

BAB I

PENDAHULUAN

1.1. Latar Belakang

Di era konten media yang melimpah, sistem rekomendasi yang dipersonalisasi menjadi semakin penting dalam membantu pengguna menemukan film yang sesuai dan menarik perhatian mereka [1]. Dengan semakin meningkatnya jumlah platform digital dan layanan streaming, pengguna dihadapkan pada ribuan pilihan film yang tersedia setiap saat. Situs IMDB mencatat lebih dari 690 ribu judul film dengan berbagai genre telah diproduksi hingga tahun 2024 (IMDB, 2024), yang mencerminkan laju produksi konten yang sangat tinggi. Kondisi ini menyebabkan fenomena informasi berlebih (*information overload*), yaitu keadaan dimana pengguna kesulitan menemukan film yang sesuai dengan preferensi mereka secara efisien karena terlalu banyak pilihan yang tersedia. Ketidakseimbangan antara jumlah konten dan kemampuan pengguna dalam melakukan seleksi membuat pengalaman pengguna menjadi kurang optimal. Untuk mengatasi permasalahan tersebut, dibutuhkan suatu sistem rekomendasi yang mampu menyaring dan menyajikan informasi secara personal dan relevan, dengan mempertimbangkan preferensi individu pengguna serta dinamika tren film yang terus berubah dari waktu ke waktu.

Sistem rekomendasi merupakan teknik perangkat lunak yang dirancang untuk memberikan saran *item* atau produk yang relevan kepada pengguna berdasarkan preferensi atau perilaku pengguna lainnya [2]. Teknologi ini memainkan peran krusial dalam meningkatkan kepuasan pengguna dengan menyajikan konten yang sesuai secara otomatis tanpa perlu pencarian manual. Sistem rekomendasi banyak diimplementasikan dalam berbagai platform, misalnya pada *e-commerce* untuk merekomendasikan produk, media sosial untuk menyarankan teman atau konten, layanan musik digital seperti *Spotify* untuk menyarankan lagu, dan aplikasi pemutar film seperti *Netflix* atau *Disney+* untuk menyajikan film yang relevan. Peran sistem rekomendasi dalam mempermudah proses pengambilan keputusan serta meningkatkan keterlibatan pengguna menjadikannya sebagai komponen penting dalam pengembangan sistem berbasis data.

Namun, sistem rekomendasi juga menghadapi beberapa tantangan teknis yang cukup signifikan, yaitu masalah *cold start* dan *sparsity* [3]. Masalah *cold start* terjadi

ketika sistem harus memberikan rekomendasi kepada penonton atau pengguna baru yang belum memiliki data historis, sehingga sistem tidak memiliki informasi yang cukup untuk melakukan prediksi. Hal serupa juga terjadi pada *item* baru, seperti film yang baru dirilis, di mana belum tersedia interaksi pengguna sebelumnya. Sementara itu, *sparsity* mengacu pada kondisi kurangnya data penilaian karena sebagian besar pengguna hanya memberikan sedikit rating, sehingga matriks interaksi pengguna-item menjadi jarang terisi. Kedua masalah ini menyebabkan akurasi dan jangkauan sistem rekomendasi menjadi terbatas karena minimnya informasi yang dapat dijadikan dasar dalam pengambilan keputusan. Oleh karena itu, diperlukan pendekatan tambahan yang mampu memperkaya informasi dan memperbaiki kinerja sistem dalam kondisi data yang minim.

Berbagai pendekatan telah dikembangkan untuk meningkatkan kinerja sistem rekomendasi, termasuk pemanfaatan algoritma optimasi sebagai salah satu teknik untuk meningkatkan performa dari sistem rekomendasi. Salah satu algoritma yang umum digunakan dalam bidang ini adalah Algoritma Genetika (AG), yang meniru proses evolusi biologis dalam mencari solusi optimal. AG telah banyak dimanfaatkan dalam proses pencarian nilai *similarity* antar pengguna maupun *item*, serta dalam pencarian parameter optimal seperti *centroid* pada algoritma *clustering* [4]. Namun, meskipun AG cukup efektif, algoritma ini memiliki kelemahan dalam hal waktu komputasi, terutama ketika diterapkan pada dataset berskala besar yang memerlukan banyak iterasi dan populasi solusi. Oleh karena itu, dibutuhkan alternatif algoritma optimasi yang mampu memberikan hasil yang lebih cepat dan efisien.

Sebagai alternatif, Harmony Search (HS) muncul sebagai algoritma optimasi yang lebih efisien dan adaptif dalam berbagai jenis permasalahan optimasi [5]. Terinspirasi dari proses improvisasi musik dalam menciptakan harmoni yang indah [6], HS menggunakan prinsip pencarian solusi yang menyerupai proses pencarian nada musik terbaik oleh sekelompok musisi. Algoritma ini menawarkan keunggulan dalam kecepatan pencarian solusi optimal tanpa membutuhkan waktu yang lama, serta memiliki kemampuan eksplorasi dan eksplorasi yang seimbang [5]. Dalam konteks *clustering*, HS dapat digunakan untuk menentukan posisi *centroid* secara lebih optimal, yaitu menghindari inisialisasi acak yang sering menjadi kelemahan pada metode Fuzzy C-Means (FCM) [7]. Dengan mengoptimalkan posisi awal *centroid*, HS mampu meningkatkan stabilitas hasil clustering serta mengurangi kemungkinan konvergensi terhadap solusi lokal yang kurang representatif terhadap struktur data.

Clustering adalah metode *unsupervised learning* yang digunakan untuk mengelompokkan data ke dalam beberapa *cluster* berdasarkan kemiripan fitur. Teknik ini bertujuan untuk menemukan struktur alami dalam data tanpa adanya label atau target keluaran yang diketahui. Pada data yang kompleks atau bersifat kategorikal, metode seperti Fuzzy C-Means (FCM) lebih disukai karena memungkinkan satu data menjadi anggota beberapa *cluster* dengan derajat keanggotaan tertentu [8]. Hal ini tentu sangat sesuai dalam sistem rekomendasi, di mana batas antar kelompok pengguna atau item sering kali tidak jelas dan dapat tumpang tindih. FCM memungkinkan representasi data yang lebih fleksibel dan realistik, terutama ketika preferensi pengguna bersifat majemuk dan tidak bisa diklasifikasikan secara tegas ke dalam satu kelompok saja.

Untuk menghasilkan rekomendasi yang akurat, metode *filtering* seperti collaborative filtering, content-based filtering, dan context-based filtering umum digunakan [3]. Masing-masing metode memiliki pendekatan yang berbeda dalam menyusun rekomendasi, namun juga memiliki keterbatasan masing-masing. Misalnya, collaborative filtering yang sangat bergantung pada interaksi pengguna akan kesulitan menangani kasus *cold start*, sedangkan content-based filtering yang menggunakan atribut konten terkadang gagal menangkap preferensi implisit atau kesamaan tidak langsung antar *item*. Oleh karena itu, pendekatan hybrid filtering dapat diadopsi dengan menggabungkan kelebihan beberapa metode *filtering*, yang bertujuan untuk meningkatkan akurasi dan relevansi hasil rekomendasi [9]. Hybrid filtering mampu menggabungkan analisis perilaku pengguna dengan informasi konten secara menyeluruh, sehingga dapat memberikan hasil rekomendasi yang lebih menyeluruh dan presisi.

Berdasarkan uraian tersebut, pada usulan penelitian ini akan dibangun sistem rekomendasi film berbasis hybrid filtering yang menggabungkan collaborative filtering dan content-based filtering. Collaborative filtering menganalisis kesamaan antar pengguna atau antar *item* berdasarkan histori interaksi pengguna, sedangkan content-based filtering memanfaatkan atribut pada konten film, seperti genre atau sinopsis, dalam menyusun rekomendasi yang sesuai dengan preferensi pengguna. Di sisi lain, proses clustering terhadap data pengguna dan film dilakukan menggunakan Fuzzy C-Means (FCM). Penggunaan FCM memungkinkan fleksibilitas dalam pengelompokan data dengan derajat keanggotaan majemuk. Guna mengatasi kelemahan FCM pada inisialisasi *centroid* yang bersifat acak dan rentan terhadap lokal

optima, algoritma Harmony Search (HS) diterapkan sebagai mekanisme optimasi dalam penentuan posisi *centroid* secara lebih optimal dan efisien [10]. Integrasi teknik hybrid filtering, Fuzzy C-Means, dan Harmony Search ini, diharapkan dapat menghasilkan sistem rekomendasi film yang tidak hanya akurat dan relevan dengan preferensi pengguna, tetapi juga mampu mengatasi tantangan-tantangan dalam sistem rekomendasi seperti *cold start*, *sparsity*, dan keterbatasan efisiensi dalam menangani dataset berskala besar.

1.2. Rumusan Masalah

Rumusan masalah dalam usulan penelitian ini yaitu :

1. Bagaimana cara mengatasi masalah cold start dan sparsity pada sistem rekomendasi agar dapat menghasilkan rekomendasi menjadi lebih akurat sesuai dengan preferensi pengguna?
2. Bagaimana pemanfaatan algoritma harmony search untuk meningkatkan performa proses clustering pada metode fuzzy c-means dalam sistem rekomendasi?
3. Bagaimana penerapan pendekatan hybrid filtering yang menggabungkan metode collaborative filtering dan content-based filtering dapat menghasilkan rekomendasi yang lebih efektif dan relevan?

1.3. Tujuan

Tujuan dari penelitian ini yaitu menerapkan algoritma Harmony Search untuk optimasi Fuzzy C-Means dalam sistem rekomendasi film berbasis Hybrid Filtering.

1.4. Manfaat

Manfaat yang akan diperoleh dari penelitian ini sebagai berikut, yaitu :

1. Membantu pengguna mendapatkan rekomendasi film yang lebih relevan dengan preferensi pengguna secara cepat dan akurat.
2. Menyediakan solusi inovatif dalam pengembangan sistem rekomendasi yang efisien untuk dataset besar dan kompleks.
3. Memberikan kontribusi ilmiah dalam penerapan algoritma Harmony Search pada sistem rekomendasi berbasis hybrid filtering.

1.5. Batasan Masalah

Agar penelitian lebih terfokus, maka ditetapkan beberapa batasan masalah sebagai berikut :

1. Sistem rekomendasi hanya difokuskan untuk domain film dan tidak mencakup jenis item lainnya misalnya musik, buku, atau produk-produk *e-commerce*.

2. Data film bersumber dari dataset publik, yaitu MovieLens, atau IMDb yang mencakup informasi seperti judul film, genre, tahun rilis dan rating pengguna pada film.
3. Penelitian ini hanya akan fokus pada pemanfaatan data yang bersifat eksplisit, seperti rating film oleh pengguna, tahun rilis film dan genre film tanpa menggunakan data tambahan seperti ulasan pengguna atau metadata lain yang lebih kompleks.
4. Penelitian ini berfokus pada pengembangan sistem rekomendasi yang dirancang untuk mengatasi permasalahan *cold start* dan *sparsity* pada data pengguna.
5. Evaluasi sistem rekomendasi dalam penelitian ini dilakukan menggunakan metrik kuantitatif, yaitu pengukuran tingkat akurasi prediksi dengan menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE), serta pengukuran kualitas rekomendasi menggunakan metrik *Precision* dan *Recall*. Evaluasi tidak mencakup uji kepuasan pengguna atau penilaian berdasarkan studi pengguna secara langsung.

Halaman ini sengaja dikosongkan