

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.....	xi
DAFTAR TABEL	xii
DAFTAR SIMBOL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Metode Penelitian.....	5
1.7 Sistematika Penulisan.....	6
BAB II LANDASAN TEORI	8
2.1 Data Mining.....	8
2.2 Text Mining.....	10
2.3 Natural Language Processing (NLP).....	11
2.4 Text Pre-Processing.....	11
2.4.1. Case Folding	11
2.4.2. Tokenisasi.....	12
2.4.3. Stopword Removal	12
2.4.4. Removal Punctuation & Numbers.....	13
2.4.5. Lemmatisasi.....	13
2.5 Word Embedding	14

2.5.1	Word2Vec.....	16
2.5.2	Glove	17
2.5.3	FastText	18
2.6	N-gram dalam Pemrosesan Teks	21
2.7	Dot Product Vektor	23
2.8	Python.....	23
BAB III	METODE PENELITIAN	24
3.1	Dataset.....	25
3.2	Text Pre-Processing.....	26
3.2.1	Case Folding.....	27
3.2.2	Tokenizing.....	28
3.2.3	Stopword Removal	28
3.2.4	Remove Punctuation & Numbers.....	28
3.2.5	Lemmatizing.....	29
3.3	FastText.....	29
3.4	Skip-Gram	30
3.5	Cosine Similarity.....	33
BAB IV	EKSPERIMEN DAN ANALISIS.....	35
4.1	Validasi Metode	35
4.2	Penerapan model pada korpus utama terjemahan Al-Qur'an.....	55
4.3	<i>Pre-Processing</i>	57
4.4	Training Data.....	58
BAB V	PENUTUP.....	72
5.1	KESIMPULAN	72
5.2	SARAN	73
	DAFTAR PUSTAKA.....	74
	RIWAYAT HIDUP	77
	LAMPIRAN.....	80