

ABSTRAK

Ulasan Google Maps menjadi sumber informasi penting dalam mengevaluasi kualitas layanan destinasi wisata karena memuat pengalaman pengunjung dalam bentuk teks dan rating numerik. Namun, jumlah ulasan yang besar, tidak terstruktur, dan beragam membuat identifikasi isu layanan sulit dilakukan secara manual. Penelitian ini bertujuan mengidentifikasi topik dominan pada ulasan Google Maps Kebun Binatang Surabaya berdasarkan segmen rating pengunjung, menentukan coherence score pada model LDA terpilih, serta menganalisis peran segmentasi rating terhadap ketajaman hasil pemodelan topik. Data dikumpulkan melalui web scraping dalam rentang satu tahun terakhir. Setelah filtering, diperoleh 1.180 ulasan yang dikelompokkan menjadi segmen rendah 119 ulasan, sedang 78 ulasan, dan tinggi 983 ulasan. Data diproses melalui preprocessing teks, pembentukan bigram dan trigram, representasi Bag of Words, pemodelan Latent Dirichlet Allocation (LDA), evaluasi coherence score, serta visualisasi wordcloud dan pyLDAvis. Hasil menunjukkan model terpilih pada segmen tinggi menghasilkan 3 topik dengan coherence score 0,5042, segmen sedang menghasilkan 2 topik dengan coherence score 0,4114 berdasarkan interpretabilitas, dan segmen rendah menghasilkan 4 topik dengan coherence score 0,4743. Skenario tanpa segmentasi menghasilkan coherence score 0,5314 dengan 2 topik, tetapi topiknya masih bersifat umum. Segmentasi rating membantu memperjelas tema kepuasan, evaluasi campuran, dan keluhan spesifik pengunjung.

Kata Kunci: *Google Maps*, *Latent Dirichlet Allocation*, pemodelan topik, segmentasi rating, ulasan pengunjung.

ABSTRACT

Google Maps reviews serve as a key source of information for evaluating the quality of service at tourist destinations, as they capture visitor experiences in the form of text and numerical ratings. However, the large volume of unstructured and diverse reviews makes it difficult to identify service issues manually. This study aims to identify dominant topics in Google Maps reviews of the Surabaya Zoo based on visitor rating segments, determine the coherence score of the selected LDA model, and analyze the role of rating segmentation in the accuracy of topic modeling results. Data were collected via web scraping over the past year. After filtering, 1,180 reviews were obtained and grouped into low-rating (119 reviews), medium-rating (78 reviews), and high-rating (983 reviews) segments. The data underwent text preprocessing, bigram and trigram formation, Bag of Words representation, Latent Dirichlet Allocation (LDA) modeling, coherence score evaluation, and visualization using word clouds and pyLDAvis. The results show that the selected model for the high segment yielded 3 topics with a coherence score of 0.5042; the medium segment yielded 2 topics with a coherence score of 0.4114 based on interpretability; and the low segment yielded 4 topics with a coherence score of 0.4743. The scenario without segmentation yielded a coherence score of 0.5314 with 2 topics, but the topics remained general in nature. Rating segmentation helped clarify themes related to satisfaction, mixed evaluations, and specific visitor complaints.

Keywords: Google Maps, Latent Dirichlet Allocation, topic modeling, rating segmentation, visitor reviews.