

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Menonton film merupakan bentuk hiburan bagi masyarakat, baik di bioskop maupun melalui kanal *streaming online* (Dewanto & Mulyadi, 2020). Perkembangan industri film modern tidak hanya ditandai dengan bertambahnya jumlah produksi, tetapi juga dengan bertambahnya jumlah genre yang ada, dan setiap tahunnya banyak film yang diproduksi dari berbagai sumber (Yusuf et al., 2023). Film telah menjadi hiburan yang dapat dinikmati oleh semua orang, tanpa memandang golongan usia (Perdana et al., 2019).

Banyaknya film yang tayang menimbulkan tantangan bagi penonton untuk memilih film yang sesuai dengan preferensi dan minat mereka (Robbani & Sulianta, 2024). Misalnya, jika seseorang ingin menonton film berdasarkan pendapatan film tertinggi dan biaya produksi terbesar, maka rekomendasinya tidak lagi hanya dilihat dari rating penonton lainnya. Sehingga butuh waktu lama untuk menemukan film yang sesuai dengan minat pengguna karena banyaknya jumlah film, dengan begitu dibutuhkanlah suatu sistem yang dapat merekomendasikan film (Mah et al., 2023). Untuk merekomendasikan film berdasarkan minat pengguna, penelitian ini mengimplementasikan algoritma *skyline query* untuk mendapatkan hasil rekomendasi film sesuai preferensi pengguna.

*Skyline query* adalah algoritma yang bertujuan untuk mengidentifikasi sekumpulan objek yang telah ditentukan sesuai dengan kriteria pengguna yang berbeda-beda (Vlachou et al., 2022). *Skyline* meningkatkan *query database* dengan memperkenalkan mekanisme untuk menyaring poin menarik dari kumpulan data berdasarkan kriteria dominasi. Dengan menggabungkan operasi *skyline*, sistem *database* dapat secara efisien mengekstrak poin yang tidak didominasi dan mengurangi kompleksitas hasil *query* (Grasmann et al., 2022). Proses penentuan objek *skyline* dilakukan dengan cara membandingkan setiap objek dalam dataset dengan semua objek lain untuk menilai dominasi. Objek dikategorikan sebagai objek yang tidak didominasi oleh objek lain jika nilai objek tersebut pada semua atributnya lebih baik, dan minimal satu atribut pada objek tersebut lebih baik dari objek lainnya (Annisa, 2021). *Skyline query*

mengembalikan sekumpulan objek *skyline* dengan memfilter objek yang didominasi oleh objek lain dari kumpulan data multi-kriteria. Hal ini memungkinkan pengguna untuk menentukan preferensi individu dalam rentang nilai ideal untuk suatu kriteria (Dzolkhifli et al., 2022). Dengan menggunakan algoritma ini, film dapat direkomendasikan sesuai preferensi pengguna yang memiliki multi-kriteria. Misalnya, preferensi film yang disukai seseorang adalah film dengan popularitas tinggi, rata-rata penilaian yang bagus, biaya produksi yang mahal, serta rating yang memuaskan.

Keunggulan algoritma *skyline* terletak pada efektivitasnya dalam menghasilkan rekomendasi yang dipersonalisasi. Cara kerja *skyline* adalah dengan melihat dominasi, yaitu unggul dari semua atribut atau minimal satu atribut (Purwayoga & Nurkholis, 2023). Salah satu jenis algoritma *skyline query* adalah *Sort Filter Skyline (SFS)*. SFS menggunakan nilai fungsi entropi untuk mengurutkan kumpulan data, guna mengurangi perbandingan dominasi. Dalam penelitian ini, objek (film) akan dibandingkan berdasarkan atribut dan jumlah datanya (Annisa, 2021). Algoritma ini juga telah banyak digunakan dalam bidang penelitian seperti visualisasi *database*, penambahan data, dan pengambilan keputusan multi-kriteria serta untuk memproses data dalam jumlah besar (Du et al., 2020). *Skyline* umumnya diterapkan pada industri pariwisata (Ghosh et al., 2021), namun dalam penelitian ini, penerapan algoritma SFS akan digunakan untuk mencari rekomendasi film berdasarkan preferensi individu.

Penelitian yang dilakukan oleh Munir dan Asqia (Munir & Asqia, 2021) mengimplementasikan metode *sort filter skyline* untuk memberikan rekomendasi pemilihan tempat makan berdasarkan preferensi wisatawan. Dalam penelitiannya, menggunakan tiga aspek kriteria sebagai preferensi, yaitu produk, pelayanan, dan pengelolaan. Ketiga aspek tersebut menjadi acuan dalam preferensi seseorang untuk memilih atau menentukan rating dari sebuah usaha kuliner. Berdasarkan ketiga aspek tersebut, dibuatlah empat atribut untuk menyajikan rekomendasi pemilihan tempat makan, yaitu: makanan, suasana, layanan, dan kesehatan. Sedangkan implementasi dalam penelitian ini akan memuat lima aspek kriteria rekomendasi film, yaitu: anggaran yang dikeluarkan untuk membuat film, pendapatan film, popularitas film, *rating*, dan durasi lamanya film. Dari aspek

tersebut, dibuatlah tujuh atribut yang menjadi landasan rekomendasi film, yaitu: *id, budget, popularity, revenue, runtime, vote average, dan vote count*.

Berdasarkan latar belakang tersebut, penelitian ini akan implementasi algoritma *Sort Filter Skyline* (SFS) dalam membantu merekomendasikan film berdasarkan preferensi individu.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dapat diambil beberapa rumusan masalah sebagai berikut :

1. Bagaimana mengimplementasikan algoritma *Sort Filter Skyline* (SFS) untuk merekomendasikan film berdasarkan preferensi individu?
2. Bagaimana mengukur kompleksitas waktu pemrosesan algoritma *Sort Filter Skyline* (SFS) berdasarkan karakteristik sebaran data, jumlah data dan jumlah atribut?

## 1.3 Tujuan Penelitian

Berikut adalah beberapa tujuan penelitian yang akan dilakukan :

1. Mengetahui hasil rekomendasi film dari penerapan algoritma *Sort Filter Skyline* (SFS) sesuai preferensi individu
2. Mengukur kompleksitas waktu eksekusi algoritma *Sort Filter Skyline* (SFS) berdasarkan karakteristik sebaran data, jumlah data dan jumlah atribut

## 1.4 Batasan Penelitian

Adapun batasan masalah yang ditetapkan adalah sebagai berikut :

1. Dataset yang digunakan untuk merekomendasikan film adalah data *real* yaitu data TMDb (API, 2024) dan data yang digunakan untuk pengujian kompleksitas waktu eksekusi adalah data sintetis yang memiliki sebaran data independen, korelasi dan anti-korelasi
2. Jumlah data *real* film yang diolah untuk rekomendasi sebanyak 4803 data dan jumlah data untuk pengujian kompleksitas waktu eksekusi menggunakan data sintetis sebanyak 50.000 data
3. Atribut data yang digunakan meliputi *id, popularity, budget, revenue, runtime, vote\_average, vote\_count*

### 1.5 Manfaat

Beberapa manfaat yang dapat diambil dari penelitian ini adalah :

1. Sistem yang dihasilkan dapat menyediakan rekomendasi yang lebih personal yang menyediakan film-film yang sesuai dengan preferensi unik pengguna
2. Implementasi algoritma ini dapat meningkatkan skalabilitas sistem rekomendasi, sehingga dapat menangani jumlah data yang lebih besar tanpa penurunan kinerja
3. Penelitian ini dapat membuka jalan bagi pengembangan metodologi baru dalam sistem rekomendasi, yang dapat diterapkan pada berbagai domain lain seperti musik, buku, dan produk *e-commerce*

