
Abstrak

Twitter merupakan salah satu situs jejaring sosial populer yang diluncurkan pertama kali pada tahun 2006. Layanan ini memungkinkan pengguna untuk menyebarkan informasi secara *realtime*. Namun, informasi yang didapatkan tidak selalu berdasarkan fakta dan terkadang dengan kesengajaan digunakan untuk menyebarkan rumor yang menyebabkan ketakutan pada publik. Sehingga diperlukan upaya deteksi untuk mengatasi serta mencegah penyebaran rumor pada Twitter. Banyak penelitian dilakukan terkait deteksi rumor, namun terbatas pada Bahasa Inggris maupun Bahasa China. Pada penelitian ini penulis membangun sistem untuk melakukan deteksi rumor berbahasa Indonesia berdasarkan implementasi metode klasifikasi SVM dan metode pembobotan TF-IDF sebagai seleksi fitur. Pengumpulan data dilakukan pada periode November 2019 hingga Februari 2020 menggunakan metode *crawling by keyword* dengan proses pelabelan secara manual. Data penelitian menggunakan topik seputar pemerintah dan tren dengan 47.449 record serta kombinasi fitur berdasarkan pengguna dan tweet. Tahapan penelitian meliputi proses pengumpulan data pada situs jejaring sosial Twitter yang kemudian dilakukan preprocessing terdiri dari proses *case folding*, *url removal*, normalisasi, *stopword removal* dan *stemming*. Tahapan selanjutnya yaitu seleksi fitur, pemodelan N-Gram, klasifikasi serta evaluasi menggunakan *confusion matrix*. Berdasarkan hasil penelitian, sistem mendapatkan kinerja yang baik pada skenario pengujian menggunakan 10 % data testing serta fitur unigram dengan nilai akurasi tertinggi sebesar 78.71%. Adapun fitur Twitter yang memiliki pengaruh terhadap deteksi rumor meliputi jumlah following, jumlah like dan mention.

Kata Kunci : SVM, TF-IDF, Crawling, twitter, rumor
