

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam kehidupan sehari-hari, orang-orang mengandalkan rekomendasi dari orang lain yang berupa kata-kata dari orang tersebut, berita dari media sosial maupun media elektronik, *survey*, dan yang lainnya untuk memilih atau membeli barang yang diinginkan. Sistem rekomendasi dapat membantu memilih buku, artikel, film, musik, barang elektronik, dan yang lainnya untuk menemukan barang atau informasi yang paling menarik dan bermanfaat.

Content-based filtering dan *collaborative filtering* adalah pendekatan yang sering digunakan dalam sistem rekomendasi[1]. *Content-based filtering* membuat profil untuk user berdasarkan item yang sebelumnya pernah di *rate* oleh user. Sehingga sistem dapat merekomendasikan item baru yang belum pernah di *rate* oleh user berdasarkan profil yang sudah ada. Pada *collaborative filtering* sistem mengaitkan user dengan grup user yang memiliki kesamaan yang sama dengan user tersebut[1]. Sehingga sistem bisa menghasilkan rekomendasi berdasarkan rating user pada item yang pernah di *rate*.

Salah satu masalah yang dihadapi pada sistem rekomendasi adalah *cold-start problem*, di mana pada sistem rekomendasi yang baru saat user dan item yang terdapat pada sistem tersebut masih sedikit maka akan susah untuk memberikan rekomendasi. *Cold-start problem* dapat dikategorikan dalam tiga tipe yaitu merekomendasikan item yang sudah ada untuk user baru, merekomendasikan item baru untuk *user* yang sudah ada, dan merekomendasikan item baru untuk user baru.

Banyak algoritma sudah dibuat untuk mengatasi masalah *cold-start* seperti Pazzani[11] mengajukan *hybrid method* untuk merekomendasikan item berdasarkan voting dari empat algoritma: *user-user collaborative filtering*, *content-based*, *demographic-based*, dan *collaboration via content*, Basilico dan Hofmann[8] mengajukan *online perceptron algorithm* bersama dengan kombinasi *multiple kernel functions* yang menyatukan *collaborative* dan *content-based filtering*. Tetapi metode tersebut tidak langsung mengatasi masalah *cold-start* itu sendiri[12], karena target user setidaknya sudah melakukan beberapa rating terhadap item yang akan direkomendasikan. *Pairwise preference regression* merupakan salah satu metode yang langsung menyerang masalah *cold-start*. *Pairwise preference regression* dapat memberikan rekomendasi tidak hanya pada user yang tidak memiliki *historical rating* tetapi juga hanya memiliki sedikit informasi demografi. Data yang digunakan untuk studi kasus ini adalah data yang user-nya memiliki informasi demografis seperti *gender*, *age*, dan *occupation*, itemnya memiliki informasi seperti tahun rilis item dan genre, dan data rating user terhadap item misalnya data *movie rating*.

1.2 Perumusan Masalah

Masalah yang akan diangkat pada penelitian ini adalah bagaimana membangun sistem rekomendasi yang dapat memanfaatkan user rating, user atribut dan item atribut menggunakan *pairwise preference regression* untuk memberikan rekomendasi item yang sudah ada kepada user baru, item baru kepada user yang sudah ada, dan item baru kepada user baru.

1.3 Tujuan

Tujuan dari penelitian ini adalah :

1. membangun profil untuk pasangan user/item dari user, item dan rating,
2. membangun *predictive model* yang dapat memberikan rekomendasi film,
3. mengevaluasi performansi dari sistem rekomendasi menggunakan pendekatan Normalized Discounted Cumulative Gain (nDGC).

1.4 Batasan Masalah

Batasan masalah dari penelitian ini adalah :

1. data yang digunakan ialah data *movie set* dari MovieLens,
2. fokus penelitian adalah pada penanganan *cold-start problem* menggunakan *pairwise preference regression*.

1.5 Metodologi penyelesaian masalah

Metodologi penyelesaian masalah pada penelitian ini adalah sebagai berikut:

- a. Mengumpulkan referensi dan studi literatur
Penelitian ini dimulai dengan proses pengumpulan referensi mengenai *collaborative filtering*. Kemudian referensi tersebut dipelajari, dipahami serta dianalisis, masalah apa yang muncul pada penelitian sebelumnya. Setelah *pairwise preference regression* terpilih sebagai solusi dari masalah *cold-start*, dipelajari lebih dalam lagi agar dapat menemukan solusi dari masalah yang timbul.
- b. Pendefinisian masalah dan studi kelayakan
Masalah yang muncul pada penelitian sebelumnya yaitu susahya memberikan rekomendasi kepada user baru ataupun item baru karena belum adanya *historical rating* dari user tersebut ke item yang akan direkomendasikan. Kemudian dianalisis prioritas dari masalah ini, apakah sangat penting atau tidak.
- c. Analisis dan perancangan kebutuhan sistem
Perancangan sistem dimulai dari mengumpulkan data user, item set, dan rating user yang kemudian data user dan item set dijadikan pasangan vektor menggunakan *pairwise preference regression*.
- d. Pengumpulan dan penyusunan dataset
Basis data yang digunakan pada kasus ini adalah data *movie set* dari MovieLens. Data set ini dipilih karena berisi informasi user (informasi demografi) dan film yang nantinya akan digunakan dalam pemberian rekomendasi.
- e. Implementasi sistem
Implementasi sistem dilakukan untuk membangun sebuah sistem yang dapat merekomendasikan item kepada user baru. Dimulai dengan membuat profil pasangan item dan user, selanjutnya membuat model

prediktif menggunakan *pairwise preference regression* untuk membuat rekomendasi item kepada user.

- f. Pengujian, analisis dan evaluasi
Pengujian dilakukan terhadap dataset yang telah disiapkan sebelumnya. Dataset berupa data user, item set, dan rating user. Pengujian menggunakan pendekatan nDCG. Hasil pengujian dianalisis serta dievaluasi untuk menyimpulkan hasil penelitian.
- g. Penyusunan laporan tugas akhir
Sebagai bentuk dokumentasi dari penelitian ini disusunlah laporan tugas Akhir yang berupa buku TA.

1.6 Sistematika Penulisan

Berikut adalah sistematika penulisan dari tugas akhir ini :

- a. Pendahuluan
Pada bab ini terdiri dari latar belakang, perumusan masalah, tujuan, batasan masalah, metodologi penyelesaian masalah dan sistematika penulisan tugas akhir ini.
- b. Tinjauan Pustaka
Bab ini menjelaskan teori-teori pendukung yang digunakan untuk menyelesaikan masalah pada tugas akhir ini seperti *Collaborative Filtering* dan *Pairwise Preference Regression*
- c. Analisis Perancangan Sistem
Pada bab ini membahas tentang analisis kebutuhan sistem dan perancangan sistem *Pairwise Preference Regression*.
- d. Pengujian dan Analisis
Pada bab ini menguraikan bagaimana hasil pengujian sistem yang telah dibangun dan analisis terhadap hasil pengujian sistem yang diperoleh.
- e. Kesimpulan dan Saran
Bab ini berisi kesimpulan yang didapat dari hasil penelitian tugas akhir ini dan saran untuk penelitian selanjutnya.