

# **Hotel Room Allocation Model for Maximizing Positive Reviews: A Sentiment Analysis Framework**

**Literature Review**

**Charlie Martin**

**Word Count - 2107**

## **Introduction**

Online reviews are of increasing importance for the hospitality and tourism industry. This is supported in the work of Schuckert, Liu and Law (2015), who found that online reviews play a critical role in sales for the industry. To add to this, Kaplan and Haenlein (2010) concluded that online reviews have become the word of mouth of the digital age. In the words of Duan et al. (2008), word of mouth has long been recognised as the most influential resource of transmission in society. Clearly, the importance of online reviews as a resource has been well established. For hotels, extracting maximum information and value from this resource can provide a competitive edge, but extracting meaningful insights from a catalogue of unorganised personal experiences is no mean feat.

The aim of this project is to provide a framework for extracting insights from online reviews of European hotels with a sentiment analysis model. Additionally, this project will propose a new perspective for gaining insights via analysis on both the hotel and reviewer attributes. Furthermore, hotel room allocation will be discussed with the suggestion of including said insights in future room allocations models. This project will suggest a room allocation model with the aim of maximizing positive reviews, which as shown in the literature discussed will benefit hotels via increased revenue. This is the literature review for the mentioned project, and will consist of an in-depth literature review, followed by the research questions and hypotheses.

## **Literature review**

### **Review Impact**

To start the literature review, this first section will go into further detail of the importance of online reviews. The importance of reviews has been discussed and analysed in many industries. For example, in a study of the restaurant industry by Zhang et al. (2010), where it is noted that online reviews are now an emerging phenomenon and are playing an increasingly important role in consumer purchase decisions. This study concluded that consumer generated ratings are positively associated with the popularity of said restaurants, increasing their reputation and revenue, and had a focus on a certain reviewer attribute, being a consumer, or a critic. These attributes were found to significantly affect the impact of the review. To take this further, Lee and Ro (2016) studied the ability of online reviews to change consumer attitudes and concluded that online reviews are so powerful they can change consumer attitudes from previous experiences. This study also took consumer attributes into account, namely their prior experience and knowledge levels.

In Zhu and Zhang's (2010) work, the goal of the study was to provide positive reconciliation of previous mixed results. This study was on the video game industry specifically, and concluded again that both product and consumer characteristics moderate the influence of online consumer reviews on product sales and suggest that previous contradictory results are likely due to data deficiencies and variations. This conclusion is widely accepted, being cited over 3000 times since.

The impact of online reviews in the hotel industry cannot be overstated. As found by Vermeulen and Seegers (2009) 84% of hotel bookers had their choices impacted by what they saw in online reviews. This study also noted that all reviews, both positive and negative, benefited the hotel via increased brand exposure. This is contradicted by the work of Sparks and Browning (2011), in which it is found that consumers tend to be more influenced by early negative information, particularly if it is easy to process. However, there is more evidence to support that more reviews are always better, such as in the study on Belgium hotels by De Pelsmacker (2018) where it is found that review volume drives room occupancy.

The key takeaway here is that reviews are of increasing importance, and to note the trend of the inclusion of reviewer attributes. From restaurants to gaming, many sectors have included consumer attributes in review analysis, with significant results – all except the hotel industry.

### **Fake Reviews**

For businesses, user reviews can play a significant role in determining the revenue for an organisation. This is confirmed in the work of Mohawesh et al. (2021), whose study on reviews found they directly impact a company's reputation and profitability. This study also notes the dangers of fake reviews often used to exploit consumer purchasing decisions. The prevalence of fake reviews is supported in another study by He et al. (2022), in which an extensive study is carried out into the market for fake reviews on amazon, which is found to be growing. This study also finds that these fake reviews have a significant impact on short term product and company performance. With the growing problem of fake reviews impact company performance in addition to analysis results, various fake review detection methods have been studied. Elmogy et al. (2021) proposes a supervised machine learning method using a combination of natural language processing and the K nearest neighbour classification model to detect fake reviews. This study finds the approach to be successful in fake review detection. This strategy is also shown to be successful in the work of Salminen et al. (2022), who successfully deploy classification techniques to detect fake reviews with near perfect accuracy. For this study, fake reviews are likely to be present and could potentially expose analysis results to bias, ruining the integrity of the dataset and analysis results. To prevent this, a supervised machine learning classification strategy will be deployed to remove fake reviews in data processing.

### **Sentiment Analysis**

Sentiment analysis is a relatively new term and is one of the fastest growing research areas in computer science, partly due to online reviews. As noted by As Mäntylä, Graziotin, and Kuutla (2018), the availability of subjective texts online has caused an outbreak in computer-based sentiment analysis, with 99% of related papers being published after 2004. This study provided a review of some of the most cited and influential studies in the field, some of which will be discussed. One of the most cited papers on sentiment analysis is Mining and Summarizing Customer reviews. In this study, Hu and Liu (2004) mention some of the key problems with the increasing popularity of reviews, namely too many reviews to analyse and reviews being too long to digest. This study discusses techniques for counteracting these problems, such as sentiment classification and feature selection with machine learning. Hu and Liu conclude that the proposed techniques are very promising and will be of increasing importance in the future for both consumers and manufacturers.

A key feature of sentiment analysis is to differentiate between the review sentiment and review rating. This is found in the work of Pang et al. (2002) which found that a crucial characteristic of the rapidly growing online discussion sites is the overall sentiment. This is further supported by Hu et al. (2014) where it is suggested that the previous focus on purely numeric ratings is due to the difficulty of extracting sentiment. In this study of amazon reviews, it is found that the review ratings in fact have no direct impact on sales, but indirectly impact sales through the sentiment. These studies highlight the need for sophisticated, computer-based analysis on sentiment to utilize the growing resource of reviews. This study also found that sentiment classification can be beneficial specifically in business intelligence applications, with machine learning techniques out-performing human baselines. One critical part of applying this kind of analysis is feature selection, which will be discussed next.

## **Feature Selection and Machine Learning Algorithms**

Feature selection is a crucial part of sentiment analysis, as found in many studies. For example, Harish and Revanasiddappa (2017) stated that feature selection is a strategy that can be used to increase categorization accuracy, effectiveness, and computational efficiency. Furthermore, in a review of the literature, Asgher et al. (2014) found that feature proper feature selection techniques play a key role in identifying relevant attributes for analysis and increasing accuracy. This review also found that feature selection is usually categorized within 4 main types, with NLP (natural language processing) being the most common. This is the only category of feature selection relevant to this literature review, and thus the only one mentioned. Specifically, this project will only use bag of words and TF-IDF feature selection.

One study by Jing et al. (2002) notes the effectiveness of TF-IDF for feature extraction, with an accuracy of 88% using Naïve Bayes, compared to 76% using IDF. Bag of words is one of the most common feature extraction methods. As noted by Qader et al. (2019) its strength lies in its simplicity and therefore applicability to a wide range of classification models while still producing strong results. Building on this, an extensive study was carried out by Ahuja et al. (2019) where feature extraction methods were tested with various machine learning algorithms and concluded that TF-IDF performs on average 3-4% better. This study also concluded that logistic regression performed better for sentiment analysis, which is supported in a similar study by Prabhat and Khullar (2017).

## **Sentiment Analysis in the Travel and Tourism Industry**

Sentiment analysis is certainly a popular topic within this industry, and many studies have been carried out. Generally, the focus of these studies is to identify key features of both positive and negative reviews and prescribe advice to hotels accordingly. Results vary between studies, for example Tran et al. (2019) found that location and restaurants were the two most important aspects for positive reviews, which is contradicted by Sodanil (2016), where it was found that the room was significantly more impactful on positive reviews than restaurants. Most notable for this project is a study by Chang et al (2020) on luxury hotel reviews. As mentioned in this study, customers for luxury hotels tend to focus more on details as upscale amenities are expected, making these reviews the ideal dataset for such analysis. It is also mentioned in this study that future research may benefit from further investigation of customer profiles, which is a gap in the literature of this field in general, as previously noted. Hence, one of the goals of this project.

It is also important to touch on the literature surrounding hotel room allocation. Hotels are in a unique position of seeing customer profiles before they arrive, allowing for additional opportunity to manage resources, namely room allocations. The literature surrounding this notes the importance of room allocation, and even refers to it as a “statistics game” in a study by Song et al. (2010). This study proposes a static game model to solve the hotel room inventory problem. Additionally, Aydin and Birbil (2018) propose a mathematical model to optimize hotel room allocation with long stay booking. These studies all note that correct hotel allocation and resource management can benefit hotels via increased profitability and note the importance of prioritizing bookings accordingly. Linking back to the review impact section, one study mentioned found that review volume drives room occupancy, which could significantly impact a room allocation model. Thus far, there have been no links between hotel room allocation and the new resource of online word of mouth. This project will aim to create a quick and practical framework for hotel managers to use, where all bookings are weighted and thus prioritised based on the results from the sentiment analysis. This will allow hotels to allocate rooms accordingly to maximize positive reviews, or minimize negative ones, which has been shown to benefit the organisations sales, reputation, and profitability.

## **Summary**

To summarize, it is well established within the literature that online reviews are a growing resource and only becoming more important to various industries. It is also well established that sentiment analysis is an effective way of analysing reviews. The key theme regarding analysis of online reviews in general is that reviewer attributes are often considered, except in the hotel industry. Additionally, various machine learning methods have been studied with different feature selection methods with varying results. Different methods will need to be tested and evaluated regarding the dataset for the hotel industry. Finally, there is a missing link between online reviews and hotel room allocation strategy and sentiment analysis may be useful in joining these two bodies of literature.

## **Research Questions and Hypotheses**

Using the previous literature review and looking at the gaps in the literature, 3 main research questions for this project have been identified. These are:

- Impact of some common guest attributes on overall review sentiment
- Key attributes in both positive and negative reviews for luxury hotels
- Most accurate feature extraction method and machine learning algorithm combination

This project will be a quantitative study; thus, the following hypotheses are proposed.

Primary Hypothesis:

- H0 – Reviewer attributes have no impact on review sentiment.
- H1 – Reviewer attributes impact review sentiment.

Other hypotheses:

- Location and room are most important for positive reviews.
- Customer service is the biggest indicator for negative reviews.
- Bookings being 'leisure' or 'business' will significantly impact review sentiment.

## References

- Ahuja, R., Chug, A., Kohli, S., Gupta, S. and Ahuja, P., 2019. The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, 152, pp.341-348.
- Asghar, M.Z., Khan, A., Ahmad, S. and Kundi, F.M., 2014. A review of feature extraction in sentiment analysis. *Journal of Basic and Applied Scientific Research*, 4(3), pp.181-186
- Aydin, N. and Birbil, S.I., 2018. Decomposition methods for dynamic room allocation in hotel revenue management. *European journal of operational research*, 271(1), pp.179-192.
- Chang, V., Liu, L., Xu, Q., Li, T. and Hsu, C.H., 2023. An improved model for sentiment analysis on luxury hotel review. *Expert Systems*, 40(2), p.e12580.
- De Pelsmacker, P., Van Tilburg, S. and Holthof, C., 2018. Digital marketing strategies, online reviews and hotel performance. *International Journal of Hospitality Management*, 72, pp.47-55.
- Duan, W., Gu, B. and Whinston, A.B., 2008. Do online reviews matter?—An empirical investigation of panel data. *Decision support systems*, 45(4), pp.1007-1016
- Elmoghy, A.M., Tariq, U., Ammar, M. and Ibrahim, A., 2021. Fake reviews detection using supervised machine learning. *International Journal of Advanced Computer Science and Applications*, 12(1).
- Harish, B.S. and Revanasiddappa, M.B., 2017. A comprehensive survey on various feature selection methods to categorize text documents. *International Journal of Computer Applications*, 164(8), pp.1-7.
- He, S., Hollenbeck, B. and Proserpio, D., 2022. The market for fake reviews. *Marketing Science*, 41(5), pp.896-921.
- Hu, M. and Liu, B., 2004, August. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).
- Hu, N., Koh, N.S. and Reddy, S.K., 2014. Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision support systems*, 57, pp.42-53.
- Jing, L.P., Huang, H.K. and Shi, H.B., 2002, November. Improved feature selection approach TFIDF in text mining. In *Proceedings. International Conference on Machine Learning and Cybernetics* (Vol. 2, pp. 944-946). IEEE.
- Kaplan, Andreas M., and Michael Haenlein. "Users of the world, unite! The challenges and opportunities of Social Media." *Business horizons* 53, no. 1 (2010): 59-68.
- Lee, S.H. and Ro, H., 2016. The impact of online reviews on attitude changes: The differential effects of review attributes and consumer knowledge. *International Journal of Hospitality Management*, 56, pp.1-9.
- Mäntylä, M.V., Graziotin, D. and Kuuttila, M., 2018. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, pp.16-32.

Mohawesh, R., Xu, S., Tran, S.N., Ollington, R., Springer, M., Jararweh, Y. and Maqsood, S., 2021. Fake reviews detection: A survey. *IEEE Access*, 9, pp.65771-65802.

Pang, B., Lee, L. and Vaithyanathan, S., 2002. Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.

Prabhat, A. and Khullar, V., 2017, January. Sentiment classification on big data using Naïve Bayes and logistic regression. In *2017 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-5). IEEE.

Qader, W.A., Ameen, M.M. and Ahmed, B.I., 2019, June. An overview of bag of words; importance, implementation, applications, and challenges. In *2019 international engineering conference (IEC)* (pp. 200-204). IEEE.

Salminen, J., Kandpal, C., Kamel, A.M., Jung, S.G. and Jansen, B.J., 2022. Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services*, 64, p.102771.

Schuckert, M., Liu, X. and Law, R., 2015. Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, 32(5), pp.608-621.

Sodanil, M., 2016. Multi-language sentiment analysis for hotel reviews. In *MATEC Web of Conferences* (Vol. 75, p. 03002). EDP Sciences.

Song, J., Parlar, M. and Yuan, Q., 2010, August. Decision Making of Hotel Room Allocation: A Statistic Game. In *2010 International Conference on Internet Technology and Applications* (pp. 1-4). IEEE.

Sparks, B.A. and Browning, V., 2011. The impact of online reviews on hotel booking intentions and perception of trust. *Tourism management*, 32(6), pp.1310-1323.

Tran, T., Ba, H. and Huynh, V.N., 2019. Measuring hotel review sentiment: An aspect-based sentiment analysis approach. In *Integrated Uncertainty in Knowledge Modelling and Decision Making: 7th International Symposium, IUKM 2019, Nara, Japan, March 27–29, 2019, Proceedings 7* (pp. 393-405). Springer International Publishing.

Vermeulen, I.E. and Seegers, D., 2009. Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism management*, 30(1), pp.123-127.

Zhang, Z., Ye, Q., Law, R. and Li, Y., 2010. The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, 29(4), pp.694-700.

Zhu, F. and Zhang, X., 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2), pp.133-148.