

Sentiment Analysis pada Movie Review Berbahasa Inggris Menggunakan Support Vector Machine Classifier (SVM) dan N-gram Model

Kartiko Nurhada¹, Wicaksono¹, Adiwijaya², Mahendra Dwifabri P³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹kartikonurhadaw@students.telkomuniversity.ac.id, ²adiwijaya@telkomuniversity.ac.id,

³mahendradp@telkomuniversity.ac.id

Abstrak

Teks lebih mudah dipahami karena menggunakan bahasa yang digunakan manusia. Walaupun begitu, terkadang beberapa teks juga memiliki makna tersendiri yang perlu dipelajari lebih dalam. Sentimen adalah pandangan yang didasarkan oleh perasaan tertentu terhadap sesuatu. Kita mungkin sering menjumpai teks sentimen ketika berada di dunia maya, salah satunya ketika mencari informasi ulasan terhadap sesuatu. Metode yang mudah dilakukan untuk mengolah informasi tersebut dinamakan analisis sentimen. Analisis sentimen adalah metode untuk mengelompokkan sentimen menjadi bermakna positif, negatif, atau netral. Dengan adanya metode yang membantu mengolah informasi menjadikan kita lebih mudah dalam menyerap dan mendapatkan informasi yang kita cari. Penelitian ini bertujuan untuk mengimplementasikan metode *feature selection N-gram* untuk meningkatkan nilai akurasi model yang dibangun berdasarkan penelitian yang pernah dilakukan oleh R.Maulana (2020), yang menunjukkan peningkatan nilai akurasi pada model dengan menggunakan proses *feature selection*. Penelitian dilakukan dengan menjalankan tiga skenario yang menguji apakah metode *feature extraction TF-IDF* memiliki pengaruh pada nilai akurasi dan perbandingan metode *preprocessing stemming* dan *lemmatization*. Pada penelitian ini digunakan dataset berupa teks *movie review* berbahasa Inggris dan menggunakan metode klasifikasi *support vector machine* (SVM). Berdasarkan penelitian dari ketiga skenario diatas, didapatkan nilai *accuracy* tertinggi sebesar 88,10 % melalui skenario *feature extraction TF-IDF* dan *feature selection N-Gram* dengan parameter Unigram + Bigram.

Kata kunci : Sentimen, SVM, N-gram