ANALYSIS OF HOTEL VISITOR REVIEWS USING LDA-BASED TOPIC MODELING METHOD

Bunga Aisyah Ravissangeta Yansari*[1], Nurvita Trianasari[2]
Universitas Telkom, Indonesia*[12]
bungaaisyah@student.telkomuniversity.ac.id*[1],
nurvitatrianasari@telkomuniversity.ac.id[2]

Abstract: In the current era, many challenges must be faced by tourism businesses in the field of hospitality. The customer experience is an important factor in creating a hotel's reputation because customers can provide feedback online. Agoda is a company that provides an online booking system with online review features including reviews from visitors to Hotel Golden Flower Bandung. This research aims to identify topics that are often discussed related to the experience of staying visitors at The Golden Flower Hotel Bandung and to identify marketing tactics that must be done by the hotel. Secondary data is obtained by doing web scrapping using the Parsehub application. The data sources used are all golden flower hotel Bandung visitor reviews available on Agoda's website. In this research, the method used is LDA-Based Topic modeling which is intended to find out the topic covered in the overall visitor review on Agoda's website is an expression of satisfaction with hotel services and facilities. Golden Flower Hotel Bandung can maintain the quality of services and facilities that get a lot of expressions of visitor satisfaction and improve services and facilities that get a lot of expressions of disappointment from the visitor experience as well as offer new strategies that are more responsible and conduct customer-centered marketing.
Keywords: Business; Marketing Tactics; Online Review; Topic Modeling

INTRODUCTION

The Un's World Tourism Organization points out that the contribution of the tourism sector is one of the large sectors in foreign exchange acquisition. But the coronavirus outbreak that cripples global travel is causing huge losses to the world tourism industry. In terms of size, global travel market dynamics, geographic spread of COVID-19, and its potential economic impact, UNWTO estimates international tourist arrivals could decrease to 3 scenarios by 58%, 70% to 78% by 2020 globally (Unwto.org, 2020). According to Bps.go.id (2020) that the number of foreign tourists coming to Indonesia in early 2020 has decreased. The main cause of the decline in the number of foreign tourist visits is the coronavirus outbreak that occurred in the last week of January 2020. This has an impact on the Level of Room Stay (LRS) of star classification hotels in Indonesia which also decreased in January 2020 (Jabar.bps.go.id, 2020). However, in August 2020 LRS star and non-star hotels in West Java experienced an increase. The highest LRS by star hotel class is recorded in 4-star hotels, while the lowest LRS occurs in 1-star hotels which means the classification of 4-star hotels is quite much in demand by visitors during coronavirus pandemic (Jabar.bps.go.id, 2020).

In the current era, there are many challenges that businesses in the field of hospitality must face. Customer experience is an important factor in creating a hotel's reputation because customers can provide feedback online (Kompasiana.com, 2019). Currently, the hospitality business utilizes the development of technology and tourism as a factor in business development. Many websites or sites provide online booking in the hospitality business. Online booking is part of the activities of a person who makes a booking or reservation for a product or service through online media (Putra, 2017).

ALMANA
Jurnal Manajemen dan Bisnis

With the online booking system, several websites provide online review features in the hospitality business. Online reviews give customers information about products and services based on other customer experiences, reviews can appear on internet platforms such as websites and social media (Zhao et al., 2015). The online review of various websites can evaluate a hospitality business. An online review shows that different individual attributes and factors are used when writing a hotel review. The process of making consumer decisions in writing a review is under the influence of important factors related to the characteristics of the hotel's products or services, information received, stay experience, and other personal preferences (Baruca & Civre, 2015).

Golden Flower Hotel Bandung is a 4-star hotel one of the parts of KAGUM Hotel Management where KAGUM Hotel Management is a hotel management chain company that has a high dedication and specializes in overseeing the manufacture, construction, and travel of various excellent quality hotels by the high standardization embraced by the company (Golden-flower.co.id, 2020). Agoda is one of the largest growing online travel booking platforms in the world. The company, founded in 2005, is growing rapidly in Asia and was taken over by the world's largest online room booking provider, Booking Holdings Inc. in 2007. Agoda provides more than 2 million accommodation properties, including apartments, villas, houses. Agoda's website is supported by over 15 million reviews from travelers including one of the reviews from visitors to Golden Flower Hotel Bandung (Agoda.com, 2020). This research aims to identify topics often discussed by visitors to Golden Flower Hotel Bandung based on their stay experience reviews. Topic modeling is a specific short-term method used in text mining which is one way to identify the topic of a text or document (Alamsyah et al., 2018). This research will offer a solution in conducting topic modeling analysis of the review data provided by visitors of Golden Flower Hotel Bandung through Agoda's website. Analysis using topic modeling aims to find out what topics often appear, making it easier to know consumer perceptions to get customer insights from the visitor's stay experience and to identify marketing tactics that must be done by the hotel.

Big Data is a field dedicated to the analysis, processing, and storage of large data sets that often come from different sources. In particular, big data handles different requirements, such as the merging of several unreceived data sets, the processing of large unstructured amounts of data, and the collection of hidden information in an (Erl et al., 2016)

Text mining is the process of filtering actionable insights from text. Text mining has enterprise value and can contribute to various business units. In particular, text mining can be used to identify actionable social media uploads for customer service organizations. Text mining has marketing implications for measuring the significance of a (Kwartler, 2018)

Topic modeling is a method of extracting information from body text to see the "topic" of what is happening in all documents. Certain topics in the modeling topic are expected to appear more in relevant documents and not in irrelevant documents (Beysolow II, 2018)

Latent Dirichlet Allocation (LDA) is one of the most popular algorithms for topic modeling, each document contains one or more topics of different proportions, and each topic represents a similar word mix. The LDA is an unattended text mining approach. Therefore, time-consuming manual annotations of data can be avoided. LDA is a hidden topic structure by finding a mix of words related to each topic and defining topics that describe each document. In the LDA the most important parameters of the initial model are determined by the number of topics (Zou, 2018).

Customer insight is the process of collecting market information from various sources, ranging from traditional markets to observing customer conversations through

social media about the company and its services. Customer insight is not determined because of the company's ability to understand the needs and desires of consumers alone, but more than that it requires the integration of skills and creativity that touch aspects of the consumer's humanitarian spirit. Customer insight is also used to define consumer preferences (Beehaqi, 2017).

The online review has benefits as one of the simple and easy ways to find information either product or service, this review can be obtained from the hall and also from online consumers who write their opinions on a website. Consumers are not only influenced by the consistency of recommendations, but stated that positive and negative reviews can increase consumer awareness of a company, especially positive reviews can improve the consumer's stance on the company, but this effect is stronger on companies that are less known to consumers. Besides that trust in e-reviews has a positive and significant influence on accommodation options (Utami et al., 2015).

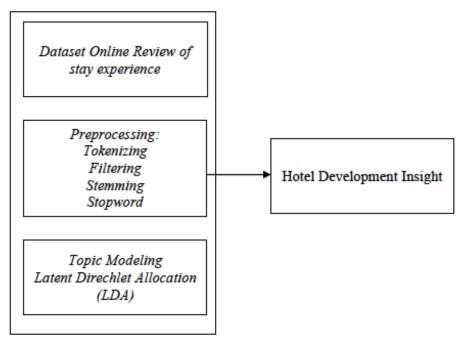Based on the explanation above, the framework is available in Figure 1.

Figure 1. Framework
Source: (Muthia, 2017)

METHODS

In this research, the method used is a qualitative method used in research with natural object conditions and not with experimental object conditions. Indrawati (2015) stated in her book the description and data involved in qualitative research methods cannot be directly measured but qualitative data quantification is done by providing code or categories.

The research used includes a descriptive type of research that aims to determine factors or variables that can measure an object or a specific field (such as a person, organization, product, or brand) but does not yet know the relationship between those factors or variables. Descriptive research is used to observe and describe research subjects or problems without affecting or manipulating variables in any way (Indrawati, 2015).

In conducting the research, researchers researched with the implementation time of Cross-Sectional, a study that can be done where data is collected only once, perhaps for a few days, weeks, or months to answer research questions. The purpose of this study is to collect relevant data when it comes to finding answers to research questions (Sekaran & Bougie, 2016). The data collection stage in this study was conducted at one time, namely on October 17, 2020.

The population data used in this study is all user data on Agoda's website. The samples used in this study used visitor review data from Golden Flower Hotel Bandung with a sample of 2,295 vulnerable reviews from January 1, 2011 – to October 4, 2020. The data was retrieved using the parsehub web scraping application. The data collected amounted to 2,295 comments and preprocessing data has been changed to 2,072 comments. Here are the data analysis techniques in this study:

Data collection is taken from all comments of Golden Flower Hotel Bandung visitors available on Agoda's website. The data is collected using the parsehub scrapping web application.

Figure 2. Data Collection
Source: Processed data (2021)

The process of cleaning and pre-processing data where in this stage the data is obtained in the form of text consisting of data attributes in the form of usernames, user reviews, and user date reviews. Here are the steps performed in the preprocessing data:
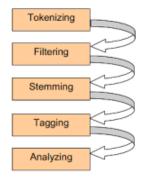

Figure 3. Stages of Data Processing
Source: (Yulian, 2018)

In the picture above Tokenizing is the stage of text segmentation into words and sentences; Filtering is the stage of retrieving important words from token results; Stemming is the word root search stage of each word resulting from filtering; Tagging is the stage of finding the initial form or root of each past tense word or stemming result word.

Analyzing is the stage of determining how far the connection between words between documents is.

Figure 4. Data Preprocessing
Source: Processed data (2021)

Topic Modeling: preprocessing data is processed using the Rstudio application which directly determines the number of modeling topics to be analyzed later. Processed data generates modeling of topics related to the discovery of structures or topics in a set of documents derived from the appearance of words in a document.

RESULTS AND DISCUSSION

The data used in this case study is reviewed on Agoda's website from January 1, 2011 – to October 4, 2020. Golden Flower Hotel Bandung visitor review data is selected based on the vulnerable time specified. There are 2,295 data in the data retrieval period, 223 data cleaning phases are considered irrelevant in the study so that the total data that has gone through the preprocessing stage and will be processed as much as 2,072 net data.

Table 1. Preprocessing Data Results

Before Preprocessing	After Preprocessing
2.295	2.072

Source: Processed data (2021)

With the Topic Modeling method that has been processed using the RStudio application, the dataset will be seen as the whole topic discussed by visitors and users of Agoda's website about Golden Flower Hotel Bandung. The following are the overall results of the LDA-based modeling topic:
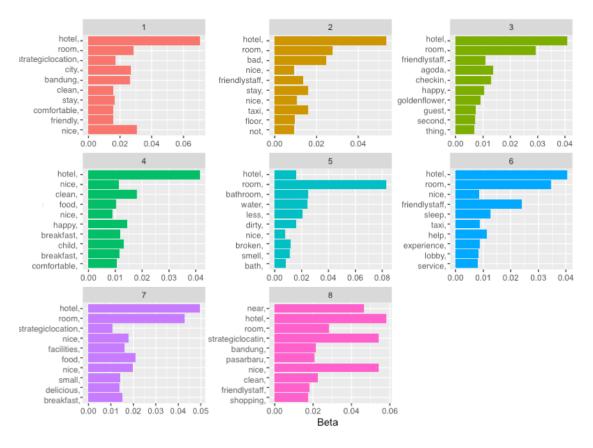
ALMANA
Jurnal Manajemen dan Bisnis

Figure 5. LDA Topic Modeling Results
Source: Processed data (2021)

Based on the Topic Modeling in Figure 5, from January 2011 to October 4, 2020, with the following topics:

Topic 1: Golden Flower Hotel Bandung has a strategic location, city, Bandung, clean, comfortable.

Topic 2: Golden Flower Hotel Bandung has nice rooms, friendly staff, good.

Topic 3: Golden Flower Hotel Bandung check-in, Agoda, happy.

Topic 4: Golden Flower Hotel Bandung is nice, clean, child, happy, comfortable.

Topic 5: Golden Flower Hotel Bandung has a bathroom, less, water, dirty, broken, smelly.

Topic 6: Golden Flower Hotel Bandung has rooms, nice, friendly staff, service, taxis, help.

Topic 7: Golden Flower Hotel Bandung has a strategic location, good, facilities, good, food, good.

Topic 8: Golden Flower Hotel Bandung is near, Pasar Baru, shopping, strategic location, Bandung, good, clean.

The results of data processing of various visitor *reviews* of Hotel Golden Flower Bandung on *Agoda's website* showed the dominant result of expressing satisfaction with hotel facilities and services. Visitors give *reviews* according to the experience felt at Hotel Golden Flower Bandung. Golden Flower Hotel Bandung is strategically located in Bandung, making it easy for visitors to reach places in Bandung. A clean and comfortable hotel, having nice rooms with friendly and kind staff will affect the loyalty as well as the duration of stay of hotel visitors. Through Agoda, visitors find it helpful to check in and make them happy because they can book hotels easily and quickly. In addition to adult

visitors Hotel, Golden Flower Bandung also has children who feel happy and comfortable with the hotel's facilities and services.

However, visitors feel the hotel has shortcomings due to the hotel bathroom which lacks water, dirty, damaged, and smelly. Also, Golden Flower hotel has additional advantages according to visitors, namely good facilities and good food, being near Pasar Baru, and shopping in   Bandung. So that this hotel can be the choice of tourists who want to visit and stay in Bandung.

In general, from the processed products can be stated that the hotel is appreciated by visitors because it is in a strategic location of Bandung City, is a clean and comfortable hotel, friendly staff, good facilities, and good food. Please note, however, that the facilities services offered such as problems in the hotel bathrooms that are lacking in water, dirty, damaged, and smelly still have room for repair, as a management recommendation.

This research shows how hotel visitor reviews can be utilized to support hotel manager decision-making, either by identifying problems treating improvements or by validating marketing strategies made to promote hotels (Calheiros et al., 2017). Both of these can be implemented to identify improvements regarding hotel bathroom problems, marketing strategies can be carried out by marketing the area the hotel is in because the advantages of this hotel through the perception of visitors are the strategic location of this hotel. Identification for fixes regarding hotel bathroom issues, the marketing strategy can be done by marketing the area of the hotel is located, because the advantage of this hotel through the perception of visitors is the strategic location of the hotel.

Based on the above findings, visitors willing to give their review without any request (contrary to the survey) has proven to be a useful means to get reliable feedback. However, reading and processing all hotel visitor reviews in the traditional way will take a long time. Therefore, text analysis techniques are valuable studies to help extract knowledge from a large series of reviews. Also, topic modeling can connect lexicon sentiment with relevant categories in hospitality to uncover topics of interest to help develop new marketing strategies and be able to identify the strengths and weaknesses of ongoing strategies (Rossetti et al., 2015).

CONCLUSION

The results of this study show the topic covered in the overall visitor review on Agoda's website is an expression of satisfaction with hotel services and facilities. It is explained by topics such as hotels are in a strategic location of The city of Bandung, is a clean and comfortable hotel, has good rooms with friendly and good staff, through the Agoda website visitors feel checked in with pleasure, child visitors feel happy and comfortable, have good facilities and good food, are near Pasar Baru, shopping Bandung. However, some facilities must be improved by the hotel management concerning this hotel having less bathrooms, water, dirty, broken, and smelly. In this study, Golden Hotel Flower Bandung needs to place its products to create an impression on consumers. The placement of this product is also an event to promote the product to visitors who have not felt the experience of staying at the hotel. In addition to Hotel website marketing activities can also be done through social media so that Golden Flower Hotel Bandung can maintain relationships with every visitor or prospective visitor by offering a new work system that is more responsible and conducting customer-centered marketing.

REFERENCES

Agoda.com. (2020). *Tentang Agoda*. Agoda.Com.

Alamsyah, A., Rizkika, W., Nugroho, D. D. A., Renaldi, F., & Saadah, S. (2018). Dynamic large scale data on Twitter using sentiment analysis and topic modeling case study: Uber. *2018 6th International Conference on Information and Communication Technology, ICoICT 2018*, *0*(c), 254–258. https://doi.org/10.1109/ICoICT.2018.8528776

Baruca, P. Z., & Civre, Ž. (2015). How do guests choose a hotel? *Academia Turistica - Tourism and Innovation*, *5*(1), 75–84.

Beehaqi, M. (2017). Customer Insight: Paradigma dan Konsep Strategi. *Jurnal Fokus Bisnis*, *16*(2), 45–58.

Beysolow II, T. (2018). Applied Natural Language Processing with Python. *Applied Natural Language Processing with Python*, 77–119. https://doi.org/10.1007/978-1-4842-3733-5

Bps.go.id. (2020). Jumlah kunjungan wisman ke Indonesia Januari 2020 mencapai 1,27 juta kunjungan. Www.Bps.Go.Id.https://www.bps.go.id/pressrelease/2020/03/02/1712/jumlah-kunjungan-wisman-ke indonesia-januari-2020-mencapai-1-27-juta-kunjungan-.html

Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment Classification of Consumer-Generated Online Reviews Using Topic Modeling. *Journal of Hospitality Marketing and Management*, *26*(7), 675–693. https://doi.org/10.1080/19368623.2017.1310075

Erl, T., Khattak, W., & Buhler, P. (2016). *Big Data Fundamentals Concepts, Drivers & Techniques*. Service Tech Press.

Golden-flower.co.id. (2020). *Golden Flower Hotel Bandung*. Golden-Flower. Co.Id. https://golden-flower.co.id/

Indrawati. (2015). *Metode Penelitian Kualitatif Manajemen dan Bisnis Konvergensi Teknologi Informasi dan Komunikasi*. PT Refika Aditama.

Jabar.bps.go.id. (2020). *Tingkat Penghunian Kamar Hotel Agustus 2020 Sebesar 34,95 persen*. Jabar.Bps.Go.Id. https://jabar.bps.go.id/pressrelease/2020/10/01/801/tingkat-penghunian-kamar-hotel-agustus-2020-sebesar-34-95-persen.html

Kompasiana.com. (2019). *Potensi Bisnis Perhotelan di Era Industri 4.0 Makin Luas dan Tanpa Batasan Klasifikasi*. Www.Kompasiana.Com.

Kwartler, T. (2018). *Text Mining in Practice with R*. Wiley.

Muthia, D. A. (2017). Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Mengunakan Algoritma Naive Bayes. *Jurnal ilmu Pengetahuan Dan Teknologi Komputer*, *2*(2), 39–45.

Putra, F. K. K. (2017). Analisis Informasi Situs Web Hotel Bintang 4 di Kota Bandung. *Tourism and Hospitality Essentials*, *7*(1), 7–20.

Rossetti, M., Stella, F., Cao, L., & Zanker, M. (2015). Analyzing User Reviews in Tourism with Topic Models. *Information and Communication Technologies in Tourism 2015*, 47–58. https://doi.org/10.1007/978-3-319-14343-9_4

Sekaran, U., & Bougie, R. (2016). *Research Methods for Business: A Skill-building Approach*. Trento Srl.

Unwto.org. (2020). *IMPACT ASSESSMENT OF THE COVID-19 OUTBREAK ON INTERNATIONAL TOURISM*. Www.Unwto.Org. https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism

Utami, V. K., Tanujaya, D., & Jokom, R. (2015). Analisa Faktor Yang Penting Dari Dimensi Electronic Word of Mouth (Ewom) Bagi Konsumen Dalam Memilih Hotel Di Situs Online Travel Agent (Ota). *Jurnal Manajemen Pemasaran*, *9*(2), 78–85. https://doi.org/10.9744/pemasaran.9.2.78-86

Yulian, E. (2018). Text Mining dengan K-Means Clustering pada Tema LGBT dalam Arsip Tweet Masyarakat Kota Bandung. Jurnal Matematika "MANTIK," 4(1), 53–58. https://doi.org/10.15642/mantik.2018.4.1.53-58

Zhao, X. (Roy), Wang, L., Guo, X., & Law, R. (2015). The influence of online reviews on online hotel booking intentions. *International Journal of Contemporary Hospitality Management*, *27*(6), 1343–1364. https://doi.org/10.1108/IJCHM-12-2013-0542

Zou, C. (2018). Analyzing research trends on drug safety using topic modeling. *Expert Opinion on Drug Safety*, *17*(6), 629–636. https://doi.org/10.1080/14740338.2018.1458838