

ABSTRAK

Penelitian ini bertujuan untuk menganalisis dan mengelompokkan ulasan pengguna aplikasi *Cookpad* menggunakan algoritma *K-Means* dan *K-Medoids*. Data yang digunakan berupa ulasan pengguna yang diperoleh melalui proses *scraping*, kemudian melalui tahap praproses teks seperti *cleaning*, *slang fixing*, *stopword removal*, *stemming*, *tokenizing*, hingga pembobotan menggunakan TF-IDF. Selanjutnya dilakukan reduksi dimensi untuk menyederhanakan kompleksitas data sebelum proses klasterisasi. Evaluasi performa dilakukan dengan membandingkan nilai *Silhouette Score* dan ketahanan terhadap *outlier* menggunakan algoritma DBSCAN. Hasil pengujian menunjukkan bahwa algoritma *K-Medoids* memiliki performa klasterisasi yang lebih baik dibandingkan *K-Means*. Pada jumlah klaster $k = 3$, *K-Medoids* menghasilkan nilai *Silhouette Score* sebesar 0,2966 (tanpa *outlier*) dan 0,2788 (dengan *outlier*), sedangkan *K-Means* memperoleh nilai 0,2795 (tanpa *outlier*) dan 0,2625 (dengan *outlier*). Analisis ketahanan terhadap *outlier* menunjukkan bahwa kedua algoritma memiliki daya tahan yang relatif setara, dengan total 680 data terdeteksi sebagai *outlier*. Berdasarkan hasil klasterisasi, sebagian besar ulasan pengguna termasuk dalam kategori positif, tercermin dari kata-kata dominan seperti "bagus", "sangat membantu", dan "suka". Sistem yang dikembangkan mampu melakukan klasterisasi secara otomatis, serta dapat dijadikan alat bantu dalam analisis persepsi pengguna aplikasi berbasis teks..

Kata Kunci : *Clustering, Cookpad, K-Means, K-Medoids, Silhouette Coefficient.*

ABSTRACT

This study aims to analyze and cluster user reviews of the Cookpad application using the K-Means and K-Medoids algorithms. The dataset consists of user comments obtained through web scraping, which were processed using text preprocessing techniques including cleansing, slang fixing, stopword removal, stemming, tokenization, and TF-IDF weighting. Dimensionality reduction was then applied to simplify the data before the clustering process. Performance evaluation was conducted by comparing the Silhouette Score and resilience to outliers using the DBSCAN algorithm. The results show that the K-Medoids algorithm outperformed K-Means in terms of clustering quality. With $k = 3$ clusters, K-Medoids achieved a Silhouette Score of 0.2966 (without outliers) and 0.2788 (with outliers), while K-Means scored 0.2795 (without outliers) and 0.2625 (with outliers). Both algorithms demonstrated relatively similar resilience to outliers, with a total of 680 outlier data points detected. The clustering results revealed that most user reviews were positive, indicated by frequent keywords such as “helpful”, “great”, and “like”. The developed system successfully performed automated clustering and can be used as a tool to analyze user sentiment from textual reviews.

Keywords : *Clustering, Cookpad, K-Means, K-Medoids, Silhouette Coefficient.*