

## ABSTRAK

Rexzy Tarnando, 92217024

“PERBANDINGAN TINGKAT KINERJA ALGORITMA NAÏVE BAYES, K-NEAREST NEIGHBOR DAN LOGISTIC REGRESSION PADA KLASIFIKASI TEKS ATAS SENTIMEN ORANG DI TWITTER (STUDI KASUS: TRANSJAKARTA)”.

Media sosial saat ini merupakan media untuk mengungkapkan hal-hal yang dipikirkan dan dirasakan oleh masyarakat. Salah satu hal yang banyak dibicarakan adalah tanggapan dari konsumen produk atau jasa. Hal ini sangat berguna bagi perusahaan untuk mengetahui tingkat kepuasan atas produk atau jasanya. Twitter merupakan salah satu media sosial yang paling banyak digunakan oleh pengguna. Dengan fakta ini, sangat menarik bagi perusahaan untuk menggunakan data di Twitter untuk kemajuan perusahaan secara umum dalam hubungan pelanggan. Dalam penelitian ini dilakukan analisis terhadap sentimen masyarakat terhadap penggunaan Transjakarta. Kajian ini membagi sentimen komunitas menjadi tiga kelas, positif, netral dan negatif. Untuk data diambil dari Twitter dengan hasil penelitian dari bulan Juni sampai Juli 2019 dengan membagi data menjadi data latih dan data uji. Jumlah data latih 144 tweet dan data uji 36 tweet. Kemudian untuk klasifikasi teks menggunakan 3 algoritma yaitu naïve bayes, k-nearest neighbor dan regresi logistik. Kemudian setelah diperoleh hasil, selanjutnya membandingkan tingkat kinerja ketiga metode dengan mencari nilai skor f1 tertinggi menggunakan rumus micro average. Rata-rata mikro dipilih karena itu yang terbaik untuk menghitung kumpulan data yang tidak seimbang. Hasil penelitian menunjukkan metode naïve bayes memiliki skor f1 terbaik dengan nilai 0.861. Untuk nilai skor f1 terbesar berikutnya adalah metode regresi logistik dengan nilai skor f1 0,833, dan yang terakhir adalah metode k-terdekat tetangga dengan nilai skor f1 0,806.

Kata Kunci: Naïve Bayes, K-Nearest Neighbor, Logistic Regression, F1 Score, Sentiment Analysis, Transjakarta  
(xii + 68 + Appendix)

Bibliography (1994 - 2019)