

BAB I

PENDAHULUAN

Pada Pendahuluan ini berisikan tentang Latar Belakang untuk melakukan penelitian, Rumusan Masalah yang ditemukan, Batasan Masalah yang menjadi batas penelitian, Tujuan dan Manfaat dari penulisan laporan akhir ini.

1.1. Latar Belakang

Kecerdasan Buatan (*Artificial Intelligence*) merupakan salah satu bagian dari ilmu komputer yang mempelajari bagaimana membuat mesin (komputer) dapat melakukan pekerjaan seperti dan sebaik yang dilakukan manusia (Dahria, 2008). Dengan kecerdasan buatan, kita dapat dimudahkan untuk membuat sebuah sistem untuk membantu pekerjaan manusia. Contoh kecil yang dapat diterapkan kecerdasan buatan adalah Sistem Rekomendasi.

Sistem rekomendasi adalah suatu sistem yang menyarankan informasi yang berguna atau menduga apa yang akan dilakukan pelanggan untuk mencapai tujuannya, misalnya seperti memilih produk tertentu. Sehingga pelanggan memilih produk dapat lebih efektif dalam menentukan produk yang diinginkannya (Arief, 2016). Sistem rekomendasi juga dapat membantu menaikkan minat dalam memilih suatu hal yang berkaitan dengan hal yang sedang dicari saat ini. Seperti contoh kita sedang berselancar dalam platform *e-commerce*, ketika kita sedang mencari sebuah barang dan ketika kita sudah melihat deskripsi atau sudah memilih barang tersebut, maka tampilan di bawah barang tersebut atau pada halaman utama platform tersebut akan menampilkan dan menyarankan barang yang berkaitan dengan barang yang sedang kita lihat atau pilih sekarang.

Perkembangan teknologi informasi dan telekomunikasi didunia mengalami peningkatan yang sangat tajam, hal ini dapat dilihat bahwa sekarang ini banyak sekali kegiatan manusia yang membutuhkan teknologi informasi maupun komunikasi, tidak terkecuali dalam bidang musik maupun film. Film merupakan audio visual yang memiliki banyak sekali genre, dari genre komedi, drama, horor, aksi dan masih banyak lagi lainnya. Dikarenakan banyaknya jumlah film yang tayang didunia, dan setiap orang memiliki genre favorit sendiri-sendiri, maka dari

itu dibutuhkan sebuah rekomendasi dari penonton film lainnya dengan melihat *Rating* film apa yang menjadi pilihan terbanyak. Diharapkan dengan adanya sistem rekomendasi ini, *User* mendapatkan rekomendasi film yang lebih baik dari data yang diberikan *user* untuk diproses menjadi rekomendasi film yang baik.

Beberapa metode dapat diterapkan dalam sistem rekomendasi seperti algoritma *Frequent Pattern Growth*, *Apriori*, *K-Nearest Neighbor*, *SVM* dan lain sebagainya. Pada dasarnya, metode tersebut digunakan sebagai menghitung kemungkinan untuk mengetahui minat dari *User*. Dari metode pada Algoritma 1 adalah metode yang akan digunakan adalah penggabungan dari *Frequent Pattern Growth* dan *K-Nearest Neighbor* dikarenakan *Frequent Pattern Growth* mempunyai keunggulan yaitu lebih cepat dalam menemukan *support* untuk *itemsets* yang dicari karena menggunakan *FP-tree* untuk mencari relasi antar *itemsets* sedangkan algoritma *Apriori* harus mencari relasi antar item dan *itemsets* sehingga sedikit memakan waktu (Kavitha, 2016). Untuk algoritma *K-Nearest Neighbor* digunakan karena lebih sederhana diimplementasikan dibandingkan *SVM* karena metode *SVM* memiliki 5 *cross validation* sehingga lebih rumit diimplementasikan (Danny, 2018).

Pembuatan sistem rekomendasi film dengan algoritma *FP-Growth* dan *K-NN* sangat penting di era digital ini. Algoritma *FP-Growth* membantu memahami preferensi pengguna melalui pola asosiasi, sedangkan *K-NN* memastikan rekomendasi lebih personal dengan mempertimbangkan kesamaan antarpenonton. Dengan kombinasi keduanya, sistem ini mampu memberikan saran film yang lebih sesuai, meningkatkan kepuasan penonton, dan mendukung pertumbuhan industri film dalam memenuhi kebutuhan penonton yang beragam.

Metode yang digunakan peneliti yang pertama adalah Algoritma *FP-Growth* yang merupakan pengembangan dari algoritma *Apriori*. Algoritma *Frequent Pattern Growth* adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data (Yuyun, 2015). Metode dari algoritma *FP-Growth* dapat digunakan sebagai sistem rekomendasi secara umum karena metode tersebut dapat menunjukkan obyek yang paling sering muncul dalam pencarian *User* maupun dalam tampilan halaman utama pada setiap *User*, karena pada dasarnya

setiap *User* mempunyai sifat dan kondisi yang berbeda-beda. Pada algoritma *FP-Growth* menggunakan konsep Pembangunan *tree*, yang biasa disebut *FP-tree*, dalam pencarian *frequent itemsets* bukan menggunakan *generate candidate* seperti yang dilakukan pada algoritma Apriori. Dengan menggunakan konsep tersebut, algoritma *FP-Growth* menjadi lebih cepat daripada algoritma Apriori. Metode *FP-Growth* digunakan karena merupakan salah satu metode pendekatan yang dapat memproses secara ringkas dengan menangani masukan dari *User* secara langsung dan membentuk serta menemukan pola frekuensi yang tinggi sehingga pola tersebut didapatkan tanpa melakukan sebuah *training* seperti biasanya. Pola tersebut akan membentuk sebuah nilai dengan nama *support* di mana nilai ini dapat digunakan pada proses pembentukan rekomendasi selanjutnya dengan metode *Association Rules*. Metode *Association Rules* merupakan lanjutan dari metode *FP-Growth* yang memanfaatkan dari nilai *support* tersebut untuk diproses sebagai sistem rekomendasi dari metode ini.

Selain itu, Peneliti menggunakan metode klasifikasi yaitu Algoritma *K-Nearest Neighbor* (KNN). Algoritma KNN merupakan algoritma yang digunakan untuk mengklasifikasikan objek baru berdasarkan atribut dan *training sample* (Chandra, 2019). Prinsip kerja KNN adalah dengan mencari jarak terdekat antara data yang digunakan dengan K tetangga (*Neighbor*) terdekatnya dalam data *training*. Metode KNN digunakan karena memiliki beberapa kelebihan, di antaranya yaitu dapat menghasilkan data yang lebih akurat dan efektif apabila memiliki *training* data yang cukup besar. Namun, metode ini juga memiliki beberapa kekurangan, seperti biaya komputasi yang cukup tinggi karena diperlukan perhitungan jarak *query instance* pada keseluruhan *training sample*. Dengan penerapan metode algoritma yang berbeda, diharapkan dapat membuat sebuah sistem rekomendasi yang lebih baik lagi sehingga *User* merasakan keakuratan dari sistem rekomendasi yang dibuat.

Kedua algoritma tersebut menggunakan *Collaborative filtering* yang menggunakan data dari *feedback User*. *Collaborative filtering* merupakan proses penyaringan atau menggunakan opini orang lain (Arif, 2016). *Collaborative filtering* melakukan penyaringan data berdasarkan kemiripan karakteristik konsumen sehingga mampu memberikan informasi yang baru kepada konsumen

karena sistem memberikan informasi berdasarkan pola satu kelompok konsumen yang hampir sama. Perbedaan minat pada beberapa anggota kelompok menjadikan sumber informasi baru yang mungkin bermanfaat bagi anggota kelompok lainnya. Dalam penelitian yang dilakukan saat ini, usaha untuk membuat sebuah sistem yang menyambungkan 2 Algoritma dengan menggunakan *Association Rule*.

Analisis asosiasi atau *Association Rule mining* adalah teknik data *mining* untuk menemukan aturan asosiatif antara suatu kombinasi item (Destiyati, 2015). Analisis asosiasi dikenal juga sebagai salah satu teknik data *mining* yang menjadi dasar dari berbagai teknik data *mining* lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan tiga parameter, *support* (nilai penunjang) yaitu jumlah kemunculan atau persentase kombinasi item tersebut dalam *database*, *Confidence* (nilai kepastian) yaitu banyaknya persentase kemunculan kombinasi antar item, serta *lift*.

Dari latar belakang tersebut, penggunaan metode *FP-Growth* dan *K-Nearest Neighbor* dapat digunakan secara bersamaan dengan menggunakan sebuah metode penggabungan yaitu *Collaborative Filtering* yang di mana kedua metode tersebut tidak dapat disatukan secara langsung namun harus dilakukan beberapa penyesuaian sehingga dari metode satu dengan yang lain dapat terhubung dengan baik.

1.2. Perumusan Masalah

Dari latar belakang yang sudah dijabarkan, maka dapat dirumuskan masalah yang akan dihadapi yaitu:

1. Bagaimana penerapan *Collaborative filtering* dari metode *FP-Growth* dengan KNN dalam merumuskan rekomendasi dari sebuah film?
2. Bagaimana hasil evaluasi *Collaborative filtering* dari metode *FP-Growth* dengan KNN pada data yang disediakan?

1.3. Tujuan Penelitian

Tujuan yang diharapkan dari perumusan masalah tersebut, adalah sebagai berikut:

1. Mengimplementasikan *Collaborative filtering* untuk metode *FP-Growth* dan KNN dalam merekomendasikan film.
2. Mengetahui seberapa baik performa dari penggabungan antara algoritma *FP-Growth* dengan algoritma KNN pada saat menyaring data dari *User* untuk diproses, agar ditemukan rekomendasi terbaik dari kedua algoritma tersebut.

1.4. Manfaat Penelitian

Berdasarkan tujuan penelitian sebelumnya, adapun manfaat yang diperoleh yaitu:

1. Mengetahui penerapan *Cascade* dengan menggabungkan antara algoritma *FP-Growth* dengan algoritma KNN sebagai sistem rekomendasi dari film.
2. Dapat digunakan sebagai bahan evaluasi dan referensi dalam penelitian-penelitian selanjutnya. Khususnya penelitian pada sistem rekomendasi pada subyek lainnya yang membutuhkan algoritma untuk memberikan rekomendasi.

1.5. Batasan Masalah

Beberapa batasan masalah yang digunakan agar pembahasan dalam penelitian ini tidak melebar dari topik adalah sebagai berikut:

1. Data yang digunakan pada penelitian ini adalah data sekunder, (data yang tidak diambil secara langsung) himpunan data tersebut bernama *The Movies Dataset* yang diambil dari situs resmi *Kaggle* (Rounak, 2017)
2. Algoritma yang digunakan adalah gabungan dari dua algoritma yaitu *Frequent Pattern Growth* dan *K-Nearest Neighbor*
3. Jumlah data yang digunakan adalah 45.572 film yang berbentuk data dengan ulasan dari *User* sebanyak 100.004 buah dari 671 *User*.
4. Membatasi nilai rekomendasi dari metode terpilih yaitu dengan maksimal 10 rekomendasi atau dengan nilai rekomendasi minimal 3.0