

Implementasi Sistem Rekomendasi Film Menggunakan Algoritma Matrix Factorization Berbasis Flask untuk Personalisasi Konten di Era Disrupsi Digital

Satya Agni Prema Raswidhyantoro¹, Mohammad Idhom^{2*}, Andreas Nugroho Sihananto³

^{1,3} Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur

19081010109@upnjatim.ac.id

andreas.nugroho.jarkom@upnjatim.ac.id

² Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur

*Corresponding author email: idhom@upnjatim.ac.id

Abstrak— Perkembangan teknologi digital telah menyebabkan peningkatan signifikan pada volume informasi yang dihasilkan setiap hari. Dalam konteks layanan hiburan daring, pengguna sering mengalami kesulitan menemukan film sesuai preferensi pribadi akibat banyaknya pilihan konten yang tersedia (*information overload*). Penelitian ini mengusulkan sistem rekomendasi film berbasis algoritma Matrix Factorization (MF) dengan pendekatan *Collaborative Filtering* untuk memberikan rekomendasi personal. Dataset yang digunakan adalah MovieLens 100K, yang telah menjadi standar evaluasi dalam penelitian sistem rekomendasi. Model dilatih menggunakan metode optimisasi Stochastic Gradient Descent (SGD) dengan pembagian data 80:20 antara *training* dan *testing*. Evaluasi dilakukan menggunakan metrik Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE). Hasil penelitian menunjukkan bahwa model mencapai RMSE = 0.9026 dan MAE = 0.6943, dengan korelasi sebesar 0.5018, menandakan prediksi yang cukup akurat terhadap rating aktual pengguna. Penelitian ini difokuskan pada pembangunan model dan rancangan sistem, sedangkan implementasi aplikasi web akan menjadi tahap lanjutan.

Kata Kunci— Sistem rekomendasi, Matrix Factorization, Collaborative Filtering, MovieLens, Flask.

I. PENDAHULUAN

Perkembangan teknologi informasi di era disrupsi digital telah mengubah secara signifikan cara manusia mengakses, mengonsumsi, dan merekomendasikan konten digital. Layanan hiburan daring seperti Netflix dan Disney+ kini menawarkan ribuan judul film dan serial [1]. Melimpahnya pilihan tersebut seringkali menyebabkan pengguna kesulitan menentukan tontonan yang sesuai dengan preferensi pribadi. Fenomena ini dikenal dengan istilah *information overload*, di mana jumlah informasi yang terlalu besar justru menyulitkan pengguna dalam pengambilan keputusan.

Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah sistem rekomendasi, yaitu teknologi yang secara otomatis memberikan saran kepada pengguna berdasarkan pola perilaku dan preferensi mereka [3]. Metode ini telah menjadi elemen penting dalam berbagai

platform digital, mulai dari layanan streaming hingga e-commerce.

Di antara berbagai pendekatan sistem rekomendasi, Collaborative Filtering (CF) merupakan metode paling populer karena mampu memanfaatkan kesamaan perilaku antar pengguna atau antar item untuk menghasilkan rekomendasi [4]. Namun, metode CF tradisional memiliki beberapa keterbatasan, terutama ketika data bersifat jarang (*data sparsity*) atau ketika pengguna dan item baru belum memiliki cukup informasi (*cold start problem*)[6].

Sebagai solusi terhadap kelemahan tersebut, dikembangkan pendekatan *Matrix Factorization* (MF) yang mampu menemukan hubungan laten antara pengguna dan item melalui proses dekomposisi matriks rating [1]. Pendekatan ini terbukti lebih efisien pada dataset berskala besar dan menjadi salah satu metode terbaik pada kompetisi Netflix Prize [7]. Dataset MovieLens [2] digunakan dalam penelitian ini karena ketersediaannya yang luas, kualitas datanya yang terjaga, serta posisinya sebagai *benchmark* standar dalam pengujian algoritma sistem rekomendasi.

Namun, sebagian besar penelitian terdahulu masih berfokus pada evaluasi model secara teoritis tanpa menunjukkan integrasi nyata dalam sistem aplikasi. Berdasarkan kondisi tersebut, penelitian ini berfokus pada pembangunan dan evaluasi model sistem rekomendasi film berbasis *Matrix Factorization*, yang hasilnya diharapkan dapat menjadi fondasi bagi pengembangan sistem rekomendasi film berbasis web pada tahap selanjutnya.

II. TINJAUAN PUSTAKA

Bagian ini membahas teori-teori dan penelitian terdahulu yang menjadi dasar dalam pengembangan sistem rekomendasi film. Tinjauan ini mencakup pembahasan mengenai konsep dasar sistem rekomendasi, algoritma *Collaborative Filtering* dan *Matrix Factorization*, serta penelitian-penelitian yang relevan dengan penerapan metode tersebut pada dataset MovieLens. Pemahaman terhadap konsep dan penelitian sebelumnya diperlukan untuk memperkuat landasan teoritis serta

menentukan pendekatan yang paling sesuai dalam penelitian ini.

A. Penelitian Terdahulu

Berbagai penelitian sebelumnya menunjukkan efektivitas algoritma *Matrix Factorization* (MF) dalam meningkatkan akurasi sistem rekomendasi film. Fiagbe [7] mengembangkan sistem rekomendasi berbasis MF menggunakan dataset MovieLens dan membuktikan bahwa metode ini lebih akurat dibandingkan pendekatan berbasis konten. Saputra [10] melakukan perbandingan antara metode *user-based collaborative filtering* dan MF, dan hasilnya MF memberikan nilai RMSE yang lebih rendah. Arviyanto dan Soebroto [9] menerapkan MF pada produk industri lokal dan membuktikan fleksibilitas algoritma ini di luar domain film. Sementara Kim et al. (2022) dan Gao et al. (2025) memperkenalkan variasi regularisasi diversifikasi serta multi-view clustering, yang meningkatkan keberagaman hasil rekomendasi.

TABEL I

PENELITIAN TERDAHULU TENTANG SISTEM REKOMENDASI FILM

Peneliti	Metode	Dataset	Hasil Utama
Fiagbe (2023) [7]	Matrix Factorization	MovieLens	MF lebih akurat dibanding content-based
Saputra (2022) [10]	MF vs User-based CF	MovieLens	MF memiliki RMSE lebih rendah
Arviyanto & Soebroto (2025) [9]	MF untuk produk industri	Data Lokal	MF efektif di domain non-film
Kim et al. (2022)	Diversely Regularized MF	MovieLens	Rekomendasi lebih bervariasi
Gao et al. (2025)	Dynamic Multi-View MF	MovieLens	Akurasi meningkat pada data heterogen

Sebagian besar penelitian terdahulu hanya berfokus pada pengujian model tanpa integrasi ke aplikasi nyata. Penelitian ini mengisi celah tersebut dengan membangun model MF yang siap diintegrasikan pada tahap selanjutnya.

B. Sistem Rekomendasi

Sistem rekomendasi merupakan teknologi yang memberikan saran personal kepada pengguna berdasarkan pola perilaku atau preferensi mereka [3].

Sistem ini membantu pengguna menavigasi banyaknya pilihan konten digital dan meningkatkan keterikatan pengguna terhadap platform [5].

C. Collaborative Filtering

Collaborative Filtering (CF) memanfaatkan kesamaan antar pengguna atau antar item untuk memprediksi preferensi baru [6]. Metode ini terbagi menjadi *user-based* dan *item-based* CF. Kelemahan CF klasik adalah data sparsitas dan cold start problem [4]. Untuk mengatasi hal tersebut, dikembangkan pendekatan berbasis faktor laten seperti *Matrix Factorization* yang lebih efisien dan akurat [1].

D. Matrix Factorization

Matrix Factorization diperkenalkan oleh Koren et al. [1] pada kompetisi Netflix Prize dan terbukti sangat efektif untuk sistem rekomendasi berskala besar. MF bekerja dengan mendekomposisi matriks rating R menjadi dua matriks faktor laten P (pengguna) dan Q . Dengan teknik optimisasi seperti *Stochastic Gradient Descent* (SGD), sistem dapat menemukan representasi laten yang meminimalkan kesalahan prediksi [5]. Bobadilla et al. [4] menunjukkan bahwa MF memberikan akurasi yang lebih tinggi dibandingkan CF tradisional.

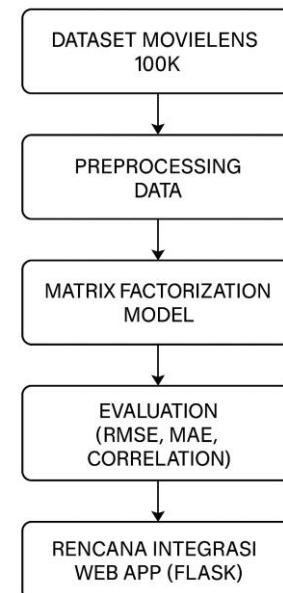
E. Dataset MovieLens

Dataset MovieLens dikembangkan oleh Harper dan Konstan [2] dan digunakan secara luas sebagai *benchmark* penelitian sistem rekomendasi. Dataset ini berisi jutaan data rating dari ribuan pengguna terhadap film, dengan versi 100K, 1M, hingga 20M. Versi 100K digunakan dalam penelitian ini karena berukuran cukup besar untuk merepresentasikan interaksi nyata namun tetap efisien untuk penelitian akademik.

F. Tools Pendukung

Penelitian ini menggunakan bahasa pemrograman Python dengan pustaka ilmiah seperti NumPy, Pandas, dan Scikit-learn. Framework Flask dirancang sebagai tahap pengembangan lanjutan karena bersifat ringan dan fleksibel untuk integrasi model Machine Learning [11].

III. METODOLOGI PENELITIAN



Gbr. 1 Kurva Training dan Validation Loss

Penelitian ini dirancang menggunakan pendekatan kuantitatif eksperimental, dengan fokus pada pembangunan model sistem rekomendasi film berbasis *Matrix Factorization* (MF). Tahapan penelitian dilakukan secara berurutan sebagaimana ditunjukkan pada Gambar 3, yang meliputi empat tahap utama:

1. pengumpulan dan *preprocessing* data,

2. pelatihan model menggunakan algoritma MF,
3. evaluasi performa model, serta
4. perancangan konseptual integrasi model ke dalam aplikasi berbasis Flask.

Alur penelitian ini bertujuan memastikan bahwa model yang dibangun tidak hanya akurat secara statistik, tetapi juga siap diadaptasikan pada sistem rekomendasi film berbasis web.

A. Dataset

Dataset yang digunakan adalah MovieLens 100K, yang berisi 100.836 data rating dari 610 pengguna terhadap 9.742 film [2]. Setelah preprocessing, jumlah rating akhir menjadi 90.258 dengan tingkat *sparsity* sebesar 95,94%.

B. Algoritma Matrix Factorization

Model MF dibangun dengan mendekomposisi matriks rating R menjadi dua matriks faktor laten P (pengguna) dan Q (film):

$$R \approx P \times Q^T$$

Keterangan:

R : rating hasil prediksi untuk pengguna ke-iii dan film ke-jjj

P : vektor faktor laten pengguna ke- u

Q : vektor faktor laten film ke- i

Fungsi loss yang diminimalkan adalah:

$$\begin{aligned} L = \sum_{(u,i) \in K} & (r_{ui} - p_u^T q_i)^2 + \lambda (\|p_u\|^2 \\ & + \|q_i\|^2) \end{aligned}$$

Keterangan:

L : fungsi *loss* total yang akan diminimalkan

r_{ui} : rating aktual pengguna ke- u terhadap film ke- i

λ : parameter regularisasi untuk mencegah *overfitting*

$\|p_u\|^2, \|q_i\|^2$: norma kuadrat dari masing-masing vektor faktor laten

Optimisasi dilakukan menggunakan *Stochastic Gradient Descent* (SGD).

C. Desain Sistem

Penelitian ini merancang sistem rekomendasi dengan tiga komponen utama:

1. Lapisan Data: pengambilan dan pembersihan dataset MovieLens.
2. Lapisan Model: pelatihan model MF menggunakan parameter $k=50$, $learning rate = 0.01$, $regularization = 0.1$.

TABEL II

PENELITIAN TERDAHULU TENTANG SISTEM REKOMENDASI FILM

Parameter	Nilai	Keterangan
Jumlah Faktor Laten (k)	50	Dimensi vektor pengguna dan film
Tingkat Pembelajaran (Learning Rate)	0.01	Kecepatan pembaruan bobot
Regularisasi (λ)	0.1	Mencegah <i>overfitting</i>
Ukuran Batch	256	Jumlah data per iterasi
Epoch	100 (penghentian dini pada 16)	Iterasi maksimum pelatihan

3. Lapisan Aplikasi (Rencana): rancangan integrasi model ke dalam aplikasi berbasis Flask sebagai tahap lanjut.

D. Evaluasi Model

Evaluasi dilakukan menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE).

RMSE sensitif terhadap kesalahan besar:

$$RMSE = \sqrt{\frac{1}{N} \sum_{(u,i) \in K} (r_{ui} - \hat{r}_{ui})^2}$$

Sedangkan MAE menghitung rata-rata kesalahan absolut:

$$MAE = \frac{1}{N} \sum_{(u,i) \in K} |r_{ui} - \hat{r}_{ui}|$$

Keterangan:

r_{ui} : nilai rating aktual ke- u

\hat{r}_{ui} : nilai rating hasil prediksi ke- u

N : jumlah total data uji

$RMSE$: akar dari rata-rata kuadrat kesalahan antara prediksi dan aktual

MAE : rata-rata selisih absolut antara prediksi dan aktual

Selain itu, dihitung koefisien korelasi *Pearson* untuk mengukur kesesuaian antara prediksi dan rating aktual.

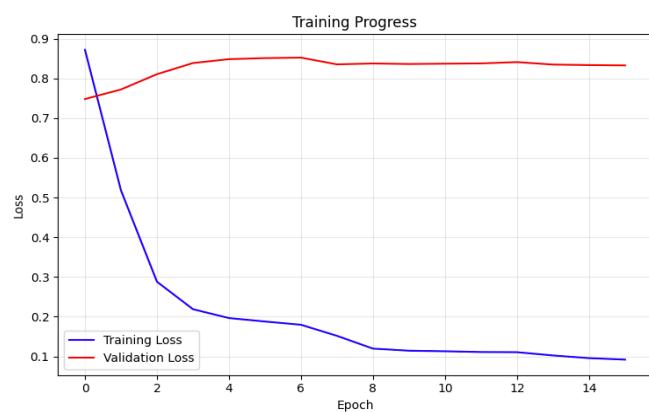
IV. HASIL DAN PEMBAHASAN

Model dilatih menggunakan pembagian data 80% *training*, 10% *validation*, dan 10% *testing*. Proses pelatihan berhenti otomatis pada epoch ke-16 dengan mekanisme *early stopping*.

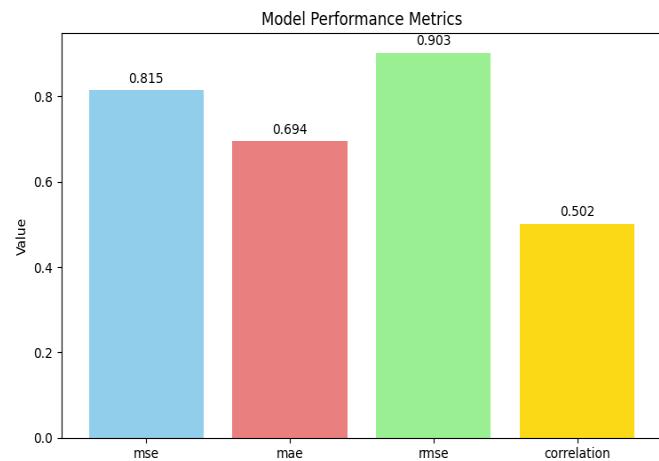
Hasil evaluasi model adalah:

- MSE: 0.8146
- MAE: 0.6943
- RMSE: 0.9026
- Korelasi: 0.5018

Nilai RMSE di bawah 1.0 menunjukkan bahwa model dapat memprediksi rating film dengan kesalahan yang rendah. Nilai korelasi positif sebesar 0.50 menunjukkan bahwa model mengenali hubungan laten antar pengguna dan film.



Gbr. 2 Kurva Training dan Validation Loss



Gbr. 3 Perbandingan nilai MSE, MAE, RMSE, dan Korelasi yang menunjukkan tingkat kesalahan prediksi rendah dan ketabilan model.

Hasil tersebut konsisten dengan penelitian Fiagbe [7] dan Saputra [10], yang melaporkan RMSE sekitar 0.88–0.94 pada dataset MovieLens 100K. Hasil eksperimen menunjukkan bahwa model telah mencapai tingkat konvergensi yang stabil dan kesalahan prediksi yang rendah, sehingga dapat dijadikan *baseline* untuk penelitian lanjutan dengan variasi parameter atau metode *hybrid*.

Dengan demikian, model MF yang dibangun memiliki performa kompetitif dan dapat menjadi dasar pengembangan sistem rekomendasi berbasis web di tahap berikutnya.

V. KESIMPULAN

Penelitian ini berhasil membangun model sistem rekomendasi film menggunakan algoritma *Matrix Factorization* dengan pendekatan *Collaborative Filtering*. Model menghasilkan RMSE = 0.9026, MAE = 0.6943, dan korelasi = 0.5018, yang menunjukkan prediksi cukup akurat terhadap data *sparse*. Penelitian ini masih berfokus pada pembangunan model dan rancangan sistem, tanpa tahap implementasi aplikasi web secara langsung.

VI. RENCANA PENGEMBANGAN SELANJUTNYA

Tahap berikutnya dari penelitian ini akan berfokus pada integrasi model Matrix Factorization ke dalam aplikasi berbasis Flask agar dapat diakses secara interaktif oleh pengguna. Selain itu, penelitian lanjutan akan mengeksplorasi pendekatan Hybrid Filtering dengan menggabungkan *Collaborative* dan *Content-Based Filtering* untuk meningkatkan relevansi rekomendasi. Pengembangan tambahan juga dapat mencakup analisis *cold start problem* serta penerapan algoritma optimisasi lain seperti Adaptive SGD atau Alternating Least Squares (ALS) untuk meningkatkan kecepatan dan akurasi model.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Pembangunan Nasional “Veteran” Jawa Timur, khususnya Program Studi Informatika, yang telah memberikan dukungan fasilitas dan bimbingan dalam penyusunan penelitian ini. Ucapan terima kasih juga disampaikan kepada dosen pembimbing yang telah memberikan arahan dan masukan selama proses penelitian, serta kepada rekan-rekan mahasiswa yang turut membantu dalam pengujian dan validasi model.

Penulis juga menyampaikan apresiasi kepada panitia Seminar Nasional Informatika Bela Negara (SANTIIKA) 2025 atas kesempatan yang diberikan untuk mempublikasikan hasil penelitian ini..

REFERENSI

- [1] Y. Koren, R. Bell, and C. Volinsky, “Matrix factorization techniques for recommender systems,” *IEEE Computer*, vol. 42, no. 8, pp. 30–37, 2009.
- [2] F. M. Harper and J. A. Konstan, “The MovieLens datasets: History and context,” *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 4, pp. 1–19, 2015.
- [3] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems Handbook*, 2nd ed., Springer, 2015.
- [4] J. Bobadilla, F. Ortega, and A. Gutiérrez, “Comprehensive evaluation of matrix factorization models for collaborative filtering recommender systems,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 6, pp. 15–23, 2024.
- [5] C. C. Aggarwal, *Recommender Systems: The Textbook*, Springer, 2016.
- [6] S. Rendle, L. Zhang, and Y. Koren, “On the difficulty of evaluating baselines: A study on recommender systems,” *arXiv preprint*, 2019.
- [7] R. Fiagbe, “Movie recommender system using matrix factorization,” *Master’s Thesis*, University of Central Florida, 2023.

- [8] A. Abdolmaleki et al., "An optimal context-aware content-based movie recommendation using genetic algorithms," *Journal of Information Science*, 2024.
- [9] F. B. Arviyanto and A. A. Soebroto, "Sistem rekomendasi model sandal pada home industri menggunakan collaborative filtering dan matrix factorization," *Jurnal PTIIK UB*, vol. 9, no. 1, 2025.
- [10] R. Saputra, "Implementasi sistem rekomendasi film dengan metode Matrix Factorization," *Jurnal Teknologi Informasi UB*, 2022.
- [11] M. Grinberg, *Flask Web Development: Developing Web Applications with Python*, 2nd ed., O'Reilly Media, 2018.