

DAFTAR ISI

HALAMAN PERNYATAAN.....	ii
HALAMAN PENGESAHAN	ii
HALAMAN PERSEMBAHAN	v
RIWAYAT HIDUP	viii
MOTTO	ix
ABSTRACT	x
ABSTRACT	xi
KATA PENGANTAR.....	xii
DAFTAR ISI.....	xiv
DAFTAR TABEL	xviii
DAFTAR GAMBAR.....	xix
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II	6
TINJAUAN PUSTAKA.....	6
2.1 Email dan Spam.....	6

2.1.1	Email / Surat Elektronik.....	6
2.1.2	Spam dan Ham	6
2.2	Data Mining.....	7
2.2.1	Pengertian Data Mining	7
2.2.2	Fungsi Data Mining	8
2.2.3	Pengelompokan Data Mining.....	8
2.2.4	Tahapan Data Mining.....	9
2.3	Klasifikasi Text	11
2.3.1	Proses Klasifikasi.....	12
2.4	Algoritma Klasifikasi	13
2.4.1	CART (Classification and Regression Trees).....	13
2.4.2	C4.5.....	15
2.4.3	C5.0.....	17
2.5	Software Pendukung (<i>RapidMiner</i>).....	19
2.6	Tableu.....	20
2.7	Penelitian Terdahulu.....	21
BAB III.....		35
METODOLOGI PENELITIAN		35
3.1	Metode Penelitian.....	35
3.2	Teknik Pengumpulan Data	35
3.3	Tahapan Penelitian	36
3.3.1	Studi Literatur	37
3.3.2	Pengumpulan Data	37
3.3.3	Pre Processing	37
3.3.4	Ekstraksi Fitur	41

3.3.5	Pembagian Data	42
3.3.6	Implementasi Algoritma CART, C4.5, C5.0.....	43
3.3.7	Visualisasi Hasil.....	44
3.3.8	Perbandingan Kinerja Algoritma	45
3.3.9	Perhitungan Manual Algoritma (CART, C4.5 dan C5.0)	46
BAB IV		79
HASIL DAN PEMBAHASAN		79
4.1	Pengumpulan Data	79
4.2	Pre-processing	80
4.2.1	Cleaning Data (Missing Values dan Duplikat Data).....	80
4.2.2	Text Preprocessing	83
4.3	Pembagian Data.....	94
4.4	Implementasi Algoritma Klasifikasi.....	95
4.4.1	Algoritma CART (Classification and Regression Tree)	95
4.4.2	Algoritma C4.5.....	98
4.4.3	Algoritma C5.0.....	101
4.5	Visualisasi Hasil	104
4.5.1	Akurasi	105
4.5.2	Precision.....	106
4.5.3	Recall.....	107
4.5.4	Waktu Eksekusi Algoritma	108
4.6	Perbandingan Kinerja Algoritma.....	109
4.7	Pembahasan	110

BAB V.....	111
KESIMPULAN DAN SARAN	111
5.1 Kesimpulan.....	111
5.2 Saran	111
DAFTAR PUSTAKA.....	112
LAMPIRAN.....	115