

DAFTAR ISI

HALAMAN PERSETUJUAN.....	ii
PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PRAKATA	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Metodologi Penelitian	6
1.7 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	8
BAB III LANDASAN TEORI.....	16
3.1. <i>Natural Language Processing</i>	16
3.1.1 <i>Tokenization</i>	17
3.1.2 <i>Stopwords Removal</i>	17
3.1.3 <i>Stemming</i>	18
3.2 <i>Gibbs Sampling Dirichlet Multinomial Mixture</i>	18
3.2.1 <i>Movie Group Process</i>	19

3.2.2	<i>Gibbs Sampling</i>	19
3.2.3	<i>Dirichlet Multinomial Mixture</i>	19
3.2.4	<i>UMass Coherence Score</i>	21
3.3	<i>Word Embedding</i>	21
3.3.1	<i>FastText</i>	21
3.2.2	<i>Bidirectional Encoder Representations from Transformers</i>	22
3.4.	<i>Temporal Convolutional Network</i>	24
3.4.1	<i>Dilated Causal Convolution</i>	25
3.4.2	<i>Residual Block</i>	26
3.4.3	<i>Epoch</i>	27
3.4.4	<i>Learning Rate</i>	27
3.4.5	<i>Dropout</i>	28
3.4.6	<i>Batch Size</i>	28
3.4.7	<i>Tanh Activation Function</i>	28
3.4.8	<i>ReLU Activation Function</i>	28
3.4.9	<i>Softmax Activation Function</i>	29
3.4.10	<i>Categorical Crossentropy Loss Function</i>	29
3.5	<i>Confusion Matrix</i>	29
3.5.1	<i>Precision</i>	30
3.5.2	<i>Recall</i>	31
3.5.3	<i>F1-score</i>	31
3.5.4	<i>Accuracy</i>	31
BAB IV ANALISIS DAN PERANCANGAN		32
4.1	Deskripsi Penelitian	32
4.2	Tahapan Penelitian	32

4.2.1	Studi Literatur	33
4.2.2	<i>Data Collecting</i>	34
4.2.3	<i>Data Preprocessing</i>	34
4.2.4	<i>Topic Modelling</i>	36
4.2.5	<i>Data Labelling</i>	37
4.2.6	Pengimplementasian <i>Word Embedding</i>	39
4.2.7	Pengembangan Model & <i>Training</i>	40
4.2.8	Implementasi	42
4.2.9	Evaluasi Model.....	43
BAB V IMPLEMENTASI.....		44
5.1	Preparasi <i>Environment</i> dan <i>Dataset</i>	44
5.1.1	<i>Import Dependencies</i>	44
5.1.2	Pemuatan <i>Dataset</i>	45
5.2	<i>Data Preprocessing</i>	46
5.2.1	<i>Tokenization</i>	46
5.2.1	<i>Text Cleaning</i>	46
5.2.3	<i>Stopwords Removal</i>	47
5.2.4	<i>Normalization</i>	47
5.2.5	<i>Stemming</i>	48
5.3	<i>Topic Modelling</i>	48
5.3.1	Pengembangan Model <i>Gibbs Sampling Dirichlet Multinomial Mixture</i> 48	
5.3.2	<i>Hyperparameter Tuning Gibbs Sampling Dirichlet Multinomial Mixture</i> 50	
5.3.3	Implementasi Model <i>Gibbs Sampling Dirichlet Multinomial Mixture</i> 50	

5.4	<i>Data Labelling</i>	51
5.4.1	Analisis Label pada Setiap <i>Cluster</i>	51
5.4.2	Verifikasi Label pada Setiap Pesan.....	52
5.4.3	Representasi <i>One-Hot-Encoding</i> Label.....	53
5.4.4	Proses <i>Stratified Train-Test Split</i> pada <i>Dataset</i>	54
5.5	Implementasi <i>Word Embedding</i>	54
5.5.1	Representasi Kata dengan <i>FastText</i>	54
5.5.2	Representasi Kata dengan <i>IndoBERT</i>	55
5.6	Pengembangan Model & <i>Training</i>	55
5.6.1	Pengembangan Model <i>Temporal Convolutional Network – FastText</i> 55	
5.6.2	<i>Training Temporal Convolutional Network – FastText Model</i>	57
5.6.3	Pengembangan Model <i>Temporal Convolutional Network – IndoBERT</i>	58
5.6.4	<i>Training Temporal Convolutional Network – IndoBERT Model.</i>	59
5.7	Evaluasi Model.....	61
5.7.1	<i>Confusion Matrix</i>	61
BAB VI HASIL DAN PEMBAHASAN		64
6.1	Hasil <i>Data Preprocessing</i>	64
6.2	Hasil <i>Topic Modelling</i>	64
6.2.1	Hasil <i>Hyperparameter Tuning Gibbs Sampling Dirichlet Multinomial Mixture</i>	64
6.3	Hasil <i>Data Labelling</i>	67
6.3.1	Hasil Analisis Label	67
6.3.2	Hasil Verifikasi Label	69
6.4	Hasil Pengembangan Model <i>Temporal Convolutional Network</i>	71

6.4.1	Hasil Arsitektur Model.....	71
6.4.2	Hasil <i>Training</i> dan <i>Hyperparameter Tuning</i>	74
6.4.3	Hasil Pengujian Model.....	78
BAB VII KESIMPULAN DAN SARAN		84
7.1	Kesimpulan.....	84
7.2	Saran	85
DAFTAR PUSTAKA		86
LAMPIRAN		89
Lampiran 1. Hasil <i>Hyperparameter Tuning GSDMM</i>		89
Lampiran 2. Hasil <i>Hyperparameter Tuning Temporal Convolutional Network – FastText</i>		107
Lampiran 3. Hasil <i>Hyperparameter Tuning Temporal Convolutional Network – IndoBERT</i>		112