

Analisis Sentimen Terhadap Project Kereta Cepat Jakarta - Bandung Menggunakan Metode SVM

Muhammad Daffa Dhiyaulhaq¹, Dr. Putu Harry Gunawan, S.Si., M.Si., M.Sc.²

Abstrak

Pertumbuhan web memberikan kontribusi yang sangat besar terkait konten yang dibuat pengguna seperti umpan balik dari pengguna, pendapat dan ulasan. Pada saat ini, banyak sekali ulasan tentang suatu hal atau berita diberbagai media platform, akan tetapi masih banyak juga ulasan yang belum meyakinkan atau bias. Pembangunan Kereta Cepat Jakarta-Bandung merupakan ikon sekaligus momentum bagi Indonesia untuk memodernisasi transportasi massal di era kemajuan yang terus menerus. Sentimen analisis adalah salah satu solusi bidang penelitian yang berbasis teks yang cocok untuk membahas masalah kepuasan berdasarkan ulasan pengguna. Dalam penelitian ini sistem akan dibuat dengan kalimat ulasan dari pengguna dan menghasilkan output berupa kelas positif dan negatif. Metode yang digunakan oleh penulis adalah pengklasifikasian menggunakan metode Support Vector Machine (SVM) dan fitur ekstraksi Word2Vec. Selain itu dilakukan perbandingan nilai akurasi antara metode Support Vector Machine, metode Naïve Bayes dan fitur ekstraksi TF-IDF. Data yang diteliti berasal dari beberapa website berita yang berisi tentang ulasan pengguna terhadap Kereta Cepat Jakarta-Bandung. Metode ini digunakan karena merepresentasikan kata-kata dalam suatu vector, selain itu proses training yang lebih cepat jika dibandingkan dengan fitur ekstraksi lainnya. Dari penelitian ini menghasilkan performansi *accuracy*, *precision*, *recall*, dan *f1-score* yaitu *accuracy* sebesar 82.74%, *precision* sebesar 75.68%, *recall* sebesar 97.67%, dan *f1-score* sebesar 85.28%. Hasil tersebut diperoleh menggunakan *hyperparameter tuning* terbaik yaitu ('C': 10, 'gamma':0.1, 'kernel': 'rbf'). Kemudian pada skenario kedua dilakukan perbandingan dengan metode Naïve Bayes. Didapatkan bahwa hasil akurasi metode Support Vector Machine menggunakan fitur ekstraksi TF-IDF mendapatkan hasil performansi yang lebih baik dan stabil dari pada ke 3 hasil performansi lainnya, yaitu sebesar 86.90%.Maka penulis simpulkan bahwa metode Support Vector Machine menggunakan fitur ekstraksi TF-IDF lebih baik jika dibandingkan dengan metode Naïve Bayes dan fitur ekstraksi Word2vec.

Kata Kunci: analisis sentimen, SVM, word2vec, ulasan