

# Artificial Intelligence and Visual Analytics: A Deep-Learning Approach to Analyze Hotel Reviews & Responses

Chih-Hao Ku  
Lawrence Tech.  
University  
[cku@ltu.edu](mailto:cku@ltu.edu)

Yung-Chun Chang  
Taipei Medical  
University  
[changyc@tmu.edu.tw](mailto:changyc@tmu.edu.tw)

Yichung Wang  
Newcastle  
University  
[yi-chuan.wang@newcastle.ac.uk](mailto:yi-chuan.wang@newcastle.ac.uk)

Chien-Hung Chen  
Taipei Medical  
University  
[m946106007@tmu.edu.tw](mailto:m946106007@tmu.edu.tw)

Shih-Hui Hsiao  
Lawrence Tech.  
University  
[shsiao@ltu.edu](mailto:shsiao@ltu.edu)

## Abstract

*With a growing number of online reviews, consumers often rely on these reviews to make purchase decisions. However, little is known about managerial responses to online hotel reviews. This paper reports on a framework to integrate visual analytics and machine learning techniques to investigate whether hotel managers respond to positive and negative reviews differently and how to use a deep-learning approach to prioritize responses. In this study, forty 4- and 5-star hotels in London with 91,051 reviews and 70,397 responses were collected and analyzed. Visual analyses and machine learning were conducted. The results indicate most hotels (72.5%) showing no preference to respond to positive and negative reviews. Our proposed deep-learning approach outperformed existing algorithms to prioritize responses.*

## 1. Introduction

Online travel agencies (OTAs) such as Booking.com, Expedia, and TripAdvisor provide a platform where users can share subjective opinions, recommendations and ratings about their travel and accommodation experiences. Today, TripAdvisor, the largest travel platform in the world [49], has over 630 million reviews and opinions with an average of 455 million of monthly unique visitors<sup>1</sup>. In TripAdvisor, a managerial response becomes the final conversation on the review because only one registered manager can create such last response<sup>2</sup>. That means, managerial responses can have a significant impact on other potential guests who plan to book a hotel.

In hospitality management, the customer-generated content such as hotel ratings and reviews could be a valuable source for identifying the consumption patterns and trends due to its active and real-time natures.

Managerial responses to customer reviews may enhance existing customers' loyalty and turn unhappy customers into loyal customers [34]. In a recent study, Schuckert et al. [37] conducted a content analysis to analyze 50 articles from 2004 to 2013 relevant to hospitality and tourism online reviews and found that existing studies focus predominantly on analyzing secondary data, discovering the relationship between online reviews and sales including customer satisfaction [33], and opinion mining of online reviews. They further point out the limitations of current studies, which include the use of simple variables such as overall ratings and the number of reviews [11] for data analyses. Furthermore, prior research has focused mostly on a stand-alone fashion [43], analyzing either online reviews or managerial responses, which generate limited insights of the interrelated relationship between online reviews and managerial responses [46].

This study fills the research gaps by taking dimensions such as aspect ratings, types of travelers, and time to respond to reviews into our data analyses. A novel approach to integrate deep-learning models and visual analytics techniques is then proposed. The overall results can be used to improve customer relationship management, make self-improvements [34] for response management managers, and generate decision-making information for travelers.

The remainder of the paper is organized as follows. First, starting with a brief literature review of online reviews and ratings, managerial response, artificial intelligence, and natural language processing. Second, drawing on our initial visual analytics, we select representative hotels to develop machine learning models. To achieve this, we develop an analytical framework including data crawling, visual analytics, and machine learning. We then conduct 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.

<sup>1</sup> TripAdvisor Q1 2018 results, <http://ir.tripadvisor.com/static-files/90504393-5f9d-4ac8-a5d7-b67d8b4adb60>

<sup>2</sup> The Final Word: TripAdvisor Management Responses, <https://www.tripadvisor.com/TripAdvisorInsights/w637>

## 2. Related work

### 2.1 Online hotel reviews and ratings

Today, tourists usually read online reviews to plan their trip and decide where to stay [30]. Traditional word-of-mouth (WOM) communications, oral messages between persons, have evolved into electronic word-of-mouth (eWOM) communications, online messages between users [30]. The proliferation of eWOM has been identified as a strong impact on consumers' purchase decision [30], revisit intentions [50], search behaviors [21], and online sales [33]. Consumers tend to search and compare tourism and hospitality products and services to reduce uncertainty [10] and potential risks associated with purchase [39].

An average review rating is an important indicator leading to hotel sales. Noone and McGuire [32] examined the relation between online reviews and online hotel booking and found that higher average review ratings lead to higher numbers of hotel bookings. However, the overall ratings may not reveal customers' real satisfaction, and more nuances of response strategies should be further studied [37].

Hotel star ratings are used to classify hotels based on their quality approved by national or local governments or independent organizations [29]. The rating system classifies hotels differently in each country<sup>3</sup>. Prior studies have revealed that a positive correlation between a hotel's star rating and hotel's sales [19]. Martin-Fuentes [29] collected a sample of more than 14,000 hotels in 100 cities from Booking.com and TripAdvisor to examine the star-rating classification system of hotels, room price, and user satisfaction measure from user ratings. The analysis result confirms that hotel stars indicate the overall quality of hotels and a hotel price is related to hotel stars and user satisfaction.

### 2.2 Managerial response

Due to an exponential increase in online reviews being available in social media platforms, managerial responses have become a new form of customer relationship management (CRM) [13]. Law et al. [24] analyzed 111 hospitality-related articles from March to August 2017 and found that hospitality CRM research has grown from a marketing to social CRM concept. They further point out that technology plays an indispensable role in such process and artificial intelligence can generate new knowledge in this rapidly growing field and thus foster customer relationship [24].

Today, managing online reviews for hotels has become an important task for hospitality management [25] and scholars have urged hotel managers to respond to online reviews proactively [39]. Existing studies [28] reveal that deficient service does not but rather improper responses lead to dissatisfied customers because most customers recognize imperfect service [38]. More recently, Sparks et al. [39] adopted an experimental approach based on Kardes's consumer inferences theory [18] to examine organizational responses to negative eWOM and found that a timely response yielded favorable customer inferences.

According to the service recovery theory [3], managerial responses to negative reviews can identify service failures and enhance customer satisfaction [45]. Kim et al. [19] collected online reviews and responses from 128 hotels in 45 states in the U.S. and found that overall ratings and responses to negative comments are the most salient predictor of hotel performance. Seeking effective approaches to manage eWOM, especially negative ones, is a widely recognized challenge for hospitality management [39].

### 2.3 Artificial intelligence (AI)

Understanding and responding to massive online reviews is a time-consuming and exhausting task. A customer review may contain both positive and negative information, which make the in-depth analysis of online reviews even more challenging. To automate the analysis process with several millions of data records, AI techniques can be used.

AI enables machines (computers) to perform intelligent and cognitive processes. The popular subfields of AI include search and planning, reasoning and knowledge representation, perception, computer vision, machine learning, and natural language processing [36]. For example, AI and opinion mining techniques [37] have been used to facilitate data analysis and identify the reviews that require immediate attention from a review management manager. If managers can respond to reviews successfully and promptly, they have an opportunity to turn unsatisfied customers into loyal customers [34].

Machine learning and data mining techniques enable computers to learn and thus make predictions or discover hidden patterns from the collected data. Machine learning includes supervised [1], which requires a set of predefined categories or tagging labels, and unsupervised methods, which does not require data labeling. Supervised machine learning algorithms such

---

<sup>3</sup> What do star ratings for hotels mean?, <https://www.tripadvisor.com/hc/en-us/articles/200614057-What-do-star-ratings-for-hotels-mean->

as Naïve Bayes [17], regression analysis [52], decision trees [41], k-nearest neighbors [9], and Support Vector Machine (SVM) [5] have been used to conduct sentiment analysis and classification for online reviews and tourism research. For example, Dey et al. [9] used Naïve Bayes and K-Nearest Neighbor to perform sentiment analysis of hotel and movie reviews and Chang et al. [5] adopted a novel SVM approach to conduct aspect-based sentiment analysis of hotel reviews and visualize the result.

One of the popular unsupervised learning algorithms is the k-means clustering algorithm [22], which group data points that are found to possess similar features. Zhang and Yu [51], for example, use a Word2Vec tool, a deep-learning tool proposed by Google, k-means clustering algorithm, and ISODATA, a clustering algorithm based on k-means, to conduct the experiments of sentiment analysis on hotel reviews and found a slight performance improvement by using Word2Vec together with ISODATA. Neural networks, imitating the function of the human brain, can be either a supervised or unsupervised approach [40]. Chong et al. [7], for example, use a neural network approach to investigate the use of online reviews, online promotional strategies, and sentiments from user reviews to predict product sales and found a positive relationship as a result.

## 2.4 Natural language processing (NLP)

NLP is an important technique to process textual data and advance research fields such as digital government, management science, political science, marketing, and hospitality management, which concerned with consumer and public opinions. NLP is an interdisciplinary field of AI, computer science, and computational linguistics, which can be used to process digital text or speech [27]. Liu et al. [27] analyzed 238 articles in Information Systems (IS) between 2004 and 2015 and revealed that an evident increasing trend of NLP research in IS.

NLP applications such as sentiment analysis, topic modeling, and document summarization require preprocessing tasks for structuring the text and extracting features [42]. Widely used NLP tasks includes tokenization, part-of-speech (POS) tagging, stop words removal [26], noun phrase extraction, named entity recognition, stemming, parsing, coreference resolution, and disambiguation [42]. Popular NLP tools and frameworks used in research projects to analyze user reviews include Java-based tools such as general architecture for text engineering (GATE)<sup>4</sup>, Mallet<sup>5</sup>,

CoreNLP<sup>6</sup>, and openNLP<sup>7</sup>, Python-based tool such as NLTK<sup>8</sup>, and R packages<sup>9</sup>.

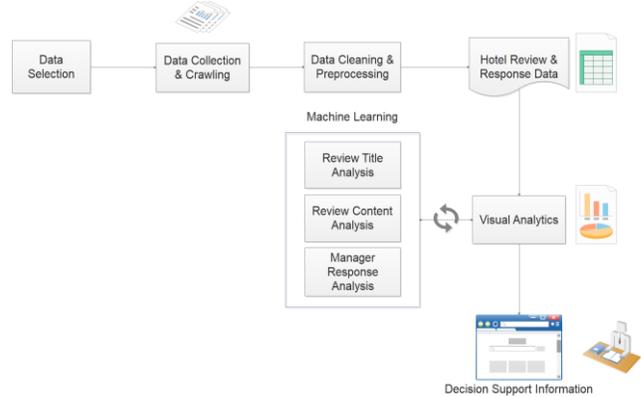
In tourism research, NLP techniques have been applied to conduct sentiment analysis [2] and topic modeling [15]. For example, Guo et al. [15] use topic modeling - Latent Dirichlet Analysis (LDA), which combines machine learning and NLP techniques to extract dimension of customer satisfaction from 266,544 online reviews for 25,670 hotels in 16 countries. Both Chang et al. [5] and Akhtar et al. [2] have used more advanced NLP techniques to detect aspect-based sentiment from hotel reviews and ratings, which extract fine-grained opinions toward hotel reviews.

This study differs from previous tourism research on analyzing hotel reviews and responses. We integrate visual analytics and deep-learning techniques and include nuances such as time, types of travelers, aspect ratings, sentiments of reviews in addition to overall ratings, star, the number of reviews to gain deeper insights.

## 3. Methodology

We develop an analytical framework (see Figure 1) by integrating five major components: data selection, data collection & crawling, data cleaning & preprocessing, visual analytics, and machine learning. Next, deep-learning models were developed to analyze hotel reviews including review titles and contents and managerial responses using novel machine learning and NLP techniques.

### 3.1 Analytical framework



**Figure 1. Analytical framework for hotel review & response analysis**

<sup>4</sup> GATE, <https://gate.ac.uk/>

<sup>5</sup> Mallet, <http://mallet.cs.umass.edu/>

<sup>6</sup> CoreNLP, <https://stanfordnlp.github.io/CoreNLP/>

<sup>7</sup> openNLP, <https://opennlp.apache.org/>

<sup>8</sup> NLTK, <https://www.nltk.org/>

<sup>9</sup> CRAN Task, <https://cran.r-project.org/web/views/NaturalLanguageProcessing.html>

### 3.1.1 Data selection

According to Mastercard Global Destination Cities Index 2017<sup>10</sup>, London has been one of the most popular cities for international travelers among 132 top destination cities. London is being selected for this study because it is an English-speaking city and the most visited city in Europe and Northern America cities based on the volume of visitors.

London is among the 18 Hilton Top Destinations<sup>11</sup> listed in Hilton's official website. Each top destination includes detailed guidance to the city and a list of hotels. We selected 43 Hiltons hotels, up to 25 miles from London, in 2017 based on Hilton's website<sup>12</sup>. Only 4- and 5-stars hotels were selected, because luxury hotels are more likely to provide better experience and service to guests. Three Hampton hotels have lower or no star ranking are eliminated from our data analysis. This results in a total of 40 hotels used in this study.

### 3.1.2 Data collection & crawling

Relying on an automatic web crawler, we collected user review, manager response, and hotel rating data from TripAdvisor. For each hotel, we collected hotel name, star, the number of excellent, good, average, poor, and terrible reviews, an average of a price range, hotel address, amenities, type of rooms, and description<sup>13</sup>. For each hotel review, we collected the hotel name, review title, review content, manager response, overall rating, aspect ratings, types of travelers, and review date. The data were collected from the earliest date (January 2010) that the rating data were available for the selected hotels to the date (October 2017) that data analyses were conducted. A total of 91,051 reviews were collected. Among them, 70,397 reviews contain managerial responses, resulting in an overall 77% response rate.

### 3.1.3 Data preprocessing

The collected raw data cannot be used for visual analytics and machine learning immediately. To preprocess the collected data, we first join two datasets - hotel data and review data to have a holistic view of the data. Here are the major steps used to preprocess diverse types of data.

- Keyword extraction. To simplify the short phrase such as 'traveled with family' and 'traveled with a coupled', we extract the five keywords Family, Solo, Couple, Friends, Business to represent types

of travelers and Unknown is used to fill up unknown travelers.

- Value extraction. We split and extract multi-value data such as aspects "[Value-5|Location-5|Rooms-5|Cleanliness-5|Service-5]" for our data analysis. In this example, the numerical value 5 is extracted to the corresponding aspects Value, Location, Rooms, Cleanliness, and Service.
- Data classification. TripAdvisor uses a 1 (terrible) - 5 (excellent) scale, a bubble rating, for visitors to rate each hotel. A 4 - 5 rating indicates a positive review, while a 1 - 3 rating indicates a negative review. A similar classification approach has been used in existing studies [43]. An additional, binary dimension 'Response or Not' is also created to indicate whether a review contains a managerial response.

## 3.2 Visual analytics and discussion

Visual analytics tools are useful to facilitate exploratory data analysis [6]. Social media data tend to be large and unstructured and contain multiple dimensions. In this study, we use exploratory data analysis (EDA) which includes an iterative process to examine summary statistics and data visualizations [35] with more than 20 dimensions such as sentiment, aspect rating, managerial response, review time, response time, latitude, longitude, and type of traveler. The purpose of this visual analytics is to produce sensemaking through rapid analysis, synthesis, deduction, induction, barnstorming, and refinement [6].

### 3.2.1 Clustering analysis

Öğüt and Taş [33] conduct a study to compare hotels from London and Paris on booking and reveal that hotel star ratings significantly affect room prices and customer ratings. That is, star ratings are correlated with room prices and frequently used to rate hotel quality. Glauber Eduardo [12] further discovers that cleanliness, location, and facilities in 8,000 hotels worldwide are relevant to hotel quality and price differences. Ye et al. [47] found that a good reputation is related to a higher hotel price. We repeat the clustering analysis based on the result of the number of clusters and between- and within-group sum of squares. A larger value of the between-group sum of squares indicates a better separation between clusters, while a smaller value of the

<sup>10</sup> Mastercard Destination Cities Index, <https://newsroom.mastercard.com/wp-content/uploads/2017/10/Mastercard-Destination-Cities-Index-Deck.pdf>

<sup>11</sup> Hilton Top Destinations, <http://www.hilton.com/top-destinations/>

<sup>12</sup> Hilton Hotels in London, <http://www.hilton.com/top-destinations/london-hotels>

<sup>13</sup> Sample hotel data – Conrad London St. James, [https://www.tripadvisor.com/Hotel\\_Review-g186338-d2309633-Reviews-Conrad\\_London\\_St\\_James-London\\_England.html](https://www.tripadvisor.com/Hotel_Review-g186338-d2309633-Reviews-Conrad_London_St_James-London_England.html)

within-group sum of squares indicates more cohesive clusters.

Aspect ratings, which cannot improve the clustering result, are excluded from our clustering analysis. For example, the location aspect is excluded because all hotels are in London. The facilities are not used as an input variable because there is no such rating in TripAdvisor. We, accordingly, use the overall ratings, average minimum and maximum price in U.S. dollars, and the aspect rating of cleanliness to generate a balanced result.

The default k-means algorithm with a normalized scaling is used to conduct clustering analysis for a given of 3 clusters. The algorithm then partitions the given 40 hotels into 3 clusters (see Figure 2) with blue, orange, and green stars, respectively. Each star represents a hotel and the label shows either a 4- or 5-star hotel. A medium with quartiles reveals that the average minimum and maximum price is \$103 and \$255, respectively. Table 1 shows the summary diagnosis of the clustering results. We found that three 5-star and four 4-star hotels have been placed in the same cluster (Cluster 3 - green stars) with the highest overall and cleanliness aspect ratings. The blue stars (Cluster 1) indicate hotels with higher overall and cleanliness ratings, compared to hotels with lower ratings in orange stars (Cluster 2).

The rest of aspect ratings such as the average location, service, sleep quality, and value ratings were also examined for three clusters. The overall result shows that cluster 3 has the highest aspect ratings among three clusters. Business service ratings and check-in ratings were excluded from our analysis because of null values. For example, the business service rating was no longer available on TripAdvisor.



Figure 2. Clustering analysis of 40 hotels in London

Table 1. Summary diagnostics of clustering results

Summary Diagnostics	
Number of Clusters:	3
Number of Points:	40
Between-group Sum of Squares:	7.1949
Within-group Sum of Squares:	3.5811
Total Sum of Squares:	10.776

Clusters	1	2	3
# of Clusters	23	10	7
Avg. Min Price	97.51	91.50	176.16
Avg. Max Price	230.48	214.09	413.87
Avg. Overall Ratings	4.12	3.47	4.35
Avg. Cleanliness Ratings	4.50	3.85	4.61

### 3.2.2 Response rate analysis

We are particularly interested in managerial responses in each hotel. A table is awkward for comparisons among all 40 hotels because we have to compare the number of responses and non-responses to gain a deeper insight into managerial responses. Additionally, each hotel review is classified into the positive or negative review based on overall hotel ratings. To facilitate our comparison, we use bar charts together with table calculations. A table calculation, provided by Tableau, is a transformation of values based on the dimensions within the level of details.

Figure 3 shows the number of reviews with (blue color) and without (orange color) managerial responses. Two hotels DoubleTree by Hilton Hotel Woking and Hilton London Green Park stand out in this visualization because both hotels have a relatively higher non-response rate (>80%), compared to the rest of hotels. Following by the hotels DoubleTree by Hilton Hotel London – Kensington, DoubleTree by Hilton Hotel London – Tower of London, and DoubleTree by Hilton London – Westminster have a non-response rate above 48%. The non-response rate of three 5-star hotels is below 27%. Five hotels such as Hilton London Bankside, DoubleTree by Hilton Dartford Bridge, Hilton London Angel Islington, Hilton London Olympia, and Hilton London Hyde Park have a very low non-response rate (<3%).

Next, the sentiment attribute, based on overall ratings, is added to our visual analysis. Due to a lengthy list of data, we only provide a brief description of the visualization results. Not surprisingly, two hotels DoubleTree by Hilton Hotel Woking and Hilton London Green Park have a high non-response rate (>79%) for both positive and negative reviews. Based on this visual exploratory, we classify hotels into three categories: negative-review-response preference, positive-review-response preference, and neutral preference.

We then tested on whether hotel managers have a preference for their response strategies. The one sample

t-test of the response preference shows that there is no difference between responses to positive reviews and responses to negative reviews ( $t=0.09$ ,  $p>0.05$ , 95CI: [-0.03-0.03]). We count the hotel as a neutral preference on response strategies when its z-score of the mean difference between positive response rate and negative response rate between -1 and 1. Surprisingly, 29 hotels (72.5%) are classified as a neutral preference, 6 hotels (15%) as a negative-review-response preference, and 5 hotels (12.5%) as a positive-review-response preference.

Prior research has emphasized that consumers are likely to pay more attention to negative reviews when making purchase decisions [4], specifically negative reviews could increase consumer awareness of hotels [44]. However, our result shows that 72.5% of hotels in our study have a neutral preference of response strategy, implying that the hotel managers put an equal amount of effort to respond to customers’ positive and negative reviews. As negative reviews indeed reduce purchase likelihood and sales [4], our finding suggests that managers should provide detailed strategies to respond to negative and positive reviews.

When the time dimension added to the response rate analysis, the result shows that the managerial response rates are similar in each month. However, we did notice that unknown travelers receive a relatively lower response rate (<60%), compare to other types of travelers (>75%). Therefore, travelers are recommended to specify a travel type to increase the response rate when leaving reviews.



Figure 3. The number of managerial responses & non-responses from 40 hotels in London

## 4. Machine learning

Learning managerial responses of high-quality hotels is crucial for CRM. The decision to respond to a review may increase transaction costs and labor costs [37] and not to respond to a review may lose opportunities to retain customers [48]. Liu et al. [28] recommend that hotel managers adopting targeted response management to increase hotel ratings. Therefore, it is important to prioritize the responses to online reviews. Leung et al. [25] also recommend hotel managers to respond to online reviews and encourage scholars to further investigate managerial responses [31].

In this research, we compile a hotel response dataset from 7 hotels of cluster 3 in Figure 2 since they present the highest overall and cleanliness aspect ratings. In fact, the cluster 3 also demonstrates the highest aspect ratings of all six aspects among three clusters. To learn the priority of managerial responses, we based on manager’s response time (i.e. the time difference between a hotel review and its response) to define two response types – critical and trivial. We adopt a quartile approach for response time to discriminate the priority of managerial responses. A hotel review belongs to a critical response if it is responded in the response time of Q1 and Q2. Otherwise, it is a trivial response (i.e. Q3, Q4, and non-response). Finally, 19,491 hotel reviews are kept and divided into the training set and the test set, containing 9,745 and 9,746 articles, respectively. Distribution of the dataset is shown in Table 2.

Table 2. Descriptive statistic of the hotel response dataset

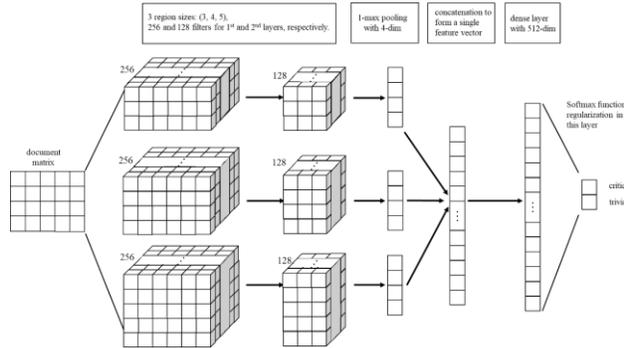
Type	# Training	# Test	Total
Critical	4,497	4,577	9,074
Trivial	5,248	5,169	10,417
Total	9,745	9,746	19,491

To detect a priority of managerial responses effectively, we proposed a deep-learning-based approach that integrated multiple convolutional neural networks [20] (CNN) for text classification. We utilize GloVe pre-trained word embeddings<sup>14</sup>, an unsupervised learning algorithm, to transfer hotel reviews as the document matrix, the rows of which are word vector representations of each token. Following Collobert and Weston [8], we can effectively treat the document matrix as an image and perform convolution on it.

Figure 4 shows the architecture of our proposed method. First, we depict three filter region sizes: 3, 4 and 5, each of which has 256 filters. Filters perform

<sup>14</sup> <https://nlp.stanford.edu/projects/glove/>

convolutions on the document matrix and generate feature maps. Next, the generated feature maps are refined through 128 filters with above different region sizes; 1-max pooling is performed over each map to capture the largest value from each feature map. Finally, we concatenate these features which are extracted from pooling layer as the input of dense layer with 512 dimensions for the penultimate layer. The final softmax layer then receives this feature vector as an input and uses it to classify the hotel review; here we assume a binary classification and hence depict two possible output states. We implemented the CNN model using Keras<sup>15</sup>, a Python deep-learning library. The maximum document length is set to 100, where longer documents were truncated, and shorter documents were padded with zeros. The 100 dimensions pre-trained word embeddings are used for document matrix generation. The training lasts for at most 100 epochs or when the accuracy of the validation sets starts to drop.



**Figure 4. Illustration of a CNN architecture for detecting priority of managerial responses**

## 6. Experimental results and discussion

A comprehensive performance evaluation of the proposed CNN-based approach with other methods is provided. Word embeddings-based approaches which represent each hotel review as the average of word embeddings (100-dimension embeddings) and classified by the SVM (denoted as SVM). Next, we further compare our method to the document modeling method that utilizes embeddings of keywords to perform text classification [16] (denote as *DKV*). In addition, the bi-directional recurrent neural network method [23] (denoted as *RNN*) is also included in the comparison. To serve as a standard for comparison, we also included the results of Naïve Bayes (denoted as *NB*) and *k*-nearest neighbors [14] (denoted as *KNN*) as baselines.

**Table 3. Performance evaluation on detecting priority of managerial responses**

Sys.	Critical	Trivial	$A^\mu$
	Precision, Recall, $F_1$ -score (%)		
NB	50.88/53.64/52.22	56.88/54.15/55.48	53.91/53.91/53.91
KNN	45.96/29.06/35.60	52.61/69.74/59.98	50.64/50.64/50.64
SVM	53.01/36.97/43.56	55.98/70.98/62.59	55.01/55.01/55.01
DKV	54.20/32.29/40.47	55.85/75.84/64.33	55.39/55.39/55.39
RNN	54.68/48.89/51.62	58.52/63.31/60.82	56.72/56.54/56.63
Our method	<b>60.25/48.12/53.51</b>	<b>61.38/71.05/65.86</b>	<b>60.85/60.82/60.56</b>

Table 3 displays the system performances for recognizing the priority of managerial responses. In general, each method in this experiment can achieve an overall  $F_1$ -score around 50%. As a baseline, the *k*-nearest neighbors method simply calculates document similarity in the bag-of-words feature space which can only accomplish a mediocre performance. The Naïve Bayes classifier is a keyword statistics-based approach which can further improve the performance with about 54%  $F_1$ -score. The word embeddings-based methods (i.e. SVM and *DKV*) is more effective in finding representative keywords, they exhibit a more evenly distributed performance among both categories. It is worth noting that the *NB* classifier indicates keyword information represented by the bag-of-words model, which is crucial in detecting the priority of managerial responses, a higher overall performance than *RNN*. Our CNN-based approach can further improve the performance through the combination of multiple CNNs, thus achieving the best overall  $F_1$ -score of 60.56%.

Our visual analysis results reveal that most hotel managers respond to positive and negative reviews equally and do not prioritize the responses. This is, in turn, lower down the performance of machine learning algorithms. We recommend that more research on response strategies should be studied and hotel managers should respond to positive and negative reviews strategically and prioritize the responses based on online review features such as sentiment, overall rating, aspect rating, and type of traveler.

## 7. Conclusion and future research

This study breaks new ground in several ways. We developed a data crawler to collect data automatically and presented a novel approach to integrating visual analytics and deep-learning models to gain insights into various aspects of hotel review and response data. The study result produces managerial, decision-making, and technical contributions. First, hotel managers can prioritize response orders and gain insights into online

<sup>15</sup> <https://keras.io/>

reviews and responses to make self-improvement. Second, the overall results also provide decision-making information for travelers to select 4-star hotels and enjoy 5-star service and environment based on our clustering analysis. Finally, we are among the first to integrate visual analytics and deep-learning models to analyze hotel reviews and responses. This can be justified by our experimental results, which indicate our proposed approach outperforms existing machine learning methods such as NB, KNN, SVM, DKV, and RNN.

There are several limitations to this study. First, this study only took the response time, response rate, and hotel rating in response data analysis. The nuances of response strategies need to be further studied. This will provide practical decision-making information for hotel managers when they respond to positive and negative reviews using different response strategies. Another limitation was that the sample represented only 40 Hilton-affiliated hotels in London. The performance of the deep-learning models needs to be tested in future research by including more cities, hotel brands, and hotels. This will provide additional insights into data analysis.

Not all eWOMs are equally important. After analyzing the textual features of the hotel reviews and managerial responses, future research can focus on the social network among those hotel questions and answers. The analytical framework applied in this study can be expanded to include social network analysis, which can show how the structure of social ties may influence the hotel reviews and manager's response strategy. For each hotel, a small social network can be generated based on the communication between customers. If one customer answers another customer's question, this indicates one directional tie between these two customers. Based on these conversation ties, a small social network of each hotel can be generated. This leads to an interesting question: if the structural cohesion of social network will influence the hotel reviews and response strategy. An equally intriguing question is whether the structural holes in each network will cause a different impact on the hotel reviews and response strategy.

### Acknowledgments

We are grateful to the anonymous reviewers for their insightful comments. This research was supported by the Ministry of Science and Technology of Taiwan under grant number MOST 106-2218-E-038 -004 -MY2 and MOST 107-2634-F-001-005.

### References

- [1] Abrahams, A.S., Fan, W., Wang, G.A., Zhang, Z. (John), and Jiao, J. An Integrated Text Analytic Framework for Product Defect Discovery. *Production and Operations Management* 24, 6 (2015), 975–990.
- [2] Akhtar, N., Zubair, N., Kumar, A., and Ahmad, T. Aspect based Sentiment Oriented Summarization of Hotel Reviews. *Procedia Computer Science* 115, (2017), 563–571.
- [3] Andreassen, T.W. Antecedents to Satisfaction with Service Recovery. *European Journal of Marketing* 34, 1/2 (2000), 156–175.
- [4] Berger, J., Sorensen, A.T., and Rasmussen, S.J. Positive Effects of Negative Publicity: When Negative Reviews Increase Sales. *Marketing Science* 29, 5 (2010), 815–827.
- [5] Chang, Y.-C., Ku, C.-H., and Chen, C.-H. Social Media Analytics: Extracting and Visualizing Hilton Hotel Ratings and Reviews from Tripadvisor. *International Journal of Information Management*, (2017).
- [6] Chen, N.-C., Brooks, M., Kocielnik, R., et al. Lariat: A Visual Analytics Tool for Social Media Researchers to Explore Twitter Datasets. *Hawaii International Conference on System Sciences 2017 (HICSS-50)*, (2017).
- [7] Chong, E., Ngai, E.W.T., Lee, F., Li, B., and Chong, A.Y.L. Predicting Online Product Sales Via Online Reviews, Sentiments, and Promotion Strategies: A Big Data Architecture and Neural Network Approach. *International Journal of Operations & Production Management* 36, 4 (2016), 358–383.
- [8] Collobert, R. and Weston, J. A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. *Proceedings of the 25th International Conference on Machine Learning*, ACM (2008), 160–167.
- [9] Dey, L., Chakraborty, S., Biswas, A., Bose, B., and Tiwari, S. Sentiment Analysis of Review Datasets Using Naive Bayes and K-NN Classifier. *International Journal of Information Engineering and Electronic Business* 8, 4 (2016), 54–62.
- [10] Fang, B., Ye, Q., Kucukusta, D., and Law, R. Analysis of the Perceived Value of Online Tourism Reviews: Influence of Readability and Reviewer Characteristics. *Tourism Management* 52, (2016), 498–506.
- [11] Gavilan, D., Avello, M., and Martinez-Navarro, G. The Influence of Online Ratings and Reviews on Hotel Booking Consideration. *Tourism Management* 66, (2018), 53–61.

- [12] Glauber Eduardo, de O.S. Worldwide Hedonic Prices of Subjective Characteristics of Hostels. *Tourism Management* 52, (2016), 451–454.
- [13] Gu, B. and Ye, Q. First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction. *Production and Operations Management* 23, 4 (2014), 570–582.
- [14] Guo, G., Wang, H., Bell, D., Bi, Y., and Greer, K. Using kNN Model for Automatic Text Categorization. *Soft Computing* 10, 5 (2006), 423–430.
- [15] Guo, Y., Barnes, S.J., and Jia, Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management* 59, (2017), 467–483.
- [16] Hsieh, Y.-L., Chang, Y.-C., Chu, C.-H., and Hsu, W.-L. How Do I Look? Publicity Mining From Distributed Keyword Representation of Socially Infused News Articles. *The 4th International Workshop on Natural Language Processing for Social Media*, Association for Computational Linguistics (2016), 74–83.
- [17] Kang, H., Yoo, S.J., and Han, D. Senti-Lexicon and Improved Naïve Bayes Algorithms for Sentiment Analysis of Restaurant Reviews. *Expert Systems with Applications* 39, 5 (2012), 6000–6010.
- [18] Kardes, F.R., Posavac, S.S., and Cronley, M.L. Consumer Inference: A Review of Processes, Bases, and Judgment Contexts. *Journal of Consumer Psychology* 14, 3 (2004), 230–256.
- [19] Kim, W.G., Lim, H., and Brymer, R.A. The Effectiveness of Managing Social Media on Hotel Performance. *International Journal of Hospitality Management* 44, Supplement C (2015), 165–171.
- [20] Kim, Y. Convolutional Neural Networks for Sentence Classification. *arXiv:1408.5882 [cs]*, Association for Computational Linguistics (2014).
- [21] King, R.A., Racherla, P., and Bush, V.D. 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 (2014), 167–183.
- [22] Kurashima, T., Iwata, T., Irie, G., and Fujimura, K. Travel Route Recommendation Using Geotagged Photos. *Knowledge and Information Systems* 37, 1 (2013), 37–60.
- [23] Lai, S., Xu, L., Liu, K., and Zhao, J. Recurrent Convolutional Neural Networks for Text Classification. *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI Press (2015), 2267–2273.
- [24] Law, R., Fong, D.K.C., Chan, I.C.C., and Fong, L.H.N. Systematic Review of Hospitality CRM Research. *International Journal of Contemporary Hospitality Management* 30, 3 (2018), 1686–1704.
- [25] Leung, D., Law, R., Hoof, H. van, and Buhalis, D. Social Media in Tourism and Hospitality: A Literature Review. *Journal of Travel & Tourism Marketing* 30, 1–2 (2013), 3–22.
- [26] Li, J., Xu, L., Tang, L., Wang, S., and Li, L. Big Data in Tourism Research: A Literature Review. *Tourism Management* 68, (2018), 301–323.
- [27] Liu, D., Li, Y., and Thomas, M.A. A Roadmap for Natural Language Processing Research in Information Systems. *Proceedings of the 50th Hawaii International Conference on System Sciences*, (2017).
- [28] Liu, X., Schuckert, M., and Law, R. Can Response Management Benefit Hotels? Evidence from Hong Kong Hotels. *Journal of Travel & Tourism Marketing* 32, 8 (2015), 1069–1080.
- [29] Martin-Fuentes, E. Are Guests of the Same Opinion as the Hotel Star-Rate Classification System? *Journal of Hospitality and Tourism Management* 29, (2016), 126–134.
- [30] Mauri, A.G. and Minazzi, R. Web Reviews Influence on Expectations and Purchasing Intentions of Hotel Potential Customers. *International Journal of Hospitality Management* 34, (2013), 99–107.
- [31] Min, H., Lim, Y., and Magnini, V.P. Factors Affecting Customer Satisfaction in Responses to Negative Online Hotel Reviews: The Impact of Empathy, Paraphrasing, and Speed. *Cornell Hospitality Quarterly* 56, 2 (2015), 223–231.
- [32] Noone, B.M. and McGuire, K.A. Pricing in a Social World: The Influence of Non-Price Information on Hotel Choice. *Journal of Revenue and Pricing Management* 12, 5 (2013), 385–401.
- [33] Ögüt, H. and Taş, B.K.O. The Influence of Internet Customer Reviews on the Online Sales and Prices in Hotel Industry. *The Service Industries Journal* 32, 2 (2012), 197–214.
- [34] Pantelidis, I.S. Electronic Meal Experience: A Content Analysis of Online Restaurant Comments. *Cornell Hospitality Quarterly* 51, 4 (2010), 483–491.
- [35] Plaisant, C., Fekete, J.-D., and Grinstein, G. Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository. *IEEE*

*Transactions on Visualization and Computer Graphics* 14, 1 (2008), 120–134.

[36] Russell, S. and Norvig, P. *Artificial Intelligence: A Modern Approach*. Pearson, Upper Saddle River, 2009.

[37] Schuckert, M., Liu, X., and Law, R. Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *Journal of Travel & Tourism Marketing* 32, 5 (2015), 608–621.

[38] Smith, A.K., Bolton, R.N., and Wagner, J. A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research* 36, 3 (1999), 356–372.

[39] Sparks, B.A., So, K.K.F., and Bradley, G.L. Responding to Negative Online Reviews: The Effects of Hotel Responses on Customer Inferences of Trust and Concern. *Tourism Management* 53, (2016), 74–85.

[40] Stalidis, G., Karapistolis, D., and Vafeiadis, A. Marketing Decision Support Using Artificial Intelligence and Knowledge Modeling: Application to Tourist Destination Management. *Procedia - Social and Behavioral Sciences* 175, (2015), 106–113.

[41] Sun, P., Cárdenas, D.A., and Harrill, R. Chinese Customers' Evaluation of Travel Website Quality: A Decision-Tree Analysis. *Journal of Hospitality Marketing & Management* 25, 4 (2016), 476–497.

[42] Sun, S., Luo, C., and Chen, J. A Review of Natural Language Processing Techniques for Opinion Mining Systems. *Information Fusion* 36, (2017), 10–25.

[43] Sun-Young Park and Jonathan P. Allen. Responding to Online Reviews: Problem Solving and Engagement in Hotels. *Cornell Hospitality Quarterly* 54, 1 (2013), 64–73.

[44] Vermeulen, I.E. and Seegers, D. Tried and Tested: The Impact of Online Hotel Reviews on Consumer Consideration. *Tourism Management* 30, 1 (2009), 123–127.

[45] Xie, K.L., Zhang, Z., and Zhang, Z. The Business Value of Online Consumer Reviews and Management Response to Hotel Performance. *International Journal of Hospitality Management* 43, Supplement C (2014), 1–12.

[46] Xie, K.L., Zhang, Z., Zhang, Z., Singh, A., and Lee, S.K. Effects of Managerial Response on Consumer eWOM and Hotel Performance: Evidence from Tripadvisor. *International Journal of Contemporary Hospitality Management* 28, 9 (2016), 2013–2034.

[47] Ye, Q., Law, R., Gu, B., and Chen, W. 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 (2011), 634–639.

[48] Yoo, K.H. and Gretzel, U. What Motivates Consumers to Write Online Travel Reviews? *Information Technology & Tourism* 10, 4 (2008), 283–295.

[49] Yu, Y., Li, X., and Jai, T.-M. (Catherine). The Impact of Green Experience on Customer Satisfaction: Evidence from Tripadvisor. *International Journal of Contemporary Hospitality Management* 29, 5 (2017), 1340–1361.

[50] Zhang, J.J. and Mao, Z. Image of All Hotel Scales on Travel Blogs: Its Impact on Customer Loyalty. *Journal of Hospitality Marketing & Management* 21, 2 (2012), 113–131.

[51] Zhang, X. and Yu, Q. Hotel Reviews Sentiment Analysis Based on Word Vector Clustering. *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCI)*, (2017), 260–264.

[52] Zhang, Z., Zhang, Z., and Yang, Y. The Power of Expert Identity: How Website-Recognized Expert Reviews Influence Travelers' Online Rating Behavior. *Tourism Management* 55, (2016), 15–24.