

# **Perbandingan Metode *Content-based*, *Collaborative* dan *Hybrid Filtering* pada Sistem Rekomendasi Lagu**

**KURNIA RAMADHAN PUTRA, MOHAMMAD ADITIYA RACHMAN**

Institut Teknologi Nasional Bandung  
Email: [kurniaramadhan@itenas.ac.id](mailto:kurniaramadhan@itenas.ac.id)

*Received* 28 September 2024 | *Revised* 22 Oktober 2024 | *Accepted* 23 November 2024

## **ABSTRAK**

*Sistem rekomendasi dapat dimanfaatkan untuk membantu pengguna menemukan item atau informasi sesuai preferensi mereka, termasuk lagu. Metode seperti Collaborative Filtering (CF), Content-Based Filtering (CBF), dan Hybrid Filtering digunakan untuk meningkatkan kualitas rekomendasi berdasarkan interaksi pengguna dan karakteristik konten. Penelitian ini membandingkan efektivitas ketiga metode tersebut dalam rekomendasi lagu menggunakan dataset dengan 68.330 entri data. Metode CF dan CBF diterapkan secara terpisah, lalu dikombinasikan dalam pendekatan hybrid untuk mengevaluasi peningkatan hasil. CF mencapai presisi 49.9%, CBF 39.5%, sedangkan hybrid CF-CBF mencatat presisi tertinggi sebesar 50.7%. Sebaliknya, hybrid CBF-CF menghasilkan presisi terendah, yaitu 38.4%. Kesimpulannya, pendekatan hybrid CF-CBF lebih unggul dalam merekomendasikan lagu sesuai preferensi pengguna dibandingkan metode lainnya.*

**Kata kunci:** *sistem rekomendasi, rekomendasi lagu, content-based filtering, collaborative filtering, hybrid filtering*

## **ABSTRACT**

*Recommender systems can be utilized to assist users in discovering items or information that align with their preferences, including music. Methods such as Collaborative Filtering (CF), Content-Based Filtering (CBF), and Hybrid Filtering enhance recommendation quality based on user interactions and content characteristics. This study compares the effectiveness of these three methods in music recommendation using a dataset containing 68,330 entries. CF and CBF were implemented separately and combined in a hybrid approach to evaluate performance improvements. CF achieved a precision of 49.9% and CBF 39.5%, while the hybrid CF-CBF approach recorded the highest precision at 50.7%. In contrast, the hybrid CBF-CF approach yielded the lowest precision, at 38.4%. In conclusion, the hybrid CF-CBF approach outperforms other methods in delivering music recommendations tailored to user preferences.*

**Keywords:** *recommendation system, song recommendation, content-based filtering, collaborative filtering, hybrid filtering*

## 1. PENDAHULUAN

Sistem rekomendasi lagu sangat penting dalam membantu pengguna menemukan lagu yang sesuai dengan selera mereka di *platform* musik digital seperti Spotify dan Apple Music **(Pesaru, dkk., 2022)**. Dengan banyaknya pilihan yang tersedia, pengguna sering kesulitan menemukan lagu yang tepat, sehingga sistem rekomendasi diperlukan untuk menyajikan pilihan lagu yang relevan. Ada beberapa metode utama yang digunakan dalam sistem rekomendasi, yaitu *Content-based Filtering*, *Collaborative Filtering*, dan *Hybrid Filtering* **(Bodduluri, dkk., 2024)**.

*Content-Based Filtering* fokus pada atribut atau fitur lagu yang disukai pengguna sebelumnya, seperti genre dan artis, untuk menemukan lagu yang mirip. Metode ini berguna saat data pengguna terbatas, namun cenderung memberikan rekomendasi yang kurang bervariasi **(Feng, dkk., 2021) (Jin, dkk., 2022)**. *Collaborative Filtering* bekerja dengan menganalisis preferensi pengguna lain yang serupa untuk memberikan rekomendasi lagu dengan memanfaatkan data interaksi seperti riwayat mendengarkan lagu tersebut. Metode ini efektif dalam menangkap preferensi kolektif tetapi kurang optimal untuk pengguna baru **(Muhammad, 2022) (Hwang, dkk., 2022)**. Di sisi lain *Hybrid Filtering* menggabungkan kedua metode tersebut untuk meningkatkan akurasi dan relevansi rekomendasi **(Whang, dkk., 2022)**. Dengan mengintegrasikan kekuatan dari *Collaborative* dan *Content-based Filtering*, *Hybrid Filtering* diharapkan dapat memberikan rekomendasi yang lebih kaya dan mengatasi masalah *cold start* dengan lebih efektif **(Lin, 2024)**.

Berdasarkan latar belakang ini, penelitian ini akan melakukan perbandingan antara metode *Collaborative Filtering*, *Content-based Filtering*, dan *Hybrid Filtering* dalam konteks sistem rekomendasi lagu. Tujuan utama dari penelitian ini adalah untuk menentukan metode mana yang paling efektif dalam memberikan rekomendasi lagu yang sesuai dengan preferensi pengguna, serta untuk mengeksplorasi potensi pengembangan lebih lanjut dalam sistem rekomendasi lagu **(Zhang, dkk., 2022) (Jaison, dkk., 2023)**. Untuk mengevaluasi hasil rekomendasinya menggunakan teknik *Precision* **(Krstinić, dkk., 2023)**.

## 2. METODOLOGI PENELITIAN

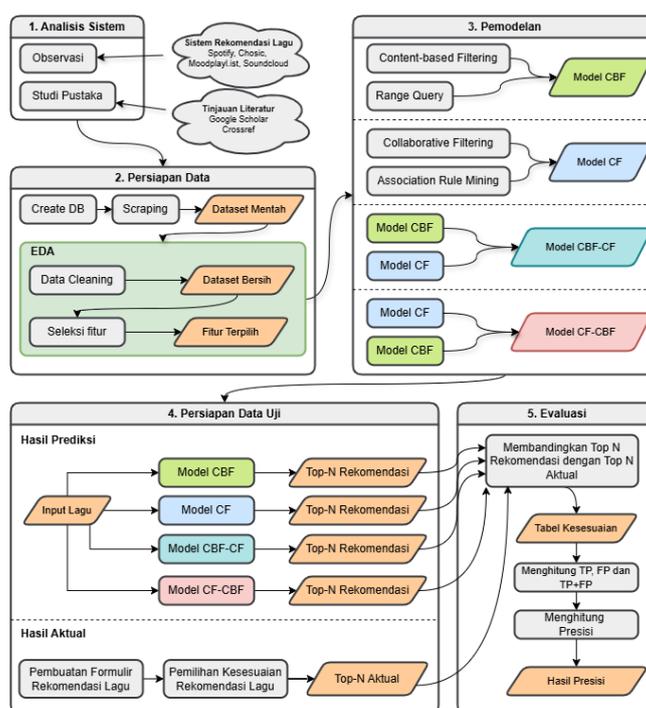
Penelitian dimulai dengan observasi yang dipadukan dengan studi pustaka untuk mengidentifikasi data yang diperlukan dalam model dan output sistem rekomendasi lagu. Selanjutnya langkah-langkah utama dari penelitian dijelaskan sebagai berikut:

- a. Penentuan Konteks Penelitian  
Sistem rekomendasi musik digunakan untuk membantu pengguna menemukan lagu-lagu yang sesuai dengan preferensi mereka. Untuk mencapai tujuan ini, tiga metode yang berbeda digunakan yaitu *Content-Based Filtering* (CBF), *Collaborative Filtering* (CF), dan *Hybrid Filtering*. Masing-masing metode memiliki pendekatan yang berbeda dalam menghasilkan rekomendasi, sehingga perlu dilakukan analisis komparatif untuk memahami kelebihan dan kekurangan dari setiap metode.
- b. Penentuan Dataset  
Penelitian ini menggunakan dataset yang terdiri dari 68.330 entri yang mencakup interaksi pengguna dengan berbagai lagu dengan kategori berdasarkan *artists*, *playlist*, dan *songs*. Data diambil dari *platform* musik untuk memastikan bahwa informasi yang digunakan representatif.
- c. Penerapan Metode Untuk Pemodelan  
Setiap metode diterapkan dengan langkah-langkah sebagai berikut:

Perbandingan Metode *Content-based*, *Collaborative* dan *Hybrid Filtering* pada Sistem Rekomendasi Lagu

- **Content-Based Filtering (CBF)**  
 Fitur Ekstraksi: menentukan karakteristik konten dari *artists*, *playlists*, dan *songs*.  
 Profil Pengguna: membangun profil pengguna berdasarkan lagu-lagu yang telah mereka dengarkan, dan merekomendasikan lagu-lagu yang memiliki kesamaan fitur.
  - **Collaborative Filtering (CF)**  
 Model Berbasis Pengguna: mengidentifikasi pengguna dengan pola interaksi yang serupa dan merekomendasikan lagu-lagu yang disukai oleh pengguna serupa.  
 Model Berbasis Item: menggunakan pola interaksi antar lagu untuk menemukan lagu-lagu yang mirip dan merekomendasikannya kepada pengguna.
  - **Hybrid Filtering**  
 Kombinasi CF-CBF: menghasilkan rekomendasi awal menggunakan CF, kemudian diperkuat dengan CBF untuk memastikan bahwa lagu-lagu yang direkomendasikan juga relevan berdasarkan karakteristik konten.  
 Kombinasi CBF-CF: menghasilkan rekomendasi awal menggunakan CBF, kemudian diperbaiki dengan CF berdasarkan interaksi pengguna dengan lagu-lagu yang direkomendasikan.
- d. Evaluasi dan Perbandingan Metode
- Pengukuran Presisi: setiap metode dievaluasi berdasarkan tingkat presisi rekomendasi yang dihasilkan, yang dihitung dengan membandingkan rekomendasi dengan preferensi aktual pengguna yang dilakukan dengan cara memberikan hasil rekomendasi kepada partisipan, kemudian partisipan memberikan umpan balik untuk selanjutnya dijabarkan pada tabel kesesuaian hasil rekomendasi.
  - Perbandingan Kinerja: hasil evaluasi dari ketiga metode dibandingkan secara sistematis untuk mengidentifikasi kelebihan dan kekurangan masing-masing metode. Analisis ini mengkaji bagaimana CF dan CBF dapat saling melengkapi dalam konteks rekomendasi, serta memahami bagaimana penggabungan kedua metode dalam pendekatan *hybrid* dapat mengatasi kelemahan masing-masing.

Langkah-langkah penelitian tersebut digambarkan pada "Gambar 1. Metodologi Penelitian".



**Gambar 1. Metodologi Penelitian**

Metodologi penelitian tersebut terdiri dari 5 tahapan utama, yaitu:

- Analisis Sistem Rekomendasi yang dilakukan melalui kegiatan observasi dan studi pustaka.
- Persiapan Data dilakukan terlebih dahulu dengan mengumpulkan data menggunakan teknik *scrapping*, selanjutnya dilakukan *data cleaning* (pembersihan data) dan seleksi fitur.
- Pemodelan dilakukan menggunakan masing-masing pendekatan yaitu CBF, CF, CBF-CF, dan CF-CBF sehingga menghasilkan top-n rekomendasi dari masing-masing model tersebut.
- Persiapan Data Uji dilakukan dengan pembuatan formulir untuk mendapatkan preferensi aktual dari pengguna.
- Evaluasi dilakukan dengan membandingkan antara hasil top-n rekomendasi dengan preferensi aktual dari pengguna, sehingga didapatkan nilai presisinya.

## 2.1. Analisis Sistem Rekomendasi

Analisis sistem dilakukan untuk memberikan pemahaman tentang hal-hal yang perlu dipersiapkan sebelum membangun model sistem rekomendasi lagu.

### 2.1.1 Observasi

Observasi dilakukan secara pengamatan secara langsung menggunakan sistem rekomendasi lagu yang ada di aplikasi *streaming* lagu, dengan cara memasukkan judul lagu dan menerima hasil dari aplikasi tersebut. Beberapa aplikasi yang digunakan yaitu Spotify, Chosic, Moodplayl.list, dan Soundcloud. "Tabel 1" menunjukkan hasil analisis kebutuhan data dan metode yang memungkinkan untuk digunakan pada rekomendasi lagu berdasarkan kegiatan observasi.

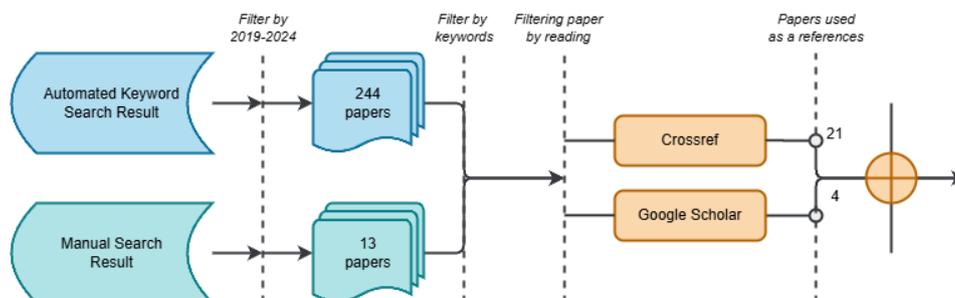
**Tabel 1. Kebutuhan Data dan Metode Rekomendasi**

No	Data	Metode Sistem Rekomendasi
1	<i>Artists</i>	CBF, CF, CBF-CF, CF-CBF
2	<i>Playlist</i>	CF, CF-CBF, CBF-CF
3	<i>Songs</i>	CBF, CBF-CF, CF-CBF

Langkah-langkah observasi ini memberikan gambaran yang jelas mengenai data yang diperlukan metode yang dapat dimanfaatkan untuk pengembangan model menggunakan data tersebut secara optimal. Selain itu, juga dilakukan pencatatan pola-pola tertentu dari rekomendasi yang dihasilkan oleh aplikasi tersebut, sehingga model yang akan dibangun dapat dioptimalkan untuk mencapai hasil rekomendasi yang lebih relevan dan akurat.

### 2.1.2 Studi Pustaka

Studi pustaka dibantu menggunakan *tool* Harzing's Publish or Perish dan repositori jurnal lainnya untuk memperdalam pemahaman tentang metodologi sistem rekomendasi, fokus pada teknik CF dan CBF periode 2019-2024, seperti yang ditampilkan pada "Gambar 2".



**Gambar 2. Studi Pustaka dari Repositori Jurnal**

Studi pustaka dilakukan untuk menggali pengetahuan dari berbagai sumber terkait sistem rekomendasi lagu. Ini mencakup tinjauan literatur tentang metode-metode yang telah digunakan dalam pengembangan sistem rekomendasi, seperti *Collaborative Filtering* dan *Content-based Filtering*.

## 2.2. Persiapan Data

Tahap persiapan data merupakan fase penting dalam pengembangan model rekomendasi, yang bertujuan untuk memastikan kualitas dan relevansi data. Langkah-langkah utama dalam persiapan data sebagai berikut:

- Pembuatan Basis Data, dilakukan dengan merancang dan membangun struktur basis data yang efisien, termasuk desain tabel, indeks, dan metode penyimpanan.
- Pengumpulan Data, dilakukan dengan mengumpulkan data melalui teknik *scrapping* dari berbagai sumber relevan untuk mendapatkan informasi yang dibutuhkan.
- Exploratory Data Analysis* (EDA), dilakukan dengan menganalisis pola, karakteristik, dan anomali dalam data melalui eksplorasi statistik dan visualisasi untuk memastikan data siap digunakan dalam pelatihan model.

Tahapan ini memastikan data yang digunakan dalam model rekomendasi akurat, relevan, dan siap untuk menghasilkan rekomendasi yang efektif

### 2.2.1 Pembuatan Basis Data

Langkah pertama dalam persiapan data adalah pembuatan basis data yang menyimpan semua informasi yang diperlukan untuk model rekomendasi. Basis data ini dirancang untuk memuat data *artists*, *playlists*, dan *songs* yang akan digunakan dalam proses pemodelan. "Tabel 2" menunjukkan hasil pembuatan basis data.

**Tabel 2. Hasil Pembuatan Basis Data**

Kode	Tabel	Deskripsi	Kolom
TBL-01	artists	Informasi artis	artist_id, artist_name, artist_genres
TBL-02	playlists	Informasi playlist	playlist_id, playlist_creator_id, playlist_original_items, playlist_items_fetched, playlist_top_artist_ids, playlist_top_genres, playlist_items, min_acousticness, max_acousticness, min_danceability, max_danceability, min_energy, max_energy, min_instrumentalness, max_instrumentalness, min_key, max_key, min_liveness, max_liveness, min_loudness, max_loudness, min_mode, max_mode, min_speechiness, max_speechiness, min_tempo, max_tempo, min_time_signature, max_time_signature, min_valence, max_valence
TBL-03	songs	Informasi lagu	song_id, song_name, artist_ids, acousticness, danceability, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, time_signature, valence

Penentuan kriteria fitur berdasarkan dataset yang ada untuk masing-masing metode dijelaskan sebagai berikut:

- Content-Based Filtering* (CBF): dataset akan difokuskan pada fitur konten dari lagu yaitu data pada TBL-03 (*songs*), tanpa mempertimbangkan interaksi pengguna. Untuk setiap lagu, atribut seperti *loudness*, *energy*, *acousticness*, *valence*, dan digunakan sebagai dasar untuk menentukan kesamaan antara lagu-lagu.

- *Collaborative Filtering* (CF): dataset akan difokuskan pada interaksi pengguna-lagu yaitu data pada BL-02 (playlists), tanpa mempertimbangkan fitur konten dari lagu. Dalam kasus ini, "playlist\_creator\_id" (pencipta playlist) dan "playlist\_items" (lagu-lagu yang ada di dalam playlist) digunakan sebagai basis untuk membangun model interaksi pengguna-lagu.
- *Hybrid Filtering*: data interaksi dan data konten yaitu data pada BL-02 (playlists) dan TBL-03 (songs) keduanya digunakan. Metode *hybrid* menggabungkan pendekatan CF dan CBF, menggunakan interaksi pengguna (dari tabel playlist) dan konten lagu (dari tabel lagu) untuk menghasilkan rekomendasi.

## 2.2.2 Pengumpulan Data Menggunakan Teknik *Scrapping*

*Scrapping* adalah teknik otomatisasi untuk mengumpulkan data dari berbagai sumber online, termasuk Spotify, yang digunakan dalam penelitian ini untuk mengumpulkan informasi lagu dengan kategori *artists*, *playlist*, dan *songs* sebanyak 68.330 entri data. Teknik ini memungkinkan peneliti memperoleh data dalam jumlah besar dan *real-time* dengan cepat, serta memfokuskan pada analisis untuk pengumpulan data yang dapat diotomatisasi (Kasica, dkk., 2021).

## 2.2.3 *Exploratory Data Analysis* (EDA)

Setelah data dikumpulkan, selanjutnya dilakukan tahapan EDA untuk memahami karakteristik data tersebut. EDA melibatkan berbagai teknik visualisasi dan analisis statistik untuk mengidentifikasi pola, tren, dan anomali yang terjadi pada data (López-Ramírez, dkk., 2024).

*Data Cleaning* adalah salah satu teknik EDA untuk identifikasi, pengoreksian, atau penghapusan kesalahan, duplikasi, atau inkonsistensi dalam data sehingga kualitas dan keakuratannya meningkat (Wang, Y., dkk., 2020)(Wang, X., dkk., 2020). Pada proses untuk memastikan penggunaan *primary key* pada *identifier* setiap tabel sebagai tindakan preventif terjadinya duplikasi item seperti *artists*, *playlists*, dan *songs*. Hasil pembersihan data dari tahap EDA ditunjukkan pada "Tabel 3".

**Tabel 3. Hasil Pembersihan Data**

No	Data	Jumlah Data	Data Hilang	Data Duplikat
1	Artists	17.333	0	0
2	Playlists	4.039	0	0
3	Songs	46.958	0	0
<b>Total</b>			<b>68.330</b>	

Seleksi fitur adalah bagian penting dari EDA, di mana hubungan antara variabel (fitur) dalam dataset dieksplorasi untuk memahami seberapa kuat keterkaitan antara satu fitur dengan fitur lainnya (Xu, dkk., 2023). Seleksi fitur dilakukan berdasarkan korelasi positif untuk menghasilkan lima fitur utama yaitu *loudness*, *energy*, *acousticness*, *valence*, dan *danceability* yang digunakan dalam model prediksi, yang dihitung menggunakan *Pearson Correlation* pada Persamaan (1).

$$r_{x,y} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{(n \sum x_i^2 - (\sum x_i)^2)(n \sum y_i^2 - (\sum y_i)^2)}} \quad (1)$$

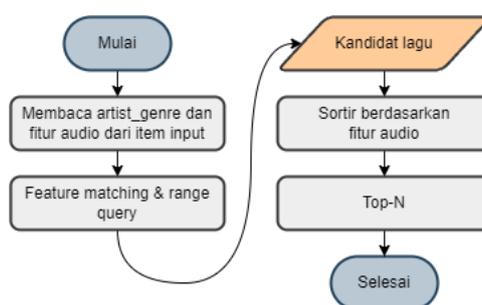
Dimana,  $r_{x,y}$  adalah korelasi antar variabel  $x$  dengan variabel  $y$ ,  $n$  adalah banyaknya sampel data,  $x_i$  adalah nilai variabel  $x$  ke- $i$ , dan  $y_i$  adalah nilai variabel  $y$  ke- $i$ . Nilai dari *Pearson Correlation* berada pada rentang -1 hingga 1, dimana nilai -1 menunjukkan korelasi yang negatif, nilai 0 menunjukkan tidak adanya korelasi, dan nilai 1 menunjukkan korelasi positif sempurna. Fitur yang telah diseleksi tersebut selanjutnya digunakan pada pemodelan.

### 2.3. Pemodelan Sistem Rekomendasi

Pemodelan sistem rekomendasi ini mencakup beberapa metode, yaitu metode tunggal seperti *Content-based Filtering* (CBF) dan *Collaborative Filtering* (CF), kemudian metode *hybrid* yaitu penggabungan metode *Content-based Filtering* diikuti dengan *Collaborative Filtering* (CBF-CF) dan *Collaborative Filtering* diikuti dengan *Content-Based Filtering* (CF-CBF).

#### 2.3.1 Pemodelan *Content-based Filtering*

Model *Content-based Filtering* (CBF) dirancang untuk memberikan rekomendasi lagu kepada pengguna berdasarkan fitur-fitur konten dari lagu yang diinputkan. Fitur-fitur ini mencakup aspek *artists*, *genre*, dan karakteristik audio dari lagu tersebut. Setiap langkah bertujuan untuk memastikan bahwa rekomendasi lagu akurat dan sesuai dengan preferensi pengguna berdasarkan data historis dari interaksi pengguna lain seperti yang ditampilkan pada "Gambar 3".



**Gambar 3. Alur Metode Content-based Filtering**

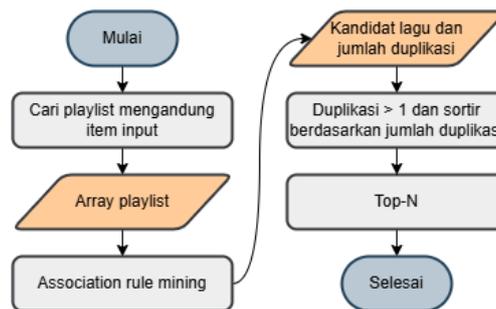
Alur kerja dari metode CBF dimulai dengan mencari *genre* dengan *feature matching* dan fitur *audio* dengan *range query* dari lagu yang diinput, dilanjutkan dengan memilih lagu yang memenuhi kriteria tertentu, dan akhirnya menyortir lagu-lagu tersebut untuk mendapatkan urutan yang paling cocok berdasarkan fitur audio. Berikut adalah contoh hasil rekomendasi yang dihasilkan dari metode *Content-based Filtering* untuk seorang pengguna seperti yang ditampilkan pada "Tabel 4".

**Tabel 4. Hasil Rekomendasi dari Metode Content-based Filtering**

Input	Hasil Rekomendasi Top-N CBF
My Chemical Romance - The Ghost of You	Nirvana - I Hate Myself And Want To Die – Demo
	Dorothy - Rest In Peace
	Thirty Seconds To Mars - The Kill
	Oasis - Champagne Supernova – Remastered
	ASIAN KUNG-FU GENERATION - 遥か彼方
Juicy Luicy - Lantas	Juicy Luicy – Lantas
	Maudy Ayunda - don't know why
	Tulus – Adaptasi
	Nick Landers - Angel of Mine
	Hivi! - Orang Ke 3
Kendrick Lamar - Not Like Us	J. Cole - Work Out
	21 Savage – monster
	Felly – Fabrics
	Childish Gambino – Heartbeat
	Migos - Narcos

### 2.3.2 Pemodelan *Collaborative Filtering*

Pada metode CF ini, teknik *Association Rule Mining* digunakan untuk memprediksi lagu yang mungkin disukai oleh pengguna berdasarkan interaksi pengguna lain saat mereka memasukkan lagu ke dalam *playlist* mereka. Metode CF dalam penelitian ini mencakup langkah-langkah utama, seperti menerima data input, mencari lagu dalam dataset *playlist*, menyimpan *playlist* relevan, dan mengekstraksi lagu-lagu terpilih. Setiap langkah bertujuan untuk memastikan bahwa rekomendasi lagu akurat dan sesuai dengan preferensi pengguna berdasarkan data historis dari interaksi pengguna lain. Pada "Gambar 4" menunjukkan alur kerja dari metode *Collaborative Filtering* dimulai dari penerimaan input, pencarian *playlist* yang relevan, hingga rekomendasi lagu yang diberikan.



**Gambar 4. Alur Metode Collaborative Filtering**

Pada Tabel 5, dijabarkan hasil rekomendasi Top-N lagu yang dihasilkan dari metode *Collaborative Filtering* untuk seorang pengguna.

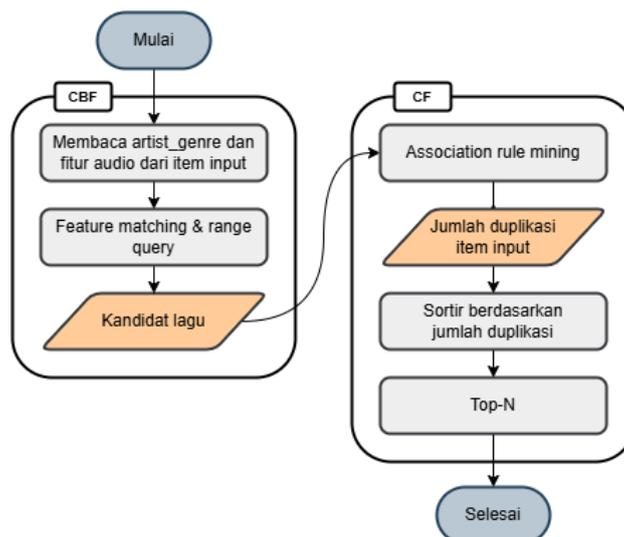
**Tabel 5. Hasil Rekomendasi dari Metode Collaborative Filtering**

Input	Hasil Rekomendasi Top-N CBF
My Chemical Romance - The Ghost of You	My Chemical Romance - Helena
	blink-182 - All The Small Things
	Green Day - American Idiot
	Fall Out Boy - Dance, Dance
	Paramore - Misery Business
Juicy Luicy - Lantas	Juicy Luicy - Tampar
	Tulus - Hati-Hati di Jalan
	Lyodra - Pesan Terakhir
	Nadin Amizah - Bertaut
	Anggi Marito - Tak Segampang Itu
Kendrick Lamar - Not Like Us	Future - Like That
	Kendrick Lamar - euphoria
	21 Savage - redrum
	Drake - Family Matters
	Travis Scott - FE!N (feat. Playboi Carti)

### 2.3.3 Pemodelan *Hybrid Filtering (CBF-CF)*

*Hybrid Filtering* dengan pendekatan CBF-CF dimulai dengan *CBF* untuk menganalisis fitur lagu dan menghasilkan rekomendasi awal. Rekomendasi ini kemudian disempurnakan dengan *CF* melalui *Association Rule Mining*, yang memperkuat hasil berdasarkan pola preferensi pengguna lain. Pendekatan ini memastikan rekomendasi relevan dan diperkuat oleh data interaksi pengguna, seperti yang ditampilkan pada "Gambar 5".

Perbandingan Metode *Content-based*, *Collaborative* dan *Hybrid Filtering* pada Sistem Rekomendasi Lagu



**Gambar 5. Alur Metode Hybrid Filtering (CBF-CF)**

Pada model CBF-CF, proses dimulai dengan pendekatan CBF, di mana sistem memfilter dan mengevaluasi lagu berdasarkan kesamaan fitur dengan preferensi pengguna, seperti genre dan artis. Setelah lagu-lagu terbaik dipilih, tahap kedua menggunakan CF untuk menghitung frekuensi kemunculan lagu-lagu tersebut dalam *playlist* pengguna lain dengan preferensi serupa. Alur kerja dari metode CBF-CF yang dimulai dengan CBF kemudian diikuti oleh CF memastikan rekomendasi lagu sesuai dengan preferensi pribadi serta didukung oleh pola perilaku pengguna lain. Dari hasil implementasi metode *hybrid* CBF-CF, hasilnya ditunjukkan pada "Tabel 6".

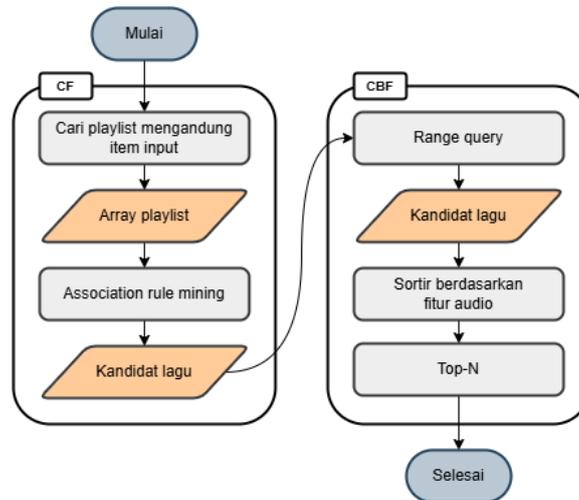
**Tabel 6. Hasil Rekomendasi dari Metode Hybrid CBF-CF**

Input	Hasil Rekomendasi Top-N Hybrid CBF-CF
My Chemical Romance - The Ghost of You	My Chemical Romance - Welcome to the Black Parade
	Yovie & Nuno - Dia Milikku
	The Neighbourhood - Flawless
	Dr. Dog - Where'd All the Time Go?
	Bon Jovi - It's My Life
Juicy Luicy - Lantas	Nadin Amizah - Beranjak Dewasa
	Lyodra - Pesan Terakhir
	Dewa - Dua Sedjoli
	Maudy Ayunda - Perahu Kertas
	Marcell - Takkan Terganti
Kendrick Lamar - Not Like Us	SZA - Saturn
	Playboi Carti - Magnolia
	Roddy Ricch - The Box
	Juice WRLD - All Girls Are The Same
	Frank Ocean - Pink + White

### 2.3.4 Pemodelan *Hybrid Filtering* (CF-CBF)

*Hybrid Filtering* dengan pendekatan linear CF-CBF menggabungkan dua pendekatan utama dalam pembuatan rekomendasi yaitu CF dan CBF. Model ini memanfaatkan kekuatan masing-masing metode untuk menghasilkan rekomendasi lagu yang lebih akurat dan relevan. CF menganalisis pola interaksi antara pengguna dan lagu menggunakan teknik *Association Rule Mining*, dengan fokus pada data historis interaksi pengguna, seperti lagu yang ada di *playlist* mereka. Sementara CBF membandingkan fitur-fitur spesifik dari lagu yang disukai pengguna

dengan lagu-lagu baru dalam basis data menggunakan operasi *Feature Matching* dan *Range Query* seperti yang ditampilkan pada Gambar 6.



**Gambar 6. Alur Metode *Hybrid Filtering* (CF-CBF)**

Alur kerja dari metode CF-CBF, dimulai dari pengumpulan data interaksi pengguna dalam CF hingga evaluasi kesamaan fitur lagu dalam CBF. Gambar 5 tersebut menunjukkan bagaimana kedua metode tersebut diintegrasikan secara bertahap untuk menghasilkan rekomendasi lagu yang tidak hanya didasarkan pada preferensi kolektif pengguna lain, tetapi juga disesuaikan dengan karakteristik unik dari lagu yang sudah disukai oleh pengguna. Dari hasil implementasi menjalankan metode *hybrid Collaborative Filtering* disertai *Content-based Filtering* (CF-CBF) hasilnya ditunjukkan pada Tabel 7.

**Tabel 7. Hasil Rekomendasi dari Metode Hybrid CF-CBF**

Input	Hasil Rekomendasi Top-N Hybrid CF-CBF
My Chemical Romance - The Ghost of You	My Chemical Romance - Helena
	Fall Out Boy - Dance, Dance
	blink-182 - All The Small Things
	Paramore - Misery Business
	All Time Low - Dear Maria, Count Me In
Juicy Luicy - Lantas	Juicy Luicy - Tampar
	Lyodra - Pesan Terakhir
	Anggi Marito - Tak Segampang Itu
	Hindia - Rumah Ke Rumah
	Hivi! - Pelangi
Kendrick Lamar - Not Like Us	Kendrick Lamar - meet the grahams
	J. Cole - Crocodile Tearz
	Joey Bada\$\$ - Love Is Only a Feeling
	Drake - First Person Shooter (feat. J. Cole)
	Rick Ross - Champagne Moments

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Hasil Evaluasi

Evaluasi sistem rekomendasi menggunakan teknik *Precision* yang mana perhitungannya menggunakan Persamaan (2).

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Dimana,

- Precision* = Menunjukkan seberapa banyak prediksi positif yang benar-benar positif dibandingkan dengan semua prediksi yang diberi label positif.
- True Positive* (TP) = Jumlah data yang benar-benar positif dan diprediksi positif oleh model.
- False Positive* (FP) = Jumlah data yang sebenarnya negatif tapi diprediksi positif oleh model.

Kuesioner dikumpulkan dan diolah menggunakan aplikasi Spreadsheet untuk dianalisis lebih lanjut. Data diorganisir dalam kolom untuk kategori TP (lagu relevan yang dipilih) dan FP (lagu relevan tetapi tidak dipilih). Pengorganisasian data ini memungkinkan perhitungan presisi sistem rekomendasi dan identifikasi kekuatan serta kelemahan sistem. Umpan balik dari responden kemudian dimasukkan ke dalam tabel kesesuaian untuk dievaluasi. Hasil analisis data kuesioner yang diperoleh dari 20 orang responden ditunjukkan pada Tabel 8.

**Tabel 8. Hasil Prediksi dan Aktual**

Responden		Model	Prediksi	Aktual		Presisi = TP / (TP+FP)
No	Nama		Top-N	TP	FP	
1	Hani Rufaidah	CF	15	8	7	8/15 = 0,53
		CBF	15	9	6	9/15 = 0,6
		CF-CBF	15	9	6	9/15 = 0,6
		CBF-CF	15	8	7	8/15 = 0,53
2	Flandi Muhammad Reza	CF	15	12	3	12/15 = 0,8
		CBF	15	10	5	10/15 = 0,67
		CF-CBF	15	10	5	10/15 = 0,67
		CBF-CF	15	9	6	9/15 = 0,6
3	Ghazi	CF	15	4	11	4/15 = 0,27
		CBF	15	3	12	3/15 = 0,2
		CF-CBF	15	9	6	9/15 = 0,6
		CBF-CF	15	4	11	4/15 = 0,27
4	Annisa Dinda Darajat	CF	15	8	7	8/15 = 0,53
		CBF	15	3	12	3/15 = 0,2
		CF-CBF	15	4	11	4/15 = 0,27
		CBF-CF	15	7	8	7/15 = 0,46
...						
20	Fauzan Addien	CF	15	10	5	10/15 = 0,67
		CBF	15	9	6	9/15 = 0,6
		CF-CBF	15	12	2	12/15 = 0,8
		CBF-CF	15	9	6	9/15 = 0,6

Selanjutnya tahap perhitungan presisi menggunakan tabel kesesuaian untuk mengukur nilai *True Actual Positive* (TAP) dan *True Predicted Positive* (TPP). Untuk menghitung rata-rata Presisi metode CBF, CF, CBF-CF, dan CF-CBF dari semua responden, menggunakan Persamaan (3).

$$\overline{Precision} = \frac{\sum_i^n Mt (R_i)}{n} \quad (3)$$

Dimana,  $\overline{Precision}$  adalah rata-rata presisi,  $i$  adalah indeks responden,  $Mt$  adalah masing-masing metode (CBF, CF, CBF-CF, dan CF-CBF),  $R$  adalah responden, dan  $n$  adalah jumlah responden sebanyak 20.

Untuk menyajikan perbandingan yang jelas antar metode, maka pada Tabel 9 merupakan tabel kesesuaian atau *confusion matrix* yang memungkinkan untuk melihat bagaimana setiap metode memprediksi rekomendasi lagu yang tepat yang mana datanya diperoleh dari Tabel 8.

**Tabel 9. Tabel Kesesuaian Prediksi**

Reponden	CBF	CF	CBF-CF	CF-CBF
1	0,6	0,53	0,53	0,6
2	0,67	0,8	0,6	0,67
3	0,2	0,27	0,27	0,6
4	0,2	0,53	0,46	0,27
...	...	...	...	...
20	0,6	0,67	0,6	0,8
Rata-Rata = $\sum Mt/n$	0.39499	0.49899	0.38413	0.50671

Hasil rata-rata presisi dari masing-masing metode disajikan pada Tabel 10.

**Tabel 9. Hasil Rata-Rata Precision dari CBF, CF, CBF-CF, dan CF-CBF**

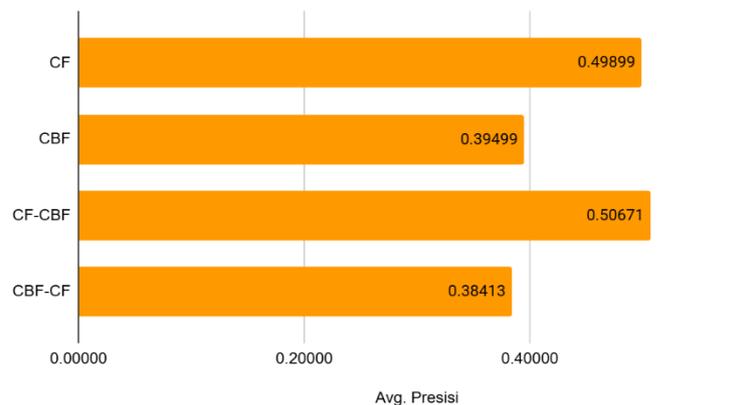
CBF	CF	CBF-CF	CF-CBF
0.39499	0.49899	0.38413	0.50671

Presisi dari berbagai metode sistem rekomendasi yang digunakan pada penelitian ini sebagai berikut:

- Metode CBF memiliki nilai presisi sebesar 0.39499 (39.5%).
- Metode CF memiliki nilai presisi sebesar 0.49899 (49.9%).
- CBF-CF memiliki nilai presisi sebesar 0.38413 (38.4%)
- CF-CBF memiliki nilai presisi sebesar 0.50671 (50.7%).

### 3.2. Pembahasan

CF-CBF memiliki presisi tertinggi, diikuti oleh CF, CBF, dan CBF-CFE. Perbandingan hasil evaluasinya divisualkan seperti yang ditunjukkan pada Gambar 7.



**Gambar 7. Alur Metode Hybrid Filtering (CBF-CF)**

Hasil perbandingan rata-rata presisi dari metode *Content-Based Filtering* (CBF), *Collaborative Filtering* (CF), serta dua pendekatan *Hybrid Filtering* (CBF-CF dan CF-CBF) menunjukkan bahwa CF-CBF memiliki presisi tertinggi di antara semua metode yang diuji, diikuti oleh CF, CBF, dan terakhir CBF-CF. Untuk memperjelas hasil ini, berikut pembahasan yang menguraikan

faktor-faktor yang mempengaruhi performa setiap metode dan bagaimana pendekatan hybrid memberikan peningkatan pada rekomendasi yang dihasilkan:

- a. CBF menghasilkan presisi yang lebih rendah (39,5%) dibandingkan CF. Hal ini disebabkan oleh pendekatan CBF yang hanya mengandalkan karakteristik konten (lagu) yang telah dipilih pengguna.
- b. Metode CF menunjukkan hasil yang cukup baik dengan nilai presisi 49,9%. CF bekerja dengan menganalisis pola preferensi dari pengguna lain yang serupa dan memberikan rekomendasi berdasarkan kesamaan interaksi antar pengguna.
- c. CF-CBF menggabungkan kekuatan dari kedua metode dengan memulai proses rekomendasi melalui pendekatan CF, lalu menyempurnakan rekomendasi berdasarkan karakteristik konten menggunakan CBF. Pendekatan ini lebih efektif karena memanfaatkan kesamaan antar pengguna terlebih dahulu dan kemudian menyaring hasil berdasarkan konten, yang memberikan hasil yang lebih personal dan relevan dengan presisi paling tinggi sebesar 50,7%.
- d. CBF-CF memulai rekomendasi dengan menganalisis karakteristik konten terlebih dahulu, lalu menggunakan CF untuk menyesuaikan rekomendasi. Pendekatan ini kurang efektif karena pengguna mungkin sudah memiliki preferensi yang jelas terhadap jenis konten tertentu, namun pendekatan ini tidak memanfaatkan kekuatan data dari pengguna lain secara optimal sejak awal dengan perolehan nilai presisi terendah (38.4%).

#### 4. KESIMPULAN

Analisis presisi menunjukkan bahwa metode hybrid CF dan CBF memberikan performa terbaik dengan presisi 0.50205, menggabungkan kekuatan kedua metode untuk menghasilkan rekomendasi lagu yang lebih relevan. Metode CF sendiri memiliki presisi 0.48727, yang cukup mendekati pendekatan hybrid, namun masih terbatas karena ketergantungan pada data interaksi dari pengguna lain. Di sisi lain, metode CBF dan kombinasi CBF-CF menunjukkan presisi yang lebih rendah, masing-masing 0.39141 dan 0.37667, yang kurang efektif dalam menangkap preferensi pengguna yang kompleks dan beragam. Hasil ini menegaskan bahwa pendekatan hybrid, dengan menggabungkan interaksi pengguna dan karakteristik lagu, mampu lebih baik dalam memahami preferensi unik pengguna, sehingga lebih akurat dalam merekomendasikan lagu yang sesuai. Dengan demikian, penelitian ini memberikan solusi terhadap masalah yang diangkat dalam latar belakang, yaitu kesulitan pengguna dalam menemukan lagu yang tepat di tengah banyaknya pilihan musik yang tersedia. Pendekatan hybrid tidak hanya meningkatkan akurasi rekomendasi, tetapi juga berpotensi meningkatkan kepuasan pengguna dengan menyediakan pengalaman personalisasi yang lebih baik.

#### DAFTAR RUJUKAN

- Pesaru, S., Sucharitha, K., Lahari, R., & Prakash, P. (2022). Music Recommendation System Using CNN Algorithm. *Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICT)*, Kannur, India, 2022, (pp. 1827-1829), doi: 10.1109/ICICT54557.2022.9917811.
- Bodduluri, K. C., Palma F., Kurti, A., Jusufi, I., & Löwenadler H. (2024). Exploring the Landscape of Hybrid Recommendation Systems in E-Commerce: A Systematic Literature Review. *IEEE Access*, 12, 28273-28296, doi: 10.1109/ACCESS.2024.3365828.

- Muhammad, M. (2022). Item Based Collaborative Filtering Based on Highest Item Similarity. *International Journal of Artificial Intelligence Research*, 6(1). <https://doi.org/10.29099/ijair.v6i1.310>.
- Hwang, S., & Park, E. (2022). Movie Recommendation Systems Using Actor-Based Matrix Computations in South Korea. *IEEE Transactions on Computational Social Systems*, 9(5), 1387-1393, doi: 10.1109/TCSS.2021.3117885.
- Feng, S., Meng, J., & Zhang, J. (2021). News Recommendation Systems in the Era of Information Overload. *Journal of Web Engineering*, 20(2), 459-470, doi: 10.13052/jwe1540-9589.20210.
- Jin, J., et al. (2022). An Agent-Based Traffic Recommendation System: Revisiting and Revising Urban Traffic Management Strategies. *IEEE Transactions on System*, 52(11), 7289-7301, Nov. 2022, doi: 10.1109/TSMC.2022.3177027.
- Wang, G., Zhang, X., Wang, H., Chu, Y., & Shao, Z. (2022). Group-Oriented Paper Recommendation With Probabilistic Matrix Factorization and Evidential Reasoning in Scientific Social Network. *IEEE Transactions on Systems*, 52(6), 3757-3771, June 2022, doi: 10.1109/TSMC.2021.3072426.
- Lin, Y. (2024). Minimizing Cold Start Time on Serverless Platforms Based on Time Series Prediction Methods. *IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT)*, (pp. 996-1001), doi: 10.1109/ICCECT60629.2024.10545789.
- Zhang, R., Tu, S., & Sun, Z. (2022). A hybrid music recommendation method based on music genes and collaborative filtering. *IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress*, (pp. 1-6), doi: 10.1109/DASC/PiCom/CBDCom/Cy55231.2022.9927924.
- Jaison C, R. R., & Rajeswari, M. (2023). Song Recommendation based on Voice Tone Analysis. *Second International Conference on Electronics and Renewable Systems (ICEARS)*, (pp. 708-712), doi: 10.1109/ICEARS56392.2023.10085001.
- Krstinić, D., Šerić, L. & Slapničar, I. (2023). Comments on MLCM: Multi-Label Confusion Matrix. *IEEE Access*, 11, 40692-40697, 2023, doi: 10.1109/ACCESS.2023.3267672.
- Kasica, S., Berret, C., & Munzner, T. (2021). Table Scraps: An Actionable Framework for Multi-Table Data Wrangling From An Artifact Study of Computational Journalism. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 957-966, doi: 10.1109/TVCG.2020.3030462.

- Wang, Y., Pan, Z., & Pan, Y. (2020). A Training Data Set Cleaning Method by Classification Ability Ranking for the  $k$ -Nearest Neighbor Classifier. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(5), 1544-1556, doi: 10.1109/TNNLS.2019.2920864.
- Wang, X., & Wang, C. (2020). Time Series Data Cleaning: A Survey. *IEEE Access*, *8*, 1866-1881, 2020, doi: 10.1109/ACCESS.2019.2962152.
- H. Xu and Y. Deng. (2018). Dependent Evidence Combination Based on Shearman Coefficient and Pearson Coefficient, *IEEE Access*, *6*, 11634-11640, doi: 10.1109/ACCESS.2017.2783320.
- López-Ramírez, G. A., Aragón-Zavala, A., & Vargas-Rosales, C. (2024). Exploratory Data Analysis for Path Loss Measurements: Unveiling Patterns and Insights Before Machine Learning. *IEEE Access*, *12*, 62279-62295, 2024, doi: 10.1109/ACCESS.2024.3394904.