

Abstrak

SMS masih menjadi salah satu pelayanan terpenting dalam media komunikasi. Namun karena biaya nya yang murah banyak yang menyalahgunakan penggunaan SMS ini yang disebut SMS *spam*. Biasa nya SMS *spam* ini dimaanfaatkan untuk melakukan promosi dan penipuan. Tidak banyak pengguna sms yang langsung percaya terhadap isi SMS ini. Tetapi ada juga yang percaya kepada isi sms tersebut.

Oleh sebab itu dibangun lah aplikasi untuk memfilter isi sms yang layak diterima oleh pengguna. Dalam tugas akhir ini dilakukan penelitian tentang SMS *Spam* Filtering. Yang bertujuan untuk mengklasifikasikan SMS menjadi dua kelas yaitu *spam* dan *ham*. Untuk melakukan penelitian ini penulis akan melewati beberapa proses. Proses pertama adalah *preprocessing* yang secara garis besar bertujuan untuk merapikan isi dari SMS tersebut. Di dalam *preprocessing* hanya akan digunakan proses *Case Folding* yaitu penyamaan karakter ke dalam huruf besar atau kecil.

Setelah semua data di *preprocessing* maka algoritma yang akan digunakan untuk proses selanjutnya adalah Dynamic Markov Compression. Algoritma ini digunakan untuk mengkompresi teks. Tahap selanjutnya adalah melakukan klasifikasi menggunakan *Minimum Cross Entropy*. Data yang akan digunakan adalah data berbahasa Indonesia. Hasil yang diinginkan untuk penelitian ini adalah nilai akurasi yang diatas 85 %.

Kata kunci : SMS Spam Filtering, Preprocessing,Algoritma Dynamic Markov Compression,Klasifikasi