

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	iii
PERNYATAAN BEBAS PLAGIASI.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xiii
INTISARI.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	3
1.5 Manfaat.....	3
BAB II TINJAUAN PUSTAKA.....	4
BAB III LANDASAN TEORI.....	9
3.1 <i>Text Mining</i>	9
3.3.1 Tahapan <i>text mining</i>	9
3.2 Sistem Temu Kembali Informasi.....	11
3.3 <i>Vector Space Model (VSM)</i>	13
3.3.1 <i>Term – frequency</i> dan <i>inverse document frequency (tf – idf)</i>	14
3.3.2 Menghitung jarak dokumen dan <i>query</i>	15
3.3.3 Menghitung similaritas <i>query</i> dan dokumen	16
3.3.4 Menghitung <i>cosine similarity</i>	17
3.4 <i>Query Expansion</i>	17
3.5 <i>Thesaurus</i>	18

3.6	<i>Association Rule</i>	20
BAB IV METODOLOGI PENELITIAN		26
4.1	Deskripsi Penelitian.....	26
4.2	Tahapan Penelitian	26
4.3	Perancangan Model	27
4.4	Evaluasi dan Pengujian.....	32
BAB V IMPLEMENTASI.....		34
5.1	Data.....	34
5.2	<i>Preprocessing</i> Data	34
5.3	<i>Vector Space Model</i>	37
5.3.1	<i>Term-frequency</i> dan <i>inverse document frequency</i>	37
5.3.2	<i>Document vector</i>	39
5.4	Perhitungan <i>Cosine Similarity</i>	40
5.5	Rangking Dokumen Temu Balik Informasi	41
5.6	Implementasi <i>Query Expansion</i>	41
5.6.1	<i>Query expansion</i> dengan <i>thesaurus</i>	42
5.6.2	<i>Query expansion</i> dengan <i>association rule fp-growth</i>	43
BAB VI HASIL DAN PEMBAHASAN		48
6.1	Pengumpulan Data.....	48
6.2	Hasil <i>Preprocessing</i> Data	48
6.3	Hasil <i>Vector Space Model</i>	49
6.3.1	<i>Term-frequency</i> dan <i>inverse document frequency</i>	49
6.3.2	<i>Document vector</i>	52
6.4	Hasil <i>Cosine Similarity</i>	52
6.5	Hasil Rangking Dokumen Temu Balik Informasi.....	53
6.6	<i>Query Expansion</i> Dengan <i>Thesaurus</i>	54
6.7	<i>Query Expansion</i> Dengan <i>Association Rule</i>	54
6.8	Hasil Evaluasi dan Pengujian	58



BAB VII KESIMPULAN DAN SARAN	74
7.1 Kesimpulan.....	74
7.2 Saran.....	74
DAFTAR PUSTAKA	76