

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMBANG	xii
INTISARI	xiii
ABSTRACT	xiv
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	2
1.3. Tujuan dan Manfaat Penelitian	2
1.4. Tinjauan Pustaka	2
1.5. Metodologi Penelitian	3
1.6. Sistematika Penulisan	3
II DASAR TEORI	5
2.1. Himpunan	5
2.1.1. Himpunan dan Elemen Himpunan	5
2.1.2. Operasi-Operasi Dasar Himpunan	6
2.2. Fungsi	6
2.3. Probabilitas Bersyarat	8
2.4. <i>Machine Learning</i>	12
2.4.1. <i>Unsupervised Learning</i>	14
2.4.2. <i>Supervised Learning</i>	14
III KLASIFIKASI TEKS DENGAN METODE NAIVE BAYES	16
3.1. Klasifikasi <i>Naive Bayes</i>	16
3.2. Klasifikasi <i>Naive Bayes</i> pada Dokumen Teks	20
IV SUPERVISED CLASSIFICATION PARSIAL PADA DOKUMEN TEKS	24
4.1. Fondasi Teori <i>Supervised Classification</i> Parsial	24

4.2. Teknik yang Diusulkan	25
4.2.1. Algoritma <i>Expectation-Maximization</i>	25
4.2.2. Tahap 1 : Reinisialisasi	27
4.2.3. Tahap 2: Membangun Pengklasifikasi Akhir Menggunakan Z , N , dan U	31
4.2.4. Memilih Pengklasifikasi	32
V HASIL EKSPERIMEN	34
5.1. Evaluasi Empiris	34
5.2. Klasifikasi Teks dengan Algoritma <i>S-EM</i>	50
VI KESIMPULAN	54
A CONTOH LAMPIRAN SKRIP PROGRAM PYTHON	57
1.1. Klasifikasi <i>Naive Bayes</i>	57
1.2. Algoritma <i>Initial Expectation-Maximization</i>	62
1.3. Algoritma <i>Spy Expectation Maximization</i>	71

DAFTAR TABEL

3.1	Dataset Pai	17
5.1	Data eksperimen	34
5.2	Dataset dengan $a = 20\%$ dan $b = 50\%$	43
5.3	Dataset dengan $a = 50\%$ dan $b = 50\%$	45
5.4	Dataset dengan $a = 50\%$ dan $b = 20\%$	46
5.5	Dataset dengan $a = 50\%$ dan $b = 20\%$	48

DAFTAR GAMBAR

3.1	data <i>20NewsGroups</i>	22
3.2	Akurasi klasifikasi <i>Naive Bayes</i> pada data <i>20NewsGroups</i>	23
4.1	Tahap reinisialisasi	29
5.1	Akurasi 3 teknik dengan $a = 20\%$ dan $b = 50\%$	44
5.2	F -score 3 teknik dengan $a = 20\%$ dan $b = 50\%$	44
5.3	Akurasi 3 teknik dengan $a = 50\%$ dan $b = 50\%$	45
5.4	F -score 3 teknik dengan $a = 50\%$ dan $b = 50\%$	46
5.5	Akurasi 3 teknik dengan $a = 50\%$ dan $b = 20\%$	47
5.6	F -score 3 teknik dengan $a = 50\%$ dan $b = 20\%$	47
5.7	Rata-rata akurasi	48
5.8	Rata-rata F -score	49