

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sistem rekomendasi (*Recommender System* atau RS) menjadi salah satu area penelitian yang penting saat ini karena membantu pengguna menemukan minat mereka di internet [1]. Minat di bidang ini tinggi karena merupakan bidang penelitian yang kaya akan masalah dan karena banyaknya aplikasi praktis yang membantu pengguna dalam menangani informasi. RS memiliki relevansi dengan pencarian informasi di berbagai bidang [2]. RS merekomendasikan banyak hal, biasanya RS digunakan untuk merekomendasikan informasi (URL, artikel berita), hiburan (buku, film, restoran), atau individu (pakar) [3]. Pada penelitian ini kami mendemonstrasikan sistem rekomendasi untuk merekomendasikan film.

Ada pendekatan yang umum untuk digunakan dalam rekomendasi film, *collaborative filtering* (CF), *content-based filtering* (CBF), dan *hybrid*. *Collaborative filtering* memprediksi item tertentu untuk pengguna aktif dari pendapat pengguna lain (*neighbor*) yang memiliki kesamaan selera seperti pengguna target. Item yang disukai *neighbor* kemudian direkomendasikan kepada pengguna target dengan harapan dia juga akan menyukainya. *Content-based filtering* memberikan rekomendasi dengan membandingkan representasi konten yang terdapat dalam suatu item, seperti *actor*, *genre*, *author*, dll yang menjadi minat pengguna dengan mengabaikan pendapat pengguna lainnya. Sedangkan *hybrid* adalah menggabungkan keduanya [4][5][6].

Sistem CBF menganalisa deskripsi item untuk mengidentifikasi item yang menarik bagi pengguna. Rekomendasi dapat dilakukan berdasarkan riwayat dari pengguna. Fitur item dan perilaku pengguna bersifat subjektif. Oleh karena itu representasi item serta pemodelan pengguna merupakan masalah utama dalam membangun CB-RS [5]. *Diversity problem* juga diakui sebagai kekurangan sistem CBF. *Diversity problem* adalah kondisi di mana hasil rekomendasi yang diberikan kurang beragam [1] dan hanya terbatas sesuai apa yang pengguna telah tonton atau disukai. *Collaborative filtering* (CF) dapat mengatasi kekurangan CBF. CF dapat merekomendasikan film dengan lebih beragam karena rekomendasi didapat dari

pengguna lain yang mirip. Namun hasil CF cenderung *random*, terkadang hasil rekomendasi tidak sesuai dengan yang pengguna inginkan [7].

Tujuan penelitian ini adalah untuk merekomendasikan film yang beragam (berdasarkan yang disukai pengguna lain) namun memiliki kemiripan preferensi sesuai dengan preferensi pengguna target. Selain itu kami menambahkan fitur demografi (usia, jenis kelamin, dan pekerjaan) pengguna, fitur rating item yang secara implisit juga merepresentasikan genre item yang disukai pengguna dan rating pengguna yang diberikan untuk item. Representasi rating pengguna dengan rating pada penelitian ini disebut dengan *genre interestingness measures* (GIM) untuk menambah akurasi kemiripan minat antar-pengguna aktif [5][8][9].

CF merupakan salah satu metode populer untuk memprediksi peringkat atau rating item berdasarkan kesamaan antar-pengguna melalui *Pearson's coefficient* di mana jarak antara dua pengguna hanya didasarkan pada rating yang diberikan pengguna. Disisi lain preferensi user tidak cukup akurat jika hanya didasarkan berdasarkan rating saja, beberapa konten deskripsi juga dibutuhkan [3][10]. Dua pengguna tidak dapat dikatakan sama hanya berdasarkan rating yang diberikan untuk item tertentu, tapi juga bergantung faktor lain yang mempengaruhi seperti latar belakang, informasi detail seperti usia, jenis kelamin, dan kesukaannya terhadap genre tertentu [8]. Untuk menghitung kesamaan antara pengguna berdasarkan data demografi dan GIM dapat dilakukan integrasi logika *fuzzy* dengan RS [5][8][11] yaitu menggunakan *Fuzzy-CF*, namun metode tersebut mengabaikan kemiripan rating antar pengguna dan atribut demografi lain. Untuk itu dibutuhkan metode untuk menambahkan similaritas dari sisi internal pengguna dengan eksternal berdasarkan interaksi pengguna dengan pengguna lain melalui riwayat rating [10].

Metode yang digunakan untuk penelitian ini menggunakan *Hybrid User-based Fuzzy Collaborative Filtering* (HU-FCF) untuk menghasilkan prediksi film terbaik sesuai kemiripan pengguna [5][8][10] yang kemudian dilanjutkan dengan tahap kedua dengan menggunakan *content-based filtering* (CBF) dengan teknik TF-IDF dan *Cosine Similarity* untuk mengurutkan kembali sesuai kemiripan genre film [7]. Penggunaan CBF tidak digabung dengan HU-FCF, tetapi dilakukan secara

bergantian. Hasil penggunaan metode ini dapat memberikan rekomendasi film yang beragam dengan memberikan rekomendasi yang juga disukai pengguna lain, serta hasil rekomendasi tersebut sesuai dengan genre film yang dipilih pengguna target.

1.2 Perumusan Masalah

Perumusan masalah pada penelitian ini adalah sebagai berikut:

1.2.1 Permasalahan

Metode *Fuzzy Collaborative filtering* tidak cukup untuk mencerminkan korelasi antar pengguna. Selain itu pengguna terkadang juga menginginkan film berdasarkan dengan genrenya.

1.2.2 Usulan Solusi

Penggunaan *Hybrid User-based Fuzzy Collaborative filtering* mengintegrasikan informasi internal dan eksternal pengguna untuk mendapatkan pengguna dengan ketertarikan paling mirip. Kemudian hasil rekomendasi HU-FCF diurutkan kembali sesuai genre film yang diinginkan menggunakan TF-IDF dan *Cosine Similarity*

1.2.3 Pertanyaan Penelitian

Pertanyaan penelitian adalah sebagai berikut:

1. Apakah metode *Hybrid User-based Fuzzy Collaborative filtering* (HU-FCF) dengan menggunakan rating, fitur demografis dan *genre interstingness measures* (GIM) lebih baik dari pada tradisional *Collaborative filtering* dengan hanya menggunakan rating saja?
2. Berapa nilai presisi rekomendasi menggunakan HU-FCF dengan kemiripan genre?

1.3 Tujuan dan Manfaat

Tujuan dan manfaat dari penelitian adalah sebagai berikut:

1.3.1 Tujuan

1. Mengetahui penerapan dan tingkat akurasi metode *Hybrid User-based Fuzzy Collaborative filtering* (HU-FCF) dalam merekomendasikan film
2. Mengetahui penerapan dan presisi metode HU-FCF dengan TF-IDF untuk mendapatkan rekomendasi dari hasil prediksi rating untuk semua pengguna,

namun kemiripan genre tetap dipertahankan sesuai item yang dipilih pengguna.

1.3.2 Manfaat

Manfaat dari penelitian ini adalah dapat membantu pengguna untuk mendapatkan rekomendasi film yang sesuai dengan preferensi pengguna.

1.4 Batasan Masalah

Pada penelitian ini terdapat batasan masalah agar pembahasan terarah dan tidak menyimpang dari tujuan penulisan. Adapun batasan masalahnya sebagai berikut:

1. Dataset yang digunakan didapat dari situs *open datasets* pada *MovieLens.org* dengan jumlah rating 100.000.
2. Metode yang digunakan pada sistem rekomendasi menggunakan *Hybird User-based Fuzzy-Collaborative filtering* dan dilanjutkan dengan kemiripan genre.
3. Metode kemiripan genre menggunakan TF-IDF dan *Cosine Similarity*
4. Demografi user menggunakan atribut usia, jenis kelamin, dan pekerjaan.
5. Tidak memperhatikan *cold start* dan *sparsity problem*
6. Fitur yang digunakan untuk rekomendasi kedua adalah genre film

1.5 Sistematika Penulisan

Sistematika penulisan Laporan Skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas latar belakang mengenai ruang lingkup yang dibahas, solusi untuk menyelesaikan permasalahan, tujuan dan manfaat dengan diterapkannya solusi yang diusulkan, batasan, serta sistematika penulisan laporan.

BAB II KAJIAN PUSTAKA

Bab ini berisikan teori mengenai metode yang dipilih dalam penelitian serta penelitian terkait yang telah ada sebelumnya.

BAB III METODE USULAN

Pada bab ini dijelaskan mengenai dataset yang akan digunakan, arsitektur sistem, algoritma metode usulan, metode penelitian, dan pengujian sistem yang akan digunakan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini dijelaskan mengenai hasil implementasi sistem sesuai arsitektur sistem, skenario dan hasil uji coba.

BAB V PENUTUP

Pada bab ini berisi tentang kesimpulan dan saran berdasarkan hasil uji coba yang telah dilakukan sebagai pengembangan penelitian terkait untuk berikutnya.