

BAB I

PENDAHULUAN

1.1 Latar Belakang

Film adalah sebuah kesenian gabungan antara seni sastra, seni peran, seni musik, dan seni rupa. [1]. Pada masa sekarang ini film telah menjadi salah satu hiburan favorit utama masyarakat. Jumlah film pertahun terhitung mencapai ribuan [2]. Dilihat dari banyaknya jumlah penonton bioskop yang terus meningkat dari tahun ke tahun. Per 2018 angka jumlah penonton bioskop di Indonesia saja telah mencapai lebih dari 50 juta penonton dengan jumlah produksi film luar negeri hingga dalam negeri sebanyak hampir 200 judul film yang telah tayang di seluruh Indonesia [3]. Setiap tahun, banyak film baru yang diproduksi oleh industri yang membuat beberapa orang bingung untuk memilih film mana yang layak untuk ditonton. Kelayakan sebuah film menjadi sebuah aspek penting untuk memberikan penilaian positif terhadap film, sehingga dapat menjadikan film tersebut laris di kalangan masyarakat luas dan penonton yang menonton juga akan merasakan kepuasan.

Pemilihan film yang bagus untuk ditonton menjadi sebuah masalah yang sering dijumpai. Kesesuaian genre, rating, alur cerita, pengambilan gambar, pemilihan musik, dan aspek lainnya merupakan sebuah poin penting yang mendasari film tersebut laris di kalangan masyarakat. Kelayakan sebuah film sangat mempengaruhi kepuasan penonton, karena kelayakan film secara langsung mempengaruhi pengalaman menonton dan reaksi emosional penonton. Saat penonton merasa terhubung secara emosional dengan cerita, karakter, atau tema yang disajikan dalam film, tingkat kepuasan mereka meningkat, dan mereka lebih

cenderung akan merekomendasikan film tersebut kepada orang lain serta ingin menontonnya kembali. Oleh karena itu, dibutuhkan sebuah aplikasi yang dapat merekomendasikan film berdasarkan selera dari sang penonton.

Beberapa penelitian menunjukkan bahwa sebuah sistem yang bisa merekomendasikan film dapat dibangun berdasarkan ulasan yang ditulis penonton tersebut. Dari ulasan yang di tulis itu, dapat diambil beberapa informasi terkait aspek-aspek kesesuaian sang penonton terhadap film tersebut. Penelitian terdahulu menunjukkan bahwa pembuatan sistem rekomendasi film dengan menggunakan metode *Collaborative Filtering*. Metode *Collaborative Filtering* adalah metode yang memberikan rekomendasi berdasarkan prefensi dari pengguna atau item lain yang mirip [4]. Dengan begitu, penonton dapat mendapatkan rekomendasi film sesuai dengan beberapa kriteria yang disukai seperti genre, durasi film, dan lain sebagainya.

Metode *Collaborative Filtering* menghasilkan rekomendasi berdasarkan keterkaitan antar *user* yang menyukai suatu item tertentu. Pada *Collaborative Filtering*, rekomendasi berdasarkan pada keterkaitan yang didapat diantara *user* yang telah merating dalam sistem. Misalnya *user* A menyukai film X dan Y, kemudian *user* B menyukai film Y maka sistem akan merekomendasikan film X kepada *user* B. *Collaborative Filtering* melakukan penyaringan data berdasarkan kemiripan karakteristik *user* sehingga mampu memberikan informasi yang baru kepada *user* karena sistem memberikan informasi berdasarkan pola satu kelompok *user* yang hampir sama [5]. Dengan begitu, rekomendasi yang diberikan menggunakan algoritma *Collaborative Filtering* dapat lebih luas cakupannya

dikarenakan algoritma ini mempertimbangkan berbagai preferensi dan pola perilaku dari banyak pengguna, bukan hanya satu individu. Dengan menganalisis data dari berbagai pengguna yang memiliki preferensi serupa, *Collaborative Filtering* mampu menghasilkan rekomendasi yang lebih relevan dan sesuai dengan selera pengguna yang bersangkutan.

Pada penelitian lain juga menggunakan metode sentimen analisis untuk menilai ulasan pengguna terkait film yang mereka tonton. Analisis sentimen merupakan salah satu teknik *Natural Language Processing* (NLP) yang menganalisis pendapat, sikap, dan emosi terhadap suatu entitas yang berupa teks [6]. Untuk algoritmanya sendiri akan menggunakan *Naïve Bayes* dikarenakan, nilai akurasi dalam pelatihan model sistem yang terbilang cukup akurat. Dari model yang telah dilatih tersebut nantinya akan diimplementasikan kedalam aplikasi berbasis website untuk mendapatkan informasi berdasarkan ulasan yang ditulis oleh pengguna setelah menonton film yang telah direkomendasikan. Model akan mendeteksi ulasan pengguna bernilai positif atau negatif yang nantinya akan digunakan sebagai media untuk rekomendasi film untuk pengguna selanjutnya yang dikombinasikan dengan algoritma *Collaborative Filtering* sebagai pendekatan utamanya. Dengan demikian, diharapkan pengguna dapat lebih mudah menemukan film yang sesuai dengan kriteria tertentu dan dapat meningkatkan pengalaman menonton film secara keseluruhan.

1.2 Tujuan dan Manfaat

1.2.1 Tujuan

1. Mengembangkan sebuah aplikasi berbasis website yang mampu merekomendasikan film-film Asia kepada pengguna berdasarkan analisis ulasan dengan menggunakan metode *Collaborative Filtering*.
2. Mengintegrasikan model analisis sentimen yang telah dilatih ke dalam aplikasi berbasis website untuk memberikan rekomendasi film-film yang sesuai dengan preferensi pengguna berdasarkan ulasan-ulasan yang telah dianalisis.
3. Membangun sebuah sistem rekomendasi yang dapat membantu pengguna dalam menavigasi melalui jumlah yang besar dari film-film Asia yang tersedia.

1.2.2 Manfaat

1. Mempermudah pengguna untuk menemukan film-film Asia yang sesuai dengan preferensi mereka.
2. Menghemat waktu untuk mencari film yang sesuai dengan selera pengguna.
3. Memperluas cakupan penonton untuk film-film Asia yang mungkin sebelumnya kurang dikenal dengan memberikan rekomendasi yang sesuai dengan preferensi individu pengguna.

1.3 Tinjauan Pustaka

Penelitian yang akan dilakukan memiliki tujuan untuk membuat sebuah aplikasi yang dapat merekomendasikan film berdasarkan ulasan yang ditulis oleh pengguna. Proses identifikasi ulasan bernilai positif atau negatif akan menggunakan algoritma multinomial naïve bayes untuk menghasilkan nilai akurasi yang akurat. Hasil dari proses tersebut akan digunakan akan disimpan yang nantinya digunakan sebagai media rekomendasi judul film yang sesuai dengan aspek penilain dari pengguna. Dengan demikian, pengguna akan mendapatkan pengalaman menonton yang lebih baik, di mana film yang direkomendasikan dapat sesuai dengan selera mereka sehingga akan mendapatkan pengalaman menonton film yang lebih baik.

Beberapa penelitian tentang sistem rekomendasi film telah banyak dilakukan sebelumnya. Seperti yang dilakukan oleh Alifia dkk. melakukan penelitian tentang Sistem Rekomendasi Film Berbasis Website dengan Metode Prototype Menggunakan Algoritma *K-Nearest Neighbors* (KNN) [7]. Penelitian tersebut dilakukan untuk membuat aplikasi yang dapat merekomendasikan film dengan memanfaatkan algoritma *K-Nearest Neighbors* (KNN) sebagai teknik data mining dalam melakukan klasifikasinya. Hasil dari penelitian tersebut berupa sistem berbasis website yang akan menampilkan 5 film yang direkomendasikan sistem. Kesimpulan dari penelitian tersebut adalah metode *content based filtering* berhasil diterapkan dalam sistem dimana sistem memberikan output berupa rekomendasi judul film berdasarkan kemiripan genre dan rating.

Penelitian lain juga dilakukan oleh Muhammad Fajriansyah dkk. yang berjudul Sistem Rekomendasi Film Menggunakan *Content Based Filtering* [3].

Penelitian tersebut membangun sebuah aplikasi rekomendasi film dengan menggunakan metode *Content Based Filtering* yang mempertimbangkan perilaku dari pengguna dari masa lalu yang kemudian diidentifikasi pola perilakunya untuk merekomendasikan barang yang sesuai dengan pola perilaku tersebut. Dari data *history* sang pengguna, nantinya akan digunakan sebagai tolak ukur sebagai media rekomendasi film yang sesuai dengan selera sang pengguna. Dengan demikian, pengguna dapat mendapatkan rekomendasi film secara efisien dan hemat waktu.

Penelitian yang dilakukan oleh Fathullah dkk. yang berjudul Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi *Naïve Bayes* dengan Fitur *Lexicon-Based* menjelaskan tentang bagaimana cara menganalisis sebuah ulasan dan rating yang diberikan oleh pengguna. Dengan adanya sentimen analisis terhadap ulasan dan rating film diharapkan memudahkan para penggemar untuk mendapatkan penilaian terhadap suatu film [1]. Penelitian tersebut mengeksplorasi bagaimana metode *Naïve Bayes*, yang dikenal efektif untuk klasifikasi teks, dapat dioptimalkan menggunakan fitur *Lexicon-Based*, yaitu pemilihan kata-kata penting berdasarkan kamus sentimen. Hal ini dilakukan untuk meningkatkan akurasi dalam klasifikasi sentimen. Pada penelitian tersebut menguji dan membandingkan akurasi, *precision*, dan *recall* dari dua pendekatan: menggunakan fitur *Lexicon-Based* dan fitur *stopwords removal*. Hasilnya menunjukkan bahwa menggunakan *Lexicon-Based* dengan memperhitungkan rating memberikan hasil yang lebih baik.

Penelitian lainnya juga dilakukan oleh Dodi Nuigraha dkk. yang berjudul Sistem Rekomendasi Film Menggunakan Metode *User Based Collaborative*

Filtering. Ide utama dalam sistem rekomendasi *Collaborative Filtering* adalah untuk memanfaatkan riwayat opini pengguna aktif lain untuk memprediksi item yang mungkin akan disukai/diminati oleh seorang pengguna [8]. Penelitian ini mendeskripsikan bagaimana sistem rekomendasi film dikembangkan dalam bentuk aplikasi website. Pengguna akan masuk melalui halaman *login*, kemudian memilih algoritma rekomendasi dan memberikan preferensi genre atau rating. Sistem kemudian akan menampilkan rekomendasi film berdasarkan preferensi pengguna. Pada Tabel 1.1 menjelaskan tentang perbandingan penelitian terdahulu.

Tabel 1.1 Perbandingan Penelitian Terdahulu.

| No | Judul | Teknologi | Hasil | Pembeda |
|----|--|----------------------------------|--|--|
| 1. | Sistem Rekomendasi Film Berbasis Website dengan Metode Prototype Menggunakan Algoritma <i>K-Nearest Neighbors</i> (KNN) 2020 | <i>K-Nearest Neighbors</i> (KNN) | Berdasarkan penelitian yang sudah dilakukan, hasil yang didapat adalah metode content based filtering berhasil diterapkan dalam sistem dimana sistem memberikan output berupa rekomendasi judul film berdasarkan kemiripan genre dan rating. Kelebihan dari sistem ini adalah : <ul style="list-style-type: none"> • Sistem mudah dioperasikan • Memiliki userfriendly pada antarmuka-nya. Kekurangan dari sistem ini adalah : <ul style="list-style-type: none"> • Hanya menggunakan 2 variabel saja, yang dimana sistem rekomendasi ini menghasilkan nilai presisi 45% dan nilai recall 60%. | Dalam penelitian ini sistem rekomendasi dibangun dengan <i>Collaborative Filtering</i> sedangkan penelitian sebelumnya menggunakan KNN sebagai metodenya. Serta menerapkan sentimen analisis untuk data yang akan digunakan untuk sistem rekomendasinya. |
| 2. | Sistem Rekomendasi Film Menggunakan | <i>Content Based Filtering</i> | Dari evaluasi yang telah dilakukan dengan metode MAP@K kepada tiga pengguna. Tingkat akurasi dari penggunaan metode content based filtering dalam sistem rekomendasi film dengan fitur judul, genre, dan | Pada penelitian ini menggunakan pendekatan <i>Content Base Filtering</i> yang berfokus pada |

| No | Judul | Teknologi | Hasil | Pembeda |
|----|--|-----------------------------------|---|---|
| | <i>Content Based Filtering</i> 2021 | | <p>sinopsis, pembobotan <i>TF-IDF</i>, dan cosine similarity dihitung mencapai 0.823254 untuk jenis kueri <i>single</i> kueri dan 0.7500556 untuk jenis kueri <i>multiple seeds</i> kueri.</p> <p>Kelebihan sistem ini adalah :</p> <ul style="list-style-type: none"> • Sistem menggunakan jenis kueri <i>multiple seeds</i> untuk meningkatkan hasil rekomendasi film. • <i>Multiple seeds query</i> cukup berpengaruh terhadap hasil rekomendasi, artinya sistem dapat memberikan rekomendasi yang lebih relevan. <p>Kekurangan sistem ini adalah :</p> <ul style="list-style-type: none"> • Penelitian ini hanya membahas tentang pelatihan model sistem rekomendasi, namun tidak menjelaskan implementasian dari model yang dibangun. | <p>satu elemen atau konten saja seperti genre. Sedangkan penelitian yang mahasiswa bangun menggunakan <i>Collaborative Filtering</i> yang elemen atau kontennya terdiri lebih dari satu. Jadi hasil rekomendasi filmnya akan lebih luas. Penelitian tersebut juga tidak menerapkan sentimen analisis.</p> |
| 3. | <p>Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi <i>Naïve Bayes</i> dengan Fitur <i>Lexicon-Based</i> 2020</p> | <i>Naïve Bayes, Lexicon-Based</i> | <p>Berdasarkan hasil pengujian, klasifikasi dengan fitur rating dan teks berdasarkan <i>stopwords removal</i> memiliki hasil <i>precision</i>, <i>recall</i> dan akurasi sebesar 0,9. Sedangkan klasifikasi dengan fitur rating dan teks berdasarkan <i>lexicon-based</i> memiliki hasil <i>precision</i>, <i>recall</i> dan akurasi sebesar 1.</p> <p>Kelebihan sistem ini adalah :</p> <ul style="list-style-type: none"> • Sistem ini memiliki hasil klasifikasi yang lebih akurat dalam membedakan sentimen. • Sistem dapat meningkatkan akurasi dalam memprediksi sentimen berdasarkan teks. <p>Kekurangan sistem ini adalah :</p> <ul style="list-style-type: none"> • Data yang digunakan sedikit, sehingga hasil uji kurang akurat | <p>Penelitian mengimplentasikan langsung sentimen analisis sebagai salah satu media rekomendasi yang dikombinasikan dengan <i>Collaborative Filtering</i> sehingga menghasilkan rekomendasi yang lebih luas.</p> |

| No | Judul | Teknologi | Hasil | Pembeda |
|----|--|---|---|---|
| 4. | Sistem Rekomendasi Film Menggunakan Metode User Based Collaborative Filtering 2021 | <i>User-Based Collaborative Filtering</i> | <p>Aplikasi menggunakan 5 bobot <i>preferences</i> pada algoritma <i>user based collaborative filtering</i> menghasilkan nilai <i>error</i> yang lebih kecil dengan nilai 0,73% dibandingkan menggunakan 3 atau 4 <i>preferences</i>.</p> <p>Kelebihan dari sistem ini adalah :</p> <ul style="list-style-type: none"> • Aplikasi ini mampu membantu pengguna untuk menampilkan list film yang sesuai dengan apa yang diinginkan oleh pengguna. • Analisis <i>user based collaborative filtering</i> dapat mempermudah melakukan perekomendasi film. <p>Kekurangan dari sistem ini adalah :</p> <ul style="list-style-type: none"> • Belum menambahkan <i>trailer</i> pada tampilan website, • Belum bisa menggabungkan 2 algoritma dalam 1 proses rekomendasi. | <p>Penelitian ini mengimplementasikan sentimen analisis untuk media rekomendasinya yang digabungkan dengan <i>Collaborative Filtering</i> sehingga menghasilkan rekomendasi yang lebih luas dan baik.</p> |

1.4 Data Penelitian

1.4.1 Dataset

Dalam penelitian ini, data diperoleh dari website MovieLens yang menyediakan data film yang bisa digunakan sebagai dataset, dimana berisi 100.000 rating, 10.329 film dengan 668 user. Namun penelitian kali ini memiliki lingkup film asia saja sehingga data film yang digunakan hanya 5.000 film saja. Film dikelompokkan menjadi 13 genre yaitu *action, adventure, crime, comedy, drama, documentary, fantasy, family, history, horror, mystery, romance, science fiction, dan thriller*. User dapat memberikan rating (skala 1 - 5) dengan rincian skala 1

adalah yang paling buruk dan skala 5 adalah skala yang paling baik terhadap film yang *user* nilai nantinya.

Selanjutnya data *review* yang diambil dengan metode *scrapping* di website IMDB untuk mendapatkan komentar orang-orang terhadap film. Data diambil dengan ketentuan minimal dari 1 film terdapat 100 komentar yang selanjutnya akan disimpan kedalam file csv. Untuk informasi seputar film seperti sinopsis, tahun rilis, dan lain sebagainya diambil dari website TMDB yang menyediakan informasi film dengan lengkap. Semua data Kemudian dibuat menjadi dataset dengan format csv agar mudah untuk diproses.

1.4.2 Pengumpulan Dataset

Pengumpulan dataset dilakukan dengan mencari data film-film yang ada di seluruh dunia. Data selanjutnya disortir untuk mendapatkan data film yang diproduksi di lingkup Asia saja. Dalam hal ini, lingkup Asia yang dimaksud merupakan negara-negara di Asia Tenggara ditambah negara Asia yang terkenal di dunia industri film yaitu Jepang, Korea Selatan, Cina, Iran, dan India. Selanjutnya mengumpulkan data review atau komentar orang-orang terhadap film tersebut. Pengumpulan dengan cara scrapping di website IMDB yang dikenal memiliki user terbanyak sehingga data review dapat mudah dikumpulkan. Informasi tambahan seperti link untuk streaming film didapat dari memanfaatkan sebuah library di Python yang bernama *simple-justwatch-api-python* yang terhubung dengan API website *Justwatch* yang berisi kumpulan informasi film berserta dengan referensi untuk menontonnya secara streaming.

1.4.3 Dataset Yang Dihasilkan

Dataset yang dikumpulkan berisi jumlah informasi terkait film seperti sinopsis, rating, genre, tahun terbit, link streaming, dan lain sebagainya. Dataset dibagi menjadi 3, yang pertama dataset movies yang berisi data film, dataset rating yang berisi rating film oleh user, dan dataset reviews yang berisi komentar user serta label dari komentar tersebut yang bernilai positif atau negatif. Dataset disimpan dengan format .csv agar saat pembuatan model sistemnya dapat di olah dengan mudah.

1.4.4 Alat Penelitian

Penelitian ini menggunakan beberapa alat untuk membangun sistem.

Beberapa alat tersebut antara lain :

1. Perangkat Keras
 - a. Laptop Acer Aspire E-14
 - b. SSD 512 GB
 - c. RAM 8 GB
 - d. Intel Core i7 7020U

2. Perangkat Lunak

Perangkat lunak yang digunakan dalam membangun sistem rekomendasi ini dapat dilihat pada Tabel 1.1 berikut ini.

Tabel 1.2 Perangkat Lunak

| No. | Alat | Fungsi |
|-----|------------|---|
| 1. | Figma | merancang tampilan <i>interface</i> sistem. |
| 2. | Tensorflow | sebagai module dalam melatih dan |

| No. | Alat | Fungsi |
|-----|-------------------------------|---|
| | | mengimplementasikan model untuk aplikasinya. |
| 3. | MySQL | sebagai media penyimpanan data hasil deteksi dan hasil rekomendasi filmnya. |
| 4. | XAMPP | media untuk meng- <i>host</i> sistem secara lokal |
| 5. | Laravel | <i>framework</i> yang dipakai untuk membangun sistem website. |
| 6. | HTML dan CSS | bahasa <i>markup</i> yang dipakai untuk membuat <i>front End</i> website. |
| 7. | Jupyter Notebook/Google Colab | teks editor yang dipakai untuk membangun kode program |
| 8. | ChatGPT ver 3.5 | membantu membangun dan menganalisa dalam implementasi kode |