

DAFTAR ISI

HALAMAN PENGESAHAN TESIS	iii
HALAMAN PERNYATAAN	iv
PRAKATA	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
INTISARI.....	xiii
ABSTRACT	xiv
BAB I	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
BAB II.....	5
TINJAUAN PUSTAKA.....	5
BAB III.....	24
LANDASAN TEORI	24
3.1 Analisis Sentimen.....	24
3.2 <i>Natural Language Processing</i> (NLP)	24
3.3 Augmentasi Data.....	27
3.3.1 Augmentasi Data dengan <i>Easy Data Augmentation</i>	27
3.3.2 Augmentasi Data dengan <i>Back-Translation</i>	29
3.3.3 Augmentasi Data dengan <i>Masked Language Modeling</i>	30
3.4 BERT.....	31
3.4.1 <i>Pre-training</i>	32

3.4.2	<i>Fine-tuning</i>	33
3.5	<i>Confusion Matrix</i>	35
3.6	Kebijakan TAPERA	36
BAB IV		37
METODOLOGI PENELITIAN		37
4.1	Studi Pustaka	37
4.2	Perancangan Sistem	37
4.2.1	Dataset	38
4.2.2	Pengumpulan Data	38
4.2.3	Pra-pemrosesan Data	38
4.2.4	Pelabelan Data	39
4.2.5	Augmentasi Data	42
4.2.6	Pemodelan	44
4.2.7	Validasi Model	46
4.2.8	Evaluasi Model	47
BAB V		49
IMPLEMENTASI		49
5.1	Lingkungan Implementasi	49
5.2	Pengumpulan Data	49
5.2.1	<i>Data Crawling</i>	49
5.2.2	Penggabungan Data	50
5.3	Pra-pemrosesan Data	51
5.3.1	Eliminasi Kolom	52
5.3.2	Pembersihan Data	52
5.3.3	Normalisasi	53
5.3.4	Penghapusan Data Duplikat	54
5.4	Pelabelan Data	55
5.4.1	Pelabelan Berbasis Lexicon	55
5.4.2	Pelabelan Manual	56
5.5	Pembagian Data	57
5.6	Augmentasi Data	58

5.6.1	<i>Synonym Replacement</i>	58
5.6.2	<i>Back-Translation</i>	62
5.6.3	<i>Masked Language Modeling</i>	66
5.6.4	<i>Masked Language Modeling dengan Cosine Similarity</i>	74
5.7	Pemodelan Data	83
5.8	Evaluasi Model.....	87
BAB VI		90
HASIL PENELITIAN DAN PEMBAHASAN		90
6.1	Dataset.....	90
6.1.1	Dataset Kebijakan Tapera.....	90
6.1.2	Dataset Benchmark	91
6.2	Hasil Augmentasi Data	92
6.2.1	<i>Synonym Replacement</i>	94
6.2.2	<i>Back-Translation</i>	97
6.2.3	<i>Masked Language Modeling (MLM) IndoBERT</i>	98
6.2.4	<i>MLM IndoBERT + Cosine Similarity</i>	100
6.3	Hasil Pemodelan.....	103
6.4	Hasil Evaluasi Model	107
BAB VII		119
KESIMPULAN DAN SARAN		119
7.1	Kesimpulan	119
7.2	Saran.....	119
DAFTAR PUSTAKA		121
LAMPIRAN		125