

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN PEMBIMBING	ii
HALAMAN PENGSAHAN PENGUJI	iii
SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR	iv
PERNYATAN BEBAS PLAGIASI.....	v
ABSTRAK.....	vi
ABSTRACT.....	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR KODE PROGRAM.....	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Tahapan penelitian.....	3
1.6.1 Rencana dan Tahapan Penelitian.....	3
1.6.2 Metode Pengembangan Sistem.....	4
1.6.3 Pengujian Sistem	5
1.6.4 Pengujian Penelitian	5
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Berita	7
2.2 <i>Web Scraping</i>	7
2.3 <i>Preprocessing</i>	8
2.3.1 <i>Case Folding</i>	8
2.3.2 <i>Remove Punctuation</i>	8
2.3.3 <i>Remove Number</i>	8
2.3.4 <i>Stopword Removal</i>	9
2.3.5 <i>Remove Short Word</i>	9
2.3.6 <i>Add Start and End Token</i>	10
2.4 <i>Word Embedding</i>	10
2.5 <i>Automatic Text Summarization</i>	12
2.5.1 <i>Extractive Text Summarization</i>	12
2.5.2 <i>Abstractive Text Summarization</i>	13
2.6 <i>Sequence to Sequence</i>	13

2.7	<i>Encoder Decoder</i>	13
2.8	<i>Deep Learning</i>	14
2.8.1	<i>Longs Short Term Memory (LSTM)</i>	15
2.8.2	<i>Gated Recurrent Unit (GRU)</i>	20
2.9	<i>Bidirectional Long Short Term Memory (BiLSTM)</i>	25
2.10	<i>Bidirectional Gated Recurrent Unit (BiGRU)</i>	27
2.11	<i>Attention</i>	27
2.12	<i>RMSProp</i>	29
2.13	<i>Early Stopping</i>	29
2.14	<i>ROUGE</i>	30
2.15	<i>State of The Art</i>	30
BAB II METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM.....		36
3.1	<i>Meetodologi Penelitian</i>	36
3.1.1	<i>Pengumpulan Data</i>	37
3.1.2	<i>Preprocessing</i>	38
3.1.3	<i>Pembuatan Model</i>	49
3.1.4	<i>Texts to Sequences</i>	50
3.1.5	<i>Padding to Sequences</i>	52
3.1.6	<i>Word Embedding</i>	53
3.1.7	<i>Bidirectional Long Short Term Memory (BiLSTM)</i>	54
3.1.8	<i>Bidirectional Gated Recurrent Unit (BiGRU)</i>	60
3.1.9	<i>Pengujian</i>	65
3.2	<i>Metodologi Pengembangan Sistem</i>	68
3.2.1	<i>Analisis Kebutuhan Sistem</i>	69
3.2.2	<i>Kebutuhan Fungsional</i>	69
3.2.3	<i>Kebutuhan Non-fungsional</i>	70
3.2.4	<i>Proses Desain</i>	70
3.2.5	<i>Perancangan Sistem</i>	70
3.2.6	<i>Pengkodean</i>	74
3.2.7	<i>Perancangan Pengujian</i>	74
BAB IV HASIL DAN PEMBAHASAN.....		75
4.1	<i>Hasil Penelitian</i>	75
4.1.1	<i>Pengumpulan Data</i>	75
4.1.2	<i>Implementasi Preprocessing</i>	78
4.1.3	<i>Pembuatan Model</i>	82
4.1.4	<i>Implementasi Sistem</i>	95
4.2	<i>Pengujian</i>	101
4.2.1	<i>Pengujian Model</i>	101
4.2.2	<i>Pengujian Sistem</i>	103
4.3	<i>Pembahasan</i>	103
BAB V KESIMPULAN DAN SARAN		105
5.1	<i>Kesimpulan</i>	105
5.2	<i>Saran</i>	105

DAFTAR PUSTAKA	107
LAMPIRAN	112

DAFTAR TABEL

Tabel 2.1 Contoh <i>Case Folding</i>	8
Tabel 2.2 Contoh <i>Remove Punctuation</i>	8
Tabel 2.3 Contoh <i>Remove Number</i>	9
Tabel 2.4 Contoh <i>Stopword Removal</i>	9
Tabel 2.5 Contoh <i>Remove Short Word</i>	10
Tabel 2.6 Contoh <i>Add Start and End Token</i>	10
Tabel 2.7 <i>State of The Art</i>	31
Tabel 2.8 <i>State of The Art</i> (lanjutan).....	32
Tabel 2.9 <i>State of The Art</i> (lanjutan).....	33
Tabel 2.10 <i>State of The Art</i> (lanjutan).....	34
Tabel 3.1 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Artikel Berita	45
Tabel 3.2 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Artikel Berita (lanjutan).....	46
Tabel 3.3 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Artikel Berita (lanjutan).....	47
Tabel 3.4 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Artikel Berita (lanjutan).....	48
Tabel 3.5 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Artikel Berita (lanjutan).....	49
Tabel 3.6 Contoh Teks Pada Tahap <i>Preprocessing</i> Data Ringkasan (<i>Summary</i>).....	49
Tabel 3.7 Hasil <i>Texts to Sequences</i>	50
Tabel 3.8 Hasil <i>Texts to Sequences</i> (lanjutan)	51
Tabel 3.9 Hasil <i>Padding to Sequences</i>	53
Tabel 3.10 Hasil <i>Preprocessing</i>	54
Tabel 3.11 Contoh Hasil FastText	54
Tabel 3.12 Hasil <i>Forget Gate</i> (f_t) tiap orde	60
Tabel 3.13 Hasil <i>Input Gate</i> (i_t) tiap orde	60
Tabel 3.14 Hasil Kandidat Baru (\check{C}_t) tiap orde	60
Tabel 3.15 Hasil <i>Cell State</i> (C_t) tiap orde.....	60
Tabel 3.16 Hasil <i>Output Gate</i> (O_t) tiap orde	60
Tabel 3.17 Hasil <i>Output Final</i> (h_t) tiap orde	60
Tabel 3.18 Hasil <i>Update Gate</i> (z_t)tiap orde	64
Tabel 3.19 Hasil <i>Reset Gate</i> (r_t) tiap orde.....	64
Tabel 3.20 Hasil <i>Reset Gate</i> (r_t) tiap orde (lanjutan)	65
Tabel 3.21 Hasil <i>Current Memory</i> (h'_t) tiap orde	65
Tabel 3.22 Hasil <i>Final Content Memory</i> (h_t) tiap orde.....	65
Tabel 3.23 Rancangan Perbandingan ROUGE-1 dan ROUGE-2.....	66
Tabel 3.24 Nilai Maksimal ROUGE-1 dan ROUGE-1	67
Tabel 3.25 Kebutuhan Perangkat Keras (<i>Hardware</i>)	70
Tabel 3.26 Kebutuhan Perangkat Lunak (<i>Software</i>).....	70
Tabel 3.27 Detail Perancangan Pengujian Sistem	74
Tabel 4.1 Kata-kata dalam Kamus <i>Stopword</i>	80
Tabel 4.2 Kata-kata dalam Kamus <i>Stopword</i> (lanjutan).....	81

Tabel 4.3 Hasil ROUGE	101
Tabel 4.4 Nilai ROUGE Tertinggi.....	103
Tabel 4.5 Pengujian Sistem	103

DAFTAR GAMBAR

Gambar 2.1 Arsitektur FastText.....	11
Gambar 2.2 <i>Extractive Text Summarization</i>	12
Gambar 2.3 <i>Abstractive Text Summarization</i>	13
Gambar 2.4 <i>Encoder Decoder</i>	14
Gambar 2.5 Jaringan LSTM	15
Gambar 2.6 Struktur LSTM.....	16
Gambar 2.7 <i>Memory Cell</i>	16
Gambar 2.8 Lapisan Sigmoid	17
Gambar 2.9 Alur Informasi pada <i>Forget Gate</i>	17
Gambar 2.10 Alur Informasi yang Melewati <i>Input Gate</i>	18
Gambar 2.11 Memperbaharui <i>Cell State</i>	19
Gambar 2.12 Alur Informasi Melewati <i>Output Gate</i>	20
Gambar 2.13 Arsitektur GRU	21
Gambar 2.14 Alur Arsitektur GRU.....	21
Gambar 2.15 <i>Update Gate</i>	22
Gambar 2.16 <i>Reset Gate</i>	23
Gambar 2.17 <i>Current Content Memory</i>	24
Gambar 2.18 <i>Final Memory</i>	25
Gambar 2.19 Arsitektur BiLSTM	26
Gambar 2.20 Koneksi Internal Lapisan BiLSTM.....	26
Gambar 2.21 Arsitektur BiGRU	27
Gambar 2.22 Arsitektur <i>Attention Mechanism</i>	28
Gambar 2.23 <i>Early Stopping Based on Cross-Validation</i>	29
Gambar 3.1 Metodologi Penelitian	36
Gambar 3.2 Hasil <i>Web Scraping Train Data</i>	37
Gambar 3.3 Hasil <i>Web Scraping Validate Data</i>	37
Gambar 3.4 Hasil <i>Web Scraping Test Data</i>	38
Gambar 3.5 <i>Flowchart Preprocessing</i>	38
Gambar 3.6 <i>Flowchart Case Folding</i>	39
Gambar 3.7 <i>Flowchart Remove Punctuation</i>	40
Gambar 3.8 <i>Flowchart Remove Number</i>	41
Gambar 3.9 <i>Flowchart Stopword Removal</i>	42
Gambar 3.10 <i>Flowchart Remove Short Word</i>	43
Gambar 3.11 <i>Flowchart Add Start and End Token</i>	44
Gambar 3.12 <i>Flowchart Texts to Sequences</i>	51
Gambar 3.13 <i>Flowchart Padding to Sequences</i>	52
Gambar 3.14 <i>Flowchart FastText</i>	53
Gambar 3.15 <i>Flowchart BiLSTM</i>	55
Gambar 3.16 <i>Flowchart BiGRU</i>	61
Gambar 3.17 <i>Flowchart ROUGE-N</i>	66

Gambar 3.18 Metodologi Pengembangan Sistem.....	69
Gambar 3.19 Arsitektur Sistem.....	71
Gambar 3.20 <i>Flowchart</i> Sistem	72
Gambar 3.21 Halaman <i>Home</i> Sistem Peringkasan Teks Abstraktif	73
Gambar 3.22 Halaman Hasil Sistem Peringkasan Teks Abstraktif	74
Gambar 4.1 File Hasil <i>Web Scraping</i>	76
Gambar 4.2 <i>Train Data</i>	77
Gambar 4.3 <i>Validate Data</i>	77
Gambar 4.4 <i>Test Data</i>	77
Gambar 4.5 Data Jika Ditampilkan.....	78
Gambar 4.6 <i>Tokenizer</i> Teks Berita	83
Gambar 4.7 <i>Tokenizer</i> Ringkasan	83
Gambar 4.8 Detail Arsitektur Model BiLSTM.....	88
Gambar 4.9 Detail Arsitektur Model BiGRU	93
Gambar 4.10 Halaman <i>Home</i>	96
Gambar 4.11 Halaman Hasil.....	96
Gambar 4.12 Hasil <i>Training</i> Model BiGRU 4.....	102
Gambar 4.13 Grafik Accuracy Model BiGRU 4	102
Gambar 4.14 Grafik Loss Model BiGRU 4	102

DAFTAR KODE PROGRAM

Kode Program 1 : <i>Web Scraping</i>	75
Kode Program 2 : <i>Web Scraping (lanjutan)</i>	76
Kode Program 3 : <i>Combine Data</i>	78
Kode Program 4 : <i>Load Dataset</i>	78
Kode Program 5 : <i>Preprocessing</i>	79
Kode Program 6 : <i>Case Folding</i>	79
Kode Program 7 : <i>Remove Punctuation</i>	79
Kode Program 8 : <i>Remove Number</i>	80
Kode Program 9 : <i>Stopword Removal</i>	81
Kode Program 10 : <i>Remove Short Word</i>	81
Kode Program 11 : <i>Add Start and End Token</i>	82
Kode Program 12 : <i>Inisialisasi Panjang Maksimal</i>	82
Kode Program 13 : <i>Split Data</i>	82
Kode Program 14 : <i>Tokenizer</i>	83
Kode Program 15 : <i>Texts to Sequences</i>	84
Kode Program 16 : <i>Padding to Sequences</i>	84
Kode Program 17 : <i>Load Pretrained Model FastText</i>	84
Kode Program 18 : <i>Get Embedding Matrix</i>	85
Kode Program 19 : <i>Model BiLSTM</i>	86
Kode Program 20 : <i>Model BiLSTM (lanjutan)</i>	87
Kode Program 21 : <i>Compile Model BiLSTM</i>	88
Kode Program 22 : <i>Training Model BiLSTM</i>	88
Kode Program 23 : <i>Training Model BiLSTM (lanjutan)</i>	89
Kode Program 24 : <i>Encoder dan Decoder Model BiSLTM</i>	89
Kode Program 25 : <i>Menyimpan Model dan Tokenizer BiLSTM</i>	89
Kode Program 26 : <i>Menyimpan Model dan Tokenizer BiLSTM (lanjutan)</i>	90
Kode Program 27 : <i>Vocab Size</i>	90
Kode Program 28 : <i>Model BiGRU</i>	90
Kode Program 29 : <i>Model BiGRU (lanjutan)</i>	91
Kode Program 30 : <i>Attention Layer</i>	91
Kode Program 31 : <i>Attention Layer (lanjutan)</i>	92
Kode Program 32 : <i>Attention Layer (lanjutan)</i>	93
Kode Program 33 : <i>Compile Model BiGRU</i>	94
Kode Program 34 : <i>Training Model BiGRU</i>	94
Kode Program 35 : <i>Endocer dan Decoder Model BiGRU</i>	94
Kode Program 36 : <i>Menyimpan Model BiGRU</i>	94
Kode Program 37 : <i>Menyimpan Model BiGRU (lanjutan)</i>	95
Kode Program 38 : <i>Import Library Implementasi Sistem</i>	95
Kode Program 39 : <i>Tampilan Halaman Home</i>	95
Kode Program 40 : <i>Fungsi Untuk Meringkas dan Menghitung Akurasi</i>	97

Kode Program 41 : Fungsi Untuk Meringkas dan Menghitung Akurasi (lanjutan).....	98
Kode Program 42 : Fungsi Untuk Meringkas dan Menghitung Akurasi (lanjutan).....	99
Kode Program 44 : Tampilan Halaman Hasil	99
Kode Program 45 : Tampilan Halaman Hasil (lanjutan).....	100
Kode Program 46 : Tampilan Halaman Hasil (lanjutan).....	101
Kode Program 47 : ROUGE.....	101

DAFTAR LAMPIRAN

Lampiran A.1 < Contoh Hasil Ringkasan BiLSTM >	112
Lampiran A.2 < Contoh Hasil Ringkasan BiGRU >	113