

Abstrak

Sekarang ini, keberadaan *Short Message Service* (SMS) yang menunjukkan tindakan penipuan semakin meningkat dan sangat mengganggu bagi pengguna SMS. SMS tersebut dikenal dengan SMS Spam. Oleh karenanya, pendeteksian atau filtering SMS Spam secara otomatis sangat diperlukan. Penelitian ini akan membangun sistem SMS Filtering yang mengklasifikasikan antara SMS Spam dengan bukan Spam (Ham). Metode yang digunakan dalam klasifikasi menggunakan Multinomial Naïve Bayes yang dikombinasikan dengan *feature weighting Term Frequency - Inverse Document Frequency* (TF-IDF). Sebelum klasifikasi data akan di *preprocessing* terlebih dahulu dengan menggunakan teknik tokenisasi, *slang handling*, *stopword*, dan *stemming*. Pembagian data untuk mendapatkan data training dan data testing pada proses pengujian dilakukan dengan menggunakan metode *cross validation*. Pengujian yang dilakukan yaitu membandingkan beberapa skenario pengujian berdasarkan penggunaan *preprocessing*nya. Hasil terbaik yang didapatkan setelah proses pengujian mendapat nilai *accuracy* 94.44% dengan pemilihan *preprocessing Slang Handling* dan *Stemming* yang selanjutnya akan diimplementasikan pada *mobile Android*. Implementasi *mobile Android* ditambahkan dengan rule berupa jika *sender* tidak ada di kontak *list user* maka SMS yang masuk akan melakukan proses klasifikasi. Hasil yang didapatkan pada aplikasi *mobile Android* dapat mendeteksi dan mengklasifikasikan SMS Spam dengan nilai *accuracy* 94.74%.

Kata Kunci: *SMS Spam Filter*, Klasifikasi, *multinomial naïve bayes*, *term frequency - inverse document frequency*.