

## Daftar Isi

Lembar Pernyataan .....	ii
Lembar Pengesahan .....	iii
Abstrak .....	iv
<i>Abstract</i> .....	v
Lembar Persembahan.....	vi
Kata Pengantar .....	vii
Daftar Isi .....	viii
Daftar Gambar.....	x
Daftar Tabel .....	xi
Daftar Istilah .....	xii
1. Pendahuluan.....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan.....	2
1.4 Batasan Masalah.....	2
1.5 Metodologi Penyelesaian Masalah.....	3
1.6 Sistematika Penulisan.....	3
2. Tinjauan Pustaka.....	5
2.1 Semantik .....	5
2.2 Keterkaitan Semantik ( <i>Semantic Relatedness</i> ).....	5
2.3 Kesamaan Kata ( <i>Word Similarity</i> ) .....	5
2.4 Latent Semantic Analysis (LSA).....	6
2.5 PMI-IR .....	6
2.6 Path-Based.....	7
2.7 Resnik dan Lin .....	8
2.7.1. Resnik .....	8
2.7.2. Lin .....	8
2.8 PMI ( <i>Pointwise Mutual Information</i> ).....	8
2.9 SOC-PMI (Second Order Co-occurrence Pointwise Mutual Information) .....	10
2.9.1 Memilih nilai $\beta$ dan $\gamma$ .....	12
2.9.2 Contoh Penerapan SOC-PMI .....	12
2.10 Simlex-999 .....	14
2.11 Brown Corpus .....	15
2.12 Wordsim-353.....	16
2.13 Miller and Charles .....	16
2.14 Korelasi Pearson.....	16
3. Perancangan Sistem .....	18
3.1 Gambaran Umum Sistem .....	18

3.2	<i>Preprocessing</i> Data .....	19
3.3	Pengitungan Bobot Kata.....	20
3.4	Spesifikasi Sistem dan Fungsionalitas Sistem .....	21
3.4.1	Spesifikasi Perangkat Keras.....	21
3.4.2	Spesifikasi Perangkat Lunak.....	21
3.4.3	Fungsionalitas Sistem .....	21
3.5	Validasi Sistem.....	21
3.5.1	Validasi Pembacaan Korpus .....	22
3.5.2	Validasi Pembacaan Kata Unik Gold Standard (Simlex-999 dan Wordsim353) .....	22
3.5.3	Validasi Pembentukan Konteks Kata dan Frekuensi Bobot Kata.....	23
3.5.4	Validasi Penghitungan Nilai SOC-PMI .....	23
3.5.5	Validasi Nilai Kesamaan (Similarity) .....	23
4.	Pengujian dan Analisis .....	25
4.1	Pengujian Sistem .....	25
4.1.1	Tujuan Pengujian .....	25
4.1.2	Dataset yang Digunakan dalam Pengujian.....	25
4.1.3	Skenario Pengujian .....	27
4.2	Metrik Evaluasi .....	27
4.3	Analisis Hasil Pengujian .....	27
4.3.1	Perbandingan Nilai Korelasi Nilai Kesamaan Semantik SOC-PMI dengan Metode Lain.....	27
4.3.2	Analisis Perbandingan Nilai Korelasi Berdasarkan Nilai Kesamaan ( <i>similarity</i> ) yang dihasilkan Sistem .....	28
4.3.3	Analisis Parameter yang Mempengaruhi Nilai Kesamaan yang dihasilkan Sistem dan Nilai Korelasinya .....	38
5.	Kesimpulan dan Saran .....	40
5.1	Kesimpulan.....	40
5.2	Saran .....	40