

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Era Digital dimana paradigma dan kebutuhan masyarakat dunia mulai bergeser dari informasi yang dibatasi oleh lembaran kertas menjadi informasi dalam bentuk digital. Internet menghidupkan sumber informasi digital dengan kemudahan akses dari berbagai perangkat teknologi informasi dan komunikasi secara cerdas, praktis dan terintegrasi. Terjadilah ledakan data dan informasi digital dimana social media, proses bisnis, publikasi-publikasi digital berperan meningkatkan volume dan pertumbuhan. Netflix perusahaan rental film yang merubah fungsi layanannya menjadi streaming video tahun 2016 tercatat memiliki 86.7 juta pelanggan yang berada di seluruh dunia dengan master Netflix Catalog yang memakan sekitar 3.14 Petabyte ruang penyimpanan cloud [1].

E-Commerce adalah pembelian, penjualan dan pemasaran barang beserta jasa melalui system elektronik, seperti televisi, radio dan jaringan computer atau internet [Wong,2010]. Berdasarkan hasil survei dari APJII tahun 2016, konten komersial yang paling sering dikunjungi masyarakat Indonesia adalah toko daring dengan jumlah 82,2 juta 62%. Selain itu, jumlah masyarakat Indonesia yang pernah bertransisi daring sebesar 84,2 juta 63,5% [safira,dkk,2017].

Perkembangan platform *e-commerce* pada abad ini sangat pesat dan saat ini banyak situs *e-commerce* di Indonesia yang menjual berbagai macam produk dan layanan.

Ledakan data dan informasi memunculkan permasalahan penemuan kembali informasi (information retrieval). Sistem rekomendasi (recommender system) adalah sistem untuk penyaringan, pemilahan item dan informasi yang mengambil preferensi dari perilaku pengguna, profil pengguna atau pendapat dari komunitas pengguna untuk membantu individu dalam mengidentifikasi konten yang menarik dan berpotensi besar untuk dipilih, dibeli atau digunakan [3], sementara collaborative filtering adalah inti dari sistem rekomendasi yang fokus pada penggunaan algoritma dan perhitungan matematika untuk mempelajari profil user, profil item dan interaksi keduanya dalam memprediksi item yang akan direkomendasikan kepada user [4]. Collaborative Filtering memecahkan sebagian masalah rekomendasi dari seorang ahli yang mengerti kebutuhan dan profil pengguna menjadi sangat dibutuhkan [5].

Saat ini, Teknik rekomendasi telah digunakan secara luas di berbagai macam domain. Rekomendasi atau system rekomendasi merupakan sebuah system yang menggunakan beberapa pengetahuan tentang daftar pengguna dan preferensinya terhadap beberapa item, dengan tujuan untuk memberi saran pada pengguna tersebut atau pada interaksi baru antara pengguna baru dengan item yang sebelumnya tidak pernah terkait. Interaksi yang terjadi dapat berupa jenis keterlibatan dalam segi hal apapun namun tetap sesuai dengan konteksnya, dan dapat bervariasi seperti opsi pembelian barang, menonton film, mendengarkan music.

System rekomendasi memiliki tujuan untuk menghasilkan rekomendasi yang berarti bagi para penggunaanya terhadap hal-hal yang mungkin menarik bagi

mereka.

Beberapa metode populer yang digunakan dalam membuat rekomendasi yaitu content based filtering, collaborative filtering dan hybrid. Content based filtering memanfaatkan interaksi antara konten item dengan profil pengguna [5] dimana yang termasuk konten item disini seperti genre, author, dll. Collaboratif filtering (CF) bekerja dengan membangun database preferensi konsumen terhadap suatu item. Seperti promosi dari mulut ke mulut, collaborative filtering memberikan prediksi rating dan personal rekomendasi berdasarkan yang disukai pengguna lain yang mempunyai selera yang sama [6]. Sementara hybrid merupakan penggabungan content based filtering dan collaborative filtering. Diantara ketiga metode diatas, perusahaan-perusahaan e-commerce raksasa seperti Amazon, netflix, pandora menggunakan metode collaborative filtering sebagai representasi dari wisdom of crowd.

Kemampuan collaborative filtering dibatasi oleh *Sparsity*. *Sparsity* yaitu terjadinya kekosongan data matriks user-item, yang disebabkan karena user merating dalam jumlah kecil dari jumlah item yang tersedia di dalam database hal tersebut terjadi ketika ada seseorang pelanggan yang hanya memberikan penilaian pada sedikit item produk, hal ini membuat perhitungan dari kesamaan dengan pelanggan lain menjadi lebih sulit dikarenakan bahwa salah satu factor penyebab dari data *Sparsity* adalah sulitnya memperoleh ulasan seperti *rating*. Oleh karena itu diperlukan suatu metode untuk memecahkan masalah tersebut agar proses Rekomendasi dapat dihasilkan dengan kualitas yang baik, Ketika bekerja pada data yang sangat besar seperti MovieLens Dataset Anwar et al. [7]. Sistem

rekomendasi tradisional kurang menunjukkan akurasi yang baik. Salah satu solusi untuk mengatasi tersebut adalah faktorisasi matrik Yuan et al. [8]. Salah satu solusi adalah melakukan imputasi yaitu proses pengisian awal terhadap data dengan metode tertentu. Dengan memanfaatkan rating film, dilakukan imputasi terhadap data untuk selanjutnya digunakan untuk mengisi kekosongan pada rating film. Penelitian ini berfokus pada penerapan proses imputasi terhadap rating film dan menganalisis pengaruhnya terhadap performansi.

Metode Ranking Based Collaborative Filtering menghasilkan performa yang lebih baik dalam mengatasi masalah *sparsity* karena dataset yang bersifat renggang, penelitian ini adalah mengimplementasikan metode Imputation Ranking Based Collaborative Filtering dan menganalisis performansi prediksi yang dihasilkan dengan Ranking Based Collaborative Filtering untuk mereduksi dimensi dataset (matriks rating) untuk mengatasi masalah *sparsity* yang dihadapi metode Ranking Based Collaborative Filtering. Berikut ini dijelaskan beberapa penelitian terkait masalah *sparsity*.

Penelitian Muhammad Iqbal A. [9] Metode Ranking Based Collaborative Filtering menghasilkan performa yang lebih baik dalam mengatasi masalah *sparsity* karena dataset yang bersifat renggang dibandingkan dengan Rating Based Collaborative Filtering. Berdasarkan evaluasi yang dilakukan, metode yang diusulkan membutuhkan waktu komputasi 71% lebih cepat dalam melakukan latihan dan 84% lebih cepat dalam membentuk rekomendasi dibandingkan dengan metode dari penelitian sebelumnya. Dilihat dari penggunaan sumber daya komputasi, metode yang diusulkan membutuhkan memori (RAM) 44% lebih rendah

dibandingkan dengan metode BordaRank serta 4,5% lebih rendah dalam memanfaatkan CPU. Sementara itu, berdasarkan matrik-matrik akurasi, metode yang diusulkan memperoleh nilai Coverage 2 kali lebih tinggi, nilai NDCG 19% lebih tinggi, namun nilai F1-score 49% lebih rendah daripada metode BordaRank. Penelitian selanjutnya dilakukan oleh Zeinab Sharifi. [10] Faktorisasi matriks adalah salah satu teknik yang paling disukai berdasarkan system rekomendasi berbasis model. Penelitian ini mengembangkan algoritma pada rekomendasi berdasarkan gagasan ini bahwa peringkat yang tidak diketahui dipengaruhi dari informasi yang diekstraksi dari peringkat yang tersedia, sehingga data perlu diproses terlebih dahulu karena ranking dari kumpulan data ini memiliki *sparsity* yang tinggi.

Berikutnya adalah penelitian Christian Wibisono. [11] Suku cadang yang dapat diganti dibengkel memiliki banyak transaksi. Tahap ini mengusulkan system pemberi rekomendasi dengan collaborative Filtering dengan sparsitas data yang tinggi. Model akan dievaluasi menggunakan matrik Mean Absolute Error (MAE), Root Mean Square Error (RMSE) dan Fraction of Concordant Pairs (FCP). Hasil model rekomendasi adalah MAE = 1,2752, RMSE = 1,4882 dan FCP = 0,4947.

Berdasarkan kondisi tersebut diatas mencoba untuk membangun sebuah system rekomendasi Collaborative Filtering untuk menyelesaikan masalah data *sparsity*. Interaksi yang terjadi antara user dengan item menggambarkan preferensi individu, sehingga tidak bergantung pada kuantitas item yang dibeli.

Model-based Collaborative Filtering diterapkan pada data interaksi pelanggan (user) dengan sebuah item sehingga dapat menghasilkan daftar rekomendasi.

Pemodelan system rekomendasi yang dibangun dapat memberikan hasil rekomendasi yang lebih baik.

1.2 Rumusan Masalah Penelitian

Data yang terus tumbuh memerlukan suatu rekomendasi sebagai alternatif untuk menemukan kembali item yang dibutuhkan. Hal tersebut karena rekomendasi yang dihasilkan memiliki kualitas yang lebih baik. Hingga penelitian ini mencoba menganalisis penanganan *sparsity*. Sehingga penelitian ini merumuskan masalah Bagaimana penanganan *sparsity* untuk meningkatkan kualitas rekomendasi ?

1.3 Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah :

1. Hanya menggunakan teknik pada collaborative filtering karena pada teknik ini item yang pernah dirating user (pengguna) akan menjadi landasan proses rekomendasi dan menggunakan Imputasi dan metode Borda.
2. Menggunakan dataset dari movielends.org.- 100k.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah menyimpulkan metode ranking based collaborative filtering yaitu Untuk meningkatkan kualitas rekomendasi dengan memecahkan masalah *sparsity* yang ada dengan menggunakan mengimplementasikan metode Imputation.

1.5 Manfaat Penelitian

1. Bagi Dunia Keilmuan

Mengaplikasikan ilmu yang selama ini diperoleh dalam bangku perkuliahan ke

dunia nyata sebagai bentuk pengabdian mahasiswa kepada masyarakat, sehingga dengan bekal ilmu pengetahuan yang dapat memberikan informasi yang bermanfaat.

2. Bagi diri sendiri

Dapat menambah pengetahuan mengenai pembuatan sebuah sistem serta pengetahuan mengenai teori-teori yang didapat pada saat perkuliahan dalam sebuah dunia kerja yang sesungguhnya. Sehingga membuat penulis menjadi lebih kritis dalam menghadapi permasalahan yang bisa saja berbeda dari teori-teori yang pernah didapat.

3. Bagi Instansi Pendidikan

Hasil penelitian dapat menambah referensi perpustakaan dan sebagai acuan bagi penelitian yang sejenis berikutnya.

1.6 Sistematika Penulisan

Bab I Pendahuluan :

Pada Bab I Pendahuluan, akan diuraikan tentang latar belakang, rumusan masalah, ruang lingkup, tujuan, manfaat penelitian, dan sistematika penulisan.

Bab II Landasan Teori :

Pada Bab II atau pada Tinjauan Pustaka, akan diuraikan tentang teori, atau pendekatan teori, proposisi dan konsep yang relevan untuk digunakan dalam menyelesaikan masalah yang telah dirumuskan, untuk mencapai tujuan yang telah ditetapkan.

Bab III Metode Penelitian :

Pada Bab III atau pada Metode Penelitian, akan diuraikan tentang rancangan dan penelitian, lokasi dan objek penelitian, sumber data, serta responden penelitian.

Bab IV Hasil dan Pembahasan :

Pada Bab IV atau Hasil dan Pembahasan, akan memuat hasil-hasil penelitian yang telah dicapai dalam pengembangan metode *ranking-based* yaitu metode Borda.

Bab ini dilengkapi dengan hasil penelitian serta evaluasi dan analisis terhadap hasil dan evaluasi yang didapatkan.

Bab V Kesimpulan dan saran :

Pada Bab V atau Kesimpulan dan saran, mengenai hasil evaluasi dan analisis serta penelitian selanjutnya yang dapat memperbaiki hasil penelitian ini.