1. Pendahuluan

Latar Belakang

Teknologi yang semakin berkembang dengan cepat dari tahun ke tahun membuat platform digital yang tersedia sebagai media informasi juga semakin berkembang. Dengan adanya kemajuan teknologi informasi ini telah mempengaruhi hampir semua aspek kehidupan manusia, dari pekerjaan hingga hiburan [1]. Sehingga membuat banyak manusia yang mulai berinteraksi di berbagai bidang untuk berbagi informasi. Salah satunya di bidang perfilman yang juga sering diinformasikan mengenai baik atau buruknya suatu film dari para penontonnya atau bisa disebut dengan mengulas suatu film. Saat ini, berbagai situs dan aplikasi media telah tersedia untuk melakukan ulasan film, salah satu yang populer dan sering dijadikan tempat melihat dan memberikan ulasan film adalah Rotten Tomatoes [2]. Situs Rotten Tomatoes adalah situs untuk melihat ulasan atau mengulas suatu film baik dari para kritikus terkenal maupun hanya seorang penonton biasa. Pemberian ulasan kepada suatu film dapat memberikan dampak yang cukup signifikan terhadap masa depan dari suatu film, tergantung dari nilai yang terkandung di ulasannya bisa bernilai positif ataupun negatif.

Dengan adanya nilai positif dan negatif dari suatu ulasan film, maka ulasannya menjadi bisa dianalisis untuk mengetahui nilai dari ulasan tersebut. Oleh karena itu, terdapat teknik yang dikenal sebagai analisis sentimen yang berguna untuk menganalisis suatu nilai tulisan. Analisis sentimen sendiri adalah bagian besar dari *Natural Language Processing* (NLP), di mana model *machine learning* akan memprediksi emosi umum yang ditampilkan dengan menganalisis informasi dari tulisan/teks menggunakan berbagai teknik [3]. Tujuan dari analisis sentimen ini adalah memeriksa apakah penonton menyukai suatu film atau tidak dengan menganalisis ulasannya [3]. Selain itu, saat dilakukan analisis sentimen terhadap ulasan film akan memberikan pemahaman yang lebih dalam mengenai film tersebut sehingga dapat berguna dalam pengambilan keputusan untuk kedepannya, seperti apakah filmnya layak untuk dibuat lanjutannya, apakah bisa dikembangkan menjadi *franchise*, bagaimana strategi pemasaran yang tepat untuk dilakukan, dan sebagainya.

Untuk melakukan analisis sentimen, sudah banyak percobaan berbagai metode *machine learning* yang dipakai dalam menganalisis sentimen terhadap ulasan film. Tentunya untuk membangun model *machine learning* yang dapat menganalisis sentimen dengan baik diperlukannya pemilihan metode yang tepat sehingga memungkinkan hasil dari analisis sentimennya memiliki akurasi yang baik. Berdasarkan penelitian sebelumnya yang dilakukan oleh Verdikha dan kawan-kawan [4], didapatkan bahwa metode DistilBERT dengan Support Vector Machine (SVM) menghasilkan nilai F1-Score tertinggi sebesar 0.7857 untuk tugas klasifikasi ujaran kebencian. Hal tersebut menunjukkan bahwa kombinasi metode DistilBERT dengan *classical machine learning*, dalam konteks ini metode SVM menunjukkan performansi yang cukup bagus sehingga memberikan peluang untuk pengembangan lebih lanjut. Kemudian, pada penelitian lainnya yang dilakukan oleh Farasalsabila dan kawan-kawan [5], didapatkan hasil dengan tingkat akurasi tertinggi sebesar 90% dari penggunaan metode BERT dan BiLSTM untuk deteksi *cyberbullying.* Namun, hasil penelitian ini menunjukan adanya *overfitting* sehingga hal ini menunjukan adanya kemungkinan pengembangan lebih lanjut untuk mengatasi *overfitting* tersebut. Berdasarkan kedua penelitian yang telah disebutkan sebelumnya, penulis melakukan pengembangan dengan menggabungkan kedua kemungkinan dari kedua penelitian tersebut yaitu mengkombinasikan DistilBERT dengan *deep learning*, dalam konteks ini yaitu BiLSTM. Jadi, pada penelitian ini akan menggunakan metode kombinasi DistilBERT dan BiLSTM untuk menganalisis sentimen ulasan film pada situs Rotten Tomatoes.

Topik dan Batasannya

Permasalahan yang dibahas pada penelitian ini adalah bagaimana hasil dari analisis sentimen menggunakan metode kombinasi DistilBERT dan BiLSTM, bagaimana kinerja metode kombinasi DistilBERT dan BiLSTM dalam menangani analisis sentimen pada ulasan film, serta bagaimana penggunaan hyperparameter tuning dapat meningkatkan kinerja metode kombinasi DistilBERT dan BiLSTM dalam menganalisis sentimen ulasan film dari situs Rotten Tomatoes. Adapun batasan dalam penelitian ini yaitu dataset yang digunakan berasal dari situs Hugging Face yang berisi ulasan dari situs Rotten Tomatoes dengan total data sebesar 10.662, yang terdiri dari 5.331 kelas positif dan 5.331 kelas negatif. Selain itu, ditambahkan juga data baru yang diambil dari situs Rotten Tomatoes sebesar 3.000 data dengan dilabeli secara manual menjadi kelas positif dan kelas negatif.

Tujuan

Penelitian ini bertujuan untuk melakukan analisis sentimen ulasan film pada situs Rotten Tomatoes menggunakan metode kombinasi dari DistilBERT dan BiLSTM, mengetahui performansi dari metode kombinasi DistilBERT dan BiLSTM saat menggunakan hyperparameter tuning maupun tidak, dan melihat keefektifan metode kombinasi DistilBERT dan BiLSTM dalam analisis sentimen terhadap ulasan film.

Organisasi Tulisan

Selanjutnya pada penelitian ini akan membahas beberapa bagian lainnya yaitu bagian 2 yang membahas mengenai studi terkait berisi teori yang mendukung penelitian ini, bagian 3 mengenai sistem yang dibangun yang berisi penjelasan dari tahap-tahap pembangunan sistem pada penelitian ini, bagian 4 merupakan hasil dan evaluasi dari penelitian yang telah dilakukan, serta pada bagian 5 berisi kesimpulan dan saran untuk penelitian kedepannya.