

## DAFTAR ISI

HALAMAN PENGESAHAN .....	ii
PERNYATAAN BEBAS PLAGIASI .....	iii
PRAKATA .....	iv
DAFTAR ISI .....	vi
DAFTAR TABEL .....	x
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN .....	xiii
INTISARI .....	xiv
<i>ABSTRACT</i> .....	xv
BAB I .....	1
1.1 Latar Belakang Masalah .....	1
1.2 Batasan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Tinjauan Pustaka .....	5
1.6 Metodologi Penelitian .....	7
1.7 Sistematika Penulisan .....	7
BAB II .....	9
2.1 Variabel Random .....	9
2.2 Probabilitas Bersyarat .....	10
2.3 Distribusi Probabilitas .....	11
2.3.1 Distribusi Gamma .....	11
2.3.2 Distribusi Beta .....	12
2.3.3 Distribusi Binomial .....	13
2.3.4 Distribusi Multinomial .....	14
2.3.5 Distribusi Dirichlet .....	14
2.4 Metode Bayesian .....	15
2.5 <i>Machine Learning</i> .....	18
2.6 Deep Learning .....	20

2.7 <i>Hyperparameter Tuning</i> .....	21
2.8 Pemrosesan Bahasa Alami ( <i>Natural Language Processing</i> ).....	22
2.8.1 <i>Text Preprocessing</i> .....	24
2.8.2 <i>Word Clouds</i> .....	25
2.9 Proses Pengumpulan Data .....	26
<b>BAB III</b> .....	28
3.1 Pemodelan Topik .....	28
3.2 <i>Feature Extraction</i> .....	29
3.2.1 Tokenisasi .....	29
3.2.2 Representasi teks .....	30
3.3 Latent Dirichlet Allocation .....	34
3.3.1 Pembangkitan Model LDA .....	35
3.3.2 Model <i>Latent Dirichlet Allocation (LDA)</i> .....	37
3.3.3 <i>Variational Bayes</i> .....	40
3.3.4 Estimasi Parameter .....	45
3.4 <i>Bidirectional Encoder Representation from Transformer (BERT)</i> .....	53
3.4.1 <i>Transformer</i> .....	53
3.4.2 Model Arsitektur BERT.....	57
3.4.3 Representasi Input .....	57
3.4.4 <i>Pre-training</i> .....	58
3.4.5 <i>Fine-tuning</i> .....	60
3.5 <i>Sentence-Bidirectional Encoder Representation from Transformer (SBERT)</i> .....	60
3.5.1 <i>Document Embeddings</i> .....	61
3.5.2 <i>Document Clustering</i> .....	64
3.5.3 <i>Topic Representation Extraction</i> .....	73
3.6 Evaluasi Model .....	82
3.6.1 <i>Coherence Measure</i> .....	82
3.7 Alur Penelitian .....	88
<b>BAB IV</b> .....	91
4.1 Deskripsi Data.....	91
4.2 Proses Pengambilan Data .....	91

4.2.1	<i>Web Crawling</i> .....	92
4.2.2	<i>Web Scraping</i> .....	93
4.3	<i>Text Pre-processing</i> .....	94
4.4	Analisis Data Eksploratif.....	99
4.4.1	Jumlah Berita Palsu Tiap Bulan .....	100
4.4.2	<i>Word clouds</i> .....	101
4.4.3	Frekuensi Tiap Kata .....	102
4.5	Representasi Teks.....	103
4.5.1	Bag of Words (BoW).....	104
4.5.2	Term Frequency-Inverse Document Frequency (TF-IDF) .....	105
4.6	Pemodelan Topik .....	106
4.6.1	Latent Dirichlet Allocation (LDA).....	106
4.6.2	<i>Sentence-Bidirectional Encoder Representation from Transformer (SBERT)</i> .....	111
4.7	Evaluasi Model .....	115
4.8	Interpretasi Topik .....	118
4.8.1	Topik 1: Konten Video Palsu Mengenai Calon Presiden Indonesia	120
4.8.2	Topik 2: Akun sosial media palsu yang mengatasnamakan pemerintah .....	122
4.8.3	Topik 3: Manipulasi fakta dan foto palsu yang berkaitan dengan akun tertentu.....	124
4.8.4	Topik 4: Piala dunia sepak bola .....	126
4.8.5	Topik 5: Facebook banyak memberikan artikel <i>hoax</i> mengenai pemerintahan Indonesia.....	128
4.8.6	Topik 6: Penyakit dan penyebabnya.....	131
4.8.7	Topik 7: Berita <i>hoax</i> banyak berisi video dan gambar berupa potongan – potongan .....	133
4.8.8	Topik 8: Vaksin untuk Covid-19.....	135
4.9	Eksplorasi Topik .....	137
4.9.1	Pengelompokan Topik.....	137
4.9.2	Jumlah Berita <i>Hoax</i> Pada Setiap Topik dan Kelompok .....	140
BAB V	.....	142
5.1	Kesimpulan.....	142



5.2 Saran.....	143
DAFTAR PUSTAKA.....	144
LAMPIRAN .....	150