

1. PENDAHULUAN

1.1 Latar Belakang

Industri perfilman di Indonesia pada saat ini mengalami perkembangan yang pesat, film buatan anak bangsa dapat bersaing dengan film impor. Perusahaan film-film di Indonesia banyak menggunakan media sosial sebagai sarana promosi film mereka. Perkembangan teknologi membuat komunikasi menjadi lebih mudah dengan adanya sosial media seperti Facebook, Twitter, Path, Instagram dan sejenisnya. Tren pergaulan semakin cepat berganti, kini media sosial sudah menjadi bagian dari kebutuhan sehari-hari. Media sosial merupakan sebuah media dalam jaringan, dimana penggunaannya dapat mudah berinteraksi tanpa batasan jarak ataupun waktu. Percakapan smartphone mendukung perkembangan media sosial begitu cepat, salah satu media sosial yang berkembang pesat adalah Twitter.

Twitter merupakan sebuah *microblogging* yang berkembang begitu pesat, diluncurkan pada 13 Juli 2006 [1] dan jumlah pengguna di Indonesia pada Maret 2015 mencapai 50 juta pengguna [2]. Twitter merupakan salah satu media sosial yang banyak digunakan berbagai kalangan, dalam Twitter terdapat istilah *follower*, *tweets*, *direct message*, dan lainnya. Sebuah tweet mengandung pesan singkat dengan maksimum karakter mencapai 140 karakter. Dalam Twitter setiap akun memiliki data yang dapat diambil seperti jumlah *follower*, jumlah *following*, jumlah *like*, jumlah *tweet*, jumlah *retweet* dan lainnya, karena hal tersebut Twitter dapat dimanfaatkan untuk menggali data yang terdapat didalamnya.

Dalam penelitian ini digunakan teknik *data mining* dan *machine learning*, yang mana *data mining* adalah proses pencarian informasi yang berguna secara otomatis, sedangkan *machine learning* merupakan cabang ilmu yang fokus mempelajari pola dari data berdasarkan fitur-nya. Algoritma klasifikasi yang ada di *machine learning* seperti *Naïve Bayes*, *Random Forrest*, *Decision Tree*, *Support Vector Machine (SVM)*[3].

Dalam penelitian ini penulis mencoba melakukan klasifikasi tren film menggunakan algoritma *Support Vector Machine* dengan melihat jumlah penonton, algoritma ini menggunakan *hyperplane* untuk memisahkan antara kelas. Kemungkinan ada beberapa *hyperplane*, tetapi SVM berusaha untuk menemukan yang terbaik melalui pengukuran margin *hyperplane* dan mencari titik maksimalnya [5]. Hal yang membuat beda dengan penelitian sebelumnya adalah ruang lingkup dari film Indonesia, kemudian pengambilan data dari Twitter yang dilakukan secara otomatis dengan penambahan fitur *following official*, *sentiment positif*, *sentiment negative*, *following production house*, *following produser*, *following sutradara*, *following artis 1*, *following artis 2* dan *following artis 3*.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian tugas akhir ini adalah :

- A. Bagaimana menerapkan pre-processing terhadap data yang diambil dari Twitter hingga menghasilkan keluaran, dan dijadikan masukan yang sesuai untuk diolah dengan algoritma SVM ?
- B. Mencari tahu metode mana yang memiliki akurasi terbaik dengan membandingkan metode SVM dengan metode *Logistic Regression* dan metode *Simple Logistic* ?
- C. Mencari fitur yang berpengaruh terhadap terhadap klasifikasi berdasarkan pemilihan fitur yang digunakan ?

1.3 Batasan Masalah

Adapun batasan-batasan masalah yang digunakan dalam menyederhanakan permasalahan dalam penelitian ini antara lain :

- A. Data yang dipakai didapat dari Twitter, yaitu melalui tweet pengguna yang berkaitan dengan data film yang diperoleh dari badan perfilman Indonesia.
- B. Metode klasifikasi yang digunakan adalah algoritma SVM yang akan dibandingkan dengan metode *logistic regression* dan *simple logistic*.

1.4 Tujuan

Dalam menjawab bentuk permasalahan yang ada, maka tujuan dari penelitian ini antara lain :

- A. Mengetahui bagaimana proses *pre-processing* terhadap data yang didapat untuk diolah dengan algoritma SVM.
- B. Mendapatkan tingkat akurasi terhadap data yang diolah menggunakan metode klasifikasi dengan algoritma SVM.
- C. Mengetahui metode terbaik untuk mengklasifikasi.

1.5 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika penulisan sebagai berikut:

- A. BAB I PENDAHULUAN
Bab ini berisi uraian Tugas Akhir ini secara umum, meliputi latar belakang masalah, perumusan masalah, tujuan, dan sistematika penulisan.
- B. BAB II KAJIAN PUSTAKA
Bab ini membahas mengenai teori-teori pendukung yang digunakan dalam penyelesaian Tugas Akhir, yaitu *twitter*, metode *Support Vector Machine* dan perhitungan akurasi.
- C. BAB III METODE PERANCANGAN SISTEM
Bab ini menjelaskan tentang tahapan proses yang dilakukan untuk membangun sistem selama pembuatan Tugas Akhir ini. Pada bab ini juga berisi contoh implementasi sederhana dari sistem yang dibangun.
- D. BAB IV IMPLEMENTASI DAN ANALISIS HASIL
Pada bab ini akan dijelaskan tentang implementasi metode yang digunakan pada bab sebelumnya dan analisis hasil implementasi tersebut.

E. BAB V PENUTUP

Bab ini berisi kesimpulan akhir dari Tugas Akhir ini dan saran-saran untuk pengembangan lebih lanjut serta daftar referensi yang digunakan.