

DAFTAR ISI

HALAMAN PENGESAHAN	ii
PERNYATAAN BEBAS PLAGIASI	iii
PRAKATA	iv
DAFTAR ISI	v
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	x
INTISARI	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan.....	4
1.5 Manfaat.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Sistem Temu Balik Informasi atau <i>Information Retrieval</i> (IR)	5
2.2 <i>Vector Space Model</i> (VSM) dan <i>Latent Semantic Analysis</i> (LSA).....	8
BAB III LANDASAN TEORI	11
3.1 <i>Big Data</i>	11
3.2 Sistem Temu Balik Informasi (<i>Information Retrieval System</i>).....	14
3.3 <i>Topic Modeling</i>	21
BAB IV METODOLOGI PENELITIAN.....	31
4.1 Alat Dan Bahan	33
4.2 Persiapan Dataset	33

4.2.1 Pengumpulan Data.....	34
4.2.2 Penggabungan Data	36
4.3 Arsitektur Pengembangan Model <i>Information Retrieveal</i> (IR) dengan menggunakan model VSM, LSA, dan Kombinasi VSM-LSA	37
4.3.1 Pembuatan <i>Vector Space Model</i> (VSM).....	41
4.3.2 Pembuatan Model <i>Latent Semantic Analysis</i> (LSA)	52
4.3.3 Pembuatan Model Kombinasi, <i>Vector Space Model</i> (VSM) dan <i>Latent Semantic Analysis</i> (LSA).....	55
4.4 Evaluasi	58
4.4.1 Evaluasi Efektivitas	58
4.4.2 Evaluasi Efisiensi	60
4.5 Penarikan Kesimpulan.....	61
BAB V IMPLEMENTASI	62
5.1 Persiapan Dataset	62
5.2 Proses Penggabungan Data	63
5.3 Proses <i>Preprocessing</i> Data.....	65
5.4 Proses Perbandingan Statistik Data Sebelum dan Sesudah <i>Preprocessing</i>	67
5.5 Proses Pembobotan dan Analisis Dokumen	68
5.6 Penerapan Model VSM	70
5.7 Penerapan Model LSA	71
5.8 Penerapan Model Kombinasi VSM dan LSA	76
5.9 Evaluasi Model VSM, LSA, dan Kombinasi VSM+LSA.....	77
BAB VI HASIL DAN PEMBAHASAN.....	84
6.1 Hasil Perolehan <i>Dataset</i>	84
6.2 Hasil Penggabungan Data.....	85
6.3 Hasil <i>Preprocessing</i>	86
6.3.1 Hasil Penghapusan Gelar Akademis, Angka Penulis, dan Tanda Baca	86

6.3.2 Hasil <i>Tokenisasi</i>	87
6.3.3 Hasil <i>Case Folding</i>	88
6.3.4 Hasil <i>Stopword</i>	89
6.4 Hasil Pemodelan <i>Informatoin Retrieval</i>	91
6.4.1 Hasil <i>Vector Space Model (VSM)</i>	91
6.4.2 Hasil <i>Latent Semantic Analysis (LSA)</i>	104
6.4.3 Hasil Kombinasi <i>Vector Space Model (VSM)</i> dan <i>Latent Semantic Analysis (LSA)</i>	116
6.4.4 Perbandingan Kinerja Setiap Model.....	128
BAB VII PENUTUP	132
7.1 Kesimpulan.....	132
7.2 Saran	133
DAFTAR PUSTAKA.....	134
LAMPIRAN	138