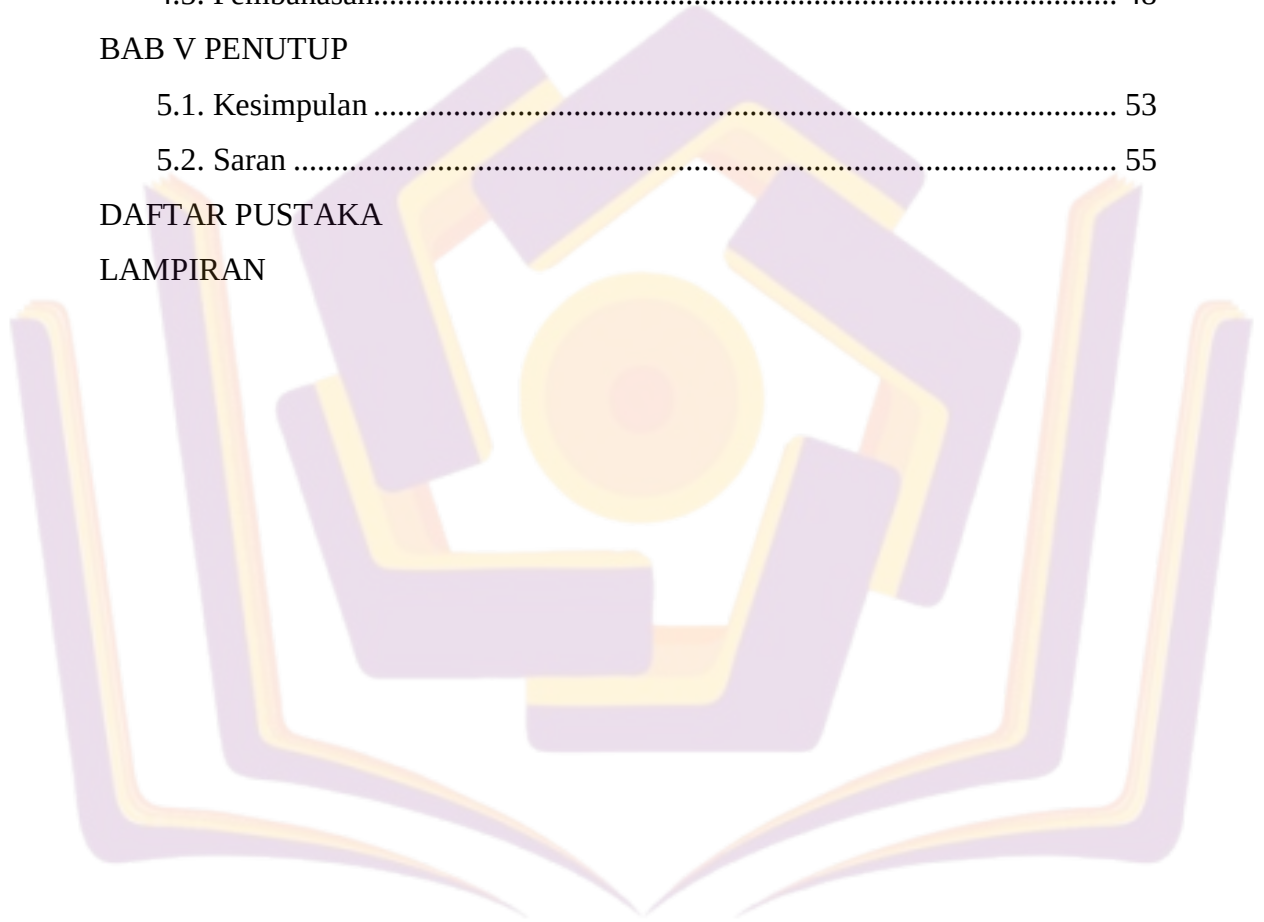


DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERNYATAAN KEASLIAN	v
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR ISTILAH	xiv
INTISARI.....	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN	
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah.....	5
1.3. Batasan Masalah	6
1.4. Tujuan Penelitian	7
1.5. Manfaat Penelitian	7
1.6. Hipotesis	8
BAB II TINJAUAN PUSTAKA	
2.1. Tinjauan Pustaka.....	10
2.2. Keaslian Penelitian.....	16
2.3. Landasan Teori.....	19
BAB III METODE PENELITIAN	
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	30
3.2. Metode Pengumpulan Data.....	31
3.3. Metode Analisis Data.....	33

3.4. Alur Penelitian	37
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	
4.1. Hasil Evaluasi Agregat	41
4.2. Analisis Performa Per-Kelas.....	43
4.3. Analisis <i>Confusion Matrix</i>	45
4.4. Dinamika <i>Training</i>	47
4.5. Pembahasan.....	48
BAB V PENUTUP	
5.1. Kesimpulan	53
5.2. Saran	55
DAFTAR PUSTAKA	
LAMPIRAN	



DAFTAR TABEL

Tabel 2.1. Matriks literatur review dan posisi penelitian Perbandingan Strategi <i>Monolingual</i> (IndoBERT) dan Berbasis Terjemahan (NLLB-200 + RoBERTa) untuk Analisis Sentimen Opini Pilpres 2024 di X.....	17
Tabel 3.1. Distribusi kelas sentimen pada dataset keseluruhan (n = 1.336)	32
Tabel 3.2. Pembagian dataset.....	33
Tabel 3.3. Konfigurasi terjemahan NLLB-200	36
Tabel 3.4. <i>Fine-tuning Hyperparameter</i>	37
Tabel 4.1. Perbandingan performa agregat pada set pengujian (n = 134).....	41
Tabel 4.2. Performa per-kelas pada set pengujian	43
Tabel 4.3. Confusion matrix IndoBERT (Strategi A)	45
Tabel 4.4. Confusion matrix RoBERTa → EN/NLLB-200 (Strategi B).....	46
Tabel Lampiran 2.1. Struktur kolom <i>dataset</i> Sentiment1.csv.....	69
Tabel Lampiran 2.2. Distribusi kelas sentimen pada dataset keseluruhan (n=1.336)	70
Tabel Lampiran 2.3. Pembagian dataset menjadi subset pelatihan, validasi, dan pengujian	70
Tabel Lampiran 2.4. Contoh cuitan per kelas sentimen (teks asli, dipotong)	71

DAFTAR GAMBAR

Gambar 3.1. Diagram alur penelitian	39
Gambar 4.1. <i>Confusion matrix</i> IndoBERT (Strategi A)	45
Gambar 4.2. <i>Confusion matrix</i> RoBERTa → EN/NLLB-200 (Strategi B)	46
Gambar 4.3. Perbandingan Metrik dan <i>Loss</i>	47



DAFTAR ISTILAH

Istilah	Singkatan	Definisi Singkatan
<i>Attention</i>	-	Mekanisme pada <i>transformer</i> yang memberi bobot relatif pada setiap token agar model fokus pada bagian <i>input</i> yang paling relevan.
BERT	<i>Bidirectional Encoder Representations from Transformers</i>	Model bahasa berbasis <i>encoder</i> <i>transformer</i> yang membaca teks secara dua arah; dilatih dengan <i>masking</i> dan <i>next-sentence prediction</i> .
<i>Confusion Matrix</i>	-	Matriks yang merangkum jumlah prediksi benar dan salah untuk setiap kelas, menjadi dasar perhitungan presisi, <i>recall</i> , dan F1.
<i>Fine-Tuning</i>	-	Pelatihan lanjutan model yang telah dilatih sebelumnya pada <i>dataset</i> spesifik tugas menggunakan bobot awal tersebut sebagai titik mulai.
<i>F1-Score</i>	-	Rata-rata harmonik dari presisi dan <i>recall</i> ; menyeimbangkan kedua metrik dalam satu angka.
IndoBERT	-	Model BERT yang dilatih khusus pada korpus bahasa Indonesia; varian yang digunakan adalah <i>indobenchmark/indobert-base-p2</i> .
NLLB-200	<i>No Language Left Behind</i> (200 bahasa)	Model terjemahan mesin neural multibahasa buatan Meta AI; varian yang digunakan adalah <i>facebook/nllb-200-distilled-600M</i> .
<i>Precision</i>	-	Proporsi prediksi positif yang benar-benar positif ($TP / (TP + FP)$).
RoBERTa	<i>Robustly Optimized BERT Pretraining Approach</i>	Varian BERT yang dilatih dengan optimasi dan data lebih besar, tanpa <i>next-sentence prediction</i> ; varian yang digunakan adalah <i>roberta-base</i> .
<i>Self-Attention</i>	-	Bentuk <i>attention</i> di mana setiap token dalam satu urutan menilai relevansi seluruh <i>token</i> lain pada urutan yang sama.
Sentimen	-	Sikap, opini, atau emosi pengirim pesan yang umumnya dikategorikan dengan konotasi positif, netral, atau negatif.
<i>Tokenizing</i> / Tokenisasi	-	Unit sub-kata hasil pemecahan teks oleh <i>tokenizer model</i> ; proses mengubah teks menjadi rangkaian <i>token</i> numerik.
<i>Transformer</i>	-	Arsitektur neural berbasis <i>attention non-recurrent</i> (Vaswani dkk., 2017) menjadi landasan bagi BERT, RoBERTa, dan IndoBERT.
NMT	<i>Neural Machine Translation</i>	Metode terjemahan bahasa otomatis berbasis satu <i>neural network</i> tiruan yang mengonversi teks secara langsung dari satu bahasa ke bahasa lain (Stahlberg, 2020).