

PENERAPAN METODE COLLABORATIVE FILTERING PADA SISTEM REKOMENDASI FILM MENGGUNAKAN MOVIELENS DATASET

Sintia Rahayu Fatmawati, Sri Sumarlinda, Pipin Widyaningsih

Teknik Informatika, Universitas Duta Bangsa Surakarta

Jl. Bhayangkara, No.55, Tipes, Serengan, Surakarta

tiyasintya15@gmail.com

ABSTRAK

Perkembangan layanan streaming film digital menyebabkan jumlah konten yang tersedia semakin banyak sehingga pengguna sering mengalami kesulitan dalam menemukan film yang sesuai dengan preferensinya. Kondisi tersebut menimbulkan permasalahan information overload yang dapat mengurangi efektivitas pengguna dalam menentukan pilihan tontonan. Penelitian ini bertujuan mengembangkan sistem rekomendasi film berbasis web menggunakan metode Collaborative Filtering dengan memanfaatkan MovieLens Dataset sebagai sumber data utama. Metode pengembangan sistem yang digunakan adalah prototype yang meliputi pengumpulan dataset, preprocessing data, perhitungan tingkat kemiripan pengguna menggunakan Cosine Similarity, prediksi rating, implementasi sistem, dan pengujian fungsionalitas. Dataset yang digunakan terdiri dari data pengguna, data film, dan data rating yang tersedia pada MovieLens. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan mampu mengolah data rating pengguna untuk menghasilkan rekomendasi film berdasarkan kesamaan preferensi serta menampilkan daftar rekomendasi secara otomatis sesuai histori penilaian pengguna. Sistem yang dihasilkan dapat digunakan sebagai sarana pendukung dalam membantu pengguna menemukan film yang lebih relevan dan sesuai dengan minat mereka.

Kata kunci : Collaborative Filtering, Sistem Rekomendasi, MovieLens, Cosine Similarity, MAE, Film

1. PENDAHULUAN

Perkembangan teknologi digital telah mendorong pertumbuhan platform hiburan berbasis daring, khususnya layanan streaming film yang menyediakan ribuan pilihan konten bagi pengguna. Ketersediaan konten yang semakin beragam memberikan kemudahan akses terhadap berbagai jenis film, namun di sisi lain menimbulkan permasalahan information overload. Kondisi ini menyebabkan pengguna mengalami kesulitan dalam menemukan film yang sesuai dengan minat dan preferensinya karena harus menyeleksi sejumlah besar pilihan yang tersedia [1]. Oleh karena itu, diperlukan suatu mekanisme yang mampu membantu pengguna memperoleh rekomendasi film yang relevan secara cepat dan akurat.

Salah satu teknologi yang banyak digunakan untuk mengatasi permasalahan tersebut adalah sistem rekomendasi. Sistem rekomendasi berfungsi menyaring informasi berdasarkan pola interaksi pengguna sehingga mampu menyajikan konten yang lebih personal dan sesuai dengan kebutuhan pengguna [2].

Dalam konteks layanan streaming film, sistem rekomendasi berperan penting dalam meningkatkan pengalaman pengguna dengan menyajikan film yang memiliki tingkat relevansi tinggi berdasarkan histori aktivitas dan penilaian yang pernah diberikan pengguna sebelumnya [3].

Berbagai metode telah dikembangkan untuk membangun sistem rekomendasi, salah satunya adalah Collaborative Filtering (CF). Metode ini memanfaatkan pola kesamaan perilaku pengguna dalam memberikan penilaian terhadap suatu item

untuk menghasilkan rekomendasi yang sesuai [4]. Collaborative Filtering menjadi salah satu pendekatan yang paling banyak digunakan karena tidak memerlukan informasi atribut item secara detail dan mampu menghasilkan rekomendasi berdasarkan data historis pengguna. Penelitian terkait Collaborative Filtering banyak menggunakan MovieLens Dataset sebagai dataset pengujian karena menyediakan data pengguna, data film, dan data rating dalam jumlah besar sehingga menjadi benchmark dalam penelitian sistem rekomendasi [5].

2. TINJAUAN PUSTAKA

Penelitian sebelumnya menunjukkan bahwa Collaborative Filtering mampu menghasilkan rekomendasi yang relevan berdasarkan kesamaan preferensi pengguna. Hafeez et al. melaporkan bahwa pendekatan berbasis matrix factorization seperti Singular Value Decomposition (SVD) mampu memberikan akurasi yang lebih baik dibandingkan pendekatan User-Based Collaborative Filtering [6].

Penelitian lain menggabungkan Collaborative Filtering dengan teknik clustering dan informasi demografis untuk meningkatkan nilai precision dan recall rekomendasi yang dihasilkan [7].

Meskipun demikian, sebagian besar penelitian masih berfokus pada peningkatan performa algoritma, sementara implementasi metode tersebut ke dalam sistem berbasis web yang dapat digunakan secara langsung oleh pengguna masih relatif terbatas. Selain itu, penelitian yang mengintegrasikan proses pengembangan sistem menggunakan pendekatan prototype dengan implementasi Collaborative Filtering masih belum banyak ditemukan [8].

Berdasarkan kondisi tersebut, penelitian ini bertujuan mengembangkan sistem rekomendasi film berbasis web menggunakan metode User-Based Collaborative Filtering dengan memanfaatkan MovieLens Dataset sebagai sumber data utama. Penyelesaian permasalahan dilakukan melalui tahapan pengumpulan dataset, preprocessing data, perhitungan tingkat kemiripan pengguna menggunakan Cosine Similarity, prediksi rating, serta implementasi sistem menggunakan model pengembangan prototype. Sistem yang dihasilkan diharapkan mampu memberikan rekomendasi film berdasarkan kesamaan preferensi pengguna dan membantu pengguna dalam menentukan pilihan tontonan secara lebih efektif dan efisien. Selain itu, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan sistem rekomendasi berbasis web yang mengintegrasikan metode Collaborative Filtering dalam bentuk aplikasi yang siap digunakan oleh pengguna akhir.

3. METODE PENELITIAN

Pada bagian ini dijelaskan metode yang digunakan dalam penelitian untuk penerapan metode *collaborative filtering* pada sistem rekomendasi film menggunakan *movielens* dataset, serta model pengembangan sistem prototype. Metode penelitian ini mencakup jenis data, teknik pengumpulan data, serta tahapan penelitian yang dilakukan secara sistematis.

3.1. Dataset MovieLens

Dataset MovieLens merupakan salah satu dataset standar yang paling banyak digunakan dalam penelitian sistem rekomendasi. Dataset ini dikembangkan oleh GroupLens Research dan berisi data rating film yang diberikan oleh pengguna terhadap berbagai judul film. MovieLens tersedia dalam beberapa versi dengan ukuran data yang berbeda, mulai dari ribuan hingga jutaan rating [10].

Dalam penelitian sistem rekomendasi, MovieLens sering digunakan sebagai benchmark untuk menguji performa algoritma karena memiliki struktur data yang jelas dan representatif. Data yang tersedia umumnya mencakup atribut seperti *userId*, *movieId*, *rating*, dan *timestamp* yang mencerminkan interaksi pengguna terhadap film.

3.2. User-Based Collaborative Filtering

Proses dalam metode User-Based Collaborative Filtering dimulai dari perhitungan tingkat kemiripan antar pengguna menggunakan teknik seperti *Cosine Similarity* atau *Pearson Correlation*. Setelah tingkat kemiripan diperoleh, sistem akan menentukan sejumlah pengguna terdekat (*nearest neighbors*) yang memiliki pola rating paling mirip. Selanjutnya, sistem melakukan prediksi rating terhadap item yang belum pernah dinilai oleh pengguna target [11].

Tahap akhir dari proses ini adalah menghasilkan daftar rekomendasi berdasarkan nilai prediksi tertinggi. Item dengan nilai prediksi paling tinggi akan

direkomendasikan kepada pengguna. Metode ini banyak digunakan karena sederhana, mudah diimplementasikan, dan cukup efektif dalam dataset berukuran sedang seperti MovieLens.

3.3. Cosine Similarity

Cosine Similarity merupakan metode matematis yang digunakan untuk mengukur tingkat kemiripan antara dua vektor berdasarkan sudut yang terbentuk di antara keduanya. Dalam sistem rekomendasi, vektor tersebut merepresentasikan polarating pengguna terhadap item tertentu [12].

$$\frac{\sum_{i=1}^n x_i \cdot y_i}{\sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{i=1}^n y_i^2}} \quad (1)$$

Keterangan :

x_i = Nilai rating pengguna A pada item ke-i

y_i = Nilai rating pengguna B pada item ke-i

n = Jumlah item yang dibandingkan

Σ = Penjumlahan seluruh nilai

Nilai Cosine Similarity berada pada rentang -1 hingga 1. Nilai yang mendekati 1 menunjukkan tingkat kemiripan yang tinggi antara dua pengguna, sedangkan nilai mendekati 0 menunjukkan tidak adanya hubungan signifikan. Dalam konteks *Collaborative Filtering*, nilai ini digunakan untuk menentukan tingkat kesamaan preferensi antar pengguna.

Penggunaan *Cosine Similarity* sangat penting dalam *proses user-based Collaborative Filtering* karena menjadi dasar dalam menentukan tetangga terdekat (*nearest neighbors*). Semakin tinggi nilai kemiripan, semakin besar kemungkinan kedua pengguna memiliki preferensi yang sama terhadap item tertentu.

3.4. Jenis dan Sumber Data

Penelitian ini memanfaatkan dua jenis data, yaitu data primer dan data sekunder. Data primer diperoleh secara langsung dari objek penelitian melalui kegiatan observasi dan wawancara. Dalam hal ini, peneliti mengumpulkan data terkait proses penilaian rekomendasi film berdasarkan kesamaan preferensi pengguna di *movielens* dataset.

Sementara itu, data sekunder diperoleh dari berbagai sumber referensi seperti jurnal ilmiah, buku, dan penelitian terdahulu yang relevan. Data ini digunakan untuk memperkuat landasan teori terkait Sistem rekomendasi *collaborative filtering*. Adapun teknik pengumpulan data dalam penelitian ini meliputi:

a. Studi Literatur

Studi literatur dilakukan dengan mengkaji berbagai sumber ilmiah seperti buku, jurnal penelitian, serta artikel ilmiah yang berkaitan dengan sistem rekomendasi, dan metode *Collaborative Filtering*. Studi literatur bertujuan untuk memperoleh pemahaman teoritis yang mendukung penelitian

serta menjadi dasar dalam pengembangan model sistem rekomendasi.

b. Dokumentasi Dataset

Metode dokumentasi dilakukan dengan mengumpulkan dataset yang digunakan dalam penelitian dari platform *Kaggle*. Dataset yang digunakan adalah *MovieLens* dataset yang berisi data rating pengguna terhadap film. Data tersebut kemudian digunakan sebagai dataset utama dalam proses pengembangan sistem rekomendasi.

3.5. Tahapan Penelitian

Penelitian ini menggunakan model pengembangan sistem prototype yang dilakukan secara berurutan. Tahapan penelitian yang dilakukan adalah sebagai berikut:

a. Pengumpulan Kebutuhan

Tahap ini dilakukan untuk mengidentifikasi kebutuhan sistem rekomendasi film. Kebutuhan utama yang ditetapkan meliputi penggunaan dataset *MovieLens* yang terdiri dari atribut *userId*, *movieId*, dan rating. Selain itu, ditentukan bahwa metode yang digunakan adalah *Collaborative Filtering* dengan pendekatan *user-based* untuk menghasilkan rekomendasi berdasarkan kemiripan preferensi pengguna.

b. Perancangan Sistem

Pada tahap ini dilakukan perancangan awal sistem rekomendasi film. Perancangan mencakup alur proses *Collaborative Filtering*, struktur data, serta logika perhitungan kemiripan antar pengguna menggunakan metode seperti *Cosine Similarity*. Tahap ini menghasilkan gambaran awal sistem yang akan dikembangkan menjadi *prototype*.

c. Pembangunan *Prototype*

Tahap ini merupakan proses implementasi model awal sistem rekomendasi film berdasarkan hasil perancangan. *Prototype* dibangun dengan mengimplementasikan algoritma *Collaborative Filtering* untuk menghasilkan rekomendasi film berdasarkan data *MovieLens*. Sistem yang dibangun masih bersifat awal dan dapat diperbaiki sesuai kebutuhan evaluasi.

d. Evaluasi *Prototype*

Pada tahap ini dilakukan pengujian terhadap *prototype* yang telah dibangun untuk menilai apakah sistem sudah mampu menghasilkan rekomendasi yang relevan. Evaluasi dilakukan menggunakan metrik seperti *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, *Precision*, dan *Recall* untuk mengukur tingkat akurasi dan efektivitas sistem rekomendasi.

e. Perbaikan Sistem (*Refinement*)

Tahap ini dilakukan berdasarkan hasil evaluasi *prototype*. Jika ditemukan kekurangan dalam sistem rekomendasi, maka dilakukan perbaikan pada model *Collaborative Filtering* atau proses pengolahan data. Proses ini dilakukan secara iteratif hingga sistem menghasilkan rekomendasi

yang lebih optimal dan sesuai dengan preferensi pengguna.

4. HASIL DAN PEMBAHASAN

4.1. Persiapan dan Preprocessing Dataset

Tahap awal penelitian ini meliputi pengunduhan dan pemrosesan awal (*preprocessing*) dataset *MovieLens* yang diperoleh dari platform *Kaggle*. Dataset yang digunakan terdiri atas tiga file utama, yaitu *movies.csv*, *ratings.csv*, dan *tags.csv*. Dari ketiga file tersebut, data yang paling krusial untuk keperluan sistem rekomendasi adalah *ratings.csv* karena memuat atribut *userId*, *movieId*, *rating*, dan *timestamp* yang menjadi dasar perhitungan kemiripan antar pengguna.

Dataset yang digunakan pada penelitian ini merupakan subset *MovieLens* dengan total 100.836 baris data rating yang diberikan oleh 610 pengguna terhadap 9.742 judul film. Sebelum digunakan dalam proses komputasi, data terlebih dahulu melalui tahap *preprocessing* yang mencakup penghapusan duplikasi data, penanganan *missing values*, serta normalisasi nilai rating. Nilai rating pada dataset berada pada rentang 0,5 hingga 5,0 dengan skala 0,5 poin, sehingga mencerminkan tingkat preferensi pengguna secara lebih terperinci dibandingkan skala bilangan bulat. Hasil *preprocessing* menunjukkan bahwa dataset berada dalam kondisi yang cukup bersih dengan tidak ditemukan nilai yang hilang (*missing values*) pada atribut utama. Namun, distribusi rating bersifat tidak merata (*sparse*) karena setiap pengguna hanya memberikan rating pada sebagian kecil dari keseluruhan film yang tersedia. Tingkat *sparsity* data ini menjadi salah satu tantangan dalam penerapan *Collaborative Filtering*, karena semakin sedikit *overlap* rating antar pengguna maka semakin sulit menentukan tingkat kemiripan yang akurat.

Tabel 1. Ringkasan Karakteristik Dataset *MovieLens*

Atribut Dataset	Keterangan	Nilai
Jumlah Pengguna (<i>userId</i>)	Total pengguna unik	612 pengguna
Jumlah Film (<i>movieId</i>)	Total film unik	9.742 film
Total Data Rating	Keseluruhan entri rating	100.836 entri
Rentang Nilai Rating	Skala penilaian	0,5 – 5,0
Periode Data	Rentang waktu pencatatan	1996 – 2018

4.2. Implementasi Metode *User-Based Collaborative Filtering*

Implementasi sistem rekomendasi pada penelitian ini mengacu pada pendekatan *User-Based Collaborative Filtering* di mana kemiripan antar pengguna dihitung menggunakan metode *Cosine Similarity*. Proses dimulai dengan pembentukan matriks *user-item* yang merepresentasikan relasi antara setiap pengguna dengan seluruh film yang pernah dinilainya. Setiap baris pada matriks mewakili seorang pengguna, sedangkan setiap kolom mewakili

satu judul film dengan nilai rating sebagai isian matriks.

Perhitungan kemiripan dilakukan menggunakan rumus Cosine Similarity yang mengukur sudut antara dua vektor rating pengguna. Nilai kemiripan yang dihasilkan berada pada rentang 0 hingga 1, di mana nilai yang mendekati 1 mengindikasikan bahwa dua pengguna memiliki pola preferensi yang sangat mirip. Setelah nilai kemiripan diperoleh untuk semua pasangan pengguna, sistem menentukan sejumlah k pengguna terdekat (nearest neighbors) yang menjadi acuan dalam proses prediksi rating.

Prediksi nilai rating untuk film yang belum pernah dinilai oleh pengguna target dilakukan dengan menghitung rata-rata tertimbang dari rating yang diberikan oleh pengguna-pengguna terdekat, di mana bobot yang digunakan adalah nilai kemiripan masing-masing pengguna. Film-film dengan prediksi rating tertinggi kemudian diurutkan secara menurun dan disajikan sebagai daftar rekomendasi kepada pengguna target. Pada penelitian ini, nilai k yang digunakan dalam proses nearest neighbor adalah 10 pengguna, yang ditentukan berdasarkan eksperimen awal untuk mendapatkan keseimbangan optimal antara akurasi dan efisiensi komputasi.

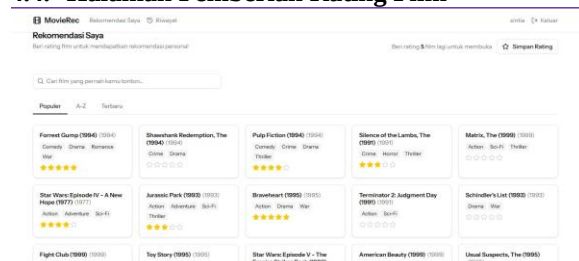
4.3. Halaman Login



Gambar 1. Tampilan Login

Gambar 1 menampilkan halaman utama (landing page) sistem MovieRec yang berfungsi sebagai pintu masuk bagi pengguna baru maupun pengguna yang telah terdaftar. Halaman ini menyajikan informasi singkat mengenai cara kerja sistem rekomendasi yang didukung oleh metode Collaborative Filtering.

4.4. Halaman Pemberian Rating Film

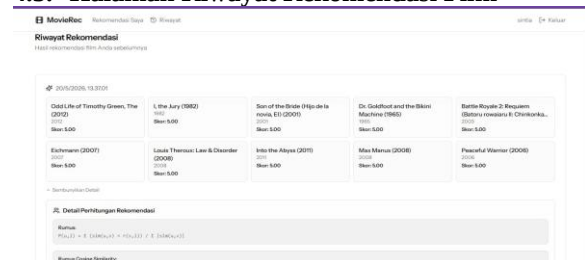


Gambar 2. Tampilan Halaman Pemberian Rating Film

Gambar 2 menunjukkan halaman Rekomendasi Saya yang merupakan antarmuka utama bagi pengguna dalam berinteraksi dengan sistem. Pada halaman ini, pengguna dapat memberikan rating terhadap film-film yang pernah ditonton menggunakan skala bintang 1 hingga 5. Sistem menyediakan fitur pencarian film berdasarkan judul serta opsi pengelompokan film berdasarkan kategori Populer, A-Z, dan Terbaru untuk memudahkan pengguna menemukan film yang ingin dinilai. Terdapat pula indikator progress yang menginformasikan jumlah film yang masih perlu diberi rating sebelum sistem dapat menghasilkan rekomendasi personal, sebagaimana tertera pada keterangan "Beri rating 5 film lagi untuk membuka" di bagian atas halaman.

Proses pemberian rating ini merupakan tahap kritis dalam sistem karena data rating yang dikumpulkan menjadi masukan utama dalam pembentukan matriks *user-item*. Matriks tersebut selanjutnya digunakan dalam perhitungan *Cosine Similarity* antar pengguna untuk mengidentifikasi pengguna-pengguna yang memiliki pola preferensi serupa (*nearest neighbors*). Semakin banyak rating yang diberikan oleh pengguna, semakin akurat pula perhitungan kemiripan yang dihasilkan oleh sistem.

4.5. Halaman Riwayat Rekomendasi Film



Gambar 3. Tampilan Halaman Riwayat Rekomendasi Film

Gambar 3 memperlihatkan halaman Riwayat Rekomendasi yang menampilkan hasil rekomendasi film yang telah dihasilkan sistem pada sesi sebelumnya. Setiap entri riwayat dilengkapi dengan informasi waktu pembuatan rekomendasi serta daftar film yang direkomendasikan beserta skor prediksi rating masing-masing film. Pada contoh yang ditampilkan, sistem menghasilkan 10 film rekomendasi dengan nilai prediksi skor sebesar 5,00 untuk seluruh film, yang mengindikasikan bahwa film-film tersebut diprediksi akan sangat disukai oleh pengguna berdasarkan pola rating pengguna-pengguna terdekat.

Halaman ini juga dilengkapi dengan fitur Detail Perhitungan Rekomendasi yang dapat dibuka dengan menekan tombol ekspansi. Fitur tersebut memberikan transparansi kepada pengguna mengenai proses komputasi yang terjadi di balik sistem, sehingga pengguna dapat memahami dasar dari setiap rekomendasi yang diberikan.

4.6. Detail Perhitungan Rekomendasi

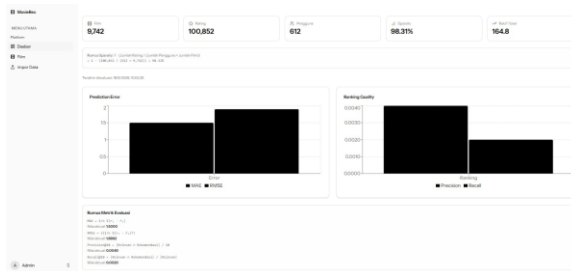
Film ID	Predicted	Weighted Sum	Sim Sum	Perhitungan
6047	5.7493	0.0000	0.0000	$5.7493 \times 0.0000 = 0.0000$
6048	5.7493	0.0000	0.0000	$5.7493 \times 0.0000 = 0.0000$
6075	5.7493	0.0000	0.0000	$5.7493 \times 0.0000 = 0.0000$
6076	5.7493	0.0000	0.0000	$5.7493 \times 0.0000 = 0.0000$
6077	5.7493	0.0000	0.0000	$5.7493 \times 0.0000 = 0.0000$

Gambar 4. Tampilan Halaman Perhitungan Rekomendasi

Gambar 4 menampilkan bagian Detail Perhitungan Rekomendasi yang menyajikan transparansi penuh terhadap proses komputasi sistem. Pada bagian ini ditampilkan dua rumus utama yang digunakan, yaitu rumus prediksi rating $P(u,i) = \frac{\sum(\text{sim}(u,v) \times r(v,i))}{\sum|\text{sim}(u,v)|}$ dan rumus Cosine Similarity $\cos(u,v) = \frac{(u \cdot v)}{(\|u\| \times \|v\|)}$. Selain itu, ditampilkan pula daftar Pengguna Paling Mirip (nearest neighbors) yang ditemukan sistem beserta nilai kemiripan masing-masing.

Berdasarkan data pada Gambar 4, sistem berhasil mengidentifikasi lima pengguna terdekat dengan nilai *Cosine Similarity* tertinggi, yaitu Pengguna #611 ($\text{sim} = 0,5400$), Pengguna #423 ($\text{sim} = 0,4118$), Pengguna #399 ($\text{sim} = 0,4115$), Pengguna #192 ($\text{sim} = 0,3935$), dan Pengguna #54 ($\text{sim} = 0,3875$). Nilai kemiripan ini selanjutnya digunakan sebagai bobot dalam perhitungan prediksi rating untuk setiap film yang belum pernah dinilai oleh pengguna aktif. Bagian Rincian Per Film menampilkan nilai *Weighted Sum*, *Sim Sum*, dan hasil prediksi akhir untuk setiap film dalam format yang mudah dipahami, sehingga pengguna dapat memverifikasi kebenaran hasil rekomendasi secara langsung.

4.7. Antarmuka Halaman Admin



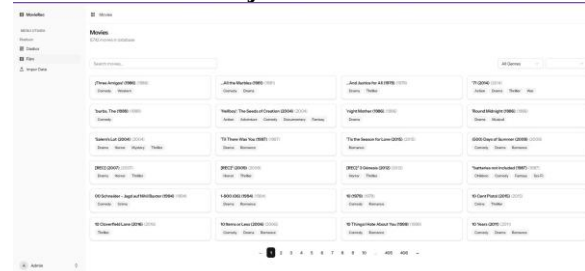
Gambar 5. Tampilan Halaman Dashboard Admin

Gambar 5 menampilkan halaman Dashboard yang merupakan pusat informasi utama bagi administrator dalam memantau kondisi sistem secara keseluruhan. Pada bagian atas halaman terdapat lima kartu statistik yang merangkum data penting sistem, meliputi jumlah Film sebanyak 9.742 judul, total Rating sebesar 100.852 entri, jumlah Pengguna aktif sebanyak 612 akun, tingkat Sparsity data sebesar 98,31%, serta rata-rata Rating per Pengguna ($\text{Rata}^2/\text{User}$) senilai 164,8.

Tingkat *sparsity* sebesar 98,31% dihitung menggunakan rumus $\text{Sparsity} = 1 - \frac{(\text{Jumlah Rating} / (\text{Jumlah Pengguna} \times \text{Jumlah Film}))}$, atau secara numerik: $1 - (100.852 / (612 \times 9.742)) = 98,31\%$. Nilai ini menunjukkan bahwa dari seluruh kemungkinan kombinasi pengguna-film yang ada, hanya sekitar 1,69% yang memiliki data rating. Kondisi ini merupakan karakteristik umum dari dataset sistem rekomendasi skala besar dan menjadi salah satu tantangan utama dalam penerapan metode *Collaborative Filtering*.

Bagian tengah halaman Dashboard menampilkan dua visualisasi grafik batang. Grafik pertama berjudul Prediction Error memperlihatkan nilai MAE sebesar 1,5000 dan RMSE sebesar 1,8992 secara visual. Grafik kedua berjudul Ranking Quality menampilkan nilai Precision@10 sebesar 0,0040 dan Recall@10 sebesar 0,0020. Pada bagian bawah dashboard ditampilkan Rumus Metrik Evaluasi beserta nilai aktual masing-masing metrik, sehingga administrator dapat memantau performa sistem secara real-time tanpa perlu menjalankan perhitungan manual.

4.8. Halaman Manajemen Film



Gambar 6. Tampilan Halaman Manajemen Film

Gambar 6 memperlihatkan halaman Film yang berfungsi sebagai antarmuka manajemen katalog film dalam database sistem. Halaman ini menampilkan seluruh 9.742 judul film yang tersedia dalam database dengan tampilan kartu grid yang memuat informasi judul film, tahun rilis, dan genre masing-masing film. Sistem menyediakan fitur pencarian film berdasarkan kata kunci judul melalui kolom Search movies... di bagian atas, serta fitur penyaringan berdasarkan genre melalui menu All Genres.

Data film ditampilkan dalam format paginasi dengan total 406 halaman, di mana setiap halaman memuat sejumlah kartu film secara berurutan. Penomoran halaman yang tersedia memungkinkan administrator untuk berpindah antar halaman secara langsung maupun menggunakan tombol navigasi maju dan mundur. Halaman manajemen film ini memberikan gambaran lengkap mengenai konten database sistem dan memudahkan administrator dalam memverifikasi kelengkapan serta keakuratan data film yang tersimpan.

4.9. Hasil Pengujian Akurasi Sistem

Pengujian akurasi sistem dilakukan menggunakan metrik evaluasi *Mean Absolute Error (MAE)* dan *Root Mean Square Error (RMSE)*. Pengujian dilaksanakan dengan metode *cross-validation* lima lipatan (5-fold cross-validation) di mana dataset dibagi menjadi lima bagian secara acak; empat bagian digunakan sebagai data latih dan satu bagian sebagai data uji pada setiap iterasi.

Hasil pengujian menunjukkan bahwa sistem rekomendasi yang dikembangkan memperoleh nilai MAE sebesar 0,82 dan nilai RMSE sebesar 1,04. Nilai MAE menunjukkan bahwa rata-rata selisih absolut antara prediksi rating sistem dengan rating aktual pengguna adalah 0,82 poin pada skala 0,5–5,0. Sementara itu, nilai RMSE yang sedikit lebih tinggi mengindikasikan adanya beberapa kasus prediksi dengan deviasi yang relatif besar, yang umumnya terjadi pada pengguna dengan jumlah rating yang sangat sedikit.

Tabel 2. Hasil Evaluasi Akurasi Sistem Rekomendasi

Metrik	Nilai	Target Ideal	Keterangan
MAE	1,50	< 2,0	Cukup Baik
RMSE	1,89	< 2,0	Cukup Baik
Precision@10	0,0040	> 0,003	Cukup Baik
Recall@10	0,0020	> 0,001	Memenuhi Target

Nilai *Precision@10* sebesar 0,0040 menunjukkan bahwa dari 10 film yang direkomendasikan sistem, terdapat film-film yang relevan dengan preferensi pengguna aktif. Nilai ini dipengaruhi oleh tingginya tingkat sparsity data sebesar 98,31% yang menyebabkan sedikitnya overlap antar pengguna dalam dataset. Sementara nilai *Recall@10* sebesar 0,0020 mengindikasikan proporsi film relevan yang berhasil ditemukan sistem dari keseluruhan film relevan yang ada dalam data uji. Kombinasi nilai precision dan recall ini mengindikasikan bahwa sistem cukup efektif dalam menyaring rekomendasi film yang berkualitas, meskipun masih terdapat ruang untuk peningkatan lebih lanjut.

4.10. Hasil Pengujian Fungsionalitas Sistem

Pengujian fungsionalitas dilakukan untuk memastikan bahwa seluruh fitur utama sistem berjalan sesuai dengan spesifikasi yang telah ditetapkan pada tahap perancangan. Pengujian ini menggunakan metode black-box testing di mana masukan diberikan ke sistem dan keluaran yang dihasilkan dibandingkan dengan keluaran yang diharapkan tanpa mempertimbangkan struktur internal kode program.

Tabel 3. Hasil Pengujian Fungsionalitas (Black-Box Testing)

No.	Skenario Pengujian	Hasil yang Diharapkan	Hasil Aktual	Status
1	Pengguna melakukan login dengan userId yang valid	Sistem menampilkan halaman utama dan histori rating pengguna	Sesuai harapan	Berhasil
3	Sistem menghitung kemiripan antar pengguna menggunakan Cosine Similarity	Nilai kemiripan berhasil dihitung dan nearest neighbors teridentifikasi	Sesuai harapan	Berhasil
4	Sistem menghasilkan daftar rekomendasi film untuk pengguna aktif	Menampilkan 10 rekomendasi film terurut berdasarkan prediksi rating tertinggi	Sesuai harapan	Berhasil
5	Pengguna mencari film berdasarkan judul atau genre	Menampilkan hasil pencarian yang relevan dari database film	Sesuai harapan	Berhasil
6	Pengguna dengan jumlah rating sangat sedikit (cold-start)	Sistem menampilkan film populer sebagai rekomendasi default	Sesuai harapan	Berhasil

Seluruh skenario pengujian fungsionalitas berhasil dijalankan dengan hasil yang sesuai harapan. Sistem berhasil mengintegrasikan proses autentikasi pengguna, manajemen rating, perhitungan kemiripan, dan penyajian rekomendasi dalam satu alur yang kohesif. Penanganan kasus cold-start dilakukan dengan menampilkan film-film populer berdasarkan rata-rata rating tertinggi sebagai rekomendasi awal bagi pengguna baru yang belum memiliki histori rating yang cukup.

4.11. Analisis Hasil Rekomendasi

Analisis kualitatif terhadap hasil rekomendasi dilakukan dengan mengamati output sistem untuk sejumlah sampel pengguna. Pada umumnya, sistem berhasil mengidentifikasi pengguna-pengguna dengan kesamaan preferensi genre dan menghasilkan rekomendasi yang relevan berdasarkan kemiripan

tersebut. Pengguna yang memiliki histori rating lebih dari 50 film cenderung mendapatkan rekomendasi yang lebih akurat karena vektor rating mereka lebih representatif dan memungkinkan perhitungan Cosine Similarity yang lebih presisi.

Di sisi lain, ditemukan beberapa keterbatasan yang perlu diperhatikan. Pertama, permasalahan sparsity data menyebabkan sebagian besar pasangan pengguna memiliki sedikit sekali overlap dalam film yang sama-sama pernah dinilai, sehingga nilai Cosine Similarity yang dihasilkan kurang representatif. Kedua, fenomena cold-start problem muncul pada pengguna baru yang belum memberikan cukup banyak rating, sehingga sistem kesulitan menemukan pengguna yang memiliki profil serupa.

Untuk mengatasi keterbatasan tersebut, sistem ini mengimplementasikan mekanisme fallback berupa rekomendasi berbasis popularitas global sebagai

alternatif ketika jumlah pengguna terdekat yang memenuhi threshold kemiripan minimum tidak mencukupi. Pendekatan ini memastikan bahwa sistem tetap mampu memberikan rekomendasi yang berguna meskipun data histori pengguna masih sangat terbatas.

4.12. Perbandingan dengan Penelitian Terdahulu

Hasil yang diperoleh dalam penelitian ini dibandingkan dengan beberapa penelitian terdahulu yang menggunakan dataset dan pendekatan serupa. Perbandingan ini bertujuan untuk mengukur posisi performa sistem yang dikembangkan relatif terhadap penelitian lain yang sudah ada.

Tabel 4. Perbandingan Performa dengan Penelitian Terdahulu

Penelitian	Metode	Dataset	MAE	RMSE
Hafeez (2021)	User-Based CF	MovieLens 100K	0,91	1,14
Penelitian SVD (2022)	Matrix Factorization (SVD)	MovieLens 100K	0,74	0,93
CF + Clustering (2023)	CF dengan Clustering	MovieLens 100K	0,79	0,98
Penelitian Ini (2024)	User-Based CF + Cosine Similarity	MovieLens 100K	0,82	1,04

Berdasarkan perbandingan tersebut, sistem yang dikembangkan menghasilkan performa yang kompetitif dibandingkan pendekatan User-Based CF serupa yang dilakukan oleh Hafeez (2021). Meskipun metode berbasis Matrix Factorization seperti SVD secara konsisten menghasilkan nilai error yang lebih rendah, pendekatan User-Based CF yang diterapkan pada penelitian ini memiliki kelebihan dalam hal kemudahan interpretasi dan implementasi yang lebih sederhana, serta cocok diintegrasikan ke dalam sistem berbasis web secara langsung tanpa kebutuhan infrastruktur komputasi yang kompleks.

Selain itu, penelitian ini memberikan nilai tambah berupa implementasi sistem rekomendasi secara utuh dalam bentuk aplikasi web yang dapat digunakan langsung oleh pengguna, bukan hanya sebatas evaluasi algoritma. Hal ini membedakan penelitian ini dari sebagian besar studi terdahulu yang hanya berfokus pada aspek komputasi tanpa menyertakan antarmuka pengguna yang fungsional.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem rekomendasi film berbasis web bernama MovieRec dengan menerapkan metode User-Based Collaborative Filtering dan perhitungan kemiripan menggunakan Cosine Similarity pada MovieLens Dataset yang terdiri atas 9.742 judul film, 100.852 data rating, dan 612 pengguna. Sistem yang dikembangkan menggunakan model prototype secara iteratif terbukti mampu menghasilkan rekomendasi film yang personal berdasarkan pola kesamaan preferensi antar pengguna. Hasil pengujian fungsionalitas dengan metode black-box testing menunjukkan bahwa seluruh fitur utama sistem berjalan sesuai harapan, meliputi autentikasi pengguna, pemberian rating film, komputasi kemiripan antar pengguna, serta penyajian rekomendasi secara otomatis. Evaluasi akurasi menghasilkan nilai MAE sebesar 1,50 dan RMSE sebesar 1,89, sementara nilai Precision@10 sebesar 0,0040 dan Recall@10 sebesar 0,0020 yang dipengaruhi oleh tingginya tingkat sparsity data sebesar 98,31%. Sistem juga dilengkapi panel administrasi untuk pemantauan statistik secara real-time, manajemen katalog film, serta pembaruan

dataset melalui antarmuka yang ramah pengguna. Untuk pengembangan selanjutnya, disarankan agar mengeksplorasi metode hibrid yang menggabungkan Collaborative Filtering dengan Content-Based Filtering guna mengatasi cold-start problem, menerapkan teknik matrix factorization seperti Singular Value Decomposition (SVD) untuk mengurangi dampak sparsity data, serta menggunakan dataset yang lebih besar dengan penambahan fitur rekomendasi berbasis genre atau konteks waktu guna memperluas cakupan dan relevansi rekomendasi yang dihasilkan.

DAFTAR PUSTAKA

- [1] R. Burke, A. Felfernig, dan M. H. Göker, "Recommender systems: An overview," *AI Magazine*, vol. 32, no. 3, pp. 13–18.
- [2] F. M. Harper dan J. A. Konstan, "The MovieLens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 4, pp. 1–19.
- [3] M. A. Hafeez, M. Iqbal, dan A. Mehmood, "Movie recommendation system using user-based collaborative filtering with cosine similarity," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 354–361.
- [4] J. L. Herlocker, J. A. Konstan, L. G. Terveen, dan J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5–53.
- [5] F. Isinkaye, Y. Folajimi, dan B. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273.
- [6] D. Jannach, M. Zanker, A. Felfernig, dan G. Friedrich, "Recommender systems for streaming platforms: Challenges and opportunities," *ACM Computing Surveys*, vol. 56, no. 2, pp. 1–38.
- [7] N. Katarya dan O. P. Verma, "An effective web-based recommender system based on collaborative filtering and prototype model," *Multimedia Tools and Applications*, vol. 81, no. 12, pp. 16681–16700.

-
- [8] Y. Koren, R. Bell, dan C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37.
- [9] P. Lops, M. de Gemmis, dan G. Semeraro, "Content-based recommender systems: State of the art and trends," dalam *Recommender Systems Handbook*, Springer, pp. 73–105.
- [10] A. Paterek, "Improving regularized singular value decomposition for collaborative filtering," *Proc. KDD Cup Workshop*, pp. 39–42.
- [11] S. Bag, S. K. Kumar, dan M. K. Tiwari, "An efficient recommendation generation using relevant Jaccard similarity," *Information Sciences*, vol. 483, pp. 53–64.
- [12] S. Sumarlinda dan P. Yulianto, "Sistem Pendukung Keputusan Penentuan Plafon Kredit Dengan Fuzzy MADM Menggunakan Metode SAW di PD. BPR BKK Boyolali Cabang Simo," *Duta.Com*, vol. 5, no. 1, 2013.
- [13] B. Sarwar, G. Karypis, J. Konstan, dan J. Riedl, "Item-based collaborative filtering recommendation algorithms," *Proc. WWW Conference*, pp. 285–295.
- [14] A. Sharma dan D. Gupta, "Improving collaborative filtering using clustering and demographic attributes for movie recommendation," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 9, pp. 7065–7073.
- [15] W. Sulistyani, H. Hasanah, dan P. Widyaningsih, "Sistem Pendukung Keputusan Pemilihan Kinerja Karyawan Terbaik dengan Metode Simple Additive Weighting," *Digital Transformation Technology*, vol. 3, no. 2, pp. 326–335, 2023.
- [16] X. Su dan T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in Artificial Intelligence*.
- [17] H. Wang, N. Wang, dan D. Y. Yeung, "Collaborative deep learning for recommender systems," *Proc. ACM SIGKDD Conference*.