

ABSTRAK

Jejaring media sosial menjadi sangat populer di kalangan masyarakat umum. Hal ini sering kali menyebabkan media sosial disalah gunakan untuk kepentingan negatif, salah satunya digunakan untuk menyebarkan pesan – pesan yang mengandung spam. Spam tersebut dari tahun ke tahun terus meningkat jumlahnya. Berdasarkan penelitian di tahun 2013, tercatat bahwa peningkatan kemunculan akun spam 355% lebih cepat dibandingkan dengan akun pada umumnya. *Spam* secara luas dapat didefinisikan sebagai penggabungan dari berbagai elemen – elemen yang berkaitan dengan perilaku komersial, psikologi penerima, konteks hukum yang luas, pertimbangan ekonomi, dan juga masalah teknik. Definisi *spam* sering kali diartikan berbeda – beda oleh setiap negara. Bahwa pada kenyataannya *spam* dapat dibagi menjadi dua karakteristik utama yaitu primer dan sekunder. Karakteristik primer merupakan karakteristik – karakteristik utama yang harus ada dalam suatu pesan untuk menandai bahwa suatu pesan tersebut adalah *spam*, sedangkan karakteristik sekunder merupakan karakteristik – karakteristik yang sering kali dikaitkan dengan *spam*, akan tetapi tidak selamanya karakteristik tersebut dibutuhkan untuk menandai bahwa suatu pesan tersebut adalah *spam*. Pada penelitian ini diusulkan *TF-IDF* dan algoritma *Multilayer Perceptron*, penggunaan algoritma *Multilayer Perceptron* karena memiliki kemampuan dan performa pembelajaran yang lebih cepat dari algoritma lain. Algoritma *Multilayer Perceptron* dapat diterapkan pada perangkat lunak ini dengan dilakukan beberapa tahapan, meliputi : pengumpulan data, proses preprocessing, proses pembobotan atau *tf-idf*, membuat model, dan melakukan train. Dari permasalahan di atas peneliti melakukan penelitian tentang *text classification* untuk mendeteksi spam di media sosial twitter menggunakan algoritma *Multilayer Perceptron* yang bertujuan untuk mendapatkan tweet netral atau yang tidak memiliki sentimen emosi dimana sering muncul pada *tweet* bertipe *spam*. Pada penelitian ini telah dilakukan analisis, berdasarkan hasil analisis, perancangan dan pembahasan yang dilakukan dapat diperoleh kesimpulan, pada penelitian ini, bahwa *tf orthography* dan *Multilayer Perceptron* lebih unggul dari yang lain dengan menghasilkan hasil akurasi tertinggi dengan 95% akurasi, 100% presisi dan 91% *recall* dalam mengklasifikasi spam pada Twitter. Algoritma *Multilayer Perceptron* memerlukan waktu training yang lebih singkat jika dibandingkan dengan algoritma yang lain. Metode *TF-IDF* dan *Multilayer Perceptron* dapat digunakan untuk melakukan klasifikasi dalam pendeteksian *spam* dan bukan *spam* pada twitter.

Kata kunci : Spam, Klasifikasi, *TF-IDF*, *Multilayer Perceptron*, Spam, Cross Validasi, Confusion Matrix