

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
PERNYATAAN BEBAS PLAGIASI	ii
DAFTAR ISI.....	iii
DAFTAR TABEL.....	vi
DAFTAR GAMBAR	vii
DAFTAR SIMBOL	viii
INTISARI	ix
<i>ABSTRACT</i>	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Penelitian	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
BAB III LANDASAN TEORI.....	12
3.1 Data Mining	12
3.2 <i>Text Mining</i>	12
3.3 <i>Text Preprocessing</i>	12
3.3.1 <i>Case-folding</i>	13
3.3.2 Tokenisasi	13
3.3.3 <i>Abbreviation</i>	13
3.3.4 <i>Stopwords removal</i>	13
3.3.5 <i>Stemming</i>	13
3.3.6 <i>Normalization</i>	14
3.3.7 <i>Feature Extraction</i>	14
3.4 <i>Text Labelling</i>	14
3.5 <i>Natural Language Processing</i>	15
3.6 Analisis Sentimen	15
3.7 Analisis Sentimen Berbasis Aspek	15
3.8 Data Imbalance	15
3.9 Model Transformator, BERT, dan IndoBERT	16
3.10 Evaluasi Model	17
BAB IV METODOLOGI PENELITIAN.....	20

4.1	Gambaran Umum.....	20
4.2	CRISP-DM.....	21
4.2.1	<i>Data Understanding</i>	21
4.2.2	<i>Data Preparation</i>	21
4.2.3	<i>Data imbalance</i>	21
4.2.4	<i>Modelling</i>	22
4.2.5	Evaluasi Model.....	26
BAB V	IMPLEMENTASI.....	27
5.1	Lingkungan Implementasi.....	27
5.2	Pengumpulan Data.....	27
5.2.1	Proses <i>Crawling Data</i>	27
5.2.2	Penggabungan Dataset <i>Crawling</i>	28
5.3	Implementasi Pelabelan Aspek Menggunakan Lexicon.....	28
5.3.1	Pelabelan Aspek Menggunakan Lexicon.....	28
5.3.2	<i>Exploratory Data Analysis (EDA)</i> Dataset Aspek.....	29
5.4	Implementasi <i>Preprocessing</i> Dataset Aspek.....	32
5.4.1	<i>Preprocessing</i> untuk Model IndoBERT.....	33
5.4.2	<i>Preprocessing</i> untuk Model BiLSTM.....	33
5.5	Implementasi Pembagian Data (Dataset Split).....	34
5.6	Implementasi Model Klasifikasi Aspek.....	35
5.6.1	Implementasi Model IndoBERT untuk Klasifikasi Aspek.....	35
5.6.2	Implementasi Model BiLSTM untuk Klasifikasi Aspek.....	39
5.7	Implementasi Pelabelan Sentimen dan EDA Data Mentah.....	42
5.7.1	Pelabelan Sentimen Menggunakan <i>Lexicon</i>	42
5.7.2	<i>Exploratory Data Analysis (EDA)</i> Data Mentah.....	45
5.8	Implementasi <i>Preprocessing</i> dan EDA Dataset Sentimen.....	46
5.8.1	<i>Preprocessing</i> Sentimen untuk IndoBERT.....	46
5.8.2	<i>Preprocessing</i> Dataset Sentimen untuk BiLSTM.....	47
5.8.3	<i>Exploratory Data Analysis (EDA)</i> Dataset Sentimen Setelah <i>Preprocessing</i>	47
5.9	Implementasi Pembagian Dataset Sentimen.....	49
5.10	Implementasi Model Klasifikasi Sentimen.....	50
5.10.1	Implementasi Model Sentimen IndoBERT.....	50
5.10.2	Implementasi Model Sentimen BiLSTM.....	52
BAB VI	HASIL DAN PEMBAHASAN.....	55
6.1	Karakteristik Dataset dan Permasalahan Data.....	55
6.1.1	Distribusi Label Aspek.....	55

6.1.2 Distribusi Label Sentimen.....	56
6.1.3 Permasalahan Dataset dan Dampaknya terhadap model.....	57
6.2 Hasil Klasifikasi Model Aspek	57
6.2.2 Hasil Klasifikasi Aspek Model IndoBERT	58
6.2.3 Hasil Klasifikasi Aspek Model BiLSTM.....	60
6.3 Hasil Klasifikasi Sentimen Menggunakan IndoBERT	61
6.3.1 Skenario Dataset Sentimen Negatif 100%	62
6.3.2 Skenario Dataset Sentimen Negatif 50%	63
6.3.3 Skenario Dataset Sentimen Negatif 10%	64
6.3.4 Skenario Tanpa Label Netral	65
6.3.5 Analisis Pengaruh Penggunaan <i>Class Weighting</i>	66
6.4 Hasil Klasifikasi Sentimen Menggunakan BiLSTM.....	69
6.5 Perbandingan Kinerja Model IndoBERT dan BiLSTM.....	71
6.5.1 Perbandingan Akurasi Keseluruhan.....	71
6.5.3 Perbandingan <i>F1-Score</i> per Kelas Sentimen.....	73
6.6 Pembahasan dan Keterbatasan Penelitian	75
6.6.1 Dampak Penggunaan <i>Lexicon-Based Labelling</i>	75
6.6.2 Tidak Adanya Label “ <i>True Neutral</i> ”.....	75
6.6.3 Ketergantungan Model terhadap Distribusi Data.....	76
6.6.4 Potensi Bias Sentimen Negatif.....	76
BAB VII KESIMPULAN DAN SARAN.....	77
7.1 Kesimpulan	77
7.2 Saran	78
DAFTAR PUSTAKA	80