

BAB I PENDAHULUAN

1.1 Latar Belakang

Saat ini, seiring dengan pesatnya pertumbuhan jumlah buku yang tersedia, menemukan buku yang sesuai dengan minat dan preferensi pembaca dapat menjadi kesulitan dan tantangan tersendiri. Sebelum adanya internet, untuk mendapatkan rekomendasi buku yang sesuai dengan selera baca dan kebutuhan, kita harus datang ke perpustakaan dan bertanya kepada petugas perpustakaan [1]. Hal tersebut kemudian menimbulkan masalah baru, dimana rekomendasi dari mereka pastinya terbatas oleh pengetahuan mereka dan terbiaskan oleh selera personal. Lalu muncullah pengembangan Sistem Rekomendasi yang dapat menyajikan referensi seperti buku untuk dibeli, film untuk ditonton, barang untuk dibeli, dan lain-lain. Tujuan sistem rekomendasi adalah untuk menyajikan konten yang efektif dan relevan kepada pengguna di dalam sebuah *platform* [2].

Secara garis besar, ada dua jenis sistem rekomendasi, yaitu yang dipersonalisasi (*personalized*) dan tidak dipersonalisasi (*non-personalized*) [3]. Sistem rekomendasi yang dipersonalisasi membangun rekomendasi dari data riwayat dan preferensi pengguna, sedangkan sistem rekomendasi yang tidak dipersonalisasi menyediakan rekomendasi secara umum kepada semua pengguna dengan melihat apa yang sedang populer di sebuah *platform* [4]. Contoh dari sistem rekomendasi yang tidak dipersonalisasi adalah *Demographic Filtering*, sedangkan untuk sistem rekomendasi yang dipersonalisasi contohnya adalah *Content-based Filtering* dan *Collaborative Filtering*, serta penggabungan antara keduanya yaitu *Hybrid Filtering* [2].

Dalam pendekatan *Content-based Filtering*, preferensi pengguna memiliki peran penting. Profil preferensi pengguna akan dibangun berdasarkan riwayat interaksi pengguna dengan *item* di waktu lampau. Deskripsi dari *item* dijabarkan dengan bantuan kata kunci. Selanjutnya algoritma yang digunakan dalam Content-

based Filtering membandingkan profil pengguna dengan item lalu mencari item lain dengan angka kemiripan tertinggi yang nantinya direkomendasikan kepada pengguna [2], [5].

Ketika tidak ada data riwayat interaksi pengguna dengan item yang cukup, maka sistem tidak bisa membangun profil preferensi pengguna dan menyebabkan prediksi yang tidak akurat. Permasalahan tersebut sudah sangat familiar di dalam sistem rekomendasi sampai-sampai diberi nama sendiri, yaitu *Cold-start problem* [6]. Beberapa metode telah tersedia untuk mengatasi masalah tersebut, seperti dengan metode *Hybrid Filtering*, yaitu penggabungan antara *Content-based Filtering* dengan *Collaborative Filtering* untuk menutupi kekurangan dari kedua metode tersebut [7]. Atau dengan mengeksploitasi demografi dari data pengguna seperti gender, umur, negara, dan lain-lain sehingga mendapatkan karakteristik pengguna.

Beberapa peneliti telah menguji berbagai metode untuk mengatasi masalah *Cold-start*. Seperti penelitian oleh Natarajan et al. (2020) dengan metode *Linked Open Data* dengan bantuan database DBPedia [8]. Zhu et al. (2021) juga telah melakukan penelitian dengan menggunakan metode *Meta Scaling* dan *Meta Shifting Network* untuk *warm-up id embedding* [9]. Atau penelitian yang dilakukan oleh Nasy'an Taufiq Al Ghifari et al. (2021) yang menggunakan analisis bibliometrik dengan dataset artikel dari Scopus [10]. Ketiga penelitian tersebut berhasil memberikan hasil kinerja yang bagus dengan akurasi rekomendasi yang meningkat.

Penelitian ini akan berfokus pada pembuatan sistem rekomendasi *Content-based Filtering* dengan *dataset* buku. Penelitian ini juga akan menggunakan data pengguna yang diambil dari media sosial Threads untuk mengatasi masalah *cold-start*. Profil pengguna akan dibangun berdasarkan utas terakhir dari *username* yang tertentu. Lalu data buku dan profil pengguna akan diolah menjadi *vector* menggunakan salah satu metode *word embedding* yaitu *Word2Vec*. Lalu kedua

vector dihitung nilai kemiripannya, nilai yang didapat digunakan untuk mencari *item* yang paling relevan dengan utas pengguna [11].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana membuat sistem rekomendasi dengan metode *Content-based Filtering* pada dataset buku dengan algoritma *Word2Vec*?
2. Bagaimana mengatasi masalah *Cold-start* dengan menggunakan sumber data dari sosial media?

1.3 Batasan Masalah

Agar penelitian ini mencapai sasaran dan tujuan yang diharapkan maka penelitian ini diberikan batasan masalah sebagai berikut:

1. Menggunakan metode *Content-based Filtering* untuk membuat sistem rekomendasi.
2. Algoritma yang diujikan adalah *Word2Vec*.
3. Data buku diambil dari dari website *goodreads.com* dengan metode *scraping*.
4. Data buku yang digunakan hanya buku literatur Indonesia.
5. Data utas diambil dari pengguna media sosial *Threads* berbahasa Indonesia.
6. Data yang diambil hanya dari profil yang bersifat *public*.
7. Pengambilan data dibatasi maksimal 25 utas.
8. Implementasi tampilan website dilakukan dengan *framework* *Gradio* dan hanya dapat diakses dari server lokal.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh penulis dalam penelitian ini adalah menghasilkan sistem rekomendasi *Content-based Filtering* dengan algoritma *Word*

Embedding (Word2Vec).

1.5 Manfaat Penelitian

a. Manfaat Teoritis

Penelitian ini memberikan gambaran tentang upaya penerapan teknik *Content-based Filtering* dengan algoritma *word2vec* dan *Cosine Similarity* dalam perancangan sistem rekomendasi menggunakan dataset buku yang diambil dari website goodreads.

b. Manfaat Praktis

Penelitian dapat digunakan sebagai sumber referensi untuk peneliti lain dalam perancangan sistem rekomendasi buku dan bagaimana menangani *cold-start problem*.

1.6 Sistematika Penulisan

a. Bab I Pendahuluan

Bab I berisikan mengenai latar belakang permasalahan, rumusan masalah, batasan masalah, maksud dan tujuan dari pembahasan, metodologi penelitian dan sistematika penulisan.

b. Bab II Landasan Teori

Bab II berisi mengenai kajian pustaka dan dasar teori yang digunakan sebagai landasan dalam melaksanakan penelitian ini.

c. Bab III Metode Penelitian

Bab III berisi mengenai alat bahan yang dibutuhkan dan bagaimana alur penelitian dilakukan.

d. Bab IV Hasil dan Pembahasan

Bab IV berisi mengenai penjabaran hasil penelitian yang telah dilakukan.

e. Bab V Penutup

Bab V berisi mengenai kesimpulan dan saran yang dapat dilakukan bagi peneliti lain yang ingin meneliti hal yang sama kedepannya.