

Review Paper

Sentiment Polarity for Restaurant Business: An Integrated Approach

Manik Rakhra¹, Neha Verma², Shubham Kalihari³

^{1,3}Department of Computer Science and Engineering

Lovely Professional University Phagwara, Punjab-14411, India

²KRM DAV College Nakodar

Manik Rakhra, rakhramanik786@gmail.com.

Handling Editor: Mohammad Shabaz

Copyright © 2021 Manik Rakhra, Neha Verma and Shubham Kalihari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Website and social networking outlets have become more and more popular places for people to voice their opinions on different issues, in particular their frustrations with brands and corporations. Similarly, emotion analysis is starting to be incorporated into corporations owing to the existence of an abundance of opinionated content from the digital channels provided by customers. The written narratives may be written in an optimistic, a negative, or a more analytical style. Travellers' ratings are an important influence on whether the hotel is a beginning or a long-term client. It helps the hotels aware of their worth in the industry, and improves their position by helping customers better define their qualities in a competitive setting. This research's key goal is to identify ways to improve hotel opinion analyses. Lexical-based combination and deep learning will yield the most appropriate results. Firstly, it uses predefined terms to look for polarity in the lexicon dictionary, and then employs the machine learning techniques as the final layer. The article discusses what is yet to be done and what has been done in the analysis so far.

Keywords : Judgment Mining, Sentiment Analysis, Unsupervised Learning

1. Introduction

Nowadays, the internet is taking on the role of a major hub that provides numerous venues for people to share their thoughts and feelings on various topics that intrigue them. Owing to the availability of vast amounts of data from customers, businesses are moving into social networking for decision making. Hotel proprietors are expanding their usage of social media to increase their perception of guests' desires, fine-tune their service, and to stay successful in the industry. Similarly, people are opting for the internet as a source of accommodations since it is easier than having the physical kind. With the recent availability of the enormous amount of user-generated data from the Internet, sentiment analysis has become more common. If you are speaking of sentiment, you may call it a feeling, or a professional study of feedback can assign it a "yes" or "no" to whether it is good or bad. While sentiment analysis has taken a great leap forward as a few years ago in the field of social media analytics, emotion processing has now advanced as well, extracting actionable information from sentiment analysis results, such as finding consumers' emotions about hotels, is still in the forefront. The visitors' input helps the owners make strategic choices based on their feelings. Unsupervised approaches depend on finding the polarity and subjectivity of terms. The controlled methodology is also known as machine learning-based and uses sets of positive, negative, and neutral classifier attributes, as well as labeled training data to generate the classifier. This customer' online reviews usually reflect their feelings such as they are satisfied, neutral, or conflicted about their stay. Lexicon-based and machine learning are proposed as a combination of approaches. When this is finished, the remainder of the document is positioned

like this. brief study linked to online hotel feedback, various levels of sentiment appraisal, and a composite review method for sentiment assessment. You will find out more about the situation in Section 3, the solution in Section 4, and what will be accomplished in the analysis in Section 3 and how they will be done in Section 4. Section 5 is focused on discovering the meanings consumers ascribe to hotel ratings. Section 6 reviews the research's relevance to the industry's employees and customers. In the technique portion, the specifics of data collection and interpretation are presented, and then the approach is shown. In the light of these findings, as well as the ideas that will be developed in the future, the consequences and constraints of this analysis have been identified and the focus is laid on future endeavors.

2. Problem formulation

Semantics or sentiment analysis or opinion mining is a form of data analysis that analyses what people feel or think or their ideas about various issues and products. Gupta et al. has said in their study [2] that the needs and wishes of the visitors must be met to succeed in running a hotel. Since guest reports are so plentiful, it is impossible for the hotel to manually do this analysis, which means it takes a whole team of people to evaluate them, which means they would spend a lot of time.

Tathgir and Jaswal claimed that several methods to sentiment analysis must be applied to extract feelings as well as current issues of sentiment classification. To study the diverse ways in which customers communicate on the Internet, there were also found to be a large number of unstructured sentiment analyses problems, such as improper grammar, informal

language, connections, hashtags, acronyms, abbreviation, domain acronyms, texting slang, satire, words with different meanings, disambiguation, and hoaxes [all these occurred because of the unstructured nature of customers' online communication]. Khan et al. claimed that algorithms function well in supervised learning, but lack of labeled training data lead to sparsity; but the results in unsupervised learning are different.

3. Research Background

For hotels, understanding what visitors want and need is a significant achievement. computational tools like sentiment analysis are gaining in popularity to handle massive volumes of unstructured data for the general public on social media if on the other hand, "judgement mining" is the study of people's judgement, "and perception" of individuals The new and most recent literature on sentiment analysis and opinion mining in online reviews is provided here for your convenience. With the adoption of social networking in the hospitality, everyone in the sector would utilize social media as a tool to provide their visitors with up-to-date and accurate expectations. This is approved by practice and research Customer ratings play a huge role in assessing the quality of the hotel-perceptions, because judgments affect future visitors A lot of writing content is now found on social media about how people perceive products, so sentiment analysis is becoming more and more common for understanding their consumer perceptions If emotions are of a stronger or weaker, it is not necessary to indicate their polarity in the writing. A hotel staff member uses mood analysis is used to identify attitudes of guests. using the sentiment analysis tool to gather written views on four-centered five-star hotels in

China and TripAdvisor to come up with the opinion In the other side, the study by Barbosa et al. [12] found that TripAdvisor review scores were also closely aligned with other means of evaluating hotels in seven countries using various ranking methods. Online feedback may be really valuable for both the guests and hotel management. There are four main levels of sentiment on which writing on websites is evaluated, namely: optimism, certainty, neutrality, certainty negative, negative, negativity, and skepticism. The qualitative analysis of sentiment may be studied in three ways: at the document or entire text stage, one or more words, or one or more individuals [13] The key aspects of sentiment analysis can vary depending on which tools you use. The documents are defined as being at the phrase, clause, and keyword stage, and are further refined as tasks. This is the most direct designation of all, when you choose to designate whether a review is favorable, bad, or none at all. as an example, when customers leave a comment, the mechanism determines if the product's general rating is positive or negative Therefore, it is believed that the text concerns only one subject and not papers that include or discuss several items. In the analysis, two machine learning methods generated findings that significantly outperformed previously published classifications of movie reviews. Sentiment analysis assumes sentences for individual products or services are sought, and judged to be impartial in the process. each item is thought of as a different, and as opposed to every other On the one hand, I would feel insulted if he said, for example Jait and Ahuja studied how to improve productivity of sentiment analysis by classifying messages according to whether they were good or not. Features must be located at a level one or more tiers above that describes the features of the design

rather than at one or more classes inside the design hierarchy

The features examined were too far down on the text and sentence stage, but not at a sufficient detail for citizens to determine whether they were appropriate for their needs. It has a more accurate analysis; it is said to have aspects as well. It is a piece of creative work to define and remove the product's critical features from the source info. Recent research [that] has shown that new sentence-based features are better at distinguishing sentiments. It is also possible to categorize emotions using controlled or unsupervised means, but the outcome of this depends on the proper elimination of the collection of elements. Due to the vast number of consumer feedback that need to be handled for new technologies, there is a demand for a blended solution to achieve improved results in a shorter time frame. Generally, a mixed strategy has greater overall consistency and success. By way of a lexical-based technique that has been paired with machine learning, classification has been an order of magnitude more accurate relative to machine learning. According to the study, the hybrid strategy outperforms the traditional strategy. Similarly, Lalji and Deshmukhukh [24] reports that the lexical method increases the metrics while preserving the words' accuracy, and has proven effective in increasing efficiency. Combining supervised and unsupervised approaches offers an abundance of options, so it is conceivable that more than one way or methodology might yield various outputs. The study conducted by Ahmad et al. discusses several methods of emotion analysis along with their benefits and weaknesses. A hybrid methodology to conduct sentiment analysis has been shown to be the most effective. Analysis technique is a way to carry out a study. The method was CRISP-DM, which refers to data

selection, processing, polarity detection, and classification that will be described in Section 5. it is critical to select the correct way that is best to interpret online hotel feedback in order to utilize and overcome problems that tend to extract sentiments. In the following sections, the various parts of the system are summarized for data gathering and data analysis.

4. Research Methodology

The method was CRISP-DM, which refers to data selection, processing, polarity detection, and classification that will be described in Section 5. it is critical to choose the correct way that is best to interpret online hotel feedback in order to utilize and overcome problems that tend to extract sentiments. In the following sections, the various parts of the system are summarized for data gathering and data analysis.

4.1 Communication gathering

The basic approach to this line of investigation is contextual. This method is normally employed to discern meanings, provide accounts of ideas, to get an understanding of definitions, or get an idea of your personal experiences. Overall, this method suits the study allows hotel guests to write web reviews to be in order to get valid results. The actual data is employed for this qualitative study. However, in general, consumers are encouraged to leave a positive or negative review of their hotel stay to get a deeper understanding of the treatment that they had, as well as more knowledge of others so that their thoughts will help the hotel to have more facilities. We estimate that a consumer survey requires less than ten minutes of time. When it comes to face-to-face communication, often people choose to

post thoughts on other websites. By using those programs, the ratings may be pulled out of the websites (API). People who are surveyed with open-ended questions have a better ability to articulate their emotions and opinions, but they're impossible to measure.

4.2 A new and enhanced data analysis system

Often, data processing has a substantial effect on the rest of the study. In qualitative analysis, thoughts and perceptions take the shape of evidence rather than objective facts, which are far more popular in quantitative research. Data review is the study of scanning and classifying for information, followed by presenting the results to the researcher and others. Data planning, data reduction, emotion interpretation, and drawing conclusions were given more attention in this research than in order to improve performance. The data can be tested using a random sample selection of data, which reduces the amount of time needed to do the study. The text analysis will determine if the data is sorted, evaluated, and then it will be shown whether the moods are negative, neutral, or optimistic. And after that, the conclusions and analyses of the results will be done. By design, open-ended surveys provide a higher level of knowledge of the target customer's thoughts and emotions. A problem with the results is that certain feedback written by consumers would not be applicable to the same product or service.

4.3 Overview of the approach

We collect the data directly from the hotel database or from its website API. The consumer gets to determine how much data is needed, and can thus receive the feedback via the application programming interface. Forums, or blogs, it's possible to gather data

as well. After data preparation and retrieval; we need to convert the data to the requested format before entering it into the database. Abbreviations, emoticons, emojis, and so on are all potential sources of noise as you use data-gathering tools online. In short, it involves text and non-textual details, plus identifying and ignoring irrelevant material. It is also critical to screen or pre-process the extracted data before carrying out research in order to remove the extraneous factors which make analysis easier. is the ability to tell whether or not anything would turn into a positive or a negative review based on the tone of the post. Any sentence in the study is searched to find its polarity and replaced by the polarity in the lexicon. When working with students, delete all the words that don't have much meaning, but keep the sentences that do.

5. Conclusion

Social networking has become a convenient means for consumers to speak their minds, and companies, particularly hotels, can benefit from this by giving weight to their customers' input. Hotels must be equipped with methods to handle a lot of data and customer care solutions that make them more effective. This paper offered a statistical method for classifying hotel feedback to further concentrate on the hotel's consumer value and customer loyalty in order to detect each guest's propensity to make good or bad online reviews. When using a lexical method, the accuracy of the sentiment analysis would be strong, but the specificity of the sentiment is low, and as a result, the output would be inadequate. Supervised and non- and unsupervised learning is combined. The study proposes a strategy that would allow the hotels to track, capture, and evaluate guests' online feedback online. a problem of the study is that it uses only English documents, but does not use those

published in other languages Further analysis is needed that incorporates reviews published in various languages is needed. Managing a hotel's online profile would be one of the study fields that needs to be addressed in the future. Systems are critical for hotels to develop because they must keep an eye on the public's reactions on the internet. Furthermore, language-specific comparisons may be predicted to yield markedly different effects

References

- [1] N. Divyashree, K. L. Santhosh, and J. Majumdar, "Opinion Mining and Sentimental Analysis of TripAdvisor.in for Hotel Reviews", *International Research Journal of Engineering and Technology*, vol. 4, no. 11, pp. 1462- 1467, 2017.
- [2] S. Gupta, S. Jain, S. Gupta, Shruti, A. Chauhan, "Opinion Mining For Hotel Rating Through Reviews Using Decision Tree Classification Method", *International Journal of Advanced Research in Computer Science*, vol. 9, no. 2, pp. 180-184, 2018.
- [3] T. Najam, K. B. Nowshath, and Vazeerudeen, "Sentiment Analysis of tweets using Supervised Machine Learning", *Journal of Applied Technology and Innovation*, vol. 3, no. 1, pp. 30-32, 2019.
- [4] X. Tian, R. Tao, W. He, and V. Akula, "Mining Online Hotel Reviews: A Case Study from Hotels in China", in *Americas Conf. on Information Systems*, San Diego, CA, 2016, pp. 1-8.
- [5] P. Choudhari, and S. V. Dhari, "Sentiment Analysis and Machine Learning Based Sentiment Classification: A Review", *International Journal of Advanced Research in Computer Science*, vol. 8, no. 3, pp. 1051-1056, 2017.
- [6] F. H. Khan, U. Qamar and S. Bashir, "A Semi-Supervised Approach to Sentiment Analysis using Revised Sentiment Strength based on SentiWordNet", *Knowledge and Information Systems*, vol. 51, no. 3, pp. 851-872, 2017.
- [7] P. Phillips, S. Barnes, K. Zigan, and R. Schegg, "Understanding the Impact of Online Reviews on Hotel Performance: An Empirical Analysis", *Journal of Travel Research*, vol. 56, no. 2, pp. 235-249, 2017.
- [8] W. He, X. Tian, R. Tao, W. Zhang, G. Yan, and V. Akula, "Application of social media analytics: A case of analyzing online hotel reviews", *Online Information Review*, vol. 41, no. 7, pp. 921-935, 2017.
- [9] S. Yordanova, and Kabakchieva, "Sentiment Classification of Hotel Reviews in Social Media with Decision Tree Learning", *International Journal of Computer Applications*, vol. 158, no. 5, pp. 1-7, 2017.
- [10] X. Zhang, Y. Yu, H. Li, and Z. Lin, "Sentimental Interplay between Structured and Unstructured User-Generated Contents-An Empirical Study on Online Hotel Reviews", *Online Information Review*, vol. 40, no. 1, 2016.
- [11] O. Hoeber, L. Hoeber, M. El Meseery, K. Odoh, and R. Gopi, "Visual Twitter Analytics (Vista) Temporally changing sentiment and the discovery of emergent themes within sport event tweets", *Online Information Review*, vol. 40, no. 1, pp. 25-41, 2016.
- [12] R. R. L. Barbosa, S. Sanchez-Alonso, and M. A. Sicilia- Urban, "Evaluating hotels rating prediction based on sentiment analysis services", *Aslib Journal of Information Management*, vol. 67, no. 4, pp. 392-407, 2015.
- [13] V. K. Bongirwar, "A Survey on Sentence Level Sentiment Analysis", *International Journal of Computer Science*

Trends and Technology, vol. 3, no. 3, pp. 110-113, 2015.

[14] S. Behdenna, F. Barigou, G. Belalem, "Document Level Sentiment Analysis: A Survey", EAI Endorsed Transactions on Context-aware Systems and Applications, vol. 4, no. 13, pp. 1-8, 2018.

[15] S. Kolkur, G. Dantal, and R. Mahe, "Study of Different Levels for Sentiment Analysis", International Journal of Current Engineering and Technology, vol. 5, no. 2, pp. 768-770, 2015.

[16] P. Patil, "Sentiment Analysis Levels and Techniques: A Survey", International Journal of Innovations in Engineering and Technology, vol. 6, no. 4, pp. 523-528, 2016.

[17] A. Tripathy, A. Anand, and S. K. Rath, "Document-level sentiment classification using hybrid machine learning approach", Knowledge and Information Systems, vol. 53, no. 3, pp. 805-831, 2017.

[18] A. Jaitly, S. Ahuja, "Improving The Accuracy For Sentence Level Sentiment Analysis", International Journal of Advanced Research in Computer Science, vol. 9, no. 4, pp. 37-41, 2018.

[19] S. L. Gupta, and A. S. Baghel, "Efficient Feature Extraction in Sentiment Classification for Contrastive Sentences", International Journal of Modern Education and Computer Science, vol. 10, no. 5, pp. 54-62, 2018.

[20] B. Keith, E. Fuentes, and C. Meneses, "A Hybrid Approach for Sentiment Analysis Applied to Paper Reviews", in Proc. of ACM SIGKDD Conf., Halifax, Nova Scotia, Canada, Aug. 2017, pp. 1-10.

[21] M. Ahmad, S. Aftab, I. Ali, and N. Hameed, "Hybrid Tools and Techniques for Sentiment Analysis: A Review", International Journal of Multidisciplinary Sciences and Engineering, vol. 8, no. 1, pp. 28-33, 2017.

[22] C. P. Ashish, and K. M. Patel, "Sentiment Analysis Using Hybrid Approach: A Survey", International Journal of Engineering Research and Applications, vol. 5, no. 1, pp. 73-77, 2015.

[23] P. K. Jaswal, and J. S. Tathgir, "Sentiment Analysis of Social Media Data using Hybrid Approach", Journal of Economic Development, Management, IT, Finance and Marketing, vol. 10, no. 2, pp. 1-6, 2018.

[24] T. K. Lalji, and S. N. Deshmukh, "Twitter Sentiment Analysis Using Hybrid Approach", International Research Journal of Engineering and Technology, vol. 3, no. 6, pp. 2887-2890, 2016.