

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Film merupakan salah satu media hiburan bagi masyarakat luas. Film sendiri dapat juga berarti sebuah industri, yang mengutamakan eksistensi dan ketertarikan cerita yang dapat mengajak banyak orang terlibat. Semakin berkembangnya industri perfilman dalam negeri maupun luar negeri dari berbagai rumah produksi, berbanding lurus dengan persaingan dalam menghasilkan karya film yang menarik dan berkualitas untuk para penikmat film diseluruh dunia. Dengan kemajuan teknologi yang pesat sekarang ini, seluruh informasi tentang film–film tersebut sudah tersedia di internet. Semakin banyak informasi yang tersedia di internet, maka akan semakin sulit juga untuk menemukan informasi yang sesuai dengan kebutuhan kita. Jika informasi tersebut diolah dengan baik maka akan di dapatkan nilai tambah dari informasi tersebut, misalkan saja mengetahui apa yang di pikirkan orang lain dari informasi yang ada di internet tersebut [1].

Akan muncul masalah, apabila seseorang atau sebuah organisasi ingin mengetahui respon masyarakat terhadap sebuah film yang tersedia di internet, misalkan di situs Youtube.<sup>1</sup>Youtube adalah situs berbagi video yang memungkinkan pengguna untuk berbagi video yang dimilikinya[2]. Banyak pengguna mengunggah sebuah video atau cuplikan film (*trailer*) yang

---

<sup>1</sup> <http://www.youtube.com>

dimilikinya, kemudian pengguna lain dapat memberikan tanggapan tentang film tersebut dalam bentuk opini berupa pengalaman baik maupun buruk. Dari opini tersebut apabila kita olah dengan baik akan didapatkan sebuah manfaat yang berguna.

*Sentiment analysis* merupakan bagian dari *opinion mining*, adalah proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi [3]. Dilakukan untuk mengetahui sikap seorang pembicara atau penulis berkaitan dengan topik yang dibahas dari komentar yang di berikan. Komentar dari seseorang yang dipengaruhi oleh emosional penulis (*sentiment analysis*) nantinya akan diklasifikasikan ke dalam kelompok sentiment baik itu positif atau negatif. Salah satu algoritma untuk mengklasifikasikan komentar positif atau negatif adalah algoritma *K-Nearest Neighbor (K-NN)*. *K-NN* adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada *K-NN*. Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel latih. Diberikan titik uji, akan ditemukan sejumlah *K* objek (titik *training*) yang paling dekat dengan titik uji. Klasifikasi menggunakan *voting* terbanyak di antara klasifikasi dari *K* objek. Algoritma *K-NN* menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari sample uji yang baru [4].

Penggunaan *K-NN* dalam penelitian bidang analisis sentimen telah banyak digunakan, salah satunya ialah untuk mengkategorikan keluhan-keluhan yang diterima oleh salah satu universitas di Indonesia melalui sebuah fasilitas *E-*

*Complaint* yang dimilikinya. Karena begitu banyak data masuk dan SDM yang menanganinya terbatas, maka diperlukan pengelompokan data-data tersebut sehingga tau mana yang diprioritaskan terlebih dahulu [5].

Berdasarkan gambaran di atas, diperlukan adanya sebuah sistem yang dapat mengklasifikasikan review film berbahasa Inggris ke dalam dua *sentiment* yaitu positif dan negatif pada review film atau *trailer bergenre* aksi di situs youtube.com guna mempermudah user mengetahui kualitas dari film yang akan dinikmati dikarenakan terdapat banyak film yang diproduksi hanya untuk bersaing dengan rumah produksi film lainnya tanpa memperdulikan kualitas film itu sendiri. Kemudian pada penelitian ini penulis menggunakan *Key-Nearest Neighbor (K-NN)* sebagai metode klasifikasi tersebut.

## **1.2. Rumusan Masalah**

Berdasarkan pada permasalahan yang telah dijelaskan pada bagian latar belakang, maka rumusan masalah dapat disusun sebagai berikut:

- a. Bagaimana implementasi dari pengembangan *sentiment analysis* pada review film *bergenre* aksi di youtube dengan metode *K-Nearest Neighbor*.
- b. Berapa tingkat akurasi yang dihasilkan dari percobaan pada penerapan klasifikasi menggunakan metode *K-Nearest Neighbor* ini.

## **1.3. Batasan Masalah**

Agar memudahkan penyusunan skripsi ini dan tidak melenceng dari yang sudah direncanakan, maka penulis memberi beberapa batasan :

- a. Data yang di analisa hanya komentar yang berbahasa Inggris dalam film dan trailer bergenre aksi dari youtube.
- b. Hanya menggunakan metode *K-NN* dalam pengklasifikasiannya.
- c. Hanya menggunakan *sentiment* positif dan negatif.

#### **1.4. Tujuan**

Sesuai dengan latar belakang dan rumusan masalah yang sudah dijelaskan, tujuan dari penelitian ini adalah sebagai berikut:

- a. menerapkan metode klasifikasi *K-Nearest Neighbor* pada analisis *sentiment review* film dan *trailer bergenre* aksi di situs youtube
- b. Mengetahui tingkat akurasi dari penerapan klasifikasi dokumen komentar pada situs youtube menggunakan metode *K-Nearest Neighbor*.

#### **1.5. Manfaat**

Dari hasil penelitian ini diharapkan akan memberikan beberapa manfaat kepada pembaca dan penulis. Manfaat yang diharapkan adalah sebagai berikut :

1. Bagi Penulis
  - a. Menerapkan ilmu yang diperoleh dari Program Teknologi Informatika Universitas Dian Nuswantoro Semarang
  - b. Mendapatkan pengetahuan lebih dalam tentang perancangan dan pengembangan *sentiment analysis* pada review film dengan metode *K-NN*.
2. Bagi Pengguna
  - a. Memudahkan pengguna mengetahui lebih awal kualitas film yang akan dilihat.

- b. Dengan mengetahui jenis film yang mendapat banyak respon positif akan memudahkan produser dalam memproduksi film.