



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
INTISARI	xiii
ABSTRACT	xiv
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Batasan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Tinjauan Pustaka	3
1.5. Metode Penulisan	5
1.6. Sistematika Penulisan	5
II DASAR TEORI	7
2.1. Berita	7
2.2. Hoaks	7
2.3. Probabilitas	8
2.4. Variabel Random	9
2.5. Bias	11
2.6. Matriks	12
2.6.1. Pengertian Matriks	12
2.6.2. Jenis-jenis Matriks	12
2.6.3. Operasi Matriks	14
2.6.4. Transpose Matriks	15
2.7. Vektor	16
2.7.1. <i>Dot Product</i>	17
2.8. Uji Proporsi Satu Populasi	17
2.9. <i>Machine Learning</i>	18
2.10. <i>Deep Learning</i>	19
2.11. Fungsi Aktivasi	21
2.12. <i>Natural Language Processing (NLP)</i>	22



2.13. <i>Text Preprocessing</i>	22
2.14. <i>Word Cloud</i>	23
2.15. <i>N-Grams</i>	24
III BERT dan INDOBERT	25
3.1. <i>Klasifikasi</i>	25
3.2. <i>Recurrent Neural Network (RNN)</i>	26
3.3. <i>Long Short Term Memory (LSTM)</i>	26
3.4. <i>Bidirectional Long Short Term Memory (BiLSTM)</i>	32
3.5. <i>Transformer</i>	35
3.5.1. <i>Encoder</i>	36
3.5.2. <i>Decoder</i>	36
3.5.3. <i>Attention Mechanism</i>	36
3.6. <i>Bidirectional Encoder Representations from Transformers (BERT)</i>	40
3.6.1. <i>Representasi input BERT</i>	41
3.6.2. <i>Tahap Pre-Training</i>	43
3.6.3. <i>Tahap Fine Tuning</i>	44
3.7. <i>Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT)</i>	45
3.8. <i>Perbandingan LSTM, BiLSTM, BERT, dan IndoBERT</i>	46
3.8.1. <i>Arsitektur</i>	46
3.8.2. <i>Efisiensi</i>	46
3.8.3. <i>Ukuran dan Kompleksitas Metode</i>	46
3.9. <i>Hyperparameters</i>	47
3.10. <i>Evaluasi Performa</i>	48
3.11. <i>Perhitungan Matematis Sederhana pada Transformer</i>	50
3.12. <i>Alur Penelitian</i>	55
IV STUDI KASUS	58
4.1. <i>Alat Analisis</i>	58
4.2. <i>Deskripsi Data</i>	58
4.3. <i>Text Preprocessing</i>	60
4.3.1. <i>Case Folding</i>	60
4.3.2. <i>Cleaning Text</i>	61
4.3.3. <i>Stop Words Removal</i>	61
4.3.4. <i>Stemming</i>	62
4.4. <i>Eksplorasi Data</i>	62
4.4.1. <i>Jumlah Data</i>	62
4.4.2. <i>Jumlah Kata dalam Berita</i>	64
4.4.3. <i>Word Cloud</i>	65
4.4.4. <i>N-Grams</i>	67



4.5. Pengaturan Metode	69
4.6. Performa Klasifikasi	70
4.6.1. BERT	70
4.6.2. IndoBERT	70
4.7. Perbandingan Performa Klasifikasi	71
4.8. <i>Confusion Matrix</i>	72
4.8.1. BERT	73
4.8.2. IndoBERT	74
4.8.3. LSTM	75
4.8.4. BiLSTM	76
V PENUTUP	77
5.1. Kesimpulan	77
5.2. Saran	78
DAFTAR PUSTAKA	79
A Lampiran Rangkuman Dataset	85
B Lampiran Kode Google Colab	88