

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah memberikan perubahan yang signifikan dalam berbagai bidang kehidupan, salah satunya pada industri musik. Perkembangan teknologi digital telah mengubah cara masyarakat memperoleh dan menikmati lagu dari media fisik menjadi layanan musik digital yang dapat diakses dengan lebih mudah dan cepat. Keberadaan platform musik digital memungkinkan masyarakat mengakses banyak lagu dengan beragam judul, artis, dan genre. Platform musik digital saat ini juga telah menyediakan fasilitas yang memudahkan pengguna dalam mengakses dan mencari lagu yang diinginkan. Meskipun demikian, banyaknya lagu yang tersedia serta beragamnya karakteristik setiap lagu dapat menimbulkan kesulitan ketika pengguna ingin menemukan lagu lain yang relevan dengan lagu tertentu. Relevansi yang dimaksud dalam penelitian ini mengarah pada lagu yang memiliki kemiripan karakteristik dengan lagu yang dijadikan acuan. Permasalahan muncul ketika pendengar telah memiliki sebuah lagu sebagai acuan, tetapi belum mengetahui lagu lain yang mempunyai karakteristik serupa sehingga informasi seperti judul atau artis dari lagu lain yang akan dicari juga belum diketahui.

Kondisi tersebut berkaitan dengan *Information Retrieval* (IR), yaitu proses memperoleh informasi yang sesuai dengan kebutuhan dari sekumpulan informasi yang tersedia (Murtopo dkk., 2022). Dalam pencarian lagu, pengguna umumnya memasukkan kata kunci seperti judul lagu atau nama artis untuk menemukan lagu yang telah diketahui, cara tersebut dapat digunakan ketika pengguna sudah mengetahui informasi mengenai lagu yang ingin dicari. Namun, ketika pengguna ingin menemukan lagu lain yang memiliki karakteristik serupa dengan lagu acuan, pencarian menggunakan kata kunci menjadi kurang praktis karena informasi mengenai lagu lain yang akan dicari belum diketahui. Oleh karena itu, dibutuhkan suatu mekanisme yang tidak hanya membantu pengguna menemukan lagu tertentu, tetapi juga dapat memberikan pilihan lagu lain yang memiliki keterkaitan dengan lagu acuan. Salah satu mekanisme yang dapat digunakan untuk memenuhi

kebutuhan tersebut adalah sistem rekomendasi, yang merupakan salah satu penerapan *Information Retrieval* dalam menyediakan informasi yang relevan sesuai dengan kebutuhan pengguna.

Sistem rekomendasi merupakan suatu sistem yang membantu pengguna dalam menemukan dan memilih item yang sesuai dengan preferensi, minat, atau kebutuhannya dari berbagai pilihan yang tersedia. (Saifudin & Widiyaningtyas, 2024) Secara umum, sistem rekomendasi dapat dibangun menggunakan beberapa pendekatan, antara lain *Collaborative Filtering*, *Content-Based Filtering*, dan *Hybrid Filtering*. *Collaborative Filtering* memanfaatkan pola interaksi atau preferensi pengguna terhadap suatu item, sedangkan *Content-Based Filtering* memanfaatkan karakteristik yang dimiliki oleh item untuk menghasilkan rekomendasi. Adapun *Hybrid Filtering* menggabungkan lebih dari satu pendekatan dalam proses rekomendasi (Widiyaningtyas dkk., 2021). Pada data lagu yang memiliki informasi berupa karakteristik audio, pendekatan berbasis konten dapat digunakan karena rekomendasi dapat ditentukan berdasarkan karakteristik yang dimiliki oleh lagu itu sendiri. Selain penggunaan pendekatan rekomendasi, kandidat lagu yang akan diproses juga dapat dikelompokkan terlebih dahulu ke dalam kategori tertentu melalui proses klasifikasi sehingga proses rekomendasi dapat dilakukan pada kelompok lagu yang lebih sesuai. Penerapan klasifikasi dan sistem rekomendasi pada data musik telah dilakukan dalam beberapa penelitian sebelumnya dengan metode dan tujuan yang berbeda, termasuk pemanfaatan *Random Forest* sebagai salah satu metode klasifikasi.

Penelitian Alvina & Yamasari (2025) menggunakan *Random Forest* dan *Decision Tree* pada rekomendasi lagu berbasis genre dengan memanfaatkan fitur audio serta memperoleh rata-rata akurasi sebesar 0,8301 pada *Random Forest* dan 0,7965 pada *Decision Tree*. Hasil tersebut menunjukkan kemampuan *Random Forest* dalam melakukan klasifikasi genre berdasarkan karakteristik audio lagu. Sementara itu, Nasution (2024) menerapkan *Random Forest* pada sistem rekomendasi musik menggunakan pendekatan *Collaborative Filtering* dan memperoleh nilai *Mean Absolute Error* (MAE) sebesar 0,56. Kedua penelitian tersebut menunjukkan bahwa *Random Forest* telah digunakan dalam klasifikasi genre maupun dalam sistem rekomendasi, tetapi dengan fungsi dan pendekatan

yang berbeda. Pada penelitian Alvina & Yamasari (2025), *Random Forest* digunakan untuk menentukan genre lagu, sedangkan Nasution (2024) menerapkannya pada sistem rekomendasi berbasis *Collaborative Filtering* yang memanfaatkan data interaksi pengguna. Dengan demikian, masih terdapat ruang penelitian untuk memanfaatkan hasil klasifikasi genre menggunakan *Random Forest* sebagai tahap penentuan kelompok kandidat sebelum proses rekomendasi dilakukan berdasarkan karakteristik lagu itu sendiri.

Berdasarkan permasalahan dan penelitian terdahulu tersebut, penelitian ini menggunakan *Random Forest Classifier* dan *Content-Based Filtering* sebagai pendekatan yang saling berkaitan dalam menghasilkan rekomendasi lagu. *Random Forest Classifier* digunakan untuk menentukan genre lagu berdasarkan karakteristik audio sehingga diperoleh kelompok lagu yang menjadi kandidat dalam proses rekomendasi. Hasil klasifikasi tersebut kemudian digunakan pada *Content-Based Filtering* untuk menghasilkan rekomendasi berdasarkan kemiripan karakteristik lagu pada kelompok genre yang telah ditentukan. Dengan demikian, *Random Forest* berperan dalam mengarahkan kandidat lagu berdasarkan genre, sedangkan *Content-Based Filtering* berperan dalam menentukan lagu yang relevan berdasarkan karakteristik lagu. Berdasarkan uraian tersebut, penelitian ini mengambil judul "Penerapan Algoritma *Random Forest Classifier* pada Sistem Rekomendasi Lagu Berbasis *Content-Based Filtering*".

1.2 Rumusan Masalah

Dari penjelasan diatas, maka diperoleh beberapa rumusan masalah sebagai berikut:

1. Bagaimana algoritma *Random Forest Classifier* dan metode *Content-Based Filtering* digunakan dalam sistem rekomendasi lagu?
2. Bagaimana hasil evaluasi sistem rekomendasi lagu berdasarkan nilai *Precision* dan *Recall* dalam mengukur kualitas rekomendasi?

1.3 Tujuan Penelitian

Berikut ini adalah tujuan dari penelitian:

1. Untuk membangun sistem rekomendasi lagu dengan menerapkan algoritma *Random Forest Classifier* untuk mengklasifikasikan genre lagu berdasarkan karakteristik audio, serta menerapkan metode *Content-Based Filtering* untuk menghasilkan rekomendasi lagu berdasarkan kemiripan karakteristik pada kelompok genre yang dihasilkan.
2. Mengevaluasi kualitas dari sistem rekomendasi lagu yang telah dibangun berdasarkan hasil pengujian menggunakan metrik *Precision* dan *Recall*.

1.4 Manfaat Penelitian

Diharapkan penelitian Tugas Akhir ini dapat memberi manfaat antara lain bagi:

1. Hasil penelitian ini diharapkan dapat menjadi bahan pertimbangan dalam pengembangan dan peningkatan fitur rekomendasi lagu pada layanan musik digital agar rekomendasi yang diberikan lebih sesuai dengan karakteristik lagu.
2. Penelitian ini diharapkan dapat menambah wawasan dan pemahaman mengenai penerapan *Random Forest Classifier* untuk klasifikasi genre serta *Content-Based Filtering* dalam pengembangan sistem rekomendasi lagu.

1.5 Batasan Penelitian

Batasan dari penelitian ini mencakup:

1. Penelitian ini memanfaatkan data sekunder yang diperoleh dari *dataset* Spotify yang dipublikasikan secara terbuka melalui *platform* Kaggle pada tahun 2022. Dataset tersebut berjumlah 114.000 data lagu.
2. Atribut audio yang digunakan meliputi *danceability*, *energy*, *loudness*, *speechiness*, *acousticness*, *instrumentalness*, *liveness*, *valence*, dan *tempo* sebagai variabel input dalam proses klasifikasi genre menggunakan algoritma *Random Forest Classifier* serta sebagai dasar perhitungan kemiripan lagu pada sistem rekomendasi berbasis *Content-Based Filtering*.

3. Proses pengklasifikasian genre lagu dilakukan menggunakan algoritma *Random Forest Classifier*, sedangkan proses rekomendasi lagu dilakukan menggunakan pendekatan *Content-Based Filtering*.
4. Evaluasi sistem dibatasi pada dua tahap, yaitu pengujian performa klasifikasi *Random Forest* menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, serta pengujian performa hasil rekomendasi menggunakan metrik *Precision* dan *Recall*.

