

Daftar Isi

Abstrak	i
Abstract	ii
Lembar Persembahan	iii
Kata Pengantar	v
Daftar Isi	vi
Daftar Tabel	viii
Daftar Gambar	ix
Daftar Istilah	x
1. Pendahuluan	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan.....	2
1.4 Batasan Masalah.....	2
1.5 Metodologi Penyelesaian Masalah.....	3
2. Landasan Teori.....	4
2.1 SMS (Short Message Service).....	4
2.1.1 SMS Spam dan Ham.....	4
2.2 Text Mining.....	4
2.3 Klasifikasi.....	5
2.4 Text Preprocessing.....	5
2.4.1 Case Folding.....	6
2.4.2 Tokenisasi.....	6
2.4.3 Penghapusan Stopword.....	6
2.4.4 Stemming.....	7
2.5 Feature Extraction.....	10
2.6 Naive Bayes Classifier.....	10
2.6.1 Posterior Probabilities.....	10
2.6.2 Class-conditional Probabilities.....	11
2.6.3 Prior Probabilities.....	12
2.6.4 Evidence.....	13
2.7 Algoritma FP-Growth.....	14
2.7.1 Pembangunan FP-Tree.....	14
2.7.2 Penerapan Algoritma FP-Growth.....	14
2.8 Frequent Itemset / Frequent Pattern.....	15
2.9 Evaluasi Klasifikasi SMS.....	16
3. Deskripsi Sistem.....	17
3.1 Deskripsi Umum Sistem.....	17
4. Pengujian dan Analisis.....	30
4.1 Lingkungan Implementasi Sistem.....	30

4.1.1	Spesifikasi Perangkat Keras.....	30
4.1.2	Spesifikasi Perangkat Lunak.....	30
4.2	Pengujian Sistem.....	30
4.3	Skenario Pengujian.....	30
4.3.1	Proses Pengujian.....	33
4.4	Analisis Hasil Pengujian.....	33
4.4.1	Analisis Proses Pengujian Pertama.....	33
4.4.2	Analisis Proses Pengujian Kedua.....	34
4.4.3	Analisis Proses Pengujian Ketiga.....	36
4.4.4	Analisis Proses Pengujian Keempat.....	37
4.4.5	Analisis Proses Pengujian Kelima.....	39
4.4.6	Analisis Proses Pengujian Keenam.....	40
4.4.7	Analisis Proses Pengujian Ketujuh.....	42
4.4.8	Analisis minimum support pada FP-Growth.....	43
4.4.9	Analisis Perbandingan Kedua Metode.....	45
5.	Kesimpulan dan Saran.....	47
5.1	Kesimpulan.....	47
5.2	Saran.....	47
6.	Daftar Pustaka.....	48
	LAMPIRAN A : Daftar Stopwords.....	49
	LAMPIRAN B : Daftar Fitur Hasil Ekstraksi SMS Corpus Big v.0.1.....	51
	LAMPIRAN C : Daftar Fitur Hasil Ekstraksi SMS Spam Collection v.0.1..	55
	LAMPIRAN D : Daftar Kata Slang (Tidak Baku).....	59
	LAMPIRAN E : Dataset berdasarkan karakteristik.....	62