

OPTIMASI SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *HYBRID FILTERING*

Septa Hendrawan¹, Yakub^{2*}

^{1,2} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

*Corresponding Author, email: yakub@ubd.ac.id

ABSTRAK

Transformasi digital dalam industri hiburan telah mendorong kebutuhan akan sistem rekomendasi yang cerdas dan adaptif, khususnya untuk platform streaming berbasis mikro seperti konter pulsa MJ CELL, untuk mengatasi permasalahan tersebut, sistem rekomendasi menjadi komponen penting dalam membantu pengguna menemukan konten yang relevan dan sesuai minat. Sistem rekomendasi umumnya dibangun menggunakan dua pendekatan utama, yaitu *Collaborative Filtering* (CF) dan *Content-Based Filtering* (CBF). CF memanfaatkan pola interaksi antar pengguna untuk memprediksi preferensi, sementara CBF menggunakan kemiripan konten berdasarkan atribut seperti genre dan synopsis. Penelitian ini bertujuan mengembangkan sistem rekomendasi film anime menggunakan pendekatan Hybrid Filtering yang menggabungkan metode *Collaborative Filtering* (CF) berbasis algoritma *Singular Value Decomposition* (SVD) dan *Content-Based Filtering* (CBF) berbasis TF-IDF dan *Cosine Similarity*. Penelitian dilakukan menggunakan dataset sekunder dari MyAnimeList melalui Kaggle, yang terdiri dari lebih dari 73.000 pengguna dan 12.000 judul anime. Proses pengembangan mengikuti metodologi CRISP-DM dengan tahapan *preprocessing*, pelatihan model, dan evaluasi kinerja. Evaluasi dilakukan menggunakan metrik RMSE, MAE, *Precision@10*, *Recall@10*, dan *F1-Score@10*. Hasil menunjukkan bahwa model CF memiliki performa prediktif terbaik (RMSE = 1,08), sementara model *Hybrid* mampu menjaga keseimbangan antara akurasi dan cakupan rekomendasi (*Precision@10* = 82,77%, *F1@10* = 0,6168). Sistem juga diimplementasikan dalam bentuk prototipe web menggunakan *Flask*, dengan fitur autentikasi pengguna, rating, dan tampilan rekomendasi personal. Penelitian ini menyimpulkan bahwa pendekatan *Hybrid Filtering* efektif untuk meningkatkan kualitas rekomendasi, serta relevan diterapkan pada *platform streaming* skala kecil yang memiliki keterbatasan data pengguna.

Kata kunci: *Collaborative Filtering*, *Content-Based Filtering*, *Hybrid Filtering*, Sistem Rekomendasi

I. PENDAHULUAN

Perkembangan teknologi digital telah membawa perubahan signifikan dalam pola konsumsi hiburan masyarakat, terutama dengan meningkatnya penggunaan layanan streaming yang menawarkan akses cepat, fleksibel, dan personal terhadap konten digital (Velamentosa & Zuliarso, 2025). Salah satu tantangan yang muncul dari melimpahnya konten adalah fenomena *paradox of choice*, yaitu ketika terlalu banyak pilihan justru menyulitkan pengguna dalam menentukan konten yang sesuai

dengan preferensi mereka (Wiputra & Shandi, 2021). Untuk mengatasi permasalahan tersebut, sistem rekomendasi menjadi komponen penting dalam membantu pengguna menemukan konten yang relevan dan sesuai minat. Sistem rekomendasi umumnya dibangun menggunakan dua pendekatan utama, yaitu *Collaborative Filtering* (CF) dan *Content-Based Filtering* (CBF). CF memanfaatkan pola interaksi antar pengguna untuk memprediksi preferensi, sementara CBF menggunakan kemiripan konten berdasarkan atribut seperti genre dan sinopsis (Jena et al., 2022). Meskipun efektif, pendekatan CF memiliki kelemahan dalam menangani kasus pengguna baru (*cold-start*), sedangkan CBF cenderung menghasilkan rekomendasi yang terlalu serupa (*over-specialization*) dan kurang bervariasi (Putri & Faisal, 2023). Untuk mengatasi keterbatasan masing-masing pendekatan, metode Hybrid Filtering dikembangkan dengan menggabungkan keunggulan CF dan CBF. Pendekatan ini terbukti lebih adaptif dalam kondisi data terbatas dan mampu memberikan rekomendasi yang lebih akurat dan beragam (Arfisko & Wibowo, 2022). Hal ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa sistem hybrid mampu mengurangi masalah *data sparsity* serta meningkatkan relevansi dan keberagaman rekomendasi (Sarhan et al., 2024).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun sistem rekomendasi film anime berbasis hybrid filtering dalam simulasi *platform streaming* berskala mikro pada konter MJ CELL. Dengan memanfaatkan algoritma *Singular Value Decomposition* (SVD) pada CF dan teknik TF-IDF serta *cosine similarity* pada CBF, sistem diharapkan dapat memberikan rekomendasi personal yang akurat meskipun dalam kondisi data yang minim.

II. METODOLOGI

Penelitian ini menggunakan pendekatan hybrid yang menggabungkan *Collaborative Filtering* (CF) dan *Content-Based Filtering* (CBF) dalam membangun sistem rekomendasi film anime berbasis web. Seluruh tahapan pengembangan sistem mengikuti metodologi CRISP-DM, yang mencakup proses dari pemahaman data hingga evaluasi model (Grotentraast et al., 2024).

2.1 Data Understanding

Dataset yang digunakan dalam penelitian ini adalah *Anime Recommendation Database* yang tersedia di Kaggle. *Dataset* terdiri atas dua file utama:

1. *anime.csv*: Berisi metadata 12.294 judul anime seperti ID, judul, genre, tipe, rating, dan jumlah anggota (members).
2. *rating.csv*: Berisi lebih dari 7 juta data interaksi pengguna, mencakup *user_id*, *anime_id*, dan skor rating (1–10).

Kedua *dataset* digabung berdasarkan *anime_id* untuk menghasilkan basis data gabungan yang digunakan pada pemodelan CF dan CBF.

2.2 Pra-Pemrosesan Data

Data rating dengan nilai -1 dihapus karena menunjukkan tidak adanya interaksi eksplisit. Selain itu, entri dengan nilai kosong atau tidak lengkap pada kolom penting seperti *genre*, *rating*, atau judul juga dibuang. Kolom teks seperti *genre*, *type*, dan judul dinormalisasi ke huruf kecil, menghapus tanda baca, serta menghilangkan entri genre seperti *unknown*. Hanya pengguna yang memiliki minimal 50 interaksi (rating) yang dipertahankan. Hal ini bertujuan mengurangi masalah *sparsity* dan meningkatkan kualitas prediksi dalam pendekatan CF (Silaparasetty, 2020). Setelah pembersihan, kedua *dataset* digabung menghasilkan file *merge_clean.csv*, yang digunakan untuk pelatihan model *hybrid*.

2.3 Pemisahan Data

Dataset dibagi menjadi dua subset: data pelatihan (80%) dan data pengujian (20%) menggunakan metode *holdout* per pengguna. Dengan teknik ini, setiap pengguna tetap muncul pada kedua subset, menghindari bias *cold-start* saat evaluasi.

2.4 Penyetelan Hyperparameter

Parameter yang diuji meliputi *n_factors*: [50, 100, 200], *n_epochs*: [20, 50, 100], *lr_all*: [0.002, 0.005, 0.01], *reg_all*: [0.02, 0.05, 0.1]. Penyetelan dilakukan dengan *RandomizedSearchCV* dan validasi silang 5-fold untuk meminimalkan nilai RMSE (Pradana & Wibowo, 2024).

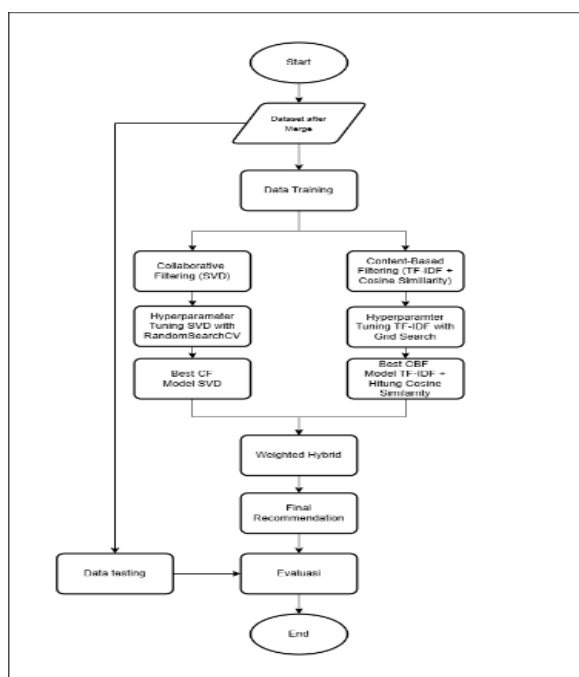
Model CBF dibangun dengan menggunakan teknik TF-IDF Vectorizer yang diterapkan pada kombinasi judul, genre, dan tipe anime. Parameter yang disetel antara lain *min_df*, *max_df*, *ngram_range*, dan *max_features*, untuk menjaga keseimbangan antara akurasi dan efisiensi komputasi (Lestari, 2024).

2.5 Pelatihan Model

Model CF dilatih menggunakan algoritma *Singular Value Decomposition* (SVD) dari pustaka surprise, menghasilkan matriks interaksi laten antara pengguna dan anime (Girsang et al., 2020). Model CBF dibangun dengan teknik TF-IDF vectorizer yang diterapkan pada kombinasi judul, genre, dan tipe anime. Kemiripan antar item dihitung menggunakan cosine similarity (Pramesti & Santiyasa, 2022). Model hybrid dibangun dengan pendekatan *weighted hybrid*, yang menggabungkan skor prediksi dari CF dan CBF menggunakan rumus :

$$Score_{Hybrid} = \alpha \cdot Score_{CF} + (1 - \alpha) \cdot Score_{CBF}$$

Nilai bobot α ditentukan berdasarkan hasil evaluasi performa masing-masing model, dengan tujuan untuk menyeimbangkan keakuratan prediksi dan relevansi rekomendasi yang diberikan (Mukti & Mardhiyah, 2022).



Gambar 1. Proses Pelatihan Model

Evaluasi sistem dilakukan melalui dua pendekatan utama. Pertama, evaluasi prediksi rating menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE), untuk mengukur akurasi numerik prediksi terhadap rating aktual (Pramarta & Baizal, 2022). Kedua, evaluasi rekomendasi Top-N menggunakan metrik *Precision@10*, *Recall@10*, dan *F1-Score@10*, untuk mengukur relevansi rekomendasi yang dihasilkan. Ketiga metrik ini dihitung berdasarkan 10 rekomendasi anime teratas yang diberikan kepada setiap pengguna, sebagaimana dilakukan pula dalam studi (Valentino & Setiawan, 2024).

III. HASIL DAN PEMBAHASAN

3.1 Evaluasi Model Rekomendasi

Tiga pendekatan sistem rekomendasi dievaluasi menggunakan metrik RMSE, MAE, *Precision@10*, *Recall@10*, dan *F1-Score@10*. Model *Collaborative Filtering* (CF) menggunakan algoritma SVD dan menunjukkan akurasi prediksi terbaik dengan RMSE 1.0802, MAE 0.8093, serta *Precision@10* sebesar 82.92%. Model *Content-Based Filtering* (CBF) yang berbasis TF-IDF dan *cosine similarity* memiliki akurasi lebih rendah, namun relevan untuk menangani *cold-start user* (Pramesti & Santiyasa, 2022).

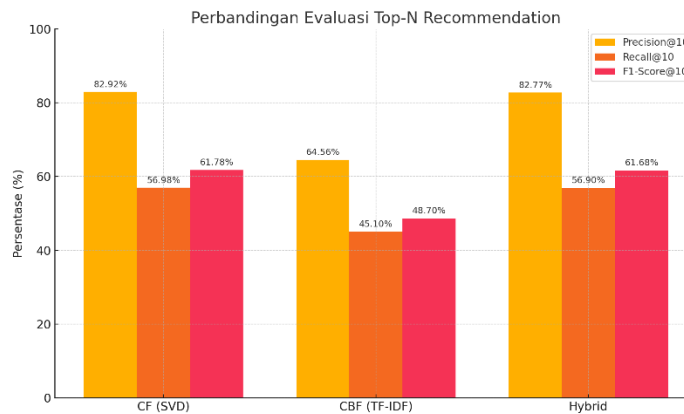
Model *Hybrid Filtering*, yang menggabungkan hasil prediksi CF dan CBF dengan bobot 0.8:0.2, menghasilkan performa seimbang (RMSE 1.5836, MAE 1.3477, *Precision@10* 82.77%, dan *F1-Score* 61.68%). Meskipun sedikit di bawah CF dalam prediksi rating, model *hybrid* lebih adaptif terhadap variasi data pengguna.

3.2 Visualisasi Perbandingan Kinerja Model

Untuk memudahkan analisis, Tabel 1 dan Gambar 2 menyajikan perbandingan langsung dari ketiga model berdasarkan semua metrik evaluasi.

Tabel 1. Perbandingan Kinerja Semua Model

Model	RMSE	MAE	Precision@10	Recall@10	F1-Score@10
CF (SVD)	1,080	0,809	82.92%	56.98%	61.78%
CBF	5,874	5,659	64.56%	45.10%	48.70%
Hybrid	1,584	1,348	82.77%	56.90%	61.68%



Gambar 1. Visualisasi Evaluasi Precision, Recall, dan F1-Score

IV. SIMPULAN

Penelitian ini berhasil membangun sistem rekomendasi film anime berbasis *hybrid filtering* yang menggabungkan metode *Collaborative Filtering* (SVD) dan *Content-Based Filtering* (TF-IDF + *cosine similarity*). Sistem dikembangkan untuk simulasi platform streaming skala kecil seperti MJ CELL yang tidak memiliki histori pengguna. Model *hybrid* yang dikembangkan menunjukkan kinerja yang seimbang antara akurasi prediksi dan cakupan rekomendasi, dengan hasil evaluasi Precision@10 sebesar 82.77% dan F1-Score@10 sebesar 0.6168. Hal ini menunjukkan bahwa sistem mampu memberikan rekomendasi yang relevan dan konsisten, bahkan dalam kondisi data yang terbatas.

DAFTAR PUSTAKA

- Arfisko, H. H., & Wibowo, A. T. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering dan Content-Based Filtering. *E-Proceeding of Engineering*, 9(3), 2149–2159. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18066>
- Girsang, A. S., Al Faruq, B., Herlianto, H. R., & Simbolon, S. (2020). Collaborative Recommendation System in Users of Anime Films. *Journal of Physics: Conference Series*, 1566(1), 012057. <https://doi.org/10.1088/1742-6596/1566/1/012057>
- Grotentraast, J., Bukhsh, S. F., Bouali, N., & Moreira, J. R. (2024). *MSc Computer*

- Science Final Project Adapted CRISP-DM approach for recommendation system development for most suitable open-source ETL tool.*
- Jena, A., Jaiswal, A., Lal, D., Rao, S., Ayubi, A., & Sachdeva, N. (2022). *Recommendation System For Anime Using Machine Learning Algorithms.*
- Lestari, P. (2024). Sistem Rekomendasi untuk Maksimalisasi Industri Film dengan Metode Demographic Filtering dan Content Based Filtering. *Jurnal Ilmu Komputer Dan Informatika*, 4(1), 1–10. <https://doi.org/10.54082/jiki.104>
- Mukti, K. T., & Mardhiyah, I. (2022). Sistem Rekomendasi Pembelian Lisensi Film Menggunakan Pendekatan Hybrid Filtering (Studi Kasus : Film Animasi Jepang). *JURSISTEKNI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 4(3), 127–139. <https://doi.org/https://doi.org/10.52005/jursistekni.v4i3.116>
- Pradana, M. A., & Wibowo, A. T. (2024). *Filtering With Word2Vec and Restricted Boltzmann.* 9(1), 231–241.
- Pramarta, A., & Baizal, A. (2022). Hybrid Recommender System Using Singular Value Decomposition and Support Vector Machine in Bali Tourism. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 7(2), 408–418. <https://doi.org/10.29100/jipi.v7i2.2770>
- Pramesti, D. A. P. D., & Santiyasa, I. W. (2022). Penerapan Metode Content-Based Filtering dalam Sistem Rekomendasi Video Game. *Jnatia*, 1(1), 229–234.
- Putri, H. D., & Faisal, M. (2023). Analyzing the Effectiveness of Collaborative Filtering and Content-Based Filtering Methods in Anime Recommendation Systems. *Jurnal Komtika (Komputasi Dan Informatika)*, 7(2), 124–133. <https://doi.org/10.31603/komtika.v7i2.9219>
- Sarhan, A. M., Ayman, H., Wagdi, M., Ali, B., Adel, A., & Osama, R. (2024). Integrating machine learning and sentiment analysis in movie recommendation systems. *Journal of Electrical Systems and Information Technology*, 11(1), 53. <https://doi.org/10.1186/s43067-024-00177-7>
- Silaparasetty, N. (2020). Machine Learning Concepts with Python and the Jupyter Notebook Environment. In *Machine Learning Concepts with Python and the Jupyter Notebook Environment: Using Tensorflow 2.0*. Apress. <https://doi.org/10.1007/978-1-4842-5967-2>
- Valentino, N., & Setiawan, E. B. (2024). Movie Recommender System on Twitter



- Using Weighted Hybrid Filtering and GRU. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(2), 159–172. <https://doi.org/10.22219/kinetik.v9i2.1941>
- Velamentosa, D., & Zuliarso, E. (2025). SISTEM REKOMENDASI FILM MENGGUNAKAN METODE CONTENT-BASED FILTERING. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2), 2918–2922. <https://doi.org/10.36040/jati.v9i2.13251>
- Wiputra, M. M., & Shandi, Y. J. (2021). Perancangan Sistem Rekomendasi Menggunakan Metode Collaborative Filtering dengan Studi Kasus Perancangan Website Rekomendasi Film. *Media Informatika*, 20(1), 1–18. <https://doi.org/10.37595/mediainfo.v20i1.54>