

Customers' Satisfaction and Preferences Using Sentiment Analysis on Traveloka: The Case of Yogyakarta Special Region Hotels

Novi Diah Wulandari^{1*}, M Helmi Zain Nuri², Lilis Kurniasari³

¹Management Department, Universitas Nahdlatul Ulama, Yogyakarta, Indonesia

²Computer Engineering, Universitas Nahdlatul Ulama, Yogyakarta, Indonesia

³Electrical Engineering, Universitas Nahdlatul Ulama, Yogyakarta, Indonesia

*Corresponding Author Email: d.wulandari@unu-jogja.ac.id

Abstract. The rapid development of internet use today shows a shift in technology that is increasingly leading to online-based media. Consumers tend to browse the completeness of product and service information via the internet by making purchases online because of limited time and perceived convenience. We analysed some of the features that hotel customers need using customer reviews. This study uses hotel customer reviews on the Traveloka website. The purpose of this study was to determine the important factors that influence tourists in choosing hotels based on hotel customer reviews. We use HBDScan algorithm and machine learning approaches to identify hotel customer segments. We found that the segmentation of tourist preferences and satisfaction is an important stage in behavioural analysis to improve the quality of hotel products and services. This analysis can improve the understanding of hotel managers about different market segments according to the level of customer preference and satisfaction. The findings of this study will help managers to set priority instructions for enhancing appropriate hotel features and use online customer reviews to improve customer satisfaction and hotel performance.

Keywords: *Customers Satisfaction, Traveloka*

1. INTRODUCTION

The rapid development of technology makes entrepreneurs continue to innovate in improving services in the service sector. Information technology helps solve a problem, increases effectiveness, efficiency, and productivity [1], [2]. In business, information technology is a means of communication and information dissemination [2], [3]. The e-commerce business based on Business-to-Consumer e-commerce (B2C) is very growing, including in the field of tour & travel services. Competition for tour & travel services is very tight, they try to improve service and convenience for consumers in obtaining information.

One of the e-commerce services Business to Consumer that is popular in tour & travel services is Traveloka, which is a leading company in Southeast Asia that provides travel services on one platform with a total of 9.01 million visitors registered on Similar web[4]–[6]. Prospective customers can easily access Traveloka to get detailed information on hotel facilities that will be visited in a location. Traveloka also provides detailed information about the service quality hotel in form of online reviews from previous visitor. Online reviews can assist managers in analysing customer behaviour accurately and comprehensively in accordance with online ranking standards [7], [8].

Nowadays most of the customers convey their personal preferences through online reviews. Thus, online review data is a new prospect for detecting the market by listening to customer voices [7], [9]–[11]. The question in this study is how the hotel rating obtained based on customer reviews can help managers identify customer preferences and segment them according to their level of satisfaction. Online customer ratings provide information about consumer behaviour and decisions and hotel performance in innovative ways [6]. With the online ranking will help hotel managers identify their customers in order to improve hotel performance [7], [12]. Clustering user review is one of the methods for analysis customer preferences from various points of view regarding hotel valuation. Hotel customers can evaluate various hotel features after check-in, such as the facilities provided, and room cleanliness. The technique used in this study, a vector drawn from doc2vec and clustered with HDBSCAN [13]. This research aims to:

- Identify customer preferences and satisfaction based on reviews on Traveloka.
- Develop a hotel decision support system, using a clustering approach.

2. LITERATURE REVIEW

2.1. Online Review

Online reviews have an important role for consumers. Generally, consumers will read online reviews before deciding to buy or order a product. One of the benefits of having online reviews is that it makes it easier for consumers to make purchasing decisions. There are 6 requirements for online reviews to be effective, namely the use of online reviews, review expertise, online review timeliness, online review volume, online review frequency and comprehensive [14].

Online reviews are a two-way communication tool that generates data on target travel locations, travel facilities and hotels [7], [15], [16]. As an information-dense industry, tourism requires online review information about consumer behaviour [17], [18]. Online reviews have an important role in determining hotel satisfaction and preference. On its website, Traveloka provides a complete selection of features such as hotel information, consumer reviews and ratings based on the results of previous consumer online reviews. In his research, Ahani has conducted a study that evaluates and calculates the preferences of tourists on the Canary Islands using the TrpAdvisor online review website [7]. To help hotel managers identify consumers and determine effective marketing strategies to meet consumer satisfaction, this research focuses on evaluating and segmenting the Traveloka user market so that consumers can plan their trips effectively.

2.2. Consumer Satisfaction

Business success is greatly influenced by customer loyalty. Customer loyalty has a positive relationship with customer satisfaction and hotel services [19]. [20] Bowen in their survey found a relationship between customer loyalty and customer sentiment. The correlation between loyalty and customer

sentiment depends on two important points [21]. Customers will reach the highest point of satisfaction followed by loyalty, and vice versa [22]. Understanding customer sentiment really helps a manager to grow customer loyalty.

Customer satisfaction is very influential on loyalty, and the desire of customers to promote it [23]. [24] have studied the relationship between hotel performance and business value in terms of consumer assessments. In their research, [25] explain the impact of reviews on hotel sales and their profitability by developing a mathematical model. [26] also explored consumer reviews. The ranking of online review sites as well as the number of reviews has a positive relationship with the average size of each online order transaction [27]. In this study, we will identify hotel services that may affect the level of customer satisfaction either positively, neutral or negatively.

2.3. Consumer Preference

According to [17] consumer preferences indicate consumer preference for a product, both goods and services. Preference analysis aims to find out what consumers like and dislike and will determine the order of importance of a product attribute [28]. Consumer reviews are a tool to identify consumer preferences, including in the field of tourism, especially hotels and restaurants [29]. Product reviews on a web provide information about customer preferences. The results show that the tourism product reviews available on the website contain valuable information about preferences. Twitter is a popular medium for voicing opinions on various topics. In his research [30] identified the strength of the specific sentiments of informal texts. Furthermore, [31] analyzed the sentiment on Tweets for hotels, the results show how Twitter data can be analyzed to rank hotels based on reviews.

3. METHODOLOGY

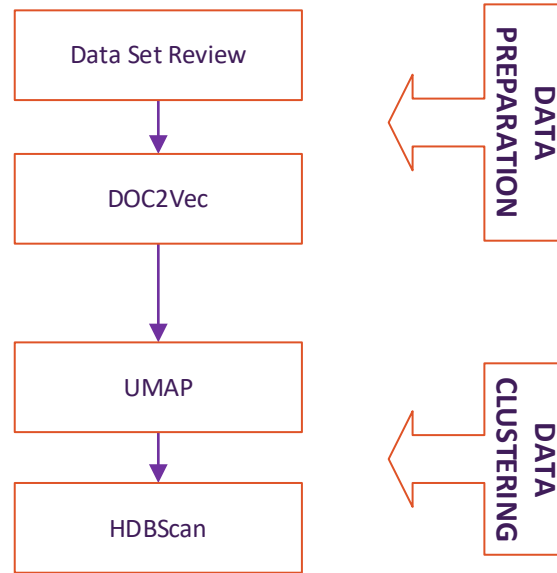


FIGURE 1. Perposed framework

In this study, machine learning techniques were developed. Machine learning is used to help achieve an understanding of the diversity of hotel user satisfaction. We use the data set from hotel user reviews from the Traveloka website. we used HBDSscan as a clustering technique [32], [33]. HBDSscan is a hierarchical density-based grouping algorithm, with the aim of obtaining the benefits of hierarchical grouping, its density and confinement. HBDSscan with its hierarchical approach will make it easier to detect which clusters are the most significant [33].

3.1. Data set review

We use hotel customer reviews on the Traveloka website. we took 1.8 million reviews from 3081 hotels listed in the Special Region of Yogyakarta on the Traveloka website. We use scraping and scrawling techniques to retrieve review data from hotel users. Review data contains positive, negative, and neutral reviews. Data review is used to create vector document. 10.000 reviews were used as data sets and randomly selected.

3.2. Data clustering using HBDSscan

Clustering is a machine learning technique that is often used in segmenting markets [7]. This learning technique tries to find patterns from a dataset by

looking for similar groups of data to reveal meaningful relationships between components [8], [34]. Clustering is one of the effective ways to reduce data dimensions [35].

Based on the design HBDSscan algorithm does not require a lot of parameter adjustment [36]. In this study, we used customer reviews as a data set of 10,000 reviews. we don't yet know how many problem areas the data set is, which means we don't know the number of clusters yet. In this case, an algorithm is needed that is able to detect the problem areas contained in the dataset. HBDSscan is able to detect these problem areas [32], [37].

4. RESULT AND DISCUSSION

4.1. Data Preparation

We used 1.8 million reviews to be trained using doc2vec and filtered out misspelled words and abbreviations. From the results of the review, we chose 10 thousand reviews which were used as a data set with less than 20 words. To make doc2vec we used gensim with a matrix of 136, 281 x 300. Then we reduced the dimensions of the matrix from 300 to 5 dimensions using UMAP to see the proximity of the vector (distance matrix).

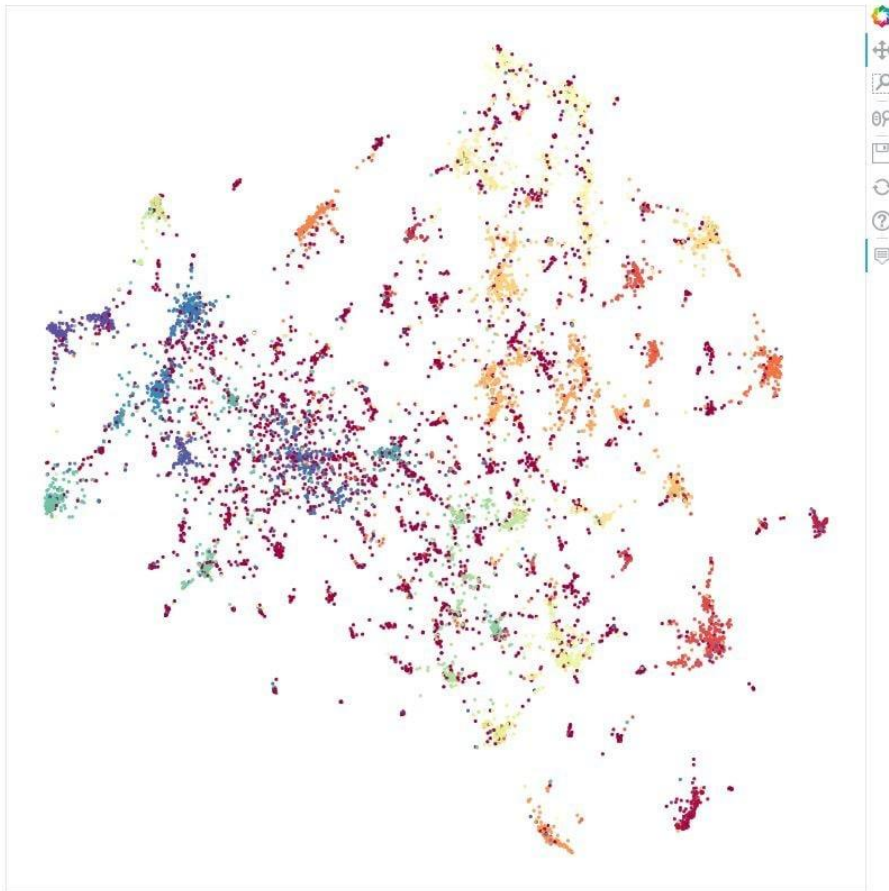


FIGURE 2. UMAP projection of the model

From the figure 2, Doc2vec has several interesting clusters. This shows that the Doc2vec model positions similar document vectors close together. Figure 3 shows an orange dot which is a

collection of reviews specifically talking about services.



FIGURE 3. Doc2vec

4.2. Clustering

Figure 4 shows a bar chart with cluster persistence from the use of different values with the consine distance model visualized in Figure 2. There are 44

clusters formed. The 5 largest clusters that need special attention are taken, namely the 44,38, 42, 15 and 24 clusters. The 5 clusters will be analyzed to determine what policies a hotel manager will pursue.

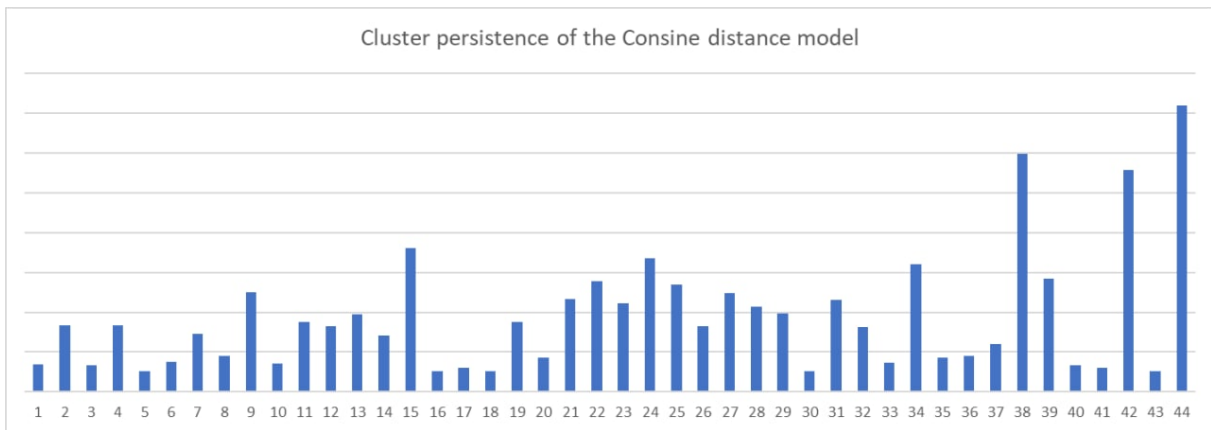


FIGURE 4. Cluster persistence of the consine distance model

Based on the results of the crowling review data, it was found 44 clusters with 62.2% of customers giving positive reviews, 28.5% negative reviews and

9.3% neutral figure 5 From these results it can be concluded that hotel consumers are satisfied with hotels in DI. Yogyakarta.

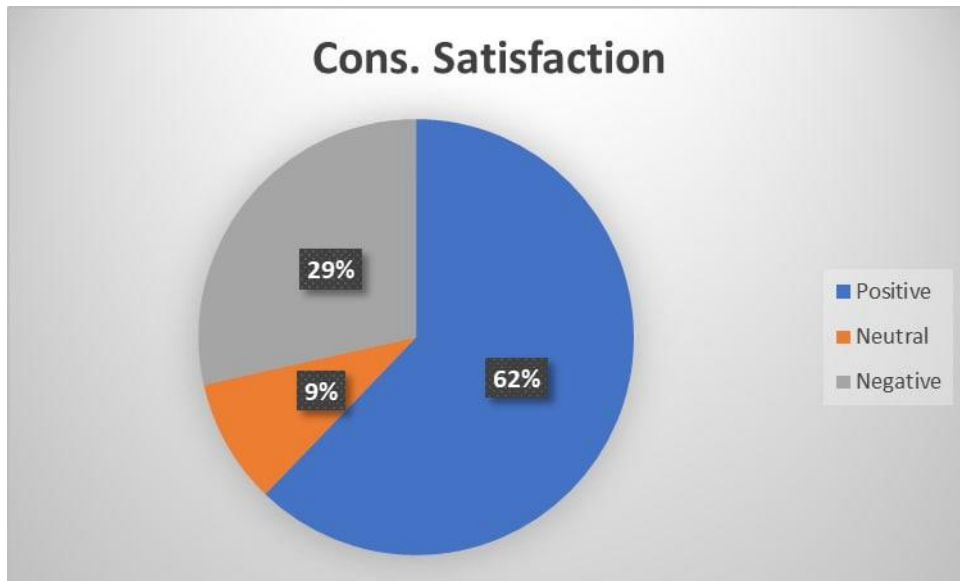


FIGURE 5. Customer satisfaction

Big 5 clusters formed showed in figure 6, namely the cluster, location, room facilities, general, services and meal. Room facilities require special attention, because based on the results of the review above, it can be seen that 94% of the reviews give negative ratings. Services have the highest percentage of positive reviews at 93.9%, this illustrates that consumers are very satisfied with service hotel in DIY. Hotel location is also an important part of consumer assessment, 35.9% of

consumers gave a positive review, 30.4% said it was neutral while the remaining 33.6% gave a negative review, because the percentages are evenly distributed, it is necessary to consider hotels in choosing a strategic location. In the general cluster, it appears that 86.9% of consumers give positive reviews, so it means that hotel services in DIY are very satisfying.

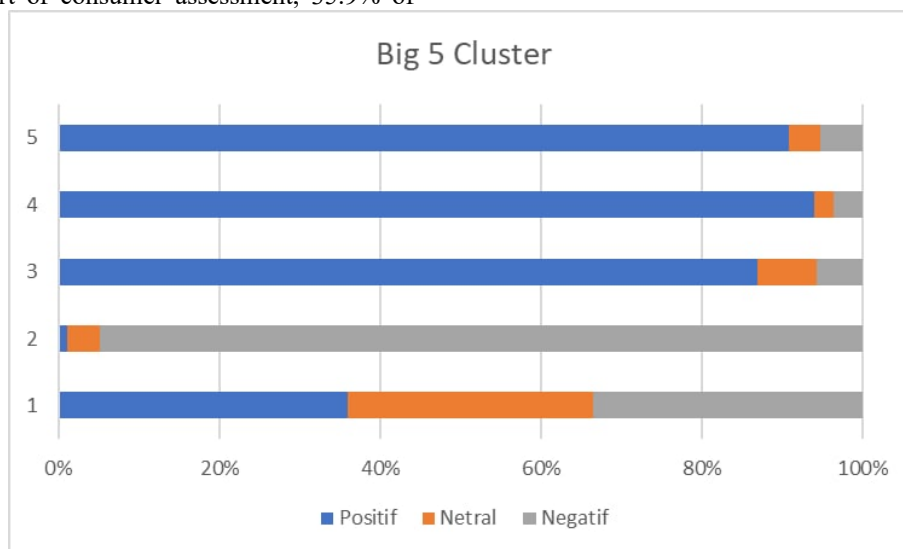


FIGURE 6. Big 5 clusters

From the above results we can see that the consumer preference in looking for hotels is the service. The best service at the hotel begins with the front office, they must be able to look attractive, neat

and polite with a cheerful facial appearance and a sweet smile must be made when guests arrive. One part of the waiter that is no less important is customer service, who must provide the best possible

explanation for guest questions using formal and polite language. Likewise, when guests complain about other problems, of course customer service must be able to provide the best problem solving for all guests. No less important in providing Service Excellence in hotels is the provision of the most complete and best facilities for guests. Rooms must

be neat and clean so that guests feel at home in the room. In addition, the presentation of food and supporting facilities in the hotel must also be considered so that guests are completely satisfied with hotel services figure 7.

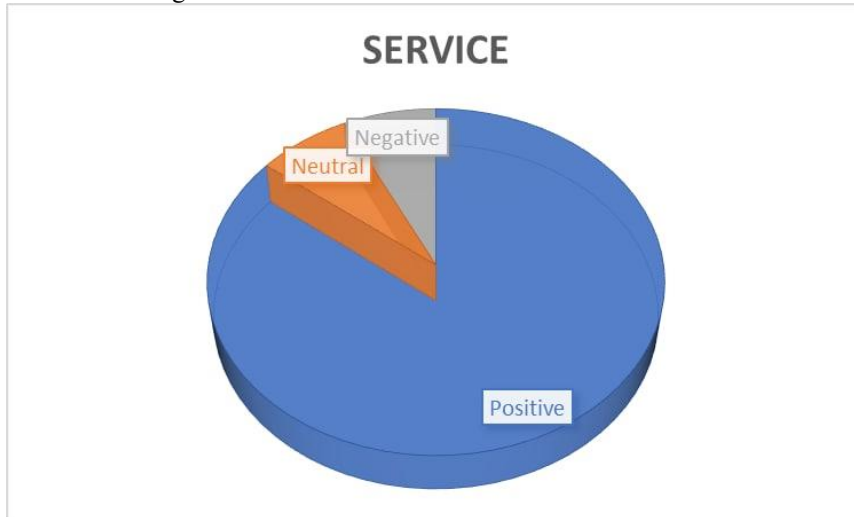


FIGURE 7. Services

Table 1 shows some of the previous studies that had the same theme using several different methods. This means that online user reviews are very

interesting to study. Online user reviews can be analyzed into a consideration for hotel managers in improving the quality of hotel services.

TABLE 1. Previous studies

No	Title	Method	Result
1	Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews	Technical attributes	Engagement of customer reviews positively affects overall customer satisfaction; online reviews are essential for hoteliers to better understand customer behaviour and improve hotel performance.
2	Relationship between customer sentiment and online customer ratings for hotels - An empirical analysis	Naïve Bayes classification technique	There is a positive effect of polarity of customer sentiment on customers for hotels across all categories. However, it does not completely affect customer ratings.
3	Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels	Multi-Criteria Decision-Making (MCDM) and soft computing approach	Travelers have different levels of satisfaction and preferences.
4	Customers' Satisfaction and Preferences Using Sentiment Analysis on Traveloka: The Case of Yogyakarta Special Region Hotels.	HBDScan algorithm and machine learning approaches	The results showed that 62% of customers gave positive reviews, it can be concluded that hotel customers were satisfied with hotels in DI Yogyakarta. In this study, it was found that the consumer preference in looking for hotels is service. The results also showed that 94% gave negative reviews for room facilities.

5. CONCLUSION

For hotel managers, customer satisfaction and preferences are important in determining the marketing strategy to be used. This study aims to see the extent to which consumer satisfaction and preferences obtained based on the results of reviews on Traveloka using HDBScan. The results showed that 62% of customers gave positive reviews, it can be concluded that hotel customers were satisfied with hotels in DI Yogyakarta. Previous research has also been conducted by Kucukusta and Li [38], [39], which stated that consumer satisfaction and preferences are the main keys in the tourism and hospitality industry. In this study, it was found that the consumer preference in looking for hotels is service. The results also showed that 94% gave negative reviews for room facilities. This can provide input to the manager to pay more attention to room facilities.

REFERENCES

- [1] E. K. Januarti and D. Priantinah, "Reviewing the Development of Information, Service, and Website Quality towards Traveloka Customer Satisfaction in Indonesia," *Petra Int. J. Bus. Stud.*, vol. 1, no. 2, pp. 63–69, 2018, doi: 10.9744/ijbs.1.2.63-69.
- [2] M. Syafik and R. Tanamal, "Analysis Of Service Quality Influence On Ecommerce Customer Satisfaction (Case Study Traveloka)," in *Proceedings of The 2017 International Conference on Technology and Applications*, 2017, pp. 2580–7072.
- [3] J. Jauhari, "Upaya Pengembangan Usaha Kecil Dan Menengah (Ukm) Dengan Memanfaatkan E-Commerce," *J. Sist. Inf.*, vol. 2, no. 1, pp. 159–168, 2010, doi: 10.4135/9781412964623.n403.
- [4] "Website Performance." https://pro.similarweb.com/#/website/world-wide-overview/traveloka.com/*999/3m?webSource=Total (accessed Nov. 15, 2019).
- [5] R. Watrionthos, S. Suryadi, D. Irmayani, M. Nasution, and E. F. S. Simanjorang, "Sentiment Analysis Of Traveloka App Using Naïve Bayes Classifier Method," *Int. J. Sci. Technol. Res.*, vol. 8, no. 07, pp. 786–788, 2019.
- [6] L. Kurniasari and A. Setyanto, "Sentiment analysis using recurrent neural network-lstm in bahasa Indonesia," *J. Eng. Sci. Technol.*, vol. 15, no. 5, pp. 3242–3256, 2020.
- [7] A. Ahani *et al.*, "Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels," *J. Retail. Consum. Serv.*, vol. 51, no. August 2018, pp. 331–343, 2019, doi: 10.1016/j.jretconser.2019.06.014.
- [8] M. Nilashi, O. Ibrahim, E. Yadegaridehkordi, S. Samad, E. Akbari, and A. Alizadeh, "Travelers decision making using online review in social network sites: A case on TripAdvisor," *J. Comput. Sci.*, vol. 28, pp. 168–179, 2018, doi: <https://doi.org/10.1016/j.jocs.2018.09.006>.
- [9] R. J. Brodie, A. Ilic, B. Juric, and L. Hollebeck, "Consumer engagement in a virtual brand community: An exploratory analysis," *J. Bus. Res.*, vol. 66, no. 1, pp. 105–114, 2013, doi: <https://doi.org/10.1016/j.jbusres.2011.07.029>.
- [10] M. Sawhney, G. Verona, and E. Prandelli, "Collaborating to create: The Internet as a platform for customer engagement in product innovation," *J. Interact. Mark.*, vol. 19, no. 4, pp. 4–17, 2005, doi: <https://doi.org/10.1002/dir.20046>.
- [11] G. L. Urban and J. R. Hauser, "'Listening In' to Find and Explore New Combinations of Customer Needs," *J. Mark.*, vol. 68, no. 2, pp. 72–87, 2004, doi: 10.1509/jmkg.68.2.72.27793.
- [12] K. Berezina, A. Bilgihan, C. Cobanoglu, and F. Okumus, "Understanding Satisfied and Dissatisfied Hotel Customers: Text Mining of Online Hotel Reviews," *J. Hosp. Mark. Manag.*, vol. 25, no. 1, pp. 1–24, Jan. 2016, doi: 10.1080/19368623.2015.983631.
- [13] S. Tahvili, L. Hatvani, M. Felderer, W. Afzal, M. Saadatmand, and M. Bohlin, "Cluster-Based Test Scheduling Strategies Using Semantic Relationships between Test Specifications," in *2018 IEEE/ACM 5th International Workshop on Requirements Engineering and Testing (RET)*, 2018, pp. 1–4.
- [14] Y. Zhao, X. Xu, and M. Wang, "Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews," *Int. J. Hosp. Manag.*, vol. 76, pp. 111–121, 2019, doi: <https://doi.org/10.1016/j.ijhm.2018.03.017>.
- [15] Y. Narangajavana, L. J. Callarisa Fiol, M. Á. Moliner Tena, R. M. Rodríguez Artola, and J. Sánchez García, "The influence of social

- media in creating expectations. An empirical study for a tourist destination,” *Ann. Tour. Res.*, vol. 65, pp. 60–70, 2017, doi: <https://doi.org/10.1016/j.annals.2017.05.002>.
- [16] M. Sigala, “Social CRM Capabilities and Readiness: Findings from Greek Tourism Firms,” 2016, pp. 309–322.
- [17] P. Kotler and I. Kevin Keller, *Marketing management*, 14th ed. New Jersey 07458: Prentice Hall, 2012.
- [18] O. Appel, F. Chiclana, J. Carter, and H. Fujita, “A Consensus Approach to the Sentiment Analysis Problem Driven by Support-Based IOWA Majority,” *Int. J. Intell. Syst.*, vol. 32, no. 9, pp. 947–965, Sep. 2017, doi: 10.1002/int.21878.
- [19] K. Jay and S. Dwi, “Customer loyalty in the hotel industry: the role of customer satisfaction and image,” *Int. J. Contemp. Hosp. Manag.*, vol. 12, no. 6, pp. 346–351, Jan. 2000, doi: 10.1108/09596110010342559.
- [20] J. T. Bowen and S. L. Chen, “The relationship between customer loyalty and customer satisfaction,” *Int. J. Contemp. Hosp. Manag.*, vol. 13, no. 5, pp. 213–217, 2001, doi: 10.1108/09596110110395893.
- [21] A. J. Coyne and A. H. Orr, “Patterns of Speciation in Jurassic Gryphaea,” *Evolution (N. Y.)*, vol. 43, no. 2, pp. 362–381, 1989, doi: 10.1017/S0094837300007107.
- [22] T. A. Oliva, R. L. Oliver, and I. C. MacMillan, “A Catastrophe Model for Developing Service Satisfaction Strategies,” *J. Mark.*, vol. 56, no. 3, pp. 83–95, Jul. 1992, doi: 10.1177/002224299205600306.
- [23] J. M. Getty and K. N. Thompson, “The Relationship Between Quality, Satisfaction, and Recommending Behavior in Lodging Decisions,” *J. Hosp. Leis. Mark.*, vol. 2, no. 3, pp. 3–22, Apr. 1995, doi: 10.1300/J150v02n03_02.
- [24] K. L. Xie, Z. Zhang, and Z. Zhang, “The business value of online consumer reviews and management response to hotel performance,” *Int. J. Hosp. Manag.*, vol. 43, pp. 1–12, 2014, doi: <https://doi.org/10.1016/j.ijhm.2014.07.007>.
- [25] Q. Ye, R. Law, and B. Gu, “The Impact of Online User Reviews on Hotel Room Sales,” *Int. J. Hosp. Manag. - INT J HOSP Manag.*, vol. 28, pp. 180–182, Mar. 2009, doi: 10.1016/j.ijhm.2008.06.011.
- [26] J. Zhang, Q. Ye, R. Law, and Y. Li, “The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews,” *Int. J. Hosp. Manag.*, vol. 29, pp. 694–700, Dec. 2010, doi: 10.1016/j.ijhm.2010.02.002.
- [27] E. N. Torres, D. Singh, and A. Robertson-Ring, “Consumer reviews and the creation of booking transaction value: Lessons from the hotel industry,” *Int. J. Hosp. Manag.*, vol. 50, pp. 77–83, 2015, doi: 10.1016/j.ijhm.2015.07.012.
- [28] A. C. Sentoso and A. S. Poniman, “Analisa Faktor Yang Menjadi Preferensi Konsumen Dalam Memilih Coffee Shop Di Surabaya,” *J. Hosp. dan Manaj. Jasa*, vol. 3, no. 1, pp. 1–11, Jan. 2015, Accessed: Nov. 24, 2020. [Online]. Available: <http://publication.petra.ac.id/index.php/manajemen-perhotelan/article/view/2824>.
- [29] E. Marrese-Taylor, F. Bravo-Marquez, J. Velasquez, and Y. Matsuo, *Identifying Customer Preferences about Tourism Products using an Aspect-Based Opinion Mining Approach*, vol. 22. 2013.
- [30] S. Kiritchenko, X. Zhu, and S. Mohammad, “Sentiment Analysis of Short Informal Text,” *J. Artif. Intell. Res.*, vol. 50, Aug. 2014, doi: 10.1613/jair.4272.
- [31] K. Philander and Y. Zhong, “Twitter sentiment analysis: Capturing sentiment from integrated resort tweets,” *Int. J. Hosp. Manag.*, vol. 55, pp. 16–24, May 2016, doi: 10.1016/j.ijhm.2016.02.001.
- [32] R. J. G. B. Campello, D. Moulavi, A. Zimek, and J. Sander, “Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection,” *ACM Trans. Knowl. Discov. Data*, vol. 10, no. 1, Jul. 2015, doi: 10.1145/2733381.
- [33] F. Öhrström, “Cluster Analysis with Meaning : Detecting Texts that Convey the Same Message,” Linköping University, 2018.
- [34] D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim, “A literature review and classification of recommender systems research,” *Expert Syst. Appl.*, vol. 39, no. 11, pp. 10059–10072, 2012, doi: <https://doi.org/10.1016/j.eswa.2012.02.038>.
- [35] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, “Recommendation systems: Principles, methods and evaluation,” *Egypt. Informatics J.*, vol. 16, no. 3, pp. 261–273,

- 2015, doi:
<https://doi.org/10.1016/j.eij.2015.06.005>.
- [36] L. McInnes, J. Healy, and S. Astels, “hdbscan: Hierarchical density based clustering,” *J. Open Source Softw.*, vol. 2, Mar. 2017, doi: 10.21105/joss.00205.
- [37] J. Smailović, J. Kranjc, M. Grčar, M. Žnidaršič, and I. Mozetič, “Monitoring the Twitter sentiment during the Bulgarian elections,” *Proc. 2015 IEEE Int. Conf. Data Sci. Adv. Anal. DSAA 2015*, 2015, doi: 10.1109/DSAA.2015.7344886.
- [38] D. Kucukusta, “Chinese travelers’ preferences for hotel amenities,” *Int. J. Contemp. Hosp. Manag.*, vol. 29, Jul. 2017, doi: 10.1108/IJCHM-09-2016-0511.
- [39] G. Li, R. Law, H. Q. Vu, and J. Rong, “Discovering the hotel selection preferences of Hong Kong inbound travelers using the Choquet Integral,” *Tour. Manag.*, vol. 36, pp. 321–330, Jun. 2013, doi: 10.1016/j.tourman.2012.10.017.