

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital dalam beberapa tahun terakhir telah mengubah cara masyarakat mengakses dan menikmati musik. Konsumsi musik yang sebelumnya banyak dilakukan melalui pembelian album fisik atau unduhan digital kini semakin bergeser ke layanan *streaming*. Laporan *International Federation of the Phonographic Industry (IFPI)* menunjukkan bahwa pendapatan musik rekaman *global* terus tumbuh dan *streaming* menjadi salah satu sumber pendapatan utama industri musik modern. Pada laporan 2024, pendapatan musik rekaman *global* tahun 2023 meningkat 10,2%, sedangkan laporan berikutnya menunjukkan *streaming* tetap menjadi kontributor utama pendapatan musik *global* (IFPI, 2024). Di Indonesia, peningkatan penetrasi internet juga memperkuat pergeseran konsumsi musik ke layanan digital karena Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) melaporkan tingkat penetrasi internet Indonesia tahun 2024 mencapai 79,5% (APJII, 2024).

Salah satu platform *streaming* musik yang banyak digunakan secara global adalah *Spotify*. *Spotify* menyediakan katalog musik yang sangat besar dan dapat diakses oleh pengguna dari berbagai negara. *Spotify Newsroom* menyebutkan bahwa *Spotify* memiliki lebih dari 100 juta lagu dan beroperasi di lebih dari 184 pasar, dengan jumlah pengguna aktif bulanan yang telah melampaui 761 juta pengguna (*Spotify*, 2026). Banyaknya koleksi lagu memberi keleluasaan kepada pengguna untuk mengeksplorasi musik, tetapi juga dapat menimbulkan permasalahan berupa *choice overload*, yaitu kondisi ketika pengguna mengalami kesulitan menentukan pilihan karena terlalu banyak alternatif yang tersedia. Fenomena ini menunjukkan bahwa pilihan dalam jumlah besar dapat menarik perhatian pengguna, tetapi juga dapat membuat proses pengambilan keputusan menjadi lebih sulit dan menurunkan kepuasan terhadap pilihan (Iyengar & Lepper, 2000). konsep tersebut pada konteks layanan *streaming* berbasis langganan dan menjelaskan bahwa kelimpahan konten pada platform *streaming* dapat menyebabkan pengguna menghadapi proses pemilihan konten yang kompleks (Longo, 2021). Oleh karena itu, dalam konteks musik *streaming*, sistem

rekomendasi diperlukan untuk membantu pengguna menyaring banyaknya pilihan lagu dan menemukan lagu yang relevan secara lebih efisien.

Permasalahan tersebut menjadikan sistem rekomendasi sebagai komponen penting dalam platform musik digital. Sistem rekomendasi berfungsi membantu pengguna menemukan item yang relevan berdasarkan informasi tertentu, baik berupa data interaksi pengguna maupun karakteristik konten item. Pada domain musik, sistem rekomendasi dapat digunakan untuk menyarankan lagu, artis, album dan genre yang berpotensi disukai pengguna. Sistem rekomendasi berperan sebagai sistem penyaring informasi yang membantu pengguna menemukan item yang sesuai dari kumpulan pilihan yang besar. Dengan demikian, sistem rekomendasi musik dapat membantu proses pencarian lagu menjadi lebih efisien dan personal (Isinkaye et al., 2015).

Secara umum, sistem rekomendasi dapat dibangun dengan beberapa pendekatan, seperti *collaborative filtering*, *content-based filtering*, dan *hybrid recommender system*. *Collaborative filtering* memberikan rekomendasi berdasarkan pola interaksi pengguna atau item, sedangkan *content-based filtering* memberikan rekomendasi berdasarkan kemiripan karakteristik konten item. *Hybrid recommender system* menggabungkan dua atau lebih pendekatan untuk mengurangi kelemahan masing-masing metode (Adomavicius & Tuzhilin, 2005; Burke, 2002). Dalam konteks penelitian ini, pendekatan *content-based filtering* dipilih karena sistem rekomendasi yang dibangun berfokus pada karakteristik konten lagu, bukan pada riwayat interaksi pengguna.

Pada sistem rekomendasi musik, pendekatan yang bergantung pada data interaksi pengguna sering menghadapi tantangan *cold-start* dan data *sparsity*. *Cold-start* terjadi ketika sistem sulit memberikan rekomendasi kepada pengguna baru atau item baru karena belum tersedia data interaksi yang cukup. Sementara itu, *data sparsity* terjadi ketika interaksi pengguna terhadap item sangat sedikit dibandingkan jumlah keseluruhan item yang tersedia. *Collaborative filtering* memiliki kelemahan pada kondisi data yang jarang dan pengguna atau item baru. Oleh karena itu, pendekatan berbasis konten menjadi alternatif yang relevan ketika data interaksi pengguna belum memadai (Su & Khoshgoftaar, 2009).

Content-Based Filtering (CBF) merupakan pendekatan rekomendasi yang

menyarankan item berdasarkan kemiripan karakteristik atau konten item. sistem rekomendasi berbasis konten membangun representasi item berdasarkan atribut atau fitur yang melekat pada item tersebut, kemudian menggunakan representasi tersebut untuk menghasilkan rekomendasi. Dalam konteks musik, konten lagu dapat berupa metadata seperti Judul Lagu, Nama Artis, Nama Album dan Genre serta fitur audio seperti *danceability*, *energy*, *valence*, *tempo*, *loudness*, *acousticness* dan *instrumentalness*. Dengan pendekatan ini, sistem dapat memberikan rekomendasi lagu berdasarkan kemiripan konten tanpa bergantung pada riwayat interaksi pengguna (Lops et al., 2011).

Penelitian ini menggunakan *dataset Spotify* dari Kaggle yang berisi metadata lagu dan fitur audio dan data hasil *scrapping* dari *spotify web developer*. *Dataset* tersebut relevan untuk penerapan *CBF* karena setiap lagu memiliki atribut yang dapat direpresentasikan sebagai fitur. Metadata lagu dapat diolah sebagai fitur teks menggunakan *TF-IDF*, sedangkan fitur audio digunakan sebagai fitur numerik setelah melalui proses normalisasi. *TF-IDF* merupakan metode pembobotan kata yang digunakan untuk mengukur tingkat kepentingan suatu istilah dalam dokumen terhadap kumpulan dokumen, sedangkan normalisasi diperlukan agar fitur numerik dengan rentang berbeda berada pada skala yang sebanding (Manning et al., 2009).

Dalam penelitian ini, fitur metadata berbasis *TF-IDF* dan *audio features* yang telah dinormalisasi digabungkan menggunakan pendekatan *Feature Fusion*. *Feature Fusion* atau *early fusion* merupakan proses penggabungan beberapa fitur dari sumber berbeda ke dalam satu representasi fitur gabungan sebelum proses analisis dilakukan. *early fusion* menggabungkan fitur dari beberapa modalitas pada tahap awal sehingga model dapat memanfaatkan informasi gabungan tersebut. Dengan demikian, sistem tidak menghitung kemiripan metadata dan audio secara terpisah, melainkan menggabungkan kedua jenis fitur terlebih dahulu ke dalam satu vektor fitur gabungan (Pawłowski et al., 2023).

Setelah vektor gabungan terbentuk, sistem menghitung kemiripan antar lagu menggunakan *cosine similarity*. *Cosine similarity* mengukur kedekatan dua vektor berdasarkan nilai *cosinus* sudut di antara keduanya. *cosine similarity* banyak digunakan dalam model ruang vektor untuk menghitung kemiripan antar dokumen atau item (Manning et al., 2009). *cosine similarity* sebagai *normalized dot product*

antar-vektor (*Scikit-Learn*, 2026c). Dalam penelitian ini, lagu acuan dibandingkan dengan seluruh lagu lain dalam dataset, kemudian lagu dengan nilai *similarity* tertinggi diurutkan sebagai hasil rekomendasi *Top-N*.

Penelitian ini mengacu pada penelitian Chirasmayee yang membangun sistem rekomendasi lagu menggunakan pendekatan *content-based filtering*, *TF-IDF vectorization*, fitur dari *Spotify*, dan *cosine similarity*. Namun, penelitian ini memiliki perbedaan penting. Berbeda dengan penelitian Chirasmayee yang menggunakan *sentiment analysis* dan *Flask* sebagai *framework web*, penelitian ini tidak menggunakan analisis sentimen dan menggunakan *Streamlit* sebagai media demo sistem. Selain itu, penelitian ini difokuskan pada penggabungan metadata berbasis *TF-IDF* dan *audio features* melalui *feature fusion* untuk menghasilkan rekomendasi lagu berdasarkan kemiripan konten (Chirasmayee, 2022).

Meskipun penelitian sebelumnya telah menerapkan *content-based filtering* dalam sistem rekomendasi musik, masih terdapat ruang pengembangan pada aspek penggabungan fitur metadata dan *audio features* secara lebih terfokus. Penelitian Chirasmayee telah menggunakan *TF-IDF vectorization*, *sentiment analysis*, fitur dari *Spotify API*, normalisasi, dan *cosine similarity* dalam membangun sistem rekomendasi lagu berbasis *playlist*. Namun, penelitian tersebut lebih menekankan pada penggunaan *sentiment analysis* dan pengembangan sistem berbasis *Flask*, sehingga belum secara khusus memfokuskan pembahasan pada proses *feature fusion* antara vektor metadata berbasis *TF-IDF* dan *audio features* yang telah dinormalisasi ke dalam satu representasi fitur gabungan. Oleh karena itu, penelitian ini berupaya mengisi celah tersebut dengan menerapkan pendekatan *feature fusion* atau *early fusion* untuk menggabungkan kedua jenis fitur sebelum dilakukan perhitungan *cosine similarity*. Dengan pendekatan tersebut, penelitian ini diharapkan dapat menunjukkan bagaimana penggabungan metadata dan *audio features* dapat digunakan untuk menghasilkan rekomendasi musik berdasarkan kemiripan konten lagu, terutama ketika data interaksi pengguna tidak tersedia.

Berdasarkan uraian tersebut, penelitian ini diarahkan untuk membangun sistem rekomendasi musik berbasis *content-based filtering* yang mampu memberikan rekomendasi lagu berdasarkan kemiripan metadata dan fitur audio. Sistem yang dibangun diharapkan dapat membantu pengguna menemukan lagu yang memiliki

karakteristik serupa dengan lagu acuan, terutama ketika data interaksi pengguna tidak tersedia. Evaluasi sistem dilakukan menggunakan metrik *Precision@N*, *MAP@N*, dan *NDCG@N* melalui dua pendekatan *ground truth*, yaitu kesamaan metadata berdasarkan nama artis atau nama album, serta kesamaan karakteristik audio. Penggunaan dua pendekatan evaluasi tersebut bertujuan untuk menilai kualitas rekomendasi dari sisi kesamaan identitas musik dan kesamaan karakteristik audio lagu (Cremonesi et al., 2010; Shani & Gunawardana, 2011).

1.2 Rumusan Masalah

1.2.1 Permasalahan

Permasalahan utama dalam penelitian ini adalah banyaknya koleksi lagu pada platform *streaming* musik yang dapat menyebabkan pengguna mengalami kesulitan dalam menemukan lagu sesuai preferensi. Kondisi tersebut berkaitan dengan *choice overload*, yaitu keadaan ketika terlalu banyak pilihan justru membuat proses pengambilan keputusan menjadi lebih sulit. pilihan dalam jumlah besar dapat menarik perhatian, tetapi juga dapat menurunkan kemudahan pengambilan keputusan dan kepuasan terhadap pilihan (Iyengar & Lepper, 2000). kelimpahan konten pada layanan *streaming* dapat membuat proses pemilihan konten menjadi kompleks. Selain itu, tidak semua sistem rekomendasi dapat diterapkan secara optimal ketika data interaksi pengguna, seperti *rating*, riwayat pemutaran, atau *playlist* pengguna tidak tersedia (Longo, 2021). Kondisi tersebut berkaitan dengan permasalahan *cold-start* dan *data sparsity*, yang menjadi salah satu kelemahan pada sistem rekomendasi berbasis interaksi pengguna (Su & Khoshgoftaar, 2009). Oleh karena itu, dibutuhkan sistem rekomendasi yang dapat membantu pengguna menyaring banyaknya pilihan lagu dan menemukan lagu yang relevan secara lebih efisien.

1.2.2 Metode Usulan

Metode yang diusulkan dalam penelitian ini adalah *Content-Based Filtering (CBF)* untuk membangun sistem rekomendasi musik berdasarkan kemiripan konten lagu. Pendekatan ini dipilih karena sistem tidak bergantung pada riwayat interaksi pengguna, melainkan memanfaatkan karakteristik lagu yang tersedia dalam *dataset*. Konten lagu yang digunakan dalam penelitian ini terdiri dari dua jenis fitur, yaitu fitur teks berbasis metadata dan fitur numerik berbasis *audio features*.

Fitur teks diperoleh dari metadata lagu, seperti judul lagu, nama artis, nama album dan genre. Metadata tersebut diproses melalui tahap *text preprocessing*, seperti *cleaning text*, *case folding*, *tokenization*, dan *stopword removal*. Setelah itu, metadata diubah menjadi representasi vektor numerik menggunakan metode *TF-IDF* (*Term Frequency–Inverse Document Frequency*). Representasi ini digunakan agar informasi teks pada metadata dapat dihitung secara matematis dalam proses rekomendasi.

Sementara itu, fitur audio diperoleh dari atribut numerik yang tersedia dalam *Spotify dataset*, seperti *danceability*, *energy*, *loudness*, *acousticness*, *instrumentalness*, *valence* dan *tempo*. Karena setiap fitur audio memiliki rentang nilai yang berbeda, maka dilakukan proses normalisasi menggunakan agar seluruh fitur berada pada skala yang sebanding. Proses normalisasi diperlukan agar fitur dengan nilai besar, seperti *tempo* tidak mendominasi proses perhitungan kemiripan.

Setelah fitur metadata berbasis *TF-IDF* dan fitur audio yang telah dinormalisasi terbentuk, kedua jenis fitur tersebut digabungkan menggunakan pendekatan *feature fusion*. Pada tahap ini, vektor *TF-IDF* metadata dan vektor *audio features* digabungkan menjadi satu vektor fitur gabungan. Dengan demikian, setiap lagu direpresentasikan tidak hanya berdasarkan identitas tekstualnya, tetapi juga berdasarkan karakteristik audionya.

Kemiripan antar lagu kemudian dihitung menggunakan *cosine similarity* terhadap vektor fitur gabungan hasil *feature fusion*. Lagu acuan yang dipilih pengguna akan dibandingkan dengan seluruh lagu lain dalam *dataset*. Sistem kemudian menghitung nilai *similarity*, mengurutkan lagu berdasarkan nilai kemiripan tertinggi, dan menampilkan hasil rekomendasi dalam bentuk *Top-N recommendation*. Lagu dengan nilai *similarity* tertinggi dianggap sebagai lagu yang paling mirip dengan lagu acuan.

Pendekatan ini dipilih karena memiliki beberapa alasan berikut:

1. Tidak membutuhkan riwayat interaksi pengguna, sehingga dapat digunakan pada kondisi pengguna baru atau ketika data histori pengguna tidak tersedia.
2. Memanfaatkan dua jenis informasi lagu, yaitu metadata sebagai representasi identitas lagu dan *audio features* sebagai representasi karakteristik musikal lagu.

3. Menggunakan *feature fusion*, sehingga metadata *TF-IDF* dan *audio features* dapat digabungkan ke dalam satu representasi fitur yang lebih lengkap.
4. Menggunakan *cosine similarity*, sehingga sistem dapat menghitung tingkat kemiripan antar lagu berdasarkan vektor fitur gabungan.
5. Menghasilkan rekomendasi *Top-N*, yaitu daftar lagu dengan nilai kemiripan tertinggi terhadap lagu acuan.
6. Mudah diimplementasikan dan dievaluasi, karena metode ini dapat dibangun menggunakan *Python*, *Scikit-learn*, dan *Streamlit*, serta dievaluasi menggunakan *Precision@N*, *MAP@N*, dan *NDCG@N*.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang dan permasalahan yang telah dijelaskan, maka pertanyaan penelitian dalam Penelitian ini dirumuskan sebagai berikut:

1. Bagaimana menerapkan *feature fusion* untuk menggabungkan metadata *TF-IDF* dan *audio features*?
2. Bagaimana membangun sistem rekomendasi musik berbasis *CBF* menggunakan metadata dan *audio features*?
3. Bagaimana kinerja sistem rekomendasi musik berdasarkan metrik *Precision@N*, *MAP@N*, dan *NDCG@N* pada evaluasi *ground truth* metadata serta *ground truth* karakteristik audio?

1.3 Tujuan & Manfaat

1.3.1 Tujuan Penelitian

Tujuan dalam penelitian ini adalah untuk:

1. Menerapkan *feature fusion* untuk menggabungkan metadata *TF-IDF* dan *audio features*.
2. Membangun sistem rekomendasi musik berbasis *CBF* menggunakan metadata dan *audio Features*.
3. Mengevaluasi kinerja sistem rekomendasi menggunakan *Precision@N*, *MAP@N*, dan *NDCG@N* berdasarkan dua pendekatan *ground truth*, yaitu kesamaan metadata dan kesamaan karakteristik audio

1.3.2 Manfaat

a. Manfaat Teoritis

1. Memberikan kontribusi kajian akademik mengenai penerapan metode

Content-Based Filtering (CBF) pada sistem rekomendasi musik berbasis konten dengan memanfaatkan *dataset Spotify*.

2. Menambah referensi penelitian terkait penggunaan *TF-IDF* dalam merepresentasikan metadata lagu, seperti judul lagu, nama artis, nama album dan genre sebagai fitur teks dalam sistem rekomendasi musik.
3. Memberikan pemahaman mengenai pemanfaatan *audio features Spotify*, seperti *danceability*, *energy*, *valence*, *tempo*, *loudness*, *acousticness*, *instrumentalness* sebagai dasar representasi karakteristik musikal lagu.
4. Memberikan kontribusi terhadap penerapan *feature fusion* dalam sistem rekomendasi musik, yaitu penggabungan fitur teks berbasis metadata *TF-IDF* dan fitur audio yang telah dinormalisasi ke dalam satu vektor fitur gabungan.
5. Menjadi referensi akademik mengenai evaluasi sistem rekomendasi musik berbasis *Top-N recommendation* menggunakan metrik *Precision@N*, *MAP@N*, dan *NDCG@N* dengan pendekatan *ground truth* berbasis kesamaan metadata dan kesamaan karakteristik audio.

b. Manfaat Praktis

1. Menghasilkan prototipe sistem rekomendasi musik yang mampu memberikan rekomendasi lagu berdasarkan kemiripan konten lagu tanpa bergantung pada riwayat interaksi pengguna.
2. Membantu pengguna menemukan lagu yang memiliki karakteristik serupa dengan lagu acuan, baik dari sisi metadata maupun fitur audio, sehingga proses pencarian lagu menjadi lebih efisien.
3. Memberikan alternatif solusi terhadap permasalahan *choice overload* pada layanan *streaming* musik, yaitu kondisi ketika pengguna kesulitan memilih lagu karena terlalu banyak pilihan yang tersedia.
4. Memberikan acuan bagi pengembang atau peneliti dalam membangun sistem rekomendasi musik berbasis konten dengan menggabungkan metadata *TF-IDF* dan *audio features* menggunakan pendekatan *feature fusion*.
5. Menghasilkan demo sistem rekomendasi musik berbasis *Streamlit* yang dapat digunakan untuk menampilkan input lagu acuan dan daftar

rekomendasi lagu *Top-N* secara interaktif.

1.4 Batasan-batasan

Agar penelitian lebih terarah dan sesuai dengan ruang lingkup yang telah ditentukan, batasan penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini diperoleh dari *dataset* publik pada platform *Kaggle*, yaitu *dataset Audio Features and Lyrics of Spotify Songs* oleh Muhammad Nakhaee. *Dataset* tersebut digunakan sebagai sumber data utama dalam proses pembangunan dan pengujian sistem rekomendasi musik dan hanya menggunakan musik berbahasa inggris saja. Data bisa diakses melalui [tautan berikut ini](https://www.kaggle.com/search?q=Audio+features+and+lyrics+of+Spotify+songs).
<https://www.kaggle.com/search?q=Audio+features+and+lyrics+of+Spotify+songs>
2. Penelitian ini hanya menggunakan pendekatan *Content-Based Filtering (CBF)*. Oleh karena itu, penelitian ini tidak membahas dan tidak mengimplementasikan metode *Collaborative Filtering*, *Hybrid Recommender System*, maupun pendekatan rekomendasi berbasis interaksi pengguna secara penuh.
3. Sistem rekomendasi yang dibangun menggunakan dua jenis fitur utama, yaitu fitur teks berbasis metadata dan fitur numerik berbasis *audio features*. Fitur teks yang digunakan dibatasi pada metadata tertentu, seperti judul lagu, nama artis, nama album dan genre. Sementara itu, fitur audio yang digunakan dibatasi pada atribut numerik yang tersedia dalam *dataset*, seperti *danceability*, *energy*, *valence*, *tempo*, *loudness*, *acousticness* dan *instrumentalness*.
4. Representasi fitur teks dibentuk menggunakan metode *TF-IDF (Term Frequency-Inverse Document Frequency)*. Penelitian ini tidak menggunakan representasi teks berbasis *embedding* seperti *Word2Vec*, *Doc2Vec*, *FastText*, *BERT*, maupun model bahasa berbasis *deep learning* lainnya.
5. Penggabungan fitur teks dan fitur audio dilakukan menggunakan pendekatan *feature fusion*, yaitu menggabungkan vektor metadata berbasis TF-IDF dan vektor *audio features* yang telah dinormalisasi menjadi satu

- vektor fitur gabungan. Penelitian ini tidak menggunakan *fusion similarity* atau penggabungan skor *similarity* secara terpisah menggunakan bobot α .
6. Sistem rekomendasi yang dibangun bersifat *item-to-item recommendation*, yaitu sistem menghasilkan rekomendasi lagu berdasarkan kemiripan terhadap satu lagu acuan yang dipilih pengguna. Penelitian ini tidak membangun sistem rekomendasi personal jangka panjang berbasis profil pengguna atau riwayat pemutaran pengguna.
 7. Evaluasi sistem rekomendasi dilakukan secara *offline* menggunakan metrik *Precision@N*, *MAP@N*, dan *NDCG@N*. *Ground truth* dalam penelitian ini dibagi menjadi dua pendekatan, yaitu *ground truth* berbasis *metadata* dan *ground truth* berbasis karakteristik audio. *Ground truth* berbasis *metadata* ditentukan berdasarkan kesamaan nama artis atau nama album antara lagu acuan dan lagu hasil rekomendasi. Sementara itu, *ground truth* berbasis audio ditentukan berdasarkan kemiripan karakteristik audio lagu. Kedua pendekatan ini digunakan sebagai *proxy* relevansi karena *dataset* tidak menyediakan label relevansi eksplisit seperti *rating* pengguna, riwayat pemutaran, atau penilaian langsung dari pengguna.
 8. Penelitian ini tidak menggunakan analisis sentimen, meskipun jurnal acuan utama menggunakan *sentiment analysis* pada nama lagu. Penelitian ini hanya berfokus pada *metadata* lagu dan *audio features*.
 9. Penelitian ini tidak melakukan pengujian langsung kepada pengguna dalam bentuk real-world user study atau survei kepuasan pengguna. Evaluasi hanya dilakukan secara kuantitatif berdasarkan metrik evaluasi sistem rekomendasi.

1.5 Sistematika Proposal

Untuk memberikan gambaran yang sistematis dan terstruktur mengenai arah penelitian ini, proposal skripsi disusun dalam beberapa bab utama sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang penelitian, perumusan masalah, metode usulan, pertanyaan penelitian, tujuan penelitian, manfaat penelitian, batasan penelitian, serta sistematika penulisan. Pada bagian latar belakang dijelaskan permasalahan banyaknya pilihan lagu pada layanan *streaming* musik yang dapat

menyebabkan pengguna mengalami kesulitan dalam menemukan lagu yang sesuai dengan preferensinya. Selain itu, bab ini juga menjelaskan alasan pemilihan metode *Content-Based Filtering* dengan memanfaatkan metadata lagu dan *audio features* sebagai dasar dalam membangun sistem rekomendasi musik.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori-teori yang digunakan sebagai dasar dalam penelitian. Pembahasan meliputi sistem rekomendasi, sistem rekomendasi musik, *Content-Based Filtering*, metadata lagu, *Spotify audio features*, *text preprocessing*, *TF-IDF*, normalisasi dan *feature scaling*, *feature fusion*, *cosine similarity*, *Top-N recommendation*, evaluasi sistem rekomendasi, *Python*, *Scikit-learn*, dan *Streamlit*. Selain itu, bab ini juga memuat penelitian terdahulu yang relevan sebagai acuan dan pembandingan dalam pengembangan sistem rekomendasi musik berbasis konten.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan penelitian yang dilakukan secara teknis dan sistematis. Pembahasan meliputi jenis dan sumber data, tahapan pengumpulan data, *preprocessing* metadata lagu, pembentukan fitur teks menggunakan *TF-IDF*, normalisasi *audio features*, penggabungan fitur menggunakan *feature fusion*, perhitungan kemiripan menggunakan *cosine similarity*, serta proses pembentukan rekomendasi *Top-N*. Bab ini juga menjelaskan rancangan implementasi sistem menggunakan *Streamlit* serta skenario evaluasi sistem menggunakan metrik *Precision@N*, *MAP@N*, dan *NDCG@N*. Evaluasi dilakukan berdasarkan dua pendekatan *ground truth*, yaitu *ground truth* berdasarkan kesamaan metadata lagu dan *ground truth* berdasarkan kemiripan karakteristik audio.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi dan evaluasi sistem rekomendasi musik yang telah dibangun berdasarkan metode yang dijelaskan pada Bab III. Pembahasan meliputi hasil pengumpulan data, hasil *preprocessing* metadata lagu, hasil pembentukan vektor *TF-IDF*, hasil normalisasi *audio features*, hasil penggabungan fitur melalui *feature fusion*, serta hasil perhitungan *cosine similarity*. Selain itu, bab ini juga menampilkan hasil rekomendasi *Top-N* berdasarkan lagu acuan yang dipilih oleh pengguna. Hasil evaluasi menggunakan *Precision@N*, *MAP@N*, dan *NDCG@N* dibahas untuk mengetahui tingkat relevansi rekomendasi yang

dihasilkan oleh sistem. Pada bagian ini, hasil penelitian dianalisis untuk mengetahui sejauh mana sistem rekomendasi musik berbasis *Content-Based Filtering* dengan penggabungan fitur metadata dan *audio features* mampu menghasilkan rekomendasi lagu yang relevan.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran dari hasil penelitian yang telah dilakukan. Kesimpulan memuat jawaban atas pertanyaan penelitian dan tujuan penelitian, terutama terkait penerapan metode *Content-Based Filtering*, penggunaan *feature fusion*, serta kinerja sistem rekomendasi berdasarkan hasil evaluasi. Sementara itu, bagian saran berisi rekomendasi untuk pengembangan penelitian selanjutnya, seperti penggunaan atribut tambahan, perluasan dataset, penerapan metode lain, serta evaluasi langsung kepada pengguna untuk menilai kualitas rekomendasi secara lebih luas.