

Abstrak

Pesan Singkat atau yang dikenal dengan SMS (*Short Message Service*) merupakan layanan pertukaran pesan antar pengguna layanan tersebut. Semakin banyaknya pengguna layanan SMS, tidak sedikit pihak yang memanfaatkannya untuk mendapatkan keuntungan, yaitu dengan menyebarkan SMS sampah, atau dikenal dengan SMS *spam*. Oleh karena itu, pada penelitian tugas akhir ini, penulis melakukan pengklasifikasian terhadap SMS yaitu kelas *spam* maupun *ham*. Pengklasifikasian SMS tersebut dengan menggunakan pendekatan *Collaborative Naïve Bayes* yang berorientasi pada rekomendasi beberapa pengguna dan *Content-Based Naïve Bayes* dengan melihat konten pada SMS. Data rekomendasi didapatkan dengan menyebarkan 300 SMS kepada pengguna. Untuk *Content-Based* dibutuhkan *preprocessing* sehingga konten SMS menjadi seragam, memiliki informasi penting, dan mempercepat proses komputasi. *Preprocessing* yang digunakan adalah *slang handling*, *stopword removal*, dan *stemming*. Pengujian dilakukan dengan membagi SMS menjadi data latih dan data uji sesuai dengan pembagian data *cross validation* yaitu *5-fold* dan *10-fold*. Hasil pengujian yang dilakukan menghasilkan tingkat akurasi sebesar 97.12% untuk *5-fold* dan 97.28% untuk *10-fold*.

Kata kunci: *Pengklasifikasian, Collaborative, Content-based, Naïve bayes, Preprocessing.*