

1. Pendahuluan

Latar Belakang

Keberhasilan atau kegagalan sebuah film bergantung pada ulasan yang didapatkan, dimana terdapat ribuan ulasan yang mempengaruhi *rating* dari sebuah film. Pada tahun 2016 misalnya, menurut Motion Picture of America Dodd yang mengatakan bahwa rata-rata enam ratus film diproduksi setiap tahunnya dan akan terus bertambah [1]. Bisa dibayangkan berapa film yang telah ada saat ini dengan jumlah ulasan pada setiap filmnya, ditambah lagi kemudahan menonton film dengan menggunakan layanan *streaming* yang tersedia. Selain itu, film bukan hanya merupakan sumber hiburan atau *entertainment*, tetapi juga merupakan sumber utama dari pemasaran global. Ketertarikan masyarakat pada suatu film dapat diperoleh dari pendapat maupun opini yang dibagikan pada salah satu media sosial yang populer yaitu Twitter [2].

Twitter merupakan media yang mewadahi ribuan sampai jutaan opini yang disebut dengan *tweet* dan dapat dimanfaatkan untuk mengumpulkan opini masyarakat terhadap suatu film. Untuk mengetahui pengaruh dari opini tersebut, maka dilakukannya analisis sentimen pada opini yang telah didapatkan. Analisis sentimen memiliki fokus utama yaitu untuk memprediksi polaritas suatu opini yang diungkapkan dalam suatu bahasa yang kemudian akan diklasifikasikan ke dalam kategori positif atau negatif. Dengan demikian sebuah film dapat dikatakan berhasil atau gagal dari opini yang telah diklasifikasikan tersebut[3,4].

Pada penelitian sebelumnya, peneliti ingin membuat beberapa model klasifikasi untuk memprediksi keberhasilan film *box office* dimana hasil akurasi yang didapatkan setiap model kemudian dibandingkan antara satu dengan yang lainnya, diantara lain adalah Naïve bayes 59.07%, SVM 60.39%, KNN 60.76%, dan Random Forest 22.97% [5]. Oleh sebab itu, pada penelitian ini berfokus untuk membangun model klasifikasi yang optimal dengan menggabungkan algoritma SVM, Naïve Bayes, Decision Tree menjadi Hybrid. Metode Hybrid ini juga merupakan *gap* dari penelitian sebelumnya. Selain itu, penelitian ini juga berfokus untuk menghasilkan model yang dapat mengklasifikasikan opini atau *tweet* dengan optimal.

Topik dan Batasannya

Berdasarkan latar belakang yang telah disampaikan sebelumnya, maka rumusan masalah dari penelitian ini adalah mengukur perfromansi dan akurasi dari penggabungan metode SVM, Naïve Bayes, Decision Tree secara Hybrid dalam pengklasifikasian *tweet review* film. Selain itu, untuk mengetahui bagaimana cara menghasilkan model klasifikasi yang optimal. Sedangkan untuk batasan masalah pada penelitian ini menggunakan sebanyak 1023 data *tweet* pengguna Twitter mengenai *review* film Nevertheless, Riverdale, Vincenzo, dan Never Have I Ever yang diambil pada bulan Juli 2021.

Tujuan

Tujuan dari penelitian tugas akhir ini adalah untuk mengetahui sentimen mengenai film yang sedang populer pada tahun 2021 dan untuk membangun model analisis sentimen dan klasifikasi *review* film menggunakan metode Hybrid yang terdiri dari algoritma klasifikasi SVM, Naïve Bayes, dan Decision Tree.

Organisasi Tulisan

Setelah ini akan dijelaskan mengenai studi terkait penelitian serupa beserta dengan hasil yang diperoleh. Lalu, pada bagian tiga akan dijelaskan system yang dibangun peneliti dan juga penjelasan mengenai gambaran beserta teori-teori terkait penelitian ini. Pada bagian empat yaitu evaluasi, akan dijelaskan mengenai hasil pengujian yang didapatkan yaitu perhhitungan akurasi, *precision*, *recall*, *F1-Score*, dan analisis dari hasil pengujian. Bagian akhir adalah kesimpulan dari hasil penelitian dan juga saran atau masukan untuk penelitian selanjutnya.