

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
PERNYATAAN	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
INTISARI	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Pembatasan Masalah	3
1.3 Tujuan Penulisan	3
1.4 Tinjauan Pustaka	3
1.5 Metode Penulisan	5
1.6 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI.....	7
2.1. Matriks	7
2.2. Distribusi Variabel Random.....	10
2.2.1. Distribusi Bernoulli.....	10
2.2.2. Distribusi Binomial.....	11
2.2.3. Distribusi <i>Uniform</i>	12
2.3. Himpunan	12
2.4. Peluang Bersyarat.....	13
2.5. Kejadian Independen.....	14
2.6. Aturan Bayes	16
2.7. <i>Maximum A Posterior</i>	17
2.8. Klasifikasi.....	18
2.9. <i>Unsupervised and Supervised Learning</i>	19

2.10.	<i>Machine Learning</i>	20
2.10.1.	Fitur.....	20
2.10.2.	Pembobotan.....	21
2.11.	<i>Data Mining</i>	22
2.12.	<i>Text Mining</i>	23
2.12.1.	Pengertian <i>Text Mining</i>	23
2.12.2.	Proses <i>Text Mining</i>	24
2.13.	Berita.....	26
2.14.	<i>Laplace's Rule of Succession</i>	28
2.15.	<i>Additive Smoothing</i>	30
BAB III KLASIFIKASI TOPIK MENGGUNAKAN NAÏVE BAYES CLASSIFIER, NAÏVE BAYES MULTINOMIAL CLASSIFIER, DAN MAXIMUM ENTROPY		34
3.1.	Klasifikasi topik.....	34
3.2.	Data.....	35
3.3.	Pemilihan Fitur.....	35
3.4.	<i>K-Fold Cross Validation</i>	36
3.5.	Matriks Pasangan Dokumen.....	37
3.6.	<i>Naïve Bayes Classifier</i>	38
3.6.1.	Model probabilitas untuk <i>Naïve Bayes</i>	38
3.6.2.	Klasifikasi dokumen.....	40
3.6.3.	Model <i>Naïve Bayes Classifier</i> untuk dokumen Bernoulli.....	40
3.6.4.	Model <i>Naïve Bayes Multinomial Classifier</i>	46
3.7.	<i>Maximum Entropy</i>	56
3.7.1.	Definisi <i>Entropy</i>	57
3.7.2.	Prinsip <i>Maximum Entropy</i>	58
3.7.3.	Model probabilitas <i>Maximum Entropy</i>	58
3.7.4.	Estimasi parameter <i>Maximum Entropy</i>	65
3.7.5.	Algoritma klasifikasi dengan <i>Maximum Entropy</i>	66
BAB IV STUDI KASUS		68
4.1.	Deskripsi Kasus.....	68
4.2.	Deskripsi Data.....	69
4.3.	<i>Data Pre-Processing</i>	73
4.3.1.	Pembagian <i>Fold</i> data.....	73
4.3.2.	<i>Case Folding</i>	74

4.3.3.	Pemisahan kata (<i>Tokenization</i>)	74
4.3.4.	Penyaringan (<i>Filtrasi</i>)	74
4.3.5.	Implementasi pemilihan fitur	76
4.3.6.	Pembuatan matriks pasangan fitur-dokumen	76
4.4.	<i>Naïve Bayes Classifier</i>	77
4.5.	<i>Naïve Bayes Multinomial Classifier</i>	79
4.6.	<i>Maximum Entropy</i>	80
4.7.	Menentukan Topik Suatu Dokumen	82
4.7.1.	Hasil klasifikasi topik dari aspek metode dan jumlah token.....	83
4.7.2.	Hasil klasifikasi topik dari metode dan jumlah data	85
4.7.3.	Hasil klasifikasi topik dari aspek metode dan keseragaman jumlah data	86
4.7.4.	Hasil klasifikasi topik dari aspek metode	88
BAB V PENUTUP		91
5.1.	Kesimpulan.....	91
5.2.	Saran.....	92
DAFTAR PUSTAKA		94
LAMPIRAN.....		97