

**PERINGKASAN TEKS BERITA
DALAM BAHASA INDONESIA SECARA ABSTRAKTIF
MENGUNAKAN METODE ATTENTION-BASED
BILSTM (*BIDIRECTIONAL LONG SHORT TERM MEMORY*)**

Digunakan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Srata-1
Pada Jurusan Teknik Informatika



Oleh :

Hanny Putri Gayatri
09021382025143

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2023**

LEMBAR PENGESAHAN SKRIPSI

**PERINGKASAN TEKS BERITA
DALAM BAHASA INDONESIA SECARA ABSTRAKTIF
MENGUNAKAN METODE ATTENTION-BASED
BILSTM (*BIDIRECTIONAL LONG SHORT TERM MEMORY*)**

Oleh :

Hanny Putri Gayatri
09021382025143

Palembang, 16 Juli 2024

Pembimbing I



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

Pembimbing II



M. Qurhanul Rizqie, M.T., Ph.D.
NIP 198712032022031006

Mengetahui,

Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Senin tanggal 15 Juli 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Hanny Putri Gayatri
NIM : 09021382025143
Judul : Peringkasan Teks Berita Dalam Bahasa Indonesia Secara Abstraktif
Menggunakan Metode Attention-Based BiLSTM (*Bidirectional Long Short Term Memory*)

dan dinyatakan LULUS.

1. Ketua Penguji

Kanda Januar Miraswan, S.Kom., M.T.
NIP 199001092019031012



2. Penguji I

Novi Yushiani, S.Kom., M.T.
NIP 198211082012122001



3. Pembimbing I

Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003



4. Pembimbing II

M. Ourhanul Rizqie, M.T., Ph.D.
NIP 198712032022031006



Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002



HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Hanny Putri Gayatri
NIM : 09021382025143
Program Studi : Teknik Informatika
Judul : Peringkasan Teks Berita Dalam Bahasa Indonesia Secara
Abstraktif Menggunakan Metode Attention-Based
BiLSTM (*Bidirectional Long Short Term Memory*)


Hasil Pengecekan Software iThenticate/Turnitin : 9%

Menyatakan bahwa Laporan Proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya akan bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 17 Juli 2024



Hanny Putri Gayatri
NIM 09021382025143

MOTTO DAN PERSEMBAHAN

Motto :

“Allah tidak membebani seseorang
melainkan sesuai dengan
kesanggupannya...”

[2:286]

Kupersembahkan karya tulis ini

kepada :

- Keluarga
- Dosen Pembimbing
- Teman Seperjuangan
- Fakultas Ilmu Komputer
Universitas Sriwijaya

ABSTRACT

News is a form of information published to the public, allowing for a better understanding of the world around them. The advancement of internet technology has significantly increased the growth of Indonesian-language news sites and created a surge in information availability. Reading the entire news article takes a lot of time and makes it difficult to receive all the information quickly and accurately, leading to the concept of automatic text summarization. The BiLSTM (Bidirectional Long Short-Term Memory) algorithm is an artificial neural network architecture commonly used in text summarization tasks, especially with the addition of an attention layer that helps focus on important parts of the text. Abstractive summarization, which works by understanding and generating new sentences from the original document, is used in this study. Evaluations show that the BiLSTM model achieved a ROUGE-1 score of 0.0594, while the Hybrid model achieved a ROUGE-1 score of 0.1316.

Keywords: *News, Technology, Automatic Summarization, Bidirectional Long Short-Term Memory, Attention, Abstractive, ROUGE*

Supervisor I



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

Palembang, 16 July 2024

Supervisor II



M. Qurhanul Rizqie, M.T., Ph.D.
NIP 198712032022031006

Approve,
Head of Informatics Engineering Department



Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002

ABSTRAK

Berita merupakan salah satu bentuk informasi yang dipublikasikan kepada masyarakat luas yang memberikan pemahaman yang lebih baik tentang dunia sekitarnya. Kemajuan teknologi internet meningkatkan pertumbuhan situs berita berbahasa Indonesia dan menciptakan lonjakan informasi. Berita yang dibaca secara keseluruhan memakan banyak waktu dan sulit untuk menerima semua informasi secara cepat dan akurat, sehingga muncul konsep peringkasan teks otomatis. Algoritma BiLSTM (*Bidirectional Long Short-Term Memory*) adalah salah satu arsitektur jaringan saraf tiruan yang biasanya digunakan dalam tugas peringkasan teks, terutama dengan penambahan layer *attention* yang membantu fokus pada bagian penting dari teks. Peringkasan secara abstraktif, bekerja dengan memahami dan menghasilkan kalimat baru dari dokumen asli, digunakan dalam penelitian ini. Evaluasi menunjukkan model BiLSTM mendapatkan skor ROUGE-1 0.0594, sementara model Hybrid mendapatkan skor ROUGE-1 0.1316.

Kata Kunci : Berita, teknologi, peringkasan otomatis, BiLSTM, attention, Abstraktif, ROUGE

Pembimbing I



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

Palembang, 16 Juli 2024

Pembimbing II



M. Qurhanul Rizqie, M.T., Ph.D.
NIP 198712032022031006

Mengetahui,
Ketua Jurusan Teknik Informatika




Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002

KATA PENGATAR

Dengan mengucapkan puji dan syukur kepada Allah SWT atas rahmat, nikmat dan karunia-Nya sehingga penyusun dapat menyelesaikan tugas akhir ini dengan judul “Peringkasan Teks Berita Dalam Bahasa Indonesia Secara Abstraktif Menggunakan Metode Attention-Based BiLSTM (Bidirectional Long Short Term Memory)” tepat pada waktunya. Tugas akhir ini disusun sebagai salah satu syarat untuk menyelesaikan studi pendidikan program Strata-1 pada Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam penyusunan tugas akhir ini, penyusun tidak terlepas dari hambatan dan tantangan. Namun, dengan adanya bimbingan, dukungan, bantuan serta doa yang baik dari berbagai pihak selama proses nya sehingga tugas akhir ini dapat diselesaikan dengan baik. Oleh karena itu, pada kesempatan ini penyusun ingin mengucapkan terima kasih secara khusus kepada :

1. Kedua orang tua ku yang tersayang, saudara-saudara ku serta keluarga besar yang terus memberikan semangat dan dukungan dalam segi moril maupun materi kepada penyusun.
2. Abi dan mami yang selalu memberikan semangat dan dukungan dalam segi moril maupun materi kepada penyusun.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. M. Fachrurrozi, S.Si., M.T. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Alvi Syahrini Utami, M.Kom, selaku Dosen Pembimbing I yang telah banyak memberikan bantuan, arahan, masukan dan saran serta ilmu dalam penyelesaian tugas akhir ini.
6. Bapak M. Qurhanul Rizqie, M.T., Ph.D, selaku Dosen Pembimbing II yang juga telah banyak memberikan bantuan, arahan, masukan dan saran serta ilmu dalam penyelesaian tugas akhir ini.

7. Bapak Dr. Abdiansah, S.Kom., M.Cs., selaku Dosen Pembimbing Akademik yang selama ini telah banyak membantu dalam kegiatan pembelajaran selama perkuliahan di Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Seluruh dosen Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya yang telah memberikan ilmu pengetahuan, masukan dan saran, motivasi selama masa pembelajaran di perkuliahan.
9. Mbak Wiwin Juliani dan Mbak Rika Muliani selaku Admin Jurusan Teknik Informatika Bilingual yang telah banyak membantu dalam kelancaran administrasi selama masa perkuliahan.
10. Teman-teman TI Bilingual B angkatan 2020 yang telah menemani penyusun selama masa perkuliahan yang tidak dapat disebutkan satu per satu.
11. Semua teman-teman organisasi baik di HMIF 2021 maupun DPM KM Fasilkom 2022 yang telah mau berbagi ilmu dan pengalaman serta menerima penyusun sebagai *partner* selama satu tahun kepengurusan.
12. Teman-teman “Mamba”, Retno dan Febi yang telah berbagi keluh kesah, motivasi, semangat serta canda tawa selama masa perkuliahan.
13. Tim HireHub khususnya Sub Unsri yang telah berbagi keluh kesah, motivasi, semangat serta canda tawa selama penyusunan tugas akhir ini.
14. Teman-teman TBRangers dan team Divisi IT dan Database MSS yang telah memberikan motivasi dan semangat selama penyusunan tugas akhir ini.
15. Semua pihak yang telah membantu penyusun dalam menyelesaikan Tugas Akhir ini yang mungkin belum dituliskan satu per satu.

Penyusun menyadari dalam penulisan tugas akhir ini masih terdapat banyak kekurangan serta keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata, semoga penelitian dari tugas akhir ini dapat berguna dan bermanfaat bagi kita semua.

Palembang, 25 Mei 2024



Hanny Putri Gayatri

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRISPI.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
BAB I PENDAHULUAN.....	I-1
1.1. Pendahuluan.....	I-1
1.2. Latar Belakang.....	I-1
1.3. Rumusan Masalah.....	I-4
1.4. Tujuan Penelitian.....	1-5
1.5. Manfaat Penelitian.....	1-5
1.6. Batasan Masalah.....	I-6
1.7. Sistematika Penulisan.....	I-6
1.8. Kesimpulan.....	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1. Pendahuluan.....	II-1
2.2. Landasan Teori.....	II-1
2.2.1. Berita.....	II-1
2.2.2. <i>Natural Language Processing</i>	II-2
2.2.3. Peringkasan Teks Otomatis.....	II-4

2.2.4. <i>Data Preprocessing</i>	II-5
2.2.5. <i>Word2Vec</i>	II-6
2.2.6. <i>Sequence-to-Sequence</i>	II-7
2.2.7. <i>Bidirectional Long Short Term Memory (BiLSTM)</i>	II-8
2.2.8. <i>Attention Mechanism</i>	II-13
2.2.9. <i>Recall-Oriented Understudy for Gisting Evaluation (ROUGE)</i>	II-14
2.2.10. <i>Rational Unified Process (RUP)</i>	II-16
2.3. Penelitian Terkait.....	II-17
2.4. Kesimpulan.....	II-19
BAB III METODOLOGI PENELITIAN.....	III-1
3.1. Pendahuluan.....	III-1
3.2. Pengumpulan Data.....	III-1
3.2.1 Jenis dan Sumber Data.....	III-1
3.2.2 Metode Pengumpulan Data.....	III-1
3.3. Tahapan Penelitian.....	III-3
3.3.1. Identifikasi Masalah.....	III-3
3.3.2. Melakukan Studi Literatur.....	III-4
3.3.3. Membuat Kerangka Kerja.....	III-4
3.3.4. Perancangan Arsitektur Model.....	III-8
3.3.5. Melakukan Pengujian dan Evaluasi.....	III-9
3.3.5.1. Kriteria Pengujian	III-9
3.3.5.2. Format Data Pengujian	III-9
3.3.5.3. Alat yang Digunakan dalam Penelitian	III-11
3.3.5.4. Evaluasi Hasil Pengujian	III-11
3.4. Metode Pengembangan Perangkat Lunak.....	III-12
3.4.1. Fase Insepsi.....	III-12
3.4.2. Fase Elaborasi.....	III-12
3.4.3. Fase Konstruksi.....	III-13
3.4.4. Fase Transisi.....	III-13
3.5. Manajemen Proyek Penelitian.....	III-14

3.6. Kesimpulan.....	III-17
BAB IV METODOLOGI PENELITIAN.....	IV-1
4.1. Pendahuluan.....	IV-1
4.2. <i>Rational Unified Process</i>	IV-1
4.2.1. Fase Insepsi.....	IV-1
4.2.1.1. Pemodelan Bisnis.....	IV-1
4.2.1.2. Kebutuhan Sistem.....	IV-2
4.2.1.3. Analisis dan Perancangan.....	IV-3
4.2.2. Fase Elaborasi.....	IV-6
4.2.2.1. Pemodelan Bisnis.....	IV-6
4.2.2.2. Perancangan Data.....	IV-6
4.2.2.3. Perancangan Antarmuka.....	IV-7
4.2.2.4. <i>Activity Diagram</i>	IV-8
4.2.2.5. <i>Sequence Diagram</i>	IV-9
4.2.3. Fase Konstruksi.....	IV-11
4.2.3.1. <i>Class Diagram</i>	IV-11
4.2.3.2. Implementasi Kelas.....	IV-11
4.2.3.3. Implementasi Antarmuka.....	IV-12
4.2.4. Fase Insepsi.....	IV-13
4.2.4.1. Pemodelan Bisnis.....	IV-13
4.2.4.2. Rencana Pengujian.....	IV-13
4.2.4.3. Pengujian.....	IV-14
4.3. Kesimpulan.....	IV-15
BAB V HASIL DAN ANALISIS PENELITIAN.....	V-1
5.1. Pendahuluan.....	V-1
5.2. Hasil Penelitian.....	V-1
5.2.1. Konfigurasi Percobaan.....	V-1
5.2.1. Data Hasil Konfigurasi.....	V-1
5.2.1. Analisis Hasil Penelitian.....	V-1

5.3. Kesimpulan.....	V-15
BAB VI KESIMPULAN DAN SARAN.....	VI-1
6.1. Kesimpulan.....	VI-1
6.2. Saran.....	VI-1
DAFTAR PUSTAKA.....	xvi
LAMPIRAN.....	xix

DAFTAR TABEL

	Halaman
Tabel III-1. Pengujian Hasil Pelatihan Model.....	III-10
Tabel III-2. Evaluasi Data Hasil Pengujian.....	III-11
Tabel III-3. Evaluasi Hasil Pengujian.....	III-12
Tabel III-4. <i>Work Breakdown Structure</i> (WBS).....	III-15
Tabel IV-1. Kebutuhan Fungsional.....	IV-2
Tabel IV-2. Kebutuhan Non Fungsional.....	IV-3
Tabel IV-3. Definisi <i>Use Case</i>	IV-4
Tabel IV-4. Skenario Use Case Memasukkan Teks.....	IV-4
Tabel IV-5. Skenario Use Case Melihat Informasi Sistem.....	IV-5
Tabel IV-6. Implementasi Kelas.....	IV-11
Tabel IV-7. Rencana Pengujian.....	IV-14
Tabel IV-8. Hasil Pengujian.....	IV-14
Tabel V-1. Konfigurasi <i>Hyperparamater</i>	V-2
Tabel V-2. Hasil Pengujian Pada Konfigurasi.....	V-3
Tabel V-3. Hasil Konfigurasi Pengujian Model Seq2Seq BiLSTM.....	V-4
Tabel V-4. Konfigurasi Hasil Pengujian Model Seq2seq Hybrid.....	V-6
Tabel V-5. Evaluasi Pada Model Seq2Seq BiLSTM.....	V-8
Tabel V-6. Evaluasi Pada Model Seq2Seq Hybrid.....	V-9
Tabel V-7. Hasil Evaluasi Nilai ROUGE.....	V-10

DAFTAR GAMBAR

	Halaman
Gambar II-1. <i>NLP Pipeline</i> (Chen, 2023).....	II-5
Gambar II-2. Arsitektur LSTM Cell (Singhal, 2020).....	II-8
Gambar II-3. Arsitektur <i>Bidirectional LSTM</i> (Yuliska, 2022).....	II-11
Gambar II-4. Arsitektur RUP (Ambler, 2010).....	II-16
Gambar III-1. Tahapan Penelitian.....	III-3
Gambar III-2. Kerangka Kerja Penelitian.....	III-4
Gambar IV-1. <i>Use Case Diagram</i>	IV-3
Gambar IV-2. Antarmuka Menu Home.....	IV-7
Gambar IV-3. Antarmuka Menu About.....	IV-7
Gambar IV-4. <i>Activity Diagram</i> Memasukkan Teks.....	IV-8
Gambar IV-5. <i>Activity Diagram</i> Melihat Informasi Sistem.....	IV-8
Gambar IV-6. <i>Sequence Diagram</i> Memasukkan Teks.....	IV-10
Gambar IV-7. <i>Sequence Diagram</i> Melihat Informasi Sistem.....	IV-10
Gambar IV-8. <i>Class Diagram</i>	IV-11
Gambar IV-9. Tampilan Menu Home.....	IV-12
Gambar IV-10. Tampilan Menu About.....	IV-13
Gambar V-1. Accuracy dan Loss Pada Model Seq2Seq BiLSTM.....	V-6
Gambar V-2. Accuracy dan Loss Pada Model Seq2Seq Hybrid.....	V-8

BAB I

PENDAHULUAN

1.1. Pendahuluan

Pada bab pendahuluan ini akan membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan.

1.2. Latar Belakang

Berita merupakan salah satu bentuk dari penyajian informasi yang dipublikasikan kepada masyarakat luas. Melalui berita, masyarakat dapat memperoleh pemahaman yang lebih baik tentang dunia sekitarnya. Hal ini didukung dengan adanya perkembangan teknologi yang memudahkan akses manusia untuk mendapatkan informasi. Internet sebagai salah satu teknologi yang berperan dalam penyebaran informasi, telah menjadi salah satu media yang banyak digunakan untuk membaca berita oleh pengguna internet di Indonesia. Sebab lebih dari 50% pengguna internet Indonesia memanfaatkan media ini untuk membaca berita (Andriani, *et al.*, 2019).

Kemajuan teknologi internet secara nyata berpengaruh terhadap pertumbuhan situs berita berbahasa Indonesia dan menciptakan lonjakan besar dalam ketersediaan informasi. Membaca keseluruhan isi berita akan memakan banyak waktu dan sulit untuk menerima semua informasi secara cepat dan akurat. Hal ini yang akhirnya memunculkan sebuah ide konsep peringkasan secara otomatis, dimana sebuah sistem akan mengekstrak informasi penting dari sebuah

teks. Salah satu bidang yang mampu dalam mengatasi permasalahan ini ialah *Automatic Text Summarization* (Peringkasan Teks Otomatis) (Mandar dan Gunawan, 2018).

Menurut Elsaid, *et al.*, (2022) peringkasan teks otomatis adalah suatu proses yang bertujuan secara otomatis menghasilkan ringkasan atau cuplikan teks yang singkat dari teks sumber yang lebih panjang. Dengan memanfaatkan teknologi *Natural Language Processing* sebuah sistem peringkasan dapat menghasilkan ringkasan yang lebih mudah dipahami, menghemat waktu pembaca namun tetap mengandung informasi yang sama. Dalam peringkasan teks sendiri secara umum terdapat dua pendekatan, yaitu ekstraktif dan abstraktif. Peringkasan ekstraktif adalah salah satu teknik peringkasan dengan memilih kalimat penting dari dokumen asli, dengan cara mengekstrak kalimat berdasarkan fitur statistik dan linguistik. Sedangkan peringkasan abstraktif bekerja dengan cara memahami dan menghasilkan kalimat baru dari dokumen asli sebagai hasil ringkasan, oleh karena itu pendekatan abstraktif terbilang lebih kompleks sebab serupa dengan ringkasan yang dilakukan oleh manusia (Badry, *et al.*, 2013). Namun didalam beberapa penelitian disebutkan satu pendekatan lainnya yaitu pendekatan *hybrid*, dimana pendekatan ini merupakan gabungan dari pendekatan ekstraktif dan abstraktif (Liu, *et al.*, 2021).

Algoritma BiLSTM (*Bidirectional Long Short-Term Memory*) merupakan salah satu jenis arsitektur jaringan saraf yang merupakan pengembangan dari model arsitektur LSTM. BiLSTM memiliki dua lapisan yang beroperasi dalam arah berlawanan dari kiri ke kanan dan sebaliknya, memungkinkan model menjadi

sangat baik dalam mengenali pola kalimat karena setiap kata diproses secara sekuensial (Fadli dan Hidayatullah, 2021). Penambahan layer *attention* (perhatian) memiliki peran penting dalam meningkatkan kinerja BiLSTM yang digunakan dalam tugas peringkasan teks, yaitu memungkinkan model untuk fokus pada bagian-bagian penting dari teks sumber selama proses peringkasan. Hal ini meningkatkan kemampuan model dalam memfokuskan perhatian pada informasi yang paling relevan dalam suatu teks (Jiang, *et al.*, 2021).

Penelitian di bidang peringkasan teks sudah dimulai sejak abad pertengahan ke-20, yang pertama kali dipelopori oleh Luhn (1958) (Mustaqhfiri, *et al.*, 2011). Beberapa penelitian lain juga telah dilakukan dalam kurun beberapa tahun ke belakang, seperti penelitian yang dilakukan oleh Wazery, *et al.* (2022), yang berjudul *Abstractive Arabic Text Summarization Based on Deep Learning*. Pada penelitian ini peringkasan teks dilakukan dengan masukan berupa dua jenis dataset berita *single document* menggunakan bahasa Arab. Algoritma *Deep Learning* seperti LSTM, GRU dan BiLSTM juga digunakan untuk memproses ringkasan teks dengan penambahan layer *attention*, kemudian hasil dari ketiga model ini akan dibandingkan pada tahap evaluasi. Pada tahap evaluasi model, peneliti menggunakan metrik evaluasi ROUGE dan BLEU. Pada dataset AHS nilai *F1-score* tertinggi yang didapatkan ROUGE-1 pada model GRU, LSTM dan BiLSTM adalah 46.44, 50.93, 51.49 sedangkan pada BLEU skor tertinggi didapatkan oleh masing-masing model adalah 0.34, 0.39 dan 0.41. Pada dataset AMN nilai *F1-score* tertinggi yang didapatkan ROUGE-1 pada model GRU, LSTM dan BiLSTM adalah 33.13, 38.2 dan 44.28 sedangkan pada BLEU skor

tertinggi masing-masing model adalah 0.35, 0.36 dan 0.41. Hasil keseluruhan menunjukkan bahwa model BiLSTM selalu mencapai skor tertinggi dan kinerja terbaik dibandingkan kedua metode lainnya.

Dalam penelitian lain yang dilakukan oleh Yuliska dan Syaliman (2022), melakukan peringkasan menggunakan metode BiLSTM dengan pendekatan ekstraktif yang menghasilkan ringkasan relevan dari sekumpulan dokumen berdasarkan sebuah *input* kueri. Dataset yang digunakan adalah *DUC 2005-2007* yang berisi dokumen teks berbahasa Inggris. Pada tahap evaluasi menggunakan ROUGE, peneliti melakukan sebuah eksperimen generasi *summary* menggunakan model CNN+LSTM sebagai model *baseline*, yang kemudian akan dibandingkan hasil evaluasi dengan model BiLSTM yang telah dibangun sebelumnya. Hasilnya menunjukkan nilai ROUGE-1, ROUGE-2 dan ROUGE-L pada model CNN+LSTM adalah 41.89, 10.24, 17.93 sedangkan pada model BiLSTM adalah 43.43, 11.40, 18.67. Secara keseluruhan dapat dilihat bahwa metode BiLSTM ini memiliki performa yang lebih baik dibandingkan dengan model *baseline* CNN+LSTM.

Berdasarkan hasil yang ditunjukkan dari penelitian di atas metode BiLSTM telah terbukti berhasil dalam menghasilkan sebuah ringkasan teks yang baik, maka penulis memutuskan untuk menggunakan metode ini pada penelitian peringkasan teks berita dalam bahasa Indonesia secara abstraktif.

1.3. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, fokus permasalahan pada penelitian ini adalah menghasilkan sebuah ringkasan teks pada berita dalam

bahasa Indonesia secara abstraktif menggunakan metode *Attention-Based BiLSTM (Bidirectional Long Short Term Memory)*. Maka, rumusan masalah pada penelitian ini adalah :

1. Bagaimana cara mengembangkan sebuah sistem peringkasan teks berita dalam bahasa Indonesia secara abstraktif menggunakan metode *attention-based BiLSTM*?
2. Bagaimana hasil kinerja *attention-based BiLSTM* dalam melakukan peringkasan teks berita dalam bahasa Indonesia secara abstraktif berdasarkan evaluasi menggunakan metrik ROUGE?

1.4. Tujuan Penelitian

Sesuai rumusan masalah, maka tujuan dari penelitian ini adalah sebagai berikut :

1. Menghasilkan sebuah sistem peringkasan teks berita dalam bahasa Indonesia menggunakan metode *attention-based BiLSTM (Bidirectional Long Short Term Memory)*.
2. Mengetahui kualitas ringkasan yang dihasilkan berdasarkan hasil evaluasi metrik ROUGE terhadap prediksi peringkasan teks secara abstraktif dalam bahasa Indonesia menggunakan metode *attention-based BiLSTM*.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Membuat informasi menjadi lebih mudah diakses oleh masyarakat yang memiliki keterbatasan waktu dalam membaca teks berita yang panjang.
2. Dengan adanya sistem peringkasan teks berita dalam bahasa Indonesia secara abstraktif menggunakan metode *attention-based BiLSTM* ini

diharapkan dapat membantu meringkas jutaan teks berita yang ada secara efektif dan berkualitas.

1.6. Batasan Masalah

Pada penelitian ini diperlukan batasan-batasan agar tujuan penelitian dapat tercapai. Adapun batasan masalah yang dibahas pada penelitian ini adalah :

1. Data masukan yang digunakan berupa artikel berita dengan jenis *single document*.
2. Dataset menggunakan data yang diambil dari IndoSum (*Indonesian Text Summarization*) dalam format .jsonl dan tersedia secara publik.

1.7. Sistematika Penulisan

Sistematika penulisan skripsi ini mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini menguraikan mengenai pokok-pokok pikiran yang melandasi pembuatan penelitian, seperti latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas dasar-dasar teori yang digunakan dalam penelitian, seperti definisi peringkasan teks, algoritma attention-based, metode BiLSTM dan lainnya. Serta beberapa kajian literatur mengenai penelitian lain yang relevan dengan penelitian ini.

BAB III. METODOLOGI PENELITIAN

Pada bab ini menjelaskan langkah-langkah yang akan dikerjakan dalam melakukan penelitian. Seperti pengumpulan data, analisis data serta perancangan sistem yang akan dibangun. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci mengacu pada suatu kerangka kerja.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini membahas mengenai analisis dan perancangan perangkat lunak yang akan digunakan sebagai alat penelitian. Terdiri dari proses analisis kebutuhan, perancangan, implementasi, dan pengujian perangkat lunak.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini diuraikan hasil pengujian berdasarkan langkah-langkah yang telah direncanakan. Tabel hasil pengujian serta analisisnya disajikan sebagai basis dari kesimpulan yang akan diambil dalam penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab sebelumnya dan juga saran yang diberikan berdasarkan hasil dari penelitian.

1.8. Kesimpulan

Pada penelitian ini akan menjelaskan bagaimana cara melakukan peringkasan teks secara abstraktif menggunakan metode attention-based BiLSTM. Penjelasan mengenai mekanisme sistem dalam melakukan peringkasan teks serta mengetahui banyak keluaran yang dihasilkan oleh sistem.

DAFTAR PUSTAKA

- Adelia, R., Suyanto, S., & Wisesty, U. N. (2019). Indonesian abstractive text summarization using bidirectional gated recurrent unit. *Procedia Computer Science*, 157, 581–588.
- Andriani, D., Indriati, & Furqon, M. T. (2019). Peringkasan Teks Otomatis Pada Artikel Berita Hiburan Berbahasa Indonesia Menggunakan Metode BM25. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3).
- Anwar, A. (2014). A Review of RUP (Rational Unified Process). In Ashraf Anwar *International Journal of Software Engineering (IJSE)* (Issue 5).
- Arikunto, S., Supardi, & Suhardjono. (2021). *Penelitian Tindakan Kelas: Edisi Revisi*.
- Baru, C., Institute of Electrical and Electronics Engineers, & IEEE Computer Society. (2019). *IEEE International Conference on Big Data: proceedings: Dec 9 - Dec 12, 2019, Los Angeles, CA, USA*.
- Barash, Y., Konen, E., Klang, E., & Sorin, V. (2020). Deep Learning for Natural Language Processing in Radiology—Fundamentals and a Systematic Review. *Journal of the American College of Radiology*, 17(5), 639–648.
- Effendy, E., Tursina, I., Aini, L., & Rizky, M. A. (2023). Dasar-Dasar Penulisan Berita. *Jurnal Dakwah Islam*, 7(1).
- Elsaid, A., Mohammed, A., Ibrahim, L. F., & Sakre, M. M. (2022). A Comprehensive Review of Arabic Text Summarization. *IEEE Access*, 10.
- Fadli, H. F., & Hidayatullah, A. F. (2021). Identifikasi Cyberbullying pada Media Sosial Twitter Menggunakan Metode LSTM dan BiLSTM.
- Ivanedra, K., & Mustikasari, M. (2019). Implementasi Metode Recurrent Neural Network pada Text Summarization dengan Teknik Abstraktif. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 6(4).
- Jiang, J., Zhang, H., Dai, C., Zhao, Q., Feng, H., Ji, Z., & Ganchev, I. (2021). Enhancements of Attention-Based Bidirectional LSTM for Hybrid Automatic Text Summarization. *IEEE Access*, 9.

- Jbene, M., Raif, M., Tigani, S., Chehri, A., & Saadane, R. (2022). User Sentiment Analysis in Conversational Systems Based on Augmentation and Attention-based BiLSTM. *Procedia Computer Science*, 207.
- Kostadinov, S. (2019). Understanding Encoder-Decoder Sequence to Sequence Model. *Towards Data Science*.
- Li, G., Li, F., Ahmad, T., Liu, J., Li, T., Fang, X., & Wu, Y. (2022). Performance evaluation of sequence-to-sequence-Attention model for short-term multi-step ahead building energy predictions. *Energy*, 259, 124915.
- Liu, W., Gao, Y., Li, J., & Yang, Y. (2021). A Combined Extractive with Abstractive Model for Summarization. *IEEE Access*, 9.
- Lopez, M. M., & Kalita, J. (2017). Deep Learning applied to NLP. <http://arxiv.org/abs/1703.03091>
- Mandar, G., & Gunawan, G. (2017). Peringkasan dokumen berita bahasa indonesia menggunakan metode cross latent semantic analysis. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 3(2).
- Mohammed Badry, R., Sharaf Eldin, A., & Saad Elzanfally, D. (2013). Text Summarization within the Latent Semantic Analysis Framework: Comparative Study. *International Journal of Computer Applications*, 81(11).
- Mustaqhfiri, M., Abidin, Z., & Kusumawati, D. R. (2011). Peringkasan Teks Otomatis Berita Berbahasa Indonesia Menggunakan Metode Maximum Marginal Relevance.
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, 18(5).
- Rahimi, S. R., Mozhdehi, A. T., & Abdolahi, M. (2018). An overview on extractive text summarization. *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation, KBEI 2017*, 2018-January.
- Setyawan, H., & Waryanto, H. (2016). Peringkasan Teks Bahasa Indonesia Menggunakan Modified Discrete Differential Evolution Algorithm.

- Sorin, V., Barash, Y., Konen, E., & Klang, E. (2020). Deep Learning for Natural Language Processing in Radiology—Fundamentals and a Systematic Review. *Journal of the American College of Radiology*, 17(5), 639–648.
- Subowo, E., Adi Artanto, F., Putri, I., & Umaedi, W. (2022). BLTSM untuk analisis sentimen berbasis aspek pada aplikasi belanja online dengan cicilan.
- Veningston, K., Venkateswara Rao, P. V., & Ronalda, M. (2022). Personalized Multi-document Text Summarization using Deep Learning Techniques. *Procedia Computer Science*, 218, 1220–1228.
- Wazery, Y. M., Saleh, M. E., Alharbi, A., & Ali, A. A. (2022). Abstractive Arabic Text Summarization Based on Deep Learning. *Computational Intelligence and Neuroscience*, 2022.
- Xiaoyan, L., & Raga, R. C. (2023). BiLSTM Model With Attention Mechanism for Sentiment Classification on Chinese Mixed Text Comments. *IEEE Access*, 11, 26199–26210.