

INTISARI

Penelitian ini berjudul "Klasterisasi Topik Ulasan pada Aplikasi Ojek Online dengan Algoritma K-Means Clustering". Seiring dengan meningkatnya penggunaan layanan transportasi ojek online seperti Grab dan Gojek, jumlah ulasan pengguna yang tersedia di platform digital semakin bertambah. Ulasan tersebut mengandung berbagai opini dan pengalaman pengguna, tetapi bersifat tidak terstruktur, sehingga sulit dianalisis secara manual. Penelitian ini bertujuan untuk mengelompokkan ulasan pengguna terhadap layanan Grab dan Gojek berdasarkan tema utama yang sering muncul dalam ulasan menggunakan metode K-Means Clustering. Penelitian ini dibatasi pada ulasan pengguna yang diperoleh dari Google Play Store dan hanya berfokus pada analisis teks menggunakan teknik Text Mining dan machine learning.

Proses penelitian melibatkan pengumpulan data ulasan pengguna dari Google Play Store, yang kemudian diproses melalui tahapan Preprocessing data seperti Handling Duplikat, Data Cleaning, Lowercasing, Normalized, Slangword, Abbreviation, Tokenization, stopword removal, Stemming dan TF-IDF untuk representasi data numerik. Algoritma K-Means Clustering diterapkan untuk mengelompokkan ulasan berdasarkan kemiripan topik, sedangkan Silhouette Score digunakan untuk mengevaluasi kualitas kluster dan menentukan jumlah kluster optimal.

Hasil penelitian menunjukkan bahwa jumlah kluster optimal untuk dataset Grab dan Gojek adalah tiga dengan masing-masing kluster mencerminkan tema utama ulasan pengguna. Pada dataset Gojek, kluster dengan skor tertinggi berisi ulasan terkait pengalaman pengguna dan kemudahan aplikasi. Sedangkan pada dataset Grab, ulasan lebih banyak membahas penggunaan aplikasi dan promosi. Nilai Silhouette Score tertinggi untuk Gojek adalah 0.5059 pada kluster 0, sementara untuk Grab adalah 0.4704 pada kluster 2 yang menunjukkan kualitas kluster yang baik. Penelitian ini menyimpulkan bahwa metode K-Means Clustering efektif dalam mengidentifikasi pola ulasan pengguna terhadap layanan Grab dan Gojek.

Kata kunci: K-Means Clustering, analisis ulasan, Grab, Gojek, Silhouette Score

ABSTRACT

This research is titled "Topic Clustering of Reviews on Online Ride-Hailing Applications Using the K-Means Clustering Algorithm." With the increasing use of online ride-hailing services such as Grab and Gojek, the number of user reviews available on digital platforms has also grown. These reviews contain various opinions and user experiences, but they are unstructured, making manual analysis challenging. This study aims to group user reviews of Grab and Gojek services based on the main themes frequently found in the reviews using the K-Means Clustering method. This research is limited to user reviews obtained from Google Play Store and focuses solely on text analysis using Text Mining and machine learning techniques.

The research process involves collecting user review data from Google Play Store, which is then processed through data Preprocessing stages such as handling duplicates, Data Cleaning, Lowercasing, normalization, slang word and Abbreviation handling, Tokenization, stopword removal, Stemming, and TF-IDF for numerical data representation. The K-Means Clustering algorithm is applied to group reviews based on topic similarity, while the Silhouette Score is used to evaluate Cluster quality and determine the optimal number of Clusters.

The results show that the optimal number of Clusters for the Grab and Gojek datasets is three, with each Cluster representing key themes of user reviews. In the Gojek dataset, the highest-scoring Cluster contains reviews related to user experience and app convenience, whereas in the Grab dataset, reviews focus more on app usage and promotions. The highest Silhouette Score for Gojek is 0.5059 in Cluster 0, while for Grab, it is 0.4704 in Cluster 2, indicating good Cluster quality. This study concludes that the K-Means Clustering method is effective in identifying user review patterns for Grab and Gojek services.

Keywords: K-Means Clustering, Review Analysis, Grab, Gojek, Silhouette Score