

DAFTAR ISI

HALAMAN PENGESAHAN PEMBIMBING	iii
HALAMAN PENGESAHAN PENGUJI.....	iv
SURAT PERNYATAAN.....	v
KARYA ASLI TUGAS AKHIR.....	v
PERNYATAAN BEBAS PLAGIASI.....	vi
ABSTRAK.....	vii
ABSTRACT	viii
KATA PENGANTAR	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Tahapan Penelitian.....	4
1.7 Sistematika Penulisan	5
BAB II.....	7
TINJAUAN LITERATUR.....	7
2.1 <i>Short Message Service (SMS)</i>	7
2.2 Spam Pada SMS	7
2.3 SMS Spam Penipuan (<i>Smishing</i>).....	7
2.4 <i>Natural Language Processing (NLP)</i>	8
2.5 BERT	9
2.5.1 Arsitektur Model	9
2.5.2 Representasi Input	10
2.5.3 Mekanisme <i>Pre-training</i>	11
2.5.4 <i>Fine-tuning</i> pada BERT.....	12
2.5.5 Lapisan Klasifikasi dan Softmax.....	12
2.6 IndoBERT.....	13
2.7 IndoBERTweet.....	14
2.8 Multilingual BERT	15
2.9 <i>Transfer Learning</i>	15
2.10 <i>Fine-tuning</i>	16
2.11 <i>Text Preprocessing</i>	17
2.11.1 <i>Case Folding</i>	17

2.11.2	Pembersihan Teks	17
2.11.3	Tokenisasi	18
2.11.4	Normalisasi Teks	18
2.12	Stratified K-Fold Cross Validation	18
2.13	Pertimbangan Teks <i>Preprocessing</i>	20
2.14	<i>Confusion Matrix</i>	20
2.14.1	Matriks Evaluasi.....	21
2.15	Penelitian Terdahulu	22
BAB III	27
METODOLOGI PENELITIAN	27
3.1	Pengambilan Data.....	28
3.1.1	Dataset Utama.....	28
3.1.2	Dataset Uji Eksternal	29
3.2	<i>Preprocessing</i>	30
3.3	Stratified 5-Fold Cross Validation	35
3.4	Tokenisasi BERT	35
3.4.1	Model yang Digunakan	36
3.5	<i>Fine-tuning</i>	36
3.5.1	Tipe Tugas	36
3.5.2	Jumlah kelas	37
3.5.3	Hyperparameter Utama.....	37
3.5.4	Mekanisme <i>Fine-tuning</i>	37
3.5.5	Klasifikasi Kelas (<i>Softmax</i> dan <i>Argmax</i>).....	39
3.6	Evaluasi Model.....	40
3.6.1	Evaluasi Internal	40
3.6.2	Evaluasi Eksternal	42
3.6.3	Perhitungan Rata - Rata dan Standar Deviasi.....	43
3.6.4	Pengukuran waktu pelatihan.....	45
BAB IV	46
HASIL DAN PEMBAHASAN	46
4.1	Implementasi	46
4.1.1	Pengambilan Dataset	46
4.1.2	Data <i>Preprocessing</i>	48
4.1.3	Stratified 5-Fold Cross Validation	50
4.1.4	Tokenisasi	51
4.1.5	Implementasi Model yang Digunakan.....	53
4.1.6	Konfigurasi Pelatihan	54

4.2	Pengujian	55
4.2.1	Pengujian Model Menggunakan Cross Validation	56
4.2.2	Pengujian Model Tanpa Cross Validation.....	58
4.2.3	Pengujian Waktu Pelatihan	59
4.2.4	Pengujian <i>Robustness</i>	59
4.2.5	Analisis Kesalahan Prediksi Model.....	60
4.3	Pembahasan	62
BAB V	64
PENUTUP	64
5.1	Kesimpulan.....	64
5.2	Saran	64
DAFTAR PUSTAKA	66