ABSTRAK

Instagram adalah aplikasi berbasis seluler yang memungkinkan pengguna untuk mengambil foto atau gambar, menerapkan berbagai alat manipulasi untuk mengubah tampilan gambar, dan membaginya secara instan dengan teman-teman di berbagai situs jejaring social. Dengan kemampuannya berbagi foto atau pun gambar, Instagram dapat dengan cepatnya menjangkau generasi yang lebih muda, dan terlihat lebih menarik daripada layanan jejaring sosial lainnya bagi berbagai kalangan masyarakat termasuk *public figure*. Namun dibalik kelebihan yang dimilikinya, instagram juga memiliki suatu kekurangan yaitu adanya komentar mengganggu yang termasuk pada kategori spam pada foto atau video yang telah diunggah. Keberadaan komentar yang tidak memiliki relevansi dengan foto atau video unggahan dan memiliki tujuan untuk kepentingan pribadi atau golongan tanpa seizin pemilik posting, seperti komentar yang mempromosikan produk atau layanan, merugikan pengguna instagram khususnya pemilik posting dan membuat pengguna lainnya merasa tidak nyaman. Beberapa solusi untuk mengatasi hal tersebut juga sudah tersedia, namun semuanya masih dengan cara manual yang pastinya akan memerlukan waktu yang lama. Penelitian ini bertujuan untuk menerapkan algoritma *support vector machine* (SVM) dan *naïve bayes* (NB) untuk mendeteksi spam pada Instagram dan mengetahui perbandingan tingkat akurasi kedua metode tersebut. Metode yang digunakan adalah Metode SVM dengan kernel *linear* dan *Multinomial* NB untuk mendeteksi komentar spam pada instagram. Hasil pengujian menggunakan Balanced Dataset menunjukkan bahwa metode SVM dengan kernel linear lebih unggul dari metode Multinomial NB dengan hasil akurasi tertinggi yang didapatkan metode SVM kernel linear sebesar 85,94% dan metode Multinomial NB sebesar 82.29%. Hasil pengujian menggunakan Unbalanced Dataset menunjukkan kelemahan metode SVM dalam menangani dataset yang tidak seimbang, dengan akurasi sebesar 87.68%, dimana metode Multinomial Naïve Bayes mendapatkan tingkat akurasi sebesar 89.10%. Namun dengan *tuning parameter* C yang tepat ($C=10$), metode SVM dengan kernel Linear dapat mengimbangi metode Multinomial NB dengan akurasi sebesar 89.10%. Penelitian ini dilakukan berdasarkan penelitian terdahulu yang mempunyai karakteristik sama dalam tema penelitian. Namun penelitian ini memiliki perbedaan dari segi dataset, teknik preprocessing, metode penelitian, dan hasil penelitian. Data pada penelitian ini didapatkan dari *Instagram* menggunakan library *instaloader* yang ada di github.

Kata kunci : Spam, Instagram, Analisis Sentimen, Klasifikasi, SVM, Support Vector Machine, Kernel, Linear, Naïve Bayes, NB, Multinomial Naïve Bayes.