

1. Pendahuluan

Latar Belakang

Pesatnya perkembangan teknologi mempengaruhi pertumbuhan dan penyebaran informasi. Teknologi internet berperan penting dalam penyebaran informasi. Informasi teks yang terdapat pada web dapat dibagi menjadi dua kategori yaitu fakta atau opini. Fakta merupakan informasi yang berdasarkan objektivitas sementara opini merupakan ekspresi *sentiment* dari penulisnya[3].

Pendapat yang diungkapkan di internet akan menjadi informasi yang penting, salah satunya informasi ulasan suatu film. Pengguna internet dapat dengan mudah menemukan informasi tertulis tentang sebuah film sebagai referensi film untuk ditonton. Dalam satu bulan, 69% penonton dapat menonton satu film dan 17,5% penonton dapat menontonnya dua kali film yang sama [1]. Informasi tentang ulasan film tersebar di internet. Akibatnya, banyak pecinta film kesulitan mencari informasi yang mereka butuhkan. Karena setiap orang memiliki opininya masing-masing, informasi yang didistribusikan cenderung tidak merata. Berdasarkan permasalahan yang dihadapi oleh para penikmat film maka penelitian analisis sentimen sangat akan mempermudah pengguna. Analisis sentimen adalah jenis pemrosesan bahasa alami untuk melacak mood publik tentang produk atau topik tertentu [2]. Tujuannya untuk menentukan polaritas dalam suatu movie review berdasarkan data yang membicarakan tentang suatu movie. Teknik analisis sentimen digunakan untuk mengekstrak dan mengelola data kemudian memasukkannya ke dalam kategori negatif atau positif. Analisis sentimen termasuk golongan supervised learning. Data movie review yang sudah dikumpulkan dijadikan data training dan data testing. Selanjutnya dilakukan proses klasifikasi.

Untuk mengetahui secara langsung apa maksud dari sebuah kalimat dalam hasil review, diperlukan algoritma yang akan memproses dan memetakan informasi tersebut menjadi sebuah vektor yaitu dengan menggunakan metode Word2vec. Pada penelitian sebelumnya Word2Vec banyak digunakan untuk melakukan klasifikasi berdasarkan vektor dari data teks [4]. Sementara pada penelitian ini Word2Vec digunakan untuk melihat vektor dari sebuah paragraf atau dokumen yang kemudian akan diklasifikasi dengan menggunakan metode *Support Vector Machine* (SVM). Dalam penelitian ini, metode SVM digunakan untuk menganalisa sentimen dari hasil review dalam forum website IMDB. *Sentiment Analysis* juga merupakan bidang ilmiah yang dapat menganalisis opini, sentimen, evaluasi, penelitian, penilaian, sikap, dan emosi orang tentang entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atribut lainnya[5]. Terdapat tiga level dalam *sentiment analysis* yaitu *coarse grained sentiment analysis*, *fine-grained sentiment analysis* dan *adjective sentiment analysis* [6]. Penulis memilih metode Word2Vec untuk melakukan klasifikasi, karena metode ini dapat membantu komputer untuk mengidentifikasi kombinasi kata yang akan diklasifikasikan. Pada penelitian sebelumnya[4], penggunaan Word2Vec memiliki hasil akurasi yang baik. Proses klasifikasi dilakukan dengan menggunakan metode SVM karena adanya penelitian sebelumnya dengan menggunakan metode SVM dengan hasil akurasi yang cukup baik yaitu 85.6%[7].

Topik dan Batasannya

Topik dan batasan masalah dalam penelitian tugas akhir ini adalah mengetahui hasil pengujian performansi dari model berdasarkan skenario dalam penelitian ini. Batasan masalah pada penelitian ini adalah dilakukan analisis sentimen dengan menggunakan Word2Vec dengan metode SVM. Data yang digunakan adalah data *review* film berbahasa Inggris yang berasal dari website IMDB, dengan jumlah 10.000 data ulasan dengan label positif dan negatif.

Tujuan

Tujuan dalam penelitian ini adalah membangun sistem analisis sentimen pada data ulasan film berbahasa Inggris menggunakan Word2Vec dengan metode SVM. Serta, mengetahui performansi dari sistem penelitian yang dibangun.

Organisasi Tulisan

Pada bagian selanjutnya pada bab 2 akan membahas studi literatur yang berkaitan dengan penelitian ini. Pada bab 3 akan membahas teori dan rancangan alur sistem penelitian. Dan pada bab 4 dan bab 5 akan membahas hasil dan analisis pada penelitian ini dan juga kesimpulan penelitian.