

ABSTRAK

Media sosial telah menjadi alat bagi masyarakat untuk mengekspresikan pendapat dan komentar mereka. Dalam penelitian ini, kami menggunakan pendekatan analisis sentimen untuk mengklasifikasikan komentar-komentar di Instagram. Metodologi yang kami gunakan meliputi pembuatan kamus data media sosial berbasis bahasa Indonesia, pengumpulan data media sosial, pra-pemrosesan data, pelatihan dan klasifikasi, serta penerapan model di media sosial. Kami mengkategorikan komentar-komentar tersebut ke dalam beberapa kategori, yaitu kesehatan, pendidikan, administrasi, ekonomi, infrastruktur, kewarganegaraan, sosial, dan lingkungan. Kami membangun sebuah kamus data media sosial berbasis bahasa Indonesia yang diambil dari portal berita, yang terdiri dari 2.205 catatan. Selain itu, kami juga menyaring 800 komentar di Instagram. Selanjutnya, kami melakukan pra-pemrosesan dataset dan melatihnya menggunakan dua model: regresi logistik multinomial dan gradient boosting. Hasil eksperimen kami menunjukkan bahwa model Gradient Boosting menunjukkan akurasi yang lebih tinggi dibandingkan dengan Regresi Logistik Multinomial, baik pada dataset pelatihan maupun pengujian: masing-masing 98,07% dan 73,11%. Hasil ini menunjukkan bahwa Gradient Boosting lebih dapat diandalkan daripada Regresi Logistik Multinomial untuk dataset kecil kami. Namun, kami mengidentifikasi bahwa model kami salah mengklasifikasikan komentar tertentu ke dalam kategori yang berbeda. Hasil ini membutuhkan analisis lebih lanjut untuk meningkatkan akurasi dan ketepatan kemampuan model.

Kata Kunci: media sosial, analisis sentimen, regresi logistik, gradient boosting.