

BAB I PENDAHULUAN

1.1 Latar Belakang

Salah satu perkembangan teknologi yang sangat berkembang saat ini adalah sistem rekomendasi. Sistem rekomendasi (*recommender system*) adalah aplikasi pemberi rekomendasi berupa informasi yang dibutuhkan oleh pengguna berdasarkan *feedback* dari pengguna lain [1]. Sistem rekomendasi yang sering dijumpai adalah sistem rekomendasi film. Film banyak digemari oleh masyarakat. Namun, banyaknya informasi film di internet membuat seseorang kebingungan untuk menentukan film yang akan ditontonnya. Untuk itu dibutuhkan suatu sistem rekomendasi yang membantu *user* untuk mendapatkan informasi mengenai film yang mungkin disukainya [2].

Sistem rekomendasi merupakan sistem yang baru, apabila dibandingkan dengan sistem pendukung keputusan (SPK). Sistem pendukung keputusan adalah sistem yang membantu *user* untuk pengambilan keputusan tentang suatu hal. Sedangkan rekomendasi adalah sistem yang digunakan untuk merekomendasikan sesuatu. Biasanya ditujukan untuk individu yang kurang pengetahuan mengenai hal tersebut. Perbedaannya adalah SPK hanya memberikan 1 keputusan sedangkan rekomendasi membantu *user* untuk memberikan keputusan dengan beberapa pilihan yang diberikan oleh sistem rekomendasi.

Sistem rekomendasi ini membutuhkan suatu data *rating*, yang kemudian diolah untuk menghasilkan daftar rekomendasi. Data yang telah diperoleh dapat berbentuk data eksplisit atau implisit. Data yang akan digunakan adalah data eksplisit. Data eksplisit adalah data yang cara mendapatkannya jelas. Contohnya *user* yang memberikan nilai secara langsung (*rating*) kepada suatu item, baik itu film, musik atau lainnya [2]. Data eksplisit ini bisa didapatkan secara gratis di *movielens* (<https://grouplens.org/datasets/movielens/100K/>). Data yang digunakan pada penelitian ini adalah 100.000 *rating*.

Sistem rekomendasi sangat bermanfaat untuk *e-commerce* (toko online) dalam merekomendasikan produk-produknya kepada *user*. Adapun website yang telah menerapkan sistem rekomendasi, yaitu: *ebay.com*, *yahoo.com* yang dikenal dengan *myYahoo*, *amazon.com* dan sebagainya [3].



Sistem rekomendasi membutuhkan metode untuk memudahkan dalam merekomendasikan item. Metode yang cukup terkenal adalah *collaborative filtering*. Sistem *collaborative filtering* adalah metode yang digunakan untuk memprediksi kegunaan item berdasarkan penilaian pengguna sebelumnya [4]. *Collaborative filtering* merupakan salah satu algoritma untuk menyusun *recommender system* dan telah terbukti memberikan hasil yang sangat baik. Beberapa pendekatan yang dapat digunakan dalam *collaborative filtering* adalah *memory based*, *model-based* dan *hybrid recommendation system* [5]. *Memory based* sering dikenal dengan metode *collaborative filtering* tradisional. Pada penelitian ini lebih fokus pada metode *collaborative filtering* tradisional. Keunggulan dari *collaborative filtering* tradisional adalah kemudahan dalam implementasi dan sangat efektif [5]. Selain itu teknik ini memiliki kelebihan yaitu sistem tidak memerlukan konten untuk dianalisa, tidak memerlukan banyak waktu untuk pengembangan bahasa, analisa dokumen, pengembangan bahasa *parsing* dan *stemming* [6].

Collaborative filtering tradisional dibagi menjadi 2, yaitu : metode *user-based (user-to-user)* dan *item-based collaborative filtering (item-to-item)*. Namun metode ini memiliki kekurangan yaitu performa akan menurun seiring bertambahnya data yang menyebabkan skalabilitas [7]. Untuk itu diterapkan teknik *clustering* untuk mengatasi skalabilitas. Untuk teknik *clustering* bisa menggunakan metode *k-means clustering*. Algoritma *k-means clustering* sangat baik digunakan ketika data dalam skala besar. Algoritma ini lebih baik jika dibandingkan dengan algoritma SOM dan *hierarchical* [8]. Namun *k-means clustering* juga memiliki kelemahan yaitu ditemukannya beberapa model *clustering* yang berbeda dan tergantungnya akurasi terhadap jumlah *cluster* [9]. Adapun penelitian yang telah menerapkan penggabungan *clustering* dan metode tradisional yaitu metode gabungan *user clustering* dan *user-based collaborative filtering*. Hasil penelitian ini menunjukkan bahwa metode gabungan *user clustering* dan *user-based collaborative filtering* lebih akurat dibandingkan *collaborative filtering* tradisional [10]. Seperti yang telah dijelaskan metode tradisional *collaborative filtering* ada 2, yaitu: *user-based* dan *item-based collaborative filtering*. Metode *item-based collaborative filtering* ini lebih akurat dibandingkan metode *user-based collaborative filtering* [11]. Untuk itu pada penelitian ini akan membuat sistem rekomendasi film secara *offline* dengan menggunakan metode *item clustering* dan *item-based collaborative filtering*. Metode *item clustering*



digunakan untuk mengelompokkan item. Sedangkan metode *item-based collaborative filtering* digunakan untuk mencari item yang *similar* (sama), memprediksi *rating* dan merekomendasikan item.

Dari uraian yang telah dijelaskan di atas, diharapkan penerapan metode gabungan *item clustering* dan *item-based collaborative filtering* lebih akurat dibandingkan dengan metode *user-based* dan *item-based collaborative filtering*. Dan metode gabungan *item clustering* dan *item-based collaborative filtering* lebih akurat dibandingkan metode gabungan *user clustering* dan *user-based collaborative filtering* karena metode *item-based collaborative filtering* lebih akurat dibandingkan metode *user-based collaborative filtering*.

1.2 Rumusan Masalah

Dari latar belakang diatas, dapat diambil rumusan masalah, yaitu:

1. Apakah metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering* lebih akurat dibandingkan metode *User-Based* dan *Item-Based Collaborative Filtering*?
2. Apakah metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering* lebih akurat dibandingkan metode gabungan *User Clustering* dan *User-Based Collaborative Filtering*?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Tujuan dari penelitian ini yaitu:

1. Untuk mengetahui perbandingan akurasi dari metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering* dengan metode *User-Based* dan *Item-Based Collaborative Filtering*
2. Untuk mengetahui perbandingan akurasi dari metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering* dengan metode gabungan *User Clustering* dan *User-Based Collaborative Filtering*

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini yaitu:

1. Dapat mengetahui perbandingan akurasi dari metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering*



dengan metode *User-Based* dan *Item-Based Collaborative Filtering*

2. Dapat mengetahui perbandingan akurasi dari metode gabungan *Item Clustering* dan *Item-Based Collaborative Filtering* dengan metode gabungan *User Clustering* dan *User-Based Collaborative Filtering*

1.4 Batasan Masalah

Penelitian ini mempunyai batasan masalah yaitu:

- a. Data yang digunakan pada penelitian ini adalah kumpulan data *movielens* (<https://grouplens.org/datasets/movielens/100K/>)
- b. Kriteria yang digunakan pada matriks *rating user-item* adalah *item* dan kriteria pada matriks *rating item-user* adalah *user*
- c. Sistem rekomendasi secara *offline*
- d. Menggunakan aplikasi matlab 2015b
- e. Tidak menghitung *cold-start problem* baik pada *user* atau *item*
- f. Pengukuran akurasi rekomendasi menggunakan *precision*, *recall* dan *f1score*
- g. Pengujian sistem rekomendasi menggunakan 5 *fold*

1.5 Metodologi Pengerjaan Tugas Akhir

Penelitian ini dilakukan dalam beberapa tahapan, sebagai berikut:

a. Studi Literatur

Pada tahap studi literatur ini, mengumpulkan dan mempelajari dari berbagai literatur yaitu buku, jurnal, skripsi ataupun paper yang berhubungan dengan Sistem Rekomendasi, *Collaborative Filtering*, Metode *Item-based Collaborative Filtering* dan Metode *K-means Clustering*.

b. Pengumpulan Data

Dalam sistem rekomendasi tentunya membutuhkan data yang nantinya akan diolah. Data yang digunakan adalah data *rating* eksplisit (*rating user*). Data yang digunakan didapat dari *movielens*.

c. Analisa dan Perancangan Sistem

Melakukan analisa pada sistem yang akan dibuat berdasarkan studi literatur dan data yang telah didapat. Setelah menganalisa, kemudian merancang sistem yang akan dibangun.



d. Implementasi Sistem

Pada tahap implementasi ini akan mengimplementasikan dari perancangan yang sudah dibuat sebelumnya. Pengimplementasian ini harus sesuai dengan perancangan yang telah dibuat agar tujuan dari pembuatan sistem terwujud.

e. Evaluasi dan Pengujian Sistem

Setelah tahap implementasi selesai maka ada 2 hal yang harus dilakukan, yaitu: menguji sistem dan evaluasi sistem. Pengujian sistem ini penting karena sistem yang terkadang kita anggap telah sempurna ternyata ada saja kesalahan yang mungkin ada. Untuk itu pengujian ini penting. Untuk pengujian menggunakan *k-fold cross validation*. Setelah itu lakukan evaluasi sistem yaitu mengecek apakah sistem yang dibuat telah sesuai harapan dan tidak ada yang eror. Jika sudah sempurna lakukan tahapan selanjutnya.

f. Dokumentasi dan Penarikan Kesimpulan

Pada tahap ini dilakukan dokumentasi hasil analisa dan perancangan yang digunakan untuk dituangkan ke dalam laporan akhir dalam bentuk skripsi.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan pada laporan penelitian ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Pada bagian ini menjelaskan dasar-dasar dari penulisan laporan tugas akhir, yaitu tentang latar belakang penelitian, rumusan masalah, tujuan dan manfaat, batasan masalah, metodologi penelitian, dan sistematika penulisan.

BAB II : STUDI PUSTAKA

Bagian ini berisi teori-teori penunjang dalam pengerjaan sistem dan penulisan laporan tugas akhir ini. Teori-teori tersebut meliputi sistem rekomendasi, *collaborative filtering*, metode *item-based collaborative filtering*, *k-means clustering*, dll.

BAB III : ANALISA DAN PERANCANGAN SISTEM

Bagian ini menguraikan desain atau rancangan dari sistem yang akan dibuat. Mulai dari kebutuhan *software* dan *hardware* yang dibutuhkan pada sistem. Analisa tersebut terdiri dari arsitektur input proses output, *diagram system*.



BAB IV : IMPLEMENTASI

Bagian ini akan menjelaskan tentang implementasi sistem yang telah dibuat. Terdapat *source code* dari beberapa bagian program yang penting. Dan melakukan pengujian sistem berdasarkan skenario pengujian yang telah ditentukan.

BAB V : PENUTUP

Bagian ini berisi kesimpulan dari hasil uji coba yang telah dilakukan serta saran yang dapat digunakan untuk pengembangan sistem rekomendasi selanjutnya.

