

**PERBANDINGAN COSINE SIMILARITY DAN  
JACCARD SIMILARITY UNTUK ITEM-BASED  
COLLABORATIVE FILTERING DALAM  
MENENTUKAN REKOMENDASI BUKU**



**Oleh :**

**Raihan Rafi Listyan Putra**

**21081010226**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA  
TIMUR  
2024**

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Dalam era digital saat ini, informasi dan konten semakin mudah diakses melalui platform online. Salah satu platform yang berkembang pesat adalah platform penyedia buku digital dan fisik, seperti Google Books, Goodreads, dan aplikasi perpustakaan *online*. Di tengah banyaknya pilihan buku yang tersedia, sering kali pengguna merasa kesulitan dalam memilih buku yang sesuai dengan minat mereka. Hal ini menyebabkan kebutuhan akan sistem rekomendasi yang dapat membantu pengguna menemukan buku yang sesuai dengan preferensi pribadi mereka.

Sistem rekomendasi berbasis data menjadi solusi untuk masalah ini. Salah satu metode yang umum digunakan dalam sistem rekomendasi adalah *Collaborative Filtering (CF)*. Collaborative Filtering dapat dibagi menjadi dua jenis, yaitu *item-based* dan *user-based*. Dalam penelitian ini, sistem rekomendasi buku yang dikembangkan menggunakan metode *Item-Based Collaborative Filtering (IBCF)*, yang memberikan rekomendasi berdasarkan kesamaan antara item yang telah diberikan rating oleh pengguna.

### **1.2 Rumusan Masalah**

Masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut.

- Bagaimana membangun sistem rekomendasi buku menggunakan metode *Item-Based Collaborative Filtering*?
- Apa saja tantangan dan keuntungan dalam implementasi *Item-Based Collaborative Filtering* untuk rekomendasi buku?
- Bagaimana performa sistem rekomendasi ini dalam memberikan saran buku yang relevan bagi pengguna?

### **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut.

- Mengembangkan sistem rekomendasi buku berbasis *Item-Based Collaborative Filtering*.
- Menganalisis performa sistem rekomendasi buku menggunakan evaluasi berbasis akurasi.
- Memberikan rekomendasi buku yang relevan dengan preferensi pengguna berdasarkan data yang ada.

#### **1.4 Batasan Masalah**

Penelitian ini dibatasi pada pengembangan sistem rekomendasi buku dengan menggunakan metode *Item-Based Collaborative Filtering* pada dataset buku dari platform tertentu (misalnya, Goodreads atau dataset yang disediakan oleh pengguna). Sistem ini hanya akan mengandalkan data rating buku dari pengguna dan tidak akan mempertimbangkan faktor lain seperti metadata buku (*genre*, penulis, dll).

#### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

- Menyediakan sistem rekomendasi buku yang dapat membantu pengguna dalam menemukan buku sesuai preferensi mereka.
- Memberikan kontribusi dalam bidang pengembangan sistem rekomendasi berbasis *collaborative filtering*.
- Membantu pengembang platform buku untuk mengimplementasikan rekomendasi yang lebih efektif.

#### **1.6 Sistematika Penulisan**

Penelitian ini disusun dengan sistematika sebagai berikut:

##### **Bab 1 Pendahuluan**

Menyajikan latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan.

##### **Bab 2 Tinjauan Pustaka**

Membahas teori-teori yang mendasari penelitian ini, termasuk pengertian sistem rekomendasi, collaborative filtering, serta studi terkait sebelumnya.

### **Bab 3 Metodologi Penelitian**

Menguraikan metodologi yang digunakan dalam penelitian, termasuk teknik implementasi *Item-Based Collaborative Filtering* dan evaluasi sistem.

### **Bab 4 Hasil dan Pembahasan**

Menyajikan hasil eksperimen dan analisis performa sistem rekomendasi.

### **Bab 5: Kesimpulan dan Saran**

Menyimpulkan temuan penelitian dan memberikan saran untuk penelitian lebih lanjut.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Sistem Rekomendasi**

Sistem rekomendasi adalah sebuah sistem yang membantu pengguna menemukan produk, informasi, atau layanan yang sesuai dengan preferensi mereka berdasarkan data yang ada. Sistem ini banyak diterapkan dalam berbagai domain seperti *e-commerce*, layanan *streaming*, dan media sosial. Rekomendasi dapat dilakukan dengan berbagai pendekatan sebagai berikut:

- *Content-Based Filtering*, berdasarkan kesamaan konten atau fitur produk.
- *Collaborative Filtering*, berdasarkan perilaku pengguna lain yang memiliki preferensi serupa.
- *Hybrid*, gabungan dari metode *collaborative* dan *content-based filtering*.

#### **2.2 Collaborative Filtering**

*Collaborative filtering* (CF) adalah metode yang paling umum digunakan dalam sistem rekomendasi. Metode ini bekerja berdasarkan asumsi bahwa pengguna yang memiliki preferensi yang sama di masa lalu akan terus memiliki preferensi yang sama di masa depan. *Collaborative filtering* dapat dibagi menjadi dua jenis utama, yaitu:

- *User-Based Collaborative Filtering* memberikan rekomendasi berdasarkan kesamaan antara pengguna.
- *Item-Based Collaborative Filtering* memberikan rekomendasi berdasarkan kesamaan antara item (produk, buku, film, dll.).

#### **2.3 Item-Based Collaborative Filtering**

*Item-Based Collaborative Filtering* berfokus pada menghitung kesamaan antar item berdasarkan rating yang diberikan oleh pengguna. Jika dua item memiliki pola rating yang serupa dari banyak pengguna, maka item tersebut dianggap mirip. Algoritma ini kemudian akan memberikan rekomendasi item yang mirip dengan item yang telah disukai atau di-rating oleh pengguna.

Metode tersebut digunakan karena *item* cenderung lebih stabil daripada daripada *user-based* CF yang melibatkan preferensi pengguna yang bisa berubah-ubah. Beberapa algoritma yang digunakan untuk menghitung kesamaan antar item adalah *Cosine Similarit* dan *Jaccard Similarity*.

## 2.4 Studi Terkait

Berbagai penelitian telah dilakukan untuk mengembangkan sistem rekomendasi berbasis *collaborative filtering* di berbagai domain. Beberapa penelitian yang relevan dengan penelitian ini, yaitu:

- Penelitian oleh Rosita, A. et al. (2022) yang membangun sistem rekomendasi buku Perpustakaan Universitas Mulawarman menggunakan *Item-Based Collaborative Filtering* dengan pendekatan *cosine similarity* dan menghasilkan nilai MAE mencapai 0.568 dan nilai MAPE yang dicapai sebesar 22%.
- Penelitian yang dilakukan oleh Yusmar, A. (2020) dalam melakukan Implementasi metode *collaborative filtering* dengan *item-based* untuk rekomendasi rumah makan menggunakan pendekatan algoritma *adjusted cosine similarity* yang menghasilkan nilai *precision* tertinggi sebesar 0,732 dengan nilai MAE terkecil adalah 0,623.

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sistematika Penelitian**

Penelitian yang dilakukan memiliki tujuan untuk membangun sistem rekomendasi buku menggunakan metode *Item-Based Collaborative Filtering*. Metodologi yang digunakan terdiri dari beberapa tahapan, yaitu pengumpulan data, praproses data, implementasi algoritma IBCF, serta evaluasi sistem.

#### **3.2 Pengumpulan Data**

Data yang digunakan dalam penelitian adalah *dataset* buku dan rating pengguna dari Amazon atau *dataset* serupa yang berisi informasi tentang buku yang sudah diberikan rating oleh pengguna. *Dataset* ini memuat informasi mengenai ID buku, ID pengguna, dan rating yang diberikan oleh pengguna terhadap buku tersebut.

#### **3.3 Praproses Data**

Tahap praproses data meliputi:

- Membersihkan Data, menghapus data yang tidak lengkap maupun terdapat duplikasi.
- Normalisasi Rating, menormalkan rating agar berada dalam rentang tertentu.
- Membangun Matriks Rating yang akan menyimpan rating dari setiap pengguna untuk setiap buku yang menjadi *input* utama dalam algoritma *Item-Based Collaborative Filtering*.

#### **3.4 Algoritma Item-Based Collaborative Filtering**

Implementasi *Item-Based Collaborative Filtering* dilakukan dengan langkah-langkah berikut:

- Menghitung Kesamaan Antara Item  
Menggunakan metode seperti *Cosine Similarity* dan *Jaccard Similarity* untuk mengukur kesamaan antara buku-buku yang ada berdasarkan rating yang diberikan oleh pengguna.
- Memberikan Rekomendasi

Berdasarkan kesamaan antara item, sistem akan memberikan rekomendasi buku kepada pengguna yang mirip dengan buku yang telah mereka beri rating tinggi.

### 3.5 Evaluasi Sistem

Evaluasi dilakukan dengan menggunakan metrik-metrik berikut:

- Precision@K  
Mengukur seberapa banyak rekomendasi yang diberikan relevan bagi pengguna.
- Recall@K  
Mengukur seberapa banyak buku yang relevan ditemukan di antara semua rekomendasi yang tersedia.
- F1-Score@K  
Merupakan kombinasi antara *precision* dan *recall*.

### 3.6 Implementasi Sistem

Sistem rekomendasi ini akan diimplementasikan menggunakan bahasa pemrograman **Python** dan pustaka **Pandas** untuk manipulasi data, **Scikit-learn** untuk perhitungan kesamaan, dan **Flask** (jika ingin mengembangkan aplikasi web).