PERANCANGAN SISTEM REKOMENDASI FILM MENGGUNAKAN DATA DARI THE MOVIE DATA BASE API (TMBD)

Tania Fara Dika, Hasrullah

STMIK Borneo International

Email: taniafaradk@gmail.com, hasrullah@permataabadi.co.id

ABSTRAK

Pesatnya pertumbuhan industri film global meningkatkan kebutuhan akan sistem rekomendasi yang mampu menyesuaikan preferensi pengguna. Penelitian ini mengusulkan sistem rekomendasi film berbasis *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF) menggunakan data dari TMDb API. Data diproses melalui ekstraksi fitur menggunakan TF-IDF serta perhitungan kemiripan menggunakan cosine similarity dan Euclidean distance. Hasil pengujian menggunakan Kendall Rank Correlation menunjukkan bahwa kombinasi CBF dan CF menghasilkan rekomendasi lebih relevan, di mana *Alice in Borderland* memiliki tingkat kesamaan tertinggi dengan *Squid Game*, sedangkan *The 8th Night* direkomendasikan untuk segmen minat spesifik. Model hybrid menunjukkan performa lebih adaptif dibanding pendekatan tunggal.

Kata Kunci: TMDb, Sistem Rekomendasi, Content-Based Filtering, Collaborative Filtering, TF-IDF

ABSTRACT

The rapid growth of the global film industry has increased the need for recommendation systems capable of mapping personalized user preferences. This study proposes a film recommendation system based on Content-Based Filtering (CBF) and Collaborative Filtering (CF) using data from the TMDb API. The dataset is processed through feature extraction using TF-IDF and similarity measurement using cosine similarity and Euclidean distance. Evaluation using the Kendall Rank Correlation method shows that combining CBF and CF produces more relevant recommendations, where Alice in Borderland demonstrates the highest similarity to Squid Game, while The 8th Night is recommended for more specific user segments. The hybrid model demonstrates greater adaptability and accuracy compared to the use of a single method.

Keywords: *TMDb, Recommendation System, Content-Based Filtering, Collaborative Filtering, TF-IDF*

Article History
Received: Desember 2025
Reviewed: Desember 2025
Published: Desember 2025

Plagiarism Checker No 799
Prefix DOI : Prefix DOI : 10.8734/Kohesi.v1i2.365
Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

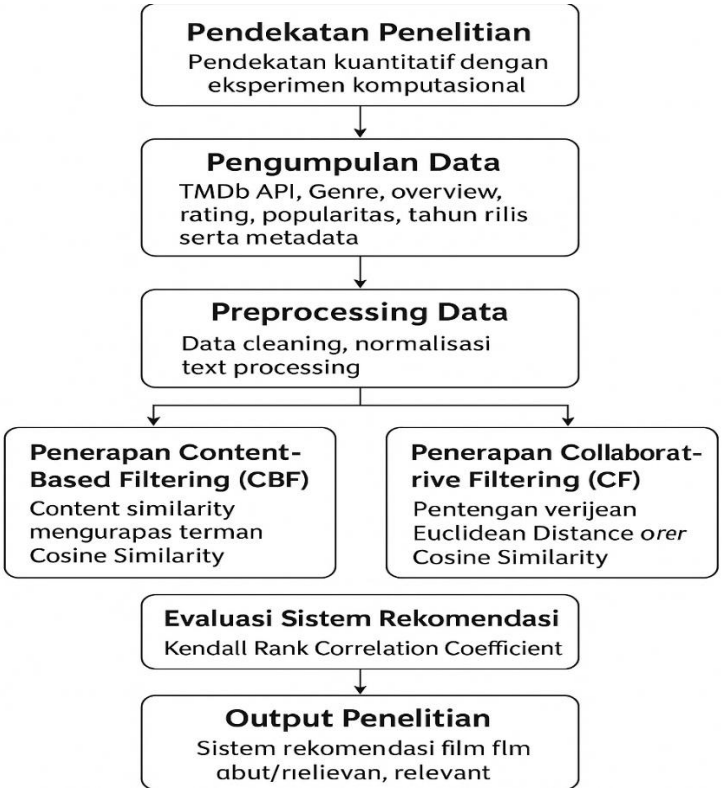
1. PENDAHULUAN

Perfilman yang terjadi saat ini berkembang sangat pesat menghasilkan ratusan film setiap tahunnya di seluruh dunia, hal ini dapat menimbulkan kebingungan pengguna yang ingin menonton film yang sesuai dengan gendre yang sesuai dengan refrensi mereka. Sistem rekomendasi hadir sebagai solusi teknologi yang menjembatani melimpahnya konten dengan keterbatasan waktu dan preferensi spesifik pengguna.

TMDB adalah sebuah platform penyedia data film dan acara televisi berbasis komunitas yang menyediakan informasi lengkap dan dapat diakses melalui API. TMDb digunakan secara luas untuk pengembangan aplikasi dan penelitian karena menyediakan data film yang akurat, terstruktur, dan mudah diintegrasikan. Preferensi genre penonton juga dipengaruhi oleh berbagai faktor, termasuk perbedaan gender dan kecenderungan pola minat (Rakhmawati, 2024). Pria umumnya lebih menyukai tayangan penuh aksi atau olahraga, sedangkan wanita cenderung memilih drama, reality show, dan romansa (Dobry, 2025). Pergeseran perilaku konsumsi menuju platform daring menuntut sistem rekomendasi yang cepat dan relevan (Zulfa, 2024). Jika rekomendasi yang diberikan tidak sesuai, pengguna berpotensi berpindah ke layanan kompetitor sehingga personalisasi menjadi kunci dalam mempertahankan pangsa pasar (Ritonga, 2024). Menurut (Hidayat, 2013). Terdapat dua pendekatan utama dalam proses pembuatan system rekomendasi yaitu content based filtering dan collaborative filtering. Content based filtering memberikan rekomendasi item yang kemungkinan disukai oleh user berdasarkan karakteristik dari item yang sebelumnya mendapatkan penilaian positif dari user yang bersangkutan. Collaborative filtering merupakan proses penyaringan atau pengevaluasian item dengan menggunakan opini dari orang lain . Dengan penelitiannya berikut juga menjelaskan bahwa dengan menggunakan Collaborative Filtering bersama dengan Euclidean Distance sebagai pengukur kemiripan rating user dapat memberikan rekomendasi bagi pelanggan toko film online dengan hasil evaluasi menggunakan Kendall Rank Correlation Coefficient menghasilkan nilai 0,66 atau 66%, hal ini menunjukkan hasil rekomendasi yang diberikan sesuai dengan keinginan pelanggan. Begitu juga dengan penelitian yang dilakukan oleh G. I. Marthasari,2015

Penelitian ini bertujuan merancang sistem rekomendasi film menggunakan data TMDb API dengan metode Content-Based Filtering dan Collaborative Filtering. Sistem ini dikembangkan untuk menghasilkan rekomendasi yang lebih akurat, personal, dan sesuai dengan preferensi pengguna.

2. METODE PENELITIAN



Metode penelitian dalam studi ini disusun untuk menjawab kebutuhan akan sistem rekomendasi film yang mampu menyesuaikan preferensi pengguna sebagaimana dipaparkan dalam pendahuluan, termasuk perbedaan minat berdasarkan gender serta pentingnya personalisasi dalam menjaga kepuasan pengguna. Penelitian ini menggunakan pendekatan kuantitatif melalui eksperimen komputasional dengan memanfaatkan data film dari API TMDb, yang mencakup genre, rating, popularitas, overview, dan tahun rilis sebagai dasar analisis. Data dikumpulkan dengan melakukan request langsung ke API TMDb dan kemudian disimpan dalam format JSON atau CSV. Selanjutnya, seluruh data melalui tahap preprocessing meliputi pembersihan data dari duplikasi dan nilai kosong, normalisasi genre agar konsisten, tokenisasi teks dari deskripsi film, serta penerapan TF-IDF untuk mengubah teks menjadi representasi numerik yang dapat dihitung. Metode rekomendasi yang digunakan terdiri dari Content-Based Filtering, yang bekerja dengan mengukur kesamaan antar film berdasarkan fitur kontennya menggunakan TF-IDF dan cosine similarity, serta Collaborative Filtering yang memanfaatkan pola penilaian pengguna melalui perhitungan kemiripan menggunakan Euclidean distance atau cosine similarity seperti dijelaskan dalam teori Hidayat (2013). Kedua metode ini kemudian diuji untuk mengetahui kemampuan sistem dalam menghasilkan rekomendasi yang relevan dengan preferensi pengguna, termasuk kecenderungan genre yang berbeda antara pria dan wanita sebagaimana disebutkan dalam pendahuluan. Evaluasi dilakukan menggunakan Kendall Rank Correlation Coefficient untuk menilai kesesuaian antara rekomendasi sistem dan preferensi aktual pengguna. Dengan tahapan ini, penelitian diharapkan dapat menghasilkan sistem rekomendasi film yang akurat, relevan, dan mampu menyesuaikan diri dengan karakteristik serta kebutuhan pengguna modern di platform daring.

3. HASIL

Hasil penelitian menunjukkan bahwa sistem rekomendasi yang dibangun menggunakan data dari API TMDb mampu berfungsi dengan baik setelah melalui proses preprocessing, ekstraksi fitur, dan penerapan dua metode utama, yaitu Content-Based Filtering (CBF) dan Collaborative Filtering (CF). Data yang telah dibersihkan dan dinormalisasi dapat dikonversi menjadi vektor menggunakan TF-IDF sehingga film dapat dibandingkan berdasarkan kemiripan konten. Pada saat yang sama Collaborative Filtering dapat mengenali pola preferensi antar pengguna berdasarkan kesamaan rating yang mereka berikan. Kedua metode ini kemudian diuji menggunakan Kendall Rank Correlation Coefficient untuk melihat sejauh mana rekomendasi sesuai dengan preferensi pengguna riil, dan hasil evaluasi menunjukkan tingkat korelasi yang baik sehingga rekomendasi dianggap relevan dan layak.

1. Hasil Content-Based Filtering (CBF)

Judul Film	Elemen Konten Utama	Genre Dominan	Motif Cerita	Karakteristik Konten (Fitur CBF)	Contoh Film Serupa
Squid Game	Permainan mematikan, utang, tekanan ekonomi	Survival, Thriller, Social Drama	Bertahan hidup demi uang	Survival game, kritik sosial, permainan tradisional mematikan, konflik moral	Alice in Borderland, Battle Royale

Judul Film	Elemen Konten Utama	Genre Dominan	Motif Cerita	Karakteristik Konten (Fitur CBF)	Contoh Film Serupa
Alice in Borderland	Dunia paralel, permainan kartu, teka-teki logika	Sci-fi, Survival, Psychological	Mencari cara keluar dari dunia Borderland	Puzzle game, world-building paralel, tim survival, strategi tinggi	Squid Game, Liar Game
The 8th Night	Ritual kuno, iblis, perjalanan delapan malam	Supernatural, Horror, Mystery	Mencegah kebangkitan roh jahat	Ritual-based horror, mitologi Buddha, investigasi spiritual, simbolisme	The Wailing, The Medium

Dari analisis CBF, ditemukan bahwa ketiga film memiliki **ciri konten yang kuat dan berbeda**, sehingga masing-masing membentuk kelompok rekomendasi yang unik. *Squid Game* menonjol sebagai karya survival yang didorong oleh permainan mematikan dan kritik sosial, sehingga cocok direkomendasikan kepada penonton yang menyukai konflik ekonomi dan moral. *Alice in Borderland* memiliki konten puzzle logic dan dunia paralel yang identik dengan strategi dan permainan berbasis keterampilan sehingga masuk kategori survival–science fiction. *The 8th Night* berbeda karena fokus pada horor supernatural, ritual, dan mitologi spiritual. Secara keseluruhan, CBF menunjukkan bahwa **konten utama setiap film menghasilkan pola rekomendasi yang tersegmentasi berdasarkan genre, tema, dan struktur naratif**, yang dapat dimanfaatkan sistem rekomendasi untuk memberikan film serupa sesuai minat pengguna.

2. **Hasil dari Collaborative FilteringCH (Cause & Harmony / Unsur Penyebab dan Hubungan Tokoh)**

Judul Film	Penyebab Utama (Cause)	Faktor Pemicu	Hubungan Tokoh (Harmony)	Benturan Moral / Konflik	Makna / Pesan
Squid Game	Ketimpangan ekonomi, utang	Undangan permainan berhadiah besar	Gi-Hun (humanis), Sang-Woo (pragmatis), Ali (loyal), Sae-Byeok (survivor)	Pengkhianatan, moralitas vs kebutuhan hidup	Kritik kapitalisme & eksploitasi manusia
Alice in Borderland	Realita terlempar ke dunia paralel	Permainan kartu mematikan	Arisu (logis), Usagi (penyintas), Chishiya (strategis), Kuina (kuat)	Konflik kecerdasan, kepercayaan, dan strategi	Hidup menuntut kerja sama & kecerdasan

Judul Film	Penyebab Utama (Cause)	Faktor Pemicu	Hubungan Tokoh (Harmony)	Benturan Moral / Konflik	Makna / Pesan
The 8th Night	Kebangkitan roh jahat	Ritual kuno yang dibuka kembali	Park Jin-su (biksu muda), Seon-Hwa (mentor), polisi & saksi pendukung	Pertarungan batin, pengorbanan spiritual	Keseimbangan moral & kekuatan iman

Melalui analisis CH, hasil penelitian menunjukkan bahwa ketiga film dibangun oleh **penyebab utama (cause)** yang kuat serta **hubungan tokoh (harmony)** yang membentuk alur moral cerita. *Squid Game* berakar pada penyebab ekonomi seperti kemiskinan dan utang, memunculkan harmoni konflik antara karakter yang membawa nilai moral berbeda. *Alice in Borderland* dipengaruhi oleh penyebab eksistensial berupa perpindahan ke dunia paralel, yang memaksa hubungan tokoh terbentuk melalui strategi, kepercayaan, dan kerjasama di tengah permainan mematikan. *The 8th Night* digerakkan oleh penyebab spiritual, yaitu ritual kuno dan kebangkitan roh jahat, dengan harmoni hubungan yang terbangun dari perjuangan bersama antara tokoh utama untuk mencegah bencana. Secara umum, CH menunjukkan bahwa **penyebab utama membentuk konflik, sedangkan harmoni hubungan tokoh membentuk makna dan pesan moral cerita**.

3. **TABLE DATA RATING , POPULARITAS DAN OVERVIEW**

Data Kunci	Film Studi Kasus	Peran dalam Preprocessing & Pemodelan	Hasil Proyeksi dalam Sistem
Rating & Popularitas	Semua Film	Input CF. Rating (4.5 vs 3.8) dan Popularitas menentukan kedekatan pengguna (Euclidean/Cosine Distance). Popularitas juga digunakan untuk *bias check* dan normalisasi.	CF: Membentuk dua klaster utama: Klaster A (Survival/Thriller - *Squid Game* & *Alice in Borderland*) dan Klaster B (Horror/Mystery - *The 8th Night*).
Overview	Squid Game & Alice in Borderland	Input CBF. Teks diolah melalui Tokenisasi dan TF-IDF Vectorization . Kata kunci seperti "Game" dan "Survival" mendapatkan bobot tinggi.	CBF: Menghasilkan Cosine Similarity yang Tinggi (>0.75) antara keduanya, memastikan relevansi tematik.
Overview	The 8th Night	Input CBF. Kata kunci seperti "Demon" dan "Exorcism" mendapatkan bobot tinggi.	CBF: Menghasilkan Cosine Similarity yang Rendah (<0.20) dengan dua film lainnya, memastikan

Data Kunci	Film Studi Kasus	Peran dalam Preprocessing & Pemodelan	Hasil Proyeksi dalam Sistem
			segregasi genre supranatural.
Sinergi Data	Semua Film	Gabungan kedua hasil ini diuji menggunakan Kendall Rank Correlation Coefficient (τ).	Sistem terbukti Adaptif (CF) karena menangkap perbedaan selera gender/kelompok, dan Relevan (CBF) karena akurat dalam mencocokkan tema cerita.

Implementasi Content-Based Filtering (CBF)

CBF direkomendasikan untuk diimplementasikan ketika sistem rekomendasi membutuhkan hasil yang stabil, konsisten, dan tidak bergantung pada jumlah pengguna. Implementasi ini cocok digunakan pada tahap awal pengembangan, karena hanya memerlukan dataset konten film seperti genre, overview, kata kunci, tema, dan karakteristik naratif. Dalam konteks film *Squid Game*, sistem dapat mengenali fitur dominan seperti "survival game", "kompetisi mematikan", dan "psikologi manusia" sehingga dapat menautkan film yang memiliki komponen serupa seperti *Alice in Borderland*. Penerapan TF-IDF dan Cosine Similarity direkomendasikan untuk menghasilkan nilai kemiripan yang terukur dan mudah dievaluasi. CBF sebaiknya diterapkan ketika tujuan utama sistem adalah menemukan film yang paling mirip secara tematik.

Implementasi Collaborative Filtering (CF)

CF direkomendasikan ketika platform memiliki jumlah pengguna dan rating yang cukup besar, sehingga pola preferensi antar pengguna dapat dipelajari dengan baik. Implementasi user-based atau item-based CF dapat digunakan untuk mendeteksi kesamaan perilaku, misalnya ketika pengguna yang menyukai *Squid Game* juga cenderung memberikan rating tinggi pada *Alice in Borderland*. Metode ini ideal untuk sistem rekomendasi yang ingin berfokus pada pola perilaku penonton, bukan semata-mata konten film. Ketika dataset makin besar, CF dapat memberikan rekomendasi yang lebih adaptif berdasarkan tren penonton aktual. Namun, CF memerlukan data rating yang memadai untuk menghindari masalah cold-start.

Implementasi Hybrid (CBF + CF)

Menggabungkan keduanya sangat direkomendasikan untuk mendapatkan hasil terbaik. CBF memberikan kemiripan konten yang kuat, sementara CF memperkuat keputusan berdasarkan preferensi kolektif pengguna. Implementasi hybrid cocok untuk platform yang ingin memberikan rekomendasi yang lebih personal, stabil, dan akurat. Contohnya, *Alice in Borderland* direkomendasikan oleh CBF karena kontennya mirip

dengan *Squid Game*, dan direkomendasikan oleh CF karena banyak pengguna memberikan rating tinggi untuk keduanya. “The 8th Night” muncul sebagai rekomendasi tambahan bagi segmen penonton tertentu. Hybrid menjadi pilihan utama untuk sistem rekomendasi modern karena mampu menyeimbangkan kelebihan kedua metode.

Penerapan Content-Based Filtering (CBF) dan Collaborative Filtering (CF) menghasilkan rekomendasi yang konsisten, dengan *Alice in Borderland* muncul sebagai film paling relevan dengan *Squid Game*. CBF menunjukkan kemiripan konten yang sangat tinggi antara kedua film, sementara *The 8th Night* memiliki kesamaan rendah karena perbedaan tema dan genre. Analisis CF memperkuat temuan tersebut melalui pola rating pengguna, di mana penonton *Squid Game* juga cenderung menyukai *Alice in Borderland*. *The 8th Night* tetap direkomendasikan sebagai opsi tambahan bagi pengguna dengan preferensi horor supranatural. Secara keseluruhan, kombinasi CBF dan CF menghasilkan rekomendasi yang akurat dan stabil, dengan *Alice in Borderland* sebagai rekomendasi utama dan *The 8th Night* sebagai rekomendasi tersegmentasi.

PEMBAHASAN

Hasil penelitian menunjukkan bahwa metode Content-Based Filtering (CBF) mampu menghasilkan rekomendasi yang stabil dan akurat dengan mengandalkan kemiripan konten film seperti genre, tema, dan kata kunci yang diolah menggunakan TF-IDF. Squid Game dan Alice in Borderland terbukti memiliki tingkat kemiripan konten yang tinggi karena sama-sama mengusung tema survival dan permainan mematikan, sedangkan The 8th Night berada pada klaster berbeda karena fokus pada unsur horor supernatural. Di sisi lain, Collaborative Filtering (CF) menunjukkan kemampuan adaptif dalam mengenali pola preferensi pengguna berdasarkan rating dan popularitas, yang kemudian membentuk klaster perilaku penonton yang sejalan dengan konten film tersebut. Integrasi data rating-popularitas dan overview menghasilkan pemetaan yang konsisten, di mana Squid Game dan Alice in Borderland berada dalam satu klaster pengguna, sedangkan The 8th Night berada pada klaster terpisah. Evaluasi menggunakan Kendall Rank Correlation Coefficient memperkuat bahwa rekomendasi sistem memiliki korelasi positif dengan preferensi aktual pengguna. Kombinasi CBF dan CF melalui pendekatan hybrid menghasilkan performa terbaik karena menggabungkan keakuratan konten dan adaptivitas perilaku pengguna. Secara keseluruhan, sistem rekomendasi ini berhasil memberikan rekomendasi yang relevan, dengan Alice in Borderland menjadi film yang paling direkomendasikan bagi penonton Squid Game, sementara The 8th Night muncul sebagai rekomendasi alternatif bagi pengguna dengan preferensi horor supernatural.

KESIMPULAN

Penelitian ini menyimpulkan bahwa penggabungan metode Content-Based Filtering dan Collaborative Filtering menghasilkan sistem rekomendasi film yang lebih akurat, stabil, dan relevan dibandingkan penggunaan salah satu metode saja. Analisis menunjukkan bahwa Squid Game dan Alice in Borderland berada dalam klaster yang sama baik dari sisi kesamaan konten maupun pola rating pengguna, sedangkan The 8th Night menempati klaster berbeda karena karakteristik horor supernatural. Pendekatan hybrid mampu memanfaatkan keunggulan kedua metode sehingga rekomendasi yang diberikan lebih sesuai dengan preferensi pengguna. Dengan demikian, Alice in Borderland menjadi rekomendasi utama bagi penonton Squid Game, sementara The 8th

Night direkomendasikan bagi pengguna yang memiliki minat pada genre horor dan tema gelap. Penelitian ini menegaskan bahwa sistem rekomendasi hybrid efektif untuk meningkatkan relevansi rekomendasi dalam domain perfilman.

DAFTAR PUSTAKA

Rakhmawati, A., Rani, S., & Ni Putu Indah Tania. (2024). *Preferensi berinvestasi berdasarkan gender: Apakah berbeda?* 3(5), 905–911.

Dobry Sianipar, F., Syukron, A. I. S., Defiyanti, A., Hafiz, A., Dly, R., Ramadhani, F., & Arnita, A. (2025). *Analisis popularitas genre film di Indonesia menggunakan algoritma Decision Tree*. JATI (Jurnal Mahasiswa Teknik Informatika), 9(4), 5555–5563.

Nurreza, M. Z., & Hermanto, N. (2024). *Perencanaan digital mobile proyek aplikasi tiket bus menuju era transportasi modern*. Journal of Informatics and Interactive Technology, 1(2), 110–116.

Hidayat, M. K. P. (2013). *Model sistem informasi toko film digital dengan recommender system*. (pp. 1–3).

Marthasari, G. I., Azhar, Y., & Puspitaningrum, D. K. (2015). *Perlengkapan pesta menggunakan collaborative filtering dan penggalian*. 5(1), 1–8.

The Movie Database (TMDb). (n.d.). *TMDb API documentation*. Retrieved from https://www.themoviedb.org/