

DAFTAR ISI

LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI.....	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN.....	xiv
DAFTAR SIMBOL	xv
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Metode Penelitian.....	5
1.6 Sistematika Penulisan.....	5
BAB II	7
LANDASAN TEORI.....	7
2.1 <i>Natural Language Processing (NLP)</i>	7

2.2	<i>Preprocessing</i>	9
2.2.1	Pembersihan teks.....	10
2.2.2	Tokenisasi	10
2.2.3	<i>Stopword Removal</i>	11
2.2.4	<i>Normalisasi</i>	12
2.2.5	<i>Case Folding</i>	13
2.2.6	<i>Stemming</i> atau lemmatisasi.....	14
2.3	<i>Semantic Similarity</i>	14
2.4	<i>Word Embedding</i>	16
2.5	<i>Word2Vec</i>	17
2.5.1	<i>Skip-Gram</i>	18
2.5.2	<i>Continuous Bag of Words (CBOW)</i>	19
2.5.3	<i>Hierarchical Softmax</i>	21
BAB III		22
ANALISIS SEMANTIC SIMILARITY ANTAR KATA MENGGUNAKAN MODEL WORD2VEC		22
3.1	Pengumpulan Data.....	23
3.2	<i>Preprocessing</i>	24
3.2.1	<i>Cleaning Dataset</i>	24
3.2.2	Tokenisasi	24
3.2.3	<i>Stopword Removal</i>	25
3.2.4	<i>Case Folding</i>	26
3.3	<i>Train Word2vec Model</i>	26
3.4	Pengukuran <i>Similarity</i> dengan <i>Cosine Similarity</i>	28
BAB IV		30
STUDI KASUS DAN ANALISIS		30

4.1	<i>Dataset Penelitian</i>	30
4.2	<i>Training Model Word2Vec</i>	32
4.3	<i>Evaluasi Hasil Training</i>	33
4.4	<i>Analisis Kualitatif Top-N Similar Words</i>	38
4.5	<i>Visualisasi t-SNE</i>	42
BAB V		45
KESIMPULAN DAN SARAN		45
5.1	Kesimpulan.....	45
5.2	Saran	46
DAFTAR PUSTAKA		47
LAMPIRAN		

