CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain
DOI: https://doi.org/10.4995/CARMA2026.2026.22447

Using Sentiment Analysis to Detect Operational Issues in a Holiday Rental Company

Mafalda Ponte[1], Ana Catarina Nunes[2], Pedro Ramos[3], Francisco Cruz[4], Paula Vicente[1]

[1]Instituto Universitário de Lisboa (Iscte-IUL), Business Research Unit (BRU-IUL), Lisboa, Portugal, [2]CMAF-CIO, Universidade de Lisboa, Portugal, [3]Instituto Universitário de Lisboa (Iscte-IUL), Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura (ISTAR-Iscte), Lisboa, Portugal, [4]Feels Like Home Group, Lisboa, Portugal.

How to cite: Ponte, M; Nunes, A. C.; Ramos, P.; Cruz; F.; Vicente, P. (2026). Using Sentiment Analysis to Detect Operational Issues in a Holiday Rental Company. In: 8th International Conference on Advanced Research Methods and Analytics (CARMA 2026). Valencia, 1-3 July 2026. https://doi.org/10.4995/CARMA2026.2026.22447

Abstract

Online reviews provide a wealth of information and feedback that businesses can use to help identify issues that require short term action and aid medium/long term decision-making. Numerical information such as global valoration and scores in specific evaluation aspects can be complemented by the insights from textual fields of the reviews. This study uses text analytics to address the large volume of online review data for a holiday rental company. Specifically, sentiment analysis is used to assess the polarity of the different text fields and topic detection is used to identify the underlying topics of the text data. This analysis allowed to identify relevant topics that are not covered by the evaluation aspects or correspond to specific issues related to more general evaluation aspects. Some reviews were also identified with large discrepancies between sentiment analysis and actual global valoration, which warrants attention from decision-makers.

Keywords: *Text Analytics; Text Mining; Topic Detection; Sentiment Analysis; Small and Medium Businesses*

1. Introduction

Customer feedback is vital for businesses to find and enhance their strengths, as well as identify and correct their weaknesses. This is especially important for small and medium businesses, which need to make an extra effort to differentiate themselves from competition. Online reviews provide a wealth of information that can be used for this purpose. Quantitative information, such as the global valoration of the reviews, can give a general overview of the customers'

perception of the service, with sustained downward trends in these indicators signaling issues that must be addressed. Text information provides a more nuanced understanding of customers' feedback, both in terms of sentiment towards the business, as well as in terms of topics that may be relevant to decision-making; be it operational and short-term or strategic and medium/long-term. Even for small and medium businesses there is availability of such data, but its treatment raises challenges. Online reviews data tends to be high volume, new data arriving constantly to the point of manual treatment being time consuming and unrealistic. On the other hand, this type of data also tends to be unstructured, in the sense that the information available is organized differently depending on the source, which must be taken into account when processing it automatically.

In this context, this paper provides a case study on using Natural Language Processing (NLP) tools, namely topic modeling and sentiment analysis, together with quantitative analysis to derive actionable insights from online reviews pertaining to Feels Like Home (FLH), a holiday rental company. Section 2 provides a brief literature review on the application of NLP to mining online reviews. Section 3 details the main challenges of the data available for study and the measures that were taken to deal with such challenges. Section 4 expands on the methodologies that were applied and main results that were obtained. Finally, Section 5 describes the main conclusions of this work, and gives some insights on how they can be put into practice.

2. Literature Review

Natural Language Processing is a branch of Artificial Intelligence (AI) that is used to derive patterns from human language. One of its most common techniques is Text Analytics, which includes the tasks of text acquisition and aggregation, text preprocessing, text exploration and text modeling (Hair et al., 2021).

Text data can be obtained from several sources, such as reviews from the company's e-commerce website, third-party review websites, surveys, call center transcripts or social media posts. The data needed for study depends on the business question(s) the study is trying to address. This type of data usually needs especially rigorous cleaning and preprocessing. Punctuation, variations in capital letters, typos, emojis and special characters add noise to the data and pose challenges to its interpretation. Thus, text preprocessing includes steps such as removal of special characters, whitespace and commonly used words (stopwords), segmentation of the text into words (tokenization), spelling correction, reducing words into their root form (stemming and lemmatization), determining the part of speech for each word (part-of-speech tagging) and converting text into numerical form (text vectorization). Text exploration gives a first overview of the contents of the text, using graphic representations such as Word Cloud and Term Frequency Chart to display the most frequent words (or root words) in the text data. For more informative representations, decisions must be made regarding part-of-speech (focusing

on nouns, verbs and adjectives and ignoring adverbs, for instance), and stopwords (manually expanding the stopword list to include the company name and other naturally frequent words that do not add meaning to the analysis). Finally, the preprocessed text data is used to build models to derive insights for decision-making.

Bi et al., (2024) provide a comprehensive review on text mining in the context of hospitality. The authors expand on common methodologies remarking that the two more common across the literature are topic modeling and sentiment analysis. Topic modeling/detection algorithms generate clusters of (key)words representative of the text, referred to as topics, in order to minimize variance within a given topic and maximize dissimilarity between different topics. Popular topic detection models are the Latent Dirichelet Allocation algorithm (LDA) (Pritchard et al., 2000 and Blei et al., 2003) and the Latent Semantic Analysis algorithm (Deerwester et al., 1990). The former generates the clusters considering correlations between words and ignoring their ordering in the text, while the latter uses Single Value Decomposition as an aide for cluster generation, by identifying words with similar meanings and thus reducing dimensionality. Finally, sentiment analysis algorithms try to assess the underlying emotions of a text through a polarity scale from –1 to +1, where negative values correspond to negative emotions and positive values correspond to positive emotions.

Bi et al., (2024) also provide a comprehensive list of tasks and business questions that can be tackled in the context of text analytics for tourism and hospitality. These include understanding tourist satisfaction, loyalty, decision-making, but also behaviors and trends across companies. Relevant stakeholders in these studies tend to be hotels and other agents related to hospitality (see, for instance, Le et al. (2025) and Zhao et al., (2024)), but other tourism-related agents, such as airlines, also derive insights from text analytics (see, for instance, Lucini et al. (2020)).

3. Data and Methodology

This section presents a description of the source data and the methodology followed in the study.

The dataset used in this study is the result of combining reviews from the platforms Booking, Airbnb, HomeAway, and Feels Like Home, concerning the reviews of the accommodations managed by FLH collected between November 2017 and October 2025. A total of 70,724 reviews were considered, distributed across the four platforms, as shown in Table 1 (number of reviews) and Figure 1 (percentage of reviews).

These variations do not affect the assignment of reviews to their respective accommodation units, since the accommodation unit identifier is used consistently across all platforms. However, the way text data is stored varies across platforms. Text may be stored in "Title", "Positive Comment" and "Negative Comment" fields, "Title" and "Comment" fields, or just in the "Comment" field.

Table 1. Distribution of the reviews by platform.

Platform	# Reviews
Booking	36,105
Airbnb	33,045
HomeAway	1569
Feels Like Home	5

It is also important to note that different behaviors can be seen across users, with some filling all text fields in their review, some only providing a title and other users not providing any text feedback. These latter cases were not discarded since many of them still provided relevant quantitative information, not only in terms of global valoration, but also with more specific evaluation aspects, namely "cleanliness", "accommodation", "location", "service", "comfort", and "value for money".

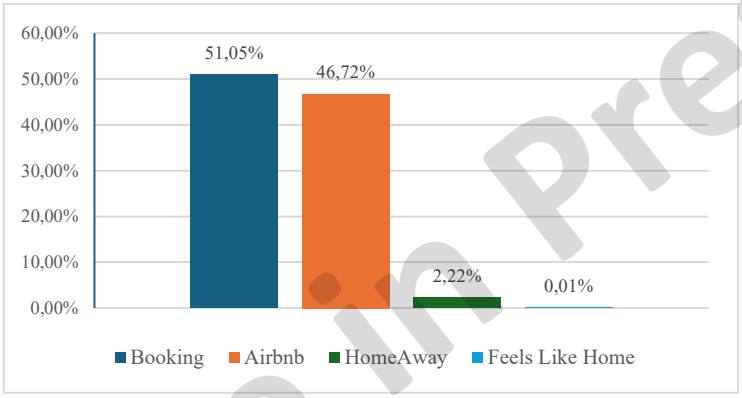

Figure 1. Distribution of the reviews by platform.

The database includes identifiers for 21 different languages, but this field proved to be unreliable. In fact, reviews with a Portuguese identifier were found containing comments in English, as well as reviews with an English identifier containing comments in Spanish, French and German. To simplify data analysis and model comparison, a filter was applied to the database to consider only reviews written in English. Since the language identifier is not reliable, a language detector (Stahl, 2004) was applied to the whole dataset to isolate these reviews. This allowed for the study of 5050 titles, 6372 positive comments, 5848 negative comments and 9628 neutral comments. These four review components were analysed separately, both in terms of preprocessing and exploratory analysis, as well as in terms of modelling. For the exploratory analysis the study was filtered to only include nouns and verbs, since these are more informative and help to identify themes that may be raised in the reviews and which are not captured by quantitative fields. The analysis showed that "apartment" and "location" were consistently part of the top-3 words in the Term Frequency Chart across all four

components, being the two most mentioned words in most of them. However, in order to not bias the analysis, it was decided not to include these words in the stopwords list. The name of the company under study was not one of the 20 most frequently mentioned words in any of the review components, and thus was also not considered a stopword.

4. Methods and Experimental Results

Text mining was applied in two perspectives: topic detection and sentiment analysis. The algorithms applied and the main results are respectively described in Sections 4.1 and 4.2.

4.1. Topic Detection

Topic detection was implemented using the Python library gensim (Rehurek & Sojka, 2011) to apply Latent Dirichlet Allocation (LDA) algorithm. The two extreme models were selected among those available: en_core_web_sm and en_core_web_trf. The former is optimized for computational efficiency, whereas the latter uses transformers to improve its accuracy. Preliminary tests were applied to the reviews of single accommodation units and the two models yielded very similar results in terms of topics, with en_core_web_trf requiring significantly greater computational times. Therefore, en_core_web_sm was selected to perform topic detection for all English reviews, which allowed to identify interpretable topics on all review components, with several of them raising themes that are different from the evaluation aspects, or are covered by the evaluation aspects but address more specific issues.

Figure 2 and Table 2 illustrate the topic detection from the entire text of the "Title" field of all the reviews in the dataset.

```
[(0,
  '0.118*"property" + 0.097*"value" + 0.068*"money" + 0.066*"need" + 0.044*"restaurant" + 0.030*"bit" + 0.023*"improvement" +
0.022*"friend" + 0.019*"improve" + 0.018*"clean"'),
 (1,
  '0.367*"apartment" + 0.068*"heart" + 0.040*"area" + 0.020*"need" + 0.019*"stop" + 0.019*"spot" + 0.019*"lago" + 0.018*"equip" +
0.017*"leg" + 0.017*"clean-"'),
 (2,
  '0.254*"place" + 0.088*"porto" + 0.087*"city" + 0.076*"stay" + 0.038*"time" + 0.026*"staff" + 0.025*"visit" + 0.024*"restaurant" +
0.023*"shop" + 0.021*"room"'),
 (3,
  '0.457*"location" + 0.143*"stay" + 0.133*"apartment" + 0.021*"host" + 0.019*"break" + 0.012*"space" + 0.012*"facility" + 0.011*"close"
+ 0.009*"funchal" + 0.009*"go"'),
 (4,
  '0.094*"home" + 0.086*"feel" + 0.072*"experience" + 0.048*"communication" + 0.044*"use" + 0.038*"appartment" + 0.032*"madeira" +
0.028*"center" + 0.027*"punctual" + 0.026*"disappoint"'),
 (5,
  '0.112*"view" + 0.060*"rental" + 0.056*"holiday" + 0.048*"management" + 0.046*"noise" + 0.032*"feeling" + 0.032*"accommodatie" +
0.029*"district" + 0.027*"encounter" + 0.027*"insalubrity"'),
 (6,
  '0.207*"lisbon" + 0.110*"apartment" + 0.087*"stay" + 0.034*"issue" + 0.030*"centre" + 0.024*"book" + 0.021*"window" + 0.021*"goede"
+ 0.019*"sight" + 0.019*"vacation"'),
 (7,
  '0.175*"recommend" + 0.120*"love" + 0.052*"lot" + 0.047*"stay" + 0.046*"night" + 0.033*"day" + 0.033*"town" + 0.032*"budget" +
0.031*"make" + 0.027*"thing"'),
 (8,
  '0.096*"locate" + 0.088*"stay" + 0.088*"family" + 0.063*"accommodation" + 0.056*"enjoy" + 0.051*"price" + 0.032*"terrace" +
0.026*"bathroom" + 0.024*"beat" + 0.023*"sound"'),
 (9,
  '0.082*"service" + 0.057*"people" + 0.053*"check" + 0.045*"come" + 0.045*"group" + 0.038*"get" + 0.029*"work" + 0.022*"customer" +
0.020*"expect" + 0.019*"street"')]
```

Figure 2. Topic detection using the en_core_web_sm model.

Figure 2 shows the results obtained when 10 is given as input for the number of topics to detect. This output was generated using the *en_core_web_sm* model, which took approximately 24 seconds of computation time. The *en_core_web_trf* model produced an output with only minor differences and the same interpretation, requiring more than 3 minutes of computation time, substantially more than for *en_core_web_sm*.

Each of the 10 topics may be resumed in the 10 insights presented in Table 2.

Table 2. List of possible insights resulting from the topic detection from Figure 2.

Topic number	Possible meaning
1	Value for money
2	Location and centrality
3	Closeness to shops and restaurants
4	Space and facilities
5	Experience with staff and communication
6	View, noise, and hygiene
7	Problem management and booking system
8	Nightlife
9	Noise and rooms
10	Service and check-in

4.2. Sentiment Analysis

Sentiment Analysis was carried out using Python library textblob (Loria, 2018). This library returns a text polarity score on a scale from –1 to +1, with negative values translating negative emotions and positive values translating positive emotions. For illustrative purposes, consider the neutral comment in Table 3, for which the *textblob* library-returned a polarization score of 0.26, translating an overall positive emotion, although with some mixed negative aspects.

Table 3. Example of neutral comment.

Neutral Comment
"Perfect location and nice apartment. Had some trouble with the wifi but decided not to have it fixed as we were constantly out and about. The apartment has all amenities. Apartments owned by private persons do have a more personal touch than apartments owned by large companies. The apartment was lovely and the communication with the company just fine. Just missed the personal touch. Simple things as a bottle of water in the fridge or basic salt and pepper in the kitchen were absent. Otherwise great stay."

Sentiment Analysis was carried out, once again considering all English reviews, and a score was computed for all reviews. For the sake of comparison, all polarity scores were converted to match the scale of the global valoration in the channel where the review was posted, and the percent deviation from the global valoration was computed for each review. The component

where the polarity scores most closely match the global valoration is the Title, which is not surprising in the sense that the title, when used, serves as a short and more subjective summary of the customer experience, being less prone to mixed emotions. Figure 3 summarizes the distribution of the deviation between the polarity scores and the global valorations for the Title.

Note that even for the Title there are very extreme deviations. These correspond to cases where the customer provides both positive and negative aspects in their review, but one or several aspects dominate their experience, translating in an extreme global valoration. These large variations should be handled with care, especially in cases where the polarity is significantly larger than the global valoration, since they may indicate immediate issues in the service.

count	5050.000000
mean	3.560565
std	59.617176
min	-800.000000
25%	-4.166667
50%	12.500000
75%	28.571429
max	100.000000

Figure 3. Summary of the distribution of the deviation between the polarity scores and the global valorations for the Title.

5. Conclusions

The aim of this study was to use text analytics to extract insights from online reviews for a holiday rental company. For this effect, the English reviews among the 70,724 available reviews between November 2017 and October 2025, from platforms Booking, Airbnb, HomeAway, and Feels Like Home were processed and analyzed through topic detection and sentiment analysis, this analysis being applied to titles, positive comments, negative comments and neutral comments. Topic detection seems to be promising as a complement to the quantitative valoration given in general and in specific evaluation aspects, helping to identify additional relevant themes that are mentioned in the reviews. Sentiment analysis resulted in polarity scores that do not always align with the global valorations, especially in situations where positive and negative aspects are mentioned in the review but the global valorations are extreme. Such cases merit reporting to the company for closer observation.

Future work includes extending the study to all reviews after translating them to English. Its insights will be used to develop a machine learning-powered alert system to report operational and strategical issues to the holiday rental company.

References

Bi, J.-W., Zhu, X.-E., & Han, T.-Y. (2024). Text Analysis in Tourism and Hospitality: A Comprehensive Review. *Journal of Travel Research*, 63(8), 1847-1869. https://doi.org/10.1177/00472875241247318

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022. https://dl.acm.org/doi/10.5555/944919.944937

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6), 391-407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9

Hair Jr., J. F., Harrison, D. E., & Ajjan, H. (2021). *Essentials of marketing analytics*. New York: McGraw-Hill.

Le, H. T. M., Phan-Thi, T. A., Nguyen, B. T., & Nguyen, T. Q. (2025). Mining online hotel reviews using big data and machine learning: an empirical study from an emerging country. *Annals of Tourism Research Empirical Insights*, 6(1), 100170. https://doi.org/10.1016/j.annale.2025.100170.

Loria, S. (2018). textblob Documentation. *Release 0.15*, *2*. https://pypi.org/project/textblob/

Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., & Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, *83*, 101760. https://doi.org/10.1016/j.jairtraman.2019.101760

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155 (2), 945–959. https://doi.org/10.1093/genetics/155.2.945.

Rehurek, R., & Sojka, P. (2011). Gensim–python framework for vector space modelling. NLP

Stahl, P. M. (2004). lingua-language-detector 2.2.0. https://pypi.org/project/lingua-language-detector/

Zhao, R., Hao, Y., & Li, X. (2024). Business analysis: User attitude evaluation and prediction based on hotel user reviews and text mining. arXiv preprint arXiv:2412.16744. https://doi.org/10.48550/arXiv.2412.16744

Acknowledgements

This work is supported by the Horizon Europe project INITIATE (Grant agreement ID 101136775).