

BAB I

PENDAHULUAN

1.1. Latar Belakang

Penelitian seputar sistem rekomendasi dalam satu dekade terakhir ini berkembang dengan pesat. Lebih dari 200 artikel penelitian yang membahas tentang sistem rekomendasi sudah terpublikasi (Beel, 2013). Seiring bertambahnya artikel-artikel tersebut, algoritma rekomendasi yang dikembangkan dan diinovasikan sesuai dengan kebutuhan masing-masing sistem pun semakin banyak. Berdasarkan sudut pandang penyaringan informasinya, algoritma rekomendasi yang sering digunakan dalam sistem rekomendasi terbagi menjadi tiga jenis, yaitu : *Content-Based Filtering* (CB), *User-Based Collaborative Filtering* (UCF) dan *Item-Based Collaborative Filtering* (ICF).

Beberapa peneliti mengembangkan ketiga algoritma penyaringan tersebut secara terpisah, namun ada juga peneliti yang menggabungkan lebih dari satu algoritma dengan tujuan mendapatkan hasil rekomendasi yang terbaik (Hu et al., 2011), (De Campos et al., 2010), (Chen et al. 2010). Untuk mengetahui apakah algoritma yang dikembangkan mendapatkan hasil yang terbaik atau tidak, para peneliti menggunakan beberapa teknik evaluasi seperti *Mean Absolute Error* (MAE), *Precision Value*, *Recall*, dan lainnya. (Beel, 2013), (Jin, 2006), (Claypool, 1999).

Dalam penerapannya pada sistem rekomendasi, UCF seringkali dipasangkan dengan ICF karena memiliki banyak kesamaan dalam teknik

penyaringan dan bahkan dalam hal kekurangan dan kelebihan. UCF mengasumsikan bahwa untuk mencari sesuatu yang akan disukai oleh pengguna A, maka harus dicari pengguna lain dengan kecenderungan yang sama. Pengguna lain yang memiliki ketertarikan / kecenderungan yang sama dengan pengguna A disebut dengan *neighbor*. Segala sesuatu yang disukai oleh *neighbor* akan menjadi sumber rekomendasi untuk pengguna A. Sedikit berbeda dengan UCF, ICF menampung segala *item* yang telah disukai oleh pengguna A terlebih dahulu dan kemudian mulai menelusuri *item* lain yang akan disukai oleh pengguna A. Perbedaan ICF dan UCF terletak pada titik awal dalam mencari rekomendasi untuk seorang pengguna dan jenis *neighbor*-nya. (Lops et al., 2011), (Sarwar et al., 2001).

Peran *neighbor* sangatlah penting dalam sistem rekomendasi, selain berperan sebagai sumber rekomendasi, ada juga data lain pada *neighbor* yang nantinya akan digunakan untuk menghitung nilai prediksi. Dalam hal ini data tersebut adalah nilai *rating neighbor* terhadap *item* rekomendasi. Nilai prediksi ini nantinya digunakan oleh para peneliti sebagai tolak ukur seberapa “baik” algoritma rekomendasi yang sedang dikembangkan.

Pada penelitian ini, peneliti mencoba memanfaatkan sebuah metode pengukuran dengan tujuan bisa mendapatkan *neighbor* terbaik, sebagai dasar pencarian rekomendasi dan nilai prediksi yang akurat. Metode pengukuran yang digunakan adalah *Pearson-Correlation Coefficient*. Dengan memanfaatkan metode pengukuran ini, peneliti ingin membuktikan seberapa besar pengaruh relasi *neighbor* terhadap keakuratan rekomendasi dan prediksi.

Keakuratan rekomendasi dan prediksi dari penelitian ini akan dievaluasi dengan menggunakan metode evaluasi Mean Absolute Error (MAE), *Precision* dan *Recall Value*. Metode MAE digunakan untuk mengukur seberapa kuat akurasi dari nilai prediksi yang diberikan oleh algoritma peneliti dengan nilai *rating* yang sudah diberi oleh pengguna (evaluasi secara *offline*). *Precision* dan *Recall Value* berfungsi untuk mengetahui seberapa banyak rekomendasi *item* yang “tepat” diberikan untuk penerima.

1.2. Perumusan Masalah

Dari latar belakang permasalahan tersebut diatas, dapat diambil rumusan masalah sebagai berikut :

1. Bagaimana menerapkan metode *Pearson-Correlation* kedalam sistem rekomendasi?
2. Apakah nilai relasi antara pengguna terhadap *neighbor*-nya dapat mempengaruhi keakuratan rekomendasi dan prediksi?

1.3. Batasan Masalah

Pada penelitian ini lingkup permasalahan dibatasi sebagai berikut :

1. Teknik pengukuran yang digunakan untuk menentukan relasi antar pengguna sistem adalah *Pearson-Correlation*
2. Algoritma yang digunakan untuk menghitung nilai prediksi rekomendasi adalah *User-Based Collaborative Filtering* dengan melibatkan nilai kedekatan pengguna dengan *neighbor*-nya.

3. Teknik evaluasi yang digunakan untuk mengukur keakuratan nilai prediksi adalah *Mean Absolute Error* (MAE).
4. *Precision, Recall Value* digunakan untuk mengukur keakuratan rekomendasi.

1.4. Manfaat Penelitian

Penelitian ini bermanfaat untuk para peneliti ataupun *developer* yang ingin mengembangkan sistem rekomendasi. Hasil evaluasi dari algoritma prediksi, algoritma rekomendasi serta pencarian *neighbor* dapat digunakan sebagai bahan tinjauan untuk penelitian di masa mendatang.

1.5. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui apakah pemanfaatan *Pearson-Correlation* dalam pencarian *neighbor* dan melibatkannya dalam proses penawaran rekomendasi serta penghitungan nilai prediksi dapat mempengaruhi kualitas rekomendasi.

1.6. Sistematika Penulisan

Penulisan tesis ini dengan judul : *User-Based Collaborative Filtering Dengan Memanfaatkan Pearson-Correlation Untuk Mencari Neighbor Terdekat Dalam Sistem Rekomendasi*, disusun dalam lima bab, dengan sistematika penulisan sebagai berikut.

Bab pertama Pendahuluan. Bab ini menjelaskan secara singkat isi dari tesis yang berisi latar belakang, perumusan masalah, batasan masalah dan manfaat penelitian serta tujuan penelitian.

Bab kedua Tinjauan Pustaka. Bab ini akan menjelaskan teori yang mendukung penyelesaian tesis yang berisi uraian penelitian pendahuluan, algoritma *User-Based Collaborative Filtering*, *Pearson-Correlation Coefficient*, *Mean Absolute Error*, *Precision* dan *Recall Value*.

Bab ketiga Metodologi Penelitian. Bab ini menjelaskan tentang bahan atau materi penelitian, data dan langkah-langkah penelitian.

Bab keempat Analisis dan Perancangan Sistem. Bab ini memberikan penjelasan tentang analisis dan desain perangkat lunak yang akan dibangun, serta sistem yang akan diterapkan.

Bab kelima Implementasi dan Pengujian Sistem. Dalam bab ini akan dibahas implementasi dari teknik pengukuran *Pearson-Correlation* untuk mencari *neighbor* dalam sistem rekomendasi yang memanfaatkan *User-Based Collaborative Filtering* dan mengevaluasi hasil yang didapat dengan menggunakan *Mean Absolute Error* dan *Precision Value*.

Bab keenam Kesimpulan dan Saran. Bab ini berisi kesimpulan penelitian serta saran-saran yang mungkin diberikan untuk pengembangan lebih lanjut.