

# Daftar Isi

<b>ABSTRAK</b> .....	<b>I</b>
<b>ABSTRACT</b> .....	<b>II</b>
<b>LEMBAR PERSEMBAHAN</b> .....	<b>III</b>
<b>KATA PENGANTAR</b> .....	<b>IV</b>
<b>DAFTAR ISI</b> .....	<b>V</b>
<b>DAFTAR GAMBAR</b> .....	<b>VII</b>
<b>DAFTAR TABEL</b> .....	<b>VIII</b>
<b>DAFTAR ISTILAH</b> .....	<b>IX</b>
<b>1. PENDAHULUAN</b> .....	<b>1</b>
1.1    LATAR BELAKANG .....	1
1.2    PERUMUSAN MASALAH .....	2
1.3    TUJUAN .....	2
1.4    METODOLOGI PENYELESAIAN MASALAH .....	2
<b>2. LANDASAN TEORI</b> .....	<b>4</b>
2.1    TEXT MINING .....	4
2.2    TEXT PREPROCESSING.....	4
2.2.1    Parsing dan Case Folding.....	4
2.2.2    Stopword Removal .....	4
2.2.3    Stemming .....	5
2.3    REPRESENTASI DOKUMEN .....	5
2.3.1    Pembobotan TF-IDF.....	5
2.3.1.1    TF (Term Frequency) .....	6
2.3.1.2    IDF (Inverse Document Frequency).....	6
2.3.2    Normalisasi .....	6
2.4    KLASTERISASI DOKUMEN.....	6
2.4.1    Agglomerative Hierarchical Clustering .....	7
2.4.2    Ward's Method.....	8
2.4.3    Complete Linkage Method.....	12
2.5    EVALUASI.....	12
2.5.1    Cophenetic Correlation Coefficient.....	12
2.5.2    FMeasure.....	13
<b>3. PERANCANGAN</b> .....	<b>15</b>
3.1    ANALISIS SISTEM .....	15
3.2    ANALISIS KEBUTUHAN SISTEM.....	18
3.2.1    Kebutuhan fungsional .....	18
3.2.2    Spesifikasi perangkat lunak .....	18
3.2.3    Spesifikasi perangkat keras .....	18
3.2.4    Use Case diagram.....	18

3.2.5	Sequence Diagram.....	21
3.2.6	Class Diagram .....	24
<b>4.</b>	<b>PENGUJIAN DAN ANALISA .....</b>	<b>25</b>
4.1	DATA UJI YANG DIGUNAKAN .....	25
4.2	SKENARIO PENGUJIAN.....	25
4.3	HASIL PENGUJIAN DAN ANALISIS .....	31
4.3.1	Analisis parameter masukkan.....	31
4.3.1.1	Pengaruh jumlah kluster dan kriteria koleksi dokumen terhadap nilai ESS .....	31
4.3.1.2	Pengaruh jumlah dokumen terhadap nilai CPCC.....	32
4.3.1.3	Pengaruh jumlah kluster terhadap nilai FMeasure .....	33
4.3.1.4	Pengaruh metode pengelompokkan dokumen terhadap nilai CPCC .....	34
4.3.2	Analisis Performansi.....	34
4.3.2.1	Berdasarkan FMeasure.....	35
4.3.2.2	Berdasarkan waktu eksekusi pengelompokkan dokumen.....	36
<b>5.</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>37</b>
5.1	KESIMPULAN .....	37
5.2	SARAN.....	37
	<b>DAFTAR PUSTAKA.....</b>	<b>38</b>
	<b>LAMPIRAN A : HASIL PENGUJIAN KOLEKSI DENGAN WARD'S METHOD.....</b>	<b>40</b>
	<b>LAMPIRAN B : HASIL PENGUJIAN DENGAN COMPLETE LINKAGE METHOD.....</b>	<b>42</b>
	<b>LAMPIRAN C : ILUSTRASI PERHITUNGAN ESS (ERROR SUM OF SQUARE) PADA WARD'S METHOD TERHADAP DOKUMEN UJI.....</b>	<b>44</b>