

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Publikasi ilmiah merupakan suatu aktivitas menyebarluaskan karya tulis ilmiah hasil penelitian [1] yang menjadi sarana bagi para penulis artikel ilmiah sebagai pengakuan keilmuan [2]. Berdasarkan data yang dirilis oleh *Scimago Journal & Country Rank*, Indonesia menempati urutan ke-38 dunia dan urutan ke-9 di wilayah Asia dengan total 376.908 dokumen karya ilmiah yang dipublikasikan berdasarkan metrik data *Scopus* per maret 2024 [3]. Hal ini menunjukkan bahwa negara Indonesia merupakan negara yang sangat berkontribusi dalam menyebarluaskan atau mempublikasikan karya ilmiah.

Berdasarkan data tersebut, karya ilmiah sudah tersebar luas di Indonesia. Tidak hanya di Indonesia, negara lain juga berkontribusi dalam menyebarluaskan dan mempublikasikan karya ilmiah. Sebagai contoh perpustakaan digital *IEEE Xplore* yang saat ini memiliki 6.327.944 dokumen karya tulis ilmiah. Karya tulis ilmiah ini terdiri dari berbagai macam karya seperti *Conferences*, *Early Access Articles*, *Journals*, *Standards*, *Magazines*, *Courses*, dan *Books* [4].

Hal tersebut baru dari karya ilmiah saja, belum lagi dari jenis yang lain seperti buku, majalah, dan lain sebagainya. Dengan banyaknya bahan bacaan ini, menimbulkan sebuah masalah menemukan bahan bacaan tersebut. Sangat melelahkan untuk mencari satu persatu bahan bacaan dari sekian banyak bahan bacaan yang ada. Hal ini menyebabkan sebagian orang kesulitan dalam menemukan bahan bacaan yang sesuai dengan preferensi mereka [5].

Dalam upaya membantu kesulitan tersebut, manusia mengembangkan sebuah sistem yang bernama Sistem Rekomendasi. Sistem Rekomendasi merupakan suatu sistem yang dibangun dengan tujuan untuk memberikan rekomendasi kepada pengguna guna membantu mereka dalam menemukan bahan bacaan yang sesuai dengan preferensi mereka [6][7].

Menurut Joeran Beel dan peneliti lainnya, makalah penelitian tentang Sistem Rekomendasi pertama kali diperkenalkan oleh Giles dan peneliti lainnya pada tahun 1998 sebagai bagian dari proyek mesin pencari informasi sains publik

dan perpustakaan digital yang bernama *CiteSeer* [8][9]. Semenjak saat itu, ada setidaknya 217 makalah penelitian lain yang diterbitkan mengenai pendekatan rekomendasi ini [9].

Seiring berjalannya waktu, semakin banyak penelitian yang diterbitkan mengenai Sistem Rekomendasi. Sistem Rekomendasi sendiri memiliki beberapa kelebihan seperti memudahkan pengguna menemukan karya ilmiah yang mereka cari [10] dan meminimalisir waktu yang dibutuhkan dalam pencariannya [11]. Selain itu, Sistem Rekomendasi juga dapat digunakan untuk mencari kemiripan dari satu pengguna dengan pengguna lain dan memberikan rekomendasi yang sesuai dengan kemiripan tersebut [12]. Meskipun demikian, Sistem Rekomendasi masih memiliki kelemahan seperti saat tidak ada data yang cukup tentang preferensi pengguna baru (*Cold Start*) yang dapat menyebabkan hasil rekomendasi tidak relevan atau akurat [13]. Oleh karena itu, pengembangan Sistem Rekomendasi diperlukan agar rekomendasi dapat menjadi lebih akurat.

Collaborative Filtering (CF) merupakan metode tradisional yang digunakan dalam mengembangkan Sistem Rekomendasi yang menggunakan preferensi historis pengguna untuk merekomendasikan *item* kepada mereka [14]. *Collaborative Filtering* memiliki beberapa pendekatan dalam merekomendasikan *item* kepada pengguna salah satunya yaitu *User-Based Collaborative Filtering* (UBCF) [15].

User-Based Collaborative Filtering atau UBCF merupakan salah satu pendekatan pada CF yang memberikan rekomendasi dengan cara memprediksi minat pengguna target melalui informasi yang berupa peringkat atau *rating* dari pengguna tetangga terdekat (*Neighbor*) [16]. UBCF dapat menghasilkan rekomendasi yang sangat dipersonalisasi karena didasarkan pada preferensi pengguna yang serupa [16][17]. Hal ini membuatnya dapat menghasilkan rekomendasi yang relevan. Selain itu, UBCF juga lebih efektif dalam menangani *Cold Start* untuk *item* baru. Hal ini dikarenakan UBCF tidak memerlukan banyak data tentang *item* itu sendiri, melainkan fokus pada kesamaan antar pengguna.

Meskipun demikian, menggunakan UBCF memiliki kelemahan tersendiri seperti *Sparsity*. *Sparsity* ini terjadi karena ada terlalu banyak *item* yang tidak banyak di *rating* oleh pengguna. Hal ini menyebabkan sulitnya menemukan

Neighbor yang cukup mirip sehingga rekomendasi yang diberikan menjadi kurang akurat atau kurang relevan. Oleh karena itu, ada beberapa pendekatan yang dapat digunakan untuk membantu meminimalisir kelemahan dari UBCF ini salah satunya ialah dengan menggunakan pendekatan *Hybrid Approach* [18].

Hybrid Approach atau dalam Bahasa Indonesia ialah pendekatan hibrida merupakan salah satu pendekatan dalam mengembangkan Sistem Rekomendasi dengan menggabungkan dua atau lebih metode dalam memberikan rekomendasi [19]. Pengembangan Sistem Rekomendasi menggunakan pendekatan hibrida biasanya dilakukan dengan tujuan untuk menutupi kelemahan yang dimiliki oleh masing-masing metode [20]. Oleh karena itu, pendekatan hibrida merupakan salah satu pendekatan yang dapat digunakan untuk pengembangan Sistem Rekomendasi agar hasil rekomendasi yang diberikan dapat menjadi lebih akurat [21]. Berdasarkan hal tersebut, *Content-Based Filtering* ditambahkan dalam penelitian ini dengan tujuan untuk menutupi kelemahan yang terjadi pada saat menggunakan metode UBCF.

Content-Based Filtering (CBF) merupakan metode yang memberikan rekomendasi *item* berdasarkan kemiripannya dengan *item* yang disukai oleh pengguna. CBF dapat memberikan rekomendasi yang sangat dipersonalisasi berdasarkan preferensi individu pengguna. Hal ini dikarenakan sistem menganalisis konten atau karakteristik *item* yang disukai oleh pengguna dan mencocokkannya dengan *item* yang memiliki karakteristik serupa. Selain itu, metode ini juga tidak bergantung pada perilaku pengguna lain dalam menghasilkan rekomendasi [22]. Oleh karena itu, data *rating* dari pengguna lain tidak terlalu diperlukan untuk memberikan rekomendasi yang relevan. Hal ini membuatnya dapat membantu menutupi keterbatasan yang ada pada metode CF. Tentunya, dalam hal ini diperlukan sebuah teknik penggabungan yang efisien dalam mengatur penggunaan algoritma dalam kondisi tertentu.

Dalam menetapkan sebuah teknik integrasi atau penggabungan yang efisien, diperlukan sebuah kondisi di mana sistem mampu secara adaptif mengenali situasi dan perubahan dalam data serta pola interaksi yang terjadi [23]. Oleh karena itu, penelitian ini menggunakan teknik *Switching* yang memungkinkan sistem untuk berpindah secara dinamis antara *Collaborative Filtering* atau *Content-Based*

Filtering tergantung pada kondisi yang ada. Dengan demikian, sistem dapat memberikan rekomendasi yang lebih akurat dan relevan sesuai dengan keadaan atau kondisi yang sedang terjadi. Hal ini tentunya harus ditunjang dengan analisis yang baik terhadap data [24]. Analisis ini mencakup pemahaman mendalam terhadap karakteristik data, seperti distribusi data [25], pola perilaku pengguna [26] Analisis pola perilaku penggemar JKT48 pada aplikasi X dalam berkomunikasi dengan sesama penggemar, serta hubungan antara *item* dan pengguna [27].

Untuk menganalisis data secara mendalam, diperlukan sebuah metode yang dapat menggali informasi penting dalam data, mengidentifikasi pola tersembunyi, dan menghubungkan variabel-variabel yang relevan secara sistematis. Metode tersebut harus terstruktur agar dapat diterapkan secara sistematis dan menghasilkan output yang konsisten. Oleh karena itu, penelitian ini menggunakan metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang memiliki struktur yang jelas dan sistematis [28], fleksibel [29], dan iteratif [30]. Dengan demikian, jika ditemukan masalah atau kebutuhan baru, pengguna metode ini dapat kembali ke fase sebelumnya untuk melakukan perbaikan atau penyesuaian.

Dalam prakteknya, penelitian ini akan menggunakan sebagian *dataset* yang bernama *Book Recommendation Dataset* [31] yang diambil dari situs web *Kaggle* dan dirilis oleh *Möbius* sebagai sumber data utama. Dataset ini berisi informasi mengenai buku, pengguna, serta *rating* yang diberikan oleh pengguna terhadap buku-buku tersebut. Data ini akan digunakan untuk mengembangkan sistem rekomendasi yang mampu memberikan saran buku berdasarkan preferensi pengguna.

Tahapan awal penelitian mencakup eksplorasi *dataset* untuk memahami struktur data, membersihkan data dari nilai-nilai yang tidak valid atau hilang, dan melakukan analisis awal untuk mengidentifikasi pola atau karakteristik penting. Setelah itu, pendekatan hibrida akan diterapkan sesuai dengan tujuan penelitian.

Dengan menggunakan metode CRISP-DM, proses pengolahan data akan dilakukan secara bertahap mulai dari tahap pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran. Pendekatan ini diharapkan dapat memastikan bahwa sistem rekomendasi yang dikembangkan dapat memberikan rekomendasi yang akurat.

Selanjutnya, model rekomendasi diimplementasikan menggunakan platform iOS. Platform iOS dipilih karena kemudahan dalam proses pengembangan serta dukungan *SwiftData* yang memungkinkan penyimpanan data secara lokal tanpa memerlukan server. Dengan pendekatan ini, aplikasi berperan sebagai media implementasi, bukan sebagai objek utama dari proses rekayasa perangkat lunak.

1.2 Rumusan Masalah Penelitian

Berdasarkan pemaparan latar belakang diatas, daftar rumusan masalah dibuat sebagai berikut:

1. Bagaimana cara mempermudah pengguna dalam menemukan bahan bacaan yang sesuai dengan preferensi mereka?

1.3 Tujuan Penelitian

Berdasarkan perumusan masalah di atas, daftar tujuan penelitian dibuat sebagai berikut:

1. Membantu mempermudah pengguna dalam menemukan bahan bacaan yang sesuai dengan preferensi mereka menggunakan sistem rekomendasi dengan pendekatan hibrida yang menggabungkan algoritma *Collaborative Filtering* dan *Content-Based Filtering* menggunakan *Switching Technique* dan *Book Recommendation Dataset*.

1.4 Batasan Masalah Penelitian

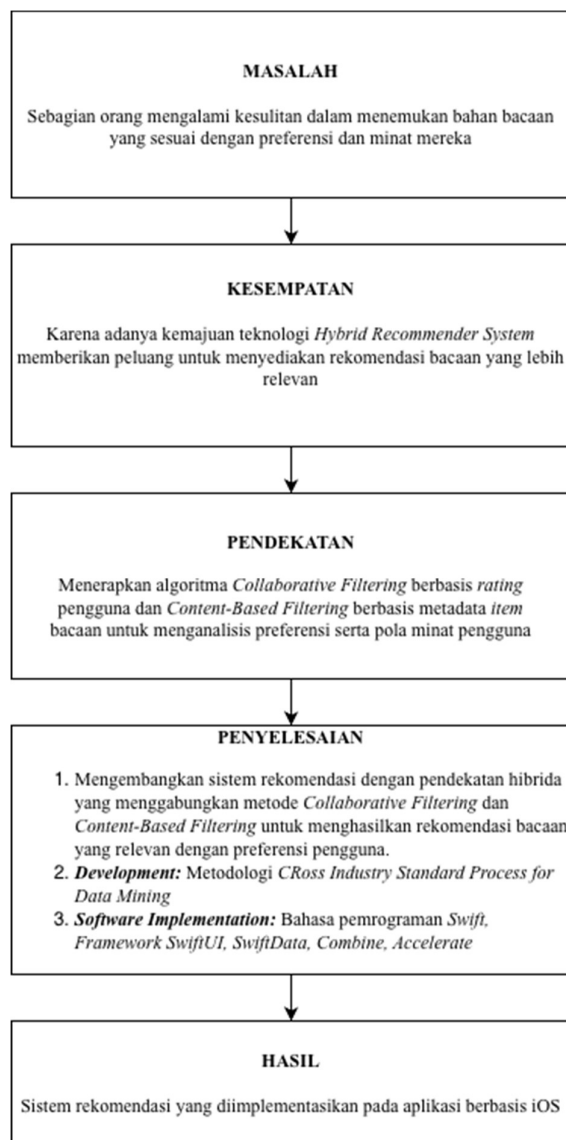
Batasan masalah pada penelitian ini sebagai berikut:

1. Penelitian ini berfokus pada pengembangan sistem rekomendasi dengan pendekatan hibrida menggunakan algoritma *Collaborative Filtering* dan *Content-Based Filtering*.
2. Evaluasi sistem rekomendasi dibatasi pada kinerja rekomendasi yang dihasilkan, menggunakan metrik *Hit Ratio* untuk menilai relevansi item bagi pengguna.
3. Aplikasi iOS hanya digunakan sebagai media implementasi.
4. Data yang digunakan dalam pengembangan sistem rekomendasi terbatas pada dataset *Book Recommendation Dataset* yang dirilis oleh *Möbius* pada platform Kaggle.

5. Sistem rekomendasi tidak mempertimbangkan faktor eksternal seperti tren pasar, promosi, atau ulasan sosial media yang dapat memengaruhi preferensi pengguna.
6. Pengujian sistem rekomendasi dilakukan dalam skala terbatas, sehingga kinerja pada skala pengguna yang lebih besar belum dianalisis.
7. Aspek keamanan dan privasi data pengguna tidak diperhatikan, fokus utama tetap pada performa rekomendasi.

1.5 Kerangka Pemikiran Penelitian

Kerangka pemikiran dalam penelitian ini ditunjukkan Gambar 1.1 dibawah ini.



Gambar 1.1 Kerangka pemikiran.

1.6 Sistematika Penulisan

Pada pembagian susunan sistem penulisan pada peneltian ini penulis membaginya menjadi lima bab yaitu:

BAB I : PENDAHULUAN

Bab ini mencakup latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan masalah, metodologi penelitian, kerangka pemikiran, dan sistematika penulisan.

BAB II : KAJIAN LITERATUR

Bab ini memuat penjelasan tentang penelitian terdahulu dengan tema serupa serta uraian mengenai teori-teori yang digunakan sebagai dasar untuk penelitian ini.

BAB III : METODE PENELITIAN

Bab ini mencakup penjelasan mengenai metode yang digunakan untuk memahami proses bisnis, menganalisis data penelitian, dan merancang sistem.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini membahas hasil implementasi metode pada data penelitian dan evaluasi sistem yang telah dibuat.

BAB V : PENUTUP

Bab ini memuat kesimpulan dari penelitian secara keseluruhan dan rekomendasi untuk pengembangan pada penelitian selanjutnya.