

1. Pendahuluan

1.1 Latar Belakang

SMS adalah suatu fasilitas untuk mengirim dan menerima suatu pesan singkat berupa teks yang berukuran kurang dari 160 karakter melalui perangkat nirkabel. Perangkat yang dimaksud adalah telepon seluler. Semakin berkembang dan meningkatnya popularitas SMS banyak pihak-pihak yang memanfaatkan hal itu untuk media iklan bahkan penipuan yang dikenal dengan istilah *spam*. Kategori di SMS dibagi 2 menjadi *spam* dan *ham*. Kalau *ham* adalah konten SMS nya berpengaruh baik terhadap pengguna.

Dalam *Feature Engineering for Mobile SMS Spam Filtering* mengatakan banyak algoritma yang dapat digunakan untuk mengatasi kasus SMS *spam* [3]. Salah satu algoritma yang digunakan adalah penggunaan filter berbasis *compression* yang mampu menangani kasus *spam filtering* dengan tambahan klasifikasi *Minimum Cross Entropy*. Salah satu kelebihan dari pengimplementasian algoritma ini adalah memiliki rasio kompresi yang tinggi tanpa perubahan pada data asli [12].

Dalam penelitian ini akan digunakan algoritma *Dynamic Markov compression* untuk kompresi dan *Minimum Cross Entropy* untuk *spam filtering*. Di dalam algoritma klasifikasi akan digunakan indexing untuk proses penempatan data nya sehingga tidak membuang waktu untuk menghitung data yang nanti nya tidak berpengaruh terhadap pengklasifikasian nya. Metode asli akan sedikit di modifikasi dalam penelitian ini untuk menyesuaikan dengan kasus yang dihadapi [6].

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, permasalahan yang diteliti dalam tugas akhir ini adalah sebagai berikut :

1. Bagaimana mengimplementasikan metode *Statistical Data Compression* dalam melakukan klasifikasi SMS *spam* dan *ham* untuk digunakan dalam SMS *Spam Filtering* ?
2. Bagaimana pengaruh penerapan metode *Statistical Data Compression* untuk SMS *Spam Filtering* ?
3. Bagaimana tingkat akurasi yang diperoleh dari metode *Statistical Data Compression* ?

1.3 Tujuan

Tujuan yang diinginkan oleh penulis dalam pengerjaan tugas akhir ini adalah sebagai berikut :

1. Mengimplementasikan metode *Statistical Data Compression* dalam melakukan klasifikasi SMS *spam* dan *ham* untuk digunakan dalam SMS *Spam Filtering*.
2. Menganalisis pengaruh penerapan metode *Statistical Data Compression* untuk SMS *Filtering*.
3. Menghitung tingkat akurasi yang didapat dari penggunaan metode *Statistical Data Compression*.

1.4 Batasan Masalah

Batasan masalah untuk tugas akhir ini adalah :

1. Data SMS bernilai *spam* dan *ham* yang digunakan untuk proses *training* diambil dari hasil penelitian sebelumnya dengan total 180 SMS [11].
2. Data melewati 2 tahap yaitu kompresi dan klasifikasi.
3. Parameter pengukuran yang digunakan *Precision*, *Recall*, dan *Accuracy*.
4. Sistem yang diperlukan menggunakan bahasa *Java* yang akan di implementasikan di android.

1.5 Metodologi Penyelesaian Masalah

Metodologi yang digunakan untuk menyelesaikan tugas akhir ini adalah sebagai berikut :

1. Studi Literatur

Tahapan mengumpulkan data, informasi dan referensi terkait dengan permasalahan yang akan diselesaikan, termasuk landasan teori, metode dan algoritma yang akan digunakan. Di kasus ini literature yang akan dicari adalah yang memiliki hubungan dengan SMS *spam*, *Data Compression Algorithm*, *Dynamic Markov Compression*, *Minimum Cross Entropy* dan *text mining*.

2. Analisis Data

Data yang dipilih merupakan data SMS yang menggunakan bahasa indonesia. Pola dari data akan coba dikenali untuk memudahkan proses yang akan dilakukan selanjutnya. Dan diharapkan dari pemilihan proses akan dapat menghasilkan nilai akurasi yang lebih baik.

3. Perancangan Sistem

Merancang sistem berdasarkan algoritma yang dipilih. Rancangan menggunakan flow chart agar lebih mudah dipahami.

4. Implementasi

Pengimplementasian mengikuti rancangan yang sudah dibuat. Implementasi di tugas akhir ini akan dibangun dengan bahasa pemrograman *java* dan aplikasi berbasis *android*.

5. Analisis Hasil

Melakukan analisis terhadap hasil yang sudah diklasifikasikan. Dan menilai hasil dari tingkat akurasi yang dihasilkan dari metode yang dipilih.

6. Pembuatan Laporan Tugas Akhir

Membuat laporan untuk mendokumentasikan penelitian dari awal sampe akhir.