

PENERAPAN COLLABORATIVE FILTERING DENGAN METODE SINGULAR VALUE DECOMPOSITION UNTUK REKOMENDASI BUKU

Radithya Putra Pratama, Yuqimuna Haq, Gadhing Nandy Pratama,

Syifa Nur Rakhmah, Findi Ayu Sariasih, Imam Sutoyo

Informatika, Universitas Bina Sarana Informatika

Teknologi Informasi, Universitas Bina Sarana Informatika

Jl. Kramat Raya No.98 Senen, Jakarta Pusat, Indonesia

ganbattesmgt@gmail.com

ABSTRAK

Perkembangan teknologi informasi telah mendorong hadirnya platform baca digital seperti Google Books dan Gramedia Digital yang menyediakan ribuan pilihan buku. Namun, kelimpahan informasi ini menimbulkan tantangan *information overload* yang membuat pengguna kesulitan menemukan buku sesuai minat. Permasalahan utama adalah metode *Collaborative Filtering* konvensional memiliki akurasi rendah (53%) pada data dengan *sparsity* tinggi dan mengalami *cold start problem* untuk pengguna baru. Penelitian ini bertujuan menerapkan metode *Collaborative Filtering* berbasis *Singular Value Decomposition* (SVD) untuk meningkatkan akurasi sistem rekomendasi buku. Metode SVD dipilih karena kemampuannya mengekstrak *latent factors* dari *user-item matrix* untuk menemukan pola tersembunyi. Pendekatan penelitian bersifat kuantitatif dengan metode eksperimen menggunakan dataset Book-Crossing dari Kaggle (271.360 buku, 278.858 pengguna, 383.842 rating). Sistem dikembangkan menggunakan Python dengan pustaka NumPy, Pandas, dan Surprise. Evaluasi menggunakan metrik *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE) menunjukkan nilai RMSE 0.7209 dan MAE 0.5588, mengindikasikan tingkat kesalahan prediksi rendah dan peningkatan akurasi signifikan. Kontribusi penelitian adalah implementasi *pure SVD* yang terbukti efektif mengatasi *data sparsity* dan menghasilkan rekomendasi personal lebih akurat. Sistem berupa prototype aplikasi web AMBACA mampu memberikan 12 rekomendasi buku berdasarkan 5 rating awal pengguna.

Kata kunci : sistem rekomendasi, Collaborative Filtering, Singular Value Decomposition, matrix factorization, rekomendasi buku, RMSE

1. PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan masyarakat, termasuk dalam cara mengakses dan mengonsumsi konten bacaan. Era digital telah melahirkan berbagai platform baca digital seperti *Google Books*, *Gramedia Digital* yang menyediakan akses ke ribuan bahkan jutaan koleksi buku hanya dalam hitungan detik. Transformasi digital ini memberikan kemudahan bagi masyarakat untuk mengakses literatur dari berbagai genre dan penulis tanpa batasan geografis maupun waktu.

Secara ideal, keberadaan platform digital ini diharapkan dapat memfasilitasi pengguna untuk menemukan buku yang sesuai dengan minat dan kebutuhan mereka dengan mudah dan cepat. Sistem yang baik seharusnya mampu memahami preferensi pengguna dan memberikan rekomendasi yang personal sehingga pengalaman membaca menjadi lebih menyenangkan dan efisien. Dengan sistem rekomendasi yang tepat, pengguna tidak perlu menghabiskan banyak waktu untuk mencari buku yang relevan di antara ribuan pilihan yang tersedia.

Namun kenyataannya, kemudahan akses ini justru menciptakan tantangan baru yang dikenal sebagai *information overload*. Pengguna mengalami kesulitan dalam menentukan buku mana yang paling sesuai dengan minat mereka di tengah ribuan pilihan yang tersedia. Kondisi ini menyebabkan proses

pencarian buku menjadi tidak efisien dan seringkali berakhir dengan keputusan yang kurang optimal. Fenomena *information overload* terjadi ketika jumlah informasi yang tersedia melebihi kemampuan pengguna untuk memproses dan mengambil keputusan dengan efektif [1].

Untuk mengatasi masalah tersebut, sistem rekomendasi (*recommender system*) mulai dikembangkan dan diterapkan di berbagai platform digital. Salah satu pendekatan yang paling populer adalah *Collaborative Filtering* (CF), yang bekerja dengan menganalisis pola kesamaan antar pengguna (*user-based*) maupun antar item (*item-based*) untuk menghasilkan rekomendasi yang relevan. Namun, metode *Collaborative Filtering* konvensional memiliki beberapa kelemahan mendasar, yaitu masalah *data sparsity* (jarangnya data rating karena tidak semua pengguna memberikan penilaian) dan *cold start problem* (kesulitan memberikan rekomendasi untuk pengguna baru yang belum memiliki cukup data rating).

Penelitian sebelumnya menunjukkan bahwa metode *Singular Value Decomposition* (SVD) memiliki keunggulan signifikan dibanding metode *Collaborative Filtering* konvensional dalam menangani data sparse [2]. Namun, implementasi *pure SVD* pada platform penyedia buku umum dengan karakteristik pengguna yang beragam dan evaluasi

komprehensif menggunakan metrik RMSE dan MAE masih terbatas.

Berdasarkan latar belakang yang telah diuraikan, dapat diidentifikasi beberapa permasalahan utama dalam penelitian ini:

- Platform penyedia buku menyediakan ribuan pilihan bacaan yang menyebabkan pengguna mengalami *information overload*, sehingga kesulitan menemukan buku yang sesuai dengan minat mereka.
- Metode *Collaborative Filtering* konvensional memiliki tingkat akurasi yang rendah (53%) ketika diterapkan pada data dengan tingkat *sparsity* tinggi, yaitu kondisi di mana sebagian besar pengguna hanya memberikan sedikit rating.
- Pengguna baru menghadapi *cold start problem*, yaitu kondisi di mana sistem kesulitan memberikan rekomendasi akurat karena minimnya data preferensi yang dapat dianalisis.
- Terdapat kesenjangan performa sebesar 32% antara metode berbasis *matrix factorization* (SVD dengan akurasi 85%) dan metode berbasis kemiripan (KNN dengan akurasi 53%), namun implementasi *pure SVD* pada platform penyedia buku umum dengan dataset beragam masih terbatas dan memerlukan validasi lebih lanjut.

Berdasarkan permasalahan yang telah diidentifikasi, penelitian ini bertujuan untuk:

- Mengimplementasikan metode *Collaborative Filtering* berbasis *Singular Value Decomposition* (SVD) pada sistem rekomendasi buku untuk meningkatkan akurasi prediksi rating.
- Menganalisis efektivitas metode SVD dalam mengatasi permasalahan *data sparsity* dan *cold start problem* pada sistem rekomendasi buku.
- Mengevaluasi akurasi dan performa metode SVD dibandingkan dengan metode *Collaborative Filtering* konvensional menggunakan metrik RMSE dan MAE pada dataset Book-Crossing.

Untuk mencapai tujuan penelitian, dilakukan beberapa tahapan penyelesaian masalah sebagai berikut:

- Pengumpulan Data: Menggunakan dataset Book-Crossing dari Kaggle yang berisi 271.360 buku, 278.858 pengguna, dan 383.842 rating eksplisit untuk memastikan data yang representatif dan beragam.
- Preprocessing Data: Melakukan pembersihan data duplikat, penghapusan entri kosong, dan normalisasi skala rating menggunakan *Min-Max Normalization* ke rentang 1-5 untuk memastikan konsistensi data.
- Implementasi Metode SVD: Menggunakan teknik *matrix factorization* untuk mendekomposisi *user-item rating matrix* menjadi *latent factors* yang merepresentasikan pola tersembunyi antara pengguna dan buku. Implementasi dilakukan menggunakan bahasa pemrograman Python

dengan pustaka Surprise yang menyediakan algoritma SVD yang telah dioptimasi.

- Evaluasi Model: Mengukur performa model menggunakan metrik RMSE dan MAE untuk mengetahui tingkat kesalahan prediksi. Perbandingan dilakukan dengan metode *Collaborative Filtering* konvensional untuk memvalidasi peningkatan akurasi.
- Pengembangan Prototype: Membangun aplikasi web bernama AMBACA menggunakan framework Flask yang memungkinkan pengguna memberikan rating pada buku dan menerima rekomendasi personal secara real-time.

Dengan rencana penyelesaian masalah yang terstruktur ini, diharapkan penelitian dapat menghasilkan sistem rekomendasi buku yang lebih akurat, mampu mengatasi *data sparsity* dan *cold start problem*, serta memberikan rekomendasi yang relevan dan personal bagi pengguna.

2. TINJAUAN PUSTAKA

2.1. Sistem Rekomendasi

Sistem rekomendasi merupakan sistem yang dirancang untuk membantu pengguna menemukan item yang relevan dari sekumpulan besar pilihan berdasarkan preferensi dan perilaku pengguna [1]. menyatakan bahwa sistem rekomendasi telah menjadi komponen penting dalam berbagai platform digital untuk mengatasi masalah *information overload* yang semakin kompleks di era digital.

2.2. Penelitian Terdahulu

Berbagai penelitian telah dilakukan untuk meningkatkan performa sistem rekomendasi menggunakan metode *Collaborative Filtering* dan *matrix factorization*. Sub-bab ini menyajikan tinjauan terhadap penelitian-penelitian relevan yang menjadi dasar pengembangan sistem rekomendasi buku berbasis SVD.

- Studi Komparatif SVD dan KNN pada Sistem Rekomendasi Buku
Sebuah studi komparatif pada sistem rekomendasi buku di perpustakaan universitas menunjukkan bahwa model berbasis SVD mencapai akurasi 85%, jauh melampaui KNN yang hanya memperoleh 53%. Perbedaan ini menegaskan keunggulan pendekatan *matrix factorization* dalam mengolah data rating yang kompleks, khususnya melalui kemampuan SVD mengekstraksi *latent factors* yang tidak ditangkap metode berbasis kemiripan. Namun, karena penelitian dilakukan pada lingkungan pengguna yang relatif homogen, generalisasi temuan ke platform dengan karakteristik pengguna yang lebih beragam masih memerlukan verifikasi lebih lanjut.[2].
- Pengembangan Extended Collaborative Filtering dengan Adaptive KNN dan SVD
Penelitian lain mengembangkan sistem *Extended Collaborative Filtering* yang mengintegrasikan adaptive KNN dengan SVD untuk mengatasi

- masalah *cold start* dan *data sparsity* [3]. Hasil penelitian menunjukkan bahwa kombinasi kedua metode mampu meningkatkan akurasi rekomendasi secara signifikan, terutama untuk pengguna baru dengan data rating terbatas. Pendekatan ini efektif dalam mengurangi efek *cold start* dengan memanfaatkan kelebihan dari kedua algoritma. Namun, pendekatan *hybrid* ini memiliki kompleksitas implementasi yang tinggi karena melibatkan multiple algorithms, sehingga sulit untuk mengisolasi kontribusi spesifik dari metode SVD itu sendiri dan meningkatkan beban komputasi sistem.
- c. Implementasi Collaborative Filtering dengan Cosine Similarity
Implementasi *collaborative filtering* untuk rekomendasi buku menggunakan *cosine similarity* dan *Euclidean distance* menunjukkan efektivitas dalam menghasilkan rekomendasi personal pada dataset berukuran kecil hingga menengah [4]. Meskipun mampu memanfaatkan pola kesamaan perilaku pengguna, metode ini masih menghadapi keterbatasan terkait *data sparsity* dan *scalability*, yang menyebabkan penurunan performa ketika jumlah pengguna maupun item meningkat secara signifikan.
 - d. Sistem Rekomendasi Hybrid dengan Semantic Relationships
Pengembangan sistem rekomendasi buku hybrid yang menggabungkan *collaborative filtering* dan *content-based filtering* berbasis *semantic relationships* terbukti mampu meningkatkan akurasi rekomendasi di perpustakaan digital [5]. Dengan memanfaatkan teknik *pattern-based* untuk menganalisis hubungan semantik antar buku, pendekatan ini berhasil mengurangi bias data serta membantu mengatasi *cold start problem* pada pengguna baru. Metode ini dinilai efektif terutama pada lingkungan perpustakaan digital dengan katalog yang terstruktur dengan baik.
 - e. Studi Komparatif Teknik Matrix Factorization
Sebuah studi komparatif komprehensif telah dilakukan terhadap berbagai teknik *matrix factorization* termasuk NMF, SVD, dan SVD++ menggunakan dataset MovieLens [6]. Hasil penelitian menunjukkan bahwa SVD memberikan performa terbaik untuk metrik RMSE dengan nilai error yang paling rendah, sementara NMF unggul dalam metrik MAE untuk beberapa kasus spesifik. Penelitian ini membuktikan bahwa SVD efektif dalam menangani data *sparse* dengan kemampuan generalisasi yang baik dan memberikan akurasi prediksi yang tinggi pada berbagai ukuran dataset, mulai dari kecil hingga besar. Teknik SVD juga menunjukkan konsistensi performa yang stabil bahkan ketika tingkat *sparsity* data meningkat.
 - f. Penerapan SVD pada Sistem Rekomendasi E-Commerce
Penelitian mengenai penerapan model SVD pada sistem rekomendasi e-commerce menunjukkan bahwa metode ini mampu meningkatkan akurasi prediksi preferensi pengguna pada data transaksi yang bersifat *sparse* dan memiliki distribusi rating tidak merata [7]. Dengan memanfaatkan dekomposisi matriks untuk mengidentifikasi pola laten dalam perilaku pembelian, SVD terbukti menghasilkan rekomendasi yang relevan dan dapat diadaptasikan ke berbagai domain rekomendasi. Namun, penelitian tersebut belum membahas secara mendalam kemampuan SVD dalam menangani *cold start problem* pada produk baru yang minim interaksi.
 - g. Optimasi SVD Berbasis Data Filling
Sebuah penelitian telah mengembangkan pendekatan optimasi algoritma SVD dengan mengintegrasikan teknik data filling untuk meningkatkan performa sistem rekomendasi [8]. Teknik data filling digunakan untuk mengisi nilai-nilai kosong dalam user-item matrix sebelum proses dekomposisi dilakukan, sehingga mengurangi efek *sparsity* yang dapat menurunkan akurasi prediksi. Penelitian ini menunjukkan bahwa proses data normalization dan hyperparameter tuning berpengaruh signifikan terhadap performa akhir sistem rekomendasi berbasis SVD. Hasil eksperimen membuktikan bahwa kombinasi data filling dengan SVD dapat meningkatkan akurasi prediksi rating dibandingkan dengan penggunaan SVD standar tanpa preprocessing tambahan. Pendekatan ini efektif dalam konteks dataset dengan tingkat *sparsity* yang sangat tinggi, di mana sebagian besar pengguna hanya memberikan rating pada sebagian kecil item yang tersedia.
 - h. Robust Matrix Completion untuk Data Rating Diskrit
Penelitian terkini mengusulkan pendekatan robust *matrix completion* yang dirancang untuk menangani data rating diskrit dan memitigasi pengaruh *fake profiles* serta *outliers* dalam sistem rekomendasi [9]. Dengan mengintegrasikan teknik deteksi anomali ke dalam proses *matrix factorization*, metode ini mampu mengidentifikasi serta mengurangi dampak rating palsu, sehingga meningkatkan akurasi dan reliabilitas prediksi, khususnya pada platform dengan tingkat interaksi pengguna yang tinggi. Kontribusi utama penelitian tersebut terletak pada pengembangan kerangka kerja yang dapat membedakan rating genuine dari rating manipulatif, sehingga kualitas data pelatihan model tetap terjaga.
Berdasarkan tinjauan penelitian terdahulu, beberapa celah penelitian teridentifikasi sebagai berikut:
 - i. Studi komparatif sebelumnya [2] hanya dilakukan pada lingkungan perpustakaan universitas dengan pengguna homogen, sehingga diperlukan validasi pada platform buku digital dengan pengguna yang lebih heterogen.

1. Pendekatan hybrid yang telah dikembangkan [3] memiliki kompleksitas tinggi, sehingga sulit mengisolasi kontribusi spesifik metode SVD terhadap peningkatan akurasi.
2. Meskipun penelitian mengenai optimasi SVD melalui *data filling* [8], *robust matrix completion* [9], serta penerapannya pada e-commerce [7] menunjukkan hasil menjanjikan, penerapan *pure SVD* pada platform penyedia buku dengan karakteristik data beragam dan evaluasi komprehensif menggunakan RMSE dan MAE masih jarang dilakukan.
3. Masih dibutuhkan penelitian yang secara khusus mengevaluasi efektivitas SVD dalam mengatasi data sparsity dan cold start problem pada konteks rekomendasi buku berskala besar.

Penelitian ini mengatasi celah tersebut dengan mengimplementasikan *pure SVD* pada aplikasi web rekomendasi buku menggunakan dataset Book-Crossing yang memiliki keragaman pengguna tinggi (278.858 pengguna unik), serta melakukan evaluasi komprehensif menggunakan RMSE dan MAE untuk membandingkan performanya dengan metode *Collaborative Filtering* konvensional.

2.3. Machine Learning

Machine Learning merupakan cabang dari *artificial intelligence* yang memungkinkan sistem belajar dan membuat prediksi berdasarkan data tanpa memerlukan pemrograman eksplisit untuk setiap tugas. Terdapat tiga paradigma utama, yaitu *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. Dalam konteks sistem rekomendasi, machine learning berperan penting dalam menganalisis pola interaksi pengguna-item untuk menghasilkan rekomendasi yang personal melalui pemanfaatan data historis seperti rating, klik, atau riwayat pembelian. Pendekatan yang umum digunakan adalah *supervised learning* karena model dilatih menggunakan data berlabel untuk memprediksi preferensi pengguna.

Penelitian terkait *matrix factorization* menunjukkan bahwa teknik ini mampu mendekomposisi matriks *user-item* menjadi representasi berdimensi lebih rendah sehingga efektif dalam menangani *data sparse* dan mengidentifikasi hubungan laten antara pengguna dan item yang tidak tampak secara langsung [10].

2.4. Collaborative Filtering

Collaborative Filtering merupakan pendekatan dalam sistem rekomendasi yang memanfaatkan pola perilaku dan preferensi pengguna dengan asumsi bahwa pengguna yang memiliki preferensi serupa pada masa lalu cenderung menunjukkan preferensi serupa di masa mendatang. Metode ini memiliki dua pendekatan utama, yaitu *user-based* yang menganalisis kesamaan antar pengguna dan *item-based* yang mengidentifikasi kesamaan antar item.

Implementasi *collaborative filtering* pada rekomendasi buku menggunakan *cosine similarity* dan *Euclidean distance* terbukti mampu menghasilkan rekomendasi personal yang efektif, meskipun masih menghadapi tantangan signifikan terkait *data sparsity* dan *scalability* seiring meningkatnya jumlah pengguna dan item [4].

Selain itu, penelitian lain mengembangkan sistem rekomendasi hybrid yang mengombinasikan *collaborative filtering* dengan *content-based filtering* berbasis *semantic relationships*, dan menunjukkan bahwa integrasi *semantic knowledge* dapat mengurangi bias data dan membantu mengatasi *cold start problem* bagi pengguna baru [5].

2.5. Singular Value Decomposition (SVD)

SVD merupakan teknik *matrix factorization* yang mendekomposisi *user-item rating matrix* ke dalam tiga matriks berdimensi lebih rendah. Metode ini efektif dalam mengatasi *data sparsity* dengan mengidentifikasi *latent factors* yang tersembunyi serta mereduksi dimensi tanpa menghilangkan informasi penting [10].

Selain itu, studi komparatif terhadap beberapa teknik *matrix factorization* seperti NMF, SVD, dan SVD++ menggunakan dataset MovieLens menunjukkan bahwa SVD memberikan performa terbaik pada metrik RMSE, sedangkan NMF lebih unggul pada metrik MAE. Temuan tersebut menegaskan bahwa SVD merupakan metode yang akurat, mampu menangani data *sparse*, dan adaptif untuk berbagai skala dataset [6].

2.6. Metrik Evaluasi

Root Mean Square Error (RMSE) dan *Mean Absolute Error (MAE)* merupakan metrik standar untuk mengevaluasi akurasi sistem rekomendasi. RMSE memberikan penalti lebih besar terhadap error yang besar dengan mengukur rata-rata kuadrat error antara prediksi dan nilai aktual, sedangkan MAE memberikan bobot yang sama untuk semua error dengan mengukur rata-rata absolut perbedaan prediksi.

Dalam penelitiannya [4] tentang rekomendasi buku menggunakan RMSE sebagai metrik utama untuk mengukur akurasi prediksi rating pada sistem *collaborative filtering*. Penelitian tersebut menunjukkan bahwa RMSE efektif dalam mengevaluasi seberapa dekat prediksi model dengan rating aktual yang diberikan pengguna. [6] melakukan analisis komparatif berbagai teknik *matrix factorization* menggunakan kombinasi RMSE dan MAE, dan menyatakan bahwa kedua metrik memberikan gambaran komprehensif tentang performa model karena RMSE lebih sensitif terhadap outlier sedangkan MAE memberikan ukuran error yang lebih stabil.

MSE dihitung menggunakan formula:

$$\sqrt{\frac{\sum (y_{\text{pred}} - y_{\text{actual}})^2}{n}}$$

dimana y_{pred} adalah nilai prediksi, y_{actual} adalah nilai aktual, dan n adalah jumlah data. Sementara MAE dihitung dengan:

$$\frac{\sum |y_{\text{pred}} - y_{\text{actual}}|}{n}$$

Nilai RMSE dan MAE yang lebih rendah mengindikasikan performa model yang lebih baik. Dalam konteks sistem rekomendasi buku, nilai RMSE di bawah 1.0 umumnya dianggap menunjukkan akurasi yang baik untuk skala rating 1-5.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen. Pendekatan ini digunakan untuk menguji efektivitas metode *Collaborative Filtering* berbasis *Singular Value Decomposition* (SVD) dalam meningkatkan akurasi sistem rekomendasi buku.

3.1. Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian terapan (applied research). Penelitian ini difokuskan pada penerapan metode *Singular Value Decomposition* (SVD) untuk meningkatkan performa sistem rekomendasi buku, dengan mengukur hasil berdasarkan tingkat akurasi menggunakan metrik RMSE dan MAE.

3.2. Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan utama, yaitu:

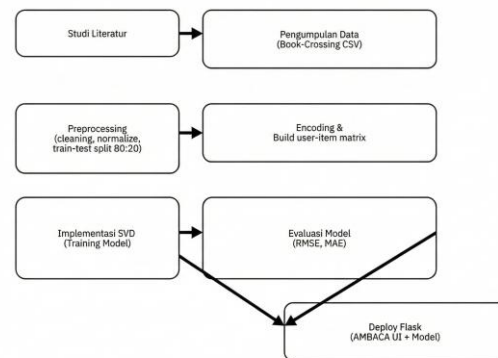
- Studi Literatur**
Mengumpulkan dan mengkaji literatur yang relevan mengenai sistem rekomendasi, metode *Collaborative Filtering*, teknik *Matrix Factorization* menggunakan SVD, serta metode evaluasi performa seperti RMSE dan MAE.
- Pengumpulan Data**
Data yang digunakan bersumber dari dataset *Book-Crossing* yang tersedia di platform Kaggle. Dataset ini berisi informasi mengenai penilaian (rating) yang diberikan pengguna terhadap buku-buku dari berbagai kategori.
- Preprocessing Data**
Meliputi proses pembersihan data (data cleaning), normalisasi rating, serta pembagian data menjadi training set dan testing set dengan rasio 80:20 untuk memastikan data siap digunakan pada tahap pelatihan dan pengujian model.
- Implementasi Metode SVD**
Pada tahap ini, model *Collaborative Filtering* berbasis SVD diimplementasikan menggunakan

bahasa pemrograman *Python* dengan bantuan pustaka *NumPy*, *Pandas*, dan *Surprise*. Proses dilakukan dengan melakukan dekomposisi matriks pengguna-item untuk menemukan *latent factors* yang menggambarkan hubungan antara pengguna dan buku.

e. Evaluasi Model

Model diuji menggunakan metrik evaluasi *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE) untuk mengukur tingkat kesalahan prediksi dan akurasi sistem rekomendasi.

Flowchart Penelitian - AMBACA (SVD Book Recommender)



Gambar 1. Flowchart Penelitian Sistem Rekomendasi Buku Berbasis SVD

Flowchart pada Gambar 1 menggambarkan alur penelitian yang digunakan dalam pengembangan sistem rekomendasi AMBACA. Proses dimulai dari studi literatur untuk memperkuat pemahaman mengenai sistem rekomendasi dan metode *Collaborative Filtering*, kemudian dilanjutkan dengan pengambilan dataset *Book-Crossing* dari Kaggle. Setelah itu, dilakukan tahap *preprocessing data*, implementasi metode SVD menggunakan Python, serta evaluasi performa model menggunakan metrik RMSE, MAE, dan MSE. Alur tersebut menunjukkan rangkaian tahapan penelitian yang sistematis, mulai dari perumusan konsep hingga pengujian hasil sistem.

3.3. Metode Pengembangan Sistem

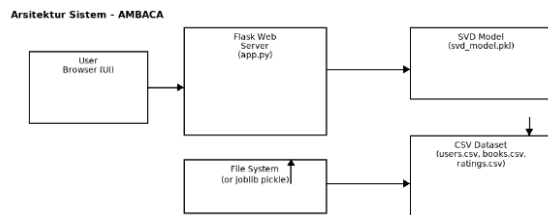
Penelitian ini menggunakan metode pengembangan perangkat lunak **Scrum** karena bersifat *iteratif* dan *adaptif*, sehingga mendukung proses pengembangan sistem secara bertahap dengan peningkatan kualitas pada setiap siklus. Scrum juga memberikan fleksibilitas dalam pengelolaan perubahan kebutuhan dan memfasilitasi evaluasi berkelanjutan terhadap performa sistem rekomendasi berbasis *Singular Value Decomposition* (SVD).

Tahapan Scrum yang diterapkan meliputi:

- Product Backlog**, yaitu menyusun daftar kebutuhan sistem seperti pengumpulan dataset, preprocessing data, implementasi metode SVD, evaluasi dengan RMSE dan MAE, serta perancangan *prototype*.
- Sprint Planning**, yaitu menentukan prioritas pekerjaan tiap sprint dan menetapkan *sprint goal*.

- c. Sprint Development, yaitu melakukan implementasi secara bertahap menggunakan *Python* dan pustaka seperti *NumPy*, *Pandas*, serta *Surprise*.
- d. Sprint Review dan Retrospective, yaitu mengevaluasi hasil setiap sprint dan memperbaiki proses pengembangan berikutnya.

Hasil akhir dari tahapan ini berupa prototipe sistem rekomendasi buku digital berbasis SVD yang mampu memberikan rekomendasi sesuai preferensi pengguna.



Gambar 2. Arsitektur Sistem Rekomendasi Buku Digital AMBACA

Arsitektur pada Gambar 2 memperlihatkan komponen utama sistem rekomendasi AMBACA yang dibangun menggunakan framework Flask. Sistem menerima input dari pengguna melalui antarmuka web, kemudian memproses data rating dan informasi buku yang tersimpan dalam file CSV. Data tersebut diolah oleh model SVD yang telah dilatih sebelumnya dan disimpan dalam format *svd_model.pkl*. Hasil prediksi kemudian dikembalikan ke pengguna dalam bentuk daftar rekomendasi buku. Alur komunikasi antar komponen berjalan secara dua arah antara antarmuka, server Flask, model, dan dataset.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Implementasi Sistem

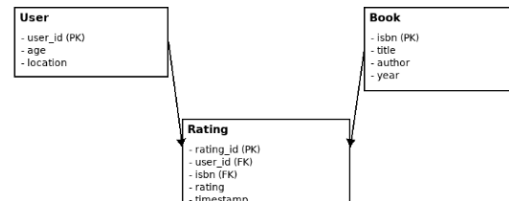
Penelitian ini menghasilkan prototipe sistem rekomendasi buku digital berbasis *Collaborative Filtering* dengan metode *Singular Value Decomposition* (SVD) yang dinamakan AMBACA (*Aku Mau Membaca*). Sistem dikembangkan menggunakan framework *Flask* berbasis web dengan bahasa pemrograman *Python* serta pustaka pendukung seperti *NumPy*, *Pandas*, dan *Surprise*.

Dataset yang digunakan merupakan dataset publik *Book-Crossing* dari Kaggle() yang berisi data pengguna, buku, dan rating. Setelah melalui tahap *preprocessing* yang mencakup pembersihan data duplikat, penghapusan entri kosong, serta normalisasi skala rating menggunakan *Min-Max Normalization* ke rentang 1–5, diperoleh total 271.360 data buku, 278.858 data pengguna, dan 383.842 data rating eksplisit yang digunakan untuk pelatihan model.

Proses *encoding* dilakukan terhadap atribut *User-ID* dan *ISBN* untuk mengonversi nilai kategori menjadi numerik, kemudian dibentuk menjadi *user-item interaction matrix* sebagai input bagi metode SVD. Model hasil pelatihan disimpan dalam format

svd_model.pkl menggunakan pustaka *Joblib*, dan dievaluasi menggunakan metrik RMSE, MAE, serta MSE untuk mengukur akurasi prediksi dan kinerja sistem rekomendasi.

ERD Dataset - Book-Crossing (simplified)



Gambar 3. Entity Relationship Diagram (ERD) Dataset Book-Crossing

ERD pada Gambar 3 memperlihatkan hubungan antara entitas *User*, *Book*, dan *Rating* yang digunakan dalam proses pembentukan *user-item matrix*. Setiap pengguna dapat memberikan satu atau lebih rating terhadap buku tertentu, sedangkan setiap buku dapat menerima banyak rating dari pengguna yang berbeda.

Untuk memberikan gambaran kuantitatif mengenai dataset yang digunakan dalam penelitian ini, disajikan statistik dataset *Book-Crossing* pada Tabel 1 berikut.

Tabel 1. Statistik Dataset Book-Crossing

Atribut	Jumlah Data	Keterangan
Pengguna (User)	278.858	Data pengguna unik dalam dataset
Buku (Book)	271.360	Data buku unik setelah pembersihan
Rating (Explicit)	383.842	Total rating eksplisit yang digunakan
Skala Rating	1–5	Hasil normalisasi Min-Max

Dari Tabel 1 dapat dilihat bahwa jumlah pengguna dan buku dalam dataset cukup besar, sedangkan data rating mengalami penyusutan signifikan setelah proses pembersihan. Hal ini menunjukkan bahwa sebagian besar data rating bersifat *sparsity* (jarang), sehingga metode SVD dipilih karena mampu menangani permasalahan tersebut melalui dekomposisi matriks.

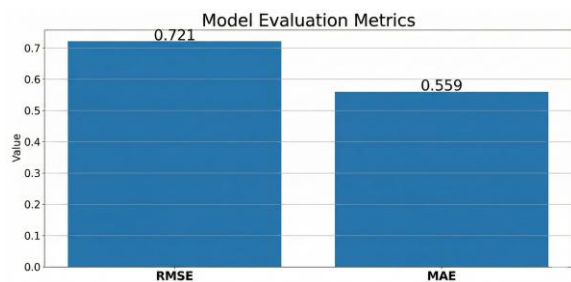
4.2. Hasil Pengujian Model

Evaluasi sistem dilakukan dengan membandingkan nilai prediksi model terhadap data aktual menggunakan tiga metrik evaluasi utama, yaitu RMSE, MAE, dan MSE. Hasil pengujian ditunjukkan pada Tabel 2.

Tabel 2. Hasil Pengujian Akurasi Model SVD

Metrik Evaluasi	Nilai
RMSE	0.7209
MAE	0.5588

Untuk memberikan visualisasi yang lebih jelas terhadap performa model, hasil evaluasi pada Tabel x.x divisualisasikan dalam bentuk grafik batang seperti yang ditampilkan pada Gambar 4.



Gambar 4. Grafik Hasil Evaluasi Akurasi Model SVD (RMSE, MAE)

Nilai RMSE sebesar 0.7209 menunjukkan bahwa selisih rata-rata antara rating prediksi dan rating aktual berada pada kisaran 0.72 poin dari skala 1–5. Sementara nilai MAE sebesar 0.5588 menunjukkan tingkat kesalahan absolut rata-rata yang relatif kecil, menandakan bahwa model SVD mampu memprediksi rating dengan tingkat deviasi rendah.

4.3. Pembahasan Hasil

Metode *Collaborative Filtering* berbasis SVD terbukti mampu mengatasi permasalahan *data sparsity* yang umum terjadi pada sistem rekomendasi. Dengan melakukan dekomposisi matriks *user-item* menjadi *latent factor*, sistem dapat mengenali pola tersembunyi antara preferensi pengguna dan karakteristik buku tanpa perlu menunggu pengguna memberikan banyak rating eksplisit.

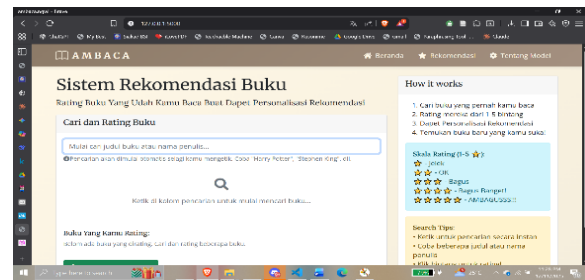
Hasil evaluasi memperlihatkan bahwa model SVD memiliki performa yang stabil dengan tingkat kesalahan prediksi di bawah 1 poin. Secara konseptual, nilai RMSE di bawah 1 sudah menandakan bahwa sistem telah menghasilkan rekomendasi yang cukup akurat untuk konteks *book recommendation*.

Selain itu, sistem juga menerapkan fitur *model confidence* yang mengindikasikan tingkat keyakinan model terhadap setiap hasil rekomendasi. Semakin tinggi nilai *confidence*, semakin besar keyakinan model bahwa pengguna akan menyukai buku yang direkomendasikan.

4.4. Hasil Implementasi Antarmuka

Antarmuka pengguna (UI) sistem AMBACA dirancang dengan pendekatan *user-friendly* dan responsif. Pengguna dapat melakukan pencarian buku berdasarkan nama penulis atau judul, memberikan rating pada buku yang pernah dibaca, serta memperoleh rekomendasi buku baru secara otomatis.

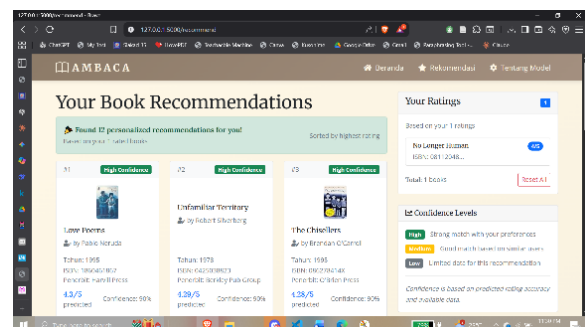
Beberapa tampilan utama sistem ditunjukkan pada Gambar 5–7.



Gambar 5. Halaman Utama Sistem AMBACA Menampilkan kolom pencarian buku serta petunjuk penggunaan sistem.



Gambar 6. Halaman Pemberian Rating Menampilkan hasil pencarian buku berdasarkan penulis dan memungkinkan pengguna memberikan rating pada setiap buku dengan skala 1–5.



Gambar 7. Halaman Rekomendasi Buku Menampilkan daftar buku rekomendasi beserta tingkat kepercayaan model (*confidence level*).

Dari hasil implementasi, sistem mampu memberikan rekomendasi personal hingga 12 buku berdasarkan lima rating awal yang diberikan pengguna. Buku-buku yang direkomendasikan menampilkan informasi lengkap seperti judul, penulis, tahun terbit, ISBN, penerbit, serta prediksi rating dan tingkat *confidence* model.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan sistem rekomendasi buku berbasis *Collaborative Filtering* menggunakan metode *Singular Value Decomposition* (SVD) melalui pengembangan sistem AMBACA. Model yang dihasilkan menunjukkan performa prediksi yang baik dengan nilai RMSE sebesar 0.7209 dan MAE sebesar 0.5588, menandakan tingkat kesalahan yang rendah serta akurasi yang stabil

dalam memprediksi rating buku. Penerapan SVD juga terbukti efektif dalam mengatasi *data sparsity* dan meningkatkan relevansi rekomendasi bagi pengguna.

Untuk pengembangan selanjutnya, disarankan integrasi pendekatan *hybrid recommendation* dengan mengombinasikan *content-based filtering* serta memperluas cakupan dataset guna menghasilkan rekomendasi yang lebih kontekstual dan adaptif terhadap dinamika preferensi pengguna. Selain itu, karena AMBACA masih berbasis web dan belum optimal untuk perangkat mobile, penelitian lanjutan diusulkan mengarah pada pengembangan aplikasi mobile native (Android/iOS) atau pendekatan *Progressive Web App* (PWA) untuk meningkatkan aksesibilitas dan pengalaman pengguna, termasuk pemanfaatan fitur notifikasi untuk rekomendasi buku terbaru.

DAFTAR PUSTAKA

- [1] O. A. S. Ibrahim, E. M. G. Younis, E. A. Mohamed, and W. N. Ismail, "Revisiting recommender systems: an investigative survey," *Neural Comput. Appl.*, vol. 37, no. 4, pp. 2145–2173, 2025, doi: 10.1007/s00521-024-10828-5.
- [2] E. Ahmed and A. Letta, "Book Recommendation Using Collaborative Filtering Algorithm," *Appl. Comput. Intell. Soft Comput.*, vol. 2023, 2023, doi: 10.1155/2023/1514801.
- [3] Sagedur Rahman, "Extended Collaborative Filtering Recommendation System with Adaptive KNN and SVD," *Int. J. Eng. Manag. Res.*, vol. 13, no. 4, pp. 105–112, 2023, doi: 10.31033/ijemr.13.4.14.
- [4] P. K. Sahoo, S. Dhanish, V. Ramana, and A. N. Kumar, "Book Recommendation using Collaborative Filtering," vol. 12, no. 04, pp. 365–369, 2023.
- [5] F. Wayesa, M. Leranso, G. Asefa, and A. Kedir, "Pattern-based hybrid book recommendation system using semantic relationships," *Sci. Rep.*, vol. 13, no. 1, pp. 1–12, 2023, doi: 10.1038/s41598-023-30987-0.
- [6] R. Alabduljabbar, "applied sciences Matrix Factorization Collaborative-Based Recommender System for Riyadh Restaurants : Leveraging Machine Learning to Enhance Consumer Choice," 2023.
- [7] W. Shalannanda, R. F. Mulia, A. I. Muttaqien, N. R. Hibatullah, and A. Firdaus, "Singular value decomposition model application for e-commerce recommendation system," *JITEL (Jurnal Ilm. Telekomun. Elektron. dan List. Tenaga)*, vol. 2, no. 2, pp. 103–110, 2022, doi: 10.35313/jitel.v2.i2.2022.103-110.
- [8] Y. Liu, F. Huang, X. Xie, and H. Huang, "Research on Singular Value Decomposition Recommendation Algorithm Based on Data Filling," *Int. J. Inf. Technol. Syst. Approach*, vol. 16, no. 3, pp. 1–15, 2023, doi: 10.4018/IJITSA.320222.
- [9] A. Archimbaud, A. Alfons, and I. Wilms, "Robust Matrix Completion for Discrete Rating-Scale Data: Coping with Fake Profiles in Recommender Systems," vol. 141, pp. 1–34, 2025, [Online]. Available: <http://arxiv.org/abs/2412.20802>
- [10] Y. Zhang, "An Introduction to Matrix factorization and Factorization Machines in Recommendation System, and Beyond," pp. 1–15, 2022, [Online]. Available: <http://arxiv.org/abs/2203.11026>