

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN I	ii
HALAMAN PENGESAHAN II	iii
KATA PENGANTAR	iv
HALAMAN PERSEMBAHAN	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	6
BAB II TINJAUAN PUSTAKA	8
2.1 BPJS Kesehatan	8
2.2 <i>Natural Language Processing</i>	9
2.3 Pra-Proses	9
2.3.1 <i>Cleaning</i>	10
2.3.2 <i>Case Folding</i>	11
2.3.3 <i>Stemming</i>	12
2.3.4 Tokenisasi	12
2.4 <i>Systematic Random Sampling</i>	13
2.5 <i>Train-Test Split</i>	14
2.6 Klasifikasi Teks	15
2.7 <i>Transformer</i>	16
2.7.1 <i>Encoder</i>	17
2.7.2 <i>Decoder</i>	18

2.7.3	<i>Transformer Attention</i>	20
2.7.4	<i>Scaled Dot-Product Attention</i>	21
2.7.5	<i>Multi-Head Attention</i>	23
2.7.6	<i>Position-Wise Feed-Forward Network</i>	25
2.7.7	<i>Embeddings dan Softmax</i>	26
2.7.8	<i>Positional Encoding</i>	26
2.8	BERT	27
2.9	<i>Pre-Training BERT</i>	28
2.9.1	<i>Masked Language Modeling (MLM)</i>	29
2.9.2	<i>Next Sentence Prediction (NSP)</i>	30
2.10	<i>Fine-Tuning BERT</i>	31
2.10.1	Representasi <i>Input</i> BERT.....	32
2.10.2	<i>Hyperparameter Tuning</i>	33
2.10.3	<i>Fine-Tuning Model</i>	34
2.11	IndoBERT	36
2.12	Evaluasi Klasifikasi.....	36
2.13	<i>Cross Validation</i>	37
2.14	Pemodelan Topik	39
2.15	BERTopic.....	40
2.16	<i>Topic Coherence</i>	41
BAB III METODOLOGI PENELITIAN		43
3.1	Jenis dan Sumber Data	43
3.2	Variabel Penelitian	44
3.3	Tahapan Analisis Data	44
3.4	Diagram Alir Analisis Data.....	47
BAB IV HASIL DAN PEMBAHASAN		48
4.1	<i>Exploratory Data Analysis (EDA)</i>	48
4.2	Pra-Proses Data	49
4.3	Sampling dan Pelabelan Manual	50
4.4	Klasifikasi Teks.....	50
4.5	Evaluasi Model.....	50
4.6	Distribusi Tipe Komentar.....	52

4.7	Pemodelan Topik	53
4.7.1	Komentar Tipe Keluhan.....	53
4.7.2	Komentar Tipe Pertanyaan.....	55
BAB V	PENUTUP.....	58
5.1	Kesimpulan	58
5.2	Saran Pengembangan	59
DAFTAR	PUSTAKA	60
LAMPIRAN	64