

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Penilaian Rating dalam sebuah film merupakan hal yang terpenting dikarenakan menggambarkan kepuasan penikmat film terhadap film yang telah tonton. Dengan kemajuan teknologi seperti sekarang kita dapat dengan mudah mengetahui penilain sebuah film dengan menggunakan *platform* penampung hasil *review* penonton yaitu *Internet movie data base* (Imdb). Imdb ini merupakan basis data daring yang berkaitan dengan informasi orang-orang yang terlibat didalam pembuatan film mulai dari aktor/aktris, sutradara, penulis sampai penata rias dan soundtrack nya[1] . Telah banyak yang memanfaatkan *platform* ini untuk menjadi referensi dalam memilih film mana yang bagus atau tidak.

Banyak informasi yang ditampilkan di situs Imdb salah satunya adalah rating film, terdapat banyak sekali rating film mulai dari yang tertinggi sampai dengan yang terenda. Rating film ini dipengaruhi oleh *review* dari para penonton dari seluruh dunia. Namun kekurangan dari imdb ini adalah hanya menampilkan *trailer* dari film-film yang akan tayang dibioskop untuk dinilai oleh para *reviews* sehingga membuat penilaian berdasarkan opini tidak objektif sebab mereka para *reviews* tidak menonton keseluruhan isi dari film, hal tersebut akan menimbulkan permasalahan ketidakakuratan dalam penilaian rating yang diberikan oleh *reviews* kepada film yang ada di *platform* dan hal-hal yang mempengaruhi para *reviews* bisa menilai suatu film hanya dengan menonton *trailernya* saja [2][3]. Opini yang diberikan oleh para *reviews* perlu dilakukan analisis mendalam dengan menonton isi keseluruhan film, untuk menentukan penilaian yang objektif terhadap sebuah film yaitu dengan menggunakan analisis sentimen.

Telah banyak sekali penelitian-penelitian sebelumnya yang sudah melakukan analisis terhadap rating film, sebagai contoh "*Movie Success Prediction using MachineLearning Algorithms and theirComparison*" dan "*Ekstraksi Ulasan Sentimen Film dari Twitter dengan Naïve Bayes pada Situs Web Media Sosial Penggemar Film*"[4][5]. Penelitian yang dilakukan menggunakan data dari imdb untuk menganalisis sentimen dengan menggunakan algoritma yang berbeda. Untuk penelitian pertama menggunakan algoritma *Random Forest Accuracy* sebesar 61%, *Ada Boost Accuracy* sebesar 49.15%, *Gradient Boost Accuracy* Sebesar 56.68%, K-

Nearest Neighbors Accuracy sebesar 44.3% dan *Support Vector Machine Accuracy* sebesar 45.88%.

Untuk penelitian kedua menggunakan algoritma *Naïve Bayes Classifier Accuracy* 79.95% dan *Rule-based System Accuracy* 80.26%. Hasil dari penelitian pertama dan kedua dapat dilihat pada tabel dibawah ini:

Tabel perbandingan akurasi Algoritma Sentimen Analisis

Algoritma penelitian I	Accuracy	Algoritma penelitian II	Accuracy
Random Forest	61 %	Naïve Bayes Classifier	79.95%
Ada Boost	49.15%	Rule-based System	80.26%
Gradient Boost	56.68%		
K-Nearest Neighbors	44.3%		
suport vektor machine	45.88%		

Pada penelitian sebelumnya nilai akurasi perbandingan tertinggi yang didapat adalah sebesar 80,26% oleh karena itu perlu dilakukan pengembangan terhadap analisis *review* film untuk mendapatkan akurasi yang lebih besar dari penelitian sebelumnya dan menyatakan bahwa rating film itu bagus atau tidak berdasarkan algoritma yang digunakan. Sebab semakin besar nilai klasifikasi dan akurasi sentimen analisis maka semakin bagus [6].

Berdasarkan latar belakang yang telah dipaparkan, penulis tertarik untuk melakukan penelitian dengan judul “**Analisis Sentimen Membandingkan Algoritma *Linear Regression* dan *Decision Tree Regression* Untuk menentukan Akurasi Peringkat Film**”.

1.2 Rumusan Masalah

Rumusan masalah yang dapat dibuat berdasarkan latar belakang yang telah diuraikan adalah sebagai berikut:

1. Bagaimana penilaian rating terhadap sebuah film?
2. Berapakah akurasi yang diperoleh dari menggunakan metode analisis sentimen dalam menentukan rating IMDb?
3. Bagaimana *Platform* IMDb melakukan rating film?

1.3 Batasan Masalah

Agar penelitian yang dilakukan menjadi lebih spesifik sesuai perumusan masalah tersebut maka batasan yang ditentukan adalah sebagai berikut:

1. Datasate yang diambil dan dianalisis hanya IMDb datasite Dari *Keaggle*.
2. Metode yang dipakai adalah algoritma sentimen analisis.
3. Sentimen analisis IMDb hanya sentimen yang berhubungan peringkat film dan akurasi peringkat film.

1.4 Tujuan Penelitian

Tujuan diadakannya penelitian yaitu untuk:

1. Mengetahui apakah metode analisis sentimen mampu secara otomatis melakukan klasifikasi peringkat film pada IMDb dengan baik.
2. Mengetahui akurasi yang diperoleh dari metode analisis sentimen untuk melakukan klasifikasi peringkat film.
3. Menentukan film mana yang bagus dan tidak berdasarkan penilaian yang objektif menggunakan metode sentimen analisis.

1.5 Manfaat Penelitian

Manfaat yang dapat diberikan penelitian ini adalah:

1. Membantu menganalisis algoritma sentimen mengatasi permasalahan pada peringkat film di IMDb.
2. Menjadi referensi bagi peneliti mendatang bila akan melakukan penelitian dengan tema terkait.
3. Membantu penonton agar lebih mudah memilih film yang terbaik berdasarkan peringkat tertinggi dan scor tertinggi di IMDb berdasarkan penilain yang akurat dan presisi.