

**PERINGKASAN TEKS OTOMATIS MENGGUNAKAN
METODE *BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS (BERT)*
PADA ARTIKEL BERITA POLITIK BERBAHASA
INDONESIA**



**OLEH
ELMA MARGARETHA BR SEBAYANG
NIM 2015101018**

**PROGRAM STUDI ILMU KOMPUTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA**

2025

**PERINGKASAN TEKS OTOMATIS MENGGUNAKAN
METODE *BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS* (BERT)
PADA ARTIKEL BERITA POLITIK BERBAHASA
INDONESIA**

SKRIPSI

**Diajukan kepada
Universitas Pendidikan Ganesha
Untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan
Program Sarjana Komputer**

**Oleh
Elma Margaretha Br Sebayang
NIM 2015101018**

**PROGRAM STUDI (S1) ILMU KOMPUTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA**

2025

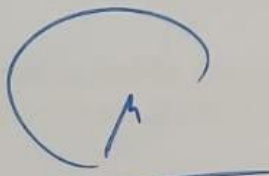
SKRIPSI

**DIAJUKAN UNTUK MELENGKAPI
TUGAS DAN MEMENUHI SYARAT-SYARAT
UNTUK MENCAPAI GELAR SARJANA KOMPUTER**

Menyetujui

Pembimbing I,

Pembimbing II,



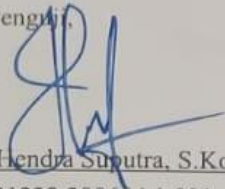
Ni Putu Novita Puspa Dewi, S.Kom., M.Cs.
NIP. 19941003 202012 2 015



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 19891026 201903 1 004

Skripsi oleh Elma Margaretha Br Sebayang
Telah dipertahankan di depan dewan penguji
Pada tanggal 4 Februari 2025

Dewan Penguji,



Dr. Putu Hendra Suputra, S.Kom., M.Cs.
NIP. 19821222 200604 1 001

(Ketua)



Dr. Ni Ketut Kertiasih, S.Si., M.Pd.
NIP. 19701118 199703 2 001

(Anggota)



Ni Putu Novita Puspa Dewi, S.Kom., M.Cs.
NIP. 19941003 202012 2 015

(Anggota)



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 19891026 201903 1 004

(Anggota)

Diterima oleh panitia Ujian Fakultas Teknik dan Kejuruan
Universitas Pendidikan Ganesha
Guna memenuhi syarat-syarat untuk mencapai gelar Sarjana Komputer

Pada :

Hari : Senin
Tanggal : 17.0.FEB.2025



Mengetahui,

Ketua Ujian

Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.
NIP. 19821111 200812 1 001

Sekretaris Ujian

I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 19891026 201903 1 004

Mengesahkan

Dekan Fakultas Teknik dan Kejuruan



Dr. Kadek Rihendra Dantes, S.T., M.T.

NIP. 19791201 200604 1 001

PERNYATAAN KEASLIAN TULISAN

Nama : Elma Margaretha Br Sebayang
Tempat, dan Tanggal Lahir : Tigabinanga, 15 Maret 2001
NIM : 2015101018
Fakultas : Teknik dan Kejuruan

Dengan ini saya menyatakan bahwa karya tulis yang berjudul **“Peringkasan Teks Otomatis Menggunakan Metode Bidirectional Encoder Representations From Transformers (Bert) Pada Artikel Berita Politik Berbahasa Indonesia”** beserta dengan seluruh isinya adalah benar-benar karya sendiri dan saya tidak melakukan penjiplakan dan pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran atas keilmuan dalam karya saya ini atau ada klaim terhadap keaslian karya saya ini.

Singaraja, 04 Februari 2025

Yang membuat pernyataan,



Elma Margaretha Br Sebayang

NIM 2015101018

MOTTO

“Apa yang Kuperbuat, engkau tidak tahu sekarang, tetapi engkau akan mengertinya kelak.”

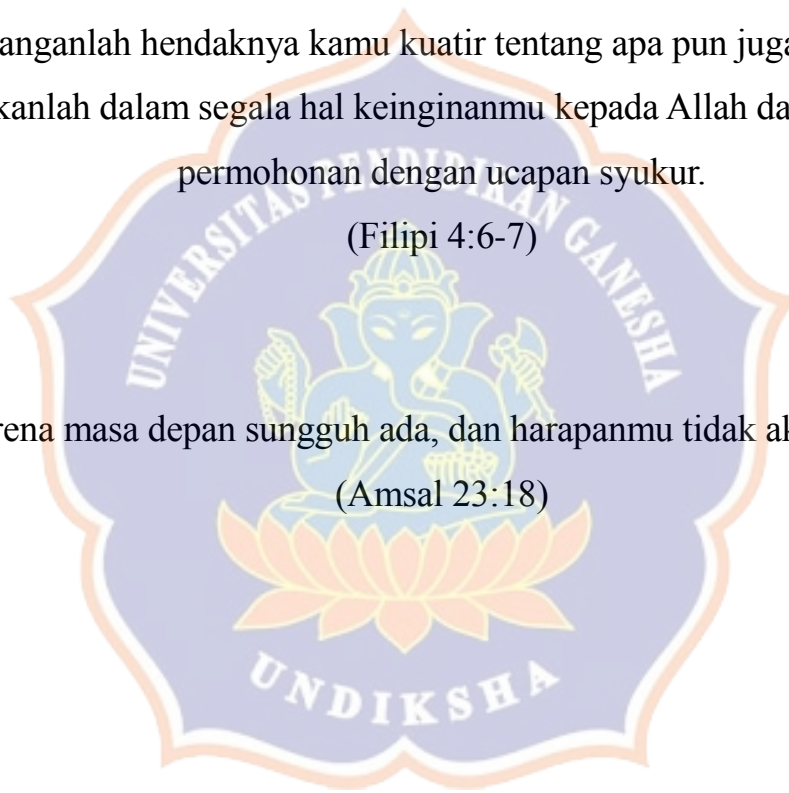
(Yohanes 13:7)

Janganlah hendaknya kamu kuatir tentang apa pun juga, tetapi nyatakanlah dalam segala hal keinginanmu kepada Allah dalam doa dan permohonan dengan ucapan syukur.

(Filipi 4:6-7)

"Karena masa depan sungguh ada, dan harapanmu tidak akan hilang"

(Amsal 23:18)



PRAKATA

Puji syukur penyusun panjatkan ke hadapan Allah subhanahu wa ta'ala karena berkat rahmat-Nya-lah, penulis dapat menyelesaikan skripsi yang berjudul **“Peringkasan Teks Otomatis Menggunakan Metode Bidirectional Encoder Representations From Transformers (Bert) Pada Artikel Berita Politik Berbahasa Indonesia”**. Skripsi ini disusun guna memenuhi persyaratan mencapai gelar sarjana komputer pada Universitas Pendidikan Ganesha.

Dalam menyelesaikan skripsi ini, penulis banyak mendapat bantuan baik berupa moral maupun material dari berbagai pihak. Untuk itu, dalam kesempatan ini, penulis mengucapkan terima kasih kepada:

1. Prof. Dr. I Wayan Lasmawan, M.Pd selaku Rektor Universitas Pendidikan Ganesha atas fasilitas-fasilitas yang telah diberikan selama penunjang perkuliahan dan kebijakan serta program-program yang dilaksanakan di Universitas Pendidikan Ganesha.
2. Dr. Kadek Rihendra Dantes, S.T., M.T. selaku Dekan Fakultas Teknik dan Kejuruan beserta jajarannya yang telah membantu mempermudah administrasi selama penyusunan skripsi ini.
3. Dr. Putu Hendra Suputra, S.Kom., M.Cs. selaku Ketua Jurusan Teknik dan Informatika serta selaku Penguji I atas motivasi dan bantuan yang telah diberikan dalam penyusunan skripsi ini.
4. I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. selaku Koordinator Program Studi Ilmu Komputer, Pembimbing II sekaligus Pembimbing Akademik atas kebijakan, motivasi, bantuan dan saran yang diberikan dalam penyusunan skripsi ini.
5. Ni Putu Novita Puspa Dewi, S.Kom., M.Cs. selaku Pembimbing I atas kebijakan, motivasi, bantuan dan saran yang diberikan dalam penyusunan skripsi ini.
6. Dr. Ni Ketut Kertiasih, S.Si., M.Pd. selaku Penguji II yang telah memberikan bimbingan, arahan dan petunjuk yang berarti dalam penyusunan skripsi ini.
7. Seluruh staf pengajar di Jurusan Teknik Informatika, Program Studi Ilmu Komputer, yang telah berbagi ilmu, memberikan motivasi, serta pengalaman

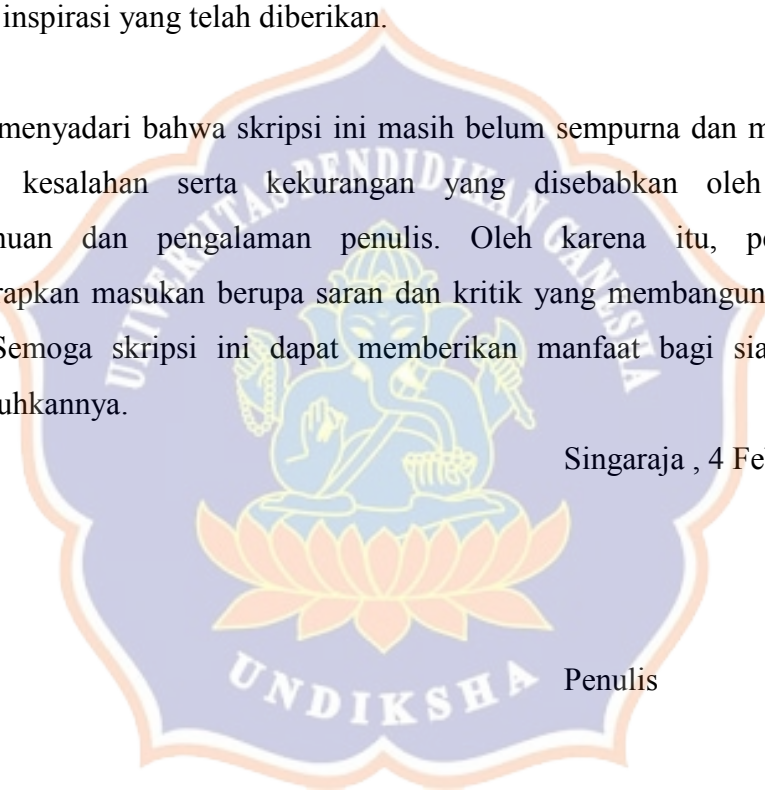
belajar yang berharga selama menempuh studi di Universitas Pendidikan Ganesha.

8. Ibu Mardiana Br Sitepu, yang telah berjuang tanpa lelah demi pendidikan penulis, sehingga penulis dapat mencapai jenjang pendidikan saat ini. Terima kasih atas doa, dukungan, dan kasih sayangnya yang tiada henti.
9. Kedua saudara kandung saya, Deoriski Febrian Sebayang dan Sani Julian Sebayang, atas setiap doa dan dukungan yang selalu diberikan dalam setiap langkah perjalanan akademik penulis.
10. Biring (Nenek), Grace Natalie, dan seluruh keluarga besar Sitepu, yang senantiasa mendoakan dan memberikan dukungan moral selama penulis menyelesaikan studi ini.
11. Viola Nikita, sahabat sejak SD hingga sekarang yang selalu setia menemani dalam setiap perjalanan hidup. Selalu hadir saat penulis membutuhkan bantuan, dengan sabar mendengarkan setiap keluhan, dan selalu menanyakan perkembangan skripsi ini. Terima kasih atas segala kebersamaan dan dukungan yang tak tergantikan.
12. Deby Enina Eninta, sahabat sejak SMA hingga sekarang yang telah mengetahui setiap perjalanan hidup penulis, khususnya dalam menempuh pendidikan ini. Terima kasih karena selalu ada, tetap mendukung, dan setia menemani hingga saat ini.
13. Junita Angelia, sahabat yang telah menemani penulis selama menempuh pendidikan di Universitas Pendidikan Ganesha. Terima kasih telah banyak meluangkan waktu untuk menemani, membantu, mendukung, dan selalu menanyakan perkembangan skripsi ini.
14. Trio Jegeg, yakni Ira dan Ruth, yang telah menjadi keluarga di tanah rantau ini. Tanpa kalian, perjalanan akademik ini mungkin akan terasa hambar. Sejak awal masa perkuliahan, kalian selalu menemani, membantu, dan menjadi pendengar setia atas setiap keluh kesah selama menjalani studi. Terima kasih atas kebersamaan, dukungan, dan persahabatan yang begitu berharga.
15. Tina, Yuda, dan Bayu, teman dekat yang ditemukan di KKN namun tetap setia menemani perjalanan akademik hingga kini. Terima kasih telah menjadi sahabat yang luar biasa.

16. Rich Girls, yakni Chintia Theresia Ginting, Ermilia Laurensya Tarigan, dan Chindi Auriska serta Junita yang selalu mendukung, memotivasi, serta mendoakan penulis dalam setiap perjalanan akademik ini semoga kita semua benar-benar menjadi rich seperti nama kita!
17. Seluruh teman-teman di Program Studi Ilmu Komputer Angkatan 2020 yang telah berbagi pengalaman, kebersamaan, serta dukungan selama penulis menempuh studi.
18. Seluruh pihak yang telah memberikan motivasi dan dukungan kepada penulis, yang tidak dapat disebutkan satu per satu. Terima kasih atas segala bantuan dan inspirasi yang telah diberikan.

Penulis menyadari bahwa skripsi ini masih belum sempurna dan mungkin masih terdapat kesalahan serta kekurangan yang disebabkan oleh keterbatasan pengetahuan dan pengalaman penulis. Oleh karena itu, penulis sangat mengharapkan masukan berupa saran dan kritik yang membangun dari berbagai pihak. Semoga skripsi ini dapat memberikan manfaat bagi siapa saja yang membutuhkannya.

Singaraja , 4 Februari 2025



Penulis

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN PERSETUJUAN DOSEN PEMBIMBING	ii
HALAMAN PERSETUJUAN DEWAN PENGUJI	iv
HALAMAN PERSETUJUAN PANITIA UJIAN	v
MOTTO	vi
PERNYATAAN KEASLIAN TULISAN	vii
PRAKATA	viii
ABSTRAK	xi
ABSTRACT	xii
DAFTAR ISI	xiii
DAFTAR TABEL	xvi
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Batasan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
BAB II LANDASAN TEORI	6
2.1 Penelitian Terkait	6
2.2 Ringkasan	9
2.3 Pengertian Teks Berita	11
3.6.1 Struktur Teks Berita	12
2.4 Peringkasan Teks Otomatis	12

2.4.1 Metode Text Summarization.....	13
2.5 BERT-SUM.....	14
2.6 DISTILBERT	16
2.7 Evaluasi dengan Recall -Oriented Understudy for Gisting Evaluation (ROUGE)	18
BAB III METODOLOGI PENELITIAN.....	20
3.1 Literature Review	22
3.2 Jenis Penelitian.....	23
3.3 Sumber Data.....	23
3.3.1 Data Primer	24
3.3.2 Data Sekunder.....	24
3.4 Proses Penyusunan Ringkasan Secara Manual	25
3.5 Data Pre-Processing dan Data Splitting	26
3.6 Modeling	27
3.6.1 BERT-SUM	28
3.6.2 DISTILBERT.....	31
3.7 Evaluasi Hasil.....	32
3.8 Cara Kerja ROUGE.....	32
BAB IV HASIL DAN PEMBAHASAN.....	34
4.1 Hasil Penelitian	34
4.1.1 Persiapan Dataset.....	34
4.1.2 Pre-Processing	37
4.2 Modeling	44
4.3 Hasil Ringkasan yang Dihasilkan Oleh Model	46
1. Dataset Politik.....	47
2. Dataset Variasi	49

3. Dataset Indosum	51
4.4 Evaluasi Kinerja Model.....	53
4.5 Analisis Hasil	55
BAB V KESIMPULAN DAN SARAN.....	57
5.1 Kesimpulan.....	57
5.2 Saran.....	58
DAFTAR PUSTAKA	59
LAMPIRAN.....	61



DAFTAR TABEL

Tabel 2.1 Kajian Penelitian Relevan	6
Tabel 4.1 Struktur Dataset.....	36
Tabel 4.2 Hasil <i>Noise Removal</i>	37
Tabel 4.3 Hasil <i>Sentence Tokenization</i>	38
Tabel 4.4 Hasil <i>Word Tokenization</i>	39
Tabel 4.5 Hasil <i>Remove Punctuation Marks</i>	39
Tabel 4.6 Hasil <i>Convert to Lowercase</i>	40
Tabel 4.7 Hasil <i>Normalize Whitespace</i>	41
Tabel 4.8 Hasil <i>Tokenisasi</i>	44
Tabel 4.9. Hasil Ringkasan Pada Dataset Politik	48
Tabel 4.10. Hasil Ringkasan Pada Dataset Variasi.....	50
Tabel 4.11. Hasil Ringkasan Pada Dataset <i>Indosum</i>	52
Tabel 4.12. Hasil Evaluasi Model Berdasarkan <i>F1-Score</i>	54



DAFTAR GAMBAR

Gambar 2.1 Metode <i>Text Summarization</i>	13
Gambar 2.2 Struktur <i>BERT-SUM Text Summarization</i>	16
Gambar 2.3 Struktur <i>DISTILBERT</i>	17
Gambar 3.1 Struktur Alur Penelitian.....	21



DAFTAR LAMPIRAN

Lampiran 1. <i>Install Packages</i>	61
Lampiran 2. <i>Import Library</i> (3)	61
Lampiran 3. <i>Import Library</i> (2)	61
Lampiran 4. <i>Load Dataset</i> (1).....	62
Lampiran 5. <i>Load Dataset</i> (2).....	62
Lampiran 6. <i>Load Dataset</i> (3).....	63
Lampiran 7. <i>Dataset Processing and Split</i>	63
Lampiran 8. <i>BERT-SUM</i> (1)	63
Lampiran 9. <i>BERT-SUM</i> (2)	64
Lampiran 10. <i>BERT-SUM</i> (3)	64
Lampiran 11. <i>BERT-SUM</i> (4).....	65
Lampiran 12. <i>BERT-SUM</i> (5)	65
Lampiran 13. <i>DISTILBERT</i> (1).....	65
Lampiran 14. <i>DISTILBERT</i> (2).....	66
Lampiran 15. <i>DISTILBERT</i> (3).....	66
Lampiran 16. <i>DISTILBERT</i> (4).....	67
Lampiran 17. <i>ROUGE Score</i> (1).....	67
Lampiran 18. <i>ROUGE Score</i> (2).....	67
Lampiran 19. <i>ROUGE Score</i> (3).....	68
Lampiran 20. <i>ROUGE Score</i> (4).....	68
Lampiran 21. <i>Jumlah Token Dataset</i>	69
Lampiran 22. <i>Result</i>	70
Lampiran 23. <i>Hasil Ringkasan Peringkas</i>	71