

Abstrak

Perkembangan teknologi informasi yang begitu cepat mampu membuat data semakin bertambah setiap harinya karena banyak orang memanfaatkan kemudahan teknologi sebagai tempat atau sarana untuk mencari informasi salah satunya adalah film. Banyaknya jumlah data yang ada pada situs *review* film membuat sulitnya mendapatkan informasi yang terkandung didalamnya. Analisis sentiment dapat menjadi solusi dalam mengolah data-data tersebut karena dapat mengklasifikasikan *review* penonton kedalam kategori positif dan negatif, serta membuat bentuk dokumen yang awalnya tidak terstruktur menjadi bentuk yang lebih terstruktur sehingga dapat memperoleh sebuah informasi penting dari kualitas sebuah film tersebut. Pada Tugas Akhir ini dilibatkan beberapa pengujian untuk mengetahui pengaruh performansi dengan membandingkan tahap *preprocessing* yang dilakukan antara *preprocessing* menggunakan *stemming* dan *preprocessing* dengan menggunakan *lemmatization*. Setelah itu melakukan perbandingan antara menggunakan seleksi fitur *Information Gain* dan tidak menggunakan seleksi fitur *Information Gain*. Hasil penelitian ini didapatkan bahwa *preprocessing* dengan menggunakan *lemmatization* dan seleksi fitur *Information Gain* mendapat hasil dengan akurasi sebesar 94.00% dibandingkan dengan tanpa menggunakan seleksi fitur *Information Gain* yaitu sebesar 92.00%.

Kata kunci : analisis sentimen, review film, preprocessing, klasifikasi, support vector machine, information gain.