

Abstraksi

Kondisi teknologi saat ini dapat secara positif mempengaruhi bagaimana internet digunakan serta bagaimana informasi, terutama informasi tentang industri film, disebarkan secara online. Banyak ulasan film yang dapat dengan mudah ditemukan berkat kemudahan ini. Ulasan film memiliki dampak yang besar dalam berbagai cara film tersedia. Berkat kemudahan berbagai informasi di internet, jumlah ulasan film menjadi beragam. Oleh karena itu, diperlukan sebuah analisis sentimen. Pada penelitian ini, metode klasifikasi yang digunakan adalah Regresi Logistik. Metode ini dipilih karena memiliki ketepatan klasifikasi yang akurat. Pada penelitian ini, Information Gain juga dipilih sebagai seleksi fitur karena cukup baik untuk melakukan pendekatan filter dalam klasifikasi. Selanjutnya untuk ekstraksi fitur, TF-IDF dipilih karena dapat mengatasi ketidakseimbangan data pada dataset. Model terbaik yang dihasilkan dari penelitian ini adalah model yang dibangun tanpa menggunakan stemming pada tahap preprocessing, tanpa menggunakan seleksi fitur information gain, dan menggunakan parameter pada Regresi Logistik yang menghasilkan nilai f1-score sebesar 76.50%.

Kata kunci: Ulasan Film, Analisis Sentimen, Information Gain, Logistic Regression