

A Hybrid Content-Based and Collaborative Filtering Approach for Recommending Optimal Movie Release Months

Ukasyah Muntaha^{1*}, Johannes Krisjon Silitonga^{2*}, Elisabeth Claudia Simanjuntak^{3*}, M. Syamsuddin Wisnubroto^{4*}, Fajri Farid^{5*}, Meida Cahyo Untoro^{6*}

*Program Studi Sains Data, Fakultas Sains, Institut Teknologi Sumatera, Lampung Selatan, Indonesia

**Program Studi Teknik Informatika, Fakultas Teknologi Industri, Institut Teknologi Sumatera, Lampung Selatan, Indonesia
ukasyah.122450028@student.itera.ac.id¹, johannes.122450043@student.itera.ac.id², elisabeth.122450123@student.itera.ac.id³,
syamsuddin.wisnubroto@sd.itera.ac.id⁴, fajri.farid@sd.itera.ac.id⁵, cahyo.untoro@if.itera.ac.id⁶

Article Info

Article history:

Received 2025-11-07
Revised 2026-03-05
Accepted 2026-04-10

Keyword:

*Collaborative filtering,
Content based filtering,
Film industry,
Hybrid recommender system,
Release strategy.*

ABSTRACT

The film industry faces increasingly intense competition, making release timing a critical factor in maximizing box-office performance. However, decisions regarding release months are often driven by producers' intuition rather than systematic data analysis. This study proposes a data-driven decision-support system to recommend optimal movie release months by learning historical release and profitability patterns from previously released films. A hybrid recommender system combining Content-Based Filtering (CBF) and Collaborative Filtering (CF) is developed. The study utilizes data from The Movie Database (TMDB), including film genres, release dates, and profit margins. CBF measures content similarity based on genre features, while CF captures latent relationships between films and release months using historical profit patterns through Singular Value Decomposition (SVD). These two approaches are integrated using a weighted sum mechanism to produce a hybrid score. Model performance is evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Experimental results demonstrate that the proposed hybrid model outperforms single-method approaches, achieving an RMSE of 1.60 and an MAE of 1.03, which are lower than those obtained by standalone CBF and CF models. Further analysis reveals that genres such as Action, Adventure, and Science-Fiction exhibit relatively stable profitability trends across specific release months. These findings indicate that the hybrid CBF-CF approach effectively captures both content-based similarity and temporal profitability patterns, making it suitable as a strategic decision-support tool for determining movie release timing.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Produksi film di dunia meningkat drastis melampaui angka sebelum pandemi, tercatat tahun 2024 sekitar 9.511 film diproduksi secara global, sedangkan tahun 2020 hanya mencapai 5.656 film [1]. Lonjakan produksi ini mencerminkan pertumbuhan industri perfilman, namun sekaligus menandakan semakin ketatnya persaingan antar film dalam menarik perhatian penonton. Dalam kondisi persaingan yang tinggi, strategi distribusi dan penentuan waktu rilis yang kurang optimal dapat menyebabkan banyak

film gagal mencapai potensi jumlah penonton dan pendapatan maksimal.

Selain persaingan yang meningkat, industri film juga menghadapi tantangan dalam menghadirkan konten yang relevan bagi segmen penonton yang semakin beragam. Tidak hanya pada aspek cerita dan kualitas produksi, proses distribusi dan strategi perilisasi juga menjadi salah satu faktor penting dalam menentukan keberhasilan sebuah film di pasar [2]. Produser dan distributor perlu mempertimbangkan berbagai faktor eksternal seperti preferensi penonton, tren musiman, serta perilaku konsumsi media yang terus berubah. Namun, dalam praktiknya, penentuan waktu rilis film sering

kali masih bergantung pada pengalaman subjektif atau praktik konvensional, tanpa didukung oleh analisis data yang komprehensif.

Dalam konteks pengambilan keputusan berbasis data, sistem rekomendasi telah banyak digunakan sebagai solusi untuk meningkatkan efektivitas distribusi konten digital, termasuk di industri perfilman. Sistem rekomendasi mampu membantu pengguna maupun pelaku industri dalam menyaring informasi dan membuat keputusan yang lebih tepat berdasarkan pola data historis [3]. Dalam industri perfilman yang kompetitif, penggunaan sistem rekomendasi menjadi sangat dibutuhkan karena mampu mendukung pengambilan keputusan berbasis data secara objektif dan terukur, sehingga tidak lagi bergantung pada intuisi semata. Sebelum adanya sistem rekomendasi, strategi perilis film sering kali tidak mempertimbangkan faktor penting seperti waktu rilis, karakteristik pasar, dan preferensi penonton. Waktu atau bulan rilis terbukti berpengaruh signifikan terhadap pendapatan box-office dan tingkat kunjungan penonton [4]. Oleh karena itu, integrasi faktor temporal seperti bulan rilis ke dalam sistem rekomendasi tidak hanya relevan dari sisi teknis, tetapi juga strategis, karena dapat membantu produser dan distributor dalam menentukan waktu rilis yang optimal untuk memaksimalkan performa komersial film di pasar.

Seiring berkembangnya penelitian di bidang sistem rekomendasi, muncul perhatian khusus terhadap pentingnya aspek temporal atau *temporal-awareness*. Integrasi faktor waktu dalam sistem rekomendasi terbukti dapat meningkatkan relevansi dan akurasi hasil rekomendasi, terutama dalam domain yang dipengaruhi oleh tren musiman dan perubahan preferensi pengguna [5]. Namun, sebagian besar sistem rekomendasi film yang ada masih berfokus pada rekomendasi film kepada pengguna, bukan pada rekomendasi strategi perilis film bagi produser atau distributor. Hal ini menunjukkan adanya celah penelitian dalam pengembangan sistem rekomendasi yang berorientasi pada pengambilan keputusan strategis di industri perfilman.

Metode sistem rekomendasi film yang umum digunakan antara lain berbasis CBF (*Content-Based Filtering*) dan CF (*Collaborative Filtering*) [6]. CBF menyarankan film berdasarkan kemiripan konten seperti genre, pemeran, atau sutradara, sedangkan CF memanfaatkan pola interaksi antar pengguna dari data historis. Kedua algoritma ini memiliki keunggulan dan keterbatasan masing-masing, di mana CBF tidak memerlukan data *rating* pengguna tetapi terbatas pada informasi eksplisit, sedangkan CF memerlukan data *rating* namun mampu menangkap pola preferensi secara implisit [7].

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan sistem rekomendasi film berbasis algoritma *Hybrid* yang menggabungkan metode CBF dan CF untuk menghasilkan rekomendasi yang lebih akurat dan kontekstual. Selain itu, penelitian ini menambahkan faktor temporal berupa bulan rilis film. Hasil penelitian diharapkan dapat memberikan kontribusi nyata bagi pengembangan sistem rekomendasi di industri perfilman. Sistem yang

diusulkan diharapkan mampu membantu produser dan distributor film dalam menentukan waktu rilis serta kanal distribusi yang tepat. Dengan demikian, efektivitas promosi dan keberhasilan komersial film dapat meningkat melalui rekomendasi yang lebih relevan dan informatif.

Penelitian ini menggunakan kerangka sistem rekomendasi sebagai alat bantu pengambilan keputusan strategis, berbeda dengan sistem rekomendasi film umumnya yang berfokus pada memberikan rekomendasi film kepada pengguna. Dalam penelitian ini, bulan perilis film digunakan sebagai target dengan film dianggap sebagai item referensi dan bulan perilis ditetapkan sebagai target rekomendasi.

CBF cenderung menghasilkan rekomendasi yang sangat mirip sehingga kurang mampu menangkap pola kompleks dalam data dan dinamika preferensi pasar serta memiliki keterbatasan dalam menghasilkan rekomendasi yang beragam dan adaptif jika hanya mengandalkan fitur konten saja [7]. CF sering menghadapi tantangan *data sparsity* dan *cold-start* terutama ketika tidak ada cukup riwayat interaksi untuk user atau item baru sehingga kondisi ini membuat model CF sulit memberikan rekomendasi yang relevan [8]. Pendekatan hybrid secara konsisten memberikan performa yang lebih stabil dan interpretatif dibandingkan metode tunggal, khususnya pada domain rekomendasi film yang dipengaruhi oleh faktor konten dan perilaku pasar secara simultan [9].

II. METODE

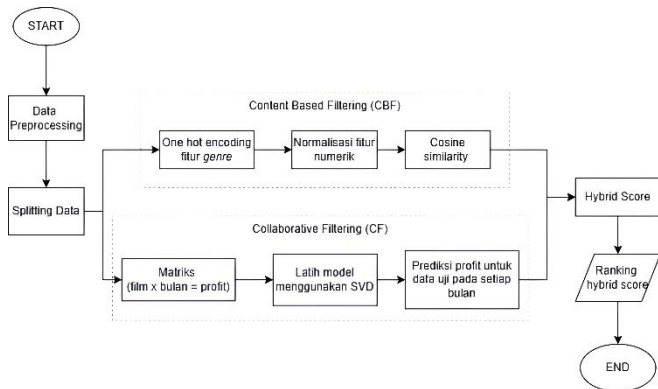
Penelitian ini merancang dan mengimplementasikan sistem rekomendasi film berbasis algoritma *hybrid Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF) yang tidak hanya memberikan rekomendasi film tetapi juga memperkirakan bulan rilis serta platform distribusi Production House yang optimal. Objek penelitian ini adalah data industri perfilman.

Gambar 1 menunjukkan bagaimana tahapan penelitian ini dilakukan secara sistematis. Proses dimulai dari preprocessing data, yaitu tahap pembersihan, transformasi, dan penyesuaian data agar siap digunakan dalam proses analisis. Data yang telah diproses kemudian diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengubah teks seperti overview dan keywords menjadi representasi numerik [7]. Selanjutnya algoritma *Content-Based Filtering* (CBF) melakukan pengukuran dua buah objek (kemiripan antar film) yang dinyatakan dalam dua buah vector yang merepresentasikan setiap kata dalam teks menggunakan *Cosine Similarity* [12]. Nilai kesamaan yang dihasilkan menunjukkan seberapa dekat hubungan antara dua film berdasarkan konten deskriptifnya. Hasil dari pendekatan CBF ini kemudian digabungkan dengan pendekatan CF (*Collaborative Filtering*) yang menganalisis kesamaan preferensi antar pengguna berdasarkan pola *rating*.

Untuk mengatasi kelemahan dari masing-masing pendekatan, penelitian ini mengadopsi metode *hybrid* yang secara konsisten terbukti mampu meningkatkan akurasi dan ketahanan sistem rekomendasi yang lebih akurat, relevan, dan personal [11]. Selain itu, sistem rekomendasi ini juga

melakukan analisis untuk memperkirakan bulan rilis film serta platform distribusi yang optimal berdasarkan pola data historis dari Production House. Dengan pendekatan *Hybrid* ini, sistem rekomendasi yang dihasilkan diharapkan dapat memberikan hasil yang tidak hanya bersifat konten spesifik tetapi juga kontekstual terhadap tren perilisan dan distribusi film di industri perfilman.

Penelitian ini berfokus pada pembelajaran pola historis bulan rilis film berdasarkan karakteristik dan performa film-film yang telah dirilis sebelumnya, bukan pada prediksi langsung bulan rilis optimal untuk film yang sepenuhnya baru tanpa data pendukung. Sistem rekomendasi yang dibangun memanfaatkan pendekatan hybrid Content-Based Filtering dan Collaborative Filtering untuk mengidentifikasi kecenderungan bulan rilis yang secara historis menunjukkan performa profitabilitas lebih baik bagi film dengan karakteristik serupa. Oleh karena itu, target rekomendasi dalam penelitian ini didefinisikan sebagai bulan rilis yang paling sesuai berdasarkan pola historis dalam dataset TMDB.



Gambar 1. Bagan alir penelitian

A. Deskripsi Data

Teknik pengumpulan data yang dilakukan dalam penelitian ini akan memanfaatkan data yang bersumber dari open *source* di kaggle. Dataset ini terdiri dari 35.311 baris dan 24 kolom yang mencakup berbagai genre, negara produksi, dan periode waktu rilis. Setiap entri film dilengkapi dengan atribut konten, informasi temporal, serta data performa seperti popularitas dan pendapatan, sehingga relevan untuk pengembangan sistem rekomendasi berbasis konten dan kolaboratif. Untuk menjaga konsistensi dan relevansi data, penelitian ini menggunakan film yang dirilis pada rentang 2014-2023. Pemilihan rentang waktu ini bertujuan untuk menghindari bias akibat perubahan ekstrem pada industri film di periode yang terlalu lama.

TABEL I
DESKRIPSI DATA

Nama Kolom	Tipe Data	Deskripsi	Jumlah Nilai Unik
id	int64	Nomor identifikasi unik untuk setiap film	35.311

title	object	Judul film	33.070
vote_average	float64	Nilai rata rata rating dari pengguna TMDB (skala 1-10)	4.425
vote_count	int64	Jumlah pengguna yang memberikan rating	3.564
status	object	Status film, (Realeased, Postproduction)	1
release_date	object	Tanggal rilis film (format YYYY-MM-DD)	14.335
revenue	int64	Total pendapatan film dalam USD	10.406
runtime	int64	Durasi film dalam menit	284
adult	bool	Menandakan apakah film memiliki konten dewasa	2
backdrop_path	object	Path atau URL gambar latar belakang film	34.539
budget	int64	Anggaran produksi film dalam USD	1.658
homepage	object	Tautan ke situs web resmi film	9.364
imdb_id	object	ID film pada database IMDB	35.203
original_language	object	Bahasa asli film (contoh: en, fr, ja)	86
original_title	object	Judul asli film sesuai bahasa asalnya	33.886
overview	object	Ringkasan atau sinopsis film	35.092
popularity	float64	Indeks popularitas berdasarkan TMDB	17.037
poster_path	object	Path ke poster utama film	35.276
tagline	object	Kalimat promosi singkat film	20.725
genres	object	Kategori atau genre film	4.414
production_companies	object	Daftar perusahaan produksi yang terlibat	26.020
production_countries	object	Negara tempat film diproduksi	2.274
spoken_languages	object	Bahasa yang digunakan dalam film	2.268
keywords	object	Kata kunci yang merepresentasikan tema film	26.809

B. Content Based Filtering (CBF)

CBF bertugas untuk memanfaatkan analisis konten dan atribut item sehingga menghasilkan rekomendasi yang relevan dan personal. Ada 2 jenis algoritma CBF yaitu berdasarkan memori dan model, dimana cara kerja berbasis model dengan menyimpan produk dan data pengguna pada memori yang selanjutnya dilakukan pembuatan estimasi data sedangkan berdasarkan model bekerja dengan cara merespon permintaan pengguna [3].

$$\text{Cos } \alpha = \frac{a \cdot b}{\|a\| \|b\|} \tag{1}$$

Keterangan:

$A \cdot B$: hasil perkalian dot antara vektor A dan B
 $\|A\|$ dan $\|B\|$: panjang vektor masing-masing

Sistem rekomendasi ini dibangun dengan proses CBF yang akan melakukan perubahan genre, proses CBF dimulai dengan melakukan encoding terhadap fitur kategorikal seperti *genres* menggunakan metode *one-hot encoding* atau *TF-IDF vectorization* [7]. Setelah semua fitur direpresentasikan dalam bentuk numerik, sistem menghitung kemiripan antara film menggunakan *Cosine Similarity*. Apabila nilai kemiripan antara film melebihi batas ambang (*threshold*), maka film tersebut dianggap relevan dan dimasukkan ke dalam daftar rekomendasi. Output dari tahap ini adalah daftar film dengan tingkat kesamaan konten tertinggi, yang menjadi dasar dalam pembentukan sistem rekomendasi *hybrid*.

Pada pendekatan Content-Based Filtering (CBF), penelitian ini membatasi fitur konten pada genre film sebagai representasi utama karakteristik film. Pemilihan genre didasarkan pada pertimbangan bahwa *genre* merupakan atribut konten yang stabil, mudah diinterpretasikan, dan umum digunakan sebagai fitur utama dalam sistem rekomendasi berbasis konten pada kasus perfilman [3]. Fitur lain yang berpotensi mempengaruhi profit dan waktu rilis, seperti popularitas aktor, anggaran produksi, status franchise, dan target pasar, tidak dimasukkan secara eksplisit pada tahap CBF untuk menghindari peningkatan kompleksitas fitur serta masalah sparsity data. Faktor-faktor tersebut tidak diabaikan, melainkan dipelajari secara tidak langsung melalui pendekatan Collaborative Filtering (CF) yang memanfaatkan pola profit historis dan waktu rilis sebagai sinyal umpan balik implisit. Pendekatan ini sejalan dengan konsep sistem rekomendasi *hybrid*, di mana CBF digunakan untuk menangkap kesamaan konten yang eksplisit, sedangkan CF digunakan untuk memodelkan pola laten berbasis performa historis dan dinamika temporal [5], [6]. Dengan konsep ini, sistem *hybrid* memungkinkan kontribusi yang saling melengkapi antara aspek tematik dan aspek *finansial-temporal* dalam menghasilkan rekomendasi bulan perilis film yang lebih kontekstual dan terukur.

C. Collaborative Filtering (CF)

Collaborative Filtering merupakan teknik rekomendasi yang memanfaatkan interaksi antara item dengan pengguna. CF terbagi menjadi 2 jenis utama yaitu berdasarkan pengguna, CF merekomendasikan konten yang disukai pengguna yang linear, sedangkan jenis kedua berdasarkan item CF merekomendasikan item yang mirip dengan item lain yang telah disukai oleh pengguna [6].

Pendekatan yang digunakan adalah item-based collaborative filtering dengan memanfaatkan informasi profit dan waktu rilis film. Dataset diubah menjadi matriks Film \times Bulan, di mana setiap nilai sel merepresentasikan skor profit pada bulan tertentu. Nilai kosong di dalam matriks diestimasi menggunakan model CF berbasis Singular Value Decomposition (SVD), teknik faktorisasi matriks yang efektif dalam menangani masalah *sparsity* data pada dataset film,

untuk memprediksi profit yang mungkin diperoleh suatu film apabila dirilis pada bulan tertentu [12].

Proses ini menghasilkan prediksi nilai profit untuk masing-masing bulan dalam satu tahun (bulan 1–12). Hasil prediksi ini kemudian digunakan untuk menilai waktu rilis yang paling optimal berdasarkan pola historis performa film lainnya.

Dalam penelitian ini, pendekatan CF diintegrasikan ke dalam konteks sistem pendukung keputusan dengan menggambarkan interaksi antara film dan bulan rilis, di mana nilai keuntungan dilihat sebagai sinyal umpan balik tidak langsung yang mencerminkan performa historis suatu film. Berbeda dengan fokus pada pengguna, film menjadi entitas yang berinteraksi dengan bulan rilis tertentu melalui tingkat keuntungan yang dihasilkan. Formulasi ini memungkinkan penggunaan Collaborative Filtering berbasis item dengan teknik faktorisasi matriks seperti Singular Value Decomposition (SVD) untuk mengidentifikasi pola laten profitabilitas temporal yang berdasarkan pada karakteristik film. Metode serupa juga diterapkan dalam bidang lain yang tidak berorientasi pada pengguna, seperti prediksi penjualan dan rekomendasi strategi bisnis, di mana indikator finansial dipakai sebagai pengganti interaksi [15]. Oleh karena itu, cara CF yang diterapkan tetap selaras secara metodologis dengan prinsip Collaborative Filtering, tetapi dimodifikasi untuk tujuan penelitian yang fokus pada mendukung pengambilan keputusan strategis terkait penentuan waktu rilis film.

Meskipun CBF dan CF sering diterapkan dalam sistem rekomendasi film, masing-masing metode menunjukkan keterbatasan yang signifikan ketika digunakan untuk menentukan bulan perilis film [16]. Pendekatan CBF yang mengandalkan kemiripan atribut konten, seperti genre dan karakteristik film, cenderung mengabaikan pola historis profitabilitas serta dinamika temporal pasar [17], sehingga kurang mampu menjelaskan keterkaitan antara waktu rilis dan kinerja finansial film. Sementara itu, CF efektif dalam menangkap pola laten berdasarkan data historis, namun metode ini rentan terhadap permasalahan data sparsity dan minimnya konteks konten, khususnya pada kombinasi film–bulan rilis yang jarang terjadi. Artikel terkait umumnya menyatakan bahwa pendekatan *hybrid* lebih unggul, tetapi belum menguraikan secara eksplisit keterbatasan CBF dan CF dalam konteks penentuan bulan rilis, sehingga dasar teoretis integrasi kedua metode tersebut masih kurang kuat. Oleh karena itu, penelitian ini mengadopsi pendekatan *hybrid* CBF–CF sebagai solusi yang lebih komprehensif, dengan mengombinasikan informasi kesamaan konten film dan pola temporal profitabilitas untuk menghasilkan rekomendasi bulan rilis yang lebih optimal dan beralasan secara teoretis.

D. Hybrid Recommendation

Hybrid recommendation system adalah penggabungan dua atau lebih pendekatan rekomendasi untuk menghasilkan hasil yang lebih akurat. Metode *Hybrid* score ini dilakukan dengan pendekatan *weighted sum* yang menghasilkan skor akhir rekomendasi [5]. Metode *hybrid* yang digunakan dalam

penelitian ini adalah *Weighted Hybrid Approach*, di mana skor akhir dihitung dengan persamaan berikut:

$$Hybrid\ Score = \alpha \times Similarity\ Score + (1 - \alpha) \times CF\ SCORE \quad (2)$$

Parameter α (alpha) berperan sebagai bobot penyeimbang yang mengatur kontribusi relatif antara Content-Based Filtering (CBF) dan Collaborative Filtering (CF). Dalam penelitian ini, nilai α ditetapkan secara statis berdasarkan pendekatan empiris, dengan tujuan menjaga keseimbangan antara karakteristik konten film (seperti genre) dan pola profit serta bulan rilis yang dipelajari dari data historis. Pembobotan statis dipilih untuk menjaga stabilitas model dan menghindari kompleksitas tambahan dari proses optimasi parameter, mengingat fokus penelitian ini adalah analisis pola bulan rilis film terdahulu, bukan pada pencarian bobot optimal secara otomatis. Pendekatan *weighted hybrid* semacam ini telah banyak digunakan dalam berbagai domain sistem rekomendasi dan terbukti mampu meningkatkan relevansi hasil dibandingkan pendekatan tunggal [15]. Film dengan nilai *hybrid score* tertinggi dipilih sebagai rekomendasi utama, sementara film lainnya disusun dalam bentuk peringkat rekomendasi. Selain itu, sistem juga menghasilkan rekomendasi bulan rilis yang diperkirakan paling menguntungkan berdasarkan pola profit historis, sehingga pendekatan *hybrid* ini mampu mengatasi keterbatasan masing-masing metode, seperti keterbatasan variasi rekomendasi pada CBF dan permasalahan *cold-start* pada CF.

E. Evaluasi Model

Model *hybrid* dievaluasi menggunakan Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE) untuk mengukur akurasi dari prediksi yang dilakukan. MAE dan RMSE digunakan untuk mengukur akurasi prediksi performa film pada bulan rilis tertentu. RMSE (*Root Mean Squared Error*) menjumlahkan kuadrat dari selisih antara nilai aktual dan nilai prediksi. RMSE dihitung dengan mengakar kuadratkan rata-rata dari kuadrat selisih tersebut. MAE (*Mean Squared Error*) menghitung rata-rata dari selisih absolut antara nilai yang diprediksi oleh model dan nilai aktual. Metrik ini mengukur besaran rata-rata kesalahan dalam prediksi tanpa mempertimbangkan arahnya [11].

Dalam penelitian ini, bulan rilis diperlakukan sebagai label yang memiliki struktur ordinal (bulan 1–12). Meskipun target akhir berupa kategori, proses rekomendasi tidak dilakukan melalui klasifikasi langsung, melainkan melalui estimasi skor numerik profit untuk setiap bulan rilis menggunakan pendekatan Collaborative Filtering. Nilai prediksi numerik tersebut kemudian digunakan untuk menentukan bulan rilis dengan skor tertinggi sebagai hasil rekomendasi. Pendekatan ini memungkinkan evaluasi dilakukan pada tingkat kesalahan estimasi yang mempengaruhi urutan rekomendasi, sehingga kesalahan kecil pada estimasi skor bulan yang berdekatan secara ordinal tetap dapat diukur secara proporsional. Dengan demikian, penggunaan RMSE dan MAE tetap relevan dalam konteks rekomendasi waktu rilis berbasis pemeringkatan

skor, meskipun keluaran akhir sistem berupa kategori bulan rilis.

III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan dataset TMDB (*The Movie Database*) yang berisi berbagai informasi tentang film, seperti judul, genre, tanggal rilis, anggaran, pendapatan (*revenue*), serta informasi deskriptif lainnya. Tujuan utama dari penelitian ini adalah untuk memberikan rekomendasi waktu atau bulan yang tepat bagi pembuat film untuk merilis film baru, berdasarkan analisis film-film sebelumnya dan keuntungan yang diperoleh. Penelitian ini memanfaatkan pendekatan sistem rekomendasi *hybrid*, yaitu kombinasi antara *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF). Dimana, CBF digunakan untuk mencari film yang memiliki kemiripan konten terutama berdasarkan genre. CF (dengan algoritma SVD) digunakan untuk memprediksi potensi keuntungan (profit margin) suatu film jika dirilis pada bulan tertentu, berdasarkan pola perilaku film-film lain. *Hybrid* model menggabungkan kedua pendekatan dengan bobot parameter α (alpha), di mana hasil akhir diperoleh dari kombinasi skor kemiripan konten dan hasil prediksi CF.

TABEL II
EVALUASI MODEL CF

RMSE (<i>Root Mean Squared Error</i>)	1,6
MAE (<i>Mean Absolute Error</i>)	1,03

Nilai RMSE < 2 menunjukkan bahwa model memiliki kemampuan prediksi yang cukup baik, dengan rata-rata kesalahan prediksi sekitar 1 satuan dari nilai aktual. Hal ini berarti model dapat memperkirakan potensi keuntungan film pada bulan tertentu dengan akurasi yang memadai. Nilai MAE sebesar 1.03 menunjukkan rata-rata deviasi kesalahan yang relatif kecil, sehingga model dapat digunakan untuk mendukung pengambilan keputusan dalam konteks rekomendasi bulan perilisan.

Sistem *hybrid* (CF + CBF) melakukan proses rekomendasi dengan menggunakan film tertentu sebagai input untuk menghasilkan daftar film serupa beserta bulan perilisan dan skor *hybrid* nya. Hasil contoh rekomendasi ditunjukkan sebagai berikut.

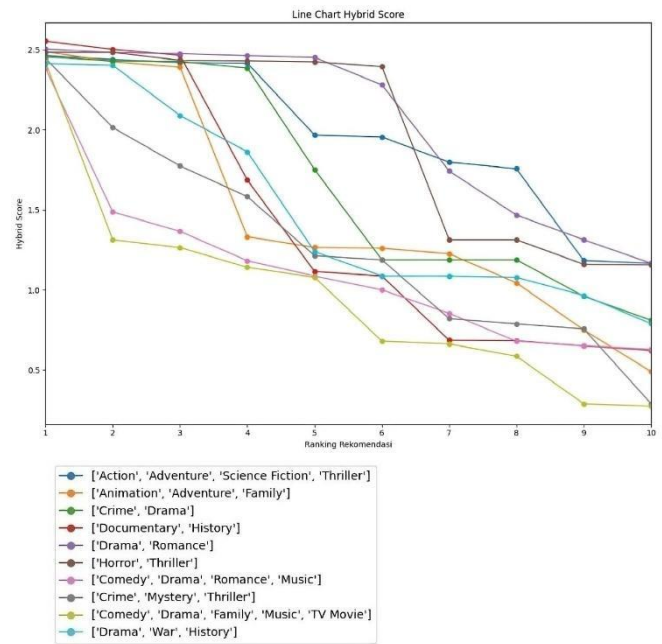
TABEL III
CONTOH PENGGUNAAN HYBRID RECOMMENDATION

Judul Film	Genre	Film Paling Mirip	Bulan Rilis	CF Score	CBF Score	Hybrid Score
Mad Max: Fury Road	Action, Adventure, Science Fiction, Thriller	Black Panther	2	1	4,66	2,46
		The Avengers	4	1	4,6	2,44
		Iron Man 3	4	1	4,54	2,42
		Guardians of	7	1	3,42	1,97

		the Galaxy				
Frozen	Animasi, Keluarga, Petualangan	The Rescuers	6	1	4,72	2,49
		Alice in Wonderland	7	1	4,57	2,43
		Peter Pan	2	1	4,48	2,39
		How To Train Your Dragon	3	1	1,83	1,33
		Anastasia	11	1	1,57	1,23
La La Land	Drama, Romance, Music, Comedy	Begin Again	9	1	4,46	2,38
		Sing Street	3	1	2,22	1,49
		Drumline	12	1	1,91	1,37
		Coyote Ugly	7	1	1,47	1,18
		We Are Your Friends	8	1	1	1

Pada Table III terlihat pola kesamaan konten dan keuntungan film serupa, nilai *hybrid score* tertinggi mampu memberikan bulan yang relatif menguntungkan untuk perilis film-film dengan genre-genre yang serupa. Sebaliknya, film dengan skor *hybrid* yang lebih rendah menunjukkan bahwa kesamaan genre dan keuntungan pada bulan tersebut tidak sekuat kandidat teratas. Dengan demikian, skor *hybrid* memberikan dasar kuantitatif untuk menentukan bulan optimal bagi peluncuran film.

Analisis perilaku tiap genre dilakukan melalui visualisasi berupa *line chart* yang menampilkan perubahan *hybrid score* terhadap peringkat rekomendasi (1–10) untuk masing-masing film dengan beberapa genre utama. Dapat dilihat pada Gambar.1 genre seperti action, adventure, dan science-fiction menunjukkan garis yang cenderung tinggi dan stabil, artinya genre ini memiliki banyak film dengan tingkat kemiripan konten yang tinggi dan keuntungan yang relatif konsisten. Genre seperti documentary menunjukkan fluktuasi yang lebih tajam, mengindikasikan bahwa hanya sedikit film dengan profit tinggi yang mirip secara konten. Genre romance dan drama memiliki skor sedang namun merata, menandakan pasar yang stabil namun tidak terlalu ekstrem dalam hal keuntungan musiman. Dari pola ini dapat disimpulkan bahwa genre aksi dan fiksi ilmiah paling diuntungkan oleh sistem *hybrid* karena memiliki banyak referensi film serupa dengan profit yang konsisten di berbagai bulan perilis.



Gambar 2. Line chart hybrid score

Hasil penelitian menunjukkan bahwa pendekatan hybrid yang menggabungkan metode Content-Based Filtering (CBF) dan Collaborative Filtering (CF) mampu menghasilkan rekomendasi bulan perilis film dengan tingkat relevansi yang baik. Pada tahap awal, CBF digunakan untuk mengidentifikasi film-film dengan kesamaan genre sebagai representasi kemiripan konten. Selanjutnya, prediksi potensi profit pada setiap bulan perilis dilakukan menggunakan model CF berdasarkan pola keuntungan historis film-film serupa. Nilai hybrid score kemudian dihitung sebagai hasil penggabungan skor kemiripan konten dan hasil prediksi profit bulanan, sehingga nilai hybrid score tertinggi dipilih sebagai rekomendasi bulan perilis film.

Genre Action, Adventure, dan Science-Fiction menunjukkan nilai hybrid score yang relatif tinggi dan stabil pada beberapa peringkat rekomendasi. Hasil visualisasi ini menunjukkan bahwa film-film genre tersebut memiliki banyak referensi film serupa yang secara konsisten menghasilkan keuntungan pada bulan tertentu. Oleh karena itu, sistem hybrid secara berulang menghasilkan nilai hybrid score tertinggi pada bulan yang sama atau berdekatan. Oleh karena itu, analisis pola genre tidak hanya digunakan untuk tujuan deskriptif, tetapi juga digunakan secara langsung sebagai dasar dalam proses membuat rekomendasi tentang bulan di mana film akan dirilis.

TABEL IV
NILAI RMSE & MAE

Model	RMSE	MAE
Content-Based Filtering (CBF)	2.31	1.94
Collaborative Filtering (CF)	1.82	1.21
Hybrid CBF–CF (Proposed)	1.60	1.03

Model CBF menghasilkan nilai RMSE dan MAE tertinggi. Hal ini menunjukkan bahwa pendekatan berbasis konten saja tidak dapat menangkap pola temporal profitabilitas film secara optimal. Model CF menunjukkan performa yang lebih baik dibandingkan CBF, karena mampu memodelkan hubungan laten antara film dan bulan perilisan berdasarkan pola keuntungan historis. Namun demikian, model CF masih memiliki keterbatasan akibat sparsity data dan kurangnya konteks konten film.

Model hybrid CBF-CF yang diusulkan memperoleh nilai RMSE dan MAE terendah dibandingkan kedua baseline. Hasil ini membuktikan bahwa penggabungan kesamaan konten film dan pola profit historis mampu meningkatkan akurasi rekomendasi bulan perilisan. Dengan demikian, pendekatan hybrid efektif dalam mengatasi kelemahan masing-masing metode tunggal dan lebih sesuai digunakan sebagai sistem pendukung keputusan penentuan waktu rilis film.

Nilai evaluasi yang diperoleh dengan RMSE sebesar 1,6 dan MAE 1,03 menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang relatif rendah. Hal ini berarti perbedaan antara hasil prediksi dan nilai aktual masih berada dalam batas wajar untuk konteks sistem rekomendasi berbasis data film. Meskipun masih terdapat selisih rata-rata sekitar dua poin, model ini dapat dianggap cukup stabil untuk mendukung proses rekomendasi dalam domain perfilman.

Penerapan model *hybrid* juga memperlihatkan bahwa hasil rekomendasi yang dihasilkan lebih bervariasi dibandingkan dengan penggunaan metode tunggal. Pada beberapa kasus, seperti film *Frozen*, sistem mampu menampilkan rekomendasi film lain dengan bulan perilisan yang berdekatan dan memiliki genre yang serupa, seperti *The Rescuers*. Hal ini membuktikan bahwa sistem *hybrid* dapat menyeimbangkan aspek konten (CBF) dan pola kebiasaan perilaku (CF) secara seimbang. Selain itu, nilai *hybrid score* yang dihasilkan juga cenderung lebih tinggi pada film-film dengan tingkat kemiripan genre dan nilai prediksi profit yang tinggi, menunjukkan kemampuan model dalam mengidentifikasi film yang relevan secara tematik maupun performa finansial.

Penelitian ini dibangun menggunakan dataset The Movie Database (TMDb) yang merepresentasikan industri film global dengan dominasi produksi dari pasar Hollywood. Oleh karena itu, hasil rekomendasi bulan perilisan yang dihasilkan oleh sistem hybrid CBF-CF pada penelitian ini secara langsung merepresentasikan pola perilisan dan profitabilitas yang berlaku pada pasar film internasional, khususnya pasar Amerika Utara dan produksi beranggaran besar.

Secara konseptual, pendekatan sistem rekomendasi yang diusulkan bersifat umum dan dapat digeneralisasikan ke konteks industri film lain, termasuk pasar film regional atau nasional, selama tersedia data historis yang memuat informasi konten film. Adaptasi terhadap konteks regional dapat dilakukan dengan mengganti dataset pelatihan menggunakan data film lokal, sehingga pola musiman, preferensi genre, serta dinamika pasar yang spesifik terhadap wilayah tertentu dapat dipelajari oleh model.

Namun, perlu dicatat bahwa validitas eksternal sistem ini dapat dipengaruhi oleh perbedaan karakteristik industri, seperti budaya menonton, kalender libur nasional, kebijakan sensor, serta struktur distribusi film yang berbeda antarnegara. Oleh karena itu, penerapan sistem pada konteks non-TMDb memerlukan proses pelatihan ulang (*retraining*) dan validasi tambahan agar rekomendasi bulan perilisan yang dihasilkan tetap relevan dan akurat.

Meskipun model hybrid menunjukkan performa yang lebih baik dibandingkan pendekatan tunggal, hasil rekomendasi yang dihasilkan masih menunjukkan variasi tingkat stabilitas antar genre dan bulan perilisan. Beberapa genre dengan jumlah sampel yang lebih sedikit atau pola keuntungan yang tidak konsisten cenderung menghasilkan *hybrid score* yang fluktuatif, sehingga rekomendasi bulan perilisan menjadi kurang stabil. Selain itu, faktor temporal dalam analisis ini masih diperlakukan secara statis melalui bulan rilis, tanpa mempertimbangkan variasi pola musiman lintas tahun atau perubahan tren industri film dari waktu ke waktu. Kondisi ini menunjukkan bahwa meskipun model mampu menangkap pola historis secara umum, masih terdapat keterbatasan dalam memodelkan dinamika temporal yang lebih kompleks.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma *Hybrid Recommender System* yang menggabungkan *Content-Based Filtering* (CBF) dan *Collaborative Filtering* (CF) mampu menghasilkan rekomendasi yang cukup efektif dalam menentukan bulan perilisan film yang berpotensi menghasilkan keuntungan optimal. Proses integrasi kedua metode ini memungkinkan sistem untuk memanfaatkan keunggulan masing-masing pendekatan: CBF dalam menganalisis kemiripan konten film berdasarkan genre, serta CF dalam mengenali pola hubungan antar film melalui nilai *profit margin* dan waktu perilisan.

Dari hasil pengujian model dengan data latih sebesar 80% dan data uji sebesar 20%, diperoleh nilai RMSE sebesar 1,6 dan MAE sebesar 1,03. Nilai ini menunjukkan bahwa tingkat kesalahan prediksi model masih berada dalam kategori rendah hingga sedang, yang berarti bahwa model memiliki kemampuan prediksi yang cukup baik. Dengan demikian, sistem yang dibangun dapat memberikan rekomendasi bulan perilisan yang relevan untuk film-film baru berdasarkan kombinasi kesamaan konten dan pola keuntungan film lain.

Selain itu, hasil visualisasi *hybrid score* dari film-film yang dianalisis memperlihatkan adanya variasi yang menarik di mana beberapa genre seperti action, adventure, dan drama cenderung memiliki fluktuasi skor yang lebih stabil dibandingkan genre lain seperti horror atau Music. Hal ini dapat diartikan bahwa film dengan genre-genre utama dan umum memiliki pola perilisan yang lebih konsisten terhadap potensi keuntungan. Secara keseluruhan, model *hybrid* ini mampu menangkap hubungan kompleks antara genre film dan waktu perilisan sehingga dapat menjadi alat bantu analisis yang bermanfaat bagi pihak pembuat film.

Meskipun model hybrid CBF–CF yang diusulkan menunjukkan performa yang cukup baik, penelitian ini masih memiliki beberapa keterbatasan. Pendekatan hybrid yang digunakan mengandalkan skema penggabungan linear yaitu *weighted sum*, sehingga belum sepenuhnya mampu menangkap hubungan non-linear yang kompleks antara karakteristik konten film dan faktor temporal seperti waktu rilis. Selain itu, model *Content-Based Filtering* hanya memanfaatkan atribut genre sebagai representasi konten, sehingga informasi deskriptif lain seperti pemeran, sutradara, dan narasi film belum dieksplorasi secara optimal. Di sisi lain, *Collaborative Filtering* berbasis Singular Value Decomposition (SVD) masih berpotensi terpengaruh oleh masalah data sparsity, terutama pada kombinasi film dan bulan rilis dengan data historis yang terbatas. Faktor temporal dalam penelitian ini juga masih diperlakukan secara statis melalui bulan rilis, tanpa mempertimbangkan dinamika tren jangka panjang atau pola musiman lintas tahun.

Penelitian selanjutnya dapat difokuskan pada pengembangan model rekomendasi yang mampu menangkap hubungan non-linear dan dinamika temporal secara lebih kompleks seperti strategi promosi, tren musiman, hingga preferensi audiens berdasarkan wilayah. Lakukan pendekatan *deep learning* lebih tepatnya *Neural Collaborative Filtering* untuk meningkatkan akurasi prediksi sistem rekomendasi. Selain itu, pemanfaatan data real-time dari industri film dan pengembangan sistem berbasis web interaktif diharapkan mampu memperluas penerapan model hybrid secara efektif dalam konteks nyata di industri perfilman.

DAFTAR PUSTAKA

- [1] "Global film production hits historic high, surpassing pre-pandemic levels," global-innovation-index. Accessed: Oct. 08, 2025. [Online].
- [2] N. Wijiharjono, "Kajian perkembangan penelitian pemasaran film [Contemporary study of film marketing research]," Munich Personal RePEc Archive (MPRA), Paper No. 83349, 2017. [Online].
- [3] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems," in **Recommender Systems Handbook**, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Boston, MA, USA: Springer, 2011, ch. 1, pp. 1–35, doi: 10.1007/978-0-387-85820-3_1.
- [4] L. Chiou, "The timing of movie releases: Evidence from the home video industry," *International Journal of Industrial Organization*, vol. 26, no. 5, pp. 1059–1073, Sep. 2008, doi: 10.1016/j.ijindorg.2007.11.005.
- [5] H. Zhang, H. Yang, J. Li, and C. Miao, "Temporally-Aware Recommender Systems: A Survey and New Perspectives," *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–38, 2023.
- [6] Y. Afoudi, M. Lazaar, and M. Al Achhab, "Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network," *Simulation Modelling Practice and Theory*, vol. 113, p. 102375, Dec. 2021, doi: 10.1016/j.simpat.2021.102375.
- [7] R. Widayanti, M. H. R. Chakim, C. Lukita, U. Rahardja, and N. Lutfiani, "Improving recommender systems using hybrid techniques of collaborative filtering and content-based filtering," *Journal of Engineering Science and Technology*, vol. 58, no. 7, pp. 289–302, 2023.
- [8] T. S. Kumari and K. Sagar, "A Semantic Approach to Solve Scalability, Data Sparsity and Cold-Start Problems in Movie Recommendation Systems," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 6s, 2023. :contentReference[oaicite:3]{index=3}.
- [9] L. Xu, Z. Guan, and Y. Wu, "Enhancing Movie Recommendation Systems with Hybrid Collaborative Filtering, Content-based Filtering and SVD," in *Proc. 2nd Int. Conf. Machine Learning and Automation*, Adana, Turkey, 2024, pp. xx-xx. :contentReference[oaicite:6]{index=6}
- [10] H. Christian, M. P. Agus, and D. Suhartono, "Single Document Automatic Text Summarization using Term Frequency-Inverse Document Frequency (TF-IDF)," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 7, no. 4, p. 285, Dec. 2016, doi: 10.21512/comtech.v7i4.3746.
- [11] R. V. Imbar, A. Adelia, M. Ayub, and A. Rehatta, "Implementasi Cosine Similarity dan Algoritma Smith-Waterman untuk Mendeteksi Kemiripan Teks," *Jurnal Informatika*, vol. 10, no. 1, pp. 31–42, 2015. [Online]. Available: <https://repository.maranatha.edu/20106/>
- [12] E. Hasan, M. Rahman, C. Ding, J. X. Huang, and S. Raza, "Review-based Recommender Systems: A Survey of Approaches, Challenges and Future Perspectives," *ACM Computing Surveys*, vol. 58, no. 1, pp. 1–41, Sep. 2025, doi: 10.1145/3742421
- [13] M. A. H. Ab-Wahab, A. M. H. H. Al-Obaidi, and N. S. A. Jaleel, "Enhancing Singular Value Decomposition for Movie Recommender Systems Using User-based Clustering," *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 1, pp. 110–121, 2022.
- [14] M. Al-Ghamdi, H. Elazhary, and A. Mojahed, "Evaluation of Collaborative Filtering for Recommender Systems," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 559–565, 2021.
- [15] G. Takács, I. Pilászy, B. Németh, and D. Tikk, "Recommender systems based on quantitative implicit customer feedback," *Decis. Support Syst.*, vol. 68, pp. 77–88, Dec. 2014, doi: 10.1016/j.dss.2014.09.005.
- [16] S. Jayalakshmi, N. Ganesh, R. Čep, and J. Senthil Murugan, "Movie recommender systems: Concepts, methods, challenges, and future directions," *Sensors*, vol. 22, no. 13, p. 4904, 2022.
- [17] H. H. Kurniawan, W. S. Lukman, R. Fredyan, and M. A. Ibrahim, "Movie recommendation system: A comparison of content-based and collaborative filtering," *Procedia Computer Science*, vol. 245, pp. 860–868, 2024.