# A System for Aspect-based Opinion Mining of Hotel Reviews

Isidoros Perikos, Konstantinos Kovas, Foteini Grivokostopoulou and Ioannis Hatzilygeroudis
*Department of Computer Engineering and Informatics, University of Patras, Patras, Greece*

Keywords: Opinion Mining, Sentiment Analysis, Online Reviews Analysis, Text Mining, Latent Dirichlet Allocation, Dependency Tree.

Abstract: Online travel portals are becoming important parts for sharing travel information. User generated content and information in user reviews is valuable to both travel companies and to other people and can have a substantial impact on their decision making process. The automatic analysis of used generated reviews can provide a deeper understanding of users attitudes and opinions. In this paper, we present a work on the automatic analysis of user reviews on the booking.com portal and the automatic extraction and visualization of information. An aspect based approach is followed where latent dirichlet allocation is utilized in order to model topic opinion and natural language processing techniques are used to specify the dependencies on a sentence level and determine interactions between words and aspects. Then Naïve Bayes machine learning method is used to recognize the polarity of the user's opinion utilizing the sentence's dependency triples. To evaluate the performance of our method, we collected a wide set of reviews for a series of hotels from booking.com. The results from the evaluation study are very encouraging and indicate that the system is fast, scalable and most of all accurate in analyzing user reviews and in specifying users' opinions and stance towards the characteristics of the hotels and can provide comprehensive hotel information.

## 1 INTRODUCTION

User web reviews on hotels and services constitute a valuable source of information. The growth of web 2.0 and the rise of social portals where users can share their experiences have changed the role of users and have transformed them to active producers of valuable content. User generated content and reviews on hotels in web travel portals are rich in opinions and experiences and can provide very indicative information for the characteristics of the services that users receive (Mellinas et al., 2016). Indeed, with the advent of Web 2.0, the social media and the web platforms, people have become more eager to express their opinions and share their experiences online regarding almost all aspects of their activities and experiences (Ravi & Ravi, 2015).

The advances on Web 2.0 services have caused substantial changes in the procedures of the tourism sector and have brought great innovations (Hu et al., 2017). Travel portals are becoming more and more necessary and helpful to users when deciding which hotel to choose and which service to acquire. The growth of the reviewing and feedback capabilities of the travel portals has rendered the internet the main means of seeking and acquiring travel information generated by other people. The user generated contents and hotel reviews on travel portals are growing in an exponential rate and can provide access to a wide pool of opinions and experiences of many other people (Bjørkelund et al., 2012; Tian et al., 2016). Users through the travel portals can communicate and share their perspectives and opinions and a great number of reviews is generated on a daily rate (Hu et al., 2017).

As the online commerce activity continuously grows, the role of online reviews is expected to become more and more important in the user decision making process (Moghaddam and Ester, 2012). These kinds of user reviews and opinions have always been an important piece of information that can greatly affect people decisions. Almost 95% of the people read online user generated hotel reviews when they are about to make a booking decision and more than one third of them find them of extreme importance in making their hotel decision (Ady, and Quadri-Felitti, 2015). This highlights the significance of reviews in the booking process and suggests that the analysis of the user reviews on travel portals and booking sites could assist both the travelers in choosing the proper hotel and services to acquire and also the hoteliers in monitoring and



Perikos I., Kovas K., Grivokostopoulou F. and Hatzilygeroudis I.
A System for Aspect-based Opinion Mining of Hotel Reviews.
DOI: 10.5220/0006377103880394
In *Proceedings of the 13th International Conference on Web Information Systems and Technologies (WEBIST 2017)*, pages 388-394
ISBN: 978-989-758-246-2

understating what their customers liked and what did not (Baka, 2016; Guo et al., 2017). Indeed, positive online user reviews on hotels can have a significant impact on other customers' decision-making process while on the other hand, complaints and negative reviews and comments could easily cause potential customers to lose loyalty and create negative electronic word-of-mouth (eWOM) (Au et al., 2009; Wu et al., 2010). Thus, the user generated reviews are useful for behavior analysis and their accurate analysis and understanding is of great importance for businesses. However, the automatic analysis of user generated content and reviews on hotels constitute a very challenging procedure.

In this paper, we present a work on the automatic aspect-based analysis of hotels reviews on the booking.com travel portal and the automatic extraction and visualization of information. An aspect based approach is followed. Initially, latent dirichlet allocation is utilized to model topic opinion and natural language processing approaches are used to specify the dependencies on a sentence level. Then, a Naive Bayes classifier is used to recognize the polarity of the user's opinion utilizing the sentence's dependency triples. An extensive evaluation study was conducted and revealed very promising results. The results indicate that the system is fast, scalable and most of all accurate in analyzing user reviews and in specifying users' opinions and stance towards the characteristics of hotels. Also results indicate that the system can provide comprehensive hotel information in a concise way. Several studies have shown that analysing and recognizing opinions in text is a quite complex problem which is acknowledged to be NLP-complete and the interpretation highly depends on the context and the background world knowledge (Shanahan et al., 2006).

The rest of the paper is structured as follows. In Section 2, literature is surveyed and related works are presented. In Section 3, the proposed method is illustrated and the system developed is described. In Section 4, the results from the experimental evaluation are presented and discussed. Finally, in Section 5 conclusions are presented and also main directions for future work are described.

## 2 BACKGROUND TOPICS AND RELATED WORK

### 2.1 Background Topics

When coming to user generated content and reviews on products and services, opinion mining has been studied manly on three different types of granularity, that are the document level, the sentence level and the feature level. Document level aims to find the general sentiment of the author of the text. For example, given a user hotel review, it determines whether the reviewer is positive or negative about the hotel. Sentence level focuses on individual sentences and aims to find whether a sentence expresses an opinion or not, and then whether the opinion is positive or negative (Liu, 2010). Studies have shown that both document level and sentence level analyses in general do not discover what exactly the user liked or not (Freitas and Vieira, 2013) (Liu, 2010). Feature-based approaches require in general some kind of manual tuning of various parameters something that can make them complex to port to other data (Moghaddam and Ester, 2011). Models that rely on latent variable models can overcome the aforementioned limitations mainly by learning the model parameters from the data in an automated approach. In this line, the design and the implementation of efficient aspect-based approaches is crucial for the accurate understanding of people thoughts and opinions towards each aspect of the hotel services they received and experienced. Indeed, although some traveling portals ask customers to express an overall rating that can be denoted in stars, focusing just on the overall rate of the hotel will not be sufficient in any case for a user to make a decision (Moghaddam and Ester, 2012).

### 2.2 Related Work

The analysis of user generated content and reviews in web travel portals has attracted the increasing attention of researchers in computer science, natural language processing and sentiment analysis mainly during the last decade. In the literature, there is an increasing research interest and many studies have been made on the design of methods and the development of systems for the mining of opinion in text. A thorough and complete description of approaches and techniques can be found in the literature (Medhat et al., 2014; Liu and Zhang, 2012; Vinodhini and Chandrasekaran, 2012). Several works study the way people express opinions and try to identify opinions in forums, social media and travel portals (Cercel and Trausan-Matu, 2014; Liu, 2015).

In the work presented in (Bjørkelund et al., 2012), authors present a system that assists tourists in choosing the appropriate hotel by visualizing polarity information on Google maps. Authors use the burst technique in order to find changes in users' opinions and visualizations represent good and bad hotels in graphical map forms. In the work presented in (Berezina et al., 2016), authors examine the underpinnings of hotel users that were satisfied and dissatisfied using text-mining approaches on the online hotel reviews. Authors achieve this mainly by comparing the online hotel reviews of satisfied users to the reviews of others and those of dissatisfied using text-link analysis of the reviews. Authors in the work presented in (Wang and Fu, 2010), present a model that extracts data from NTCIR- 6 opinion corpus. The Chi- Square metric is utilized in order to extract the subjective cues from the customer reviews which are then used to find the subjectivity density from training data. Then, a Naïve Bayes classifier is applied for the classification of subjectivity reporting satisfactory results. In the work presented in (Ye et al., 2009) authors present an approach on analyzing user reviews on travel sites using supervised machine learning approaches such as Support vector machines and Naïve Bayes. The N-gram model is used to represent text and the evaluation results report quite satisfactory accuracy which is above 80%.

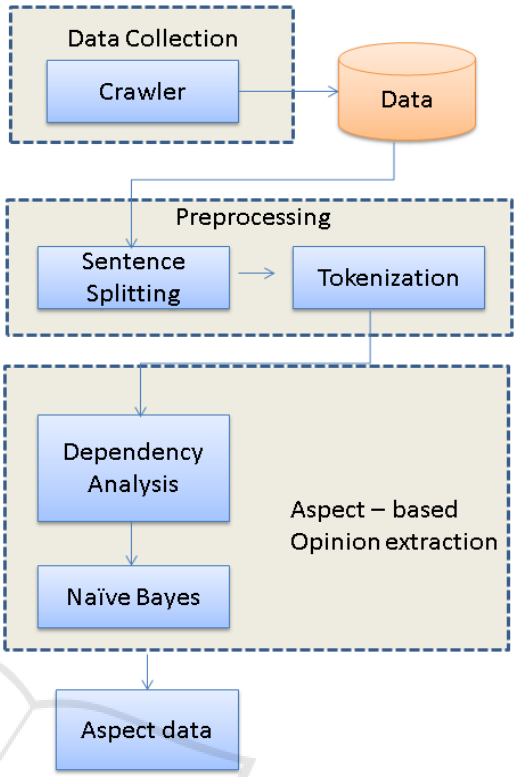


Figure 1: An overview of the workflow of the approach.

## 3 MINING OPINIONS OF HOTEL REVIEWS

In this section, the system developed to analyze user reviews on booking.com is presented and its functionality is illustrated. It overall workflow is depicted in Figure 1.

Initially, the system's crawler is used to access hotel pages on booking.com, and to collect and index the user reviews on them. The collected reviews undergo a series of analyses. Latent Dirichlet Allocation (LDA) is used to model the topic and specify main aspects. The analysis of the textual reviews is conducted on sentence level, so a given review is split in sentences. Each sentence is analyzed by Stanford parser tool and the dependency tree indicating the interaction among the sentence words is created. Sentences that contain key aspects among the dependencies of the sentence words are forwarded to the Naïve Bayes classifier, trained to recognize the existence and the polarity of users opinions towards the aspect.

### 3.1 Indexing and Preprocessing of Reviews

The first part of the system consists of the crawler developed to automatically assess, retrieve and index user reviews on the booking.com travel portal. The crawler follows a hotel-based retrieval process where hotels are accessed and users' reviews towards the hotel are retrieved and indexed. It automatically extracts all reviews on a single hotel using the hotel's information page on the booking.com portal. Then the crawler can navigate through hotels to extract all users' reviews. The textual content of the reviews are stored on the system's database. For each review, are retrieved the date that it was created, the full textual content of the review, the title of the review (if any) and also user related information such as the type of user's travel (e.g. business, leisure, solo, couple, family).

The retrieved reviews are indexed in the database of the system and natural language analysis processes are conducted. Initially, sentences of the review are split and then each sentence is separately handled by the system.

### 3.2 Topic Modelling using LDA

Latent Dirichlet Allocation is a powerful and widely used approach for the modelling of topics and it infers hidden topic structure from the reviews based on a probabilistic framework. The idea behind the method utilized in the approach is that all hotel reviews share the same topic set but each hotel's review corpus exhibits a different probabilistic mixture of those topics. In general, the LDA model assumes that there is a hidden structure which consists of a set of topics in the whole textual dataset. The LDA algorithm utilizes and relies on the co-occurrence of the words in the different reviews in order to infer the underling hidden structure. The model computes the posterior distribution of the unobserved variables in the corpus of the review. So, for a specified part of training reviews, the LDA specifies two main outputs. The first output is a set of topics which are associated with the set of words, which contribute to the topic via their weights. The second output consists of a set of reviews with a vector of weight values displaying the probability of a review containing a specific topic. After the specification of the main aspects of the reviews corpus, the analysis of each textual review is conducted.

## 3.3 Textual Analysis of Reviews

The system analyzes the structure of each sentence with the use of the Stanford parser (De Marneffe et al. 2006). The Stanford parser is used to determine the grammatical structure of a sentence and specify for each word its base form (lemma) and its grammatical role in the sentence. Also, it specifies the relationships between the sentence's words and determines the corresponding dependencies, which provide remarkable assistance in sentence analysis. The dependency tree represents the complete grammatical relations between the sentence's words in a concise tree based approach. Dependencies in general indicate the way that sentence's words are connected and interact with each other. When the sentence morphosyntactic analysis is completed and the dependency tree is created, special parts of the dependency tree and specific words are further analyzed. The dependency tree is analyzed and the relationships and types of interactions/connections between the sentence words are examined. As an example case, let us consider the sentence: "The morning breakfast was nicely presented". The dependencies of the sentence indicating the way that the sentence words are connected are the following:

"The morning breakfast was nicely presented."
((‘presented’, ‘VBN’), ‘nsubjpass’, (‘breakfast’, ’NN’)),
((‘breakfast’,‘NN’),‘det’, (‘The’, ‘DT’)),
((‘breakfast’,‘NN’),‘compound’, (‘morning’, ‘NN’)),
((‘presented’,‘VBN’),‘auxpass’, (‘was’, ‘VBD’)),
((‘presented’,‘VBN’),‘advmod’, (‘nicely’, ‘RB’)).

and the dependency tree of the sentence is depicted in Figure 2.

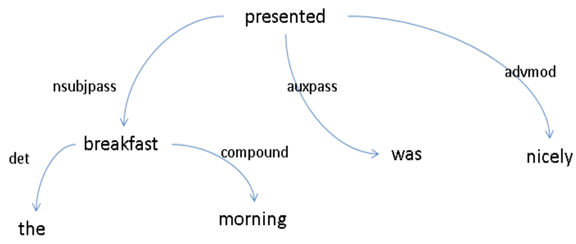


Figure 2: Dependency tree of the sentence.

The specification of the dependencies is a quite important step of the proposed approach. After the analysis of the user's review and the specification of a person's references to specific aspects of the hotel, the sentence that contains that aspects, is deeper analysed and the word dependencies are specified. These word dependencies can provide indicative clues to the machine learning algorithm in order to specify user's attitude towards the specific aspect he/she mentioned and addressed in his/her review.

## 3.4 Polarity Recognition

The recognition of the user's opinion towards each aspect is conducted using the Naïve Bayes classifier. After the analysis of a user's comment on a sentence level, and the recognition of aspect in a sentence, the sentence's words and the corresponding dependencies are given as input to the Naïve Bayes classifier in order to classify the sentence into the proper polarity category. The output of the classifier concerns the one of the three categories neutral, positive, negative that classifies the user's textual mention to the aspect. The integration of the Naïve Bayes method in the system was decided based on its performance during the experimental phase.

Naïve Bayes is a widely used model for classification and it can achieve high accuracy when it comes to text categorization. It is based on Bayes theorem and assumes that documented words are generated through a probability mechanism. The lexical units of a textual corpus are labelled with a specific category or with a specific category set and are processed computationally. During this processing, each document is treated as a bag of words, and the document is assumed not to have any internal word structure, and words do not have any interconnection. The Bayesian formula calculates the probability of a defined polarity class as:

$$P(c|s) = \frac{P(c)P(s|c)}{P(s)}$$

In the formula $P(c)$ represents the probability that a sentence belongs to category '*c'*, $P(s)$ represents the probability of the occurrence of sentence '*s'*, $P(s|c)$ represents the probability sentence '*s'* to belong to the category '*c'* and finally, $P(c|s)$ represents the probability that given the sentence '*s'*, the sentence belongs to category *c*. The term $P(s|c)$ can be calculated by taking into account that the conditional probabilities of occurrences of sentence words given category '*c'*, as follows:

$$P(s|c) = \prod_{1 \le k \le n} P(s_k\,|c)$$

where $P(s_k|c)$ is the probability that the term '$s_k$' occurs given the category '*c'*, and 'n' represents the length of the sentence '*s'*.

### 3.5 Hotel Statistics

After the analysis of all users' reviews on a hotel, the system can provide indicative aspect-based, hotel centric statistics of users' reviews. This piece of information is visualized in various graphical forms and can represent the summative opinions of users. Statistics and visualizations could be assistive to both users and hotel administrative staff to monitor their customer's opinions across the characteristics of the hotel services offered to them.

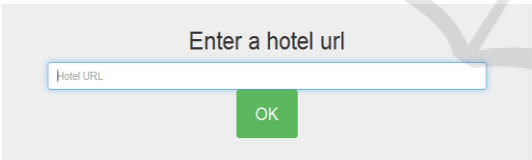


Figure 3: Analysis of the reviews for specific hotel.

Initially, given an inquiry about the analysis of a hotel's reviews on the booking.com travel portal (as presented in Figure 3), the system retrieves and analyzes each review and specifies the user's attitude in the review towards the aspects he/she mentioned in his/her review. In the context of the system's functionality the number of the aspects specified by LDA method was 15. In an example case hotel, the presentence of the aspects mentioned in the users' reviews are depicted by the system in figures such as Figure 4.

The most frequently mentioned aspect in users reviews was specified to be the room and after that the hotel location, the staff, the cleanness and the bed. Given the 15 aspects specified, the less

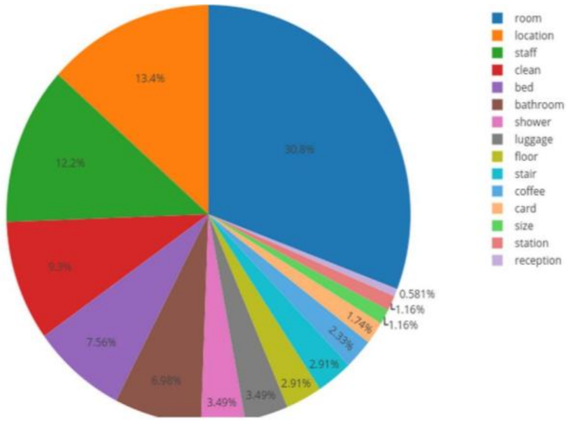


Figure 4: Percentage of aspects that were indicated in user reviews.

frequently mentioned aspect was the reception, and the (existence of nearby) station. Also, the system illustrates the rates of the polarity users' opinions.

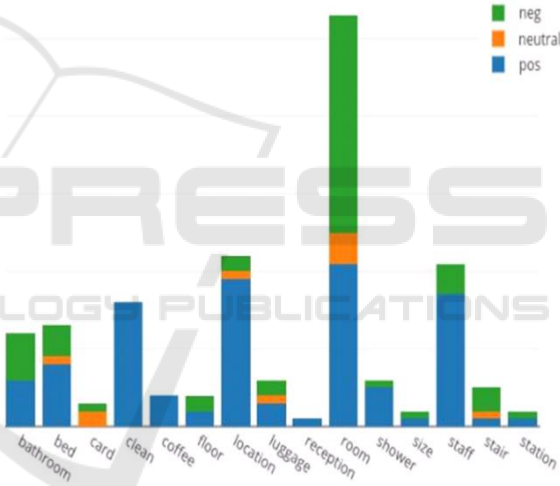


Figure 5: Rates of the polarity of users' opinions towards each aspect.

Given the above example diagram, almost half of the users' mentions to the room aspect were specified to be negative, approximately 40% were specified to be positive and only 10% were specified to be neutral.

## 4 EXPERIMENTAL EVALUATION

In this section, we present the experimental study that was designed and conducted and also the results collected. Initially, we describe the datasets that were used in the study and then, the results and the performance of the system.

## 4.1 Dataset

A dataset that consisted of user generated reviews on various hotels on booking.com was formulated. Initially, the crawler of the system accessed and obtained hotel reviews for various hotels and, for the needs of the study, a total number of 1131 reviews for 27 hotels were used to formulate the dataset. For a hotel centric scope, we indexed hotels that had more than 25 reviews written in English language.

After that, an expert human annotator was used to read each review and specify the user's opinion polarity for the aspects that he/she mentions. These annotations made by the human expert were used as gold standard for the evaluation of the system's performance.

## 4.2 Results

In the context of the study the annotated dataset of reviews that had their polarity on aspects determined by the expert, was used and the performance of the system was assessed. Given that we have a multiple class output variable, we use the following metrics: average accuracy, precision and F-measure (Sokolova and Lapalme, 2009). In general, for a given class, the precision metric is the fraction of the instances that were classified to a class and actually belong to that class, while recall is the fraction of instances that belong to class A that were correctly classified to that class. The F-measure metric combines the recall and precision values and is calculated as follows:

$$F\_measure = 2 \times \frac{prec \times rec}{prec + rec}$$

Table 1: Performance in detecting the polarity of the reviews.

| Metric | Value |
|---|---|
| Precision | 0.72 |
| Recall | 0.71 |
| F-measure | 0.72 |

The results indicate that the system's performance is quite encouraging. The system specified correctly the category of the polarity of users' mentions on aspects in almost 72% of the cases. A deeper examination of the systems performance revealed the cases that were misclassified concerned neutral opinions that were classified as polarized. Also, the system's performance in separating positive from negative mentions of aspects is very good. Results show that the system is performing quite well and it can be utilized to provide an analysis of users reviews in order to assist both hotel businesses in understanding their customers thoughts and future travelers in their decision making process.

# 5 CONCLUSIONS

In this paper, we present a work on the automatic aspect-based analysis of user reviews on the booking.com travel portal and the automatic extraction and visualization of information. An aspect based approach is followed. Initially, latent dirichlet allocation is utilized to model topic opinion and natural language processing approaches are used to specify the dependencies on a sentence level. Then machine learning approaches such as Naïve Bayes is used to recognize the polarity of the user's opinion utilizing the sentence's dependency triples. An extensive evaluation study was conducted and revealed very promising results. The results are very encouraging and indicate that the system is fast, scalable and most of all accurate in analyzing user reviews and in specifying users' opinions and stance towards the characteristics of the hotels. Also results showed that the system can provide comprehensive hotel information in a concise way.

Future work will focus and examine specific directions. Initially, a bigger scale evaluation will be conducted to provide a more complete insight of the performance of the system. Another direction for future work is to examine ensemble classification schemas that will combine discrete classifiers under different ensemble approaches with the aim to enhance the classification performance. Furthermore, another direction concerns the examination of temporal and geographical characteristics with the aim to capture the way that opinions on hotel's aspects change over time.

# REFERENCES

Ady, M., & Quadri-Felitti, D. (2015). Consumer research identifies how to present travel review content for more bookings.

Au, N., Buhalis, D., & Law, R. (2009). Complaints on the online environment—the case of Hong Kong hotels. *Information and communication technologies in tourism 2009, 73-85.*

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.

Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. *Journal of Hospitality Marketing & Management*, 25(1), 1-24.

Bjørkelund, E., Burnett, T. H., & Nørvåg, K. (2012, December). A study of opinion mining and visualization of hotel reviews. *In Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services (pp. 229-238). ACM.*

Blei, D. M., & Lafferty, J. D. (2009). *Topic models. Text mining: classification, clustering, and applications, ed. Srivastava, A.N., & Sahami M. CRC Press, Boca Raton, FL, 10(71), 34.*

Cercel, D. C., & Trausan-Matu, S. (2014). Opinion Propagation in Online Social Networks: A Survey. *In Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14) (p. 11). ACM.*

De Marneffe, M. C., MacCartney, B., & Manning, C. D. (2006, May). Generating typed dependency parses from phrase structure parses. In Proceedings of LREC (Vol. 6, No. 2006, pp. 449-454).

Freitas, L. A., & Vieira, R. (2013, May). Ontology based feature level opinion mining for portuguese reviews. *In Proceedings of the 22nd International Conference on World Wide Web (pp. 367-370). ACM.*

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.*

Han, H. J., Mankad, S., Gavirneni, N., & Verma, R. (2016). What guests really think of your hotel: Text analytics of online customer reviews. Cornell Hospitality Report, 16 (2), 3-17.

Hu, Y. H., Chen, Y. L., & Chou, H. L. (2017). Opinion mining from online hotel reviews–A text summarization approach. *Information Processing & Management*, *53*(2), 436-449.

Liu, B. (2010). Sentiment analysis and subjectivity. In *Handbook of Natural Language Processing, Second Edition* (pp. 627-666). Chapman and Hall/CRC.

Liu, B., 2015. Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press

Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463). Springer US.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113.

Mellinas, J. P., María-Dolores, S. M. M., & García, J. J. B. (2016). Effects of the Booking. com scoring system. Tourism Management, 57, 80-83.

Mellinas, J. P., María-Dolores, S. M. M., & García, J. J. B. (2016). The use of social media in hotels as indicator of efficient management. Tourism & Management Studies, 12(2), 78-83.

Moghaddam, S., & Ester, M. (2011, July). ILDA: interdependent LDA model for learning latent aspects and their ratings from online product reviews. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 665-674). ACM.

Moghaddam, S., & Ester, M. (2012, October). On the design of LDA models for aspect-based opinion mining. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 803-812). ACM.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*, *89*, 14-46.

Shanahan, J. G., Qu, Y., & Wiebe, J., 2006. Computing attitude and affect in text: theory and applications, vol. 20, Springer.

Tian, X., He, W., Tao, R., & Akula, V. (2016). Mining Online Hotel Reviews: A Case Study from Hotels in China.

Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. *International Journal*, *2*(6), 282-292.

Wang, X., & Fu, G. H. (2010, July). Chinese subjectivity detection using a sentiment density-based naive Bayesian classifier. In *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on* (Vol. 6, pp. 3299-3304). IEEE.

Wu, Y., Wei, F., Liu, S., Au, N., Cui, W., Zhou, H., & Qu, H. (2010). OpinionSeer: interactive visualization of hotel customer feedback. *IEEE transactions on visualization and computer graphics*, *16*(6), 1109-1118.

Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, *36*(3), 6527-6535.