

INTISARI

Ulasan *online* merupakan salah satu jenis *e-WOM* yang memiliki dampak signifikan dalam sektor *hospitality industry*. Pemberian ulasan, baik positif maupun negatif dapat memengaruhi minat calon *customer* dalam memutuskan menggunakan layanan pada suatu *hospitality industry*. Ulasan positif dapat memengaruhi kenaikan sebesar 4,4%, sedangkan ulasan negatif dapat memengaruhi penurunan sebesar 2,8% untuk jumlah kamar hotel yang terjual. Selain itu, adanya perbedaan perspektif kepuasan tamu hotel yang berasal dari negara dan budaya berbeda menuntut pengelola hotel untuk memahami *guest satisfaction* dalam melakukan ulasan *online*. Namun, banyaknya ulasan *online* yang tersedia di *platform* pemesanan *online* membuat pengelola hotel merasa kesulitan dalam melakukan analisis untuk memahami ulasan tersebut. Faktor lainnya adalah mahalnya biaya yang digunakan untuk melakukan analisis ulasan *online* secara manual menjadi permasalahan dalam memahami ulasan *online*. Oleh karena itu, dibutuhkan kerangka kerja untuk melakukan analisis *hotel guest satisfaction* melalui ulasan *online* dengan efektif dan efisien.

Dalam penelitian ini, digunakan analisis sentimen dengan metode *hierarchical clustering* untuk memahami *guest satisfaction* berdasarkan layanan yang diterima tamu hotel. Selain itu, untuk memahami perspektif *guest satisfaction* yang berbeda-beda dalam melakukan ulasan digunakan data hotel dari 6 benua di seluruh dunia yang didapatkan dari *platform* TripAdvisor.com. Didapatkan 13.067 data ulasan *online* tamu hotel yang berasal dari hotel di benua Afrika, Asia, Amerika Utara, Amerika Selatan, *Oceania*, serta Eropa.

Data ulasan *online* yang telah didapatkan kemudian akan melewati beberapa tahapan utama, diantaranya adalah pengelompokan data ulasan, penentuan polaritas sentimen ulasan, pemilihan dan perhitungan kosakata, klasterisasi data, dan visualisasi data penelitian. Hasil penelitian menunjukkan bahwa aspek layanan *reception* menjadi perhatian utama bagi tamu hotel asal kelompok negara *Africa* 1. Kemudian aspek *design* menjadi perhatian utama tamu hotel asal kelompok negara *America* 1, *Asia* 2, *Oceania*, dan *Europe* 2. Selain itu, terdapat aspek layanan *cleanliness* yang menjadi perhatian utama tamu hotel yang berasal dari kelompok negara *Africa* 2, *America* 2, dan *Asia* 3. Serta terdapat aspek layanan *reservation* yang menjadi perhatian utama tamu hotel yang berasal dari kelompok negara *United States America*, *Asia* 1, *Europe* 1, dan *Channel Island* dalam melakukan ulasan *online*. Pada penelitian ini juga didapatkan performansi prediksi sentimen ulasan dengan menggunakan algoritma TextBlob dengan nilai akurasi sebesar 92,74%, presisi sebesar 96,97%, *recall* sebesar 94,73%, dan *F1-score* sebesar 95,84%.

Kata kunci: Analisis Sentimen, *Hierarchical Clustering*, Ulasan Online Hotel, Natural Language Processing, Identifikasi Aspek Layanan Hotel, *Word Similarity*.

ABSTRACT

Online reviews are a type of e-WOM that has a significant impact on the hospitality industry. Giving reviews, both positive and negative can affect the interest of potential customers in deciding to use services in the hospitality industry. Positive reviews can account for a 4,4% increase, while negative reviews can account for a 2,8% decrease in the number of hotel rooms sold. In addition, the different perspectives on hotel guest satisfaction who come from different countries and cultures require hotel managers to understand guest satisfaction in conducting online reviews. However, the large number of online reviews available on online booking platforms makes it difficult for hotel managers to do a guest satisfaction analysis. Another factor is the high cost used to perform online review analysis manually, which is a problem in understanding online reviews. Therefore, a framework is needed to do hotel guest satisfaction analysis through online reviews effectively and efficiently.

In this study, sentiment analysis was used with the hierarchical clustering method to understand guest satisfaction based on the services received by hotel guests. In addition, to understand different guest satisfaction perspectives in reviewing the hotel, 6 continents hotel data around the world is used from the TripAdvisor.com platform. 13.067 hotel guest online review data were obtained from hotels in the continents of Africa, Asia, North America, South America, Oceania, and Europe.

The online review data that has been obtained will then go through several main stages, including grouping the data, determining the polarity of the review sentiment, vocabulary selection and calculation, data clustering, and visualizing. The results showed that the reception service aspect was the main concern for hotel guests from the Africa 1 group of countries. Then the design aspect became the main concern for hotel guests from the America 1, Asia 2, Oceania, and Europe 2 groups of countries. In addition, there is a cleanliness service aspect which is the main concern of hotel guests who came from the Africa 2, America 2, and Asia 3 groups of country. As well as there is a reservation service aspect that is the main concern of hotel guests who come from the United States America, Asia 1, Europe 1, and Channel Island groups of country to do online reviews. This study also obtained the performance of the review sentiment prediction using the TextBlob algorithm with an accuracy value of 92,74%, a precision of 96,97%, a recall of 94,73%, and an F1-score of 95,84%.

Keywords: Sentiment Analysis, Hierarchical Clustering, Online Hotel Reviews, Natural Language Processing, Identification of Hotel Service Aspects, Word Similarity.