

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN</b>	<b>iii</b>
<b>HALAMAN PERNYATAAN</b>	<b>iv</b>
<b>PRAKATA</b>	<b>v</b>
<b>DAFTAR ISI</b>	<b>viii</b>
<b>DAFTAR TABEL</b>	<b>x</b>
<b>DAFTAR GAMBAR</b>	<b>xii</b>
<b>INTISARI</b>	<b>xiii</b>
<b>ABSTRACT</b>	<b>xiv</b>
<b>I PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang . . . . .	1
1.2 Rumusan Masalah . . . . .	2
1.3 Batasan Masalah . . . . .	2
1.4 Tujuan Penelitian . . . . .	2
1.5 Manfaat Penelitian . . . . .	3
1.6 Metodologi Penelitian . . . . .	3
1.7 Sistematika Penulisan . . . . .	4
<b>II TINJAUAN PUSTAKA</b>	<b>6</b>
<b>III LANDASAN TEORI</b>	<b>11</b>
3.1 Twitter . . . . .	11
3.2 Normalisasi . . . . .	11
3.3 Word Representation . . . . .	13
3.4 Word Embedding . . . . .	14
3.4.1 Word2Vec . . . . .	14
3.4.2 FastText . . . . .	18
3.5 Levenshtein Distance . . . . .	18
3.6 Jaro-Winkler Similarity . . . . .	19
3.7 Tes McNemar . . . . .	20
3.8 Confusion Matrix . . . . .	21

<b>IV Analisis dan Perancangan Model</b>	<b>23</b>
4.1 Dekripsi Umum Penelitian . . . . .	23
4.2 Pengumpulan Data . . . . .	23
4.3 Analisis Sistem . . . . .	25
4.4 Rancangan Sistem . . . . .	26
4.4.1 Rancangan Preprocessing . . . . .	26
4.4.2 Rancangan Sistem . . . . .	27
4.4.3 Rancangan Proses Normalisasi . . . . .	31
4.4.4 Rancangan Perbandingan Model dan Penentuan Model . . . . .	33
4.4.5 Rancangan Pengujian . . . . .	34
<b>V IMPLEMENTASI</b>	<b>36</b>
5.1 Lingkungan Implementasi . . . . .	36
5.1.1 Spesifikasi perangkat lunak . . . . .	36
5.1.2 Spesifikasi perangkat keras . . . . .	37
5.2 Implementasi Scrapping data . . . . .	37
5.2.1 Implementasi Scrapping data Twitter . . . . .	37
5.2.2 Implementasi Scrapping data KBBI . . . . .	38
5.3 Implementasi Preprocessing . . . . .	40
5.3.1 Filter data yang tidak diperlukan dan Menghapus duplicate . . . . .	40
5.3.2 <i>Language Filtering</i> . . . . .	41
5.3.3 <i>Cleaning</i> dan lowercasing . . . . .	42
5.3.4 Membagi data . . . . .	43
5.3.5 Penyusunan Korpus . . . . .	44
5.4 Implementasi Penyusunan KBBI . . . . .	46
5.5 Implementasi Pelatihan Model . . . . .	46
5.5.1 pelatihan model Word2Vec . . . . .	46
5.5.2 Pelatihan Model FastText . . . . .	49
5.6 Implementasi Normalisasi . . . . .	52
5.6.1 Pencarian Kandidat . . . . .	52
5.6.2 Pengecekan dengan Treshold2 . . . . .	52
5.6.3 Perhitungan kedekatan kata dengan Jaro winkler dan Levenshtein . . . . .	53
5.6.4 Normalisasi . . . . .	54
5.7 Implementasi Normalisasi dengan setiap Model . . . . .	54
5.8 Implementasi Penentuan Treshold . . . . .	56

5.8.1	Implementasi Treshold1 . . . . .	57
5.8.2	Implementasi Treshold2 . . . . .	57
5.8.3	Implementasi Treshold3 . . . . .	58
5.9	Implementasi Model Akhir . . . . .	58
5.10	Implementasi Pengujian . . . . .	61
<b>VI</b>	<b>Hasil dan Pembahasan</b>	<b>63</b>
6.1	Treshold . . . . .	63
6.1.1	Treshold1 . . . . .	64
6.1.2	Treshold2 . . . . .	65
6.1.3	Treshold3 . . . . .	66
6.2	Hyperparameter . . . . .	68
6.2.1	Pengaruh <i>Learning Rate</i> . . . . .	68
6.2.2	Pengaruh ukuran window . . . . .	69
6.2.3	Pengaruh epoch . . . . .	71
6.2.4	Pengaruh ukuran vektor . . . . .	73
6.3	Evaluasi Model Menggunakan Set Parameter Terbaik . . . . .	74
6.4	Contoh Hasil Evaluasi . . . . .	77
6.5	Evaluasi dengan Data Corpus . . . . .	80
<b>VII</b>	<b>Penutup</b>	<b>82</b>
7.1	Kesimpulan . . . . .	82
7.2	Saran . . . . .	82
	<b>DAFTAR PUSTAKA</b>	<b>88</b>