

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
PERNYATAAN BEBAS PLAGIASI	iv
MOTTO	v
PRAKATA.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
INTISARI.....	xi
ABSTRACT	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat	4
1.6 Keaslian Penelitian.....	4
1.7 Metodologi Penelitian	4
1.8 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
BAB III LANDASAN TEORI.....	11
3.1 Text Mining.....	11
3.1.1 Metode klasifikasi dan <i>clustering</i>	11
3.1.2 Pra pemrosesan data.....	12
3.1.3 Pembobotan <i>term</i>	14
3.1.4 Pengukuran similaritas	15
3.2 Clustering menggunakan Agglomerative Hierarchical Clustering	17
3.2.1 <i>Complete linkage</i>	19
3.2.2 <i>Average linkage</i>	19
3.2.3 <i>Weighted linkage</i>	20
3.2.4 <i>Ward's linkage</i>	21
3.3 Silhouette score	21

3.4	N-gram	22
3.5	Bug dalam pengembangan software	23
BAB IV ANALISIS DAN RANCANGAN SISTEM.....		25
4.1	Analisis Sistem.....	25
4.2	Deskripsi Dataset	27
4.3	Pra pemrosesan.....	28
4.4	Pembentukan n-gram	32
4.5	Perhitungan Similaritas	32
4.6	Clustering Menggunakan AHC.....	35
4.7	Pengujian dan Evaluasi	35
BAB V IMPLEMENTASI.....		37
5.1	Implementasi Skenario AHC menggunakan Jaccard Coefficient.....	37
5.1.1	Implementasi pra pemrosesan data	37
5.1.2	Implementasi <i>Jaccard Coefficient</i>	40
5.1.3	Implementasi <i>Inverse Document Frequency</i>	41
5.1.4	Implementasi <i>Weighted Jaccard Coefficient</i>	44
5.1.5	Implementasi <i>cosine, euclidean</i> dan <i>manhattan</i>	44
5.1.6	Implementasi <i>n-gram</i>	45
5.1.7	Implementasi <i>Agglomerative Hierarchical Clustering</i>	46
BAB VI HASIL DAN PEMBAHASAN		48
6.1	Proses dan Hasil Clustering	48
6.2	Pengujian Perbandingan Clustering Menggunakan Agglomerative Hierarchical Clustering	52
6.2.1	Pengujian <i>AHC unigram</i>	52
6.2.2	Pengujian <i>AHC bigram</i>	53
6.2.3	Pengujian <i>AHC trigram</i>	54
6.2.4	Pengujian <i>AHC quadrigram</i>	55
6.2.5	Pengujian <i>AHC</i> dengan perhitungan jarak yang lain	56
BAB VII KESIMPULAN DAN SARAN		58
7.1	Kesimpulan	58
7.2	Saran.....	58
DAFTAR PUSTAKA		59
LAMPIRAN.....		62
Lampiran 1 Implementasi clustering menggunakan AHC.....		62