



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Film menurut Kamus Besar Bahasa Indonesia memiliki arti yaitu lakon atau cerita gambar hidup, dalam Bahasa Inggris film biasanya disebut sebagai movie [1]. Film-film yang masih baru biasanya akan tayang di bioskop atau sinema. Bioskop atau sinema di Indonesia tidak hanya menayangkan film-film yang berasal dari industri perfilman dari luar negeri contohnya Hollywood (Amerika Serikat) ataupun Bollywood (India) namun juga menayangkan film-film yang berasal dari Indonesia sendiri.

Perfilman di Indonesia sedang mengalami perkembangan. Perkembangan perfilman di Indonesia ini dapat dilihat dengan bertambahnya penonton film-film Indonesia. Berdasarkan pada situs Indonesia.Go.Id, pada tahun 2016 jumlah penonton mencapai 34,5 juta, tahun 2017 meningkat menjadi 40,5 juta, dan pada tahun 2018 menyentuh lebih dari 50 juta penonton . Bahkan film Indonesia yang diproduksi pada tahun 2018 berjumlah 200 judul, meningkat dari tahun sebelumnya yang hanya menghasilkan 143 judul saja [2].

Berdasarkan survey yang dilakukan oleh IDN Times dengan 411 responden, menunjukkan penonton di Indonesia memilih film yang akan ditonton berdasarkan beberapa faktor. Orang-orang memilih film untuk ditonton di bioskop berdasarkan *genre* ada sebesar 16,4 persen, pemain atau aktor sebesar 14,5 persen, dan sutradara sebesar 12,7 persen [3].

Untuk dapat mengetahui minat dari para penonton film Indonesia tersebut dibutuhkan penilaian sebuah film yang biasanya berbentuk *rating*. *Rating* merupakan kesimpulan yang menggambarkan kualitas dari elemen-elemen yang ada dalam sebuah film [4]. Film dengan *rating* yang tinggi akan memiliki kualitas yang baik sehingga akan diminati oleh para penontonnya. Sebaliknya, film dengan

rating yang rendah akan kurang diminati. *Rating* film ini juga dapat menentukan kesuksesan atau popularitas dari sebuah film.

Saat ini minat peneliti-peneliti tentang *rating* film Indonesia sangatlah sedikit. Sebagian besar dari mereka lebih banyak membahas tentang film di luar negeri dibandingkan dengan film-film Indonesia [4]–[8]. Dikarenakan penelitian tentang *rating* film Indonesia ini masih sangat sedikit, sehingga dibutuhkan penelitian yang membahas tentang *rating* film Indonesia.

Untuk penelitian tentang *rating* film Indonesia ini dapat dilakukan dengan menggunakan teknik *data mining*. *Data mining* merupakan proses untuk mencari pola (*pattern*) dan mendapatkan informasi yang berguna dengan menggunakan teknik atau metode tertentu. *Data mining* memiliki beberapa bidang tersendiri yang mencakup prediksi, klasifikasi, deskripsi, asosiasi, estimasi, dan clustering [9]. Prediksi merupakan suatu Teknik memperkirakan secara sistematis untuk sesuatu yang paling mungkin terjadi di kemudian hari berdasarkan data dari masa lalu maupun data yang dimiliki di saat ini [7]. Menurut KBBI, prediksi disebut sebagai peramalan atau prakiraan sebuah nilai yang belum diketahui dan untuk masa yang akan datang. Dengan prediksi ini diharapkan dapat menghasilkan informasi yang berguna untuk para pembuat film untuk meningkatkan kualitas film serta untuk perkembangan perfilman di Indonesia kedepannya.

Penelitian-penelitian untuk memprediksi *rating* film berdasarkan *rating* ini sebelumnya sudah menggunakan beberapa algoritma *data mining*. Metode yang digunakan pada penelitian-penelitian sebelumnya adalah: Naïve Bayes dengan hasil akurasi sebesar 55,80% [5], Naïve Bayes dan K-Nearest Neighbor sebesar 53,99% dan 51,26% [7]. Berdasarkan beberapa penelitian terdahulu yang sebelumnya disebutkan dapat disimpulkan bahwa algoritma Naïve Bayes merupakan algoritma yang terpopuler untuk digunakan dalam memprediksi klasifikasi *rating* pada film namun dengan hasil yang kurang baik.

Untuk penelitian ini, prediksi akan dilakukan dengan menggunakan data film dari Indonesia dari tahun 2011 hingga 2020 dengan tujuan untuk meningkatkan performa model berdasarkan penelitian terdahulu. Model akan dibuat

menggunakan algoritma Naïve Bayes dan KNN (*K-Nearest Neighbor*) dengan teknik *Feature Selection* untuk menentukan *rating* pada film Indonesia. Teknik *Feature Selection* digunakan dalam penelitian ini dikarenakan kegunaannya yang dapat memilih fitur-fitur yang paling relevan dan dapat meningkatkan performa dari model yang dibuat [10]. Hasil yang didapat dalam penelitian ini nantinya akan memberikan kontribusi terhadap penelitian pada *rating* film Indonesia.

1.2. Rumusan Masalah

Berdasarkan pada latar belakang masalah yang ada, maka dapat dibuat rumusan masalah yang akan diselesaikan pada penelitian ini dengan melakukan prediksi *rating* pada film-film Indonesia adalah sebagai berikut:

1. Bagaimana membuat model untuk prediksi *rating* film dengan algoritma *Naïve Bayes* dengan teknik *Feature Selection*?
2. Bagaimana membuat model untuk prediksi *rating* film dengan algoritma *KNN* dengan teknik *Feature Selection*?
3. Bagaimana peningkatan performa yang didapat dari model prediksi *rating* film yang dibuat menggunakan teknik *Feature Selection* dibandingkan tanpa *Feature Selection* dalam penelitian ini?

1.3. Batasan Masalah

Di bawah ini merupakan batasan masalah pada penelitian ini:

1. Penelitian ini hanya menggunakan data dari film-film Indonesia.
2. Penelitian ini hanya membuat model prediksi untuk data film Indonesia pada tahun 2011 sampai 2020.
3. Penelitian ini hanya membahas nilai akurasi pada model prediksi yang dibuat.

1.4. Tujuan dan Manfaat Penelitian

1.4.1. Tujuan Penelitian

1. Membuat model Naïve Bayes untuk prediksi *rating* film Indonesia.
2. Membuat model K-Nearest Neighbor untuk prediksi *rating* Indonesia.
3. Meningkatkan akurasi model Naïve Bayes berdasarkan penelitian terdahulu dengan menggunakan teknik *feature selection*.
4. Meningkatkan akurasi model KNN berdasarkan penelitian terdahulu dengan menggunakan teknik *feature selection*.

1.4.2. Manfaat Penelitian

1. Sebagai kontribusi dalam penelitian *rating* pada film-film Indonesia dengan menggunakan algoritma Naïve Bayes dan KNN.
2. Menjadi referensi bahan acuan untuk metode Naïve Bayes dan KNN.
3. Menambah wawasan mengenai penggunaan teknik *data mining* dengan tahapan CRISP-DM dalam melakukan prediksi *rating* film.

U M M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A