

ABSTRAK

Perkembangan *Korean Wave* meningkatkan konsumsi drama Korea, namun penambahan judul baru menyebabkan *information overload*. Karena drama Korea kaya metadata namun matriks *rating sparse*, penelitian ini membandingkan *Item-Based Collaborative Filtering* (IBCF) dan *Content-Based Filtering* (CBF) untuk rekomendasi drama Korea. Dataset dari MyDramaList terdiri atas 404 judul dan 404.000 *rating* dari 9.352 pengguna periode 2016 - 2025. IBCF menggunakan *cosine similarity* dan *weighted sum*, sedangkan CBF menerapkan *feature-augmented* dengan menggabungkan TF-IDF dari genre, sinopsis, pemain, sutradara, penulis, dan *tags* secara *horizontal* serta *augmentasi average rating* (bobot 30%). Data dibagi 80:20 per-user split, dievaluasi dengan RMSE dan MAE. IBCF optimal di K=5 menghasilkan RMSE = 1,7233 dan MAE = 1,0687, sementara CBF mencapai RMSE = 1,2969 dan MAE = 0,9199. CBF unggul dengan RMSE 24,7% dan MAE 13,9% lebih rendah, lebih cepat, dan lebih baik menangani *cold start*. IBCF unggul dalam *serendipity* karena rekomendasi lintas genre. Hasil ini menunjukkan CBF lebih praktis untuk katalog baru dengan riwayat pengguna minim. Penelitian menyimpulkan CBF dengan augmentasi rating lebih sesuai untuk dataset drama Korea yang kaya metadata namun *sparse*, dan merekomendasikan pengembangan *hybrid* di masa depan.

Kata Kunci: Sistem Rekomendasi, Content-Based Filtering, Collaborative Filtering, RMSE, MAE

ABSTRACT

The Korean Wave has increased K-Drama consumption globally, yet new titles on streaming platforms cause information overload. Since K-Dramas are metadata-rich but have sparse rating matrices, this study compares Item-Based Collaborative Filtering (IBCF) and Content-Based Filtering (CBF) for K-Drama recommendation. The dataset, scraped from MyDramaList, comprises 404 titles and 404,000 ratings from 9,352 users (2016–2025). IBCF uses cosine similarity and weighted sum, while CBF employs feature-augmentation by horizontally stacking TF-IDF vectors from genres, synopses, cast, directors, writers, and tags, with average rating augmentation (30% weight). Data was split 80:20 per-user and evaluated using RMSE and MAE. Optimal IBCF at $K=5$ achieved $RMSE = 1.7233$ and $MAE = 1.0687$, while CBF reached $RMSE = 1.2969$ and $MAE = 0.9199$. CBF outperformed IBCF with RMSE 24.7% and MAE 13.9% lower, faster computation, and better cold-start handling. Conversely, IBCF excelled in serendipity by offering cross-genre recommendations. These findings suggest CBF is more practical for new catalogs with limited user history. This study concludes that CBF with rating augmentation is more suitable for metadata-rich yet sparse K-Drama datasets, and recommends developing hybrid approaches in future research.

Keywords: recommendation system, content-based filtering, collaborative filtering, RMSE, MAE