

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metodologi Penelitian	3
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
BAB III DASAR TEORI	10
3.1 Tes Uraian	10
3.2 Representasi Kata	11
3.4 <i>Text Clustering</i>	15
3.5 <i>K-means</i>	16
3.5.1 Proses <i>K-means</i>	17
3.5.2 Kelebihan <i>K-means</i>	19
3.5.3 Kekurangan <i>K-means</i>	19
3.5.4 Pemilihan jumlah kluster dengan nilai <i>Silhouette</i>	20
3.5.5 <i>Euclidean distance</i>	20
3.5.6 <i>Initial centroid</i>	21
3.5.7 <i>Probability sampling</i>	21
3.5.8 Nilai densitas	22
3.6 Evaluasi Internal Kluster	23
3.6.1 <i>Sum Squared Error (SSE)</i>	23
3.6.2 Nilai <i>Silhouette</i>	24
BAB IV ANALISIS DAN RANCANGAN	26
4.1 Analisis Permasalahan.....	26
4.2 Analisis <i>Dataset</i>	28
4.3 Rancangan Representasi Kata	31
4.4 Rancangan Inisialisasi Pusat Kluster	33
4.3.1 Inisialisasi dengan <i>probability sampling</i>	33
4.3.2 Inisialisasi dengan perhitungan densitas	35
4.5 Rancangan Proses <i>Clustering</i>	37
4.6 Rancangan Pengujian	38
BAB V IMPLEMENTASI.....	40

5.1	Spesifikasi Perangkat Keras dan Perangkat Lunak	40
5.2	Implementasi Penyiapan Data	40
5.3	Implementasi Pengubahan Kata menjadi Vektor	41
5.4	Implementasi Inisialisasi dengan <i>Probability Sampling</i>	43
5.5	Implementasi Inisialisasi dengan Parameter Densitas	45
5.6	Implementasi <i>Clustering</i> dengan <i>K-means</i>	47
5.7	Implementasi Perhitungan SSE	48
5.8	Implementasi Perhitungan Rata-rata Nilai <i>Silhouette</i>	48
BAB VI HASIL DAN PEMBAHASAN		50
6.1	Pengujian pada Vektor 50 Dimensi	50
6.1.1	Pengujian data jawaban 1	50
6.1.2	Pengujian data jawaban 2	52
6.1.3	Pengujian data jawaban 3	54
6.1.4	Pengujian data jawaban 4	56
6.1.5	Pengujian data jawaban 5	58
6.1.6	Pengujian data jawaban 6	59
6.1.7	Hasil pengujian pada vektor 50 dimensi	61
6.2	Pengujian pada Vektor 100 Dimensi	63
6.2.1	Pengujian data jawaban 1	63
6.2.2	Pengujian data jawaban 2	65
6.2.3	Pengujian data jawaban 3	67
6.2.4	Pengujian data jawaban 4	68
6.2.5	Pengujian data jawaban 5	70
6.2.6	Pengujian data jawaban 6	72
6.2.7	Hasil pengujian pada vektor 100 dimensi	73
6.3	Hasil Pengujian Vektor 50 Dimensi dan 100 Dimensi	75
6.4	Hasil <i>Clustering</i> dengan Metode Perhitungan Densitas	78
6.4.1	Hasil <i>clustering</i> data jawaban 1	78
6.4.2	Hasil <i>clustering</i> data jawaban 2	80
6.4.3	Hasil <i>clustering</i> data jawaban 3	83
6.4.4	Hasil <i>clustering</i> data jawaban 4	86
6.4.5	Hasil <i>clustering</i> data jawaban 5	88
6.4.6	Hasil <i>clustering</i> data jawaban 6	90
6.5	Hasil Perbandingan Klaster dan Label	92
BAB VII PENUTUP		94
7.1	Kesimpulan	94
7.2	Saran	95
DAFTAR PUSTAKA		96
LAMPIRAN		98
Lampiran 1. Pertanyaan		98
Lampiran 2. Kunci Jawaban		101
Lampiran 3. Cuplikan hasil <i>clustering</i>		107