

ABSTRAK

Sistem rekomendasi lagu dikembangkan untuk memudahkan pengguna dalam menemukan musik sesuai preferensi di tengah banyaknya pilihan lagu yang tersedia pada platform *streaming*. Akan tetapi, metode rekomendasi konvensional seperti *Collaborative Filtering* dan *Content-Based Filtering* masih memiliki kelemahan, di antaranya permasalahan *cold start* dan *data sparsity* yang dapat menurunkan kualitas rekomendasi. Penelitian ini bertujuan untuk mengimplementasikan metode *Hybrid Filtering* dengan mengombinasikan *Collaborative Filtering* berbasis *Singular Value Decomposition* (SVD) dan *Content-Based Filtering*, mengevaluasi kinerja sistem menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE), serta menganalisis kemampuan metode *hybrid* dalam mengatasi kendala *cold start* dan *data sparsity*.

Penelitian ini menggunakan dataset Spotify dari Kaggle yang mencakup 2.000 lagu periode 1998–2020 dengan informasi seperti popularitas, atribut audio, dan interaksi pengguna. Tahap awal penelitian meliputi *preprocessing* data berupa penghapusan data duplikat, penyesuaian atribut, konversi *implicit feedback* menjadi *explicit rating*, dan pemisahan data menjadi 80% data latih serta 20% data uji. Metode *Collaborative Filtering* dengan algoritma SVD digunakan untuk memprediksi rating lagu, sementara *Content-Based Filtering* memanfaatkan algoritma *cosine similarity* untuk menghitung kemiripan lagu berdasarkan metadata. Hasil rekomendasi dari kedua metode digabungkan secara *linear weighted* dengan bobot 0,6 untuk *Content-Based Filtering* dan 0,4 untuk *Collaborative Filtering*. Sistem rekomendasi ini dikembangkan menggunakan model *waterfall* dan diuji menggunakan *black-box testing* untuk memverifikasi fungsi setiap fitur.

Pengujian performa model menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) menunjukkan kombinasi parameter terbaik terdapat pada $n_factors = 10$ dan $epochs = 40$ dengan nilai MAE = 0,1537 dan RMSE = 0,1820, yang menandakan tingkat kesalahan prediksi rendah. Pengujian keberagaman rekomendasi (*intra-diversity*) menunjukkan metode *hybrid* memiliki nilai tertinggi, yaitu 0,0135, dibandingkan *Collaborative Filtering* (0,0102) dan *Content-Based Filtering* (0,0087), yang berarti rekomendasi lebih bervariasi. Hasil pengujian *black-box* juga membuktikan seluruh fitur sistem dapat berjalan sesuai rancangan. Dengan demikian, metode *Hybrid Filtering* terbukti meningkatkan akurasi prediksi sekaligus memperluas variasi rekomendasi, serta efektif dalam mengatasi masalah *cold start* dan *data sparsity* pada sistem rekomendasi lagu.

Kata Kunci: Sistem rekomendasi lagu, *hybrid filtering*, *collaborative filtering*, *singular value decomposition*, *content-based filtering*

ABSTRACT

Music recommendation systems are developed to help users find songs that match their preferences among the vast selection of music available on streaming platforms. However, conventional recommendation methods such as Collaborative Filtering and Content-Based Filtering have certain limitations, including cold start and data sparsity issues that reduce recommendation quality. This study aims to implement a Hybrid Filtering method by combining Collaborative Filtering based on Singular Value Decomposition (SVD) and Content-Based Filtering, evaluate the system's performance using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), and analyze the hybrid method's ability to address cold start and data sparsity challenges.

This research uses a Spotify dataset from Kaggle consisting of 2,000 songs released between 1998 and 2020, containing attributes such as popularity, audio features, and user interactions. The preprocessing stage included removing duplicate records, adjusting attributes, converting implicit feedback to explicit ratings, and splitting the data into 80% training and 20% testing sets. Collaborative Filtering was implemented using the SVD algorithm to predict song ratings, while Content-Based Filtering used cosine similarity to calculate song similarity based on metadata. The two methods were integrated using a linear weighted hybrid approach with weights of 0.6 for Content-Based Filtering and 0.4 for Collaborative Filtering. The recommendation system was developed following the waterfall model and tested using black-box testing to verify functionality.

Performance evaluation using MAE and RMSE indicated that the best parameter combination was achieved with $n_factors = 10$ and $epochs = 40$, resulting in $MAE = 0.1537$ and $RMSE = 0.1820$, demonstrating high prediction accuracy. Intra-diversity testing showed that the hybrid method achieved the highest score (0.0135) compared to Collaborative Filtering (0.0102) and Content-Based Filtering (0.0087), indicating more diverse recommendations. Black-box testing confirmed that all system features functioned as intended. Overall, the Hybrid Filtering method successfully improves prediction accuracy and recommendation diversity while effectively addressing cold start and data sparsity issues in music recommendation systems.

Keywords: *Music recommendation system, hybrid filtering, collaborative filtering, singular value decomposition, content-based filtering*