

Abstrak

Layanan kesehatan publik BPJS Kesehatan Cabang Surabaya menjadi fokus utama dalam penilaian kepuasan masyarakat melalui ulasan digital di Google Maps. Banyaknya ulasan yang tidak terstruktur memerlukan metode analisis otomatis untuk mengekstraksi informasi penting. Penelitian ini bertujuan untuk menerapkan *topic modeling* menggunakan algoritma *Non-negative Matrix Factorization* (NMF) untuk mengidentifikasi topik utama dari 3.009 ulasan pelanggan. Tahapan penelitian meliputi pengumpulan data, *preprocessing* teks (termasuk *cleaning*, *case folding*, *tokenization*, *normalization*, *stopword removal*, dan *stemming*), representasi kata dengan TF-IDF dan N-Gram (Unigram, Bigram, Trigram), serta pemodelan topik dengan variasi jumlah topik (k). Hasil pengukuran menunjukkan bahwa model mencapai performa optimal pada jumlah topik $k=4$ dengan nilai *Coherence Score* (C_v) sebesar 0,6984. Empat topik utama yang berhasil diidentifikasi adalah (1) Pelayanan Staf Lapangan (1.308 ulasan), (2) Kepuasan dan Evaluasi Kecepatan Layanan (378 ulasan), (3) Harapan Mutu dan Perkembangan bagi Masyarakat (302 ulasan), dan (4) Kendala Operasional Antrean dan Waktu Tunggu (293 ulasan). Hasil penelitian ini memberikan implikasi praktis bagi manajemen BPJS Kesehatan Cabang Surabaya untuk memprioritaskan perbaikan pada manajemen antrean fisik dan efisiensi waktu guna meningkatkan mutu pelayanan publik secara berkelanjutan.

Kata Kunci: *Topic Modeling*, BPJS Kesehatan, Google Maps, *Non-negative Matrix Factorization*, *Coherence Score*.

Abstract

Public health services at the BPJS Kesehatan Surabaya Branch have become a primary focus in assessing community satisfaction through digital reviews on Google Maps. The vast number of unstructured reviews necessitates an automated analysis method to extract key information. This study aims to implement topic modeling using the Non-negative Matrix Factorization (NMF) algorithm to identify main topics from 3,009 customer reviews. The research stages include data collection, text preprocessing (cleaning, case folding, tokenization, normalization, stopwords removal, and stemming), word representation using TF-IDF and N-Grams (Unigram, Bigram, Trigram), and topic modeling with variations in the number of topics (k). Measurement results indicate that the model achieved optimal performance at $k=4$ topics with a Coherence Score (C_v) of 0.6984. The four main topics identified are (1) Field Staff Service (1,308 reviews), (2) Satisfaction and Service Speed Evaluation (378 reviews), (3) Quality Expectations and Development for the Community (302 reviews), and (4) Operational Queue Constraints and Waiting Time (293 reviews). These findings provide practical implications for BPJS Kesehatan Surabaya Branch management to prioritize improvements in queue management systems and time efficiency to enhance the quality of public services sustainably.

Keywords: *Topic Modeling, BPJS Kesehatan, Google Maps, Non-negative Matrix Factorization, Coherence Score.*