

DAFTAR ISI

HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN.....	iv
PRAKATA	v
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
INTISARI.....	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
BAB III LANDASAN TEORI	14
3.1 <i>Natural Language Processing (NLP)</i>	14
3.2 Pra-pemrosesan Teks	14
3.3 Augmentasi Data Teks.....	18
3.4 FastText.....	21
3.5 <i>Long Short-Term Memory (LSTM)</i>	24
3.6 Evaluasi	29
3.7 Ujaran Kebencian (<i>Hate Speech</i>)	32
3.8 Deteksi Ujaran Kebencian.....	33

BAB IV METODOLOGI PENELITIAN	34
4.1 Deskripsi Umum Penelitian	34
4.2 Tahapan Penelitian.....	34
4.3 Analisis Permasalahan	35
4.4 Pengumpulan Data.....	36
4.5 Pra-pemrosesan Teks	38
4.6 Augmentasi Data	39
4.7 Ekstraksi Fitur	40
4.8 Pengembangan Model.....	41
4.9 Evaluasi Model.....	43
BAB V IMPLEMENTASI	44
5.1 Spesifikasi <i>Hardware</i> dan <i>Software</i>	44
5.2 Implementasi Pengumpulan Data	44
5.3 Implementasi Pra-pemrosesan Teks	45
5.4 Implementasi Augmentasi Data	50
5.5 Implementasi Pembangunan dan Pelatihan Model.....	59
BAB VI HASIL PENELITIAN & PEMBAHASAN	68
6.1 Hasil Pengumpulan Data.....	68
6.2 Hasil Augmentasi Data	69
6.3 Hasil <i>Hyperparameter Tuning</i> pada Model	70
6.4 Hasil Evaluasi Model.....	73
BAB VII KESIMPULAN & SARAN	80
7.1 Kesimpulan.....	80
7.2 Saran	80
DAFTAR PUSTAKA	81

DAFTAR TABEL

Tabel 2.1 Tinjauan pustaka deteksi ujaran kebencian (<i>hate speech</i>).....	10
Tabel 3.1 Contoh <i>Case Folding</i>	15
Tabel 3.2 Contoh <i>Data Cleaning</i>	16
Tabel 3.3 Contoh <i>Text Normalization</i>	16
Tabel 3.4 Contoh <i>Stemming</i>	17
Tabel 3.5 Contoh <i>Stopwords Removal</i>	17
Tabel 4.1 Contoh Dataset Ujaran Kebencian (Susanto et al., 2024)	37
Tabel 4.2 Contoh Kalimat Hasil Augmentasi dengan EDA.....	40
Tabel 6.1 Hasil <i>Hyperparameter Tuning</i> pada Setiap <i>Trial</i>	71
Tabel 6.2 Performa Model Klasifikasi Ujaran Kebencian	73

DAFTAR GAMBAR

Gambar 3.1	Arsitektur fastText (Zolotov dan Kung, 2017).....	22
Gambar 3.2	Arsitektur LSTM (Ingolfsson, T. M., 2021)	25
Gambar 3.3	Algoritma LSTM	27
Gambar 3.4	Arsitektur BiLSTM (Mikhailov & Kashevnik, 2021)	28
Gambar 3.5	Algoritma BiLSTM	29
Gambar 3.6	<i>Confusion Matrix</i>	30
Gambar 4.1	Tahapan Penelitian.....	35
Gambar 4.2	Pesebaran Data pada Dataset IndoToxic2024 (Label 0 untuk Non Ujaran Kebencian dan Label 1 untuk Ujaran Kebencian)	36
Gambar 4.3	Arsitektur Model Deteksi Ujaran Kebencian BiLSTM	41
Gambar 5.1	Implementasi Pembagian Data (<i>Data Splitting</i>).....	45
Gambar 5.2	Implementasi <i>Case Folding</i>	46
Gambar 5.3	Implementasi <i>Data Cleaning</i>	47
Gambar 5.4	Implementasi <i>Text Normalization</i>	48
Gambar 5.5	Implementasi <i>Stemming</i>	49
Gambar 5.6	Implementasi <i>Stopwords Removal</i>	49
Gambar 5.7	Implementasi Fungsi <i>get_synonyms()</i>	51
Gambar 5.8	Implementasi Fungsi <i>synonym_replacement()</i>	52
Gambar 5.9	Implementasi Fungsi <i>add_word()</i>	52
Gambar 5.10	Implementasi Fungsi <i>random_insertion()</i>	53
Gambar 5.11	Implementasi Fungsi <i>swap_word()</i>	54
Gambar 5.12	Implementasi Fungsi <i>random_swap()</i>	54
Gambar 5.13	Implementasi Fungsi <i>random_deletion()</i>	55
Gambar 5.14	Implementasi Keseluruhan Metode EDA	57
Gambar 5.15	Implementasi Tokenisasi dan <i>Padding</i>	59
Gambar 5.16	Implementasi Pembuatan Matrix Embedding	61
Gambar 5.17	Implementasi Pembuatan Arsitektur Model.....	63
Gambar 5.18	Implementasi Pelatihan Model	64
Gambar 5.19	Implementasi <i>Hyperparameter Tuning</i>	66

Gambar 6.1 Distribusi Label pada Data Train, Data Validasi, dan Data Test (Label 0 untuk Non Ujaran Kebencian dan Label 1 untuk Ujaran Kebencian)	68
Gambar 6.2 Perbandingan Distribusi Label pada Data Train Sebelum dan Setelah dilakukan Data Augmentasi	69
Gambar 6.3 Grafik Performa Model Klasifikasi Ujaran Kebencian	75
Gambar 6.4 Perbandingan <i>Confusion Matrix</i> Model <i>Baseline</i> (a) dan Model dengan Kombinasi Operasi EDA (b)	77