



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



**IMPLEMENTASI MODEL *TEXT-TO-TEXT TRANSFER*
TRANSFORMER UNTUK PERINGKASAN ULASAN
OTOMATIS BERBASIS WEB**

SKRIPSI

RANDY AULIA ANANDA RUSLANI 2207411041

**POLITEKNIK
NEGERI
JAKARTA**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2026



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



**IMPLEMENTASI MODEL *TEXT-TO-TEXT TRANSFER*
TRANSFORMER UNTUK PERINGKASAN ULASAN
OTOMATIS BERBASIS WEB**

SKRIPSI

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan Untuk
Memperoleh Diploma Empat Politeknik**

RANDY AULIA ANANDA RUSLANI 2207411041

**POLITEKNIK
NEGERI
JAKARTA**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2026



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

SURAT PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan dibawah ini:

Nama : Randy Aulia Ananda Ruslani
NIM : 2207411041
Jurusan/Program : Teknik Informatika & Komputer / Sarjana Terapan
Studi : Teknik Informatika
Judul Skripsi : Implementasi *Text-to-Text* Transfer *Transformer*
untuk Peringkasan Ulasan Otomatis Berbasis Web.

Menyatakan dengan sebenarnya bahwa skripsi ini benar-benar merupakan hasil karya saya sendiri, bebas dari peniruan terhadap karya dari orang lain. Kutipan pendapat dan tulisan orang lain ditunjuk sesuai dengan cara-cara penulisan karya ilmiah yang berlaku.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa dalam skripsi ini terkandung ciri-ciri plagiat dan bentuk-bentuk peniruan lain yang dianggap melanggar peraturan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Depok, 21 Januari 2026

**POLITEKNIK
NEGERI
JAKARTA**

Materai
TTD

Randy Aulia Ananda Ruslani
NIM. 2207411041



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Randy Aulia Ananda Ruslani
NIM : 2207411041
Jurusan/Program : Teknik Informatika & Komputer / Sarjana Terapan
Studi : Teknik Informatika
Judul Skripsi : Implementasi *Text-to-Text* Transfer *Transformer*
untuk Peringkasan Ulasan Otomatis Berbasis Web.

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Rabu Tanggal 17, Bulan Juni, Tahun 2026, dan dinyatakan **LULUS**.

Disahkan oleh

Pembimbing I : Euis Oktavianti, S.Si., M.T.I. ()
Penguji I : Rizki Elisa Nalawati, S.T. M.T. ()
Penguji II : Dr. Ir. Dewi Yanti Liliyana, S.Kom., M.Kom. ()
Penguji III : Chairul Fikri, S.Kom., M.Kom. ()

Mengetahui :
Ketua Jurusan Teknik Informatika dan Komputer

Dr. Anita Hidayati, S. Kom., M. Kom.
NIP. 197908032003122003

Teknik Informatika dan Komputer – Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah Azza Wa Jalla atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan pada Program Sarjana Terapan Teknik Informatika dan Komputer, Politeknik Negeri Jakarta. Proses penyusunan skripsi ini menjadi pengalaman yang sangat berharga bagi penulis, karena tidak hanya menuntut pemahaman akademis, tetapi juga kemampuan berpikir kritis, kreatif, dan adaptif dalam menyelesaikan berbagai tantangan yang ditemui selama penelitian.

Penulis menyadari bahwa laporan ini tidak akan terselesaikan tanpa bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Ayah dan Ibu sosok motivator penulis yang saya sangat kasihi atas dukungan moral, doa, dan kasih sayang yang tiada henti.
2. Ibu Euis Oktavianti, S.Si., M.T.I, selaku dosen pembimbing, yang telah membimbing dan memberikan arahan dengan penuh kesabaran selama proses penulisan laporan ini;
3. Ibu Dr. Anita Hidayati, S.Kom., M.Kom., selaku ketua jurusan TIK yang telah memberikan kesempatan kepada saya untuk menempuh studi di Program sarjana terapan teknik informatika dan komputer Politeknik Negeri Jakarta;
4. PT. Telekomunikasi Indonesia, khususnya Mas Anton Suhartono, S.Stat. yang telah memberikan *resource* LLM yang sangat membantu saya dalam melakukan eksplorasi *dataset* selama pengerjaan skripsi
5. Serta sahabat dan tim eai enjinir haceespe, 'adalah pokoknya', teman-teman TI B 2022 yang saya banggakan (Aulia Putri S, Heical Chandra, Adam, dkk) yang telah menemani penulis selama masa persiapan seminar proposal.

Penulis berharap laporan tugas akhir ini dapat memberikan manfaat, baik bagi pengembangan ilmu pengetahuan, dunia industri, maupun sebagai catatan pengalaman berharga yang akan selalu dikenang. Semoga segala kebaikan dan bantuan yang diberikan dibalas oleh Allah Azza Wa Jalla dengan limpahan rahmat dan karunia-Nya



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

SURAT PERNYATAAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Politeknik Negeri Jakarta, saya bertanda tangan dibawah ini:

Nama : Randy Aulia Ananda Ruslani
NIM : 2207411041
Jurusan/Program Studi : Teknik Informatika & Komputer / Teknik Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta Hak Bebas Royalti Non-Eksklusif atas karya ilmiah saya yang berjudul:

IMPLEMENTASI TEXT-TO-TEXT TRANSFER TRANSFORMER UNTUK PERINGKASAN ULASAN OTOMATIS BERBASIS WEB

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non Eksklusif ini Politeknik Negeri Jakarta Berhak menyimpan, mengalih media-kan/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Depok, 21 Januari 2026

Materai
TTD


Randy Aulia Ananda Ruslani
NIM. 2207411041



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

IMPLEMENTASI *TEXT-TO-TEXT TRANSFER TRANSFORMER* UNTUK PERINGKASAN ULASAN OTOMATIS BERBASIS WEB

ABSTRAK

Ulasan konsumen pada platform e-commerce menjadi sumber informasi penting dalam menilai kualitas produk dan memengaruhi keputusan pembelian. Namun, jumlah ulasan yang besar dengan variasi gaya bahasa dan panjang teks menyulitkan pengguna dalam memperoleh informasi secara cepat. Penelitian ini bertujuan untuk mengembangkan sistem peringkasan ulasan produk berbasis web menggunakan model IndoT5 dengan pendekatan fine-tuning. Dataset yang digunakan berjumlah sekitar 7.000 pasangan ulasan dan ringkasan dalam bahasa Indonesia yang telah melalui tahap praproses. Evaluasi dilakukan menggunakan metrik ROUGE-L, BERTScore, penilaian ahli bahasa, serta pengujian usability melalui SUS dan NPS. Hasil penelitian menunjukkan bahwa model mampu menghasilkan ringkasan dengan ROUGE-L sebesar 0,4 dan BERTScore F1 sebesar 0,8, yang mengindikasikan kesesuaian yang baik terhadap ringkasan acuan. Penilaian ahli bahasa memperoleh rata-rata 4,3 dari 5 pada aspek keabstrakan, faithfulness, dan struktur bahasa, yang menunjukkan ringkasan informatif dan mudah dipahami. Selain itu, sistem memiliki tingkat kegunaan yang baik dengan SUS kategori B dan NPS sebesar 51% (excellent). Dengan demikian, sistem yang dikembangkan mampu menghasilkan ringkasan berkualitas dan membantu pengguna memahami ulasan secara lebih efisien.

Kata Kunci : peringkasan abstraktif, IndoT5, ulasan e-commerce, pemrosesan bahasa alami, *transformer*



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR ISI

SURAT PERNYATAAN BEBAS PLAGIARISME	i
LEMBAR PENGESAHAN	ii
KATA PENGANTAR	iii
SURAT PERNYATAAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	iv
ABSTRAK	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	ix
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan dan Manfaat	3
1.4.1 Tujuan Penelitian	3
1.4.2 Manfaat Penelitian	4
1.5 Sistematika Penulisan	4
BAB 2 TINJAUAN PUSTAKA	6
2.1 Aplikasi Peringkasan Teks	6
2.1.1 Ulasan	6
2.1.2 Peringkasan Teks	6
2.1.3 Peringkasan Ulasan Produk	7
2.2 E-Commerce	7
2.3 <i>Software Development Life Cycle (SDLC)</i>	7
2.3.1 Tahap Analisis	7
2.3.2 Tahap Perancangan	8
2.3.3 Tahap Implementasi.....	8
2.3.4 Tahap Pengujian	8
2.3.5 <i>Waterfall</i>	9
2.4 <i>Cross Industry Standard Process – Data Mining (CRISP-DM)</i>	10
2.4.1 <i>Business Understanding</i>	10
2.4.2 <i>Data Understanding</i>	10



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2.4.3	Data Preparation	10
2.4.4	Modelling	10
2.4.5	Evaluasi Model	18
2.4.6	Deployment	20
2.5	Penelitian Terkait	21
BAB 3 METODE PENELITIAN		23
3.1	Rancangan Penelitian	23
3.1.1	Tahapan Penelitian	24
3.1.2	Objek Penelitian	26
BAB 4 HASIL DAN PEMBAHASAN		27
4.1	Identifikasi Kebutuhan	27
4.1.1	Kebutuhan <i>Fine-tuning</i> Model	27
4.1.2	Kebutuhan Pengembangan Web	30
4.2	Perancangan Sistem	32
4.2.1	Perancangan Model	32
4.2.2	Perancangan Web	38
4.3	Implementasi Sistem	42
4.3.1	Implementasi Model	42
4.3.2	Implementasi Web	53
4.4	Pengujian Sistem	55
4.4.1	Deskripsi Pengujian	55
4.4.2	Prosedur Pengujian	59
4.4.3	Hasil dan Analisis Data Pengujian	59
BAB 5 PENUTUP		69
5.1	Kesimpulan	69
5.2	Saran	69
DAFTAR PUSTAKA		x
DAFTAR RIWAYAT HIDUP PENULIS		xvi
LAMPIRAN		xvii



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR TABEL

Tabel 1. Penelitian Terkait	21
Tabel 2. Kriteria Kebutuhan <i>Dataset</i>	27
Tabel 3. Kriteria Kebutuhan Model	29
Tabel 4. Kebutuhan Fungsional Web	30
Tabel 5. Kebutuhan Non-Fungsional Web	30
Tabel 6. Kebutuhan Perangkat Keras	31
Tabel 7. Kebutuhan Perangkat Lunak	31
Tabel 8. <i>Prompt Generate</i> Ulasan Sintesis	42
Tabel 9. <i>Prompt Generate</i> Ringkasan Ulasan	45
Tabel 10. Hasil Uji Kelayakan <i>Dataset</i>	47
Tabel 11. Hasil Fase Pertama (Pemilihan varian model)	48
Tabel 12. Hasil Fase Kedua (<i>Hyperparameter</i> Tuning)	49
Tabel 13. Tabel Hasil Pengujian Fine-Tuned model dengan Data Uji (Test)	52
Tabel 14. Skenario Pengujian Black Box	55
Tabel 15. Parameter Pengujian oleh Ahli	57
Tabel 16. Pertanyaan Pengujian SUS	58
Tabel 17. Pernyataan NPS	59
Tabel 18. Hasil Uji Black Box	60
Tabel 19. Hasil Pengujian Oleh Ahli Bahasa	62
Tabel 20. Hasil SUS	63
Tabel 21. Nilai SUS Tiap Responden	65
Tabel 22. Perhitungan Nilai NPS	68

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>Transformer</i>	12
Gambar 2.2 Diagram <i>Framework</i> T5	13
Gambar 2.3 Proses Pelatihan T5 pada Sisi Encoder	15
Gambar 2.4 Proses Pelatihan T5 pada Sisi Decoder	17
Gambar 2.5 Proses Update Bobot dan Perhitungan Loss Setelah Blok Decoder .	18
Gambar 3.1. Diagram Alur Penelitian.....	23
Gambar 4.1 Use Case Diagram Web	38
Gambar 4.2 <i>Activity</i> Diagram Web	39
Gambar 4.3 <i>Wireframe</i> Halaman Utama.....	41
Gambar 4.4 Grafik <i>Epoch-Loss Train</i> Percobaan Pertama.....	52
Gambar 4.5 Halaman Utama Web	54





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Ulasan konsumen pada platform *e-commerce* telah menjadi sumber informasi utama dalam menilai kualitas produk dan layanan (Wahdiniawati *et al.*, 2024; Pramudita, 2024). Berbagai penelitian menunjukkan bahwa ulasan memiliki pengaruh signifikan terhadap minat dan keputusan pembelian, di mana sekitar 76% responden dilaporkan terlebih dahulu membaca ulasan sebelum melakukan transaksi (Hildayanti *et al.*, 2024; Nurbaiti *et al.*, 2025). Namun, jumlah ulasan yang sangat besar dengan variasi gaya bahasa, struktur, dan panjang teks sering menyulitkan pengguna untuk memperoleh gambaran umum mengenai opini konsumen secara cepat (Putri *et al.*, 2024). Kondisi tersebut mendorong perlunya teknologi peringkasan ulasan agar informasi dapat disajikan secara lebih ringkas dan informatif.

Dalam peringkasan teks, terdapat dua pendekatan utama yang dapat digunakan, yaitu pendekatan ekstraktif dan pendekatan abstraktif. Kedua metode tersebut memiliki karakteristik dan mekanisme yang berbeda dalam menghasilkan ringkasan, sehingga pemilihannya perlu disesuaikan dengan sifat dan kebutuhan data ulasan *e-commerce*. (Yuliska *et al.*, 2020)

Pendekatan ekstraktif pada dasarnya masih memadai untuk teks yang bersifat faktual dan terstruktur seperti berita atau laporan resmi, karena informasi penting biasanya sudah tersaji dalam kalimat yang jelas dan runtut. Akan tetapi, ulasan konsumen cenderung tidak terstruktur, beragam, dan penuh dengan gaya bahasa informal, sehingga pendekatan ekstraktif sering kali kurang mampu menangkap makna secara utuh karena hanya memilih potongan kalimat tanpa membangun kembali konteks informasi (Yuliska *et al.*, 2020). Sementara itu, pendekatan abstraktif lebih mampu menyatukan inti informasi dari berbagai ulasan ke dalam narasi baru yang ringkas dan alami, menyerupai cara manusia merangkum sebuah teks (Yuliska *et al.*, 2020). Oleh karena itu, pendekatan abstraktif dipandang lebih



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

sesuai untuk merepresentasikan pengalaman serta penilaian konsumen pada ulasan *e-commerce*.

Meskipun pendekatan abstraktif lebih relevan, penelitian terdahulu menunjukkan bahwa tidak semua model neural mampu menghasilkan ringkasan yang optimal. Model berbasis LSTM memiliki keterbatasan dalam memahami dependensi panjang dan rentan mengalami permasalahan *catastrophic forgetting* (Bahari *et al.*, 2024). Pendekatan LSTM yang dikombinasikan dengan *self-attention* juga masih menunjukkan performa yang lebih rendah dibandingkan model *transformer* (Rosaria *et al.*, 2024).

Selain itu, model BERT2BERT yang dimodifikasi menyerupai arsitektur *encoder-decoder* seperti T5 memang mampu melakukan peringkasan abstraktif, namun kinerjanya masih berada di bawah pendekatan yang menggunakan arsitektur T5 secara langsung (Muharam *et al.*, 2025). Berbagai temuan tersebut menunjukkan bahwa arsitektur T5, khususnya IndoT5, terbukti lebih unggul dan mampu menghasilkan ringkasan yang lebih relevan serta mudah dipahami (Bahari *et al.*, 2024).

Oleh karena itu, penelitian ini membutuhkan arsitektur *transformer* seperti IndoT5 sebagai solusi dalam mengembangkan sistem peringkasan ulasan secara abstraktif. Pemanfaatan IndoT5 diharapkan mampu menghasilkan ringkasan yang ringkas, informatif, dan merepresentasikan makna utama dari berbagai ulasan secara koheren. Melalui sistem peringkasan yang dikembangkan, pengguna diharapkan dapat memahami gambaran umum opini konsumen dengan lebih cepat tanpa harus membaca seluruh ulasan satu per satu, sehingga proses pengambilan keputusan pembelian dapat dilakukan secara lebih efektif.

1.2 Rumusan Masalah

Berdasarkan Latar Belakang yang telah diuraikan diatas, maka rumusan masalah dalam penelitian ini adalah Bagaimana penerapan model Indonesia *Text-to-Text Transfer Transformer* (IndoT5) untuk peringkasan ulasan otomatis pada ulasan *e-commerce* berbahasa Indonesia.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

1.3 Batasan Masalah

Dari rumusan masalah yang telah ditetapkan, maka batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Implementasi model dilakukan menggunakan *framework* PyTorch dan *library* Hugging Face *Transformer*.
2. Metode pelatihan model dilakukan dengan teknik *fine-tuning* terhadap *pre-trained* model T5.
3. *Pre-trained* model yang digunakan difokuskan pada model T5 yang mendukung bahasa Indonesia.
4. Data yang digunakan dalam penelitian ini adalah kumpulan ulasan produk *e-commerce* Tokopedia berbahasa Indonesia, yang diperoleh dari sumber *open-source* seperti Kaggle.
5. ringkasan ulasan pada *dataset* dihasilkan menggunakan LLM GPT-4.
6. Penelitian ini hanya berfokus pada peringkasan teks abstraktif, dan tidak membahas metode peringkasan ekstraktif.
7. Evaluasi kinerja model peringkasan dilakukan menggunakan metrik evaluasi peringkasan teks yang relevan, seperti ROUGE-L, BERTScore, dan pengujian manual.
8. Penelitian ini tidak membahas analisis sentimen, klasifikasi ulasan, maupun sistem rekomendasi produk.
9. Sistem dan model yang dikembangkan dalam penelitian ini hanya untuk peringkasan ulasan berbahasa Indonesia.

1.4 Tujuan dan Manfaat

Tujuan dan manfaat dari implementasi sistem peringkasan ulasan otomatis berbasis web menggunakan model IndoT5 untuk ulasan produk *e-commerce* berbahasa Indonesia adalah sebagai berikut

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah mengimplementasikan model IndoT5 dalam sistem peringkasan ulasan *e-commerce* berbahasa Indonesia berbasis web serta mengevaluasi kinerja model dan sistem yang dibangun.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

1.4.2 Manfaat Penelitian

Adapun manfaatnya yang diharapkan dari penelitian ini adalah sebagai berikut :

1. Membantu pengguna memahami gambaran umum opini konsumen secara lebih cepat tanpa perlu membaca seluruh ulasan satu per satu.
2. Mendukung pengguna dalam proses pengambilan keputusan pembelian secara lebih efektif dan efisien.

1.5 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini terdiri dari 5 bab, yaitu :

1. BAB I PENDAHULUAN

Bab I berisikan latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat, serta sistematika penulisan yang akan menjadi landasan utama dalam peringkasan teks abstraktif otomatis berbasis *transformer* untuk kumpulan ulasan produk *e-commerce* berbahasa Indonesia.

2. BAB II TINJAUAN PUSTAKA

Bab II memuat tinjauan pustaka dan landasan teori yang menjadi dasar pembahasan dalam penelitian ini. Pada bab ini juga disajikan kajian terhadap penelitian-penelitian terdahulu yang relevan, guna memberikan gambaran umum serta memperkuat landasan ilmiah terkait topik yang diangkat dalam penelitian ini.

3. BAB III METODOLOGI PENELITIAN

Bab III membahas perencanaan dan perancangan sistem yang akan dikembangkan dalam penelitian ini. Pada bab ini dijelaskan rancangan penelitian, tahapan pelaksanaan penelitian, objek penelitian, serta model, *framework*, dan teknik yang digunakan. Selain itu, bab ini juga menguraikan teknik pengumpulan dan analisis data, jadwal pelaksanaan penelitian, serta rincian biaya yang direncanakan untuk mendukung pelaksanaan penelitian ke depannya.

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

4. BAB IV PEMBAHASAN

Bab IV membahas hasil pengembangan dan pengujian model dari penelitian yang telah dilakukan. Pada bab ini, seluruh tahapan pengujian dijelaskan secara sistematis, mulai dari deskripsi pengujian dan skenario uji model hingga pemaparan hasil yang diperoleh. Selanjutnya, data hasil pengujian model dianalisis dan dievaluasi untuk menilai kinerja model yang dikembangkan. Melalui pembahasan tersebut, diharapkan dapat diperoleh gambaran yang jelas mengenai tingkat efektivitas metode yang diimplementasikan dalam sistem.

5. BAB V PENUTUP

Bab V menyajikan rangkuman keseluruhan dari penelitian yang telah dilakukan. Bab ini memuat kesimpulan yang diperoleh berdasarkan hasil penelitian, serta saran yang dapat dijadikan acuan untuk pengembangan atau penelitian selanjutnya



**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB 2

TINJAUAN PUSTAKA

2.1 Aplikasi Peringkasan Teks

Aplikasi peringkasan teks (*text summarization*) merupakan sistem berbasis komputer yang dirancang untuk menghasilkan ringkasan teks secara otomatis dengan memanfaatkan algoritma tertentu. Sistem ini bekerja dengan menyederhanakan dokumen asli menjadi bentuk yang lebih ringkas tanpa menghilangkan informasi utama atau gagasan penting yang terkandung di dalamnya. Dalam prosesnya, teks dimasukkan sebagai *input*, kemudian diolah oleh sistem peringkasan untuk menghasilkan keluaran berupa teks yang lebih singkat namun tetap mewakili isi dokumen secara keseluruhan. Seiring dengan perkembangan teknologi, aplikasi peringkasan teks telah banyak diterapkan untuk merangkum berbagai jenis dokumen, seperti artikel, laporan, dan ulasan, serta dinilai mampu memberikan solusi yang lebih cepat dan efisien dibandingkan peringkasan yang dilakukan secara manual. (Karo *et al.*, 2024).

2.1.1 Ulasan

Menurut Kamus Besar Bahasa Indonesia (KBBI), kata ulasan secara umum sebagai nomina (kata benda) berarti kupasan, tafsiran, atau komentar terhadap suatu hal, yang menunjukkan penjelasan atau pendapat tentang sesuatu berdasarkan pengamatan atau analisis. (KBBI)

2.1.2 Peringkasan Teks

Sumarisasi teks adalah proses menghasilkan ringkasan dari dokumen panjang dengan tetap mempertahankan informasi penting dan makna utama sehingga penyajian informasi menjadi lebih singkat, padat, dan mudah dipahami. Secara umum, peringkasan teks terbagi menjadi dua pendekatan, yaitu ekstraktif yang memilih bagian penting langsung dari teks sumber, dan abstraktif yang menyusun ulang informasi untuk menghasilkan kalimat baru yang merepresentasikan isi teks. Berbagai penelitian menunjukkan bahwa pendekatan abstraktif berbasis model *transformer* memiliki keunggulan dibandingkan pendekatan ekstraktif karena



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

mampu menghasilkan ringkasan yang lebih koheren, kontekstual, dan menyerupai ringkasan buatan manusia (Nasution *et al.*, 2025).

2.1.3 Peringkasan Ulasan Produk

Peringkasan ulasan produk merupakan salah satu bentuk *multi-document summarization* yang bertujuan merangkum berbagai ulasan pengguna pada sebuah toko daring terhadap produk tertentu. Peringkasan ulasan produk ini digunakan untuk memperoleh gambaran dan pendapat pengguna mengenai suatu produk secara cepat dan akurat (Yuliska *et al.*, 2020).

2.2 E-Commerce

E-commerce atau perdagangan elektronik merupakan kegiatan jual beli barang dan jasa yang dilakukan melalui jaringan internet, di mana seluruh proses transaksi, mulai dari penyampaian informasi, pemesanan, hingga pembayaran, berlangsung secara digital tanpa adanya pertemuan langsung antara penjual dan pembeli. Melalui sistem ini, pertukaran informasi dan transaksi keuangan dapat dilakukan secara elektronik, sehingga hubungan bisnis antara pelaku usaha dan konsumen dapat berjalan lebih efisien (Efawati *et al.*, 2024).

2.3 Software Development Life Cycle (SDLC)

Software Development Life Cycle (SDLC) merupakan metode pengembangan perangkat lunak yang mencakup tahapan analisis, perancangan, implementasi, pengujian, hingga pemeliharaan, yang diterapkan secara sistematis sejak tahap inisiasi hingga perawatan sistem (Jamal *et al.*, 2022). Salah satu model SDLC yang paling umum digunakan dalam pengembangan perangkat lunak adalah model *Waterfall*.

2.3.1 Tahap Analisis

Tahapan analisis merupakan fase dalam SDLC yang bertujuan untuk mengidentifikasi dan memahami kebutuhan pengguna serta permasalahan yang akan diselesaikan oleh sistem. Pada tahap ini dilakukan pengumpulan dan pengolahan informasi terkait proses bisnis, kebutuhan fungsional, dan kebutuhan non fungsional agar sistem yang dikembangkan sesuai dengan tujuan dan harapan



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

pengguna. Hasil dari tahap analisis menjadi dasar dalam perancangan sistem sehingga dapat meminimalkan kesalahan pada tahap pengembangan selanjutnya (Setiyani *et al.*, 2021)

2.3.2 Tahap Perancangan

Tahapan perancangan (*design*) merupakan tahap dalam metode *Waterfall* yang bertujuan untuk memodelkan sistem sebelum proses implementasi dilakukan. Pada tahap ini, kebutuhan yang telah dianalisis diterjemahkan ke dalam bentuk desain sistem menggunakan beberapa diagram, seperti *activity* diagram, *use case* diagram, *sequence* diagram, dan *class* diagram, untuk menggambarkan alur proses, interaksi aktor dengan sistem, serta struktur dan relasi antar komponen sistem. Tahap perancangan ini berperan penting dalam memastikan pengembangan sistem berjalan secara terarah dan sesuai dengan kebutuhan yang telah ditetapkan (Hafidz *et al.*, 2022; Ashari *et al.*, 2023; Ramdany, 2024).

2.3.3 Tahap Implementasi

Tahapan implementasi merupakan fase dalam metode *Waterfall* yang berfokus pada penerapan hasil perancangan sistem ke dalam bentuk perangkat lunak yang dapat dijalankan. Pada tahap ini, sistem dikembangkan sesuai dengan desain yang telah dirancang, umumnya dalam bentuk aplikasi berbasis web dengan memanfaatkan *framework* pendukung seperti FastAPI dan TailwindCSS. Selain itu, pengujian fungsional juga dilakukan menggunakan metode black box testing untuk memastikan setiap fungsi sistem berjalan sesuai dengan kebutuhan dan menghasilkan keluaran yang tepat (Setiawansyah *et al.*, 2022; Aipina *et al.*, 2022; Eriya, 2020).

2.3.4 Tahap Pengujian

Tahapan pengujian merupakan tahap dalam pengembangan perangkat lunak yang bertujuan untuk memastikan bahwa sistem yang dibangun telah berjalan sesuai dengan kebutuhan dan fungsi yang diharapkan. Pengujian dilakukan dengan menjalankan skenario tertentu guna menemukan kesalahan atau kecacatan pada sistem serta menilai kesesuaian respons yang diberikan. Salah satu metode yang umum digunakan adalah black box testing, yaitu pengujian yang berfokus pada



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

fungsi sistem tanpa melihat struktur internal kode, sehingga dapat diketahui apakah sistem telah memenuhi spesifikasi yang ditetapkan atau masih memerlukan perbaikan (Eriya, 2020).

1. *System Usability Scale (SUS)*

System Usability Scale (SUS) adalah metode pengukuran yang digunakan untuk menilai tingkat *usability* (kegunaan/penggunaan) suatu sistem atau aplikasi dari sudut pandang pengguna akhir. Instrumen ini berupa kuesioner standar yang terdiri atas 10 pernyataan dengan skala Likert 5 poin, yang kemudian dihitung untuk menghasilkan skor akhir dalam rentang 0–100 yang menggambarkan persepsi pengguna terhadap kegunaan suatu sistem (Sukma *et al.*, 2023).

2. *Net Promoter Score (NPS)*

Net Promoter Score (NPS) adalah metrik atau indeks yang digunakan untuk mengukur loyalitas dan kepuasan pelanggan terhadap suatu produk, layanan, atau organisasi berdasarkan kecenderungan pelanggan untuk merekomendasikannya kepada orang lain. Pengukuran ini dilakukan dengan menggunakan satu pertanyaan utama, yakni “Seberapa besar kemungkinan Anda merekomendasikan produk/layanan/sistem ini kepada teman atau kolega?”, yang dijawab pada skala 0–10. Selanjutnya responden diklasifikasikan menjadi tiga kategori: *promoter* (nilai 9–10), *passive* (nilai 7–8), dan *detractor* (nilai 0–6). Skor NPS dihitung dengan mengurangi persentase *detractor* dari persentase *promoter*. (Charisma *et al.*, 2025).

2.3.5 *Waterfall*

Metode *Waterfall* merupakan salah satu model *Software Development Life Cycle* yang banyak digunakan karena proses pengembangannya dilakukan secara berurutan dan sistematis. Setiap tahapan SDLC harus diselesaikan terlebih dahulu sebelum melanjutkan ke tahap berikutnya, sehingga pengembangan perangkat lunak berjalan secara terstruktur dan terkontrol (Risald, 2021).



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2.4 Cross Industry Standard Process – Data Mining (CRISP-DM)

Metode CRISP-DM merupakan kerangka kerja yang digunakan sebagai standar dalam proses data *mining*, sekaligus berfungsi sebagai panduan dalam merencanakan strategi analisis serta pemecahan masalah yang diterapkan pada konteks bisnis maupun penelitian (Wibisono *et al.*, 2022).

2.4.1 Business Understanding

Business *Understanding* merupakan tahap awal dalam CRISP-DM yang berfokus pada penentuan target dan tujuan bisnis. Pada tahap ini juga disusun rencana strategis untuk mencapai tujuan tersebut, serta dilakukan pemahaman terhadap kondisi dan permasalahan bisnis yang akan dianalisis (Wibisono *et al.*, 2022)

2.4.2 Data Understanding

Tahap Data *Understanding* merupakan proses menghimpun data yang kemudian dianalisis untuk memahami karakteristik dan pola yang terkandung di dalamnya, sehingga pada tahap ini dihasilkan pengetahuan yang menjadi dasar penggunaan data dalam konteks penelitian maupun bisnis (Wibisono *et al.*, 2022).

2.4.3 Data Preparation

Preprocessing merupakan tahapan yang dilakukan sebelum proses pelatihan model dengan tujuan untuk membersihkan data, menyesuaikan format, serta mempersiapkan data agar siap digunakan (Agung *et al.*, 2023). Tahapan ini diperlukan untuk mengatasi ketidaksesuaian pada data awal sehingga data menjadi lebih seragam, yang pada akhirnya membantu algoritma sistem cerdas menghasilkan keluaran yang lebih akurat.

2.4.4 Modelling

Tahap *modelling* dalam metodologi CRISP-DM merupakan proses pemilihan dan penerapan teknik pemodelan yang paling tepat terhadap data yang telah dipersiapkan sebelumnya. Pada fase ini, peneliti atau analis menetapkan algoritma yang akan digunakan, membangun model menggunakan data pelatihan, serta melakukan penyesuaian parameter agar model yang dihasilkan selaras dengan tujuan bisnis yang ingin dicapai. Model tersebut dikembangkan untuk



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

mengidentifikasi pola atau hubungan dalam data sesuai dengan kebutuhan analisis atau prediksi, dan selanjutnya digunakan sebagai dasar pada tahap evaluasi guna menilai kinerja serta kelayakan model dalam menyelesaikan permasalahan yang dikaji. (Pratama *et al.*, 2023)

1. Python

Python merupakan bahasa pemrograman tingkat tinggi yang populer dan bersifat multifungsi, sehingga banyak digunakan oleh pengembang perangkat lunak hingga pengembangan model *machine learning* dan *deep learning*. Perkembangan penggunaan Python dalam bidang *machine learning* dan *deep learning* berlangsung sangat pesat, terutama karena ketersediaan berbagai pustaka yang mendukung proses pengembangan model sesuai kebutuhan pengembang, sehingga mampu menghasilkan performa dan akurasi yang tinggi (Alfarizi *et al.*, 2023).

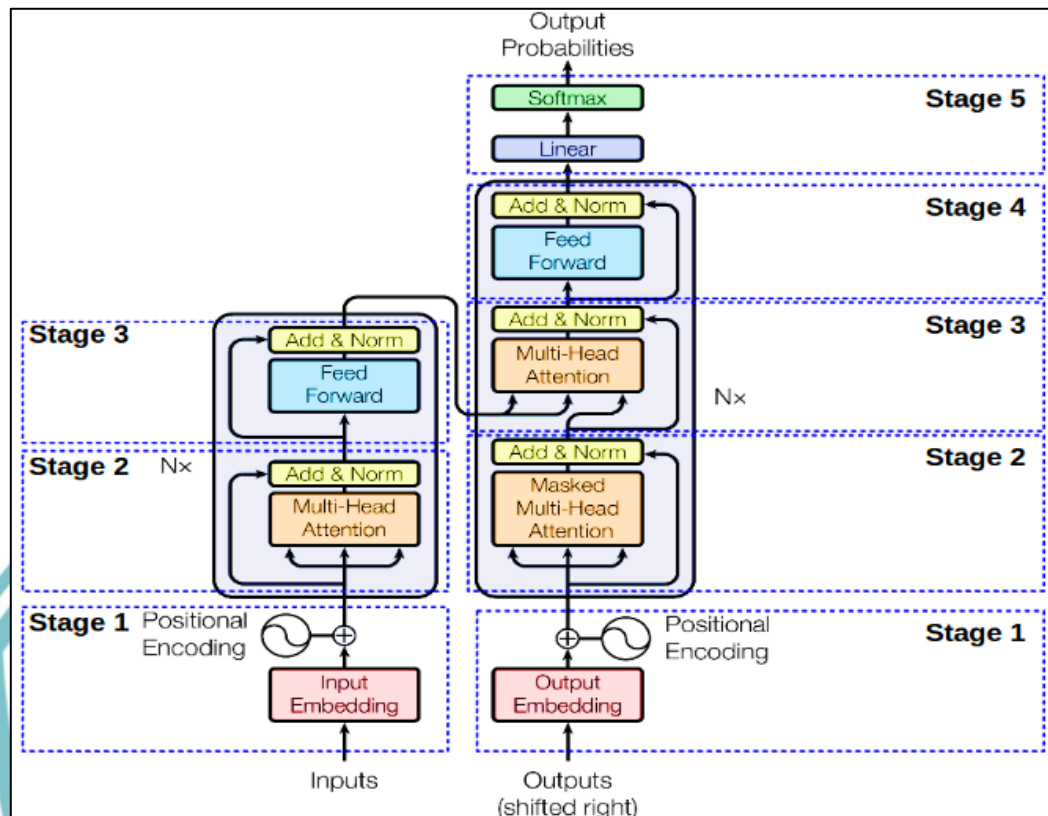
2. PyTorch

PyTorch adalah sebuah *open-source deep learning framework* yang menyediakan pustaka *tensor* teroptimasi untuk pembelajaran mendalam (*deep learning*) yang dapat dijalankan di CPU maupun GPU. *Framework* ini dirancang untuk mempercepat proses dari *research prototyping* hingga *production deployment* dengan mendukung graf komputasi dinamis dan integrasi yang kuat dengan Python, sehingga memudahkan pembangunan dan pelatihan model *neural network* yang kompleks. (PyTorch, 2024)

3. *Transformer*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Gambar 2.1 Arsitektur *Transformer*

Sumber : Vaswani, 2017

transformer adalah model *deep learning sequence-to-sequence* berbasis mekanisme *attention* yang dirancang tanpa komponen rekuren maupun konvolusional untuk mengatasi keterbatasan efisiensi pemrosesan urutan panjang serta kesulitan dalam menangkap dependensi jangka panjang. Model ini memproses seluruh token secara paralel, dimulai dengan mengubah *input* menjadi representasi vektor melalui *embedding* yang diperkaya dengan *positional encoding* sinusoidal agar informasi urutan tetap terjaga. Representasi tersebut kemudian diproses oleh *encoder* yang tersusun atas beberapa lapisan identik, di mana setiap lapisan mengombinasikan *multi-head self-attention* untuk menangkap konteks global dari berbagai sudut pandang dan *feed-forward network* untuk memperkuat representasi token, *Transformer* juga dilengkapi dengan *residual connection* dan *layer normalization* guna menjaga stabilitas pelatihan. Selanjutnya, *decoder*

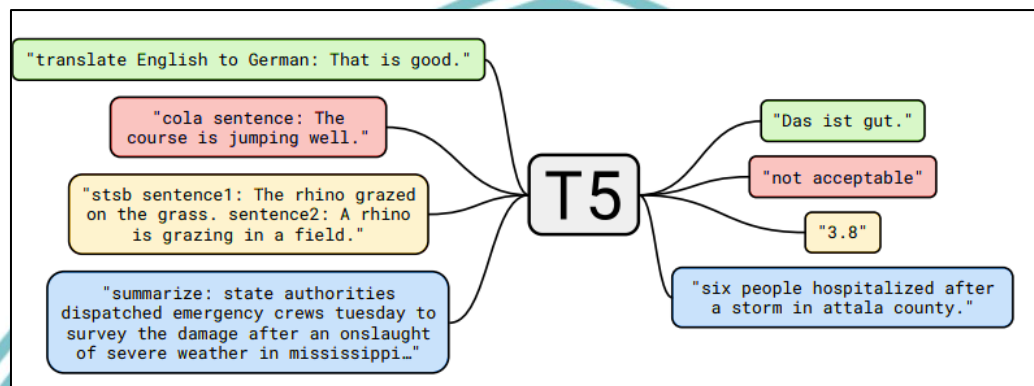


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

menghasilkan distribusi probabilitas token melalui fungsi *softmax* untuk memprediksi kata berikutnya, sehingga *Transformer* mampu memahami konteks secara menyeluruh dan menghasilkan keluaran yang akurat serta efisien. (Vaswani *et al.*, 2017)

4. *Text-to-Text Transfer Transformer (T5)*



Gambar 2.2 Diagram *Framework T5*

Sumber : Raffel, 2020

T5 merupakan pengembangan arsitektur *Transformer* yang memformulasikan seluruh tugas NLP, seperti peringkasan, penerjemahan, klasifikasi, dan tanya jawab, ke dalam satu kerangka *text-to-text*, sehingga satu model yang sama dapat digunakan untuk berbagai tugas hanya dengan menambahkan prefiks pada masukan tanpa memerlukan arsitektur khusus per tugas. Model ini memanfaatkan *relative position* bias untuk menangkap hubungan antar kata berdasarkan jarak relatifnya serta dilatih menggunakan skema span *corruption* pada korpus berskala besar C4, yang mendukung pemahaman konteks secara menyeluruh dan menghasilkan teks yang lebih koheren, khususnya pada tugas generatif. Berangkat dari pendekatan tersebut, IndoT5 dikembangkan sebagai adaptasi T5 untuk Bahasa Indonesia dengan memanfaatkan korpus C4 berbahasa Indonesia, sehingga mampu memahami struktur dan konteks kebahasaan secara lebih baik serta menghasilkan keluaran yang relevan, koheren, dan selaras secara semantik dengan karakteristik Bahasa Indonesia (Utami *et al.*, 2023; Puspitasari *et al.*, 2025; Raffel *et al.*, 2020).



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

5. Fine Tuning Model

Fine-tuning adalah proses penyesuaian model pra-latih (*pre-trained model*) dengan melatihnya kembali pada *dataset* yang lebih spesifik dan relevan untuk tugas tertentu. Dalam proses ini, bobot atau parameter model yang sudah dilatih sebelumnya pada data umum disesuaikan agar lebih optimal dalam menyelesaikan tugas lanjutan, seperti klasifikasi, analisis sentimen, atau peringkasan teks yang sesuai dengan domain atau konteks penelitian spesifik. Dengan *fine-tuning*, model yang awalnya bersifat umum dapat mengkhususkan diri pada tugas yang ditargetkan dengan *dataset* yang relatif lebih kecil, sehingga meningkatkan performa dan relevansi keluaran model untuk tugas tersebut (Nurinayah *et al*, 2025).

Secara teknis, proses pelatihan (*fine tuning*) IndoT5 diawali dengan mengubah teks masukan menjadi representasi diskret menggunakan tokenizer berbasis SentencePiece. Pada tahap ini, teks mentah dikonversi menjadi urutan token ID yang kemudian dipetakan ke dalam ruang vektor berdimensi tinggi melalui lapisan token embedding. Hasil pemetaan tersebut menghasilkan representasi vektor dengan 768 dimensi. Sebelum diproses lebih lanjut oleh encoder, representasi vektor tersebut ditambahkan dengan relative position bias untuk menangkap informasi jarak relatif antar token tanpa bergantung pada posisi absolut (Raffel *et al.*, 2020).

Vektor Representatif yang telah memuat informasi posisi tersebut kemudian diproses oleh blok encoder yang tersusun atas mekanisme Multi-Head Self-Attention dan jaringan Feed-Forward. Pada mekanisme Self-Attention, setiap token ditransformasikan menjadi tiga vektor, yaitu Query (Q), Key (K), dan Value (V), melalui perkalian dengan matriks bobot W^Q , W^K , dan W^V . Selanjutnya, hubungan antar token dihitung menggunakan mekanisme attention sebagaimana dirumuskan oleh Vaswani *et al.* (2017)

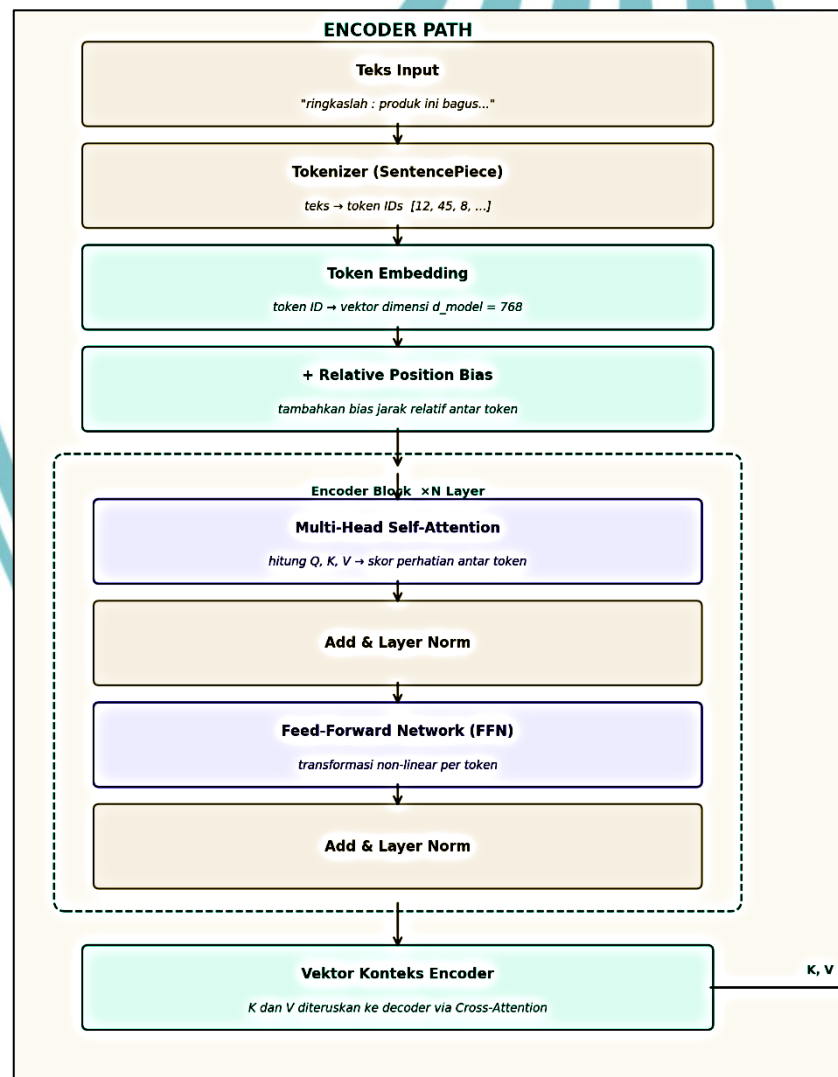
$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

output dari setiap sub-layer kemudian melewati proses Add & Layer Normalization sebelum diteruskan ke sub-layer berikutnya. Pada tahap itu, vektor hasil proses digabungkan kembali dengan vektor aslinya agar model tidak kehilangan informasi awal, lalu rentang nilainya diratakan (dinormalisasi) agar skalanya tetap stabil dan tidak "meledak" saat masuk ke lapisan berikutnya. Setelah seluruh lapisan encoder selesai diproses, dihasilkan vektor konteks yang merepresentasikan pemahaman model terhadap keseluruhan teks masukan. Vektor konteks ini selanjutnya diteruskan ke decoder dalam bentuk pasangan matriks K dan V yang akan digunakan sebagai input context injection.



Gambar 2.3 Proses Pelatihan T5 pada Sisi Encoder



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

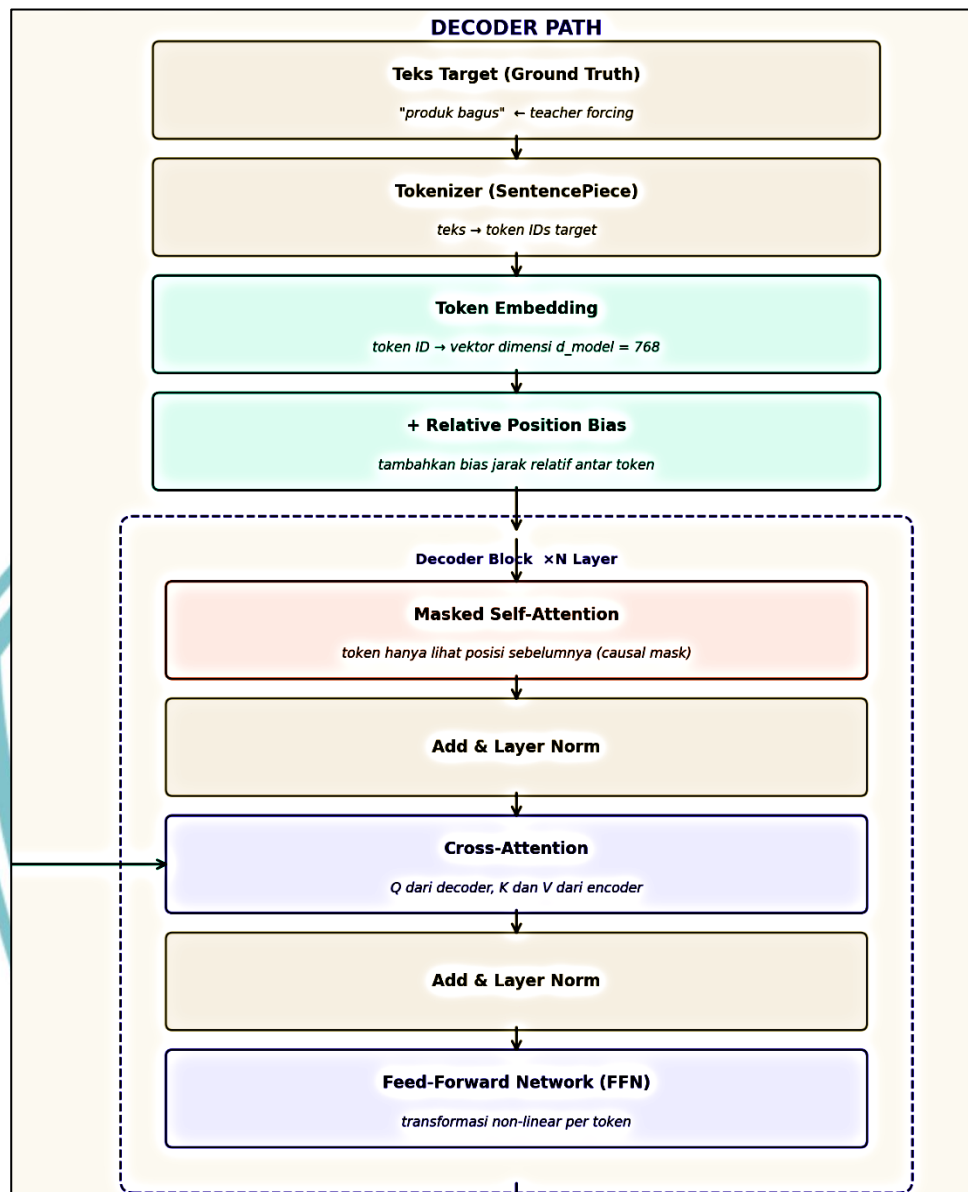
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Di sisi decoder, teks target (ground truth) juga melalui tahapan tokenisasi, token embedding, dan penambahan relative position bias sebelum masuk ke blok decoder. Proses pada decoder diawali dengan mekanisme Masked Self-Attention, di mana setiap token hanya dapat memperhatikan token-token sebelumnya dan tidak memiliki akses terhadap token pada posisi berikutnya. Pembatasan ini diterapkan menggunakan causal mask berbentuk matriks segitiga bawah yang mengisi skor perhatian pada posisi mendatang dengan nilai ∞ sebelum operasi softmax dilakukan. Dengan demikian, probabilitas untuk melihat token selanjutnya menjadi nol ketika dihitung menggunakan fungsi aktivasi softmax. Mekanisme tersebut mendukung penerapan teacher forcing, yaitu strategi pelatihan yang menggunakan token target sebenarnya sebagai masukan decoder pada setiap langkah, bukan token hasil prediksi model sebelumnya (Raffel et al., 2020)

Setelah melalui tahap Masked Self-Attention dan Add & Layer Normalization, representasi decoder diperkaya melalui mekanisme Cross-Attention. Pada tahap ini, Query (Q) berasal dari decoder, sedangkan Key (K) dan Value (V) berasal dari encoder. Kombinasi tersebut memungkinkan model menggunakan informasi (Context) dari teks input dalam proses generate output token. Hasil Cross-Attention kemudian diproses oleh jaringan Feed-Forward dan kembali melewati proses Add & Layer Normalization sebelum menghasilkan vektor akhir dari decoder

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Gambar 2.4 Proses Pelatihan T5 pada Sisi Decoder

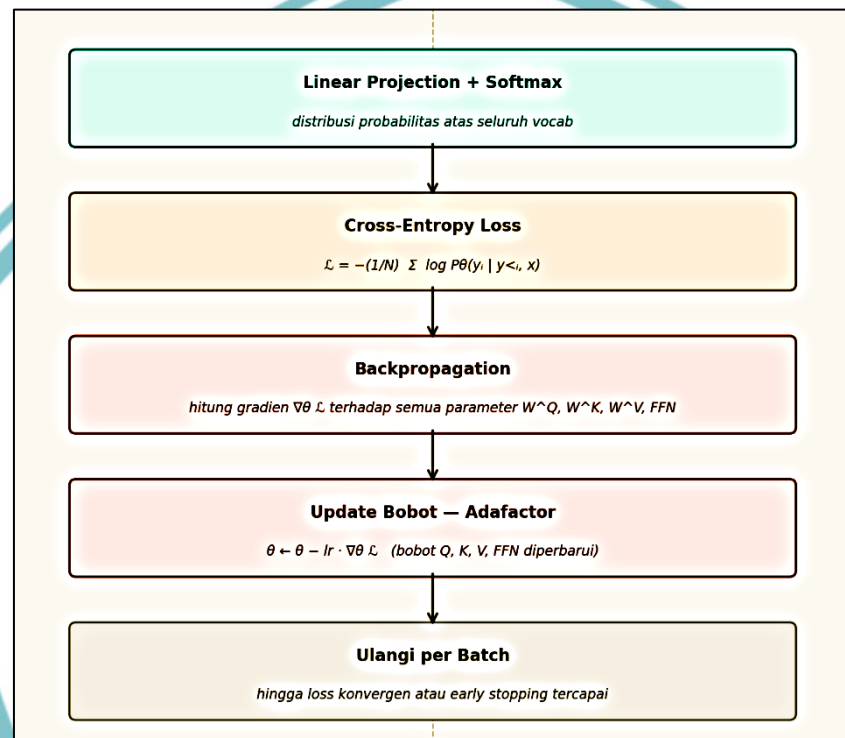
Seluruh rangkaian forward pass tersebut menghasilkan distribusi probabilitas terhadap seluruh kosakata pada setiap posisi token melalui lapisan linear projection dan fungsi softmax. Distribusi probabilitas yang dihasilkan kemudian dibandingkan dengan token target menggunakan fungsi Cross-Entropy Loss yang dihitung secara rata-rata pada setiap token. Nilai loss tersebut digunakan untuk menghitung gradien seluruh parameter model, termasuk matriks W^Q , W^K , dan W^V , serta bobot pada Feed-Forward Network (FFN), melalui algoritma backpropagation.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Selanjutnya, parameter (weight) model diperbarui menggunakan algoritma optimasi Adafactor (Raffel et al., 2020). Proses ini dilakukan secara berulang pada setiap batch data pelatihan hingga model mencapai kondisi konvergen, yang umumnya ditandai oleh penurunan nilai loss secara konsisten atau terpenuhinya kriteria early stopping.



Gambar 2.5 Proses Update Bobot dan Perhitungan Loss Setelah Blok Decoder

2.4.5 Evaluasi Model

Proses evaluasi adalah tahap di mana model yang telah dikembangkan diuji dan dinilai untuk memastikan bahwa performanya sesuai dengan standar yang telah ditetapkan pada tahap business *understanding*, sehingga dapat dipastikan model tersebut memenuhi kebutuhan bisnis

1. ROUGE-L

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) merupakan metrik evaluasi untuk menilai kualitas ringkasan otomatis dengan membandingkan kesamaan urutan kata (sequence overlap) antara ringkasan



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

hasil model dan ringkasan referensi. ROUGE-1 digunakan untuk mengukur kecocokan kata (*unigram*), ROUGE-2 untuk mengukur kecocokan dua kata berturutan (*bigram*), sedangkan ROUGE-L mengukur kesamaan berdasarkan *Longest Common Subsequence* (LCS) yang merefleksikan kesesuaian urutan kata. Nilai ROUGE yang lebih tinggi menunjukkan kualitas ringkasan yang lebih baik. Nilai rouge berada dalam rentang 0 -1, semakin mendekati 1 maka semakin baik (Bahari *et al.*, 2024).

Perhitungan ROUGE-L dapat dilihat melalui persamaan berikut,

$$\text{ROUGE-L} = \frac{\text{LCS}(X, Y)}{|Y|}$$

Dimana,

$\text{LCS}(X, Y)$ = panjang *Longest Common Subsequence* antara ringkasan model X dan referensi Y

$|Y|$ = jumlah kata pada ringkasan referensi

2. BERTScore

BERTScore merupakan metrik evaluasi untuk tugas *text generation* yang mengukur kesamaan semantik antara teks hasil model dan teks referensi dengan memanfaatkan *contextual embeddings* dari model BERT. Berbeda dengan metrik berbasis n-gram seperti ROUGE, BERTScore membandingkan representasi vektor setiap token menggunakan *cosine similarity*, sehingga mampu menangkap kesamaan makna meskipun terdapat perbedaan kata atau susunan kalimat. Proses perhitungannya dilakukan dengan mencocokkan setiap token pada teks hasil dengan token yang paling mirip pada teks referensi (*greedy matching*), lalu dihitung nilai *precision*, *recall*, dan F1-score sebagai ukuran kesamaan keseluruhan dimana Nilai BERTScore-F1 berada pada rentang 0–1, semakin mendekati 1 maka semakin baik. (Zhang *et al.*, 2020)



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Perhitungan BERTScore *precision* dapat dilihat pada persamaan berikut ,

$$P = \frac{1}{|X|} \sum_{x \in X} \max_{y \in Y} \cos(e_x, e_y)$$

Sementara BERTScore Recall,

$$R = \frac{1}{|Y|} \sum_{y \in Y} \max_{x \in X} \cos(e_y, e_x)$$

Dengan penurunan rumus F1-Score, maka didapat,

$$\text{BERTScore}_{F1} = \frac{2PR}{P + R}$$

Dimana,

X = token pada teks hasil model

Y = token pada teks referensi

e_x, e_y = embedding token dari BERT

$\cos(\cdot)$ = cosine similarity

2.4.6 Deployment

1. Fast-API

FastAPI adalah *framework* web Python modern berbasis Python *type hints* yang dirancang untuk membangun API dengan performa tinggi, mudah digunakan, cepat untuk dikembangkan, dan siap untuk produksi. FastAPI memanfaatkan fitur *type hints* Python standar untuk validasi data otomatis, dan secara otomatis menghasilkan dokumentasi API interaktif sesuai dengan spesifikasi OpenAPI (FastAPI, 2025).

2. Aplikasi Web

Web *application* adalah program yang dapat diakses melalui web (browser) yang menghasilkan konten secara dinamis berdasarkan interaksi pengguna atau data kontekstual, menggunakan teknologi web seperti HTML, CSS, dan JS (Li *et al*, 2024).



2.5 Penelitian Terkait

Tabel 1. Penelitian Terkait

Judul	Peringkasan teks otomatis abstraktif menggunakan <i>Transformer</i> pada teks bahasa Indonesia (Bahari <i>et al.</i> , 2024)	Peringkasan Teks Berita Berbahasa Indonesia Menggunakan LSTM dan <i>Transformer</i> (Rosaria <i>et al.</i> , 2024)	<i>Enhancing Abstractive Multi-Documents Summarization with Bert2Bert Model for Indonesian Language</i> (Muharam <i>et al.</i> , 2025)
Tujuan	mengukur seberapa baik performa <i>Transformer</i> pada peringkasan teks abstraktif dalam bahasa Indonesia.	Membandingkan performa LSTM+ <i>Self Attention Mechanism</i> dengan model <i>transformers</i> pada tugas peringkasan teks abstraktif dalam bahasa Indonesia.	Untuk menguji efektivitas model Bert2Bert dalam meningkatkan kualitas peringkasan <i>multi</i> -dokumen abstraktif untuk bahasa Indonesia.
Metode	Fine Tuning <i>Pre-train</i> T5 dan IndoT5 model dan mengevaluasi model dengan ROUGE 1 & 2	Fine Tuning model LSTM+ <i>Self Attn</i> , model BART <i>small</i> dan T5 <i>Small</i> . Dan mengevaluasi model dengan ROUGE 1,2 dan L	Fine <i>tuning</i> model BERT2BERT (<i>modified bertDecoder</i>) dan mengevaluasi dengan R1,R2,RL dan BERTScore
Hasil	<i>Transformer</i> , khususnya IndoT5, mampu melakukan peringkasan teks abstraktif bahasa Indonesia dengan ringkasan yang cukup dapat dipahami, meskipun masih terdapat <i>typo</i> dan kesalahan tata bahasa, dengan skor ROUGE tertinggi ROUGE-1 sebesar 0.61 dan ROUGE-2 sebesar 0.51.	Model LSTM-atensi memiliki performa yang rendah dengan nilai RL tertinggi sebesar 13.0/0.13, sementara model T5 <i>Small</i> meraih performa paling baik dengan nilai RL sebesar 66.2/0.662, R1 = 0.69, dan R2 = 0.62.	Model Modifikasi BERT2BERT memiliki performa yang cukup baik dalam peringkasan abstraktif dengan nilai R1 = 41.13/0.41, R2 = 22.85/0.23, R3 = 34.65/0.35, dan BERTScore = 73.30

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Penelitian sebelumnya menunjukkan bahwa model berbasis *Transformer* unggul dalam tugas peringkasan abstraktif berbahasa Indonesia. IndoT5 terbukti mampu menghasilkan ringkasan yang relevan dan mudah dipahami (Bahari *et al.*, 2024), sementara T5 *Small* menunjukkan performa yang lebih baik dibandingkan pendekatan LSTM–*Self Attention* pada peringkasan teks berita (Rosaria *et al.*, 2024).

Sementara itu, model BERT2BERT yang dimodifikasi untuk meniru arsitektur *encoder–decoder* seperti T5 juga menunjukkan kinerja yang cukup baik pada peringkasan abstraktif *multi-dokumen*, namun memiliki performa yang lebih rendah dibandingkan pendekatan yang secara langsung menggunakan arsitektur T5 (Muharam *et al.*, 2025).

Oleh karena itu, IndoT5 dipilih sebagai basis model yang efektif dan efisien untuk peringkasan ulasan *e-commerce* berbahasa Indonesia dan di evaluasi dengan ROUGE dan BERTScore, dengan fokus penelitian ini pada proses *fine-tuning* IndoT5 dalam merangkum ulasan Tokopedia secara informatif dan representatif.

POLITEKNIK
NEGERI
JAKARTA



Hak Cipta :

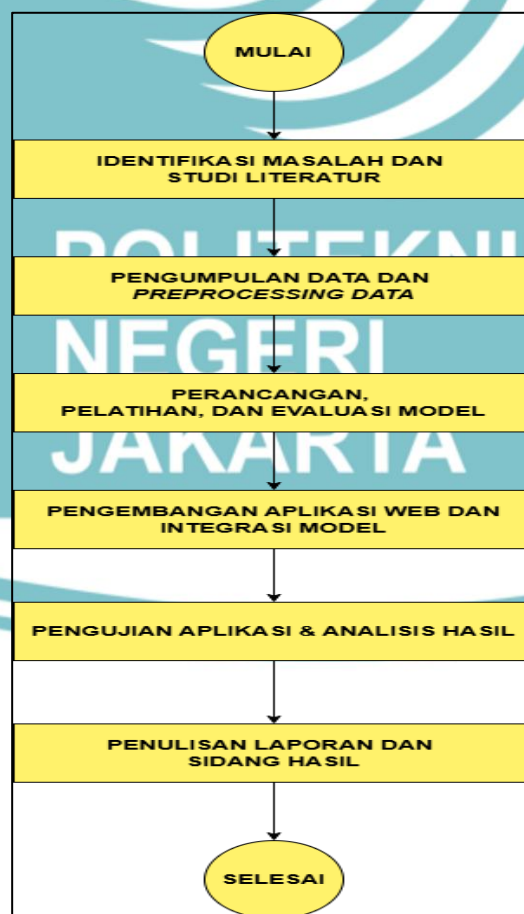
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB 3

METODE PENELITIAN

3.1 Rancangan Penelitian

Penelitian ini dirancang untuk mengimplementasikan arsitektur Indonesia *Text-to-Text Transfer Transformers* (IndoT5) dalam peringkasan teks abstraktif otomatis pada ulasan produk *e-commerce* berbahasa Indonesia melalui proses *fine-tuning* terhadap model pra-latih. Model peringkasan yang dihasilkan kemudian diintegrasikan ke dalam aplikasi berbasis web sebagai media uji coba untuk menampilkan hasil ringkasan secara langsung kepada pengguna. Evaluasi kualitas ringkasan dilakukan menggunakan metrik ROUGE-L dan BERTScore, sedangkan evaluasi terhadap sistem dilakukan dengan metode *System Usability Scale* (SUS) dan *Net Promoter Score* (NPS).



Gambar 3.1. Diagram Alur Penelitian



3.1.1 Tahapan Penelitian

Tahapan penelitian dalam implementasi model peringkasan teks abstraktif berbasis IndoT5 mengikuti bagan pada lampiran I dengan rincian sebagai berikut:

1. Identifikasi Masalah dan Studi Literatur

Tahap ini bertujuan untuk memahami permasalahan terkait penting dan banyaknya ulasan produk *e-commerce* berbahasa Indonesia yang memerlukan ringkasan agar memudahkan pengguna memahami opini umum pembeli. Pada tahap ini ditentukan fokus masalah, cakupan penelitian, dan strategi pemecahannya. Selain itu, dilakukan studi literatur dengan mengumpulkan referensi dari jurnal, artikel ilmiah, dan sumber relevan lainnya untuk memperoleh dasar teori terkait peringkasan teks abstraktif, model IndoT5, dan NLP berbahasa Indonesia.

2. Persiapan dan *Preprocessing* Data

Dataset ulasan produk *e-commerce* yang diperoleh akan melalui proses pembersihan dan *preprocessing* agar model dapat belajar secara optimal. Proses ini meliputi:

- Sintesis Ulasan – Memperbanyak ragam bentuk ulasan dengan menggunakan LLM GPT 4 yang meniru gaya ulasan produk di tokopedia
- Pembersihan Teks – Menghapus karakter khusus, tanda baca yang tidak relevan dan penghapusan frasa berulang.
- Pengelompokan dan penggabungan ulasan yang ada berdasarkan *Product ID* yang sama.
- Sintesis data ringkasan dari tumpukan ulasan yang sudah di kelompokkan dan digabungkan.
- Tokenisasi dan *Padding* – Memecah teks menjadi token sesuai format *input* IndoT5 dan menyesuaikan panjang *input* agar seragam untuk pelatihan.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

- Uji Kelayakan *Dataset* – Menyaring Baris pada *dataset* yang layak untuk digunakan sebagai data latih dan evaluasi dengan menggunakan parameter pengujian seperti *compression ratio* dan *cosine similarity*.
- Pengaturan *Dataset* – Membagi *dataset* menjadi 90% data latih, dan 10% uji untuk memastikan evaluasi model yang akurat.

3. Perancangan, Pelatihan, dan Evaluasi Model

Perancangan pelatihan dan evaluasi model dilakukan secara bertahap melalui *fine-tuning* IndoT5. Tahap awal berfokus pada pemilihan varian model terbaik antara IndoT5-small, IndoT5-base, dan IndoT5-large menggunakan *hyperparameter* yang dibuat tetap. Setelah varian terbaik diperoleh, dilakukan *hyperparameter tuning* untuk mengoptimalkan *learning rate*, *batch size*, dan jumlah *epoch*. Model terpilih kemudian dievaluasi menggunakan data uji dengan metrik ROUGE-L dan BERTScore, serta dilengkapi evaluasi manual oleh pengguna melalui aplikasi web yang dibangun guna memastikan ringkasan yang dihasilkan informatif, relevan, dan mudah dipahami.

4. Pengembangan Aplikasi Web dan Integrasi Model

Model yang telah terlatih akan diintegrasikan ke dalam aplikasi *website* yang dibangun menggunakan *Framework* python, FastAPI, sebagai sarana uji coba sehingga pengguna dapat langsung melihat ringkasan ulasan secara praktis. Implementasi web ini hanya untuk pengujian dan evaluasi, bukan untuk produksi skala besar.

5. Pengujian Aplikasi dan Analisis Hasil

Pengujian dilakukan melalui skenario uji untuk menilai akurasi, konsistensi, dan kualitas ringkasan yang dihasilkan sistem melalui uji Black Box, SUS NPS, dan UAT. Hasil pengujian lalu dianalisis dan didokumentasikan sebagai bahan laporan penelitian yang akan diseminasi.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

3.1.2 Objek Penelitian

Objek penelitian dalam penelitian ini adalah aplikasi peringkasan ulasan produk e-commerce berbahasa Indonesia dengan menggunakan model berbasis T5, objek penelitian ini akan diuji dan dievaluasi berdasarkan beberapa aspek seperti performa model dalam menghasilkan ringkasan ulasan, Tingkat kegunaan *Website* dan tingkat kepuasan pengguna terhadap sistem atau aplikasi yang telah dibuat.





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB 4

HASIL DAN PEMBAHASAN

4.1 Identifikasi Kebutuhan

Berdasarkan hasil telaah pustaka dari berbagai jurnal yang telah dikaji, diperoleh sejumlah kebutuhan yang menjadi dasar dalam penelitian ini. Kebutuhan tersebut meliputi *fine-tuning* model, pengembangan aplikasi web, serta perangkat lunak dan perangkat keras pendukung. Seluruh kebutuhan tersebut perlu dipenuhi agar proses penelitian dapat berjalan secara optimal dan sistem yang dikembangkan dapat berfungsi dengan baik.

4.1.1 Kebutuhan *Fine-tuning* Model

1. Kebutuhan *Dataset*

Dataset menjadi komponen yang sangat penting dalam proses *fine-tuning* model *Transformer* seperti IndoT5. Kriteria atau spesifikasi *dataset* yang digunakan dalam penelitian ini dapat dilihat pada tabel berikut

Tabel 2. Kriteria Kebutuhan *Dataset*

Kode	Nama	Deskripsi
KD-1	Bahasa Indonesia	Bahasa yang digunakan dalam <i>dataset</i> didominasi oleh bahasa Indonesia, dengan sisipan beberapa istilah dalam bahasa Inggris. Bagian ulasan ditulis dalam gaya bahasa yang cenderung informal, sedangkan bagian ringkasan disusun menggunakan bahasa yang lebih formal.
KD-2	Jumlah data	Jumlah data yang ditargetkan dalam penelitian ini adalah sebanyak minimal 6.000 baris, dengan komposisi yang terdiri atas 4000 data sentimen campuran, 1000



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Kode	Nama	Deskripsi
		data sepenuhnya negatif, dan 1000 data sepenuhnya positif.
KD-3	Isi Data	Berisi data dalam bentuk teks yang terdiri atas pasangan antara kumpulan ulasan dan hasil ringkasannya. Setiap kumpulan ulasan sekurang-kurangnya terdiri dari 40 ulasan.

Dataset yang digunakan dalam penelitian ini merupakan gabungan antara data hasil *scraping* asli dari situs Tokopedia dan data sintesis yang dihasilkan menggunakan *Large Language Model* (LLM). Data sintesis tersebut bertujuan untuk membentuk ringkasan dari kumpulan ulasan sekaligus memperkaya variasi data, sehingga dapat membantu meningkatkan performa model.

Jumlah minimum 6.000 baris dipilih sebagai titik tengah yang realistis antara kebutuhan model dan ketersediaan data. IndoT5 punya parameter yang besar, jadi butuh data yang cukup supaya proses *fine-tuning* stabil dan tidak mudah *overfitting*. Di sisi lain, ulasan *e-commerce* yang benar-benar natural dan relevan jumlahnya juga terbatas. Berdasarkan temuan Mehrafarin et al. (2022), *dataset* di kisaran 7.000 sampel sudah cukup membuat model *transformer* mempelajari pola pada tugas spesifiknya. Sementara itu, jumlah data yang jauh lebih kecil seperti 1.000–2.500 masih kurang optimal sehingga 6.000 data bisa dibilang sudah cukup sebagai batas bawah minimum dan tetap realistis untuk mendukung *fine-tuning* yang efektif.

Komposisi data dalam penelitian ini tidak dibuat sepenuhnya seimbang karena IndoT5 merupakan model *text-to-text* yang berfokus pada pembelajaran pola peringkasan, bukan klasifikasi sentimen. Meskipun demikian, distribusi data tetap dijaga agar tidak terlalu timpang untuk menghindari bias dalam proses pembelajaran. Proporsi data campuran dibuat



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

lebih dominan karena secara alami ulasan pada *e-commerce* umumnya mengandung kombinasi sentimen positif dan negatif.

Selanjutnya, *dataset* akan dibagi menjadi dua bagian, yaitu 90% sebagai data latih dan 10% sebagai data evaluasi. Proporsi ini dipilih karena jumlah data yang tersedia tidak terlalu besar, sehingga porsi data latih perlu dioptimalkan agar model dapat belajar secara maksimal

2. Kebutuhan Model

Dalam mengembangkan fitur AI untuk tugas peringkasan *text* ulasan berbahasa *indonesia*, maka diperlukan sebuah model dengan kriteria seperti yang dapat dilihat pada tabel berikut.

Tabel 3. Kriteria Kebutuhan Model

Kode	Nama	Deskripsi
KM-1	Bahasa Indonesia	Model yang digunakan harus dapat memahami dan menghasilkan kalimat - kalimat dalam bahasa Indonesia.
KD-2	Kemampuan <i>Text-to-Text</i>	Model yang digunakan harus dapat mengubah <i>text</i> inputan menjadi sebuah <i>text</i> lain
KD-3	Berbasis <i>Transformer</i>	Model yang digunakan merupakan model dengan arsitektur berbasis <i>transformer</i>

Dalam mengembangkan fitur AI untuk peringkasan ulasan berbahasa Indonesia, diperlukan model yang mampu memahami serta menghasilkan teks dalam bahasa Indonesia secara baik. Selain itu, model harus memiliki kemampuan *text-to-text*, yaitu mampu mentransformasikan teks masukan menjadi teks keluaran dalam bentuk ringkasan yang relevan dan terstruktur.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Model juga perlu berbasis arsitektur *Transformer* karena mekanisme *self-attention* yang dimilikinya memungkinkan pemahaman konteks secara lebih menyeluruh, termasuk pada teks yang panjang.

4.1.2 Kebutuhan Pengembangan Web

1. Kebutuhan Fungsional

Dalam mengembangkan sebuah *website* yang dapat menjadi antarmuka bagi pengguna untuk menggunakan fitur AI peringkasan ulasan Tokopedia, dibutuhkan sistem yang memiliki fungsional sebagai berikut.

Tabel 4. Kebutuhan Fungsional Web

Kode	Nama	Deskripsi
KF-1	<i>Input</i> Link	Pengguna dapat memasukan <i>link</i> tokopedia secara manual di <i>Website</i> .
KF-2	<i>Output</i> Ringkasan	Pengguna dapat melihat hasil ringkasan ulasan yang dihasilkan AI berdasarkan <i>link</i> yang dimasukan.

2. Kebutuhan Non-Fungsional

Selain memenuhi kebutuhan fungsional, sistem juga harus memperhatikan kebutuhan non-fungsional guna memastikan kenyamanan pengguna dalam mengoperasikan sistem.

Tabel 5. Kebutuhan Non-Fungsional Web

Kode	Nama	Deskripsi
KNF-1	Responsif	Antarmuka sistem web bisa menyesuaikan tata letak sesuai ukuran jenis dan perangkat pengguna.
KNF-2	Kecepatan Peringkasan	Sistem web mampu menghasilkan ringkasan dalam waktu 1–3 menit serta tetap dapat menangani beberapa



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Kode	Nama	Deskripsi
		permintaan secara bersamaan ketika terjadi <i>high traffic</i> .

3. Kebutuhan Perangkat Keras

Untuk melakukan *fine tune* model indoT5 secara efektif, dibutuhkan perangkat keras yang dapat mengakselerasi proses pelatihan dan dapat digunakan dengan spesifikasi sebagai berikut.

Tabel 6. Kebutuhan Perangkat Keras

Kode	Nama	Deskripsi
KPK-1	GPU	Tesla P100-PCIE-16GB./T4x2
KPK-2	CPU	Intel(R) Xeon(R) CPU @ 2.20GHz
KPK-3	RAM	Minimal 16 GB

4. Kebutuhan Perangkat Lunak

Untuk mengembangkan sistem, mulai dari persiapan data hingga *deployment integrated web*, dibutuhkan *software* sebagai berikut.

Tabel 7. Kebutuhan Perangkat Lunak

Kode	Nama	Deskripsi
KPL-1	Visual Studio Code	<i>Source-Code</i> Editor untuk mengoperasikan kode python beserta dengan <i>library</i> pendukung untuk data <i>preprocessing</i> dan mengembangkan sistem <i>website</i> .
KPK-2	Jupyter Notebook	Virtual <i>code</i> editor dengan akselerator GPU untuk mengoperasikan kode python beserta <i>library</i> pendukung untuk <i>fine tune</i>



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Kode	Nama	Deskripsi
		model yang disediakan oleh google dan kaggle.
KPK-3	N8N	<i>Software</i> untuk membuat proses otomatisasi sintesis <i>dataset</i> dengan LLM GPT 4.
KPK-4	<i>AWS Cloud</i>	Platform Online untuk melakukan <i>deployment</i> sistem.

4.2 Perancangan Sistem

Setelah kebutuhan penelitian berhasil diidentifikasi, untuk mengembangkan sistem, dibutuhkan sebuah perancangan sistem. perancangan sistem dibagi menjadi dua aspek utama, yaitu perancangan model dan perancangan web.

4.2.1 Perancangan Model

1. Pengumpulan dan sintesis data ulasan

Pengumpulan data dalam penelitian ini memanfaatkan dua sumber utama, yaitu data *open-source* dan data sintesis. Data *open-source* diperoleh dari platform *Kaggle* berupa *dataset* ulasan produk Tokopedia berbahasa Indonesia yang mencakup informasi teks ulasan, *rating*, sentimen, serta identitas produk

Untuk meningkatkan jumlah data sekaligus memperkaya variasi bahasa yang digunakan, dilakukan sintesis ulasan menggunakan model GPT-4 yang meniru gaya penulisan ulasan *e-commerce* melalui proses otomatisasi menggunakan *software* N8N . Kombinasi kedua sumber ini menghasilkan *dataset* yang lebih beragam, sehingga diharapkan dapat membantu model dalam memahami berbagai pola bahasa ulasan pengguna.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2. *Preprocessing* data Ulasan

Sebelum digunakan dalam proses pelatihan model, data ulasan terlebih dahulu melalui tahap *preprocessing* untuk meningkatkan kualitas dan konsistensi data. Tahapan ini meliputi pembersihan teks dari karakter khusus seperti emotikon, tanda baca yang tidak relevan, serta penghapusan frasa yang berulang. Sementara itu, kata singkatan atau kata tidak baku tetap dipertahankan dengan harapan model dapat mempelajari pola bahasa alami yang umum digunakan dalam ulasan *e-commerce*.

Selanjutnya, ulasan dikelompokkan dan digabungkan berdasarkan *Product ID* yang sama, karena model dirancang untuk meringkas sekitar 40 tumpukan ulasan menjadi satu ringkasan abstraktif, bukan meringkas ulasan secara individual. Seluruh langkah ini bertujuan memastikan data siap diproses secara optimal oleh model sekaligus meminimalkan gangguan akibat *noise* pada teks.

3. Sintesis data ringkasan

Setelah tumpukan ulasan terbentuk, dilakukan proses sintesis ringkasan menggunakan model bahasa GPT-4 untuk menghasilkan ringkasan abstraktif dari setiap kumpulan ulasan. Proses ini tidak dilakukan secara manual per baris data, tetapi diotomatisasi menggunakan *workflow automation tool* N8N, sehingga setiap tumpukan ulasan dapat diproses secara sistematis dan konsisten dalam jumlah besar.

Melalui alur otomatis tersebut, data ulasan dikirim ke GPT-4 untuk dihasilkan ringkasan yang tidak hanya mengekstrak kalimat penting, tetapi juga merangkum informasi utama secara natural dan koheren. Hasil akhirnya berupa pasangan data yang terdiri dari teks ulasan gabungan dan ringkasan abstraktifnya, yang kemudian digunakan sebagai data pelatihan model peringkasan.

4. Uji Kelayakan *Dataset*

Oleh karena dataset sebagian besar di-*generate* dengan bantuan LLM, sehingga perlu dilakukan pengujian dataset untuk memastikan kualitas



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

dataset yang digunakan selama proses pelatihan model IndoT5. Dalam tahap ini, dilakukan uji kelayakan terhadap setiap pasangan teks ulasan dan ringkasan. Proses ini menggunakan parameter evaluasi seperti *compression ratio* untuk mengukur tingkat pemadatan informasi ulasan kedalam kalimat ringkasan, *cosine similarity* untuk menilai kesesuaian semantik antara ulasan dan ringkasan, dan pengujian secara manual yang dilakukan oleh ahli linguistik. Secara matematis, *compression* rasio dihitung menggunakan persamaan berikut

$$\text{Compression Ratio} = \frac{\text{Panjang token ringkasan}}{\text{Panjang token ulasan}} \times 100\%$$

Untuk nilai kesamaan semantik, dihitung menggunakan rumus *cosine similarity* sebagai berikut

$$\text{Cosine Simmilarity } (A, B) = \frac{A \times B}{||A|| \times ||B||}$$

Dimana,

A = Representasi vektor teks ulasan

B = Representasi vektor teks ringkasan

Data yang tidak memenuhi kriteria, yaitu memiliki *compression ratio* lebih dari 50% dan nilai *similarity* kurang dari 70% akan dieliminasi untuk mencegah penurunan kualitas pelatihan model.

Penelitian terdahulu menyebutkan *Compression ratio* 50% dinilai paling optimal sebagai nilai ambang batas *compression rate* (Halimah et al., 2022). Sebaliknya, angka di atas 50% berisiko menghasilkan pola ringkasan yang kurang ideal, karena jumlah token ringkasan seharusnya lebih sedikit daripada token teks asli.

Ambang batas *similarity* sebesar 70% dipilih karena nilai di bawah angka tersebut berisiko menurunkan kualitas kesesuaian makna antara ulasan yang diringkas dengan hasil ringkasannya. Selain itu, beberapa penelitian sebelumnya juga mengategorikan rentang nilai *similarity* 0,6–1,0 sebagai tingkat kemiripan tinggi (Mardatillah et al., 2021).



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Setelah melakukan pengujian secara komputasi komputer, barulah dilakukan pengujian manual oleh ahli dan dilakukan sampling dataset karena mengingat mahalanya biaya pengujian jika menggunakan data populasi. Sampling dilakukan dengan menggunakan metode slovin.

Metode Slovin dipilih karena jumlah populasi pasangan data ulasan dan ringkasan telah diketahui secara pasti, sedangkan karakteristik kualitas populasi belum diketahui. Di sisi lain, proses validasi dilakukan secara manual oleh ahli bahasa sehingga pemeriksaan terhadap seluruh populasi kurang efisien dari segi waktu dan sumber daya. Oleh karena itu, metode Slovin digunakan untuk menentukan ukuran sampel yang representatif berdasarkan tingkat kesalahan (margin of error) yang telah ditetapkan. adapun perhitungan slovin dengan batas kesalahan sebesar 10% seperti yang dapat dilihat pada persamaan berikut.

$$n = \frac{N}{1 + N(e^2)} = \frac{N}{1 + N(0.1^2)}$$

dimana,

N = Jumlah populasi

n = jumlah sample

e = *batas toleransi error*

sub-sampel tersebut akan diuji berdasarkan dua penilaian, yaitu dari aspek representatif-abstraktif dan aspek halusinasi. Aspek representatif-abstraktif akan menilai seberapa baik ringkasan yang dihasilkan LLM menangkap poin penting dalam ulasan yang ada sehingga ringkasan sudah cukup mewakili seluruh ide pokok dari ulasan-ulasan yang ada. Sementara aspek halusinasi menilai apakah ringkasan yang dihasilkan memuat sesuatu yang tidak seharusnya akibat dari hayalan atau imajinasi dari LLM (AI).

Ahli linguistik akan menilai kedua aspek di atas secara kategoris, yaitu “Ya” atau “Tidak”. *Dataset* dinyatakan layak digunakan jika hasil penilaian sub-sampel oleh ahli linguistik minimal mencapai 50% “Ya pada kedua aspek.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Dengan demikian, proses seleksi ini bertujuan memastikan *dataset* yang digunakan relevan, konsisten, dan mampu mendukung model dalam menghasilkan ringkasan yang akurat serta representatif.

5. *Fine-Tune* Model IndoT5

Setelah data dinyatakan siap, tahap berikutnya adalah proses pelatihan model yang dilakukan dalam dua fase utama. Fase pertama berfokus pada pemilihan varian model terbaik di antara IndoT5-small, IndoT5-base, dan IndoT5-large dengan seluruh *hyperparameter* dibuat tetap agar perbandingan berlangsung objektif. Tujuannya adalah menentukan arsitektur dengan performa terbaik sebelum dilakukan optimasi lebih lanjut.

Fase kedua merupakan proses *hyperparameter tuning* yang berfokus pada eksplorasi kombinasi *learning rate* dan *batch size*. Secara matematis, pembaruan bobot pada proses optimasi mengikuti persamaan umum *gradient descent* berikut ini (Shazeer et al., 2018).

$$W_t = W_{t-1} - \alpha_t \nabla f(x_{t-1}) \quad (1)$$

Di mana,

W_t : bobot pada iterasi ke- t

α_t : *learning rate*

$\nabla f(x_{t-1})$: gradien fungsi *loss*.

Dari persamaan tersebut terlihat bahwa besar kecilnya perubahan bobot ($W_t - W_{t-1}$) secara langsung dipengaruhi oleh nilai *learning rate*. Semakin besar α_t , semakin besar langkah pembaruan bobot sehingga proses menuju konvergensi dapat berlangsung lebih cepat, namun berisiko tidak stabil. Sebaliknya, nilai yang terlalu kecil menghasilkan pembaruan bobot yang lambat sehingga memperpanjang waktu konvergensi.

Sementara itu, *batch size* memengaruhi kualitas estimasi gradien yang digunakan dalam pembaruan bobot. Pada skema *mini-batch*, gradien yang



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

dihitung pada persamaan (1) dapat dihitung melalui persamaan (2) dibawah ini (Qian & Klabjan, 2020).

$$\nabla f_B(\theta) = \frac{1}{B} \sum_{i=1}^B \nabla \ell_i(\theta) \quad (2)$$

Di mana,

B : ukuran *batch*

$\nabla \ell_i(\theta)$: turunan fungsi *loss* terhadap *weight* pada sampel ke- i

$\nabla f_B(\theta)$: estimasi gradien berdasarkan *mini-batch*

Nilai B menentukan seberapa baik gradien mewakili keseluruhan data. Batch size yang lebih besar membuat nilai gradien menjadi lebih stabil karena dihitung dari lebih banyak data (lebih representatif), sehingga pembaruan bobot W_t menjadi lebih konsisten dan tidak terlalu berubah-ubah. Sebaliknya, *batch size* yang kecil menghasilkan gradien yang lebih bervariasi karena hanya dihitung dari sedikit data. Hal ini dapat membuat perubahan bobot terjadi lebih cepat, tetapi juga berisiko membuat proses pelatihan menjadi kurang stabil sebelum mencapai konvergensi.

Selain itu, perhitungan gradien loss selama pelatihan menggunakan turunan pertama fungsi rata-rata Cross-Entropy Loss per token dengan pendekatan maximum likelihood dan teknik teacher forcing (Raffel et al., 2020). Fungsi ini mengukur seberapa yakin model dalam memprediksi setiap token target yang benar pada setiap langkah dekoding. Adapun persamaan dari fungsi rata-rata Cross-Entropy Loss per token dapat dilihat pada persamaan berikut.

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \log P_{\theta}(y_i | y_{<i}, x) \quad (3)$$

Di mana,

$\mathcal{L}$: nilai *Cross-Entropy Loss*

N : jumlah total token dalam *sequence* target

y_i : token target ke- i yang benar (*ground truth*)

$y_{<i}$: semua token target sebelum posisi i (*teacher forcing*)

x : *sequence* input dari encoder

P_{θ} : probabilitas yang diprediksi model dengan parameter θ



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

log : logaritma natural dengan basis bilangan euler.

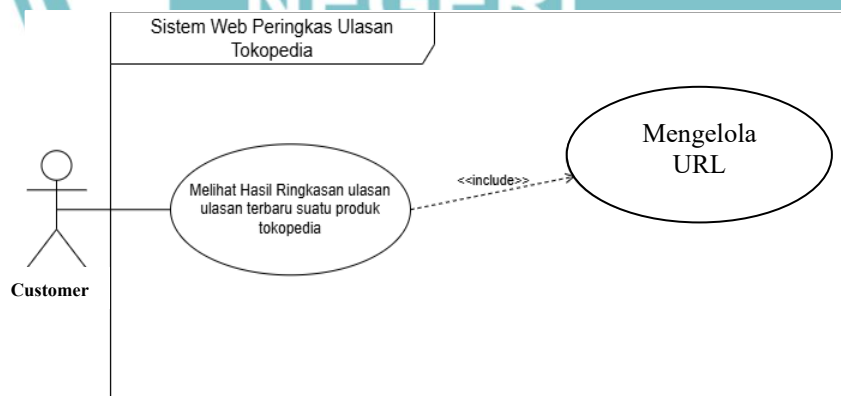
Nilai domain fungsi ini berada pada rentang 0 hingga tak hingga (∞), di mana semakin besar nilai loss maka semakin rendah tingkat keyakinan model dalam memprediksi token yang tepat, dan sebaliknya nilai loss yang mendekati nol mengindikasikan bahwa model hampir selalu memprediksi token yang benar dengan probabilitas tinggi.

Dengan demikian, kombinasi *learning rate* dan *batch size* secara langsung memengaruhi besar kecilnya perubahan W_t pada setiap iterasi. Perubahan tersebut pada akhirnya menentukan kecepatan serta stabilitas model dalam mencapai konvergensi yang berdampak pada kualitas hasil pembelajaran model. Selama *hyperparameter tuning*, Jumlah *epoch* diseragamkan untuk seluruh eksperimen, serta diterapkan mekanisme *early stopping* guna menghentikan pelatihan apabila tidak terjadi peningkatan performa selama tiga *epoch* berturut-turut. Model terbaik kemudian dievaluasi menggunakan metrik ROUGE-L dan BERTScore.

4.2.2 Perancangan Web

1. Use Case Diagram

Berdasarkan kebutuhan fungsional yang telah didefinisikan sebelumnya, berikut adalah *use case diagram* dari sistem web yang dibangun.



Gambar 4.1 Use Case Diagram Web

Diagram Use Case tersebut menunjukkan bahwa *User* berinteraksi dengan Sistem Web Peringkat Ulasan Tokopedia untuk melihat ringkasan ulasan terbaru suatu produk. Untuk dapat menampilkan hasil ringkasan, pengguna



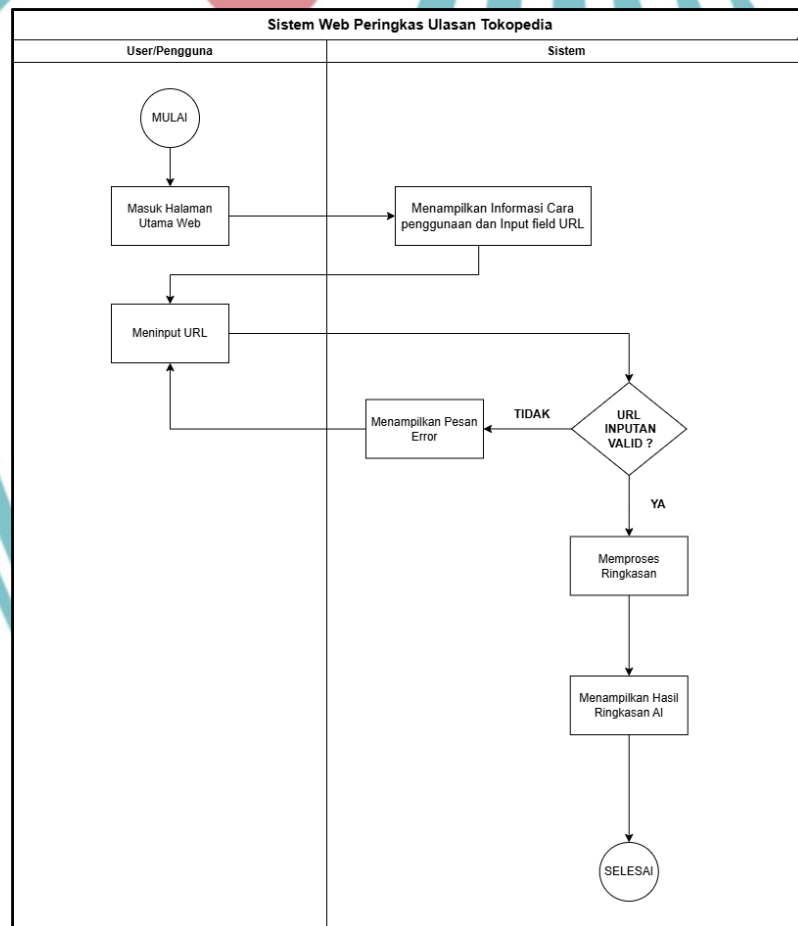
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

harus terlebih dahulu menginput URL halaman utama produk di Tokopedia, yang ditunjukkan dengan relasi `<<include>>` sebagai langkah wajib dalam proses. Selain itu, *User* juga dapat menghapus lalu menginput URL baru.

2. *Activity* Diagram

Activity Diagram adalah pengembangan dari *use case* diagram yang menggambarkan alur kerja tiap fitur pada sistem Peringkasan ulasan otomatis suatu produk di tokopedia. Diagram ini menunjukkan interaksi antara *user* dengan sistem, sehingga membantu memperjelas proses serta memudahkan pengembangan dan pemeliharaan sistem.



Gambar 4.2 *Activity* Diagram Web

Diagram *activity* tersebut menggambarkan alur kerja sistem web peringkasan ulasan Tokopedia dari interaksi antara pengguna dan sistem. Proses dimulai ketika pengguna membuka halaman utama web, lalu sistem menampilkan



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

petunjuk penggunaan serta kolom *input* URL produk. Pengguna kemudian memasukkan URL produk Tokopedia dan sistem melakukan validasi. Jika URL tidak valid, sistem menampilkan pesan *error*. Jika URL valid, sistem akan melanjutkan proses peringkasan melalui beberapa tahapan. Pertama, sistem akan mengambil maksimal 200 ulasan terbaru yang dibuat dalam kurun waktu kurang dari satu tahun. Pemilihan ini didasarkan pada pertimbangan *trade-off* antara batas maksimum token *input* saat pelatihan model dan tingkat relevansi ulasan produk. Semakin baru ulasan yang digunakan, maka ringkasan yang dihasilkan akan semakin relevan. Selanjutnya, sistem akan membersihkan data hasil *scraping* sebelum akhirnya mengirimkannya ke model AI untuk diringkas.

Selanjutnya, sistem mengecek apakah terdapat respons sukses dari AI *engine*. Jika respons berhasil diterima, sistem menampilkan hasil ringkasan kepada pengguna; jika tidak, sistem menampilkan pesan *error*. Setelah itu, proses dinyatakan selesai

3. *Wireframe User Interface*

Untuk memudahkan pengembangan antarmuka *website*, dibutuhkan sebuah *wireframe* yang dapat dijadikan sebagai *blueprint* agar pengembangan tampilan menjadi lebih terarah. Berikut ini adalah *wireframe* dari web yang akan dikembangkan.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Gambar 4.3 Wireframe Halaman Utama

Sesuai dengan kebutuhan fungsional dari web yang telah dijabarkan diatas, maka tampilan/antarmuka web dibuat sesederhana mungkin ke dalam satu *page* yang terdiri atas 2 *section*, yaitu *section* informasi dan *section generate* ringkasan. Seperti yang terlihat dalam Gambar 4.3, *section* informasi berisikan komponen-komponen informasi seputar web seperti *engine* apa yang digunakan, cara menggunakan *website*, dan *value proposition* dari *tools* AI yang digunakan. Selanjutnya, *section generate* ringkasan berisikan *field* untuk memasukan *url* dan *output textbox* yang akan berisikan hasil ringkasan ulasan produk dari *url* tokopedia yang dimasukan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

4.3 Implementasi Sistem

Rancangan sistem yang telah dirumuskan kemudian direalisasikan melalui tahap implementasi. Pada penelitian ini, proses tersebut dilaksanakan dalam dua bagian, yaitu *fine-tune* model dan pengembangan aplikasi web. Pemisahan ini dilakukan karena aplikasi web hanya dapat berjalan optimal apabila model peringkasan telah selesai di – *fine-tune* dan terbukti memiliki performa yang baik. Dengan alur pengembangan seperti ini, sistem web yang dihasilkan benar-benar menggunakan model terbaik sebagai *engine* AI utamanya

4.3.1 Implementasi Model

1. Pengumpulan dan sintesis data ulasan

Seperti yang telah dijelaskan pada tahap perancangan, data akan diperoleh melalui dua sumber, yaitu data *scraping* tokopedia dari *Kaggle* dan proses automasi sintesis ulasan. Berdasarkan eksplorasi *dataset* di *kaggle*, ditemukan 2 *dataset* utama yang relevan dengan penelitian, yaitu *dataset* ulasan tokopedia 2019 dan 2025. Dari 2 *dataset* tersebut, didapatkan 105.000 baris ulasan lengkap dengan sentimen, id produk dan deskripsi lainnya. Kemudian, dari hasil sintesis ulasan, didapatkan 3.292 data tumpukan ulasan dari kombinasi 1000 produk dan sentimen (heterogen dan homogen). Berikut ini adalah *prompt* yang digunakan untuk menghasilkan data sintesis ulasan selama proses automasi berlangsung.

Tabel 8. *Prompt Generate Ulasan Sintesis*

Input dan konfigurasi	Prompt Messages		Output
	System	User	
{\$ nama_produk} Temperature = 0.1	saat ini aku sedang membuat <i>dataset</i> untuk <i>finetune</i> model T5. bantu aku untuk membuat <i>dataset</i> itu dengan mengikuti aturan berikut : 1.) anda akan menerima	berikut adalah nama produk untuk anda ulas = {\$ nama_produk}	Tumpukan ulasan dipisahkan dengan titik



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Input dan konfigurasi	Prompt Messages	Output
	nama suatu produk. 2.) dari nama itu, buatlah tumpukan ulasan dalam bentuk paragraf yang berisikan 35 sampai 40 ulasan pembeli dari produk itu 3.) setiap ulasan harus disampaikan dengan bahasa natural dan banyak singkatan seperti khalayak umum menulis ulasan di <i>ecommerce</i> contohnya dengan kata (bgt,bgs,yg,tp) dan sebagainya 4.) tumpukan ulasan harus bersentimen campur, supaya terkesan natural dari hasil <i>scrapping</i> di tokopedia, misalnya 30 bernuansa positif dari segi bahan sementara 5 lainnya bernuansa negatif dari segi pengiriman atau warna.namun, jika produk adalah non-fisik, maka sampaikan pengalaman memakainya secara jujur 5.) terakhir, langsung berikan <i>output</i> tanpa penjelasan sebelum dan sesudah jawaban. antar ulasan dipisahkan dengan titik (.) **contoh** (barangny sdh d terima, blm dicoba tp warnany ok.	



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

<i>Input dan konfigurasi</i>	<i>Prompt Messages</i>	<i>Output</i>
	bgs sih utk harga segini mah pantes aj. ini kurirnya kacau dus ny sampe penyok, barangny jd peyot ni lain kali hati hati.) 6.) sekarang mulai buat tumpukan ulasan !	

2. *Preprocessing* data Ulasan

Setelah diperoleh data mentah sebanyak 105.000 baris ulasan produk dari Tokopedia, data tersebut terlebih dahulu melalui tahap *pre-processing*. Pada tahap ini dilakukan pembersihan data dari karakter yang tidak beraturan, kata yang berulang, serta emotikon yang berpotensi menjadi *noise*. Setelah proses pembersihan selesai, data kemudian dikelompokkan berdasarkan id produk yang sama, dengan ketentuan setiap kelompok memiliki minimal 40 baris ulasan. Hasil akhir dari proses *pre-processing* terhadap data mentah ini menghasilkan 4.324 baris data yang selanjutnya digabungkan dengan kumpulan ulasan sintetis, sehingga total keseluruhan menjadi 7.634 baris tumpukan ulasan.

3. Sintesis data ringkasan

Setiap data tumpukan ulasan selanjutnya diringkas melalui proses otomatisasi menggunakan N8N. Hasil dari proses otomatisasi peringkasan tersebut berupa data berpasangan antara tumpukan ulasan dan ringkasannya, dengan total sebanyak 7.634 pasangan data. Adapun *prompt* yang digunakan selama proses sintesis ringkasan dapat dilihat pada tabel berikut.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Tabel 9. *Prompt Generate Ringkasan Ulasan*

Input dan konfigurasi	Prompt Messages		Output
	System	User	
{ \$ tumpukan_ulasan} Temperature = 0.1	"saat ini aku sedang membuat <i>dataset</i> untuk <i>finetune</i> model T5. bantu aku untuk membuat <i>dataset</i> itu dengan mengikuti aturan berikut : 1.) anda akan menerima paragraf yang berisikan tumpukan ulasan pelanggan terhadap suatu produk yang ditulis dengan bahasa <i>indonesia</i> yang tidak baku. 2.) dari paragraf tersebut, buatlah ringkasan **ABSTRAKTIF** dari ulasan tersebut, dengan panjang ringkasan **maksimal 25 sampai 30 persen dari panjang tumpukan ulasan** . intinya, **Ringkasan ulasan tidak boleh lebih panjang dari tumpukan ulasan <i>original</i>** 3.) ringkasan harus ditulis secara naratif abstraktif dengan bahasa <i>indonesia</i> yang natural, tidak kaku dan mengalir seakan anda menjelaskan ringkasan ulasan produk kepada calon pelanggan secara	berikut adalah tumpukan ulasan untuk anda Ringkas = { \$ tumpukan_ulasan}	Pasangan Tumpukan ulasan dan ringkasan



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Input dan konfigurasi	Prompt Messages	Output
	jujur berdasarkan ulasan yang ada dan tidak mengarang. ringkasan tidak ditulis berbentuk poin-poin tetapi berbentuk narasi dan tidak dijabarkan dengan titik dua (:) 4.) langsung berikan <i>output</i> tanpa penjelasan sebelum dan sesudah jawaban. <i>output</i> format harus diawali 'produk ini (lanjut ringkaskan ulasan produk secara komprehensif,ringkas,dan jujur)'. 5.) **CONTOH OUTPUT yang benar** = 'Produk ini dinilai bagus karena barang sesuai gambar, harga bersaing, dan <i>packing</i> rapi. Meskipun ada keluhan soal pengiriman atau respon yang terkadang lambat, mayoritas pembeli tetap puas dengan fungsi jamnya dan merekomendasikan toko ini.' 6.) sekarang mulai ringkas tumpukan ulasan"	



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

4. Uji Kelayakan *Dataset*

Untuk memastikan bahwa *dataset* yang digunakan benar-benar layak sebagai data latih model, dilakukan proses uji kelayakan yang mencakup 3 aspek utama, yaitu rasio kompresi, nilai *similarity* dan pengujian manual oleh ahli linguistik. Berdasarkan hasil eksplorasi, dari total 7.634 data tumpukan ulasan, sebanyak 6789 data dinyatakan memenuhi kriteria seleksi. Data yang lolos tersebut kemudian dibagi menjadi dua bagian, yaitu 6449 data latih dan 340 data validasi.

Tabel 10. Hasil Uji Kelayakan *Dataset*

Metrik Pengujian	Nilai <i>Threshold</i>	Jumlah data tersisa
Uji <i>Compression Ratio</i>	$\leq 50\%$	7522
Uji <i>Cosine Similarity</i>	$\geq 70\%$	6789

Pada uji rasio kompresi, diperoleh rentang nilai antara 3% hingga 74%. Data dengan nilai kompresi diatas 50% dieliminasi sehingga menyisakan 7522 data dengan nilai kompresi $\leq 50\%$.

Setelah itu, 7522 data tersebut diuji *cosine similarity* dan diperoleh nilai minimum sebesar 55% dan maksimum sebesar 91%. Secara distribusi, nilai *similarity* menunjukkan pola yang cukup normal dengan rata-rata berada pada angka 77%, serta rentang kuartil (Q1, Q3) berada di antara 73% hingga 81%. Dalam tahap seleksi, data dengan nilai *similarity* dibawah 70% dieliminasi sehingga menyisakan 6789 data dengan nilai *similarity* $\geq 70\%$, *input* token ≤ 1900 token, serta token *output* ≤ 230 token.

Data yang sudah dieleminasi tersebut kemudian diambil sebagian dengan metode slovin dan didapatkan $\cong 100$ sampel. Dari hasil pengujian oleh ahli terhadap sub-sampel tersebut, ahli menyatakan bahwa pasangan ulasan-ringkasan yang ada pada data mayoritas layak untuk digunakan sebagai data latih dengan perbandingan kategori 95% “layak” dan 5% “tidak layak”.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Adapun 5% data yang tidak layak digunakan, hal tersebut disebabkan oleh adanya halusinasi pada hasil ringkasan LLM sebagai contoh terdapat narasi “packing rapi” pada ringkasan padahal dalam ulasannya tidak ada. Akan tetapi, angka halusinasi ini jauh sangat sedikit jika dibandingkan dengan yang layak sehingga data masih bisa digunakan untuk digunakan sebagai *dataset* pelatihan model IndoT5.

Terakhir, data dibagi untuk pelatihan, validasi dan pengujian dengan perbandingan 90:5:5 yang menghasilkan 6126 data latih, 323 data validasi, dan 340 data uji.

5. *Fine-Tune* Model IndoT5

Pada fase pertama, yaitu fase pemilihan varian model indoT5, konfigurasi *hyperparameter* dan jumlah data dibuat sama, yaitu *Epoch* = 6, *Batch Size* = 8, dan *Learning Rate* = 10^{-4} , sehingga fokusnya hanya pada pengaruh jumlah parameter model terhadap nilai evaluasi. Adapun hasil dari fase pertama dapat dilihat pada tabel berikut.

Tabel 11. Hasil Fase Pertama (Pemilihan varian model)

Kode	Nama Model	Jumlah Parameter	Rouge L
Varian <i>Small</i> (VS)	IndoT5-Small	60 M	0.37
Varian Base (VB)	IndoT5-Base	220 M	0.39
Varian <i>Large</i> (VL)	IndoT5-Large	770 M	0.35

Berdasarkan Tabel 4.10, model dengan performa terbaik adalah model VB dengan perolehan nilai ROUGE-L sebesar 0,39 dan BERTScore-F1 sebesar 0,80. Hasil ini menunjukkan bahwa model VB mampu menghasilkan ringkasan dengan tingkat kesesuaian struktur informasi yang cukup baik serta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

memiliki kedekatan makna yang tinggi secara semantik terhadap ringkasan *ground truth*.

Apabila dibandingkan dengan model VS, performa model VB yang lebih baik diduga dipengaruhi oleh kapasitas parameter yang lebih memadai dalam memahami karakteristik data. Teks masukan pada penelitian ini berupa tumpukan ulasan yang tidak terstruktur dan memiliki variasi gaya bahasa yang cukup beragam. Kompleksitas tersebut kemungkinan belum dapat ditangkap secara optimal oleh model VS yang memiliki sekitar 60 juta parameter, sehingga kemampuan representasinya dalam memahami pola dan konteks menjadi lebih terbatas.

Sementara itu, meskipun model VL memiliki jumlah parameter yang lebih besar, peningkatan kapasitas model tidak selalu berbanding lurus dengan peningkatan performa. Model berukuran besar umumnya memerlukan jumlah data pelatihan yang lebih banyak agar dapat belajar secara lebih optimal. (Hoffman *et al.*, 2022) Dengan demikian, model VB menunjukkan keseimbangan yang lebih tepat antara kapasitas parameter dan jumlah *dataset* yang digunakan dalam penelitian ini.

Setelah varian model terbaik didapatkan dari fase pertama, selanjutnya adalah fase kedua, yaitu *hyperparameter tuning*. Dalam fase ini, varian model terbaik akan di uji dengan beberapa kombinasi *Batch Size* dan *Learning Rate*. Adapun hasil dari fase kedua dapat dilihat pada tabel berikut.

Tabel 12. Hasil Fase Kedua (*Hyperparameter Tuning*)

No.	Experiment	Learning Rate	Batch Size	Roug eL	$\Delta\epsilon$ (<i>val</i> – <i>train</i> loss)	Epoch Stop
1	Baseline <i>Hyperparameter</i>	10^{-4}	6	0.40	0.60	15



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

No.	Experiment	Learning Rate	Batch Size	RougeL	$\Delta\epsilon$ $ (\text{val} - \text{train loss}) $	Epoch Stop
2	Memperkecil Learning Rate	10^{-5}	6	0.36	0.02	15
3	Memperbesar Learning Rate	10^{-3}	6	0.38	1.47	9
4	Memperkecil Batch Size	10^{-4}	2	0.40	0.72	11
5	Memperbesar Batch Size	10^{-4}	10	0.40	0.38	13

Berdasarkan hasil pengamatan pada tabel, variasi *learning rate* tidak menunjukkan perbedaan signifikan terhadap skor evaluasi *Rouge-L* yang berada pada kisaran 0,36–0,40. Namun, pengaruhnya terlihat jelas pada nilai $\Delta\epsilon$ ($|\text{validation loss} - \text{training loss}|$) dimana *validation loss* adalah perhitungan nilai *loss* dengan menggunakan 323 data validasi dan *training loss* menggunakan 6126 data latih (*train*). *Learning rate* yang lebih besar ($1e-3$) menghasilkan $\Delta\epsilon$ tertinggi sebesar 1,47 yang mengindikasikan gap *loss* besar serta melemahnya kemampuan generalisasi model. Sebaliknya, *learning rate* yang lebih kecil ($1e-5$) menghasilkan $\Delta\epsilon$ terendah (0,02) yang menunjukkan generalisasi lebih baik, tetapi tidak diikuti peningkatan performa (*Rouge-L* hanya 0,36), sehingga proses konvergensi cenderung lambat dan kurang optimal. *Learning rate* $1e-4$ memberikan keseimbangan terbaik dengan $\Delta\epsilon$ moderat dan skor *Rouge-L* stabil di 0,40.

Fenomena ini sesuai dengan persamaan umum *gradient descent* (1), di mana nilai *learning rate* yang terlalu besar menyebabkan pembaruan bobot menjadi terlalu agresif sehingga pelatihan tidak stabil, sedangkan *learning rate* yang



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

terlalu kecil menghasilkan pembaruan yang lebih halus namun memperlambat konvergensi. Hal ini juga didukung oleh jumlah *epoch* stop, di mana *learning rate* besar mempercepat konvergensi (misalnya berhenti pada *epoch* 7 untuk $1e-3$) tetapi dengan risiko instabilitas tinggi, sedangkan *learning rate* kecil membutuhkan *epoch* lebih banyak (hingga 15) tanpa peningkatan performa yang signifikan.

Pada variasi *batch size*, terlihat bahwa peningkatan ukuran *batch* cenderung menurunkan nilai $\Delta\epsilon$. *Batch size* 10 menghasilkan $\Delta\epsilon$ sebesar 0,38 yang menunjukkan pelatihan lebih stabil dan gap generalisasi lebih kecil. Sebaliknya, *batch size* 2 menghasilkan $\Delta\epsilon$ sebesar 0,72 yang mengindikasikan fluktuasi gradien lebih tinggi dan generalisasi yang lebih lemah. Meskipun demikian, seluruh variasi *batch size* tetap menghasilkan skor *Rouge-L* yang relatif sama (0,40), sehingga perbedaannya lebih terletak pada stabilitas pelatihan.

Temuan ini sesuai dengan persamaan gradien mini-*batch* (2), di mana ukuran *batch* yang kecil meningkatkan varians estimasi gradien sehingga pembaruan bobot menjadi lebih fluktuatif, sedangkan ukuran *batch* yang lebih besar menghasilkan estimasi gradien yang lebih stabil dan konsisten.

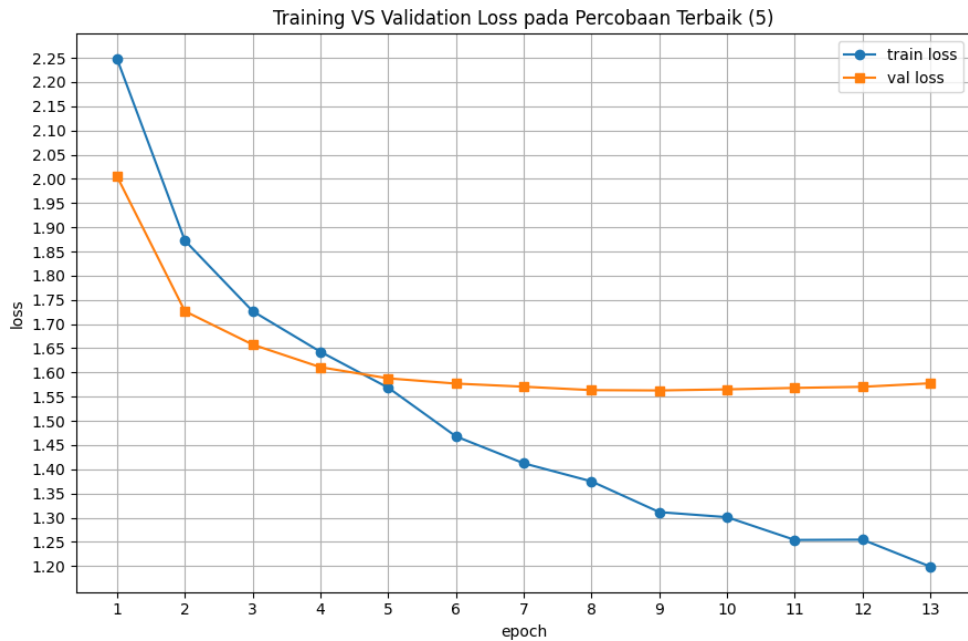
Dengan demikian, konfigurasi optimal diperoleh pada *learning rate* $1e-4$ dan *batch size* 10, karena menghasilkan skor *Rouge-L* tertinggi (0,40) dengan $\Delta\epsilon$ paling kecil (0,38), yang menunjukkan keseimbangan terbaik antara performa dan kemampuan generalisasi model.

Adapun *epoch*, jika berpacu kepada hasil proses *train* pada percobaan terbaik (ke-5), pelatihan berhenti pada *epoch* ke-13. Melalui gambar berikut, kita bisa melihat bahwa menambah *epoch* > 13 ataupun mengurangi *epoch* < 13 tidak membuat hasil pelatihan menjadi lebih baik.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Gambar 4.4 Grafik *Epoch-Loss Train* Percobaan Pertama

berdasarkan gambar tersebut, model mulai mencapai konvergensi pada *epoch* ke-9, yang ditandai dengan *validation loss* yang melandai dan pelan pelan meningkat hingga *epoch* ke-13. Sementara itu, *training loss* terus menurun sehingga $\Delta\epsilon$ semakin besar dan mengindikasikan potensi *overfitting*. Hal ini menunjukkan bahwa penambahan *epoch* setelah titik tersebut tidak lagi menurunkan *validation loss* secara signifikan dan justru berpotensi menurunkan kemampuan generalisasi model, sedangkan pengurangan jumlah *epoch* kemungkinan akan meningkatkan nilai *loss* pada kedua metrik.

Setelah melalui fase training, Model terlatih tersebut kemudian diuji dengan 340 data uji dan dengan parameter yang sama ketika model dilatih.

Tabel 13. Tabel Hasil Pengujian Fine-Tuned model dengan Data Uji (Test)

Parameter	Nilai
Rouge-L Average	0.38
BERTScore-F1	0.80

Berdasarkan nilai tersebut, hal ini mengindikasikan bahwa hasil ringkasan model T5 dengan *ground truth* memiliki kesamaan semantik yang

cukup tinggi sehingga model memiliki kemampuan dalam menghasilkan ringkasan yang tetap mempertahankan makna utama teks meskipun terjadi perbedaan dalam struktur kalimat atau pemilihan kata.

4.3.2 Implementasi Web

Setelah model terbaik didapatkan, maka model tersebut akan diintegrasikan dengan sistem web yang telah dibangun. Web dan model di-*deploy* dengan menggunakan platform *AWS Cloud* supaya dapat diakses secara publik guna mendukung pengujian sistem. Kode implementasi pengembangan web dapat diakses melalui *repository* berikut. <https://github.com/RandyAnanda2003/tokped-AI-summarizer-T5Engine> dan <https://github.com/RandyAnanda2003/tokped-summarizer-UI>.

Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

1. Halaman Utama Web Peringkasan Ulasan Tokopedia.

Review Summary Tokopedia
Powered by finetuned T5 AI 100% GRATIS

★ T5 AI - ANALISIS ULASAN TOKOPEDIA

Malas baca review, tapi takut barang bodong?

Masukkan URL produk Tokopedia dan dapatkan ringkasan ulasan yang jelas dan objektif dari ulasan-ulasan terbaru pembeli asli.

CARA MENGGUNAKAN

- 01 Salin URL Produk**
Buka produk di Tokopedia, lalu copy URL-nya dari address bar browser
- 02 Paste & Klik Generate**
Tempelkan URL di form, lalu klik tombol "Generate Summary"
- 03 Dapatkan Ringkasan**
Tunggu 1-3 menit dan ringkasan ulasan terbaru siap dibaca

Mulai analisis review sekarang
Masukkan URL produk Tokopedia untuk mendapatkan ringkasan objektif berdasarkan 40-60 ulasan pembeli terverifikasi terbaru.

URL PRODUK TOKOPEDIA

CONTOH URL

Ringkasan Ulasan
Dianalisis dari 8 ulasan pembeli

Secara umum pembeli sangat puas karena produk asli, berkualitas, dan cocok untuk gen.

- **Kualitas:** original, build quality mewah, fungsi normal, packaging rapi, ergonomis, sesuai deskripsi dan diharapkan awet.
- **Pengiriman:** pengiriman cepat dengan packing rapi dan seller responsif.

Summary berhasil dibuat dari 8 ulasan terverifikasi

KENAPA PAKAI TOOL INI?

- Cepat**
Analisis selesai dalam sekitar 1-3 menit
- AI Pintar**
Menggunakan teknologi T5 (Text-to-Text Transfer Transformer)
- 100% Gratis**
Tidak ada biaya tersembunyi, selamanya gratis
- Objektif**
Rangkuman dari 40-60 review pembeli terbaru yang terverifikasi

Dibuat sebagai tugas akhir | Teknik Informatika
Politeknik Negeri Jakarta 2026

Disclaimer: Tool ini tidak berafiliasi resmi dengan Tokopedia

Gambar 4.5 Halaman Utama Web

Sesuai dengan rancangan yang telah disusun, *website* ini terdiri dari satu halaman utama yang memuat seluruh komponen secara terintegrasi, yaitu informasi singkat mengenai sistem, *field input*, dan *field output*. Pada halaman tersebut, pengguna diberikan penjelasan bahwa tujuan *website* adalah menghasilkan ringkasan objektif



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

dari ulasan produk asli di Tokopedia. Selain itu, tersedia tiga langkah penggunaan yang ditampilkan secara sederhana, yaitu memasukkan URL produk Tokopedia, menekan tombol *generate*, lalu menunggu hingga sistem menampilkan hasil ringkasan. Sebagai ilustrasi, dimasukan contoh *input* berupa tautan produk *gaming controller* di Tokopedia dan sistem berhasil menghasilkan *output* berupa ringkasan dalam bentuk narasi berpoin seperti yang dapat dilihat pada *field output* pada gambar 4.5.

4.4 Pengujian Sistem

Setelah web dan model selesai dikembangkan, selanjutnya, Sistem ini kemudian diuji secara langsung untuk mengevaluasi kualitas sistem web dan hasil peringkasan model guna memastikan apakah sistem tersebut sudah layak digunakan. Selain itu, pengujian ini bertujuan untuk mengidentifikasi potensi kekurangan sebagai dasar evaluasi dan pengembangan lebih lanjut di masa depan.

4.4.1 Deskripsi Pengujian

Pengujian sistem ini terdiri dari 4 tipe pengujian, yaitu *Black box testing*, *System Usability Scale (SUS)*, *Net Promoter Score (NPS)*, dan pengujian oleh ahli. *Black box testing* dan pengujian oleh ahli dilakukan secara mandiri. Sedangkan *SUS* dan *NPS* dilakukan oleh pengguna akhir.

1. Black Box Testing

Guna memastikan setiap fitur pada antarmuka sistem berfungsi sebagaimana mestinya, maka disusunlah skenario pengujian *black box* untuk diuji secara manual seperti yang dapat dilihat pada tabel berikut.

Tabel 14. Skenario Pengujian Black Box

Kode Uji	Skenario Pengujian	Harapan/Ekspektasi
BB-01	Pengguna membuka halaman utama web	Sistem berhasil menampilkan halaman utama yang berisi informasi singkat, petunjuk penggunaan, dan kolom <i>input</i> URL produk



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Kode Uji	Skenario Pengujian	Harapan/Ekspektasi
BB-02	Pengguna memasukkan URL kosong lalu menekan tombol <i>generate</i>	Sistem menampilkan pesan <i>error</i> bahwa URL harus diisi dan proses tidak dilanjutkan
BB-03	Pengguna memasukkan URL yang bukan berasal dari Tokopedia	Sistem menampilkan pesan <i>error</i> bahwa URL tidak valid dan proses dihentikan
BB-04	Pengguna memasukkan format URL yang tidak sesuai (misalnya karakter acak)	Sistem menampilkan pesan <i>error</i> validasi URL dan tidak melakukan proses <i>scraping</i>
BB-05	Pengguna memasukkan URL produk Tokopedia yang valid	Sistem menerima <i>input</i> dan melanjutkan ke proses <i>scraping</i> ulasan sampai menghasilkan ringkasan di <i>field output</i> .
BB-06	Pengguna memasukkan URL Tokopedia yang tidak mengarah ke halaman suatu produk	Sistem menampilkan pesan <i>error</i> bahwa halaman tidak dapat diproses
BB-07	Pengguna memasukkan URL Produk Tokopedia yang tidak mempunyai ulasan atau kurang dari 10 ulasan.	Sistem menampilkan pesan <i>error</i> bahwa tidak ada ulasan produk yang bisa diringkaskan atau ulasan produk kurang dari 10 ulasan.
BB-08	Terjadi gangguan sistem saat proses berlangsung	Sistem akan menampilkan pesan <i>error</i> tanpa detail <i>error</i> kepada <i>enduser</i> .



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2. Pengujian Oleh Ahli

Pengujian oleh ahli Bahasa dilakukan untuk mengukur kemampuan model dalam menghasilkan sebuah ringkasan yang sesuai dengan kaidah peringkasan abstraktif. Dalam pengujian ini, ahli akan diberikan tumpukan ulasan beserta dengan hasil ringkasan yang dihasilkan oleh sistem web, lalu ahli akan menilai sesuai dengan parameter penilaian seperti yang dapat dilihat pada tabel berikut.

Tabel 15. Parameter Pengujian oleh Ahli

Kode Uji	Pernyataan	Nilai
POA -1	Model sudah cukup baik dalam membuat ringkasan abstraktif.	Penilaian dengan menggunakan skala <i>likert</i> 1-5. dengan aturan sebagai berikut.
POA-2	Ringkasan yang dihasilkan cukup representatif dengan kumpulan ulasan yang ada.	- 1 = Sangat Tidak Setuju - 2 = Tidak Setuju - 3 = Cukup Setuju - 4 = Setuju
POA-3	Apakah struktur bahasa ringkasan yang dihasilkan cukup baik sesuai dengan kaidah bahasa Indonesia .	- 5 = Sangat Setuju

3. *System Usability Scale*

Untuk mengukur tingkat kegunaan sistem, dilakukan pengujian menggunakan metode *System Usability Scale* (SUS). Pengujian ini bertujuan menilai kemudahan penggunaan dan pengalaman pengguna terhadap sistem. Adapun daftar pernyataan SUS yang digunakan dalam pengujian ditampilkan pada tabel berikut.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Tabel 16. Pertanyaan Pengujian SUS

Kode Uji	Pertanyaan
SUS-1	Saya berpikir akan menggunakan web ini lagi
SUS-2	Saya merasa web ini rumit untuk digunakan
SUS-3	Saya merasa web ini mudah untuk digunakan
SUS-4	Saya membutuhkan bantuan teknis dari orang lain agar dapat menggunakan web ini
SUS-5	Saya merasa fitur peringkasan pada web ini berjalan dengan semestinya
SUS-6	Saya merasa ada/banyak hal yang tidak konsisten baik dari segi informasi atau hasil ringkasan pada web ini
SUS-7	Saya merasa pengguna lain akan cepat memahami cara kerja dari web ini
SUS-8	Saya merasa web ini membingungkan bagi saya atau pengguna lain
SUS-9	Saya merasa tidak ada hambatan (<i>confident</i>) ketika menggunakan web ini
SUS-10	Saya perlu membiasakan diri terlebih dahulu sebelum dapat menggunakan web ini

Seluruh pernyataan SUS pada Tabel 4.13 menggunakan format jawaban yang seragam, yaitu skala *Likert* dengan rentang nilai 1 hingga 5. Skor 1 menunjukkan “Sangat Tidak Setuju”, sedangkan skor 5 menunjukkan “Sangat Setuju”

4. Net Promotor *Score*.

Untuk Mengetahui seberapa “*recommended*” *website* yang telah dibangun, berikut adalah pertanyaan NPS yang akan diujikan.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Tabel 17. Pernyataan NPS

Kode Uji	Pertanyaan
NPS-1	Seberapa besar kemungkinan Anda merekomendasikan <i>website</i> ini kepada orang lain ?

Format jawaban NPS pada Tabel 4.21 diisi dengan skala *likert* 1 sampai 10. Nilai 1 menyatakan “ Sangat Tidak Merekomendasikan”, sedangkan nilai 10 menyatakan “ Sangat Merekomendasikan”.

4.4.2 Prosedur Pengujian

Berbagai metode pengujian yang telah dijelaskan sebelumnya dilaksanakan melalui dua prosedur utama, yaitu pengujian mandiri, dan penyebaran kuesioner uji coba web. *Black box* testing dilakukan secara mandiri untuk memastikan seluruh fungsi sistem berjalan dengan baik sebelum digunakan oleh pengguna akhir.

Sementara itu, SUS, dan NPS digabungkan dalam satu kuesioner uji coba web yang diisi oleh pengguna akhir. Selain itu, Pengujian oleh ahli dilakukan dengan cara wawancara bersama dengan ahli bahasa, yaitu Ibu Farizka Humolungo selaku dosen bahasa Indonesia di Jurusan Administrasi Niaga, Politeknik Negeri Jakarta. Adapun target responden NPS dan SUS dari pengujian ini adalah pengguna *e-commerce* Tokopedia dari semua kalangan.

4.4.3 Hasil dan Analisis Data Pengujian

1. Pengujian Mandiri

- Black Box Testing

Pengujian mandiri meliputi uji *black box* dengan beberapa skenario yang telah disebutkan, skenario dianggap berhasil jika hasil sesuai dengan harapan dan gagal jika hasil tidak sesuai dengan harapan. Hasil pengujian *black box* dapat dilihat pada *table* berikut



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Tabel 18. Hasil Uji Black Box

Kode Uji	Skenario Pengujian	Harapan/Ekspektasi	Hasil
BB-01	Pengguna membuka halaman utama web.	Sistem berhasil menampilkan halaman utama yang berisi informasi singkat, petunjuk penggunaan, dan kolom <i>input</i> URL produk.	Sesuai
BB-02	Pengguna memasukkan URL kosong lalu menekan tombol <i>generate</i> .	Sistem menampilkan pesan <i>error</i> bahwa URL harus diisi dan proses tidak dilanjutkan.	Sesuai
BB-03	Pengguna memasukkan URL yang bukan berasal dari Tokopedia.	Sistem menampilkan pesan <i>error</i> bahwa URL tidak valid dan proses dihentikan.	Sesuai
BB-04	Pengguna memasukkan format URL yang tidak sesuai (misalnya karakter acak).	Sistem menampilkan pesan <i>error</i> validasi URL dan tidak melakukan proses <i>scraping</i> .	Sesuai
BB-05	Pengguna memasukkan URL produk	Sistem menerima <i>input</i> dan melanjutkan ke proses <i>scraping</i> ulasan	Sesuai



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Kode Uji	Skenario Pengujian	Harapan/Ekspektasi	Hasil
	Tokopedia yang valid.	sampai menghasilkan ringkasan di <i>field output</i> .	
BB-06	Pengguna memasukkan URL Tokopedia yang tidak mengarah ke halaman suatu produk.	Sistem menampilkan pesan <i>error</i> bahwa halaman tidak dapat diproses.	Sesuai
BB-07	Pengguna memasukkan URL Produk Tokopedia yang tidak mempunyai ulasan atau kurang dari 5 ulasan.	Sistem menampilkan pesan <i>error</i> bahwa tidak ada ulasan produk yang bisa diringkas atau ulasan produk kurang dari 5 ulasan.	Sesuai
BB-08	Terjadi gangguan sistem saat proses berlangsung.	Sistem akan menampilkan pesan <i>error</i> tanpa detail <i>error</i> kepada <i>end user</i> .	Sesuai

Berdasarkan hasil *black box* testing pada Tabel 4.21, seluruh skenario berjalan sesuai ekspektasi sehingga sistem mencapai kelayakan fungsional 100% sesuai rancangan. Namun, pengujian ini hanya mencakup skenario



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

yang telah ditentukan, sehingga masih ada kemungkinan munculnya kondisi tak terduga saat digunakan sehari-hari oleh *end user*.

- Pengujian Oleh Ahli.

Guna memvalidasi hasil ringkasan ulasan yang dihasilkan oleh model, dilakukan pengujian linguistik yang dikaji oleh Bapak Fulan bin fulan selaku dosen mata kuliah bahasa Indonesia Jurusan Administrasi Niaga, Politeknik Negeri Jakarta. Adapun hasil pengujian oleh *expertise* dapat dilihat pada tabel berikut.

Tabel 19. Hasil Pengujian Oleh Ahli Bahasa

Kode Uji	Skenario Pengujian	Nilai
POA -1	Model sudah cukup baik dalam membuat peringkasan abstraktif	5
POA-2	Ringkasan yang dihasilkan cukup representatif dengan teks aslinya	5
POA-3	Struktur bahasa ringkasan yang dihasilkan cukup baik sesuai dengan kaidah bahasa Indonesia	3

Seperti yang ditunjukkan pada Tabel 4.16, hasil pengujian oleh ahli bahasa menunjukkan bahwa ringkasan abstraktif yang dihasilkan oleh model AI memiliki kualitas yang cukup baik, baik dari aspek kesesuaian isi (*faithfulness*) maupun dari segi struktur kalimat dalam bahasa Indonesia. hanya saja, ada sedikit catatan dari ahli bahwa dalam beberapa kali percobaan terdapat redundansi frasa dalam satu kalimat yang memiliki maksud yang sama.



2. Pengujian Berbasis *User* (Kuesioner)

Setelah melakukan penyebaran kuesioner melalui media sosial dan relasi pribadi, diperoleh data hasil pengujian dari total 46 responden. Seluruh responden tersebut merupakan pengguna *e-commerce* Tokopedia dan merupakan pengguna *e-commerce* tokopedia.

- *System Usability Scale* (SUS)

Tabel 20. Hasil SUS

KODE	SUS 1	SUS 2	SUS 3	SUS 4	SUS 5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10
R1	3	1	5	2	5	1	5	1	5	2
R2	5	2	4	1	4	2	5	2	4	3
R3	4	4	4	4	4	5	4	4	5	5
R4	5	1	5	1	4	2	4	2	4	3
R5	5	3	4	1	5	3	2	3	2	5
R6	4	3	3	3	4	4	3	2	3	4
R7	5	2	5	2	4	2	5	2	2	2
R8	5	2	5	1	5	3	4	2	4	4
R9	5	5	5	5	5	5	5	5	5	5
R10	5	4	5	1	4	4	4	1	4	1
R11	5	1	5	1	5	1	5	1	5	1
R12	4	1	4	1	5	1	5	1	5	1
R13	3	1	5	1	3	3	4	2	5	2
R14	5	1	5	4	5	1	5	1	5	2
R15	2	1	5	1	4	3	4	1	4	2
R16	4	1	5	1	3	1	5	1	1	1
R17	5	1	5	1	5	2	4	2	3	2

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

KODE	SUS 1	SUS 2	SUS 3	SUS 4	SUS 5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10
R18	4	1	5	2	5	2	4	2	5	3
R19	5	1	5	1	5	1	5	1	5	1
R20	4	2	5	1	4	2	5	2	4	2
R21	4	2	4	1	5	3	4	2	4	4
R22	5	2	4	2	4	2	4	2	4	3
R23	5	5	5	1	5	1	5	1	5	1
R24	5	1	5	1	5	2	5	1	5	1
R25	5	1	4	1	5	1	4	1	5	1
R26	5	1	5	2	5	1	5	1	5	2
R27	5	1	5	1	5	1	5	1	4	2
R28	3	1	5	1	5	1	5	1	5	1
R29	3	2	3	4	3	2	3	3	3	5
R30	5	1	5	2	5	2	4	2	5	2
R31	5	2	5	2	5	2	4	2	1	2
R32	4	2	5	3	4	3	4	2	4	3
R33	4	1	5	1	5	2	5	2	5	2
R34	4	1	5	1	4	1	5	1	5	1
R35	5	2	5	2	5	5	5	2	5	5
R36	4	1	5	2	4	2	4	2	5	2
R37	4	2	4	2	4	2	4	2	4	3
R38	5	1	5	5	5	5	5	1	5	2
R39	4	1	5	1	5	1	5	1	4	3
R40	5	5	5	5	5	5	5	5	5	5



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

KODE	SUS 1	SUS 2	SUS 3	SUS 4	SUS 5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10
R41	4	5	5	5	5	5	5	5	5	5
R42	5	1	5	3	5	5	5	5	5	2
R43	4	1	5	1	4	1	4	1	5	1
R44	5	1	5	1	5	1	5	1	5	1
R45	5	1	5	3	5	1	5	2	5	3
R46	2	1	5	1	2	4	5	2	5	2
R47	2	4	2	4	4	4	2	4	2	2

Nilai hasil diatas dapat dihitung menjadi nilai SUS dengan menggunakan persamaan sebagai berikut.

$$SUS \text{ Ganjil} = (SUS1 + SUS3 + SUS5 + SUS7 + SUS9) - 5$$

$$SUS \text{ Genap} = 25 - (SUS2 + SUS4 + SUS6 + SUS8 + SUS10)$$

$$SUS = (SUS_{Ganjil} + SUS_{Genap}) \times 2.5$$

Dengan menggunakan persamaan diatas, maka didapat nilai SUS tiap-tiap responden seperti pada tabel 4.19 berikut

Tabel 21. Nilai SUS Tiap Responden

KODE	Nilai SUS
R1	90
R2	80
R3	47.5
R4	82.5
R5	57.5
R6	52.5



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

KODE	Nilai SUS
R7	77.5
R8	77.5
R9	50
R10	77.5
R11	100
R12	95
R13	77.5
R14	90
R15	77.5
R16	82.5
R17	85
R18	82.5
R19	100
R20	82.5
R21	72.5
R22	75
R23	90
R24	97.5
R25	95
R26	95
R27	95



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

KODE	Nilai SUS
R28	95
R29	47.5
R30	87.5
R31	75
R32	70
R33	90
R34	95
R35	72.5
R36	82.5
R37	72.5
R38	77.5
R39	90
R40	50
R41	47.5
R42	72.5
R43	92.5
R44	100
R45	87.5
R46	72.5
R47	35
Rata-Rata	78.67



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Berdasarkan hasil perhitungan *System Usability Scale* (SUS) yang disajikan pada Tabel 4.19, diperoleh nilai rata-rata sebesar 78,67. Nilai tersebut termasuk dalam kategori *Good* dengan grade “B”. Hal ini menunjukkan bahwa sistem web peringkasan ulasan produk otomatis yang dikembangkan telah berfungsi dengan baik serta dinilai layak untuk digunakan.

- *Net Promoter Score* (NPS)

Tabel 22. Perhitungan Nilai NPS

	1	2	3	4	5	6	7	8	9	10
NPS	0	0	1	1	0	2	6	9	14	14
Kategori	DETRACTOR						PASSIVE		PROMOTER	
Presentase	9%						32%		60%	
Nilai NPS	PROMOTER – DETRACTOR = 51%									

Berdasarkan hasil perhitungan *Net Promoter Score* (NPS) yang disajikan pada Tabel 4.20, diperoleh nilai sebesar 51% untuk web klasifikasi lirik lagu. Nilai tersebut termasuk dalam kategori NPS *Great/Excellent*. Hal ini menunjukkan bahwa secara umum pengguna merasa puas terhadap sistem yang dikembangkan. Selain itu, hasil ini juga menjadi indikasi positif bahwa sebagian besar pengguna memiliki kecenderungan untuk merekomendasikan web tersebut kepada orang lain.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB 5

PENUTUP

5.1 Kesimpulan

Penelitian ini berhasil mengimplementasikan model IndoT5 untuk meringkas ulasan-ulasan produk di *e-commerce* Tokopedia secara otomatis melalui aplikasi berbasis web dengan metode *fine-tuning* pada *pre-trained* model indoT5. Hasil pelatihan menunjukkan bahwa kualitas ringkasan yang dihasilkan sudah mampu merepresentasikan isi ulasan dengan cukup baik, baik dari segi kesesuaian informasi maupun kedekatan makna dengan ringkasan acuan (*Ground Truth*).

Hal tersebut diperkuat dengan justifikasi ahli bahasa. berdasarkan penilaian ahli bahasa Indonesia, ringkasan yang dihasilkan dinilai memiliki tingkat keabstrakan yang baik, tetap mempertahankan kesesuaian isi, serta menggunakan struktur bahasa yang jelas dan mudah dipahami.

Sementara itu, dari sisi performa sistem, aplikasi web yang dikembangkan juga menunjukkan tingkat kegunaan dan *behaviour* fitur yang baik serta mendapatkan respons positif dari pengguna sehingga secara keseluruhan sistem dapat dikatakan berhasil, layak digunakan, dan mampu memberikan ringkasan yang berkualitas.

5.2 Saran

Berdasarkan hasil penelitian yang didapatkan, berikut beberapa saran yang dapat dipertimbangkan atau diterapkan pada pengembangan sistem ke depannya :

1. Peningkatan kualitas dan kapabilitas model

Pengembangan selanjutnya dapat mempertimbangkan penggunaan model *small language* model (SLM) yang lebih mutakhir dan efisien, seperti varian model berbasis T5 terbaru atau model ringan lain seperti Gemma yang telah di-*fine-tune* untuk tugas *text summarization*. Penggunaan model yang lebih canggih diharapkan mampu meningkatkan kualitas ringkasan, baik dari sisi pemahaman konteks maupun koherensi hasil ringkasan, terutama jika didukung dengan sumber daya komputasi yang lebih memadai.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2. Memperbanyak jumlah, memperbaiki struktur bahasa, dan variasi *dataset*

Jumlah *dataset* yang digunakan dalam penelitian ini masih relatif terbatas jika dibandingkan dengan *dataset* standar seperti IndoSum yang memiliki jumlah data jauh lebih besar. Oleh karena itu, penambahan jumlah *dataset*, baik dari segi kuantitas maupun variasi data, sangat disarankan untuk meningkatkan kemampuan generalisasi model. Selain itu, peningkatan *dataset* ini juga sejalan dengan kebutuhan model yang lebih kompleks agar dapat dilatih secara optimal.

3. Pengembangan metode *hybrid* (ekstraktif–abstraktif)

Penelitian selanjutnya dapat mengembangkan pendekatan *hybrid summarization*, yaitu dengan mengombinasikan metode ekstraktif dan abstraktif. Sebagai contoh, proses awal dapat dilakukan menggunakan metode ekstraktif seperti TF-IDF atau model berbasis IndoBERT untuk memilih kalimat-kalimat penting, kemudian hasilnya diringkas kembali menggunakan model abstraktif seperti T5. Pendekatan ini diharapkan dapat meningkatkan presisi dan relevansi informasi yang dihasilkan dalam ringkasan.

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR PUSTAKA

- Aipina, D. and Witriyono, H. (2022) ‘Pemanfaatan *Framework* Laravel Dan *Framework* Bootstrap Pada Pembangunan Aplikasi Penjualan Hijab Berbasis Web’, *Jurnal Media Infotama*, 18(1), p. 2022. <https://doi.org/10.37676/jmi.v18i1.1836>
- Alfarizi, M. R. S., Al-farish, M. Z. ., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. . (2023). Penggunaan Python Sebagai Bahasa Pemrograman untuk *Machine Learning* dan *Deep Learning*. *Karimah Tauhid*, 2(1), 1–6. <https://doi.org/10.30997/karimahtauhid.v2i1.7518>
- Amal, S. and Kusnadi, K. (2022) ‘Perancangan ERP Menu Hr-Training Berbasis Odoo Menggunakan Metode SDLC Studi Kasus PT.XYZ’, *Remik*, 6(3), pp. 426–435. <https://doi.org/10.33395/remik.v6i3.11612>
- Ardyanti, C. P. R., Wibisono, Y., & Megasari, R. (2024). Peringkasan teks berita berbahasa Indonesia menggunakan LSTM dan *Transformer*. *IJAi (Indonesian Journal of Applied Informatics)*, 8(2). <https://doi.org/10.14421/jiska.2025.10.1.110-121>
- Ardyanti, C. P. Rosaria., Wibisono, Y., & Megasari, R. (2024). Peringkasan teks berita berbahasa Indonesia menggunakan LSTM dan *Transformer*. *Indonesian Journal of Applied Informatics*, 8(2). <https://doi.org/10.20961/ijai.v8i2.78856>
- Ashari, N., Darwis, D. and Kisworo, K. (2023) ‘*Game* Edukasi Pengenalan Dampak Buruk Merokok Bagi Kesehatan Berbasis Android’, *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 4(1), pp. 22–28. <https://doi.org/10.33365/jatika.v4i1.2455>
- Bahari, A., & Dewi, K. E. (2024). Peringkasan teks otomatis abstraktif menggunakan *Transformer* pada teks bahasa Indonesia. *Komputa: Jurnal Ilmiah Komputer dan Informatika*, 13(1), 83–91. <https://doi.org/10.34010/komputa.v13i1.11197>
- Charisma, U., Nazaruddin, N., Suherman, S., & Hamdy, M. I. (2025). Analisis tingkat loyalitas konsumen terhadap *e-commerce* Shopee menggunakan



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

- metode Net Promoter Score. *INTECOMS: Journal of Information Technology and Computer Science*, 8(4). <https://doi.org/10.31539/k19hh659>
- Eriya, E. (2020) ‘Sistem Manajemen Inventaris Laboratorium Otomatis Menggunakan Barcode’, *Multinetics*, 6(2), pp. 149–156. <https://doi.org/10.32722/multinetics.v6i2.3305>
- FastAPI Documentation. (2025). In FastAPI. Retrieved December 28, 2025, from <https://fastapi.tiangolo.com/>
- Hafidz, K., Irawan, M. D. and Nawar, H. D. (2022) ‘Sistem Penginputan Data Bahan Pokok pada Pasar Tradisional Sumatera Utara Berbasis Website di Disperindag Sumut’, *sudo Jurnal Teknik Informatika*, 1(3), pp. 98–107. <https://doi.org/10.56211/sudo.v1i3.27>
- Hartawan, G., Maylawati, D. S., & Uriawan, W. (2024). *Bidirectional and Auto-Regressive Transformer (BART) for Indonesian Abstractive Text Summarization*. *Jurnal Informatika Polinema*, 10(4), 535–542. <https://doi.org/10.33795/jip.v10i4.5242>
- Hildayanti, s. K., utami, e. Y., noor, l. S., elizabeth, r., & munizu, m. (2023). Pengaruh sosial media *marketing* dan *online customer review* terhadap minat beli melalui aplikasi shopee. *Jurnal ilmiah edunomika*, 8(1). <https://doi.org/10.29040/jie.v8i1.11201>
- Ivanedra, K., & Mustikasari, M. (2025). Implementasi metode *recurrent neural network* pada *text summarization* dengan teknik abstraktif. *Jurnal Teknologi Informasi dan Ilmu Komputer*. <https://doi.org/10.25126/jtiik.2019641067>
- Jamal, S., & Kusnadi, K. (2022). Perancangan ERP Menu HR-Training berbasis Odoo menggunakan metode SDLC: Studi kasus PT. XYZ. *Remik*, (3), 426–435. <https://doi.org/10.33395/remik.v6i3.11612>
- Karo Karo, I. M., Dewi, S., & Perdana, A. (2024). Implementasi *text summarization* pada *review* aplikasi digital *library System* menggunakan metode *maximum marginal relevance*. *JEKIN: Jurnal Teknik Informatika*, 4(1). <https://doi.org/10.58794/jekin.v4i1.671>



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

- Li, T., Huang, R., Cui, C., Towey, D., Ma, L., Li, Y.-F., & Xia, W. (2024). A survey on web application testing: A decade of evolution. arXiv. <http://arxiv.org/abs/2412.10476>
- Muharam, A. F., Gerhana, Y. A., Maylawati, D. S., Ramdhani, M. A., & Abdul Rahman, T. K. (2025). Enhancing abstractive multi-document summarization with Bert2Bert model for Indonesian language. JISKA, 10(1), 110–121. <https://doi.org/10.14421/jiska.2025.10.1.110-121>
- Nasution, S., Ferdiana, R., & Hartanto, R. (2025). Towards two-step fine-tuned abstractive summarization for low-resource language using transformer T5. International Journal of Advanced Computer Science and Applications, 16(2), Article 120. <https://doi.org/10.14569/IJACSA.2025.01602120>
- Nurbaiti, E., Rismawan, A. N., Fuadah, S. L., Praya, M. H., Syifa, B. N., & Nugraha, J. T. (2025). Analisis pengaruh ulasan dalam bisnis *online* untuk pembelian barang. *Journal of Governance and Public Administration*, 2(2), Maret. <https://doi.org/10.70248/jogapa.v2i2.1701>
- Nurinayah, A., & Haviana, S. F. C. (2025). Implementasi *fine-tuning* untuk prediksi solusi dari kalimat masalah pada artikel ilmiah menggunakan *large language models* (LLM). Jurnal Transistor Elektro dan Informatika (TRANSISTOR EI), 7(1), 10–17. <https://jurnal.unissula.ac.id/index.php/EI/article/view/44340>
- Pramudita, A. S. (2024). *How reviews and ratings influence consumer purchase intentions* in Indonesian e-commerce. Jurnal Logistik Bisnis, 14(2). <https://doi.org/10.46369/logistik.v14i2.3924>
- Pratama, S., Iswandi, I., Sevtian, A., & Anjani, T. P. (2023). Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma C4.5 dengan CRISP-DM. *Journal of Applied Informatics and Computing*, 7(1), 16–20. <https://doi.org/10.30871/jaic.v7i1.4998>
- Puspitasari, W., Ramdani, D., & Maulana, A. M. (2025). IndoT5 (*Text-to-Text Transfer Transformer*) algorithm for paraphrasing Indonesian language Islamic sermon manuscripts. *Khazanah Journal of Religion and Technology*, 2(2), 1093. <https://doi.org/10.15575/kjrt.v2i2.1093>



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

- Putri, N., & Hadi, E. D. (2024). Influence of online customer reviews and ratings on purchasing decisions on Shopee ecommerce with trust as mediator and emotional as moderation. *Student Journal of Business and Management*, 7(2), 149–162. <https://doi.org/10.33369/sjbm.v7i2.36329>
- PyTorch. (2024). *PyTorch: An open source machine learning framework*. <https://pytorch.org>
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). *Exploring the limits of transfer learning with a unified text-to-text transformer*. *Journal of Machine Learning Research*, 21(140), 1–67. <http://jmlr.org/papers/v21/20-074.html>
- Ramdany, S. (2024) ‘Penerapan UML *Class Diagram* dalam Perancangan Sistem Informasi Perpustakaan Berbasis Web’, *Journal of Industrial and Engineering System*, 5(1). <https://doi.org/10.31599/2e9afp31>
- Rifathurrahman, R. F., Chaniago, H., & Efawati, Y. (2024). *The effect of e-service quality on repurchase interests of the online shopping Shopee Indonesia*. *Jurnal Ekonomika dan Manajemen*, 13(1), 8–25. <https://doi.org/10.36080/jem.v13i1.2846>
- Setiawansyah, S., Lestari, D. T. and Megawaty, D. A. (2022) ‘Sistem Informasi Pkk Berbasis *Website Menggunakan Framework Codeigniter* (Studi Kasus: Kampung Purwoejo)’, *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 3(2), pp. 244–253. <https://doi.org/10.33365/jatika.v3i2.2031>
- Setiyani, L. and Tjandra, E. (2021) ‘Analisis Kebutuhan Fungsional Aplikasi Penanganan Keluhan Mahasiswa Studi Kasus: Stmik Rosma Karawang’, *Jurnal Inovasi Pendidikan dan Teknologi Informasi (JIPTI)*, 2(1), pp. 8–17. <https://doi.org/10.52060/pti.v2i01.465>
- Sukma, A. P., Yusuf, R., & Dai, R. H. (2023). Analisis pengukuran *usability* Sistem Informasi Manajemen Baznas (SIMBA) menggunakan metode *System Usability Scale (SUS)*. *Diffusion: Journal of Systems and Information Technology*, 3(2), 224–231. <https://doi.org/10.37031/diffusion.v3i2.21342>



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

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://arxiv.org/abs/1706.03762>

Wahdiniawati, S. A., Apriani, A., & Orlando, A. A. (2024). *Digital transformation and e-commerce growth: Impact on consumer behavior and SMEs*. *Dinasti International Journal of Economics, Finance & Accounting*, 5(5). <https://doi.org/10.38035/dijefa.v5i5.3438>

Wibisono, D., Ulum, F. and Megawaty, D. A. (2022) ‘Pengukuran Kualitas Sistem Informasi Kasir Terhadap Layanan Konsumen Menggunakan Metode *CrispDm* Variabel *Servqual* (Studi Kasus: Studio Foto *Archa*)’, *Jurnal Teknologi dan Sistem Informasi (JTISI)*, 3(4). <http://jim.teknokrat.ac.id/index.php/JTISI>.

Yuliska, & Syaliman, K. U. (2020). *Literatur review terhadap metode, aplikasi, dan dataset peringkasan dokumen teks otomatis untuk teks berbahasa Indonesia*. *IT Journal Research and Development (ITJRD)*, 5(1), 19–31. [https://doi.org/10.25299/itjrd.2020.vol5\(1\).4688](https://doi.org/10.25299/itjrd.2020.vol5(1).4688)

Shazeer, N., & Stern, M. (2018). *Adafactor: Adaptive learning rates with sublinear memory cost*. *arXiv*. <https://doi.org/10.48550/arXiv.1804.04235>

Qian, X., & Klabjan, D. (2020). *The impact of the mini-batch size on the variance of gradients in stochastic gradient descent*. *arXiv*. <https://arxiv.org/abs/2004.13146>

Mardatillah, U., Budiawan Zulfikar, W., Rialdy Atmadja, A., Taufik, I., & Uriawan, W. (2021, August). *Citation analysis on scientific articles using cosine similarity*. In **2021 7th International Conference on Wireless and Telematics (ICWT)** (pp. 1–6). IEEE. <https://doi.org/10.1109/ICWT52862.2021.9678402>



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Mehrafarin, H., Rajaei, S., & Pilehvar, M. T. (2022). *On the importance of data size in probing fine-tuned models*. In *Findings of the Association for Computational Linguistics: ACL 2022* (pp. 239–249). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.findings-acl.20>

Vanesha, N. A., Rizky, & Purwanto, A. (2024). Comparison between usability and user acceptance testing on educational game assessment. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 13(2), 210–215. <https://doi.org/10.32736/sisfokom.v13i2.2099>





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR RIWAYAT HIDUP PENULIS



Randy Aulia Ananda Ruslani

Lahir di Jakarta, 11 Mei 2003. Anak Sulung dari 3 bersaudara Keluarga Ade Ruslani dan Siti Nurjayanti. Lulus dari SD Dharma Karya UT pada tahun 2015, MTsN 3 Jakarta pada tahun 2018, dan MAN 11 Jakarta pada tahun 2021. Saat ini sedang menempuh studi di Program Sarjana Terapan Teknik Informatika di Politeknik Negeri Jakarta. Selama berkuliah, Penulis tertarik dan mendalami bidang *artificial intelligence*, Data, dan *backend engineering*.

**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

LAMPIRAN

Lampiran I : Berita Acara Kegiatan Demo Model dan Penilaian Hasil Ringkasan Model bersama dengan Ahli Linguistik



KEMENTERIAN PENDIDIKAN TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp: (021) 91274097, Fax : (021) 7863531, (021) 7270036 Hunting
Laman : <http://www.pnj.ac.id>, e-mail : tik.pnj@gmail.com

L..

BERITA ACARA

Pada hari ini Jumat, Tanggal 10 April 2026, telah dilaksanakan sesi wawancara dan validasi data hasil penelitian bersama :

Nama : Farizka Humolungo, S.Pd., M.A.

NIP : 199103022023212042

Jabatan : Dosen, KPS D4 BISPRO PNJ.

Bidang Keahlian : Linguistik, Bahasa Indonesia, dan Bahasa Inggris.

Sesi ini dilakukan dalam rangka validasi hasil peringkasan teks ulasan produk otomatis yang dihasilkan oleh model IndoT5 (*Text-to-Text Transfer Transformer*) yang telah melalui proses *fine-tuning* oleh peneliti. Adapun kegiatan yang dilakukan selama sesi berlangsung dapat dilihat pada poin-poin berikut.

1. Peneliti menjelaskan secara umum penelitian yang dilakukan sebagai upaya penyamaan konteks dan pemahaman antara peneliti dan ahli linguistik.
2. Ahli linguistik melakukan pengujian terhadap hasil ringkasan abstraktif yang dihasilkan oleh model, meliputi penilaian terhadap tingkat representatif ringkasan terhadap kumpulan ulasan asli suatu produk, serta evaluasi terhadap struktur bahasa Indonesia yang digunakan dalam ringkasan.
3. Ahli linguistik memberikan penilaian kuantitatif menggunakan skala *Likert* (1–5) berdasarkan parameter pengujian yang tercantum pada lembar lampiran ke-2.

Berdasarkan hasil wawancara dan validasi tersebut, ahli linguistik menyatakan bahwa hasil peringkasan model IndoT5 yang telah *di-fine-tuning* oleh peneliti dinyatakan Berhasil sebagai ringkasan abstraktif terhadap kumpulan ulasan produk di *e-commerce*.

Demikian Berita Acara ini dibuat dengan sebenarnya untuk dapat digunakan sebagaimana mestinya.

Depok, 10 April 2026

Pihak Pertama (Peneliti)  Randy Aulia Ananda Ruslani NIM. 2207411041	Pihak Kedua (Ahli Linguistik/Bahasa)  Farizka Humolungo, S.Pd., M.A. NIP. 199103022023212042
--	---



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



KEMENTERIAN PENDIDIKAN TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp: (021)91274097, Fax : (021) 7863531, (021) 7270036 Hunting
Laman : <http://www.pnj.ac.id>, e-mail : tik.pnj@gmail.com

L..

LEMBAR PARAMETER PENGUJIAN HASIL RINGKASAN

Kode Uji	Pernyataan	Nilai
POH -1	Model sudah cukup baik dalam membuat ringkasan abstraktif.	Penilaian dengan menggunakan skala likert 1-5. dengan aturan sebagai berikut.
POH-2	Ringkasan yang dihasilkan cukup representatif dengan kumpulan ulasan yang ada.	- 1 = Sangat Tidak Setuju - 2 = Tidak Setuju - 3 = Cukup Setuju - 4 = Setuju - 5 = Sangat Setuju
POH-3	Apakah struktur bahasa ringkasan yang dihasilkan cukup baik sesuai dengan kaidah bahasa Indonesia .	

Kode Uji	Skenario Pengujian	Nilai
POH -1	Model sudah cukup baik dalam membuat peringkasan abstraktif	5
POH-2	Ringkasan yang dihasilkan cukup representatif dengan teks aslinya	5
POH-3	Struktur bahasa ringkasan yang dihasilkan cukup baik sesuai dengan kaidah bahasa Indonesia	3



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Lampiran II : Sample dataset yang digunakan model dalam proses fine tune task summarization.

source	target
ok trims sukses. trimakasih barang dah sampai tapi maaf sejauh pengalaman ini respon nya lama. terima kasih barangnya sudah say terima bagus. barang sdh diterima kemasannya rapih barangnya oke. ok barangnya cuma pengirimannya kurang cepet. mantap trimakasih. thanks ya barang aman oke original. paket dah sampe tadi siang packing rapi unit sesuai harga mantap. trimakasih murmer dan toko pedia barang datang nya cepat barang sesuai gambar sukses terus murmer dan toko pedia. paket udh sampai mbak sesuai pesanan saran untuk kedepanx mbak jamx tlog dipakai bantal mbak biar lebih aman mksih. barang ok bngt sukses terus. barang dah mendarat dengan selamat gan hatur nuhun. jamnya bagus sesuai gambar thank gan smga sukses ya. jam lasebo nya keren sudah ane terima gan set digital analog hari tanggal alarm gmpang gan harga bersaing sama lapak yg sebelah ane suka tunggu next order gan. sesuai harapan jamnya bagus dan harganya juga oke banget semoga bisa jadi langganan. good product good packing recommended seller. barang sudah diterima dg selamat bungkus jg bagus pelayanan cepat rekomended banget toko sis ini diperbanyak lagi barangnya sis jadi saya akan jadi pelanggan tetap. thx gan. barang udah diterima sis makasih. thx baru sampai agak lama tapi bagus barangnya	produk ini mendapatkan banyak pujian karena barang sesuai gambar, fungsi jam lengkap seperti digital dan analog serta fitur tanggal dan alarm, packing rapi dan harga terjangkau sehingga banyak pembeli merekomendasikan dan menjadi pelanggan tetap. Beberapa mengeluh respon penjual dan pengiriman kadang lambat dan meminta pengemasan lebih aman, namun mayoritas puas dengan kualitas dan barang tiba selamat.

POLITEKNIK
NEGERI
JAKARTA



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

pesan size datang ga bisa dipake. sepatunya sesuai foto cuma lem nya kurang rapih pengiriman cepat kurir baik dan ramah. sepatunya bagus anak bujang suka size nya kecil tdk sama dgn size sepatu sekolahnya. model sporty banget sampe yg unboxing anak saya warna dan size sesuai pesanan kualitas bagus anak saya suka sepatunya buat latihan di klub basketnya kurir nya cepat dan ramah di wa langsung dtg padahal paket baru selesai disortir. sepatunya sudah diterima produknya bagus mudah an muat sama anak ku. barang datang langsung unboxing sepatunya bagus ukurannya pas sama anak saya. agak kecewa sepatu yang bagian depan rusak lemnnya kurang rapi mau copot. bentuk dan ukuran pas warna biru cakep. sepatunya bagus bgt sesuai expektasi smoga awet yaa pokoknya jadi langganan deh. sesuai di gmbarnya pas dipakai. biasa pake nomer naik ke pas alhamdulillah. masyallah bagus warnanya cantik. bentuk dan ukuran mantul real pack anakku sangat suka harga terjangkau. barang sudah sampai sesuai pesanan kk trimakasih bagus bgt. warna dan ukuran sudah sesuai untuk kualitas juga ssesuai. brng sdh sampai tetapi ukuran ny kekecilan kk tpi gk pp nanti msu sy tukar kn. bahan nya bagus sesuai sama foto ga kecewa amanah pokonya. sepatu sesuai pesanan muat di kaki anak ku umur thn insole cm. barangnya bagus tapi pesen ukuran yg datang ukuran. ukuran pas kualitas oke nyaman di kaki anak saya

produk ini cukup disukai pembeli karena modelnya sporty, warna dan gambar sesuai foto, bahan terasa oke dan anak-anak nyaman memakainya untuk sehari-hari atau latihan. Pengiriman sering cepat dengan kurir ramah dan harga terbilang terjangkau, namun ada keluhan soal ukuran yang kadang kekecilan atau tidak konsisten dan beberapa paket datang dengan lem kurang rapi atau kerusakan di bagian depan sehingga perlu penukaran. Secara keseluruhan mayoritas pembeli puas dan merekomendasikan produk ini dengan harapan kualitasnya tahan lama.

produk ini mendapat ularan positif karena

**POLITEKNIK
NEGERI
JAKARTA**



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Lampiran III : Berita Acara Kegiatan validasi sub-sampel *dataset* bersama dengan Ahli Linguistik.

Bukti Digisigner :

<https://www.digisigner.com/online/showDocument?documentId=84d1bad6-fb0b-485d-a809-7103a5b80ada&invitationId=3f2f6720-3c4e-4a6b-9695-9347d6d7f2dc>



KEMENTERIAN PENDIDIKAN TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp: (021)91274097, Fax : (021) 7863531, (021) 7270036 Hunting
Laman : <http://www.pnj.ac.id>, e-mail : tik.pnj@gmail.com

L3

BERITA ACARA

Pada hari ini, Senin, Tanggal 29 Juni 2026, telah dilaksanakan sesi validasi sampel *dataset* penelitian bersama :

Nama : Kezia Kathryn, M.Hum

Bidang Keahlian : Linguistik/Sastra Bahasa Indonesia

Sesi ini dilakukan dalam rangka uji kelayakan sampel *dataset* untuk melatih model T5 dalam tugas peringkasan teks ulasan produk secara abstraktif. Adapun kegiatan yang dilakukan selama sesi berlangsung dapat dilihat pada poin-poin berikut.

1. Peneliti menjelaskan secara umum penelitian yang dilakukan sebagai upaya penyamaan konteks dan pemahaman antara peneliti dan ahli linguistik.
2. Ahli linguistik melakukan pengujian terhadap sampel *dataset* ringkasan abstraktif yang dihasilkan oleh *Large Language Model* (LLM), meliputi penilaian terhadap keabsahan ringkasan secara abstraktif dan halusinasi ringkasan.
3. Ahli linguistik memberikan penilaian berdasarkan kategori 'Ya' atau 'Tidak' terhadap setiap parameter pengujian yang tercantum pada lembar lampiran ke-2.

Berdasarkan hasil validasi data sampel (terlampir), *dataset* yang digunakan oleh peneliti dinyatakan (Layak / Tidak Layak) Digunakan sebagai *dataset* model peringkasan abstraktif otomatis terhadap kumpulan ulasan produk di *e-commerce*.

Demikian Berita Acara ini dibuat dengan sebenarnya untuk dapat digunakan sebagaimana mestinya.

Depok, 29 Juni 2026

Pihak Pertama (Peneliti)


Randy Aulia Ananda Ruslani
NIM. 2207411041

Pihak Kedua (Ahli Linguistik/Sastra)


Kezia Kathryn, M.Hum



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



KEMENTERIAN PENDIDIKAN, TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp. (021) 91274097, Fax : (021) 7863531, (021) 7270036 Haring
Laman : <http://www.pnj.ac.id>, e-mail : tlk.pnj@gmail.com

L3

LEMBAR CONTOH PARAMETER PENGUJIAN HASIL RINGKASAN

Konteks Penelitian :

Untuk melatih model AI T5 dalam tugas peringkasan tumpukan ulasan berbahasa Indonesia, peneliti membutuhkan dataset latih yang terdiri atas tumpukan ulasan dan ringkasannya. masalahnya adalah ringkasan dari tumpukan ulasan tersebut dihasilkan oleh bantuan LLM sehingga untuk menjadi *ground truth* keabsahannya harus diuji oleh ahli linguistik. Dengan demikian, Jika dataset teruji layak digunakan, LLM akan mempelajari pola ringkasan yang benar dan sesuai dengan kaidah peringkasan abstraktif berbahasa Indonesia.

Petunjuk Penilaian :

- Ahli membandingkan antara tumpukan ulasan dengan ringkasannya. lalu memberikan penilaian berdasarkan parameter yang ada.
- **Dataset yang layak** digunakan adalah data dengan ringkasan yang representatif, abstraktif, dan tidak mengalami halusinasi
- **Dataset yang tidak layak** adalah data dengan ringkasan yang tidak representatif/tidak abstraktif/mengalami halusinasi
- **Halusinasi** adalah kondisi ketika hasil ringkasan LLM memuat sesuatu yang tidak seharusnya akibat dari hayalan atau imajinasi dari LLM (AI)
- **Representatif** adalah seberapa baik ringkasan menangkap poin poin penting dalam ulasan yang ada sehingga ringkasan sudah cukup mewakili seluruh ide pokok dari ulasan-ulasan yang ada.
- **Abstraktif** adalah seberapa baik tingkat abstraksi ringkasan yang dibuat oleh LLM
- **LLM** adalah AI Model yang digunakan, Contoh : ChatGPT

Berikut ini adalah contoh tabel penilaian :



KEMENTERIAN PENDIDIKAN, TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp. (021) 91274097, Fax : (021) 7863531, (021) 7270036 Haring
Laman : <http://www.pnj.ac.id>, e-mail : tlk.pnj@gmail.com

L3

Kode Uji	Tumpukan Ulasan	Ringkasannya oleh LLM	Apakah ringkasan representatif dan memenuhi kriteria abstraktif ?	Apakah ringkasan mengalami halusinasi ?
UKD-01	mksih ya sis pketn sampai cepet, barang cepet sampai dan responsive, lumayan tpi ada aroma tdk nyaman dari matrasnya. terima kasih barang sudah diterima dengan baik barang oke ada yang lebih tebal gan matrasnya. makasih barang sdh sampai. sesuai dgn deskripsi puas lah belanja di agan thanx lot gan. kualitas bagus seller ramah responnya cepet top deh makasih ya. barang sesuai gambar pelayanan bagus. sesuai pesanan pengiriman cepet. terimakasih untuk ds nya yaa sudah diterima dengan baik oleh customer. warna sesuai pesanan lunyuan lah. barang sudah sampai trimis warung olahraga tokped. warnanya bagus bahannya juga empuk dan nyaman sarung tas nya berguna sekali thanks. matrasnya sudah sampai barangnya bagus packingnya rapi overall memuaskan. mantap seller responsive keset kesetnya kaya perawan. sudah diterima terima kasih. brgnya sdh tiba dgn kondisi baik dan pengiriman nya cepet thanks. barang sdh saya trima cepet nyampinya tq. sipp mantap rescell. bagus harga murah tapi gak murahan. matrasx ok suka bahanx warnanya pengiriman lama di jne sesuai track sampe tgl kirim ojek tgl dan baru dpt baranya. mantap recomended seller. sip. karpet ok sesuai dengan gambar. barang diterima dengan baik terimakasih banyak. pesanan cepet sampai dan warna sesuai pesanan. terima kasih barang sudah sampai.	produk ini umumnya dinilai memuaskan karena sesuai gambar dan deskripsi, kualitas serta ketebalan matras terasa nyaman, packing rapi dan penjual ramah serta responsif sehingga pengiriman sering cepet. Beberapa pembeli mengeluhkan aroma tidak nyaman pada matras dan kadang proses pengiriman atau pemrosesan agak lama tergantung kurir. Secara keseluruhan banyak yang puas dan merekomendasikan produk ini.	Ya/Tidak	Ya/Tidak



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



KEMENTERIAN PENDIDIKAN TINGGI, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JAKARTA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
Universitas Indonesia, Jl. Prof. DR. G.A. Siwabessy, Depok 16425
Telp: (021) 91274097, Fax: (021) 7863531, (021) 7270036 Hmailing
Laman : <http://www.pnj.ac.id>, e-mail : tk.pnj@gmail.com

L3

barangnya dah sampe makasih ya. good quality good quality. barang sesuai pesanan dan sesuai gambar. lumayan memuaskan kan brg cepat sampai. bagus sesuai harapan. produk sesuai denga digambar memuaskan. ok. barang dan kecepatan kirim sangat bagus thx. kondisi barang bagus. maaf telat beri ulasan barang bagus kualitas ok. barangny sdh sampe bagus makasih. barang sudah diterima dengan baik memuaskan sesuai dengan order tapi sepertinya lama diproses pengiriman ya terima kasih sist. kualitas produk bagus ketebalannya sesuai gak licin sip			
--	--	--	--

Catatan Tambahan dari Ahli linguistik :

Secara keseluruhan, rangkuman dapat merepresentasikan isi ulasan dengan baik dan sesuai. Hanya ada satu ketidaksesuaian pada data nomor 1, yaitu rangkuman menyebutkan packing rapi, padahal informasi ini tidak tercantum pada ulasan.

Tambahan catatan soal kaidah kebahasaan, disarankan penggunaan bahasa untuk ditingkatkan. Beberapa kalimat kurang efektif, menyebabkan alurnya terasa kurang nyaman dibaca. Penggunaan kata penghubung seperti "tetapi", "namun", dan "sehingga" juga masih perlu disesuaikan agar hubungan antarkalimat lebih tepat dan koheren. Di beberapa bagian, penggunaan tanda baca, terutama titik dan koma, juga masih terdapat kesalahan. Isi rangkuman sudah dapat dipahami dengan baik, tetapi saran perbaikan pada beberapa aspek bisa menghasilkan rangkuman yang lebih rapi dan nyaman dibaca.



Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta