

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring dengan perkembangan waktu, teknologi saat ini sangat berkembang pesat di era industri 4.0. Kebutuhan informasi untuk manusia yang berbentuk fisik sudah berubah menjadi bentuk yang digital, hal ini merupakan representasi yang ada di era digital. Internet menyediakan berbagai sumber informasi berupa data digital, informasi yang diberikan disesuaikan dengan kebutuhan pengguna. Selain itu informasi berupa data digital juga mudah dalam mengaksesnya, oleh karena itu pengguna Internet meningkat setiap hari[1]. Akibatnya, jumlah data yang ditransfer melalui Internet juga meningkat. Namun, cara cepat dan efektif untuk memperoleh informasi yang berguna bagi pengguna dalam informasi yang sangat besar telah menjadi masalah yang mendesak untuk dipecahkan[2]. Banyaknya tantangan di era digital, membuat seiring berjalannya waktu memunculkan sebuah mesin rekomendasi yang nantinya akan membantu dalam melakukan pencarian sebuah informasi. Didalam mesin rekomendasi terdapat sebuah sistem yaitu sistem rekomendasi, sistem ini nantinya akan bertanggung jawab penuh atas mesin rekomendasi. Sistem rekomendasi merupakan sebuah sistem yang bekerja dengan cara menyaring fragmen informasi penting dari sejumlah besar informasi yang dihasilkan secara dinamis dalam memberi sebuah rekomendasi kepada pengguna atau user

yang berkaitan dengan item-item yang mungkin diminati[3]. Sistem rekomendasi mempunyai keuntungan bagi penyedia dan pengguna, serta pengurangan biaya transaksi dalam menemukan dan memilih barang, sistem rekomendasi juga terbukti meningkatkan kualitas dan pengambilan keputusan[4].

Terdapat 3 teknik dalam sistem rekomendasi, yaitu collaborative filtering, content-based filtering, dan hybrid filtering (penggabungan kolaboratif dan berbasis konten)[5]. Collaborative filtering menjadi salah satu metode yang biasa digunakan pada sistem rekomendasi. Cara kerja dari metode ini adalah dengan menggunakan penilaian yang diberikan oleh user lain berupa rating atau feedback dalam memprediksi item yang mungkin disukai. Dengan adanya filtering akan membantu seseorang untuk memperoleh sesuatu yang diminati. Collaborative filtering sudah digunakan dalam aplikasi platform sosial untuk menonton film yaitu Netflix. Di Netflix, mereka menggunakan sistem rekomendasi untuk personalized videos ranker (PVR), top-n ranker videos, trending now, continue to watching, video similarity, row selection dan ranking, evidence, search and related work[6]. Selain Netflix, web e-commerce seperti, Tokopedia, Amazon, dan juga aplikasi pada web buku bacaan online yang sudah menggunakan sistem ini adalah Google Books, Bookrix, Open Library, Good Reads, dan Smashwords [5].

Collaborative filtering dibagi menjadi dua pendekatan, yaitu memory-based collaborative filtering dan model-based collaborative filtering[6]. Model-based adalah pendekatan terbaik untuk menangani beberapa permasalahan di sistem rekomendasi. Permasalahan yang pertama adalah data sparsity yang mana banyak pada sistem rekomendasi komersial didasarkan pada data set yang berukuran besar, mengakibatkan matrik user-item yang akan digunakan dalam melakukan collaborative filtering sangat besar dan jarang, sehingga hal ini menjadi tantangan dalam hasil rekomendasi lalu bisa mengakibatkan error pada proses perhitungan matriks[2]. Masalah yang kedua yaitu skalabilitas yang akan mempengaruhi pada proses jalannya perhitungan dalam menghasilkan rekomendasi. Masalah yang ketiga ialah cold start problem yang mana setiap pengguna dan item baru mulai memasuki sistem, karena pada metode collaborative filtering bekerja berdasarkan interaksi pengguna dan item[7]. Contoh algoritma dari model-based adalah Alternating Least Squares (ALS), Singular Value Decomposition (SVD), dan Stochastic Gradient Descent[8]. Berdasarkan literatur, collaborative filtering tradisional tidak dapat menangani sparsity sebaik model-based. Dari beberapa metode tersebut salah satunya lebih cocok berdasarkan akurasi untuk mengatasi sparsity pada tiga dataset karakteristik yang berbeda[9].

Dalam penelitian ini penulis akan melakukan perbandingan empat model berbasis algoritma yaitu Alternating Least Squares (ALS), Singular Value Decomposition (SVD), Alternating Least Squares Weight Regularization (ALS-WR) dan SVD++ menggunakan beberapa karakteristik

dataset publik yang berbeda seperti MovieLens, Jester dan BookCrossing. Alasan penulis melakukan penelitian ini adalah untuk menemukan algoritma mana yang terbaik untuk menangani data sparsity dan algoritma mana yang cocok untuk setiap dataset yang ada.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah disampaikan, permasalahan yang akan dirumuskan yaitu bagaimana langkah melakukan eksperimen menggunakan empat algoritma dengan dua algoritma dasar dan dua algoritma terupdate serta untuk mengetahui bagaimana hasil perbandingan dari empat algoritma yang digunakan penulis sehingga bisa mengetahui dataset mana yang dapat ditangani sparsitynya.

1.3 Batasan Masalah

Dalam sebuah melakukan penelitian agar lebih rinci dan terarah, maka dari itu perlu adanya batasan masalah. Adapun batasan tersebut adalah sebagai berikut:

1. Menggunakan dataset MovieLens 1M, Jester 1M, dan Bookcrossing 1M
2. Pengujian mesin dilakukan melalui Google Collabs
3. Metode yang di gunakan pada sistem rekomendasi ialah Collaborative filtering dengan empat algoritma yaitu alternating least squares (ALS), singular value decomposition (SVD), Alternating Least Squares weight regularization (ALS-WR) dan SVD++.

1.4 Maksud dan Tujuan Penelitian

Menerapkan metode Collaborative Filtering menggunakan algoritma Alternating Least Squares (ALS), Singular Value Decomposition (SVD), Alternating Least Squares Weight Regularization (ALS-WR) dan SVD++. Sehingga mendapatkan hasil perbandingan dari rekomendasi sistem dan dapat mengetahui tingkat keakuratan tiap metode.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat untuk melakukan pengembangan sebuah sistem rekomendasi film dan juga dapat meningkatkan kualitas dari sistem tersebut.

1.6 Sistematika Penulisan

Tahapan dalam penelitian yang akan dilakukan, adalah sebagai berikut:

1.6.1 Metode Pengumpulan Data

Metode yang akan dilakukan dalam pengumpulan data antara lain:

1. Studi Literatur

Dalam tahapan ini, dilakukan pengumpulan literatur seperti jurnal ilmiah, buku, artikel, serta berbagai sumber informasi lainnya yang berkaitan dengan topik dalam penelitian ini. Sehingga dapat membantu untuk menambah pengetahuan tentang penelitian yang akan dilakukan serta dapat dijadikan sebagai referensi.

2. Metode Observasi

Dalam tahapan ini penulis melakukan pengumpulan data melalui website.

<https://www.kaggle.com> yaitu dataset movie, jester joke, dan bookcrossing.

1.6.2 Metode Analisis

Merupakan tahapan dari pendefinisian masalah yang akan dibangun.

Tahapan yang dilakukan dalam metode ini adalah:

1. Pengumpulan data

Mengambil dataset dari tiga sumber dataset yang berbeda terdiri dari MovieLens 1M dimana data tersebut berisi 1.000.209 ratings dan 4.000 Item dari 6.000 user. Dataset kedua Jester Joke berisi 1.761.439 ratings dan 150 item dari 59.132 user. Dataset ketiga dari Book Crossing yang berisi 1.149.780 ratings dan 271.379 item dari 278.858 user.

2. Membagi Dataset

Dalam tahapan ini dataset dibagi menjadi dua bagian yaitu, Data Training dan Data Testing.

3. Metode pengujian Model

Dalam tahapan ini dilakukan dengan cara mengukur kinerja dari metode yang akan digunakan yaitu ALS, ALS-WR, SVD, SVD++.

1.7 Sistematika Penulisan

Tahap ini digunakan untuk memudahkan serta mengetahui pembahasan yang di tulis oleh penulis pada skripsi ini secara menyeluruh. Adapun penulisan sistematika adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta penataan penulisan dari tugas akhir.

BAB II LANDASAN TEORI

Bab ini akan menjelaskan tentang pustaka yang menjadi tinjauan dalam pelaksanaan Tugas Akhir. Pustaka tersebut mencakup penjelasan dari judul atau masalah pada tugas akhir.

BAB III ANALISIS DAN PERANCANGAN

Bab ini membahas tentang analisis dan perancangan yang dilakukan pada penelitian didalam tugas akhir.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas mengenai implementasi serta pengujian yang dilakukan pada Sistem Rekomendasi yang akan dibuat

BAB V PENUTUP

Bab ini membahas tentang kesimpulan dari bab-bab sebelumnya dan saran dalam melakukan pengembangan penelitian kedepannya.