

## DAFTAR ISI

HALAMAN JUDUL .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN BEBAS PLAGIASI.....	iv
PRAKATA .....	v
DAFTAR ISI .....	vii
DAFTAR TABEL .....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN .....	xiii
ABSTRAK .....	xiv
ABSTRACT .....	xv
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 Penelitian .....	3
1.6 Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	9
3.1 Pemodelan Topik .....	9
3.1.1 Latent Dirichlet Allocation (LDA) .....	11
3.1.2 Proses Generatif.....	13
3.1.3 <i>Gibbs Sampling</i> .....	14
3.2 <i>Text Mining</i> .....	16
3.3 <i>Preprocessing</i> .....	16
3.3.1 <i>Case Folding</i> .....	16
3.3.2 <i>Tokenization</i> .....	17
3.3.3 <i>Normalization</i> .....	17
3.3.4 <i>Filtering</i> .....	18
3.3.5 <i>Stemming</i> .....	18
3.4 Evaluasi.....	19

3.4.1	Validasi Topik dengan <i>Perplexity</i> .....	19
3.4.2	Validasi Topik dengan <i>Topic Coherence</i> .....	20
BAB IV ANALISIS DAN PERANCANGAN SISTEM .....		21
4.1	Analisis Sistem.....	21
4.4.1	Analisis Kebutuhan Sistem.....	22
4.4.2	Arsitektur Sistem .....	22
4.2	Pengumpulan Data .....	24
4.3	<i>Preprocessing</i> .....	25
4.4	<i>N-Gram</i> .....	37
4.5	<i>Representatif Dictioanary</i> .....	38
4.6	Pembobotan.....	39
4.7	LDA .....	43
4.7.1	Validasi Model Topik .....	43
4.7.1	Pembentukan Model Topik.....	44
BAB V IMPLEMENTASI .....		46
5.1	Spesifikasi Sistem .....	46
5.2	Implementasi Pengambilan Data .....	46
5.3	Implementasi <i>Preprocessing</i> .....	48
5.3.1	Pembentukan Model Topik.....	48
5.3.2	Implementasi <i>Normalization</i> .....	48
5.3.3	Implementasi <i>Stemming</i> .....	51
5.4	Implementasi <i>N-Gram</i> .....	51
5.5	Implementasi <i>Dictionary Representation</i> .....	52
5.5.1	Implementasi Konversi Dokumen menjadi <i>Dictionary</i> .....	52
5.6	Implementasi Pembobotan.....	53
5.7	Implementasi Validasi Model .....	53
5.7.1	Klasifikasi Dokumen ke dalam Topik dengan Metode LDA ..	53
5.7.2	Eksperimen Pemodelan Topik dengan LDA .....	55
5.8	Implementasi Pemodelan Topik dengan <i>Latent Dirichlet Allocation</i> .....	57
5.8.1	Implementasi <i>Loading Dictionary</i> dan <i>Corpus</i> .....	57
5.8.2	Implementasi Pembentukan Model Topik.....	58
5.8.3	Implementasi Pendokumentasian <i>Logging</i> .....	58
5.9	Implementasi Distribusi Dokumen per Topik.....	59
5.9.1	<i>Import library</i> .....	59

5.9.2	<i>Loading Input</i> .....	59
5.9.3	Operasi Data .....	60
5.10	Visualisasi LDA.....	60
BAB VI HASIL DAN PEMBAHASAN.....		62
6.1	Validasi Model Topik .....	62
6.1.1	Penentuan Jumlah Iterasi Terhadap Nilai <i>Perplexity</i> .....	62
6.1.2	Penentuan Jumlah Iterasi Terhadap Nilai <i>Coherence</i> .....	63
6.1.3	Penentuan Jumlah Topik Terhadap Nilai <i>Perplexity</i> .....	64
6.1.4	Penentuan Jumlah Topik Terhadap Nilai <i>Coherence</i> .....	65
6.2	Pembentukan Model LDA .....	66
6.2.1	Hasil pembentukan Model LDA.....	68
6.3	Visualisasi LDA.....	69
6.3.1	Visualisasi LDA Menggunakan <i>pyLDAvis</i> .....	69
6.3.2	Visualisasi <i>WordCloud</i> .....	71
6.4	Pendefinisian Topik.....	74
BAB VII KESIMPULAN DAN SARAN .....		76
7.1	Kesimpulan .....	76
7.2	Saran .....	76
DAFTAR PUSTAKA.....		78
LAMPIRAN.....		80
Lampiran 1. Perhitungan Metode <i>Latent Dirichlet Allocation (LDA)</i> <i>Collapsed Gibbs Sampling</i> .....		81
Lampiran 2. Daftar <i>List Stopwords</i> .....		96
Lampiran 3. Hasil Pemodelan Topik .....		97
Lampiran 4. Visualisasi <i>pyLDAvis</i> .....		98