

## ABSTRAK

Informasi mengenai pendapat, penilaian, pandangan dapat banyak kita temukan di Internet. Salah satunya adalah ulasan film yang berisi pendapat seorang penonton terhadap suatu film. Semua tanggapan dari penonton yang mengulas suatu film, tidak bisa dianggap menjadi suatu sentimen, sehingga diperlukan teknik untuk menganalisis *review* tersebut sehingga memberikan informasi yang bernilai. Analisis sentimen merupakan suatu subjek pada machine learning yang memiliki tujuan untuk mengekstrak subjektif informasi dari ulasan tekstual. Metode *machine learning Support Vector Machine (SVM)* seringkali digunakan untuk melakukan analisis sentiment. Namun, banyaknya kata yang tidak diperlukan, kata yang berulang, dan banyaknya atribut menyebabkan kinerja pengklasifikasian memberikan tingkat akurasi yang rendah. Untuk mengurangi fitur yang tidak dibutuhkan, perlu menerapkan penyeleksian fitur salah satunya dengan menggunakan Preprocessing data merupakan langkah penting untuk tugas pemrosesan bahasa alami (NLP). Ini mengubah teks menjadi bentuk yang lebih mudah dicerna sehingga algoritme pembelajaran mesin dapat bekerja lebih baik. Dalam hal ini kami melakukan jenis preprocessing berikut.. Dengan Dataset Imdb ditrain menggunakan beberapa algoritma klasifikasi machine learning seperti :

*SVM : 89%, Naive Bayes : 76%, SGD Classifier : 89%, Ridge Classifier : 88%, Decision Tree : 71%, Logistic Regression : 88%, Random Forest : 85%, KNN : 81%*. Berdasarkan hasil prediksi menggunakan beberapa algoritma klasifikasi machine learning, terdapat 2 model yang dapat digunakan untuk diaplikasikan ke web, yaitu Algoritma *SVM* dan Algoritma *SGDClassifier*.