

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>HALAMAN PENGESAHAN .....</b>	<b>ii</b>
<b>PERNYATAAN BEBAS PLAGIASI .....</b>	<b>iii</b>
<b>PRAKATA .....</b>	<b>iv</b>
<b>DAFTAR ISI.....</b>	<b>vi</b>
<b>DAFTAR TABEL .....</b>	<b>ix</b>
<b>DAFTAR GAMBAR.....</b>	<b>x</b>
<b>INTISARI .....</b>	<b>xii</b>
<b>ABSTRACT .....</b>	<b>xiii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah .....	3
1.5 Manfaat Penelitian .....	4
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>5</b>
2.1 Perkembangan Model BERT .....	5
2.2 Penelitian Terkait.....	6
<b>BAB III LANDASAN TEORI.....</b>	<b>10</b>
3.1 Klasifikasi Teks .....	10
3.2 Text Preprocessing .....	10
3.2.1 <i>Case Folding</i> .....	11
3.2.2 <i>Noise Removal</i> .....	11
3.2.3 Slang dan Singkatan .....	12
3.2.4 <i>Tokenizing</i> .....	12
3.2.5 <i>Stemming</i> .....	14
3.2.6 <i>Stopwords Removal</i> .....	14
3.3 Word Embedding.....	15
3.3.1 BERT .....	15
3.3.2 Word2Vec.....	21
3.4 <i>Recurrent Neural Network (RNN)</i> .....	22

3.4.1	<i>Long Short-Term Memory (LSTM)</i> .....	24
3.5	Hugging Face.....	25
3.6	TensorFlow .....	25
3.7	Scikit-learn .....	26
3.8	<i>Hyperparameter</i> .....	26
3.8.1	<i>Epoch</i> .....	26
3.8.2	<i>Batch Size</i> .....	26
3.8.3	<i>Learning Rate</i> .....	28
3.8.4	<i>Probabilitas Dropout</i> .....	29
3.8.5	<i>Metode Pooling</i> .....	29
3.8.6	<i>Activation Function</i> .....	30
3.8.7	<i>Loss Function</i> .....	33
3.9	Confusion Matrix.....	34
<b>BAB IV ANALISIS DAN PERANCANGAN .....</b>		<b>38</b>
4.1	Deskripsi Umum Model .....	38
4.2	Tahapan Penelitian .....	38
4.3	Alat dan Bahan .....	39
4.4	Pembuatan <i>Dataset</i> .....	40
4.5	Rancangan Model Klasifikasi Teks.....	42
4.5.1	Rancangan <i>Text Preprocessing</i> .....	42
4.5.2	Rancangan Model IndoBERT-LSTM .....	43
4.5.3	Rancangan Model <i>Baseline</i> Word2Vec-LSTM.....	45
4.5.4	Rancangan Model <i>Baseline Fine-tuned</i> IndoBERT .....	46
4.5.5	Perhitungan <i>Loss</i> .....	47
4.5.6	Rancangan Evaluasi Model .....	49
<b>BAB V IMPLEMENTASI.....</b>		<b>51</b>
5.1	Pembuatan <i>Dataset</i> .....	51
5.1.1	<i>Crawling</i> Data Twitter.....	51
5.1.2	Pembersihan dan Pemisahan Data.....	53
5.1.3	Penggabungan Data .....	54
5.2	Pengembangan Model .....	55
5.2.1	Persiapan <i>Dataset</i> .....	55

5.2.2	Model IndoBERT-LSTM .....	56
5.2.3	Model <i>Baseline</i> Word2Vec-LSTM .....	59
5.2.4	Model <i>Baseline Fine-tuned</i> IndoBERT .....	63
5.2.5	Pelatihan dan Validasi Model .....	64
5.2.6	Evaluasi Model .....	65
<b>BAB VI HASIL DAN PEMBAHASAN .....</b>		<b>67</b>
6.1	Pembuatan <i>Dataset</i> .....	67
6.2	Hasil Model IndoBERT-LSTM .....	68
6.2.1	Hasil pada <i>Dataset</i> tidak Termodifikasi .....	68
6.2.2	Hasil pada <i>Dataset</i> Termodifikasi .....	71
6.3	Hasil Model <i>Baseline</i> Word2Vec-LSTM .....	73
6.3.1	Hasil pada <i>Dataset</i> tidak Termodifikasi .....	73
6.3.2	Hasil pada <i>Dataset</i> Termodifikasi .....	74
6.4	Hasil Model <i>Baseline Fine-tuned</i> IndoBERT .....	76
6.4.1	Hasil pada <i>Dataset</i> tidak Termodifikasi .....	76
6.4.2	Hasil pada <i>Dataset</i> Termodifikasi .....	78
6.5	Perbandingan Model .....	80
6.5.1	Perbandingan Penggunaan <i>Word Embedding</i> .....	81
6.5.2	Perbandingan Penggunaan Metode Pengklasifikasi .....	82
<b>BAB VII KESIMPULAN DAN SARAN .....</b>		<b>83</b>
7.1	Kesimpulan .....	83
7.2	Saran .....	84
<b>DAFTAR PUSTAKA .....</b>		<b>85</b>