

**SKRIPSI**

***ABSTRACTIVE TEXT SUMMARIZATION* PADA BERITA WEBSITE  
MEDIA CENTER PALANGKA RAYA MENGGUNAKAN  
TRANSFORMER MODEL**



**DISUSUN OLEH:**

**NOVITA NATALIA**

**213020503032**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS PALANGKA RAYA**

**2026**

**ABSTRACTIVE TEXT SUMMARIZATION PADA BERITA WEBSITE MEDIA CENTER PALANGKA RAYA MENGGUNAKAN TRANSFORMER MODEL**

**SKRIPSI**

Sebagai salah satu syarat menyelesaikan Program Strata-1 pada Jurusan Teknik Informatika Fakultas Teknik Universitas Palangka Raya

Oleh:

**Novita Natalia**

**213020503032**

**Telah dipertahankan didepan tim penguji, pada:**

Hari/Tanggal : Kamis, 07 Mei 2026

Waktu : 10.00-11.30 WIB

Tempat : Ruang Ujian 1

1. **DRS. JADIAMAN PARHUSIP, M.Kom.** : ..... (Ketua Penguji)  
NIP. 196304231985021001
2. **EFRANS CHRISTIAN., S.T., M.T.** : ..... (Pembimbing Utama/Pertama)  
NIP. 199106302019031013
3. **FELICIA SYLVIANA, S.T., M.M.** : ..... (Pembimbing Pendamping/Kedua)  
NIP. 197601182003122003
4. **DEDDY RONALDO, S.T., M.T.** : ..... (Anggota)  
NIP. 198012262008121002

Mengetahui :

**JURUSAN / PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK UNIVERSITAS PALANGKA RAYA**

Ketua Jurusan,



**ARIELA LESTARI, S.Kom., M.Cs., Ph.D.**

**NIP. 19800322 200501 2 004**

## PERNYATAAN

Dengan ini saya menyatakan dengan sebenar-benarnya bahwa dalam Skripsi ini tidak terdapat karya ilmiah yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi, serta tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam Skripsi ini dan disebutkan dalam Tinjauan Pustaka.

Palangka Raya. 07 Mei 2026



Novita Natalia

21302050303

## RIWAYAT PENYUSUN

### Data Diri

Nama : Novita Natalia  
NIM : 213020503032  
Fakultas : Teknik  
Jurusan/Program Studi : Teknik Informatika  
Tempat, Tanggal Lahir : Tumbang Habaon. 25 Oktober 2003  
Agama : Kristen  
Status Dalam Keluarga : Anak Kandung  
Anak Ke : 1 (Pertama)  
Alamat : Desa Tumbang Habaon, Kab.Gunung Mas, Kec.  
Tewah.  
No.Telepon/Hp : 0859 7164 2040



### Data Diri Orang Tua

Nama Ayah : Apollo Theodor  
Pekerjaan : Petani  
Nama Ibu : Berliana  
Pekerjaan : Petani  
Alamat Orang Tua : Desa Tumbang Habaon, Kab.Gunung Mas,  
Kec.Tewah  
No.Telepon/HP : 0877 6480 9621

### Riwayat Pendidikan

SD : SDN 1 Rangan Mumpung (Tahun Lulus 2015)  
SMP : SMP NEGRI 4 Tewah (Tahun Lulus 2018)  
SMA : SMA NEGRI 2 TEWAH (Tahun Lulus 2021)

## HALAMAN PERSEMBAHAN

Puji syukur ke hadirat Tuhan yang maha esa, atas limpahan rahmat, dan karuniaNya, sehingga penulis dapat menyelesaikan skripsi ini dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan pada jenjang S1 di Jurusan Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya. Penulis juga mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan memberikan dukungan selama proses penyusunan skripsi ini, terutama kepada:

1. Kedua orang tua dan seluruh keluarga yang selalu menjadi sumber kekuatan terbesar. Terima kasih atas doa yang tak pernah putus, kasih sayang yang tak terhingga, serta dukungan penuh yang telah diberikan selama ini.
2. Ibu Frieda, S.T., M.T., selaku Dekan Fakultas Teknik Universitas Palangka Raya.
3. Ibu Ariesta Lestari, S.Kom., M.Cs., Ph.D., selaku Ketua Jurusan Teknik Informatika Fakultas Teknik Universitas Palangka Raya.
4. Bapak Deddy Ronaldo, S.T., M.T., selaku Dosen Pembimbing Akademik dan Dosen Penguji II, atas dukungan, bimbingan dan motivasi selama masa studi.
5. Bapak Drs. Jadianan Parhusip, M.Kom., selaku ketua penguji, atas segala masukan, kritik, dan saran yang diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan lebih baik.
6. Bapak Efrans Christian, S.T., M.T., selaku Dosen Pembimbing I dan Ibu Felicia Sylviana, S.T., M.T., selaku Dosen Pembimbing II, atas segala ilmu, arahan, waktu, kesabaran, motivasi dan dukungan dalam membimbing penulis dari awal penyusunan hingga skripsi ini terselesaikan.
7. Seluruh staf Jurusan Teknik Informatika, atas bantuan dan layanan yang mendukung kelancaran studi dan penyusunan skripsi ini.
8. Sahabat dan teman-teman saya, Ayudya Octaviani, Liani Cikita S, Meity Damaria, Dita Yulianti, Hana Husna S., Aurel Diya M W, Novita Kristiana, serta semua pihak lainnya yang tidak dapat disebutkan satu per satu, yang telah

memberikan dukungan, semangat, nasihat, motivasi, dan kebersamaan yang tidak ternilai selama perjalanan ini.

9. Penulis ucapkan terimakasih kepada Yoga Pranata, terima kasih telah meminjamkan laptop sehingga penulis dapat melanjutkan dan menyelesaikan skripsi ini dengan baik. Terima kasih telah menjadi bagian dari perjalanan ini, atas doa, dukungan, perhatian, serta kehadiran yang selalu memberi semangat dalam setiap proses penyusunan skripsi. Terima kasih karena selalu menemani di setiap suka maupun lelah hingga penulis mampu menyelesaikan pendidikan ini.

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yesus Yang Maha Kuasa, karena berkat kasih, penyertaan, dan karunia-Nya, penulis dapat menyelesaikan penyusunan skripsi ini dengan baik. Penelitian ini berjudul **“Abstractive Text Summarization pada Berita Website Media Center Palangka Raya Menggunakan Transformer Model”**.

Laporan skripsi ini disusun sebagai sumber pengetahuan dan informasi bagi semua pihak, serta diharapkan dapat menjadi referensi bagi mahasiswa yang mengerjakan skripsi di bidang yang sama, sehingga dapat dipahami dan dikembangkan lebih lanjut.

Penulis menyadari bahwa laporan skripsi ini masih memiliki kekurangan. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan demi penyempurnaan laporan ini. Akhir kata, penulis berharap skripsi ini dapat memberikan manfaat bagi semua pihak, khususnya bagi mahasiswa Jurusan Teknik Informatika.

Palangka Raya, 07 Mei 2026

Novita Natalia

# **ABSTRACTIVE TEXT SUMMARIZATION PADA BERITA WEBSITE MEDIA CENTER PALANGKA RAYA MENGGUNAKAN TRANSFORMER MODEL**

**Novita Natalia (213020503032)**

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya

Kampus Tanjung Nyaho, Jl. Yos Sudarso, Palangka Raya 73111

Email: nvtnatalia25@gmail.com

## **ABSTRAK**

Perkembangan Deep Learning dalam bidang Natural Language Processing (NLP) memungkinkan peringkasan teks berita secara otomatis. Pendekatan abstraktif berbasis arsitektur Transformer terbukti lebih efektif dibandingkan metode RNN dan LSTM dalam menangkap konteks global teks, salah satunya melalui model IndoBART yang dirancang khusus untuk bahasa Indonesia. Kebutuhan akan penyajian informasi yang ringkas semakin penting pada website Media Center Palangka Raya yang memuat berita resmi pemerintah daerah dalam jumlah besar, sehingga penelitian ini memanfaatkan tambahan dataset dari Detik.com dan Multi media center untuk meningkatkan kemampuan generalisasi model.

Penelitian ini bertujuan menerapkan model Transformer IndoBART untuk abstractive text summarization pada berita Media Center Palangka Raya, Detik.com, dan Multi Media Center (MMC). Data dikumpulkan dari sekitar 6.000 URL berita dan melalui preprocessing diperoleh 5.678 data bersih, yang selanjutnya disertai gold summary serta dibagi menjadi data latih, validasi, dan uji. Model kemudian dilakukan fine-tuning dan dievaluasi menggunakan metrik ROUGE.

Hasil penelitian menunjukkan bahwa model IndoBART mampu menghasilkan ringkasan berita yang informatif dengan performa yang baik berdasarkan evaluasi ROUGE, dengan nilai ROUGE-1 sebesar 41,85 yang mencerminkan kemampuan model menangkap kata-kata penting dari ringkasan referensi, ROUGE-2 sebesar 17,92 yang menunjukkan koherensi pasangan kata dan struktur kalimat, serta ROUGE-L dan ROUGE-Lsum sekitar 29,96 yang menandakan kesesuaian alur dan struktur ringkasan dengan ringkasan manual. Selain itu, sistem berbasis web menggunakan Streamlit berhasil memfasilitasi proses peringkasan berita secara otomatis, sehingga IndoBART terbukti efektif untuk abstractive text summarization pada berita pemerintahan dan lokal berbahasa Indonesia.

**Kata Kunci:** Abstractive Text Summarization, IndoBART, Transformer, NLP, ROUGE

**ABSTRACTIVE TEXT SUMMARIZATION ON THE NEWS ON THE  
PALANGKA RAYA MEDIA CENTER WEBSITE USING THE  
TRANSFORMER MODEL**

**Novita Natalia (213020503032)**

*Department of Informatics Engineering, Faculty of Engineering, University of  
Palangka Raya*

*Tanjung Nyaho Campus, Jl. Yos Sudarso, Palangka Raya 73111*

*Email: nvtnatalia25@gmail.com*

**ABSTRACT**

*The development of Deep Learning in the field of Natural Language Processing (NLP) enables automatic news text summarization. Abstractive approaches based on the Transformer architecture have proven to be more effective than RNN and LSTM methods in capturing the global context of text, one of which is the IndoBART model specifically designed for the Indonesian language. The need for concise information presentation is increasingly important on the Media Center Palangka Raya website, which contains a large volume of official local government news. Therefore, this study incorporates additional datasets from Detik.com and Multi Media Center to enhance the model's generalization capability.*

*This study aims to implement the IndoBART Transformer model for abstractive text summarization on news articles from Media Center Palangka Raya, Detik.com, and Multi Media Center (MMC). The data were collected from approximately 6,000 news URLs and processed through preprocessing to obtain 5,678 clean data samples, which were then accompanied by gold summaries and divided into training, validation, and testing datasets. The model was fine-tuned and evaluated using the ROUGE metric.*

*The results indicate that the IndoBART model successfully generates informative news summaries with good performance based on ROUGE evaluation, achieving a ROUGE-1 score of 41.85, which reflects the model's ability to capture important words from the reference summaries, a ROUGE-2 score of 17.92, indicating coherence in word pairs and sentence structure, and ROUGE-L and ROUGE-Lsum scores of approximately 29.96, demonstrating alignment in flow and structure with the manual summaries. In addition, the web-based system using Streamlit successfully facilitated the automatic news summarization process, so that IndoBART proved effective for abstractive text summarization of Indonesian-language government and local news.*

**Keywords:** *Abstractive Text Summarization, IndoBART, Transformer, NLP, ROUGE*

## DAFTAR ISI

<b>RIWAYAT PENYUSUN .....</b>	<b>iv</b>
<b>HALAMAN PERSEMBAHAN .....</b>	<b>v</b>
<b>KATA PENGANTAR .....</b>	<b>vii</b>
<b>ABSTRAK.....</b>	<b>viii</b>
<b>ABSTRACT .....</b>	<b>ix</b>
<b>DAFTAR ISI.....</b>	<b>x</b>
<b>DAFTAR TABLE .....</b>	<b>xiii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xiv</b>
<b>BAB I .....</b>	<b>1</b>
<b>PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah .....	3
1.4. Tujuan Penelitian .....	4
1.5. Manfaat Penelitian .....	4
1.6. Sistematika Penulisan .....	5
1.7. Jadwal Penelitian .....	5
<b>BAB II.....</b>	<b>6</b>
<b>LANDASAN TEORI .....</b>	<b>6</b>
2.1. Tinjauan Pustaka.....	6
2.2. Berita.....	7
2.3. Model Transformer.....	9
2.4. Model indoBART .....	10
2.5. Metrik ROUGE.....	11
2.6. Teks Summarization .....	14
2.7. Abstractive Summarization .....	17
2.8. Natural Language Processing.....	19
2.8.1. Text Input dan Data Collection .....	20
2.8.2. Representasi Teks dan Tokenisasi .....	20
2.8.3. Teks Representation.....	21

2.8.4. Model Selection and Training .....	21
2.8.5. Evaluation .....	21
2.9. Indek Pembangunan Literasi Masyarakat .....	22
2.10. Metode Pengembangan Sistem Waterfall .....	24
2.11. Unified Modelling Language (UML) .....	26
2.11.1. Use Case Diagram .....	27
2.11.2. Activity Diagram .....	28
<b>BAB III .....</b>	<b>30</b>
<b>METODOLOGI PENELITIAN .....</b>	<b>30</b>
3.1. Studi Literatur .....	31
3.2. Pengumpulan Data .....	31
3.2.1. Sumber Data .....	32
3.2.2. Scraping Data .....	33
3.2.3. Penyusunan Data Referensi (Gold Summary) .....	35
3.3. Preprocessing Data .....	35
3.3.1. Tahap Cleaning Teks .....	36
3.3.2. Tahap Lowercasing .....	36
3.3.3. Penerapan preprocessing pada dataset .....	36
3.4. Pembagian Dataset .....	37
3.5. Tokenisasi (Tokenizer IndoBART) .....	37
3.6. Tahap Pelatihan Model (Training Stage) .....	38
3.7. Tahap Fine-Tuning Model IndoBART .....	39
3.7.1. Parameter Fine-Tuning .....	40
3.8. Evaluasi Akhir Model IndoBART .....	41
3.9. Perancangan dan Desain Sistem .....	43
3.9.1. Requirements Analysis (Analisis Kebutuhan) .....	43
3.9.2. Design (Desain Sistem) .....	44
3.9.2.1. Diagram Alir Sistem (Flowchart) .....	45
3.9.2.2. Pemodelan Sistem (Unified Modeling Language - UML) .....	46
3.9.2.3. Desain Antarmuka .....	54
3.10. Pengujian Sistem .....	61
<b>BAB IV .....</b>	<b>62</b>

<b>HASIL DAN PEMBAHASAN .....</b>	<b>62</b>
4.1. Hasil Pengumpulan Data .....	62
4.2. Hasil Penyusunan Data Referensi (Gold Summary) .....	64
4.3. Hasil PreProcessing Data .....	66
4.2.1. Hasil Cleaning Teks dan Lowercassing .....	66
4.4. Hasil Pembagian Dataset .....	67
4.5. Hasil Tokenisasi (Tokenizer IndoBART) .....	68
4.6. Hasil Tahap Pelatihan Model (Training Stage) .....	69
4.7. Hasil Tahap Fine-tuning Model IndoBart .....	70
4.7.1. Interpretasi Metrik dan Stabilisas Pelatihan .....	70
4.8. Hasil Evaluasi Akhir Model IndoBART .....	71
4.8.1. Perbandingan dengan Penelitian Terdahulu (Baseline) .....	72
4.9. Hasil Implementasi dan Pengujian Sistem .....	74
4.9.1. Implementation .....	74
4.9.2. Testing .....	88
<b>BAB V .....</b>	<b>92</b>
<b>KESIMPULAN .....</b>	<b>92</b>
5.1. Kesimpulan .....	92
5.2. Saran .....	93
<b>DAFTAR PUSTAKA .....</b>	<b>95</b>

## DAFTAR TABLE

Gambar 2. 1. Berita .....	8
Gambar 2. 2. Arsitektur Model Transformer.....	10
Gambar 2. 3. Extrative dan Abstraktif Summarization.....	16
Gambar 2. 4. Natural Language Processing .....	20
Gambar 2. 5. Tingkatan Skor Indeks Pembangunan Literasi Pembangunan Masyarakat.....	24
Gambar 2. 6. Waterfall.....	25
Tabel 3. 1. Perbandingan struktur HTML.....	35
Tabel 3. 2. Parameter Fine tuning .....	40
Tabel 3. 3. Definisi Aktor .....	47
Tabel 3. 4. Definisi Use Case .....	48
Tabel 3. 5. Scenario Utama Menginput Berita.....	49
Tabel 3. 6. Scenario Alternatif Menginput Berita .....	49
Tabel 3. 7. Scenario Utama Melakukan Terjemahan.....	50
Tabel 3. 8. Scenario Alternatif Melakukan Terjemahan.....	50
Tabel 3. 9. Scenario Utama Melakukan About .....	50
Tabel 4. 1. Hasil pengumpulan data.....	62
Tabel 4. 2. Hasil Data Scraping.....	63
Tabel 4. 3. Gold_Summary.....	65
Tabel 4.4. Perbandingan Data Sebelum dan Sesudah Preprocessing.....	67
Tabel 4. 5. Hasil Pembagian Dataset .....	68
Tabel 4. 6. Parameter Konfigurasi Tokenizer.....	68
Tabel 4. 7. Statistik Panjang Token Dataset.....	69
Tabel 4. 8. Rincian Data pada Tahap Pelatihan.....	69
Tabel 4. 9. Hasil Fine Tuning Model IndoBART .....	70
Tabel 4. 10. Hasil Evaluasi Akhir Model IndoBART.....	71
Tabel 4. 11. Perbandingan Hasil Evaluasi Model dengan Baseline.....	72
Tabel 4. 12. Black Box Testing Website Peringkasan berita	88

## DAFTAR GAMBAR

Gambar 1. 1. Jadwal penelitian Skripsi.....	5
Gambar 2. 1. Berita .....	8
Gambar 2. 2. Arsitektur Model Transformer.....	10
Gambar 2. 3. Extrative dan Abstraktif Summarization.....	16
Gambar 2. 4. Natural Language Processing .....	20
Gambar 2. 5. Tingkatan Skor Indeks Pembangunan Literasi Pembangunan Masyarakat .....	24
Gambar 2. 6. Waterfall.....	25
Gambar 3. 1. Alur Penelitian .....	30
Gambar 3. 2. Flowchart Sistem .....	45
Gambar 3. 3. Uses Case Diagram sistem Peringkasan Berita .....	47
Gambar 3. 4. Activity Diagram Peringkasan Berita.....	52
Gambar 3. 5. Storyboard Halaman Metode Input Manual Text.....	57
Gambar 3. 6. Storyboard Halaman Metode Input News Url .....	58
Gambar 3. 7. Storyboard Halaman Translation.....	59
Gambar 3. 8. Storyboard Halaman About .....	60
Gambar 4. 1. Terminal VS Code.....	75
Gambar 4. 2. Implementasi Antarmuka Utama.....	75
Gambar 4. 3. Peringkasan Input Manual .....	76
Gambar 4. 4. Hasil Ringkasan dan Terjemahan Input Manual .....	78
Gambar 4. 5. Hasil Ringkasan Bahasa Indonesia.....	80
Gambar 4. 6. Peringkasan Input URL.....	81
Gambar 4. 7. Hasil Ringkasan dan Terjemahan input URL .....	83
Gambar 4. 8. Hasil Ringkasan Input URL Bahasa Indonesia .....	84
Gambar 4. 10. Hasil Translation.....	87
Gambar 4. 12. Tampilan Halaman About.....	87

## DAFTAR PUSTAKA

- Amalia, N. (2025). Transparansi vs publisitas dalam komunikasi publik pemerintah. *Jurnal Komunikasi dan Kebijakan Publik*, 10(1), 45–58.
- Aprilyada, G., Zidan, M. A., Nurlia, Ainunisa, R. A., & Winarti, W. (2023). Peran kajian pustaka dalam penelitian tindakan kelas. *Jurnal Kreativitas Mahasiswa*, 1(2), 165–173.
- Astuti, W., Muljono, A., & Sutriawan, A. (2024). Indonesian news text summarization using mBART algorithm. *Jurnal Informatika dan Teknologi*, 15(1), 112–125.
- Badan Pusat Statistik Provinsi Kalimantan Tengah. (2024). *Indeks Pembangunan Literasi Masyarakat dan Unsur Penyusunnya Menurut Kabupaten/Kota di Provinsi Kalimantan Tengah, 2023-2024*. Diakses dari <https://kalteng.bps.go.id>.
- Bahari, R., & Dewi, S. (2024). Analisis model Transformer untuk peringkasan teks berita otomatis. *Jurnal Teknologi Informasi*, 12(2), 88–95.
- Cahyawijaya, S., Winata, G. I., Wilie, B., Vincentio, K., Li, X., Kuncoro, A., ... & Fung, P. (2021). IndoNLG: Benchmark and resources for evaluating Indonesian natural language generation. *arXiv preprint arXiv:2104.03101*.
- Destriana, R., Husain, S. M., Handayani, N., & Siswanto, A. T. P. (2021). *Diagram UML dalam membuat aplikasi Android Firebase: Studi kasus aplikasi bank sampah*. Deepublish.
- Fadilah, N., & Mony, H. (2020). Kaidah bahasa Indonesia jurnalistik dalam media online. *Jurnal Komunikasi*, 14(1), 22–34.
- Halimah, N., et al. (2022). Peringkasan teks otomatis menggunakan algoritma LexRank pada artikel berita. *Jurnal Sistem Informasi*, 11(3), 201–210.
- Hartawan, G., et al. (2024). Bidirectional and Auto-Regressive Transformer (BART) for Indonesian abstractive text summarization. *Journal of Computer Science and Informatics*, 5(2), 45–55.
- Mulyasa, E. (2021). Penerapan Teknologi Informasi dalam Jurnalisme Modern. Bandung: Remaja Rosdakarya.
- Perpustakaan Nasional Republik Indonesia. (2023). *Laporan Indeks Pembangunan Literasi Masyarakat (IPLM) Tahun 2023*. Jakarta: Perpustakaan RI.

- Singh, A., Kaur, H., & Rani, R. (2023). Automatic short text summarization techniques in social media platforms. *Future Internet*, 15(11), 311. <https://doi.org/10.3390/fi15110311>.
- Suwandi, P., Putri, R., & Sulastri, S. (2024). Inovasi pendidikan dengan menggunakan model deep learning di Indonesia. *Jurnal Pendidikan Kewarganegaraan dan Politik*, 2(2), 69–77. <https://ejournal.upi.edu/index.php/jpknp/article/view/2024-02-69>.
- Taufiq, M., & Wahyudi, A. (2023). Implementasi metode Waterfall pada pengembangan sistem informasi. *Jurnal Rekayasa Perangkat Lunak*, 9(1), 12–20.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008. <https://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>.