
Abstrak

Analisis sentimen sangat penting untuk memahami opini dan reaksi pengguna terhadap produk atau layanan. Movie review memiliki kompleksitas yang mencakup berbagai aspek seperti plot, akting, dan visual. Memahami sentimen berdasarkan review ini dapat memberikan wawasan berharga bagi pembuat film, sutradara, dan pembuat kebijakan. Namun, analisis sentimen movie review tidak dapat dikatakan mudah karena tantangan-tantangan yang dapat dihadapi. Tantangan ini bisa berupa penggunaan bahasa review yang tidak terstruktur, pemilihan fitur-fitur yang tidak tepat, dataset yang kompleks, atau metode klasifikasi yang berat.

Dalam penelitian ini, digunakan dataset IMDb yang berisi 50.000 data review serta label sentimen (positif, negatif). Sedangkan algoritma klasifikasi yang digunakan ialah Naïve Bayes, yang tepatnya ialah Multinomial Naïve Bayes. Algoritma ini didukung oleh vectorization menggunakan TF-IDF, dan feature selection menggunakan chi-squared.

Hasil penelitian menunjukkan bahwa model Multinomial Naïve Bayes dengan parameter TF-IDF yang optimal dapat menghasilkan accuracy sebesar 86.14%. Performa model tersebut meningkat menjadi 87.71% setelah mengimplementasikan chi-squared dengan parameter yang juga optimal.

Penelitian ini memberikan wawasan mengenai teknik serta tantangan dalam analisis sentimen movie review pada IMDb dengan metode klasifikasi Multinomial Naïve Bayes, serta menguji parameter pre-processing untuk menentukan dan mengevaluasi performa yang dihasilkan.

Kata kunci: Multinomial Naïve Bayes, TF-IDF, chi-squared, accuracy, IMDb