

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	i
HALAMAN PENGESAHAN PEMBIMBING	ii
HALAMAN PENGESAHAN PENGUJI.....	iii
PERNYATAAN KARYA ASLI TUGAS AKHIR.....	iv
PERNYATAAN BEBAS PLAGIASI.....	v
HALAMAN PERSEMBAHAN.....	vi
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
1.6 Tahapan Penelitian	3
1.6.1 Metodologi Penelitian	3
1.6.2 Metodologi Pengembangan Sistem.....	3
1.7 Sistematika Penulisan	3
BAB II TINJAUAN LITERATUR	5
2.1 Skripsi	5
2.2 Abstrak	5
2.3 <i>Topic Modeling</i>	5
2.4 <i>Latent Dirichlet Allocation</i>	6
2.5 <i>Gibbs Sampling</i>	7
2.6 <i>Bigram Model</i>	8
2.7 <i>Trigram Model</i>	9
2.8 <i>Coherence Score</i>	9
2.9 <i>Waterfall</i>	12
2.10 Penelitian Sebelumnya.....	13
BAB III METODOLOGI PENELITIAN	15
3.1 Metodologi Penelitian	15
3.1.1 Pengumpulan Data	15
3.1.2 <i>Preprocessing</i>	18
3.1.3 Pembuatan Model	18
3.1.4 Evaluasi	19
3.2 Metodologi Pengembangan Sistem.....	20
3.2.1 Analisis Kebutuhan Sistem	20

3.2.2	Perancangan Antar Muka.....	20
BAB IV HASIL DAN PEMBAHASAN.....		22
4.1	Hasil Penelitian	22
4.1.1	Pengumpulan Data	22
4.1.2	<i>Preprocessing</i>	27
4.1.3	Pembuatan Model	37
4.1.4	Evaluasi.....	44
4.1.5	Pengujian.....	49
4.1.6	Rancangan Sistem	54
4.2	Pembahasan.....	55
BAB V KESIMPULAN DAN SARAN		56
5.1	Kesimpulan	56
5.2	Saran.....	56
DAFTAR PUSTAKA.....		57

DAFTAR TABEL

Tabel 2.1 Penelitian Sebelumnya.....	14
Tabel 3.1 Contoh sampel dari isi dataset	17
Tabel 4.1 Daftar kategori	25
Tabel 4.2 Isi data dari <i>web scraping</i>	26
Tabel 4.3 Data abstrak	27
Tabel 4.4 Data abstrak setelah <i>remove punctuation</i>	30
Tabel 4.5 Data abstrak setelah <i>case folding</i>	31
Tabel 4.6 Data abstrak setelah <i>remove stopword</i>	33
Tabel 4.7 Data abstrak setelah <i>remove abstract in English translated</i>	34
Tabel 4.8 Data abstrak setelah <i>remove kata kunci</i>	35
Tabel 4.9 Data abstrak setelah <i>tokenizing</i>	37
Tabel 4.10 Data abstrak setelah dilakukan <i>bigram</i>	38
Tabel 4.11 Contoh perhitungan <i>bigram</i> pada kalimat pertama.....	38
Tabel 4.12 Data abstrak setelah dilakukan <i>trigram</i>	39
Tabel 4.13 Contoh perhitungan <i>trigram</i> pada kalimat pertama	39
Tabel 4.14 Data <i>dictionary</i>	40
Tabel 4.15 Data <i>corpus</i>	42
Tabel 4.16 Inisiasi jumlah topik awal	43
Tabel 4.17 Hasil perhitungan probabilitas <i>gibbs sampling</i>	43
Tabel 4.18 Hasil jumlah topik setelah <i>gibbs sampling</i>	43
Tabel 4.19 Contoh dari matriks kemunculan pasangan kata	45
Tabel 4.20 Hasil perhitungan NPMI	46
Tabel 4.21 Hasil <i>Coherence Score</i>	47
Tabel 4.22 Hasil <i>topic modeling</i> untuk kategori <i>Agriculture (General)</i>	49
Tabel 4.23 Sampel dokumen uji	50
Tabel 4.24 Hasil validasi dokumen uji	51

DAFTAR GAMBAR

Gambar 2.1 Model LDA (Blei, Ng, & Jordan, 2003)	6
Gambar 3.1 Metodologi Penelitian	15
Gambar 3.2 <i>Flowchart</i> proses <i>web scraping</i>	16
Gambar 3.3 Metodologi Pengembangan Sistem <i>Waterfall</i>	20
Gambar 3.4 Rancangan Antar Muka	21
Gambar 4.1 Isi data dari RSS 2.0.....	24
Gambar 4.2 Grafik <i>Coherence Score</i> dari kategori <i>Agriculture (General)</i>	49
Gambar 4.3 Halaman Sistem	54
Gambar 4.4 Halaman Sistem Lanjutan	54