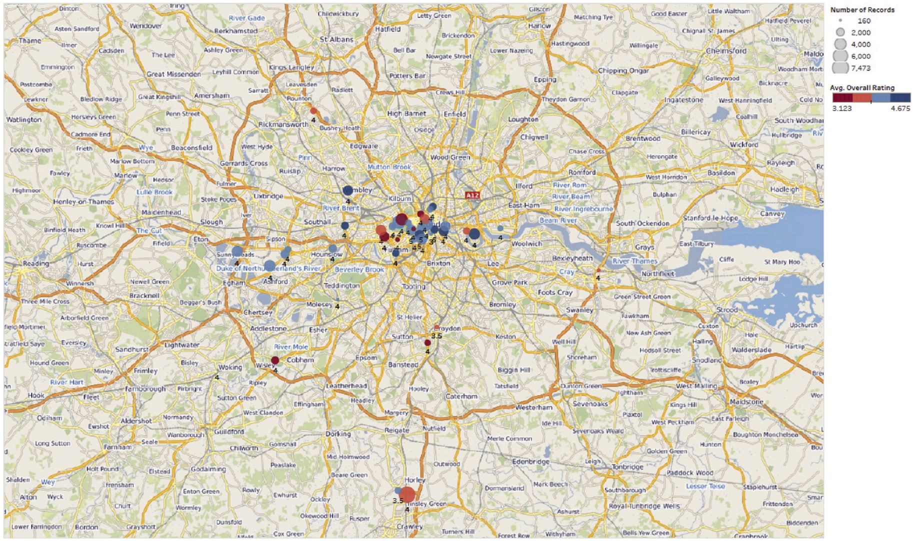

Fig. 7. Spatial analysis of the 43 hotels in London.

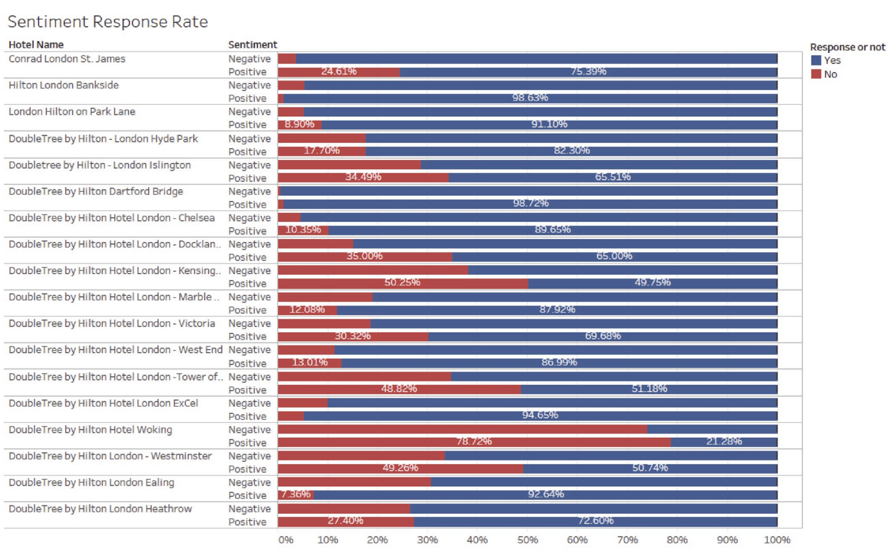

Fig. 8. Positive and negative response rates for the selected hotels in London.

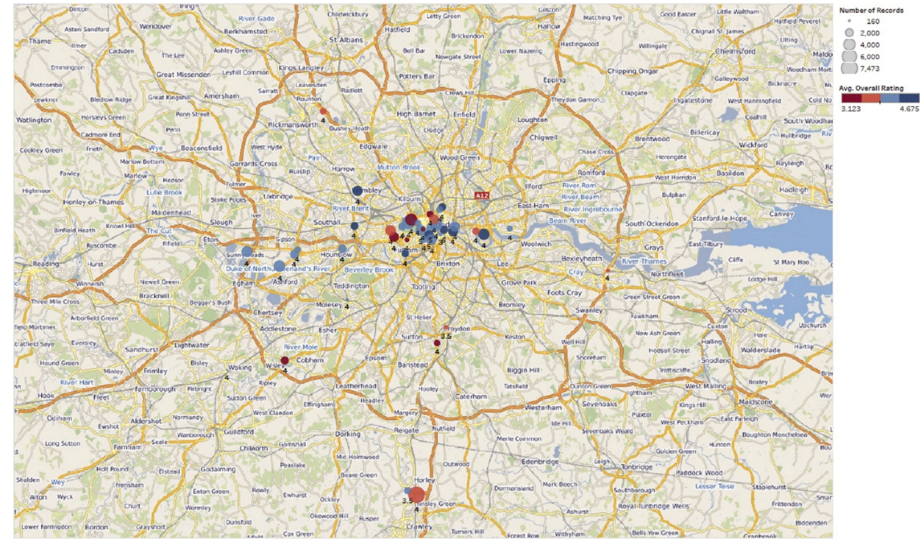

图7. 伦敦43. 酒店的空间分析。

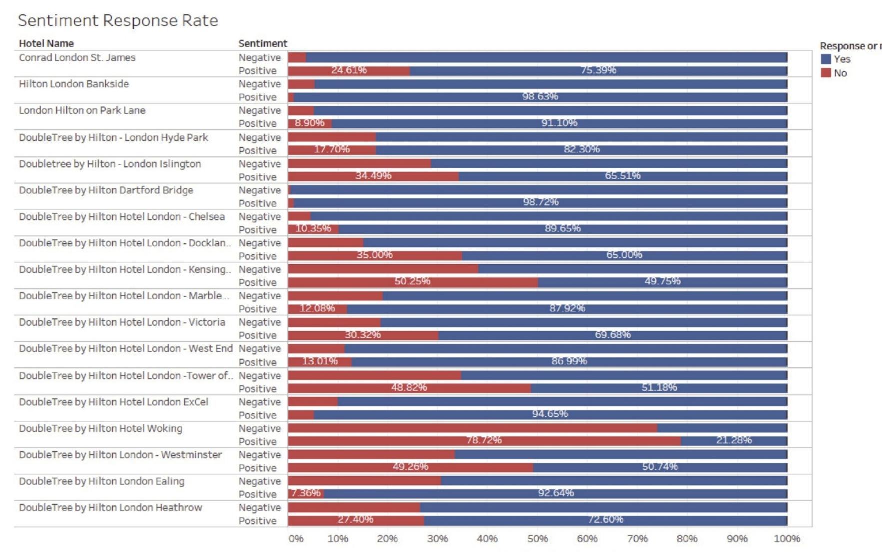

图8. 伦敦选定酒店的正面和负面回复率。

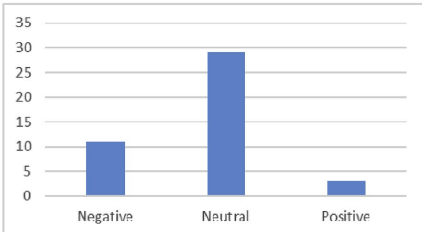

Fig. 9. A preference to respond to types of hotel reviews.

embedding, and a CNN model concatenating the hotel review features. To efficiently conduct machine learning, it is important to preprocess the raw text data, so we first transformed all words to lower case for consistency, then we filtered out the stop words such as "a" and "the," which contain very little information, and finally, we removed punctuation.

Word embeddings can be generated using the Word2Vec framework, of which there are two different models: the continuous bag-of-word based model (CBOW) and the skip-gram model. The CBOW model, which is different from the traditional bag-of-words model, predicts the current word based on a continuous distributed representation of the context. In contrast, the skip-gram model predicts words from a range before and after the current word. The quality of the skip-gram model improves as the range stretches, but this increases the computation effort (Mikolov et al., 2013).

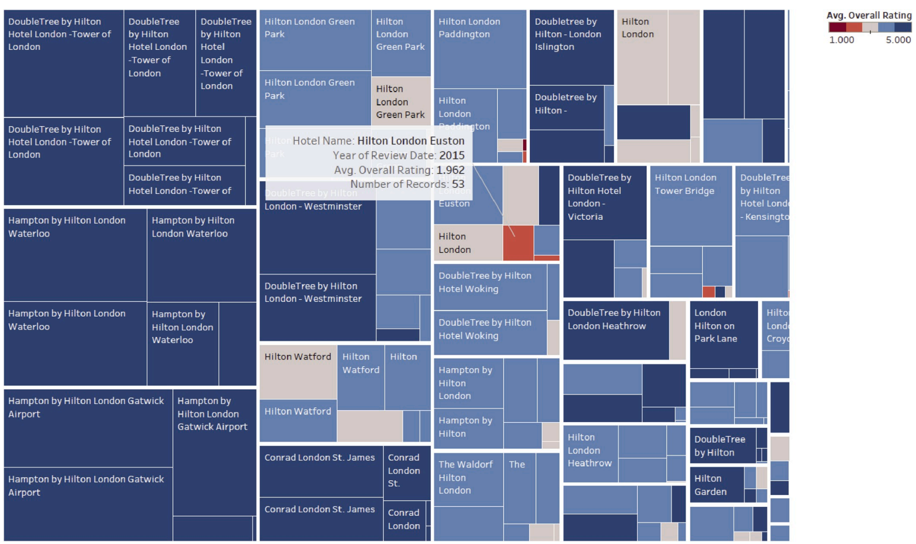

Fig. 10. Treemap analysis of 43 hotels in London with a time dimension.

4. Convolutional neural network-based multi-feature fusion for hotels responses

Learning hotel responses of high-quality hotels is crucial for hotel management. The decision to respond to a review may increase transaction and labor costs (Schuckert et al., 2015) and not to respond to a review may result in lost opportunities to retain customers (Yoo & Gretzel, 2008). As Liu et al. (2015) recommend that hotel managers adopt targeted response management to increase hotel ratings, it is important to prioritize the responses to online reviews. Leung et al. (2013) advise hotel managers to respond to online reviews and encourage scholars to further investigate hotel responses (Min et al., 2015).

To effectively detect proactive and non-proactive hotel responses, we propose a deep learning-based approach that integrates multiple CNNs (Kim, 2014, pp. 1746–1751) and multi-features for text classification. We classified proactive and non-proactive responding to a hotel review by developing a model with the three steps of preprocessing, word

In addition to the self-document embedding used by Word2Vec, recent empirical research has also regularly applied pre-trained instead of self-trained embeddings. Because researchers found that using pre-trained embeddings may accelerate performance (Liu, 2015), we utilized pre-trained word embeddings[6] of GloVe, an unsupervised learning algorithm. This algorithm transforms preprocessed hotel reviews into the document matrix, the rows of which are word vector representations of each token. Following Collobert and Weston (2008), we can effectively treat the document matrix as an image upon which we can perform convolutions. Fig. 13 shows the architecture of our proposed method, which consists of seven layers: the Input, Convolution, Max-pooling, Flattening, Concatenation, Fully-connected, and Softmax layers. The following summarizes each of the layers.

[6] https://nlp.stanford.edu/projects/glove/.

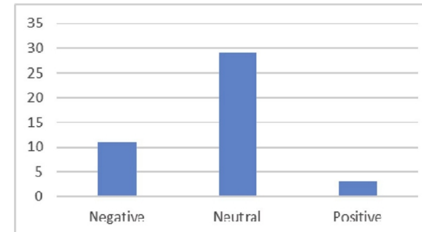

图9. 对酒店评论类型做出回应的偏好。

嵌入，以及连接酒店评论功能的CNN模型。为了有效地进行机器学习，预处理非常重要
原始文本数据，因此我们首先将所有单词转换为小写，以保持一致性，然后过滤掉包含很少信息的停止词，如"a"和"the"，最后删除标点符号。

单词嵌入可以使用Word2Vec框架生成，该框架有两种不同的模型：基于单词的连续包模型（CBOW）和skip-gram模型。CBOW模型不同于传统的单词袋模型，它基于上下文的连续分布式表示来预测当前单词。相比之下，skip-gram模型预测当前单词前后的单词。skip-gram模型的质量随着范围的扩大而提高，但这会增加计算工作量（Mikolov等人，2013年）。

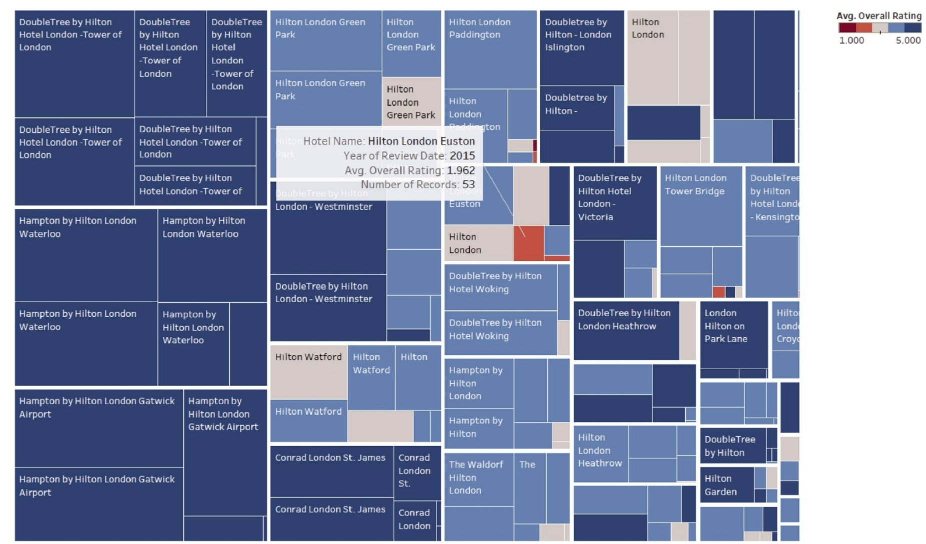

图10. 用时间维度对伦敦43. 酒店进行树状图分析。

4. 基于卷积神经网络的酒店响应多特征融合

了解高质量酒店的酒店反应对酒店管理至关重要。回应审查的决定可能会增加交易和劳动力成本（Schuckert等人，2015年），而不回应审查可能会导致失去留住客户的机会（Yoo & Gretzel, 2008年）。正如Liu等人（2015年）建议酒店经理采用针对性响应管理来提高酒店评级，重要的是对在线评论的响应进行优先排序。Leung等人（2013年）建议酒店经理回复在线评论，并鼓励学者进一步调查酒店回复（Min等人，2015年）。

为了有效检测酒店的主动和非主动响应，我们提出了一种基于深度学习的方法，该方法集成了多个CNN
（Kim，2014,第1746–1751页）和文本分类的多特征。
我们对酒店评论的主动响应和非主动响应进行了分类
通过开发一个包含三个预处理步骤的模型，word

除了Word2Vec使用的自我文档嵌入之外，最近的实证研究也经常使用预训练而非自我训练的嵌入。因为研究人员发现-
经过训练的嵌入可能会提高性能（Liu，2015），我们使用了手套的预训练单词嵌入[6]，这是一种无监督学习算法。该算法将预处理的酒店评论转换为
文档矩阵，其中的行是每个标记的字向量表示。继Collobert和Weston（2008）之后，我们可以有效地将文档矩阵视为可以执行卷积的图像。图13显示了我们提出的方法的架构，它由七层组成：输入层、卷积层、最大池、平坦层、级联层、完全连接层和Softmax层。下面总结了每一层。

[6] https://nlp.stanford.edu/projects/glove/.

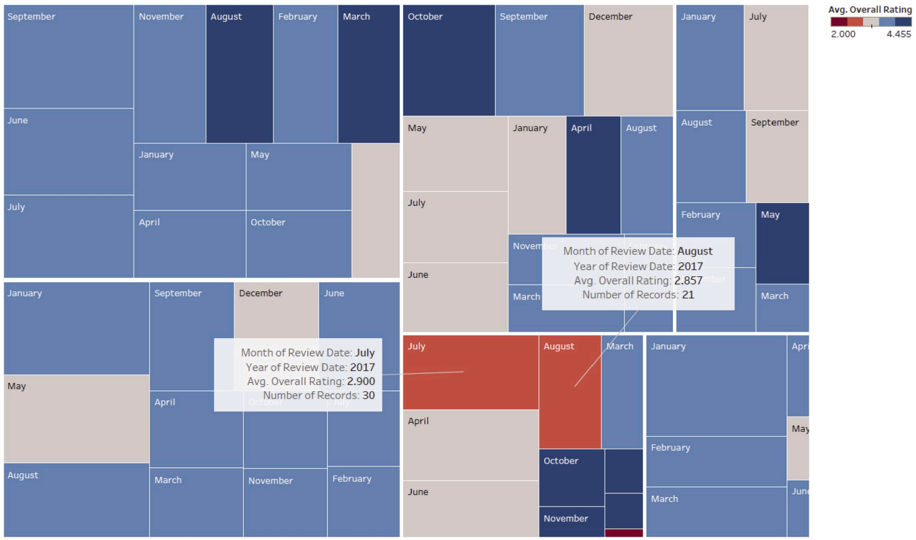

Fig. 11. Treemap analysis of hilton London Green Park with a time dimension. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4.1. Input layer

We used GloVe pre-trained word embeddings (i.e., glove.6B) to transform hotel reviews into a 300-dimension document matrix. We also set the maximum document length to 150, where longer documents were truncated, and shorter documents were padded with zeros.

4.2. Convolution layer

We acquired new feature c_i with the filter w, using the window of h words from i to $i+h$-1. Through activation function (s) and passing the bias (b), we obtained the c_i function as follows:

$$c_i = s(wx_{i:i+h-1} + b) \tag{6}$$

In this study, the activation function was set to *ReLU* and we depicted the three filter region sizes of 3, 4 and 5, with each having 256 filters. Filters perform convolutions on the document matrix and generate feature maps as follows:

$$c = [c_1, \ c_2, \ \cdots c_{n-h+1}] \tag{7}$$

4.3. Max-pooling layer

The 1-max pooling was performed over each map to capture the largest value $\hat{c} = \max\{c\}$ from each feature map.

4.4. Flattening layer

After max-pooling, we concatenated the matrices processed by three different filter kernel sizes. Then we needed to flatten the joint matrix into 1 dimension with the aim of concatenating the 1-dimension additional features.

4.5. Concatenation layer

Deep learning can unearth the latent features by itself, usually relying on language input instead of feature engineering (Young et al., 2017). However, to enhance performance, a number of feature vectors can be put together with the neural networks (Do et al., 2019). In this research, we therefore integrated three aspects of the feature set into the CNN based on the following three visual analytics findings:

- *Sentiment Score*: We investigated the impact of sentiment score for recognizing the proactive hotel responses. The overall rating and aspect ratings are included in this feature set, which is equal to 1 if the overall rating is greater than 3; otherwise, it is 0; for the aspect ratings, we considered the aspects of service and cleanliness based on the observation of visualizations. Here, we used the rating score of both aspects for the feature values.
- *Temporal Interval*: From the temporal visualizations, we observed that the proactive hotel responses were impacted by different time periods. For instance, managers often cannot promptly respond to hotel reviews on Saturday since it is the busiest day. Based on this perspective, we considered the influence of the day of the week of hotel reviews when recognizing proactive hotel responses.
- *Reviewer Profile*: A reviewer's profile is related to the credibility and trustworthiness of a review (Filieri, 2016). Therefore, proactive hotel responses can be associated with the characteristics of travelers. For instance, a manager would be more likely reply to a hotel review from a senior traveler with a more credible profile than a junior one. This is because the hotel review posted by a senior traveler is more valuable as a reference. Taking this into account, we explored the characteristics of users, including traveler types {family, friends, couple, business, and solo} and contributor levels {1–6 levels} to develop our deep learning model. Each traveler can earn different

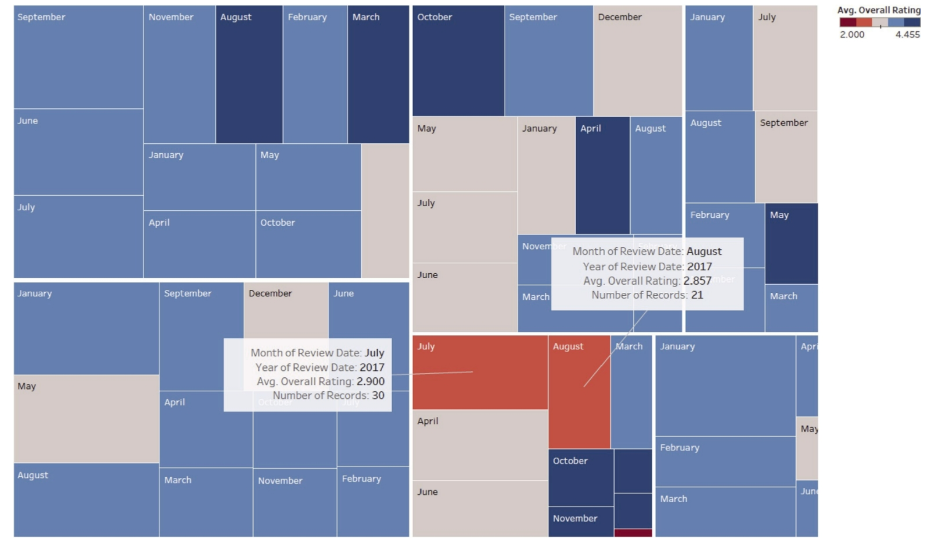

图11. 希尔顿伦敦绿色公园的时间维度树状图分析。（有关此图例中颜色的说明，请参阅本文的网络版。）

4.1. 输入层

我们使用手套预先训练的单词嵌入（即手套6B）将酒店评论转换为300维文档矩阵。我们还将最大文档长度设置为150,其中较长的文档被截断，较短的文档被填充为零。

4.2. 卷积层

我们使用从i到i þ h-1. h字窗口，通过过滤器w获得了新的特征c_i。通过激活函数（s）并通过偏置（b），我们得到了c[1. 函数，如下所示:

c_i ¼ sðwx$_{i:iþh\ 1}$ þ bÞ (6)

在这项研究中，激活函数设置为ReLU，我们描述了3、4和5的三个过滤区域大小，每个区域有256个过滤器。过滤器对文档矩阵执行卷积，并生成特征映射，如下所示:

c ¼ ½c$_1$；c$_2$；⋯ c$_n$ $_{hþ1}$ (7)

4.3. 最大池层

在每个地图上执行1-max池，以从每个特征地图中获取最大值c ¼ maxfcg。 b

4.4. 压平层

在最大池之后，我们将由三种不同大小的过滤器内核处理的矩阵连接起来。然后我们需要将关节矩阵展平为一维，目的是连接一维附加特征。

4.5. 连接层

深度学习可以自行挖掘潜在特征，通常依靠语言输入而不是特征工程（Young等人，2017年）。然而，为了提高性能，可以将许多特征向量与神经网络结合在一起（Do等人，2019年）。因此，在本研究中，我们基于以下三个视觉分析结果，将该功能集的三个方面整合到CNN中:

- 情绪评分：我们调查了情绪评分对识别酒店主动反应的影响。此功能集中包括整体评级和方面评级，如果整体评分大于3分；否则为0；对于方面评级，我们基于对可视化的观察，考虑了服务和清洁度方面。在这里，我们使用这两个方面的评分作为特征值。
- 时间间隔：我们从时间可视化中观察到酒店的主动响应受到不同时间段的影响。例如，由于周六是最繁忙的一天，经理们往往无法及时回复酒店评论。基于这一观点，我们在识别酒店的主动回应时，考虑了酒店每周评论的影响。
- 评审员简介：评审员简介与可信度和审查的可信度（Filieri，2016）。因此，酒店的主动响应可能与旅行者的特征有关。例如，与初级旅行者相比，经理更有可能以更可靠的个人资料回复高级旅行者对酒店的评论。这是因为资深旅行者发布的酒店评论更具参考价值。考虑到这一点，我们探索了用户特征，包括旅行者类型{家庭、朋友、夫妻、企业和个人}和贡献者级别{1–6级}，以开发我们的深度学习模型。每个旅行者的收入都不一样

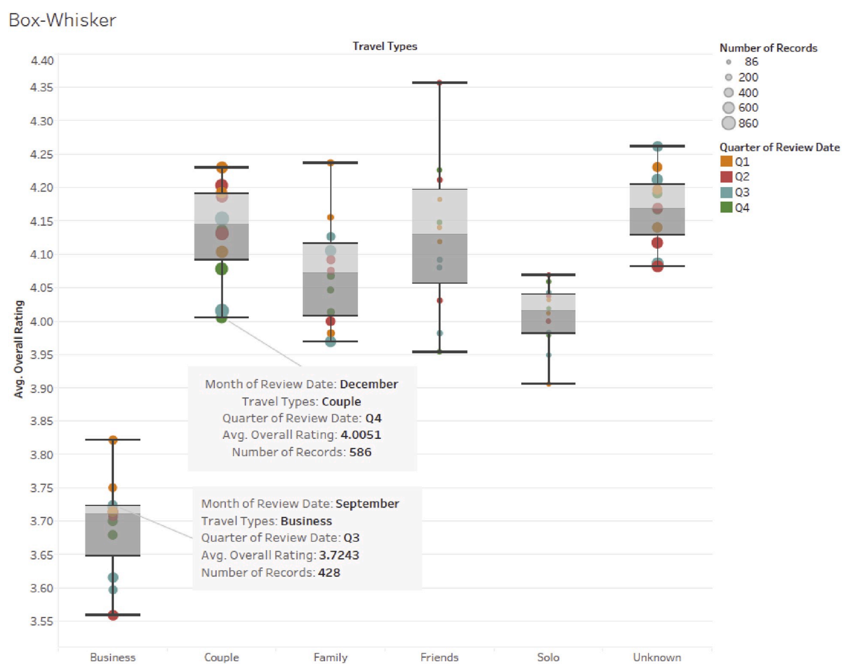

Fig. 12. Box-and-Whisker analysis of traveler types with a time dimension.

points according to the type of contributions to TripAdvisor review,[7] so to be a level 6 contributor, at least 10,000 points are required, compared to 300 points required by a level 1 contributor.

4.6. *Fully-connected layer and softmax layer*

Following the Concatenation layer, we designed two dense fully-connected layers to gradually reduce the dimension to 16. The final *softmax* layer then receives this 16-dim vector as an input and uses it to classify the hotel review; here we assumed a binary classification and hence depicted two possible output states.

The CNN model was implemented using Keras,[8] a Python deep learning library. For this, we used a binary cross-entropy as loss function and Adam as the optimizer. The batch size was set to be 256, and the training lasted for 10 epochs.

5. Experiment results and discussion

In this paper, we evaluated the algorithm performance by identifying proactive hotel responses in terms of the precision rate, recall rate, and the F_1-score, as well as the micro-average metrics to compare the average performance (Manning et al., 2008). First, we investigated the effect of additional feature sets of the hotel review that improves our multi-CNN model by adding features of sentiment score, temporal interval, and reviewer profile. Table 3 displays the system performances of our multi-CNN model (denoted as 3CNN) and the results of incrementally applying the extra three feature sets, denoted as + *SentimentScore*, +*TemporalInterval*, and +*ReviewerProfile*. Moreover, considering the response quality of different hotels is various.

In addition, we conducted an iterative k-means clustering analysis with visual analytics to generate three clusters using aspect ratings such as cleanliness and service ratings (see Table 4). The clusters were generated by selecting varying combinations of features and observing between-group and within-group sums of squares. We compared three clusters of hotels with higher, average, and lower aspect ratings to see if hotel response strategies were different. Among the three clusters, cluster 3 with eight hotels demonstrates the highest contribution level and rating values. We further investigated the relationship between each contributor level and rating by visualizing their relationships, but there were no corresponding results. Note that the average contributor level listed in Table 4 is part of the reviewer profile, which is included in the *concatenation layer* of our deep learning model.

As shown in Table 3, adding the features of sentiment score and temporal interval performs better than the 3CNN, which can improve about 2% F_1-score. This is because some hotel managers tend to only respond to the hotel reviews with an extreme sentiment, while other hotel responses are less consistent. That is, the sentiment factor can only increase the performance slightly. This echoes Park and Allen's (2013) findings as well our visualization results.

It is worth noting that the feature of temporal interval further

[7] TripCollective, https://www.tripadvisor.com/vpages/tripcollective_faqs.html.

[8] Keras, https://keras.io/.

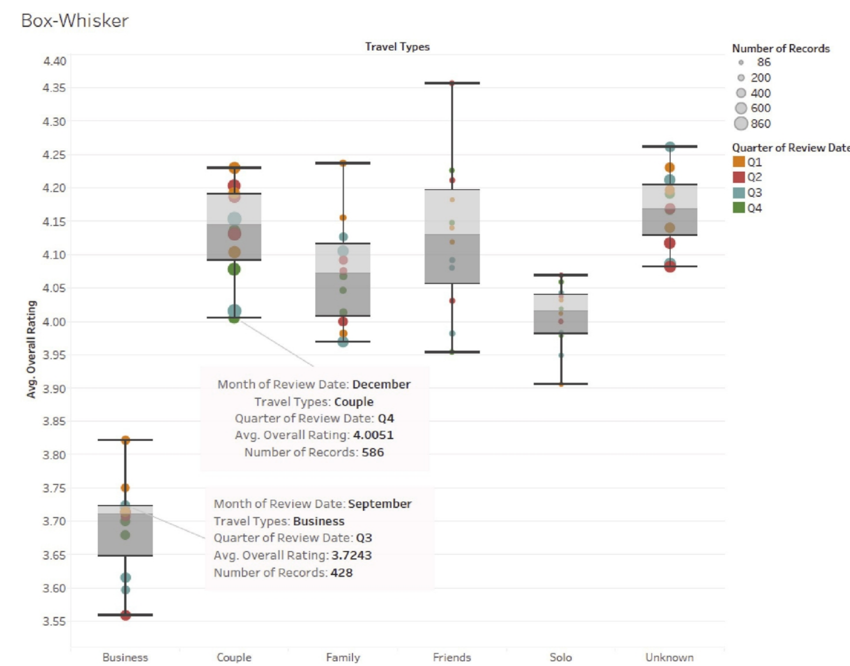

图12. 具有时间维度的旅行者类型的盒须分析。

根据TripAdvisor review的贡献类型，[7]因此，要成为6级贡献者，至少需要10000分，而1级贡献者需要300分。

4. 6. *全连接层和softmax层*

在连接层之后，我们设计了两个密集的完全连接层，以逐渐将维度减少到16. 最后一个softmax层接收这个16. 矢量作为输入，并使用它来对酒店评论进行分类；这里我们假设了一种二进制分类，因此描述了两种可能的输出状态。

CNN模型是使用Keras[8] 实现的，这是一个Python深度的模型学习图书馆。为此，我们使用二进制交叉熵作为损失函数，Adam作为优化器。批量大小设置为256,培训持续了10个时期。

5. 实验结果与讨论

在本文中，我们通过识别酒店主动响应的准确率、召回率和F_1分数，以及比较平均性能的微观平均指标来评估算法性能（Manning et al., 2008）。首先，我们调查了酒店评论的附加功能集对提高我们的服务质量的影响

多CNN模型通过添加情绪评分、时间间隔和评论者档案的特征。表3显示了我们的多CNN模型（表示为3CNN）的系统性能，以及增量应用额外三个特征集（表示为3CNN）的结果 Þ 感伤的分数，ÞTemporalInterval和RevÞwerProfile。此外，考虑到不同酒店的响应质量是不同的。

此外，我们使用可视化分析进行了迭代k均值聚类分析，以使用清洁度和服务评级等方面评级生成三个聚类（见表4）。通过选择不同的特征组合并观察组间和组内的平方和生成聚类。我们比较了三个方面评分较高、平均和较低的酒店集群，看看酒店的应对策略是否不同。在这三个集群中，拥有八家酒店的集群3表现出最高的贡献水平和评级值。我们通过可视化每个贡献者水平和评级之间的关系，进一步研究了它们之间的关系，但没有相应的结果。请注意，表4中列出的平均贡献者级别是reviewer配置文件的一部分，它包含在我们的深度学习模型的连接层中。

如表3所示，添加了情绪评分和时间间隔的性能优于3CNN,后者可以提高约2%的F_1分数。这是因为一些酒店经理往往只以极端情绪回应酒店评论，而其他酒店的回应则不那么一致。也就是说，情绪因素只能略微提高绩效。这与Park和Allen（2013）的发现以及我们的可视化结果相呼应。

值得注意的是，时间间隔的特征

[7] 集体的， https://www.tripadvisor.com/vpages/tripcollective_faqs.html.

[8] Keras, https://keras.io/.

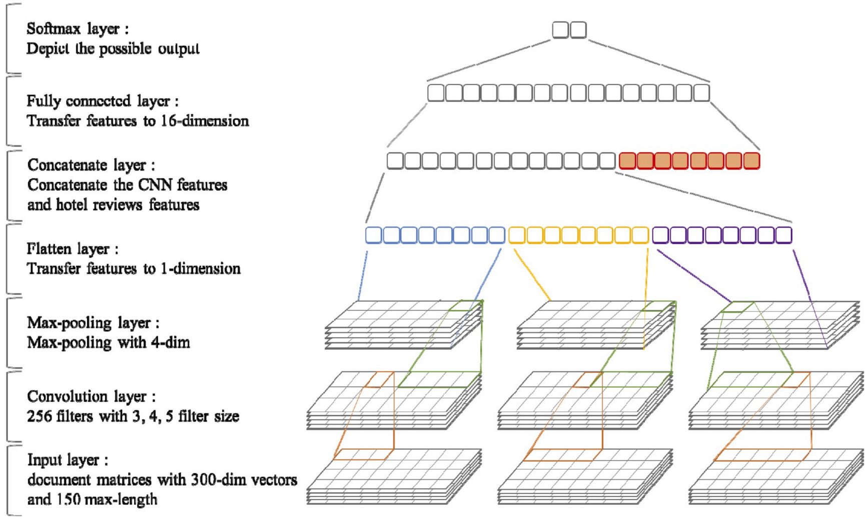

Fig. 13. Illustration of a CNN-based multi-feature fusion architecture for detecting proactive hotel responses.

Table 3
Incremental contribution of adding different features with and without clustering for detecting proactive responses.

System	Cluster	Proactive	Non-proactive	Micro-avg
		Precision, Recall, F_1-score (%)		
3CNN	without cluster	54.47/ 33.55/**41.52**	57.26/ 76.04/65.32	55.97/ 56.46/ 56.22
+SentimentScore	without cluster	56.08/ 29.11/38.33	57.08/ 80.52/66.80	56.62/ 56.84/ 56.73
+TemporalInterval	without cluster	57.98/ 31.60/40.91	57.92/ 80.43/67.35	57.95/ 57.94/ **57.94**
+ReviewerProfile	without cluster	57.47/ 28.75/38.33	57.34/ 81.83/**67.43**	57.40/ 57.37/ 57.39
3CNN	with cluster	52.61/ 32.47/40.15	57.36/ 75.55/65.21	55.25/ 56.40/ 55.82
+SentimentScore	with cluster	52.89/ 31.46/39.45	57.34/ 76.64/65.60	55.37/ 56.53/ 55.94
+TemporalInterval	with cluster	54.93/ 30.48/39.20	57.42/ 78.56/66.35	56.33/ 57.10/ 56.71
+ReviewerProfile	with cluster	55.95/ 28.93/38.14	57.52/ 80.74/67.18	56.87/ 57.47/ 57.17

improves the system performance, which indicates that proactive and non-proactive hotel responses are associated with a day of the week. Our further analysis indicates that Sunday and Monday account for the most proactive responses with 19.58% and 20.44%, respectively; this implies that nearly half of the proactive responses are made on these two days. In contrast, Thursday and Friday account the least proactive responses with 9.66% and 3.58%, respectively, which may relate to the degree of busyness preparing for the weekend peak stay time. However, the overall performance slightly decreases when the feature of the reviewer profile is added. This is because reviewer profile features for the number of proactive and non-proactive responses is not a significant difference in terms of *t*-test with a 95% confidence level. It is interesting to note that the overall results indicate there is no difference with and without clusters. The clusters cannot enhance the overall deep learning performance for at least three reasons: 1) the size of training sample, 2) the quality of clusters, and 3) the diversity of hotel types (1–5 star hotels). This is consistent with analysis results from the response rate visualizations. Therefore, we decided to conduct a further performance evaluation without clusters.

A comprehensive performance evaluation of the proposed CNN-based approach with other methods is provided in Table 5. In this experiment, the word embedding-based approaches represent each hotel review as an average of word embeddings (300-dimension embeddings) and are classified by the SVM (denoted as SVM). Next, we further compared our method to TextCNN, a well-known CNN-based text classification approach (Kim, 2014, pp. 1746–1751; denoted as *CNN*), and also to the bi-directional recurrent neural network method (Lai et al., 2015; denoted as *RNN*). To serve as baseline standards for comparison, we also included the results of Naïve Bayes (denoted as *NB*) and *k*-nearest neighbors (Guo, Wang, Bell, Bi, & Greer, 2006; denoted as *KNN*).

As a baseline, the Naïve Bayes classifier is a keyword statistics-based approach which only accomplishes a mediocre performance with a 43% F_1-score. The word embeddings-based methods (i.e., SVM) are more effective in extracting discriminative keywords and exhibit a more evenly distributed performance among both categories. Therefore, SVM further improves the performance to a 46% F_1-score. It is worth noting that the KNN simply calculates document similarity in the bag-of-words feature space which outperforms both NB and SVM. This is because the distribution of hotel review representations may be hard-divided and imperfectly independent so that NB and SVM are inferior to KNN. As

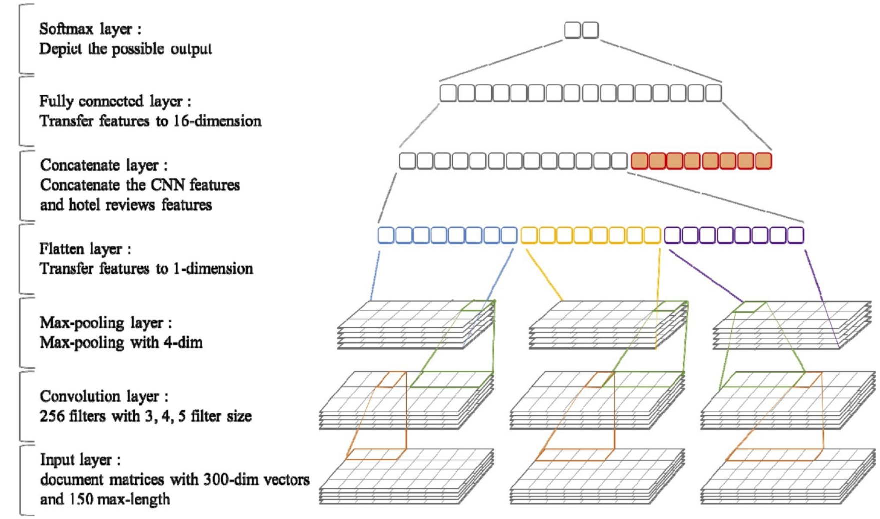

图13. 用于检测酒店主动响应的基于CNN的多功能融合架构示意图。

表3
通过添加不同的功能（包括和不包括集群）来检测主动响应的增量贡献。

系统	集群	积极主动的	Non-积极主动的	微平均
		Precision, Recall, F_1-score (%)		
3CNN	没有集群	54.47/ 33.55/**41.52**	57.26/ 76.04/65.32	55.97/ 56.46/ 56.22
þSentimentScore	没有集群	56.08/ 29.11/38.33	57.08/ 80.52/66.80	56.62/ 56.84/ 56.73
þTemporalInterval	没有集群	57.98/ 31.60/40.91	57.92/ 80.43/67.35	57.95/ 57.94/ **57.94**
þReviewerProfile	没有集群	57.47/ 28.75/38.33	57.34/ 81.83/**67.43**	57.40/ 57.37/ 57.39
		52.61/ 32.47/40.15	57.36/ 75.55/65.21	55.25/ 56.40/ 55.82
3CNN	使用集群	52.89/ 31.46/39.45	57.34/ 76.64/65.60	55.37/ 56.53/ 55.94
þSentimentScore	带簇	54.93/ 30.48/39.20	57.42/ 78.56/66.35	56.33/ 57.10/ 56.71
þTemporalInterval	使用集群	55.95/ 28.93/38.14	57.52/ 80.74/67.18	56.87/ 57.47/
þReviewerProfile	使用集群			57.17

提高了系统性能，这表明主动和非主动酒店响应与一周中的某一天相关。我们进一步的分析表明，周日和周一的反应最积极，分别为19.58%和20.44%；这意味着近一半的主动回应是在这两天做出的。相比之下，周四和周五的反应最不积极

分别为9.66%和3.58%,这可能与忙着为周末高峰停留时间做准备。然而添加reviewer配置文件的功能后，总体性能略有下降。这是因为reviewer profile的数字功能在95%置信水平的t检验中，主动和非主动反应的差异不显著。有趣的是，整体结果表明，有或没有集群没有区别。集群不能提高整体深度学习绩效，至少有三个原因：1）训练样本的规模，2）训练的时间集群的质量，以及3）酒店类型的多样性（1–5星级酒店）。

这与反应率可视化的分析结果一致。因此，我们决定在没有集群的情况下进行进一步的绩效评估。

表5. 出了所提出的基于CNN的方法与其他方法的综合性能评估。在这个实验中，基于单词嵌入的方法将每个酒店评论表示为单词嵌入的平均值（300. em- bedding），并通过SVM（表示为SVM）进行分类。接下来，我们进一步将我们的方法与著名的基于CNN的文本分类方法TextCNN（Kim，2014,第1746–1751页；表示为CNN）以及双向递归神经网络方法（Lai等人，2015；表示为RNN）进行了比较。为了作为比较的基线标准，我们还纳入了Naïve Bayes（表示为NB）的结果和k-最近邻（郭、王、贝尔、毕和格里尔，2006；表示为*KNN*).

作为基线，Naïve Bayes分类器是一种基于关键字统计的方法，其性能一般，只有43%

F_1——得分。基于单词嵌入的方法（即SVM）在提取区分性关键词方面更有效，并且在这两个类别中表现出更均匀的性能。因此，SVM进一步将性能提高到46%的F_1分数。值得注意的是，KNN只需在单词包特征空间中计算文档相似度，其性能优于NB和SVM。这是因为酒店评论表述的分布可能难以划分且不完全独立，因此NB和SVM不如KNN。像

Table 4
Three Clusters of hotels.

Clusters	Number of hotels	Avg. contributor level	Avg. overall rating	Avg. max price	Avg. min price	Avg. service ratings	Avg. cleanliness ratings
1	11	3.09	3.47	216.81	93.70	3.83	3.85
2	24	3.17	4.12	215.15	94.5	4.22	4.50
3	8	**3.34**	**4.35**	**406.11**	**165.83**	**4.41**	**4.61**

Table 5
The performance result of compared methods.

System	Proactive	Non-proactive	Micro-avg
	Precision, Recall, F_1-score (%)		
NB	38.52/40.62/39.54	46.80/44.62/45.68	42.98/42.78/42.88
SVM	39.46/26.57/31.76	50.96/65.18/57.20	45.66/47.39/46.51
KNN	46.05/**44.14**/45.07	53.91/55.81/54.84	50.29/50.44/50.36
CNN	50.10/21.62/30.21	54.93/81.60/65.66	52.71/53.97/53.33
RNN	56.01/04.90/09.01	54.35/**96.71**/**69.59**	55.11/54.41/54.76
Our method	**57.98**/31.60/**40.91**	**57.92**/80.43/67.35	**57.95**/**57.94**/**57.94**

section 2.1 mentioned, RNN is expert in sequential processing and CNN is good at extracting features. As a result, the neural network model (i.e., RNN and CNN) can further improve performance to reach about 55% and 53% respectively. In this study, our method combined with multiple CNNs was able to extract latent linguistic features from hotel reviews through the convolution and pooling layers. Multiple dense layers were adopted to refine the discriminative features for identifying the proactive and non-proactive hotel responses. Moreover, we further fused the sentiment and temporal information into our multi-CNN model. Consequently, our model achieved the best precision, recall, and F_1-scores among the compared methods.

5.1. Response and review strategies

Based on our experiment results and visual analyses, we can offer the following strategic recommendations to hotel practitioners and travelers.

Response Strategy 1: We recommend that hotel managers identify types of travelers prior to their responses. Existing studies found that an increase in hotel review ratings can boost hotel reservations (Ye et al., 2011) and high-rated hotels can increase prices (Zimmermann et al., 2018). Further, Liu et al. (2013) mentioned that no prior study had examined the different types of travelers and their expectations. Therefore, we conducted additional visual analyses on types of travelers and aspect ratings. We found that couple travelers tend to rate higher, while business travelers tend to rate lower. Moreover, solo and business travelers care about available business facilities such as computers and printers that allow them to work while staying at the hotel. Family travelers prefer faster check-in and -out experiences and better sleep quality, while friends are more concerned about locations. Based on aspect rating analyses, each hotel can improve its service and facility based on the major source of customers. For example, hotels may provide faster check-in and -out services to all travelers, adjustable firmness of mattresses and pillows to family and business travelers and encourage more couple travelers to leave reviews of their staying experience.

Response Strategy 2: We recommend that hotel managers analyze online reviews actively and respond to negative and positive reviews strategically because online reviews may influence consumers' purchasing intention (Berger et al., 2010) and hotel reputation (Proserpio & Zervas, 2017). It has been found that strategic responses can also increase the perceived helpfulness of online reviews (Liu & Park, 2015). Our visual analytics indicates the hotel managers do not appear to respond to positive and negative reviews strategically. For example, DoubleTree by Hilton Woking has a high non-response rate, which is greater than 70%, regardless of whether responding to positive and negative reviews. We recommend that hotel managers respond to negative reviews timely and actively for the following reasons. First, a timely response to negative reviews can increase travelers' trust (Sparks et al., 2016) and is an important predictor of hotel performance (Kim et al., 2015). Second, after the hotel's response, the length of negative views tends to increase, and the number of negative reviews tends to decrease, which leads to higher hotel ratings (Proserpio & Zervas, 2017).

Response Strategy 3: We recommend that hotel managers strategically identify experienced reviewers and opinion leaders based on reviewer profiles. Kwok and Xie (2016) further point out that responses to opinion leaders' reviews will proactively influence the perceived helpfulness of the reviews. Opinion leaders are important promoters of products and services (Lin et al., 2018), and as such, hotel managers should further analyze the characteristics of reviewers such as years of review, number of positive and negative reviews, and traveler types. In this way, managers can even choose opinion leaders to work with in order to better understand how to promote their hotels to different types of travelers.

Response Strategy 4: Temporal factors and patterns remain under-researched in the current literature. Our analyses indicate that certain months, such as July and August, are more likely to receive more negative reviews. We recommend that hotel managers devote more resources for monitoring and providing timely responses to online reviews (Lui et al., 2018). In addition, response strategies should be adjusted based on the day of the work week, weekends, holidays and seasons to engage different types of travelers. For example, additional response representatives are required to deal with the sudden increase of reviews during holidays, weekends, and special events.

Review Strategy 1: When traveler types are added to the response rate analysis, unknown travelers receive a relatively lower response rate (62%) compared to other traveler types (>78%). Therefore, we recommend that travelers specify a travel type such as a couple, family, or friend if they would like to receive a response from the hotel manager. Moreover, since the time dimension was also added to our analysis, we found that hotel managers tend to respond to reviews during weekdays, typically Mondays, and July was the month with the highest response rate.

6. Conclusion, limitations, and future research

Sun et al. (2017) point out that even though computer vision and speech recognition have greatly improved with deep learning techniques in recent years, deep learning-based NLP is still in its infancy. Although existing studies are valuable in analyses of online reviews and their sentiments, they shed little light on response strategies. This paper addresses this shortcoming by investigating hotel review response strategies by using smart technologies such as deep learning and visual analytics as effective tools that can assist hotel representatives in their decision-making to prioritize responses to reviews.

This data-driven study has yielded theoretical, managerial, and technical contributions. First, this study adds to a growing body of tourism research on hotel response strategies by investigating the intricate relations between hotel reviews and managerial responses. Our study complements the existing research literature by examining traveler types, aspect ratings, review sentiments, temporal factors, and reviewer profiles. Secondly, this study goes beyond ratings and sentiment analysis by analyzing the linguistic features of reviews and responses to empirically identify response strategies. Thirdly, our study complements the existing research methods such as statistical and

表4
三个酒店群。

集群	酒店数量	平均贡献者水平	平均整体评级	平均最高价格	平均最低价格	平均服务等级	平均清洁度等级
1	11	3.09	3.47	216.81	93.70	3.83	3.85
2	24	3.17	4.12	215.15	94.5	4.22	4.50
3	8	**3.34**	**4.35**	**406.11**	**165.83**	**4.41**	**4.61**

表5
比较方法的性能结果。

系统	积极主动的	非主动	微平均
	准确度、召回率、F_1-得分（%）		
NB	38.52/40.62/39.54	46.80/44.62/45.68	42.98/42.78/42.88
SVM	39.46/26.57/31.76	50.96/65.18/57.20	45.66/47.39/46.51
KNN	46.05/**44.14**/45.07	53.91/55.81/54.84	50.29/50.44/50.36
CNN	50.10/21.62/30.21	54.93/81.60/65.66	52.71/53.97/53.33
RNN	56.01/04.90/09.01	54.35/**96.71**/**69.59**	55.11/54.41/54.76
我们的方法	**57.98**/31.60/**40.91**	**57.92**/80.43/67.35	**57.95**/**57.94**/**57.94**

第2.1节提到，RNN擅长顺序处理，CNN擅长提取特征。因此，神经网络模型（即RNN和CNN）可以进一步提高性能，分别达到约55%和53%。在这项研究中，我们的方法与多个CNN相结合，能够通过卷积和合并层从酒店评论中提取潜在的语言特征。采用多个密集层来细化区分特征，以识别主动和非主动酒店响应。此外，我们进一步将情绪和时间信息融合到我们的多CNN模型中。因此，在比较的方法中，我们的模型获得了最好的精确度、召回率和F_1分数。

5. 1. 应对和审查战略

基于我们的实验结果和视觉分析，我们可以为酒店从业者和旅行者提供以下战略建议。

应对策略1：我们建议酒店经理在做出回应之前确定旅行者的类型。现有研究发现，提高酒店评论评级可以提高酒店预订量（Ye等人，2011年），而高评级酒店可以提高价格（Zimmermann等人，2018年）。此外，Liu等人（2013年）提到，之前没有研究调查过不同类型的旅行者及其期望。因此，我们对旅行者类型和方面评级进行了额外的视觉分析。我们发现情侣旅行者的评分更高，而商务旅行者的评分更低。此外，单人和商务旅行者关心可用的商务设施，如电脑和打印机，使他们能够在酒店工作。家庭旅行者更喜欢更快的入住和退房体验以及更好的睡眠质量，而朋友则更关心地点。根据方面评级分析，每家酒店都可以根据主要客户来源改进其服务和设施。例如，酒店可能会为所有旅行者提供更快的入住和退房服务，为家庭和商务旅行者提供可调节的床垫和枕头硬度，并鼓励更多情侣旅行者留下他们的住宿体验评论。

应对策略2：我们建议酒店经理进行分析
在线评论会积极应对负面和正面评论，因为在线评论可能会影响消费者的购买意愿（Berger等人，2010年）和酒店声誉（Proserpio & Zervas, 2017年）。研究发现，战略应对也可以提高在线评论的帮助感（Liu & Park, 2015）。我们的视觉分析表明，酒店经理似乎没有从战略上回应正面和负面评论。例如，希尔顿沃金的DoubleTree有很高的未回复率，无论是否回复正面和负面评论，都超过70%。我们建议酒店经理对

由于以下原因，及时积极地进行负面评论。首先，及时回应负面评论可以增加旅行者的信任（Sparks等人，2016年），也是酒店业绩的重要预测因素（Kim等人，2015年）。其次，在酒店做出回应后，负面评论的长度趋于增加，负面评论的数量趋于减少，这导致酒店评级更高（Proserpio & Zervas, 2017）。应对策略3：我们建议酒店经理根据评审员档案从战略上确定有经验的评审员和意见领袖。郭和谢（2016）进一步指出，对意见领袖评论的回应将积极影响评论的有用性。意见领袖是产品和服务的重要推动者（Lin et al., 2018），因此，酒店经理应进一步分析评论者的特征，如评论年数、正面和负面评论的数量以及旅行者类型。这样，管理者甚至可以选择与意见领袖合作，以便更好地了解如何将他们的酒店推广到不同类型的酒店旅行者。

应对策略4：当前文献中对时间因素和模式的研究仍然不足。我们的分析表明，某些月份，如7月和8月，更可能收到更多负面评论。我们建议酒店经理投入更多资源监控在线评论并及时做出回应（Lui等人，2018年）。此外，应对策略应根据工作日、周末、节假日和季节进行调整，以吸引不同类型的旅行者。例如，需要额外的响应代表来处理假期、周末和特殊活动期间突然增加的评论。

回顾策略1：将旅行者类型添加到响应中时
比率分析表明，与其他类型的旅行者相比（大于78%），未知旅行者收到的回复率相对较低（62%）。因此，如果旅行者希望收到酒店经理的回复，我们建议他们指定旅行类型，如夫妇、家人或朋友。
此外，由于时间维度也被添加到了我们的分析中，我们发现酒店经理倾向于在工作日（通常是周一）回复评论，而7月是回复率最高的月份。

6. 结论、局限性和未来研究

Sun等人（2017年）指出，尽管近年来随着深度学习技术的发展，计算机视觉和语音识别有了很大的进步，但基于深度学习的NLP仍处于起步阶段。尽管现有的研究在分析在线评论及其情绪方面很有价值，但它们对回应策略几乎没有什么帮助。本文通过使用智能技术（如深度学习和视觉分析）来研究酒店评论响应策略，以此来解决这一缺陷，这些智能技术可以帮助酒店代表在决策中优先考虑对评论的响应。

这项数据驱动的研究产生了理论、管理和技术贡献。首先，这项研究通过调查酒店评论和管理层回应之间的复杂关系，为越来越多的关于酒店回应策略的旅游研究增添了新的内容。我们的研究通过检查旅行者类型、方面评级、评论情绪、时间因素和评论人概况来补充现有的研究文献。其次，本研究超越了评级和情绪分析，通过分析评论和回复的语言特征来实证确定回复策略。第三，我们的研究补充了现有的研究方法，如统计学和统计学

econometrics models and surveys by using visual analytics, deep learning, and NLP techniques.

From a managerial perspective, hotel managers can respond to reviews strategically and to maintain a good relationship with customers. Xie et al. (2016) advise hotels to adopt managerial response strategies, and it has been shown that the speed of response, for example, is important to service recovery (Gu & Ye, 2014; Zhao et al., 2019) and can enhance customer engagement (Li et al., 2017). The temporal factor and reviewer profile can be further used for market segmentation and improvement of hotel services and facilities.

The technical contributions of this study are evident. Our proposed analytical framework shows good explanatory power and outperforms existing machine learning methods such as NB, KNN, SVM, CNN, and RNN, and, thus can be expanded to conduct deeper analyses. What sets this study apart is that we presented a novel approach to integrating visual analytics and deep learning-based NLP models to gain insights into various aspects of hotel reviews and responses. Our visual analysis results demonstrate that the use of review sentiment, temporal interval, and reviewer profile together with rating information, can help hotel managers prioritize responses and develop response strategies.

There are, however, limitations to this study. First, the dataset represented only 43 Hilton-affiliated hotels in London. Further, the research findings derived from our visual analytics and the existing study (Park & Allen, 2013) indicate that most hotel managers do not adopt clear response strategies, which leads to a lower performance of deep-learning models. To further examine the performance of deep learning models, we need to include more cities, hotel brands, and hotels in our future studies. Second, the perception from the users' perspective was not studied. It would be instructive to investigate users' perceptions and satisfaction when they read hotel responses to both positive and negative reviews. Despite the limitations addressed above, our study provides new insights to demonstrate the power of integrating visual analytics and deep learning-based NLP to analyze hotel reviews and responses.

Declaration of competing interest

None.

Appendix A. Recent research related to hotel reviews on TripAdvisor from three leading journals

Author (year)	Publication Source	Reviews/ responses	Reviewer profile, temporal factor, travel type, aspect rating	Data analysis/research method	Visualization
Zhang et al. (2020)	*Tourism Management*	Hotel reviews and responses	Not available	Text mining (topic matching), SVM, econometric models, hypothesis testing	Result display: line chart
Wang et al. (2020)	*Tourism Management*	Hotel reviews	Travel type	Term frequency-inverse document frequency (TF-IDF), Word2Vec, ratio and rating analysis	Result display: bar and line chart
Hu, Teichert, et al. (2019) and Hu, Zhang, et al. (2019)	*Tourism Management*	Hotel reviews	Travel type	Structural topic model, text analysis, topic correlation analysis	Result display: network visualization
(Liu et al. (2019))	*Tourism Management*	Hotel reviews	Not available	Sentiment analysis based on a lexicon	Result display: sentiments image
Bi, Liu, Fan, and Zhang (2019)	*Tourism Management*	Hotel reviews	Temporal factor	LDA, SVM, regression model, statistical analysis	Result display: bar and line chart
Taecharungroj and Mathayomchan (2019)	*Tourism Management*	Hotel reviews	Aspect rating	LDA, Naïve Bayes, sentiment analysis	Result display: bar and distribution charts
Hu, Teichert, et al. (2019) and Hu, Zhang, et al. (2019)	*Tourism Management*	Hotel reviews	Travel type, aspect rating	Regression analysis, hypothesis testing, text mining	Result display: line chart
Gao et al. (2018)	*Tourism Management*	Hotel reviews	Reviewer profile, travel type, aspect rating	Logistic model, hypothesis testing	Not available
Lui et al. (2018)	*Tourism Management*	Hotel reviews and responses	Temporal factor	Statistical model, hypothesis testing	Not available
Radojevic et al. (2018)	*Tourism Management*	Hotel reviews	Travel type, aspect ratings, reviewer profile	Statistical model	Result display: line chart, histogram
Marine-Roig and Ferrer-Rosell (2018)	*Tourism Management*	Hotel review titles	Not available	Content analysis, keyword analysis, cognitive analysis	Not available
Liu et al. (2017)	*Tourism Management (Research Note)*	Hotel reviews	Aspect rating	Statistical analysis and data comparison	Result display: distribution chart, line chart
Xiang et al. (2017)	*Tourism Management*	Hotel reviews	Not available	LDA, Naïve Bayes, linear regression model, sentiment analysis	Result display: bar and line chart
Li et al. (2017)	*Tourism Management*	Hotel reviews and responses	Temporal factor	Statistical model, hypothesis testing	Not available
Guo et al. (2017)	*Tourism Management*	Hotel reviews	Aspect rating	LDA, text analysis, statistical analysis	Result display: bar line, and distribution chart
Geetha, Singha, and Sinha (2017)	*Tourism Management*	Hotel reviews		Sentiment analysis, Naïve Bayes, WordNet lexicon, clustering	Result display: bar chart and word cloud
Banerjee and Chua (2016)	*Tourism Management*	Hotel reviews	Travel type	Factorial analysis, ANOVA	Result display: line chart
Baka (2016)	*Tourism Management*	Hotel reviews and responses	Not available	Case study (no data analysis)	Not available
Yang et al. (2016)	*Tourism Management*	Hotel reviews	Temporal factor	Regression model	Result display: location map
de la Peña et al. (2016)	*Tourism Management*	Hotel reviews	Not available	Econometric model	Not available
Zhang and Cole (2016)	*Tourism Management*	Hotel reviews	Not available	Regression model	Not available
Li et al. (2015)	*Tourism Management*	Hotel reviews	Travel type, temporal factor	Statistical analysis	Result display: bar and line chart
Li et al. (2013)	*Tourism Management*	Hotel reviews	Travel type, aspect rating		Not available

(*continued on next page*)

使用可视化分析、深度学习和NLP技术建立计量经济学模型和调查。

从管理的角度来看，酒店经理可以战略性地回应评论，并与客户保持良好的关系。谢等人（2016年）建议酒店采取管理响应策略，例如，响应速度对服务恢复很重要（顾和叶，2014；赵等人，2019），并可以提高客户参与度（李等人，2017年）。时间因素和评审人员概况可进一步用于市场细分和酒店服务和设施的改进。

这项研究的技术贡献是显而易见的。我们提出的分析框架显示了良好的解释能力，优于现有的机器学习方法，如NB、KNN、SVM、CNN和RNN,因此可以扩展以进行更深入的分析。这项研究的与众不同之处在于，我们提出了一种新的方法来整合视觉分析和基于深度学习的NLP模型，以深入了解酒店评论和回应的各个方面。我们的视觉分析结果表明，使用评论情绪、时间间隔、评论人简介以及评级信息可以帮助酒店管理者优先考虑响应并制定响应策略。

然而，这项研究也有局限性。首先，该数据集仅代表伦敦的43家希尔顿附属酒店。此外，我们的视觉分析和现有研究（Park & Allen, 2013）得出的研究结果表明，大多数酒店经理没有采取明确的应对策略，这导致深度学习模型的绩效较低。为了进一步检验深度学习模型的性能，我们需要在未来的研究中纳入更多的城市、酒店品牌和酒店。第二，没有研究用户的感知。当用户阅读酒店对正面和负面评论的反应时，调查他们的看法和满意度将是有益的。尽管存在上述局限性，我们的研究提供了新的见解，以证明整合视觉分析和基于深度学习的NLP来分析酒店评论和回应的能力。

竞合利益的申报

没有。

附录A.三大主流杂志关于TripAdvisor酒店评论的最新研究

作者（年）	出版物来源	评论/回应	时态评论者简介 因素、旅行类型、方面评级	数据分析/研究方法	可视化
张等（2020）	*旅游管理*	酒店评论和回应	不可用	文本挖掘（主题匹配），支持向量机，经济计量模型、假设检验	结果显示：折线图
王等人（2020年）	旅游管理	酒店评论	旅行类型	术语频率逆文档频率（TF-IDF）、Word2Vec、比率和评级分析	结果显示：条形图和折线图
胡泰彻等（2019）和胡章等（2019）	旅游管理	酒店评论	旅行类型	结构主题模型、文本分析、主题相关性分析	结果显示：网络可视化
(Liu et al. (2019))	旅游管理	酒店评论	不可用	基于词典的情感分析	结果显示: 情感意象
毕、刘、范和张（2019年）	旅游管理	酒店评论	时间因素	LDA,支持向量机，回归模型，统计分析	结果显示：条形图和折线图
Taecharungroj和Mathayomchan（2019年）	旅游管理	酒店评论	方面评级	LDA,天真的Bayes,情绪分析	结果显示：条形图和分布图
胡泰彻等（2019）和胡章等（2019）	旅游管理	酒店评论	旅行类型、方面评级	回归分析，假设检验，文本挖掘	结果显示：折线图
Gao et al. (2018)	旅游管理	酒店评论	审核人简介、旅行类型、，方面评级	逻辑模型，假设检验	不可用
Lui et al. (2018)	旅游管理	酒店评论和回复	时间因素	统计模型、假设检验	不可用
Radojevic等人（2018年）	旅游管理	酒店评论	旅行类型，方面评级，审核人档案	统计模型	结果显示：折线图、直方图
Marine Roig和Ferrer Rosell（2018）	旅游管理	酒店评论标题	不可用	内容分析、关键词分析、认知分析	不可用
Liu et al. (2017)	旅游管理（研究报告）	酒店评论	方面评级	统计分析和数据比较	结果显示: 分布图，折线图
项等（2017）	旅游管理	酒店评论	不可用	LDA,朴素贝叶斯，线性回归模型，情绪分析	结果显示：条形图和折线图
Li et al. (2017)	旅游管理	酒店评论和回复	时间因素	统计模型、假设检验	不可用
Guo et al. (2017)	旅游管理	酒店评论	方面评级	LDA,文本分析，统计分析	结果显示：条线，和分布图
吉塔、辛哈和辛哈（2017）	旅游管理	酒店评论		情绪分析、天真贝叶斯、WordNet词汇、聚类	结果显示：条形图和单词云
班纳吉和蔡（2016）	旅游管理	酒店评论	旅行类型	因子分析	结果显示：折线图
巴卡（2016）	旅游管理	酒店评论和回复	不可用	案例研究（无数据分析）	不可用
杨等（2016）	旅游管理	酒店评论	时间因素	回归模型	结果显示：位置地图
de la Pen~a et al. (2016)	旅游管理	酒店评论	不可用	计量经济模型	不可用 张和
科尔（2016）	旅游管理	酒店评论	不可用	回归模型	不可用
Li et al. (2015)	旅游管理	酒店评论	旅行类型、时间因素	统计分析	结果显示：条形图和折线图
Li et al. (2013)	旅游管理	酒店评论	旅行类型、方面评级		不可用

(下一页续)

(*continued*)

Author (year)	Publication Source	Reviews/ responses	Reviewer profile, temporal factor, travel type, aspect rating	Data analysis/research method	Visualization
				Fuzzy logic, aggregation function (Choquet Integral)	
Briggs, Sutherland, and Drummond (2007)	*Tourism Management*	Hotel reviews	Not available	Survey, ANOVA	Result display: bar chart
Radojevic et al. (2019)	*Annals of Tourism Research*	Hotel reviews	Travel type, aspect rating	Regression model	Result display: bar and line chart
Hernández, Kirilenko, and Stepchenkova (2018)	*Annals of Tourism Research*	Hotel reviews	Not available	Network analysis	Result display: Geo-visualization, bar chart
Filieri (2016)	*Annals of Tourism Research*	Hotel reviews	Not available	Interview	Not available
Phillips et al. (2019)	*Journal of Travel Research*	Hotel reviews	Not available	Data mining, sentiment analysis, decision tree	Result display: bar and distribution chart
Yang and Mao (2019)	*Journal of Travel Research*	Hotel reviews	Temporal factor	Econometric model	Not available

(*continued*)

Author (year)	Publication Source	Reviews/ responses	Reviewer profile, temporal factor, sentiment, travel type, aspect rating	Data analysis/research method	Visualization
Yang et al. (2018)	*Journal of Travel Research*	Hotel reviews	Travel type, temporal factor	Logit model, statistical analysis	Result display, spatial distribution
Mkono and Tribe (2017)	*Journal of Travel Research*	Hotel reviews	Not available	Netnographic approach	Not available
Radojevic et al. (2019)	*Annals of Tourism Research*	Hotel reviews	Travel type, aspect rating	Regression model	Result display: bar and line chart

Note: The selected articles are related to hotel reviews or responses on TripAdvisor. We are interested in the research methods, data attributes, and data analyses that have been used in analyzing textual hotel reviews. There are articles published out of Tourism Management, Journal of Travel Research, and Annals of Tourism Research. We also expanded our review to these articles to verify the consistency of our findings.

Appendix B. Recent research related to online reviews from three leading journals

Author (year)	Publication Source	Reviews/responses	Data analysis/research method	Visualization
Li and Ryan (2020)	*Tourism Management*	Online reviews	Concept analysis (Leximancer)	Result display: concept map, bar chart
Woodman et al. (2019)	*Tourism Management*	Lodge reviews	Manual review, bacteria measurement	Not available
Kirilenko et al. (2019)	*Tourism Management*	Attraction reviews	Geographical analysis	Data exploration and result display: map
Liu et al. (2019)	*Tourism Management*	Online reviews	Econometric models	Not available
Mellinas et al. (2019)	*Tourism Management*	Online reviews	Statistical model, hypothesis testing	Result display: line chart
Soler et al. (2019)	*Tourism Management*	Not available	Hedonic pricing model, descriptive analysis, regression model	Not available
Gerdt, Wagner, and Schewe (2019)	*Tourism Management*	Online reviews	Descriptive statistics, regression analysis, sentiment coding	Result display: line chart
Pantano et al. (2017)	*Tourism Management*	Online review of Empire State Building	Machine learning models using Mathematica	Result display: line chart
(Yang et al. (2018))	*Tourism Management*	Not available	Meta-analysis, hierarchical linear modelling, hypothesis testing	Result display: histogram
Liu et al. (2018)	*Tourism Management*	Attraction reviews	Statistical model, hypothesis testing	Result display: line chart
Batista e Silva et al. (2018)	*Tourism Management*	Not available	GIS and spatiotemporal mapping	Result display: density map
Su and Teng (2018)	*Tourism Management*	Negative museums reviews	Content analysis requires manual coding	Not available
Boo and Busser (2018)	*Tourism Management*	Online reviews	Automated and manual content analysis, concept extraction	Result display: concept map
Ganzaroli, De Noni, and van Baalen (2017)	*Tourism Management*	Restaurant reviews	Statistical model, hypothesis testing	Result display: bar chart
	Tourism Management	Hotel reviews		
Sparks et al. (2016)	*Tourism Management*	Negative online reviews, responses	Survey study, simulation, hypotheses testing	Result display: bar chart
Filieri, Alguezaui, and McLeay (2015)	*Tourism Management*	Not available	Hypothesis testing, structural equation modelling (SEM), Web Questionnaire (TripAdvisor)	Not available
Phillips et al. (2015)	*Tourism Management*	Positive online reviews	Artificial neural network, regression analysis	Not available
Lu and Stepchenkova (2012)	*Tourism Management*	Not available	Content analysis, chi-square analysis,	Result display: destination map
Zehrer et al. (2011)		Travel blogs	2x2 quasi experimental design, hypothesis testing	

(*continued on next page*)

(续)

作者（年）	出版物来源	评论/回应	评审员简介、时间因素、旅行类型、方面评级	数据分析/研究方法	可视化
				模糊逻辑，聚合函数（Choquet积分）	
布里格斯、萨瑟兰和德拉蒙德（2007）	旅游管理	酒店评论	不可用	调查，方差分析	结果显示：条形图
Radojevic等人（2019年）	旅游研究纪事	酒店评论	旅行类型、方面评级	回归模型	结果显示：条形图和折线图
赫娜口恩德斯、基里连科和斯特普钦科娃（2018）	*旅游研究纪事*	酒店评论	不可用	网络分析	结果显示：地理可视化，条形图
菲列里（2016）	旅游研究纪事	酒店评论	不可用	采访	不可用
菲利普斯等人（2019年）	旅游研究杂志	酒店评论	不可用	数据挖掘、情绪分析、决策树	结果显示：条形图和分布图
杨和毛（2019）	旅游研究杂志	酒店评论	时间因素	计量经济模型	不可用

(续)

作者（年）	出版物来源	评论/回应	评论者简介、时间因素、情绪、旅行类型、方面评级	数据分析/研究方法	可视化
Yang等人（2018年）	旅游研究杂志	酒店评论	旅行类型、时间因素	罗吉特模型，统计分析	结果显示，空间分发酒店评论
Mkono和部落（2017）	旅行杂志*研究*		不可用	网络图解法	不可用
Radojevic等人（2019年）	旅游研究纪事	酒店评论	旅行类型、方面评级	回归模型	结果显示：条形图和折线图

注：所选文章与TripAdvisor上的酒店评论或回复有关。我们对研究方法、数据属性和数据分析感兴趣，这些都被用于分析文本酒店评论。在《旅游管理》、《旅游研究杂志》和《旅游研究年鉴》上发表了一些文章。我们还将审查范围扩大到这些文章，以验证我们发现的一致性。

附录B.与三种主要期刊的在线评论相关的最新研究

作者（年）	出版物来源	Reviews/responses	数据分析/研究方法	可视化
李和瑞安（2020）	*旅游业管理*	在线评论	概念分析（Leximancer）	结果显示：概念图，条形图不可用
伍德曼等人（2019年）	旅游业*管理*	提交评论	手动检查，细菌测量	
基里连科等人（2019年）	旅游业*管理*	吸引力评论	地理分析	数据探索与结果显示：地图
Liu et al. (2019)	旅游管理	在线评论	计量经济模型	不可用
梅利纳等人（2019年）	旅游管理	在线评论	统计模型、假设检验	结果显示：折线图
Soler等人（2019年）	旅游管理	不可用	特征定价模型、描述性分析、回归模型	不可用
格特、瓦格纳和舍韦（2019年）	*旅游管理*	在线评论	描述性统计、回归分析、情绪编码	结果显示：折线图
潘塔诺等人（2017年）	旅游管理	帝国大厦在线评论	基于Mathematica的机器学习模型	结果显示：折线图
（Yang等人（2018））	旅游管理	不可用	元分析、分层线性建模、假设检验	结果显示：直方图
Liu et al. (2018)	旅游管理	吸引力评论	统计模型、假设检验	结果显示：折线图
巴蒂斯塔·e·席尔瓦等人（2018年）	旅游业*管理*	不可用	地理信息系统与时空制图	结果显示：密度图 负面评论
苏和滕（2018）	旅游管理		内容分析需要手动编码	不可用
Boo和Busser（2018）	旅游管理	在线评论	自动和手动内容分析，概念提取	结果显示：概念图
甘萨罗利、德诺尼和范巴伦（2017年）	*旅游管理旅游管理*	餐厅评论	统计模型、假设检验	结果显示：条形图 酒店
		评论		
Sparks等人（2016年）	旅游管理	负面的在线评论、回应	调查研究、模拟、假设测试	结果显示：条形图
菲列里、阿尔盖扎伊和麦克利（2015年）	*旅游管理*	不可用	假设检验、结构方程建模（SEM）、网络问卷（TripAdvisor）	不可用
菲利普斯等人（2015年）	旅游管理	积极的在线评论	人工神经网络，回归分析	不可用
卢和斯特普琴科娃（2012）	旅游业*管理*	不可用	内容分析，卡方分析，	结果显示：目的地地图
Zehrer et al. (2011)		旅行博客	2x2准实验设计，假设检验	

(下一页续)

(*continued*)

Author (year)	Publication Source	Reviews/responses	Data analysis/research method	Visualization
	Tourism Management			Result display: plot with Box & Whisker
Amaro, Duarte, and Henriques (2016)	*Annals of Tourism Research*	Not available	Questionnaire, k-means cluster analysis	Result display: box plot and line chart
Mkono (2016)	*Annals of Tourism Research*	Online reviews	Hermeneutic interpretation, netnography	Not available
Alaei et al. (2019)	*Journal of Travel Research*	Multiple tourism online reviews	Sentiment analysis (a review paper)	Not available
Gal-Tzur, Bar-Lev, and Shiftan (2019)	*Journal of Travel Research*	City forums	Questionnaire, statistical model	Not available
Bigné, William, and Soria-Olivas (2019)	*Journal of Travel Research*	Online reviews	Self-organizing Maps (SOM), in-depth interview, sentiment analysis	Result display: SOM and line chart
Fu, Hao, Robert, Li, and Hsu (2019)	*Journal of Travel Research*	Chinese travel news	Sentiment analysis	Result display: line chart
Vu et al. (2019)	*Journal of Travel Research*	Restaurant reviews	Text mining, sentiment analysis, statistical analysis	Result display: bar chart
Stamolampros et al. (2019)	*Journal of Travel Research*	Flight reviews	NLP, LDA, topic modelling, logistic regression	Not available
Gkritzali, Gritzalis, and Stavrou (2018)	*Journal of Travel Research*	City forum	Sentiment analysis	Result display: bar chart
Kirilenko et al. (2018)	*Journal of Travel Research*	Online reviews	Sentiment analysis	Not available
Phillips et al. (2017)	*Journal of Travel Research*	Online reviews	Hypotheses testing, statistical modelling	Result display: bar chart
Kim and Fesenmaier (2017)	*Journal of Travel Research*	Negative online reviews	Hypotheses testing, statistical analysis	Result display: bar chart
Murphy and Chen (2016)	*Journal of Travel Research*	Online reviews	Questionnaires, exploratory, observation method	Result display: bar chart
Tanford and Montgomery (2015)	*Journal of Travel Research*	Online reviews	Hypothesis testing, questionnaires, statistical analysis	Result display: line chart
Kazeminia, Del Chiappa, and Jafari (2015)	*Journal of Travel Research*	Online reviews	Content analysis, Leximancer, thematic analysis and semantic analysis	Result display: concept map

Note: The selected articles are related online textual reviews, e.g., restaurant reviews, travel reviews, and airline reviews. We are interested in the research methods and data analyses that have been used in analyzing textual reviews in tourism research. There are articles published out of Tourism Management, Journal of Travel Research, and Annals of Tourism Research. We also expanded our review to these articles to verify the consistency of our findings.

Author contributions

C.H. Chen conducted data collection; Y.C. Chang and C.H. Ku developed models, conducted data analyses, experiments, and wrote the manuscript.

Funding

This project was partially supported by the Ministry of Science and Technology of Taiwan under grant MOST 107-2410-H-038-017-MY3 and MOST 109-2634-F-001-008 for Dr. Yung-Chun Chang.

References

Al-Smadi, M., Qawasmeh, O., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2018). Deep recurrent neural network vs. Support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *Journal of Computational Science, 27*, 386–393. https://doi.org/10.1016/j.jocs.2017.11.006.

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on big data. *Journal of Travel Research, 58*(2), 175–191. https://doi.org/10.1177/0047287517747753.

Amaro, S., Duarte, P., & Henriques, C. (2016). Travelers' use of social media: A clustering approach. *Annals of Tourism Research, 59*, 1–15. https://doi.org/10.1016/j.annals.2016.03.007.

Baka, V. (2016). The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector. *Tourism Management, 53*, 148–162. https://doi.org/10.1016/j.tourman.2015.09.004.

Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management, 53*, 125–131. https://doi.org/10.1016/j.tourman.2015.09.020.

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management, 68*, 101–115. https://doi.org/10.1016/j.tourman.2018.02.020.

Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science, 29*(5), 815–827. https://doi.org/10.1287/mksc.1090.0557.

Bigné, E., William, E., & Soria-Olivas, E. (2019). Similarity and consistency in hotel online ratings across platforms. *Journal of Travel Research.* , 0047287519859705. https://doi.org/10.1177/0047287519859705.

Bi, J.-W., Liu, Y., Fan, Z.-P., & Zhang, J. (2019). Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews. *Tourism Management, 70*, 460–478. https://doi.org/10.1016/j.tourman.2018.09.010.

Boo, S., & Busser, J. A. (2018). Meeting planners' online reviews of destination hotels: A twofold content analysis approach. *Tourism Management, 66*, 287–301. https://doi.org/10.1016/j.tourman.2017.11.014.

Briggs, S., Sutherland, J., & Drummond, S. (2007). Are hotels serving quality? An exploratory study of service quality in the scottish hotel sector. *Tourism Management, 28*(4), 1006–1019. https://doi.org/10.1016/j.tourman.2006.08.015.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research [review article]. *IEEE Computational Intelligence Magazine, 9*(2), 48–57. https://doi.org/10.1109/MCI.2014.2307227.

Casaló, L. V., Flavián, C., Guinalíu, M., & Ekinci, Y. (2015a). Do online hotel rating schemes influence booking behaviors? *International Journal of Hospitality Management, 49*, 28–36. https://doi.org/10.1016/j.ijhm.2015.05.005.

Casaló, L. V., Flavián, C., Guinalíu, M., & Ekinci, Y. (2015b). Avoiding the dark side of positive online consumer reviews: Enhancing reviews' usefulness for high risk-averse travelers. *Journal of Business Research, 68*(9), 1829–1835. https://doi.org/10.1016/j.jbusres.2015.01.010.

Chaturvedi, I., Cambria, E., & Vilares, D. (2016). Lyapunov filtering of objectivity for Spanish sentiment model. In *2016 international joint conference on neural networks (IJCNN)* (pp. 4474–4481). https://doi.org/10.1109/IJCNN.2016.7727785.

Cheng, X., Fu, S., Sun, J., Bilgihan, A., & Okumus, F. (2019). *An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust.*

Chen, T., Xu, R., He, Y., & Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications, 72*, 221–230. https://doi.org/10.1016/j.eswa.2016.10.065.

Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th*

(续)

作者（年）	出版物来源	Reviews/responses	数据分析/研究方法	可视化
	旅游管理年鉴			结果显示：用方框打印 &胡须
阿马罗、杜阿尔特和亨里克 (2016)	*研究*	不可用	问卷调查，k-均值聚类分析	结果显示：方框图和折线图
Mkono (2016)	旅游研究纪事	在线评论	解释学解释	不可用
Alaei et al. (2019)	旅游研究杂志	多个旅游在线评论	情绪分析（综述）	不可用
Gal Tzur、Bar Lev和Shiftan（2019年）	*旅游研究杂志*	城市论坛	问卷调查，统计模型	不可用
比根□e、 威廉，还有索里亚·奥利瓦斯（2019）	*旅游研究杂志*	在线评论	自组织地图（SOM）、深度访谈、情绪分析	结果显示：SOM和折线图
傅、郝、罗伯特、李和徐（2019年）	*旅游研究杂志*	中国旅游新闻	情绪分析	结果显示：折线图
Vu et al. (2019)	旅游研究杂志	餐厅评论	文本挖掘、情绪分析、统计分析	结果显示：条形图飞行回
Stamolampros等人（2019年）	旅行杂志 *研究*	顾	NLP、LDA、主题建模、逻辑回归	不可用
Gkritzali、Gritzalis和Stavrou（2018年）	*旅游研究杂志*	城市论坛	情绪分析	结果显示：条形图
基里连科等人（2018年）	旅游研究杂志	在线评论	情绪分析	不可用
菲利普斯等人（2017年）	旅游研究杂志			
金和费森梅尔（2017）	旅行杂志 *研究*	在线评论	假设检验、统计建模	结果显示：条形图负面在
墨菲和陈（2016）	旅行杂志 *研究*	线评论	假设检验、统计分析	结果显示：条形图在线评
		论	问卷调查，探索性，观察法	结果显示：条形图
坦福德和蒙哥马利（2015）	*旅游研究杂志*	在线评论	假设检验、问卷调查、统计分析	结果显示：折线图
卡泽米尼亚、德尔基亚帕和贾法里（2015年）	*旅游研究杂志*	在线评论	内容分析、词汇量分析、主题分析和语义分析	结果显示：概念图

注：所选文章是相关的在线文本评论，例如餐厅评论、旅行评论和航空公司评论。我们对旅游研究中用于分析文本评论的研究方法和数据分析感兴趣。在《旅游管理》、《旅游研究杂志》和《旅游研究年鉴》上发表了一些文章。我们还将审查范围扩大到这些文章，以验证我们发现的一致性。

作者投稿

陈春晖进行了数据收集；张元庆和辜振华开发了模型，进行了数据分析、实验，并撰写了手稿。

资金来源

该项目得到了台湾科技部的部分支持，获得了MOST 107-2410-H-038-017-MY3和MOST 109-2634-F-001-008的资助。

参考文献

Al Smadi，M., Qawasmeh, O., Al Ayyoub，M., Jararweh, Y.,和Gupta, B.（2018）。深度递归神经网络与支持向量机用于基于方面的阿拉伯酒店评论情绪分析。计算科学杂志，27386-393。 https://doi.org/10.1016/j.jocs.2017.11.006.

Alaei, A. R., Becken, S.,和Stantic, B.（2019年）。旅游业情绪分析：利用大数据。旅游研究杂志，58（2），175–191。 https://doi.org/10.1177/ 0047287517747753.

阿马罗，S.,杜阿尔特，P.,和亨利克斯，C.（2016）。旅行者对社交媒体的使用：一种聚类方法。《旅游研究年鉴》，59, 1–15。 https://doi.org/10.1016/j. annals.2016.03.007.

Baka, V.（2016年）。用户生成评论的形成：回顾过去，了解未来在旅游行业管理声誉。旅游管理，53148-162。 https://doi.org/10.1016/j.tourman.2015.09.004.

Banerjee, S.,和Chua, A. Y. K.（2016）。在TripAdvisor中搜索旅行者酒店评级的模式。旅游管理，53125-131。 https://doi.org/10.1016/j. tourman.2015.09.020.

巴蒂斯塔·e·席尔瓦·F.,玛丽安·埃雷拉，M. A.,罗西纳，K.,里贝罗·巴兰科，R.,弗雷尔，S.,和斯齐亚维纳，M.（2018）。利用传统和大数据源，以高分辨率分析欧洲旅游业的时空模式。旅游管理，68101-115。 https://doi.org/10.1016/j.tourman.2018.02.020.

伯杰，J.,索伦森，A. T.,和拉斯穆森，S. J.（2010）。负面宣传的积极影响：当负面评论增加销售额时。营销科学，29（5），815–827。 https://doi.org/10.1287/mksc.1090.0557.

比根□e、 E.，William, E.，和Soria Olivas，E.（2019年）。酒店的相似性与一致性跨平台在线评级。旅游研究杂志,47287519859705 https://doi.org/10.1177/0047287519859705.

毕俊伟、刘亦永、范志平和张俊杰（2019）。群体智慧：通过在线评论进行重要绩效分析（IPA）。旅游管理，70460-478。 https://doi.org/10.1016/j.tourman.2018.09.010.

Boo, S.,和Busser, J. A.（2018）。会议策划人员对目的地酒店的在线评论：一种双重内容分析方法。旅游管理，66287-301。 https://doi. org/10.1016/j.tourman.2017.11.014.

Briggs, S., Sutherland, J.,和Drummond, S.（2007）。酒店服务质量好吗？苏格兰酒店业服务质量的探索性研究。旅游管理，28（4），1006–1019。 https://doi.org/10.1016/j.tourman.2006.08.015.

坎布里亚，E.,怀特，B.（2014）。跳跃NLP曲线：自然语言处理研究综述[综述文章]。IEEE计算智能杂志，9（2），48–57。 https://doi.org/10.1109/MCI.2014.2307227.

卡萨洛□, L. V.,弗拉维亚□n、 C., Guinalíu, M.,和Ekinci, Y.（2015a）。进行在线酒店评级计划会影响预订行为吗？《国际酒店管理杂志》，49、28–36。 https://doi.org/10.1016/j.ijhm.2015.05.005.

卡萨洛□, L. V.,弗拉维亚□n、 C., Guinalíu, M.,和Ekinci, Y.（2015b）。避免黑暗的一面积极的在线消费者评论：增强评论对高风险厌恶旅行者的有用性。商业研究杂志，68（9），1829–1835年。 https://doi.org/10.1016/j. jbusres.2015.01.010.

查图维迪，I.,坎布里亚，E.,和维拉雷斯，D.（2016）。西班牙情绪模型客观性的李雅普诺夫滤波。2016年国际神经网络联合会议（IJCNN）（第4474–4481页）。 https://doi.org/10.1109/IJCNN.2016.7727785.

郑，X.,傅，S.,孙，J.,比尔吉汉，A.,和奥库姆斯，F.（2019年）。共享经济驱动的酒店平台在线评论调查：信任的观点。

陈婷、徐若瑞、何妍、王霞（2017）。使用BiLSTM CRF和CNN通过句子类型分类改进情绪分析。专家系统与应用，72221–230。 https://doi.org/10.1016/j.eswa.2016.10.065.

Collobert, R.,和Weston, J.（2008）。自然语言处理的统一架构：具有多任务学习的深度神经网络。在25日的会议记录中

international conference on machine learning (pp. 160–167). https://doi.org/10.1145/1390156.1390177.

De Pelsmacker, P., van Tilburg, S., & Holthof, C. (2018). Digital marketing strategies, online reviews and hotel performance. *International Journal of Hospitality Management, 72*, 47–55. https://doi.org/10.1016/j.ijhm.2018.01.003.

Dickinger, A. (2011). The trustworthiness of online channels for experience- and goal-directed search tasks. *Journal of Travel Research, 50*(4), 378–391. https://doi.org/10.1177/0047287510371694.

Do, H. H., Prasad, P., Maag, A., & Alsadoon, A. (2019). Deep learning for aspect-based sentiment analysis: A comparative review. *Expert Systems with Applications, 118*, 272–299. https://doi.org/10.1016/j.eswa.2018.10.003.

Dragoni, M., Federici, M., & Rexha, A. (2018). *An unsupervised aspect extraction strategy for monitoring real-time reviews stream*. Information Processing & Management. https://doi.org/10.1016/j.ipm.2018.04.010.

Dreisbach, C., Koleck, T. A., Bourne, P. E., & Bakken, S. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics, 125*, 37–46. https://doi.org/10.1016/j.ijmedinf.2019.02.008.

Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM, 57*(6), 74–81. https://doi.org/10.1145/2602574.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management, 52*, 498–506. https://doi.org/10.1016/j.tourman.2015.07.018.

Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research, 58*, 46–64. https://doi.org/10.1016/j.annals.2015.12.019.

Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management, 51*, 174–185. https://doi.org/10.1016/j.tourman.2015.05.007.

Franke, B., Plante, J.-F., Roscher, R., Lee, E. A., Smyth, C., Hatefi, A., et al. (2016). Statistical inference, learning and models in big data. *International Statistical Review, 84*(3), 371–389. https://doi.org/10.1111/insr.12176.

Fu, Y., Hao, J.-X., Robert, Li, X., & Hsu, C. H. C. (2019). Predictive accuracy of sentiment analytics for tourism: A metalearning perspective on Chinese travel news. *Journal of Travel Research, 58*(4), 666–679. https://doi.org/10.1177/0047287518772361.

Gal-Tzur, A., Bar-Lev, S., & Shiftan, Y. (2019). Using question & answer forums as a platform for improving transport-related information for tourists. *Journal of Travel Research.* , 0047287519877254. https://doi.org/10.1177/0047287519877254.

Ganzaroli, A., De Noni, I., & van Baalen, P. (2017). Vicious advice: Analyzing the impact of tripadvisor on the quality of restaurants as part of the cultural heritage of venice. *Tourism Management, 61*, 501–510. https://doi.org/10.1016/j.tourman.2017.03.019.

Gao, B., Li, X., Liu, S., & Fang, D. (2018). How power distance affects online hotel ratings: The positive moderating roles of hotel chain and reviewers' travel experience. *Tourism Management, 65*, 176–186. https://doi.org/10.1016/j.tourman.2017.10.007.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels - an empirical analysis. *Tourism Management, 61* (Supplement C), 43–54. https://doi.org/10.1016/j.tourman.2016.12.022.

Gerdt, S.-O., Wagner, E., & Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tourism Management, 74*, 155–172. https://doi.org/10.1016/j.tourman.2019.02.010.

Get Started | Geocoding API. (2018). *Google developers*. https://developers.google.com/maps/documentation/geocoding/start.

Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019). Using social media to identify tourism attractiveness in six Italian cities. *Tourism Management, 72*, 306–312. https://doi.org/10.1016/j.tourman.2018.12.007.

Giglio, S., Pantano, E., Bilotta, E., & Melewar, T. C. (2020). Branding luxury hotels: Evidence from the analysis of consumers' "big" visual data on TripAdvisor. *Journal of Business Research*. https://doi.org/10.1016/j.jbusres.2019.10.053.

Gkritzali, A., Gritzalis, D., & Stavrou, V. (2018). Is xenios zeus still alive? Destination image of athens in the years of recession. *Journal of Travel Research, 57*(4), 540–554. https://doi.org/10.1177/0047287517705225.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management, 59*, 467–483. https://doi.org/10.1016/j.tourman.2016.09.009.

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2006). Using kNN model for automatic text categorization. *Soft Computing, 10*(5), 423–430. https://doi.org/10.1007/s00500-005-0503-y.

Gu, B., & Ye, Q. (2014). First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management, 23*(4), 570–582. https://doi.org/10.1111/poms.12043.

Hernández-Ortega, B. (2018). Don't believe strangers: Online consumer reviews and the role of social psychological distance. *Information & Management, 55*(1), 31–50. https://doi.org/10.1016/j.im.2017.03.007.

Hernández, J. M., Kirilenko, A. P., & Stepchenkova, S. (2018). Network approach to tourist segmentation via user generated content. *Annals of Tourism Research, 73*, 35–47. https://doi.org/10.1016/j.annals.2018.09.002.

Hilton London. (2017). *Find hilton hotels in London - UK*. http://www.hilton.com/top-destinations/london-hotels.

Hilton Travel. (2018). Hilton. https://travel.hilton.com/.

Hu, M. (2018). *Visualization of textual content from social media and online communities [Georgia Institute of Technology]*. https://smartech.gatech.edu/handle/1853/59828.

Huang, C. D., Goo, J., Nam, K., & Yoo, C. W. (2017). Smart tourism technologies in travel planning: The role of exploration and exploitation. *Information & Management, 54*(6), 757–770. https://doi.org/10.1016/j.im.2016.11.010.

Huang, J., Li, J., Yu, D., Deng, L., & Gong, Y. (2013). Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 7304–7308). https://doi.org/10.1109/ICASSP.2013.6639081.

Hudson, S., & Thal, K. (2013). The impact of social media on the consumer decision process: Implications for tourism marketing. *Journal of Travel & Tourism Marketing, 30*(1–2), 156–160. https://doi.org/10.1080/10548408.2013.751276.

Hu, F., Teichert, T., Liu, Y., Li, H., & Gundyreva, E. (2019). Evolving customer expectations of hospitality services: Differences in attribute effects on satisfaction and Re-Patronage. *Tourism Management, 74*, 345–357. https://doi.org/10.1016/j.tourman.2019.04.010.

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management, 72*, 417–426. https://doi.org/10.1016/j.tourman.2019.01.002.

Julia, M. (2018). *Big cities, big business: Bangkok, London and Paris lead the way in mastercard's 2018 global destination cities Index [credit card and travel]. MasterCard social newsroom*. https://newsroom.mastercard.com/press-releases/big-cities-big-business-bangkok-london-and-paris-lead-the-way-in-mastercards-2018-global-destination-cities-index/.

Kazeminia, A., Del Chiappa, G., & Jafari, J. (2015). Seniors' travel constraints and their coping strategies. *Journal of Travel Research, 54*(1), 80–93. https://doi.org/10.1177/0047287513506290.

Kim, Y. (2014). *Convolutional neural networks for sentence classification*. ArXiv:1408.5882 [Cs] http://arxiv.org/abs/1408.5882.

Kim, J., Jamie), & Fesenmaier, D. R. (2017). Sharing tourism experiences: The posttrip experience. *Journal of Travel Research, 56*(1), 28–40. https://doi.org/10.1177/0047287515620491.

Kim, W. G., Lim, H., & Brymer, R. A. (2015). The effectiveness of managing social media on hotel performance. *International Journal of Hospitality Management, 44* (Supplement C), 165–171. https://doi.org/10.1016/j.ijhm.2014.10.014.

Kim, K., Park, O., Yun, S., & Yun, H. (2017). What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management. *Technological Forecasting and Social Change, 123*, 362–369. https://doi.org/10.1016/j.techfore.2017.01.001.

Kim, H., & Stepchenkova, S. (2015). Effect of tourist photographs on attitudes towards destination: Manifest and latent content. *Tourism Management, 49*, 29–41. https://doi.org/10.1016/j.tourman.2015.02.004.

King, T. (2018). *Tableau unveils new natural language Processing and automated data Prep. Best business intelligence and data analytics tools, software, solutions & vendors*. https://solutionsreview.com/business-intelligence/tableau-unveils-new-natural-language-processing-and-automated-data-prep/. October 23.

King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing, 28*(3), 167–183. https://doi.org/10.1016/j.intmar.2014.02.001.

Kirilenko, A. P., Stepchenkova, S. O., & Hernandez, J. M. (2019). Comparative clustering of destination attractions for different origin markets with network and spatial analyses of online reviews. *Tourism Management, 72*, 400–410. https://doi.org/10.1016/j.tourman.2019.01.001.

Kirilenko, A. P., Stepchenkova, S. O., Kim, H., Li, X., & Robert,). (2018). Automated sentiment analysis in tourism: Comparison of approaches. *Journal of Travel Research, 57*(8), 1012–1025. https://doi.org/10.1177/0047287517729757.

Ku, C. H., Chang, Y.-C., Wang, Y., Chen, C.-H., & Hsiao, S.-H. (2019, January 8). Artificial intelligence and visual analytics: A deep-learning approach to analyze hotel reviews & responses. In *52nd Hawaii international conference on system sciences, maui, Hawaii*. http://scholarspace.manoa.hawaii.edu/handle/10125/59963.

Kwok, L., & Xie, K. L. (2016). Factors contributing to the helpfulness of online hotel reviews: Does manager response play a role? *International Journal of Contemporary Hospitality Management, 28*(10), 2156–2177. https://doi.org/10.1108/IJCHM-03-2015-0107.

Ladhari, R., & Michaud, M. (2015). eWOM effects on hotel booking intentions, attitudes, trust, and website perceptions. *International Journal of Hospitality Management, 46*, 36–45. https://doi.org/10.1016/j.ijhm.2015.01.010.

Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent convolutional neural networks for text classification. In *Proceedings of the twenty-ninth AAAI conference on artificial intelligence* (pp. 2267–2273). http://dl.acm.org/citation.cfm?id=2886521.2886636.

Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research, 75*, 410–423. https://doi.org/10.1016/j.annals.2019.01.014.

Lee, C. H., & Cranage, D. A. (2012). Toward understanding consumer processing of negative online word-of-mouth communication: The roles of opinion consensus and organizational response strategies. *Journal of Hospitality & Tourism Research*. https://doi.org/10.1177/1096348012451455.

Lee, J. Y., & Dernoncourt, F. (2016). *Sequential short-text classification with recurrent and convolutional neural networks*. ArXiv:1603.03827 [Cs, Stat] http://arxiv.org/abs/1603.03827.

Leung, D., Law, R., Hoof, H. van, & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. *Journal of Travel & Tourism Marketing, 30*(1–2), 3–22. https://doi.org/10.1080/10548408.2013.750919.

Levy, S. E., Duan, W., & Boo, S. (2013). An analysis of one-star online reviews and responses in the Washington, D.C., lodging market. *Cornell Hospitality Quarterly, 54* (1), 49–63. https://doi.org/10.1177/1938965512464513.

国际机器学习会议（第160–167页）。 https://doi.org/10.1145/ 1390156.1390177.

Pelsmacker, P., van Tilburg, S.,和Holthof, C.（2018年）。数字营销策略、在线评论和酒店绩效。《国际酒店管理杂志》，72, 47–55。 https://doi.org/10.1016/j.ijhm.2018.01.003.

Dickinger, A.（2011）。在线渠道对于经验和目标导向搜索任务的可信度。《旅游研究杂志》，50（4），378–391。 https://doi.org/ 10.1177/0047287510371694.

Do, H. H., Prasad, P., Maag, A.,和Alsadoon, A.（2019年）。基于方面的情绪分析的深度学习：比较综述。专家系统与应用，118272–299。 https://doi.org/10.1016/j.eswa.2018.10.003.

Dragoni, M., Federici, M.,和Rexha, A.（2018）。一种用于监控实时评论流的无监督方面提取策略。信息处理与管理。 https://doi.org/10.1016/j.ipm.2018.04.010.

Dreisbach, C., Koleck, T. A., Bourne, P. E.,和Bakken, S.（2019年）。对自然语言处理和从电子患者编写的文本数据中挖掘症状的文本进行系统回顾。《国际医学信息学杂志》，125, 37–46。 https:// doi.org/10.1016/j.ijmedinf.2019.02.008.

范文和戈登医学博士（2014）。社交媒体分析的力量。ACM通讯，57（6），74–81。 https://doi.org/10.1145/2602574.

方，B.,叶，Q.,库库斯塔，D.,罗，R.（2016）。在线旅游评论感知价值分析：可读性和评论人特征的影响。
旅游管理，52498-506。 https://doi.org/10.1016/j. tourman.2015.07.018.

菲列里，R.（2016）。什么使在线消费者评论可信？《旅游研究年鉴》，58, 46–64。 https://doi.org/10.1016/j.annals.2015.12.019.

菲列里，R.,阿尔盖扎伊，S.,和麦克莱伊，F.（2015）。为什么旅行者信任TripAdvisor？消费者对媒体信任的前因及其对推荐采纳和口碑的影响。旅游管理，51174-185。 https://doi.org/10.1016/j.tourman.2015.05.007.

Franke, B., Plante, J.-F., Roscher, R., Lee, E. A., Smyth, C., Hatefi, A.,等人（2016年）。大数据中的统计推断、学习和模型。《国际统计评论》，84（3），371–389。 https://doi.org/10.1111/insr.12176.

傅，Y.,郝，J-X.,罗伯特，李，X.,和许春华（2019）。旅游情绪分析的预测准确性：中国旅游新闻的元学习视角。旅游研究杂志，58（4），666–679。 https://doi.org/10.1177/0047287518772361.

Gal Tzur, A., Bar Lev, S.,和Shiftan, Y.（2019年）。利用问答论坛作为平台,为游客改善交通相关信息。旅游研究杂志,47287519877254 https://doi.org/10.1177/0047287519877254.

Ganzaroli, A., De Noni, I.,和van Baalen, P.（2017年）。恶意建议：作为威尼斯文化遗产的一部分，分析tripadvisor对餐厅质量的影响。旅游管理，61501-510。 https://doi.org/10.1016/j. tourman.2017.03.019.

高B.李X.刘S.方D.（2018）。权力距离如何影响在线酒店评级：连锁酒店和评论者的旅行体验的积极调节作用。
旅游管理，65176-186。 https://doi.org/10.1016/j. tourman.2017.10.007.

吉塔，M.,辛哈，P.,和辛哈，S.（2017）。酒店顾客情绪与在线顾客评分的关系实证分析。旅游管理，61（补编c），43–54。 https://doi.org/10.1016/j.tourman.2016.12.022.

Gerdt,S-O.，Wagner, E.,和Schewe, G.（2019年）。酒店业可持续性与顾客满意度之间的关系：以eWOM为数据源的探索性调查。旅游管理，74155-172。 https://doi.org/10.1016/j. tourman.2019.02.010.

开始|地理编码API。(2018). 谷歌开发者。 https://developers.google.com/maps/documentation/geocoding/start.

吉廖，S.,贝尔塔奇尼，F.,比洛塔，E.,和潘塔诺，P.（2019年）。利用社交媒体识别意大利六个城市的旅游吸引力。旅游管理，72306-312。 https://doi.org/10.1016/j.tourman.2018.12.007.

吉廖，S.,潘塔诺，E.,比洛塔，E.,和梅勒瓦尔，T. C.（2020）。豪华酒店品牌化：来自TripAdvisor上消费者“大”视觉数据分析的证据。商业研究杂志。 https://doi.org/10.1016/j.jbusres.2019.10.053.

Gkritzali, A., Gritzalis, D.,和Stavrou, V.（2018年）。塞尼奥斯·宙斯还活着吗？雅典在经济衰退时期的目的地形象。《旅游研究杂志》，57（4），540–554。 https://doi.org/10.1177/0047287517705225.

郭，Y.,巴恩斯，S. J.,贾，Q.（2017）。从在线评分和评论中挖掘意义：使用潜在dirichlet分配的游客满意度分析。旅游管理，59467-483。 https://doi.org/10.1016/j.tourman.2016.09.009.

郭，G.,王，H.,贝尔，D.,毕，Y.,和格里尔，K.（2006）。使用kNN模型进行自动文本分类。软计算，10（5），423–430。 https://doi.org/10.1007/ s00500-005-0503-y.

顾斌、叶问（2014）。社交媒体的第一步：衡量在线管理回应对客户满意度的影响。《生产和运营管理》，23（4），570–582。 https://doi.org/10.1111/poms.12043.

赫恩□andez Ortega，B.（2018年）。不要相信陌生人：在线消费者评论和社会心理距离的作用。信息与管理，55（1），31–50。 https://doi.org/10.1016/j.im.2017.03.007.

赫恩□andez，J. M., Kirilenko, A. P.,和Stepchenkova, S.（2018）。网络教学法通过用户生成的内容进行游客细分。旅游研究年鉴，73, 35–47。 https://doi.org/10.1016/j.annals.2018.09.002.

伦敦希尔顿酒店。(2017). 在英国伦敦找到希尔顿酒店。 http://www.hilton.com/top-destinations/london-hotels.

希尔顿旅行社。(2018). 希尔顿。 https://travel.hilton.com/.

胡敏（2018）。社交媒体和在线社区文本内容的可视化[乔治亚理工学院]。 https://smartech.gatech.edu/handle/1853/59828.

黄成德、顾俊杰、南坤和柳成伟（2017）。智能旅游技术在旅游规划中的作用：探索与开发。《信息与管理》，54（6），757–770。 https://doi.org/10.1016/j.im.2016.11.010.

黄，J.,李，J.,于，D.,邓，L.,龚，Y.（2013）。使用具有共享隐藏层的多语言深层神经网络进行跨语言知识转移。2013年IEEE声学、语音和信号处理国际会议（第7304–7308页）。 https://doi.org/10.1109/ICASSP.2013.6639081.

哈德逊，S.,和塔尔，K.（2013）。社交媒体对消费者决策过程的影响：对旅游营销的影响。《旅游与旅游营销杂志》，30（1–2），156–160。 https://doi.org/10.1080/10548408.2013.751276.

胡，F.,泰切特，T.,刘，Y.,李，H.,和冈迪列娃，E.（2019）。不断变化的顾客对酒店服务的期望：不同属性对满意度和再次光顾的影响。旅游管理，74345-357。 https://doi.org/10.1016/j. tourman.2019.04.010.

胡恩泰、张铁通、高宝波和百色一世（2019年）。酒店顾客抱怨什么？使用结构化主题模型进行文本分析。旅游管理，72417-426。 https://doi.org/10.1016/j.tourman.2019.01.002.

Julia, M.（2018）。大城市、大企业：曼谷、伦敦和巴黎在万事达卡2018年全球目的地城市指数[信用卡和旅游]中领先。万事达卡社交新闻编辑室。 https://newsroom.mastercard.com/press-releases/big-cities-big-b usiness-bangkok-london-and-paris-lead-the-way-in-mastercards-2018-global-destination-cities-index/.

Kazeminia, A., Del Chiappa, G.,和Jafari, J.（2015）。老年人的旅行限制及其应对策略。旅游研究杂志，54（1），80–93。 https://doi.org/10.1177/ 0047287513506290.

Kim, Y.（2014）。用于句子分类的卷积神经网络。ArXiv:1408.5882 [Cs] http://arxiv.org/abs/1408.5882.

Kim, J., Jamie）和Fesenmaier, D. R.（2017）。分享旅游体验：旅行后的体验。旅游研究杂志，56（1），28–40。 https://doi.org/10.1177/ 0047287515620491.

Kim, W. G., Lim, H.,和Brymer, R. A.（2015）。管理社交媒体对酒店绩效的有效性。《国际酒店管理杂志》，44（补编c），165–171。 https://doi.org/10.1016/j.ijhm.2014.10.014.

Kim, K., Park, O., Yun, S.,和Yun, H.（2017）。是什么让游客对旅游目的地感到消极？混合文本挖掘方法在智能目的地管理中的应用。《技术预测与社会变革》，123362-369。 https://doi.org/10.1016/j.techfore.2017.01.001.

Kim, H.,和Stepchenkova, S.（2015）。旅游照片对目的地态度的影响：显性和隐性内容。旅游管理，49, 29–41。 https:// doi.org/10.1016/j.tourman.2015.02.004.

金·T.（2018）。Tableau推出新的自然语言处理和自动化数据准备。最佳商业智能和数据分析工具、软件、解决方案和供应商。 https://solutionsreview.com/business-intelligence/tableau-unveils-new-natural-language-processing-and-automated-data-prep/.10月23日。

金·R·A.、拉赫拉·P.、布什·V·D.（2014）。关于网络口碑，我们知道什么和不知道什么：文献综述和综合。《互动营销杂志》，28（3），167–183。 https://doi.org/10.1016/j.intmar.2014.02.001.

Kirilenko, A. P., Stepchenkova, S. O.,和Hernandez, J. M.（2019）。通过网络和在线评论的空间分析，对不同来源市场的目的地景点进行比较聚类。旅游管理，72400-410。 https://doi.org/ 10.1016/j.tourman.2019.01.001.

Kirilenko, A. P., Stepchenkova, S. O., Kim, H., Li, X.,和Robert,）。(2018). 旅游业中的自动情绪分析：方法比较。《旅游研究杂志》，57（8），1012–1025。 https://doi.org/10.1177/0047287517729757.

古春晖、张春晖、王春晖、陈春晖和萧世海（2019年1月8日）。人工智能和视觉分析：分析酒店评论和回复的深度学习方法。第52届夏威夷系统科学国际会议，夏威夷毛伊岛。 http://scholarspace.manoa.hawaii.edu/handle/10125/59963.

郭，L.,谢，K. L.（2016）。影响在线酒店评论有用性的因素：经理的回应是否起作用？《国际当代酒店管理杂志》，28（10），2156–2177。 https://doi.org/10.1108/IJCHM-03-2015-0107.

拉德哈里，R.,和米肖，M.（2015）。eWOM对酒店预订意向、态度、信任和网站感知的影响。《国际酒店管理杂志》，46, 36–45。 https://doi.org/10.1016/j.ijhm.2015.01.010.

赖世安，徐立军，刘国强，赵杰（2015）。用于文本分类的递归卷积神经网络。《第二十九届AAAI人工智能会议记录》（第2267–2273页）。 http://dl.acm.org/citation.cfm?id 2886521.2886636. ¼

罗，R.,李，G.,方，D. K. C.,和韩，X.（2019）。旅游需求预测：一种深入学习的方法。《旅游研究年鉴》，75410-423。 https://doi.org/10.1016/j.annals.2019.01.014.

Lee, C. H.,和Cranage, D. A.（2012）。了解消费者对负面在线口碑传播的处理：意见共识和组织应对策略的作用。酒店与旅游研究杂志。 https://doi.org/10.1177/1096348012451455.

Lee, J. Y.和Dernoncourt, F.（2016）。使用递归和卷积神经网络的序列短文本分类。ArXiv:1603.03827 [Cs, Stat] http://arxiv.org/abs/1 603.03827.

Leung, D., Law, R., Hoof, H. van和Buhalis, D.（2013）。旅游和酒店业中的社交媒体：文献综述。《旅游与旅游营销杂志》，30（1–2），3–22。 https://doi.org/10.1080/10548408.2013.750919.

Levy, S. E., Duan, W.,和Boo, S.（2013）。对华盛顿特区住宿市场一星在线评论和回应的分析。康奈尔酒店季刊，54（1），49–63。 https://doi.org/10.1177/1938965512464513.

Li, C., Cui, G., & Peng, L. (2017). The signaling effect of management response in engaging customers: A study of the hotel industry. *Tourism Management, 62*, 42–53. https://doi.org/10.1016/j.tourman.2017.03.009.

Li, G., Law, R., Vu, H. Q., & Rong, J. (2013). Discovering the hotel selection preferences of Hong Kong inbound travelers using the choquet integral. *Tourism Management, 36*, 321–330. https://doi.org/10.1016/j.tourman.2012.10.017.

Li, G., Law, R., Vu, H. Q., Rong, J., Zhao, X., & Roy,). (2015). Identifying emerging hotel preferences using emerging pattern mining technique. *Tourism Management, 46*, 311–321. https://doi.org/10.1016/j.tourman.2014.06.015.

Lin, H.-C., Bruning, P. F., & Swarna, H. (2018). Using online opinion leaders to promote the hedonic and utilitarian value of products and services. *Business Horizons, 61*(3), 431–442. https://doi.org/10.1016/j.bushor.2018.01.010.

Li, F., Sam), & Ryan, C. (2020). Western guest experiences of a Pyongyang international hotel, North Korea: Satisfaction under conditions of constrained choice. *Tourism Management, 76*, 103947. https://doi.org/10.1016/j.tourman.2019.07.001.

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions* (1 edition). Cambridge University Press.

Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management, 71*, 337–347. https://doi.org/10.1016/j.tourman.2018.10.004.

Liu, S., Law, R., Rong, J., Li, G., & Hall, J. (2013). Analyzing changes in hotel customers' expectations by trip mode. *International Journal of Hospitality Management, 34*, 359–371. https://doi.org/10.1016/j.ijhm.2012.11.011.

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management, 47*, 140–151. https://doi.org/10.1016/j.tourman.2014.09.020.

Liu, X., Schuckert, M., & Law, R. (2015). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing, 32*(8), 1069–1080. https://doi.org/10.1080/10548408.2014.944253.

Liu, X., Schuckert, M., & Law, R. (2018). Utilitarianism and knowledge growth during status seeking: Evidence from text mining of online reviews. *Tourism Management, 66*, 38–46. https://doi.org/10.1016/j.tourman.2017.11.005.

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism Management, 59*, 554–563. https://doi.org/10.1016/j.tourman.2016.08.012.

Liu, X., Zhang, Z., Law, R., & Zhang, Z. (2019). Posting reviews on OTAs: Motives, rewards and effort. *Tourism Management, 70*, 230–237. https://doi.org/10.1016/j.tourman.2018.08.013.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management, 68*, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009.

Lui, T.-W., Bartosiak, M., Piccoli, G., & Sadhya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management, 67*, 180–190. https://doi.org/10.1016/j.tourman.2018.01.014.

Lu, W., & Stepchenkova, S. (2012). Ecotourism experiences reported online: Classification of satisfaction attributes. *Tourism Management, 33*(3), 702–712. https://doi.org/10.1016/j.tourman.2011.08.003.

Lu, W., & Stepchenkova, S. (2015). User-generated content as a research mode in tourism and hospitality applications: Topics, methods, and software. *Journal of Hospitality Marketing & Management, 24*(2), 119–154. https://doi.org/10.1080/19368623.2014.907758.

Manek, A. S., Shenoy, P. D., Mohan, M. C., & RVK. (2017). Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier. *World Wide Web, 20*(2), 135–154. https://doi.org/10.1007/s11280-015-0381-x.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (1 edition). Cambridge University Press.

Marine-Roig, E., & Ferrer-Rosell, B. (2018). Measuring the gap between projected and perceived destination images of catalonia using compositional analysis. *Tourism Management, 68*, 236–249. https://doi.org/10.1016/j.tourman.2018.03.020.

Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality Management, 34*, 99–107. https://doi.org/10.1016/j.ijhm.2013.02.012.

Ma, Y., Xiang, Z., Du, Q., & Fan, W. (2018). Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning. *International Journal of Hospitality Management, 71*, 120–131. https://doi.org/10.1016/j.ijhm.2017.12.008.

Mellinas, J. P., Nicolau, J. L., & Park, S. (2019). Inconsistent behavior in online consumer reviews: The effects of hotel attribute ratings on location. *Tourism Management, 71*, 421–427. https://doi.org/10.1016/j.tourman.2018.10.034.

Melo, A. J. D. V. T., Hernández-Maestro, R. M., & Muñoz-Gallego, P. A. (2017). Service quality perceptions, online visibility, and business performance in rural lodging establishments. *Journal of Travel Research, 56*(2), 250–262. https://doi.org/10.1177/0047287516635822.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. ArXiv:1301.3781 [Cs] http://arxiv.org/abs/1301.3781.

Min, H., Lim, Y., & Magnini, V. P. (2015). Factors affecting customer satisfaction in responses to negative online hotel reviews: The impact of empathy, paraphrasing, and speed. *Cornell Hospitality Quarterly, 56*(2), 223–231. https://doi.org/10.1177/1938965514560014.

Mkono, M. (2016). The reflexive tourist. *Annals of Tourism Research, 57*, 206–219. https://doi.org/10.1016/j.annals.2016.01.004.

Mkono, M., & Tribe, J. (2017). Beyond reviewing: Uncovering the multiple roles of tourism social media users. *Journal of Travel Research, 56*(3), 287–298. https://doi.org/10.1177/0047287516636236.

Murphy, H. C., & Chen, M.-M. (2016). Online information sources used in hotel bookings: Examining relevance and recall. *Journal of Travel Research, 55*(4), 523–536. https://doi.org/10.1177/0047287514559033.

Noone, B. M., & McGuire, K. A. (2013). Pricing in a social world: The influence of non-price information on hotel choice. *Journal of Revenue and Pricing Management, 12*(5), 385–401. https://doi.org/10.1057/rpm.2013.13.

Öğüt, H., & Taş, B. K. O. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *Service Industries Journal, 32*(2), 197–214. https://doi.org/10.1080/02642069.2010.529436.

Pan, L.-Y., & Chiou, J.-S. (2011). How much can you trust online information? Cues for perceived trustworthiness of consumer-generated online information. *Journal of Interactive Marketing, 25*(2), 67–74. https://doi.org/10.1016/j.intmar.2011.01.002.

Pantano, E., Priporas, C.-V., & Stylos, N. (2017). 'You will like it!' using open data to predict tourists' response to a tourist attraction. *Tourism Management, 60*, 430–438. https://doi.org/10.1016/j.tourman.2016.12.020.

Park, S.-Y., & Allen, J. P. (2013). Responding to online reviews: Problem solving and engagement in hotels. *Cornell Hospitality Quarterly, 54*(1), 64–73. https://doi.org/10.1177/1938965512463118.

Park, S.-Y., & Allen, J. P. (2013). Responding to online reviews: Problem solving and engagement in hotels. *Cornell Hospitality Quarterly, 54*(1), 64–73. https://doi.org/10.1177/1938965512463118.

Parkhe, V., & Biswas, B. (2016). Sentiment analysis of movie reviews: Finding most important movie aspects using driving factors. *Soft Computing, 20*(9), 3373–3379. https://doi.org/10.1007/s00500-015-1779-1.

Pearce, P. L., & Wu, M.-Y. (2018). Entertaining international tourists: An empirical study of an iconic site in China. *Journal of Hospitality & Tourism Research, 42*(5), 772–792. https://doi.org/10.1177/1096348015598202.

de la Peña, M. R., Núñez-Serrano, J. A., Turrión, J., & Velázquez, F. J. (2016). Are innovations relevant for consumers in the hospitality industry? A hedonic approach for Cuban hotels. *Tourism Management, 55*, 184–196. https://doi.org/10.1016/j.tourman.2016.02.009.

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543). http://www.aclweb.org/anthology/D14-1162.

Phillips, P., Antonio, N., de Almeida, A., & Nunes, L. (2019). The influence of geographic and psychic distance on online hotel ratings. *Journal of Travel Research*. , 0047287519858400. https://doi.org/10.1177/0047287519858400.

Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: An empirical analysis. *Journal of Travel Research, 56*(2), 235–249. https://doi.org/10.1177/0047287516636481.

Phillips, P., Zigan, K., Santos Silva, M. M., & Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network analysis. *Tourism Management, 50*, 130–141. https://doi.org/10.1016/j.tourman.2015.01.028.

Poria, S., Cambria, E., & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems, 108*, 42–49. https://doi.org/10.1016/j.knosys.2016.06.009.

Proserpio, D., & Zervas, G. (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science, 36*(5), 645–665. https://doi.org/10.1287/mksc.2017.1043.

Puts, M., Daas, P., & Waal, T. de (2015). Finding errors in big data. *Significance, 12*(3), 26–29. https://doi.org/10.1111/j.1740-9713.2015.00826.x.

Radojevic, T., Stanisic, N., & Stanic, N. (2019). The culture of hospitality: From anecdote to evidence. *Annals of Tourism Research, 79*, 102789. https://doi.org/10.1016/j.annals.2019.102789.

Radojevic, T., Stanisic, N., Stanic, N., & Davidson, R. (2018). The effects of traveling for business on customer satisfaction with hotel services. *Tourism Management, 67*, 326–341. https://doi.org/10.1016/j.tourman.2018.02.007.

Richard, H. (2018). *Best global brands | brand profiles & valuations of the world's top brands*. http://brandirectory.com/league_tables/table/hotels-50-2018.

Rojas-Barahona, L. M. (2016). Deep learning for sentiment analysis. *Language and Linguistics Compass, 10*(12), 701–719. https://doi.org/10.1111/lnc3.12228.

Rose, G., & Willis, A. (2018). *Seeing the smart city on twitter: Colour and the affective territories of becoming smart: Environment and planning D: Society and space*. https://doi.org/10.1177/0263775818771080.

Schuckert, M., Liang, S., Law, R., & Sun, W. (2018). How do domestic and international high-end hotel brands receive and manage customer feedback? *International Journal of Hospitality Management*. https://doi.org/10.1016/j.ijhm.2018.08.017.

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing, 32*(5), 608–621. https://doi.org/10.1080/10548408.2014.933154.

李聪、崔庚、彭立军（2017）。管理层响应在吸引顾客方面的信号效应：一项关于酒店业的研究。旅游管理，62, 42–53。https://doi.org/10.1016/j.tourman.2017.03.009.

李庚，罗，R., Vu, H. Q.,荣，J.（2013）。利用choquet积分发现香港入境旅客的酒店选择偏好。旅游管理，36321-330。https://doi.org/10.1016/j.tourman.2012.10.017.

李，G.,罗，R.,Vu,HQ.,荣，J.,赵，X，&罗伊，）。(2015). 使用新兴模式挖掘技术识别新兴酒店偏好。旅游管理，46311-321。https://doi.org/10.1016/j.tourman.2014.06.015.

林H-C.，布鲁宁，P. F.,斯瓦纳，H.（2018）。利用在线意见领袖来提升产品和服务的享乐和实用价值。商业视野，61（3），431–442。https://doi.org/10.1016/j.bushor.2018.01.010.

Li, F., Sam）和Ryan, C.（2020年）。朝鲜平壤一家国际酒店的西方客人体验：选择受限条件下的满意度。旅游管理，76103947。https://doi.org/10.1016/j.tourman.2019.07.001.

Liu, B. (2015). 情绪分析：挖掘观点、情绪和情绪（1版）。剑桥大学出版社。

刘洋，黄，K.,鲍，J.,陈，K.（2019）。倾听来自国内的声音：中国游客对澳大利亚旅游目的地的情绪分析。旅游管理，71337-347。https://doi.org/10.1016/j.tourman.2018.10.004.

刘绍，罗，R.,荣，J.,李，G.,和霍尔，J.（2013）。按出行模式分析酒店客户期望的变化。《国际酒店管理杂志》，34359-371。https://doi.org/10.1016/j.ijhm.2012.11.011.

刘志强和朴南生（2015）。什么是有用的在线评论？对旅游产品网站的影响。旅游管理，47140-151。https://doi.org/10.1016/j. tourman.2014.09.020.

刘克强，舒克特，M.,罗，R.（2015）。响应管理能让酒店受益吗？来自香港酒店的证据。《旅游与旅游营销杂志》，32（8），1069–1080。https://doi.org/10.1080/10548408.2014.944253.

刘克强，舒克特，M.,罗，R.（2018）。身份寻求过程中的功利主义和知识增长：来自在线评论文本挖掘的证据。旅游管理，66, 38–46。https://doi.org/10.1016/j.tourman.2017.11.005.

刘亦永，泰切特，T.,罗西，M.,李，H.,胡，F.（2017）。大数据促进大洞察：通过412784条用户评论调查酒店满意度的特定语言驱动因素。旅游管理，59554-563。https://doi.org/10.1016/j. tourman.2016.08.012.

刘X，张，Z.,罗，R.,张，Z.（2019）。在在线旅行社上发布评论：动机、奖励和努力。旅游管理，70230-237。https://doi.org/10.1016/j. tourman.2018.08.013.

李，J.,徐，L.,唐，L.,王，S.,和李，L.（2018）。旅游研究中的大数据：文献综述。旅游管理，68301-323。https://doi.org/10.1016/j. tourman.2018.03.009.

Lui,T-W.，Bartosiak, M., Piccoli, G.,和Sadhya, V.（2018年）。在线评论回应策略及其对竞争绩效的影响。旅游管理，67180-190。https://doi.org/10.1016/j.tourman.2018.01.014.

卢-W.&斯特普琴科娃，S.（2012）。网上报道的生态旅游体验：满意度属性分类。旅游管理，33（3），702–712。https://doi.org/10.1016/j.tourman.2011.08.003.

卢-W.&斯特普钦科娃，S.（2015）。作为旅游和酒店应用研究模式的用户生成内容：主题、方法和软件。《酒店营销与管理杂志》，24（2），119–154。https://doi.org/10.1080/19368623.2014.907758.

马内克，A. S.,谢诺伊，P. D.,莫汉，M. C.,和RVK。(2017). 基于基尼指数特征选择方法和支持向量机分类器的大型影评情感分析方面术语提取。万维网，20（2），135–154。https://doi.org/10.1007/ s11280-015-0381-x.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). 信息检索导论（1版）。剑桥大学出版社。

Marine Roig，E.,和Ferrer Rosell，B.（2018）。使用成分分析测量加泰罗尼亚的投影和感知目的地图像之间的差距。旅游管理，68236-249。https://doi.org/10.1016/j.tourman.2018.03.020.

Mauri, A. G.,和Minazzi, R.（2013）。网络评论会影响酒店潜在客户的期望和购买意愿。《国际酒店管理杂志》，34, 99–107。https://doi.org/10.1016/j.ijhm.2013.02.012.

马云、向志军、杜、Q、范伟（2018）。用户提供的照片对酒店评论有用性的影响：深入学习的分析方法。《国际酒店管理杂志》，71120-131。https://doi.org/10.1016/j. ijhm.2017.12.008.

梅利纳，J. P.,尼科洛，J. L.,和帕克，S.（2019）。在线消费者评论中的不一致行为：酒店属性评级对位置的影响。旅游管理，71421-427。https://doi.org/10.1016/j.tourman.2018.10.034.

梅洛，A. J. D. V. T.,赫娜ndez Maestro，R. M.,和Mun~oz Gallego，P. A.（2017）。服务农村住宿机构的质量认知、在线知名度和业务表现。旅游研究杂志，56（2），250–262。https://doi.org/ 10.1177/0047287516635822.

米科洛夫，T.,陈，K.,科拉多，G.,和迪安，J.（2013）。向量空间中单词表示的有效估计。ArXiv:1301.3781 [Cs] http://arxiv.org/abs/ 1301.3781.

Min, H., Lim Y.,和Magnini, V. P.（2015）。影响顾客对负面在线酒店评论的满意度的因素：同理心的影响、释义、、还有速度。康奈尔酒店季刊，56（2），223–231。https://doi.org/10.1177/1938965514560014.

Mkono, M.（2016）。本能的游客。《旅游研究年鉴》，57206-219。https://doi.org/10.1016/j.annals.2016.01.004.

Mkono, M.,和Tribe, J.（2017）。超越回顾：揭示旅游社交媒体用户的多重角色。《旅游研究杂志》，56（3），287–298。https://doi.org/10.1177/0047287516636236.

墨菲，H. C.,陈，M-M.（2016）。酒店预订中使用的在线信息源：检查相关性和回忆。旅游研究杂志，55（4），523–536。https://doi.org/10.1177/0047287514559033.

Noone, B. M.,和McGuire, K. A.（2013）。社交世界中的定价：非价格信息对酒店选择的影响。收入与定价管理杂志，12（5），385–401。https://doi.org/10.1057/rpm.2013.13.

Ög̃üt，H.,和Tas̃，B. K. O.（2012）。网络顾客评论对酒店业的销售和价格。《服务业杂志》，32（2），197–214。https://doi.org/10.1080/02642069.2010.529436.

潘立尧和邱俊硕（2011）。你能相信多少在线信息？消费者产生的在线信息感知可信度的线索。互动营销杂志，25（2），67–74。https://doi.org/10.1016/j.intmar.2011.01.002.

潘塔诺，E.,普里波拉斯，C-V.，和斯泰洛斯，N.（2017）。'你会喜欢的！'使用开放数据预测游客对旅游景点的反应。旅游管理，60430–438。https://doi.org/10.1016/j.tourman.2016.12.020.

Park,S-Y.，和Allen, J. P.（2013）。回应在线评论：解决问题和参与酒店。康奈尔酒店季刊，54（1），64–73。https://doi.org/ 10.1177/1938965512463118.

Park,S-Y.，和Allen, J. P.（2013）。回应在线评论：解决问题和参与酒店。康奈尔酒店季刊，54（1），64–73。https://doi.org/ 10.1177/1938965512463118.

Parkhe, V.和Biswas, B.（2016年）。电影评论的情绪分析：利用驱动因素寻找最重要的电影方面。软计算，20（9），3373–3379。https://doi.org/10.1007/s00500-015-1779-1.

Pearce, P. L.,和Wu,M-Y.（2018）。招待国际游客：对中国标志性景点的实证研究。《酒店与旅游研究杂志》，42（5），772–792。https://doi.org/10.1177/1096348015598202.

德拉庞·a·M·R·努恩斯·塞拉诺·J·a·图里奥n、J.&贝拉zquez，F. J.（2016）。是与酒店业消费者相关的创新？古巴酒店的享乐方式。旅游管理，55184-196。https://doi.org/10.1016/j. tourman.2016.02.009.

彭宁顿，J.,索谢尔，R.,曼宁，C.（2014）。手套：单词表示的全局向量。在2014年自然语言处理中的经验方法会议记录（第1532–1543页）。http://www.aclweb.org/anthology/D14-1162.

Phillips, P., Antonio, N., de Almeida, A.,和Nunes, L.（2019年）。地理和心理距离对在线酒店评级的影响。旅游研究杂志,47287519858400 https://doi.org/10.1177/0047287519858400.

Phillips, P., Barnes, S., Zigan, K.,和Schegg, R.（2017）。了解在线评论对酒店绩效的影响：一项实证分析。旅游研究杂志，56（2），235–249。https://doi.org/10.1177/0047287516636481.

Phillips, P., Zigan, K., Santos Silva, M. M.,和Schegg, R.（2015）。在线评论对瑞士酒店绩效决定因素的交互影响：神经网络分析。旅游管理，第50130-141页。https://doi.org/10.1016/j. tourman.2015.01.028.

Poria, S., Cambria, E.,和Gelbukh, A.（2016）。使用深度卷积神经网络进行观点挖掘的方面提取。基于知识的系统，108, 42–49。https://doi.org/10.1016/j.knosys.2016.06.009.

Proserpio, D.,和Zervas, G.（2017）。在线声誉管理：评估管理层回应对消费者评论的影响。营销科学，36（5），645–665。https://doi.org/10.1287/mksc.2017.1043.

Puts, M., Daas, P.,和Waal, T. de（2015）。在大数据中发现错误。重要性，12（3），26–29。https://doi.org/10.1111/j.1740-9713.2015.00826.x.

T.拉多耶维奇、N.斯坦尼西奇和N.斯坦尼奇（2019年）。好客文化：从轶事到证据。《旅游研究年鉴》，79102789。https://doi.org/10.1016/j. annals.2019.102789.

T.拉多耶维奇、N.斯坦尼奇、N.斯坦尼克和R.戴维森（2018年）。商务旅行对酒店服务顾客满意度的影响。旅游管理，67326-341。https://doi.org/10.1016/j.tourman.2018.02.007.

理查德·H.（2018）。最佳全球品牌|世界顶级品牌的品牌简介和评估。http://brandirectory.com/league_tables/table/hotels-50-2018.

罗哈斯·巴拉奥纳，L. M.（2016）。深入学习情绪分析。《语言与语言学指南针》，10（12），701–719。https://doi.org/10.1111/lnc3.12228.

罗斯·G.和威利斯·A.（2018）。在twitter上看到智慧城市：色彩和智慧的情感领域：环境和规划D：社会和空间。https:// doi.org/10.1177/0263775818771080.

舒克特，M.,梁，S.,罗，R.,孙，W.（2018）。国内外高端酒店品牌如何接收和管理客户反馈？国际酒店管理杂志。https://doi.org/10.1016/j.ijhm.2018.08.017.

Schuckert, M., Liu, X.,和Law, R.（2015）。酒店和旅游在线评论：近期趋势和未来方向。《旅游与旅游营销杂志》，32（5），608–621。https://doi.org/10.1080/10548408.2014.933154.

Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing, 21*(4), 76–94. https://doi.org/10.1002/dir.20090.

Serra Cantallops, A., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management, 36*, 41–51. https://doi.org/10.1016/j.ijhm.2013.08.007.

Soler, I. P., Gemar, G., Correia, M. B., & Serra, F. (2019). Algarve hotel price determinants: A hedonic pricing model. *Tourism Management, 70*, 311–321. https://doi.org/10.1016/j.tourman.2018.08.028.

Sparks, B. A., So, K. K. F., & Bradley, G. L. (2016). Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern. *Tourism Management, 53*, 74–85. https://doi.org/10.1016/j.tourman.2015.09.011.

Stamolampros, P., Korfiatis, N., Kourouthanassis, P., & Symitsi, E. (2019). Flying to quality: Cultural influences on online reviews. *Journal of Travel Research, 58*(3), 496–511. https://doi.org/10.1177/0047287518764345.

Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. *Information Fusion, 36*, 10–25. https://doi.org/10.1016/j.inffus.2016.10.004.

Su, Y., & Teng, W. (2018). Contemplating museums' service failure: Extracting the service quality dimensions of museums from negative on-line reviews. *Tourism Management, 69*, 214–222. https://doi.org/10.1016/j.tourman.2018.06.020.

Tableau: Business Intelligence and Analytics Software. (2018). *Tableau software*. https://www.tableau.com/.

Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management, 75*, 550–568. https://doi.org/10.1016/j.tourman.2019.06.020.

Tanford, S., & Montgomery, R. (2015). The effects of social influence and cognitive dissonance on travel purchase decisions. *Journal of Travel Research, 54*(5), 596–610. https://doi.org/10.1177/0047287514528287.

Torres, E. N., Singh, D., & Robertson-Ring, A. (2015). Consumer reviews and the creation of booking transaction value: Lessons from the hotel industry. *International Journal of Hospitality Management, 50*(Supplement C), 77–83. https://doi.org/10.1016/j.ijhm.2015.07.012.

TripAdvisor. (2018). *TripAdvisor*. Inc. Earnings Press Release. https://tripadvisor.mediaroom.com/2018-11-07-TripAdvisor-Inc-Earnings-Press-Release-Available-on-Companys-Investor-Relations-Site.

Valdivia, A., Hrabova, E., Chaturvedi, I., Luzón, M. V., Troiano, L., Cambria, E., et al. (2019). Inconsistencies on TripAdvisor reviews: A unified index between users and sentiment analysis methods. *Neurocomputing, 353*, 3–16. https://doi.org/10.1016/j.neucom.2018.09.096.

Valdivia, A., Luzón, M. V., & Herrera, F. (2017). Sentiment analysis in TripAdvisor. *IEEE Intelligent Systems, 32*(4), 72–77. https://doi.org/10.1109/MIS.2017.3121555.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management, 30*(1), 123–127. https://doi.org/10.1016/j.tourman.2008.04.008.

Vu, H. Q., Li, G., Law, R., & Zhang, Y. (2019). Exploring tourist dining preferences based on restaurant reviews. *Journal of Travel Research, 58*(1), 149–167. https://doi.org/10.1177/0047287517744672.

Wang, L., Wang, X., Peng, J., & Wang, J. (2020). The differences in hotel selection among various types of travellers: A comparative analysis with a useful bounded rationality behavioural decision support model. *Tourism Management, 76*, 103961. https://doi.org/10.1016/j.tourman.2019.103961.

Wei, W., Miao, L., Huang, Z., & Joy). (2013). Customer engagement behaviors and hotel responses. *International Journal of Hospitality Management, 33*, 316–330. https://doi.org/10.1016/j.ijhm.2012.10.002.

Woodman, C. J., Min-Venditti, A. A., Woosnam, K. M., & Brightsmith, D. J. (2019). Water quality for guest health at remote Amazon ecotourism lodges. *Tourism Management, 72*, 202–208. https://doi.org/10.1016/j.tourman.2018.11.014.

Wu, S., Roberts, K., Datta, S., Du, J., Ji, Z., Si, Y., et al. (2020). Deep learning in clinical natural language processing: A methodical review. *Journal of the American Medical Informatics Association, 27*(3), 457–470. https://doi.org/10.1093/jamia/ocz200.

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management, 58*, 51–65. https://doi.org/10.1016/j.tourman.2016.10.001.

Xiang, Z., Magnini, V. P., & Fesenmaier, D. R. (2015). Information Technology and consumer behavior in travel and tourism: Insights from travel planning using the internet. *Journal of Retailing and Consumer Services, 22*, 244–249. https://doi.org/10.1016/j.jretconser.2014.08.005.

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management, 43*(Supplement C), 1–12. https://doi.org/10.1016/j.ijhm.2014.07.007.

Xie, K. L., Zhang, Z., Zhang, Z., Singh, A., & Lee, S. K. (2016). Effects of managerial response on consumer eWOM and hotel performance: Evidence from tripadvisor. *International Journal of Contemporary Hospitality Management, 28*(9), 2013–2034. https://doi.org/10.1108/IJCHM-06-2015-0290.

Yang, Y., Mao, Z., & Eddie). (2019). Welcome to my home! An empirical analysis of airbnb supply in US cities. *Journal of Travel Research, 58*(8), 1274–1287. https://doi.org/10.1177/0047287518815984.

Yang, Y., Mao, Z., & Tang, J. (2018). Understanding guest satisfaction with urban hotel location. *Journal of Travel Research, 57*(2), 243–259. https://doi.org/10.1177/0047287517691153.

Yang, Y., Mueller, N. J., & Croes, R. R. (2016). Market accessibility and hotel prices in the caribbean: The moderating effect of quality-signaling factors. *Tourism Management, 56*, 40–51. https://doi.org/10.1016/j.tourman.2016.03.021.

Yang, Y., Park, S., & Hu, X. (2018). Electronic word of mouth and hotel performance: A meta-analysis. *Tourism Management, 67*, 248–260. https://doi.org/10.1016/j.tourman.2018.01.015.

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management, 28*(1), 180–182. https://doi.org/10.1016/j.ijhm.2008.06.011.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of E-word-of-mouth to hotel online bookings. *Computers in Human Behavior, 27*(2), 634–639. https://doi.org/10.1016/j.chb.2010.04.014.

Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism, 10*(4), 283–295. https://doi.org/10.3727/109830508788403114.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2017). *Recent trends in deep learning based natural language processing*. ArXiv:1708.02709 [Cs] http://arxiv.org/abs/1708.02709.

Zehrer, A., Crotts, J. C., & Magnini, V. P. (2011). The perceived usefulness of blog postings: An extension of the expectancy-disconfirmation paradigm. *Tourism Management, 32*(1), 106–113. https://doi.org/10.1016/j.tourman.2010.06.013.

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of beijing. *Tourism Management, 75*, 595–608. https://doi.org/10.1016/j.tourman.2019.07.002.

Zhang, Y., & Cole, S. T. (2016). Dimensions of lodging guest satisfaction among guests with mobility challenges: A mixed-method analysis of web-based texts. *Tourism Management, 53*, 13–27. https://doi.org/10.1016/j.tourman.2015.09.001.

Zhang, J. J., & Mao, Z. (2012). Image of all hotel scales on travel blogs: Its impact on customer loyalty. *Journal of Hospitality Marketing & Management, 21*(2), 113–131. https://doi.org/10.1080/19368623.2011.615017.

Zhang, X., Qiao, S., Yang, Y., & Zhang, Z. (2020). Exploring the impact of personalized management responses on tourists' satisfaction: A topic matching perspective. *Tourism Management, 76*, 103953. https://doi.org/10.1016/j.tourman.2019.103953.

Zhang, D., Xu, H., Su, Z., & Xu, Y. (2015). Chinese comments sentiment classification based on word2vec and SVMperf. *Expert Systems with Applications, 42*(4), 1857–1863. https://doi.org/10.1016/j.eswa.2014.09.011.

Zhang, Z., Zhang, Z., & Yang, Y. (2016). The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tourism Management, 55*, 15–24. https://doi.org/10.1016/j.tourman.2016.01.004.

Zhao, J., Cao, N., Wen, Z., Song, Y., Lin, Y., & Collins, C. (2014). #FluxFlow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 1773–1782. https://doi.org/10.1109/TVCG.2014.2346922.

Zhao, X., Roy), Wang, L., Guo, X., & Law, R. (2015). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management, 27*(6), 1343–1364. https://doi.org/10.1108/IJCHM-12-2013-0542.

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management, 76*, 111–121. https://doi.org/10.1016/j.ijhm.2018.03.017.

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing, 237*, 350–361. https://doi.org/10.1016/j.neucom.2017.01.026.

Zimmermann, S., Herrmann, P., Kundisch, D., & Nault, B. R. (2018). Decomposing the variance of consumer ratings and the impact on price and demand. *Information Systems Research, 29*(4), 984–1002. https://doi.org/10.1287/isre.2017.0764.

Yung-Chun Chang received the Ph.D. degree in information management from National Taiwan University, Taiwan, in 2016. He is currently an assistant professor in the Graduate Institute of Data Science at Taipei Medical University. His papers have appeared in the IEEE Transactions on Knowledge and Data Engineering, International Journal of Information Management, Database, ACL, HICSS, ICDE, etc. His current research interests include natural language processing, text mining, information retrieval, knowledge discovery, and question answering.

森·S·和勒曼·D.（2007）。你为什么告诉我这些？对网上消费者负面评论的调查。互动营销杂志，21（4），76–94。https:// doi.org/10.1002/dir.20090.

Serra Cantallops, A.,和Salvi, F.（2014）。《新消费者行为：伊渥姆和酒店研究综述》。《国际酒店管理杂志》，36, 41–51。https://doi.org/10.1016/j.ijhm.2013.08.007.

Soler, I. P., Gemar, G., Correia, M. B.,和Serra, F.（2019年）。阿尔加维酒店价格决定因素：享乐定价模型。旅游管理，70311-321。https://doi.org/10.1016/j.tourman.2018.08.028.

Sparks, B. A., So, K. K. F.,和Bradley, G. L.（2016）。回应负面在线评论：酒店回应对客户信任和担忧推断的影响。旅游管理，53, 74–85。https://doi.org/10.1016/j.tourman.2015.09.011.

Stamolampros, P., Korfiatis, N., Kourouthanassis, P.,和Symitsi, E.（2019年）。飞向质量：文化对在线评论的影响。《旅游研究杂志》，58（3），496–511。https://doi.org/10.1177/0047287518764345.

孙绍、罗灿、陈杰（2017）。评论意见挖掘系统的自然语言处理技术。信息融合，36, 10–25。https://doi.org/10.1016/ j.inffus.2016.10.004.

Su, Y.,和Teng, W.（2018）。反思博物馆的服务失败：从负面在线评论中提取博物馆的服务质量维度。旅游管理，69214-222。https://doi.org/10.1016/j.tourman.2018.06.020.

Tableau：商业智能和分析软件。(2018). Tableau软件。https://www.tableau.com/.

Taecharungroj, V.和Mathayomchan, B.（2019年）。分析TripAdvisor对泰国普吉岛旅游景点的评论。旅游管理，75550-568。https://doi.org/10.1016/j.tourman.2019.06.020.

坦福德，S.,和蒙哥马利，R.（2015）。社会影响和认知失调对旅游购买决策的影响。《旅游研究杂志》，54（5），596–610。https://doi.org/10.1177/0047287514528287.

Torres, E. N., Singh, D.,和Robertson Ring，A.（2015）。消费者评论和预订交易价值的创造：来自酒店业的经验教训。《国际酒店管理杂志》，50（补编C），77–83。https://doi.org/10.1016/j. ijhm.2015.07.012.

TripAdvisor。(2018). TripAdvisor。公司收益新闻稿。https://tripadvisor.mediaroom.com/2018-11-07-TripAdvisor-Inc-Earnings-Press-Release-Available-on-Compa nys-Investor-Relations-Site.

A.瓦尔迪维亚、E.赫拉博瓦、I.查图尔维迪、卢佐�n、M. V.,特罗亚诺，L.,坎布里亚，E.等人。(2019). TripAdvisor评论的不一致性：用户和情绪分析方法之间的统一索引。神经计算，353, 3–16。https://doi.org/10.1016/j. neucom.2018.09.096.

卢佐州瓦尔迪维亚�n、M. V.和Herrera, F.（2017）。TripAdvisor中的情绪分析。IEEE智能系统，32（4），72–77。https://doi.org/10.1109/MIS.2017.3121555.

Vermeulen,E.和Seegers, D.（2009）。久经考验：在线酒店评论对消费者考虑的影响。旅游管理，30（1），123–127。https://doi.org/10.1016/j.tourman.2008.04.008.

吴海奎、李国光、罗瑞和张勇（2019）。根据餐厅评论探索游客的用餐偏好。旅游研究杂志，58（1），149–167。https://doi.org/ 10.1177/0047287517744672.

王立军、王晓军、彭建军和王建军（2020年）。不同类型旅行者在酒店选择上的差异：一个有用的有限理性行为决策支持模型的比较分析。旅游管理，76103961。https://doi. org/10.1016/j.tourman.2019.103961.

魏、W、苗、L、黄、Z和乔伊）。(2013). 顾客参与行为和酒店回应。《国际酒店管理杂志》，33316-330。https://doi. org/10.1016/j.ijhm.2012.10.002.

Woodman, C. J., Min Venditti，A. A., Woosnam, K. M.,和Brightmith，D. J.（2019）。偏远亚马逊生态旅游度假酒店的水质，以保障客人健康。旅游管理，72202-208。https://doi.org/10.1016/j.tourman.2018.11.014.

吴淑珍、罗伯茨、达塔、杜、J、季、Z、斯、Y等（2020年）。临床自然语言处理中的深度学习：系统性综述。美国医学信息学协会杂志，27（3），457–470。https://doi.org/10.1093/jamia/ocz200.

向志强、杜、Q、马、Y、范伟（2017）。主要在线评论平台的比较分析：社交媒体分析对酒店业和旅游业的影响。旅游管理，58, 51–65。https://doi.org/10.1016/j.tourman.2016.10.001.

项，Z.,马格尼尼，V. P.,和费森梅尔，D. R.（2015）。旅游和旅游业中的信息技术和消费者行为：来自互联网旅游规划的见解。零售和消费者服务杂志，22, 244–249。https://doi.org/ 10.1016/j.jretconser.2014.08.005.

谢克力，张志强，张志强（2014）。在线消费者评论的商业价值和管理层对酒店绩效的反应。《国际酒店管理杂志》，43（补编c），1–12。https://doi.org/10.1016/j.ijhm.2014.07.007.

谢克力、张志强、张志强、辛格、A.、李世强（2016）。管理层回应对消费者eWOM和酒店绩效的影响：来自tripadvisor的证据。《国际当代酒店管理杂志》，28（9），2013–2034年。https://doi.org/10.1108/IJCHM-06-2015-0290.

杨，Y，毛，Z和埃迪）。(2019). 欢迎来到我家！美国城市airbnb供应的实证分析。旅游研究杂志，58（8），1274–1287。https://doi.org/10.1177/0047287518815984.

杨勇、毛志强、唐杰（2018）。了解客人对城市酒店位置的满意度。《旅游研究杂志》，57（2），243–259。https://doi.org/10.1177/ 0047287517691153.

Yang, Y., Mueller, N. J.,和Croes, R. R.（2016）。加勒比地区的市场可及性和酒店价格：质量信号因素的调节效应。旅游管理，56, 40–51。https://doi.org/10.1016/j.tourman.2016.03.021.

杨勇、朴世贤和胡晓霞（2018）。电子口碑与酒店绩效：荟萃分析。旅游管理，67248-260。https://doi.org/10.1016/j. tourman.2018.01.015.

叶问，罗，R.,顾，B.（2009）。在线用户评论对酒店客房销售的影响。《国际酒店管理杂志》，28（1），180–182。https://doi.org/10.1016/j.ijhm.2008.06.011.

叶问，罗，R.,顾，B.,陈，W.（2011）。用户生成内容对旅行者行为的影响：一项关于电子口碑对酒店在线预订影响的实证调查。计算机与人类行为，27（2），634–639。https://doi. org/10.1016/j.chb.2010.04.014.

Yoo, K. H.,和Gretzel, U.（2008）。是什么促使消费者撰写在线旅游评论？信息技术与旅游，10（4），283–295。https://doi.org/ 10.3727/109830508788403114.

杨，T.,哈扎里卡，D.,波里亚，S.,和坎布里亚，E.（2017）。基于深度学习的自然语言处理的最新趋势。ArXiv:1708.02709 [Cs] http://arxiv.org/abs/1 708.02709.

Zehrer, A., Crotts, J. C.,和Magnini, V. P.（2011）。博客帖子的感知有用性：预期不确定范式的延伸。旅游管理，32（1），106–113。https://doi.org/10.1016/j.tourman.2010.06.013.

张坤、陈妍、李灿（2019）。通过计算机深度学习模型分析照片的视觉内容，发现游客在旅游目的地的行为和感知：以北京为例。旅游管理，75595-608。https://doi. org/10.1016/j.tourman.2019.07.002.

张勇和科尔，S. T.（2016）。有流动性挑战的客人中住宿客人满意度的维度：基于网络文本的混合方法分析。旅游管理，53, 13–27。https://doi.org/10.1016/j.tourman.2015.09.001.

张俊杰、毛志强（2012）。旅游博客上所有酒店规模的形象：它对客户忠诚度的影响。《酒店营销与管理杂志》，21（2），113–131。https://doi.org/10.1080/19368623.2011.615017.

张X，乔S，杨Y，张Z.（2020）。探讨个性化管理反应对游客满意度的影响：主题匹配视角。旅游管理，76103953。https://doi.org/10.1016/j.tourman.2019.103953.

张德旭、徐浩、苏志强、徐勇（2015）。基于word2vec和SVMperf的中文评论情感分类。《专家系统与应用》，42（4），1857–1863年。https://doi.org/10.1016/j.eswa.2014.09.011.

张志强，张志强，杨勇（2016）。专家身份的力量：网站认可的专家评论如何影响旅行者的在线评级行为。旅游管理，55, 15–24。https://doi.org/10.1016/j.tourman.2016.01.004.

赵，J.,曹，N.,文，Z.,宋，Y.,林，Y.,和柯林斯，C.（2014）#FluxFlow：对社交媒体上传播的异常信息进行可视化分析。IEEE可视化和计算机图形学学报，20（12），1773–1782。https://doi.org/10.1109/ TVCG.2014.2346922.

赵，X，罗伊），王，L，郭，X和罗，R.（2015）。在线评论对在线酒店预订意向的影响。《国际当代酒店管理杂志》，27（6），1343–1364。https://doi.org/10.1108/IJCHM-12-2013-0542.

赵勇、徐晓霞、王明（2019）。预测总体客户满意度：来自酒店在线文本评论的大数据证据。《国际酒店管理杂志》，76111-121。https://doi.org/10.1016/j.ijhm.2018.03.017.

周，L.,潘，S.,王，J.,和瓦西拉科斯，A. V.（2017）。大数据机器学习：机遇与挑战。神经计算，237350-361。https://doi.org/10.1016/j.neucom.2017.01.026.

Zimmermann, S., Herrmann, P., Kundisch, D.,和Nault, B. R.（2018）。分解消费者评级的差异以及对价格和需求的影响。信息系统研究，29（4），984–1002。https://doi.org/10.1287/isre.2017.0764.

张永春于2016. 获得台湾国立台湾大学信息管理博士学位。他目前是台北医科大学数据科学研究所助理教授。他的论文发表在IEEE知识与数据工程学报、国际信息管理杂志、数据库、ACL、HICSS、ICDE等。他目前的研究兴趣包括自然语言处理、文本挖掘、信息检索、知识发现和问答。

Chih-Hao Ku is an Assistant Professor in the Department of Information Systems of Monte Ahuja College of Business at Cleveland State University. Dr. Ku received his M.S. and Ph.D. in Information Systems and Technology (2012) at Claremont Graduate University. His research currently focuses on deep learning, natural language processing, and visual analytics. He has published his work in the *Journal of the American Society for Information Science and Technology*, *Government Information Quarterly*, *Journal of Information Systems, and International Journal of Information Management* among others.

Chien-Hung Chen received the Master degree in Data Science from Taipei Medical University, Taiwan, in 2019. He is currently working toward the PhD degree in the Graduate Institute of Networking and Multimedia at the National Taiwan University. His research interests include natural language processing and information retrieval. He has participated in many researches such as sentiment analysis on tweets, finding Chinese poet style from poetry, sentiment dialogue system, and online question answering.

顾志浩是克利夫兰州立大学蒙特阿胡亚商学院信息系统系助理教授。Ku博士于2012年在克莱蒙特研究生大学获得信息系统与技术硕士和博士学位。他的研究目前集中于深度学习、自然语言处理和视觉分析。他在《美国信息科学与技术学会杂志》、《政府信息季刊》、《信息系统杂志》和《国际信息管理杂志》等期刊上发表了自己的研究成果。

陈建鸿于2019. 获得台湾台北医科大学数据科学硕士学位。他目前正在国立台湾大学网络与多媒体研究生院攻读博士学位。他的研究兴趣包括自然语言处理和信息检索。他参与了许多研究，如推特上的情感分析、从诗歌中寻找中国诗人风格、情感对话系统和在线问答。