ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan drama Korea berdasarkan tingkat rekomendasi dengan menggunakan algoritma K-Means. Dataset yang digunakan terdiri dari informasi rating dan jumlah vote dari berbagai judul drama, yang telah melalui tahap preprocessing untuk memastikan kualitas data. Proses klasterisasi dilakukan untuk mengidentifikasi pola dan struktur dalam data secara tidak berlabel. Hasil klasterisasi menunjukkan bahwa dua klaster optimal dapat dibentuk, masing-masing mewakili kelompok drama yang direkomendasikan dan kurang direkomendasikan. Validasi hasil klaster menggunakan metrik Silhouette Score menghasilkan nilai sebesar 0.97, yang mengindikasikan bahwa struktur klaster cukup baik dan pemisahan antar kelompok data jelas. Klaster 1 berisi drama dengan rating dan vote tinggi, sedangkan klaster 0 terdiri atas drama dengan rating dan vote rendah. Hasil ini menunjukkan bahwa algoritma K-Means efektif dalam mengelompokkan preferensi pengguna terhadap drama Korea berdasarkan data rating dan vote, serta dapat digunakan sebagai dasar sistem rekomendasi atau pengambilan keputusan dalam platform hiburan digital. Pengembangan selanjutnya dapat dilakukan dengan metode clustering lain seperti *DBSCAN*, *Hierarchical Clustering* dan penambahan fitur seperti genre untuk meningkatkan akurasi.

Kata kunci: K-Means, CRISP-DM, Clustering, Rekomendasi Film, Drama Korea.

UBP KARAWANG

ABSTRACT

This study aims to group Korean dramas based on the level of recommendation using the K-Means algorithm. The dataset used consists of rating information and the number of votes from various drama titles, which have gone through a preprocessing stage to ensure data quality. The clustering process is carried out to identify patterns and structures in the data in an unlabeled manner. The clustering results show that two optimal clusters can be formed, each representing a group of recommended and less recommended dramas. Validation of the cluster results using the Silhouette Score metric produces a value of 0.97, which indicates that the cluster structure is quite good and the separation between data groups is clear. Cluster 1 contains dramas with high ratings and votes, while cluster 0 consists of dramas with low ratings and votes. These results indicate that the K-Means algorithm is effective in grouping user preferences for Korean dramas based on rating and vote data, and can be used as a basis for a recommendation system or decision making in a digital entertainment platform. Further development can be done with other clustering methods such as DBSCAN, Hierarchical Clustering and the addition of features such as genre to improve accuracy.

Keywords: K-Means, CRISP-DM, Clustering, Movie Recommendation, Korean Drama.

UBP KARAWANG