

**IMPLEMENTASI MODEL TRANSFORMERS T5-*small*
UNTUK *TEXT SUMMARIZATION* DALAM PEMBUATAN
META DESKRIPSI SEO**

SKRIPSI

IRVAN ARIF AGUSTIAN

20210040145



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN**

SUKABUMI

2025

**IMPLEMENTASI MODEL TRANSFORMERS T5-small
UNTUK *TEXT SUMMARIZATION* DALAM PEMBUATAN
META DESKRIPSI SEO**

SKRIPSI

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh Gelar
Sarjana Komputer Pada Program Studi Teknik Informatika*

IRVAN ARIF AGUSTIAN

20210040145



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN**

SUKABUMI

2025

PERNYATAAN PENULIS

JUDUL :IMPLEMENTASI MODEL TRANSFORMERS T5-*small*
UNTUK *TEXT SUMMARIZATION* DALAM PEMBUATAN
META DESKRIPSI SEO

NAMA :IRVAN ARIF AGUSTIAN

NIM 20210040145

”Penulis menyatakan dan bertanggung jawab dengan sebenarnya bahwa skripsi ini adalah hasil karya sendiri, kecuali cuplikan dan ringkasan yang masing-masing telah dijelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka penulis bersedia untuk mempertanggungjawabkannya secara akademik sesuai dengan peraturan yang berlaku, beserta segala hak dan kewajibannya”.

Sukabumi, Agustus 2025



Materai

Irvan Arif Agustian
Penulis

PENGESAHAN SKRIPSI

JUDUL : IMPLEMENTASI MODEL TRANSFORMERS T5-*small*
UNTUK *TEXT SUMMARIZATION* DALAM
PEMBUATAN META DESKRIPSI SEO

NAMA PENULIS : IRVAN ARIF AGUSTIAN

NIM PENULIS 20210040145

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 11 Agustus 2025. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom) Sukabumi, 11 Agustus 2025

Pembimbing I

Pembimbing II

Ivana Lucia Kharisma, M.Kom

NIDN . 0429038002

Penguji

Gina Purnama Insany, S.ST, M.Kom

NIDN . 041707790

Ketua Program Studi Teknik Informatika



Alun Sujjada, S.Kom, M.T

NIDN . 0718108001

Ir. Somantri, S.T., M.Kom

NIDN . 0419128801

Dekan Fakultas Teknik Komputer dan Desain

Ir.H. Paikum S.T.,IPM., Asean Eng

NIDN. 0402037401

HALAMAN MOTTO DAN PERUNTUKAN

MOTTO

"Dalam sunyi, menemukan ruang untuk berpikir. Dalam kesendirian, menemukan kekuatan untuk menyelesaikan Skripsi. ini adalah bukti bahwa keinginan untuk tidak melakukan pekerjaan berulang adalah ibu dari segala otomatisasi. Ini kupersembahkan untuk mimpiku yang paling mulia: tidur nyenyak tanpa memikirkan revisi"

(Penulis)

"Aku terlalu malas untuk menjadi ambisius"

(Dostoevsky "Notes from Underground")

"Kemajuan tidak dicapai oleh mereka yang bangun pagi, melainkan oleh mereka yang malas dan mencoba mencari cara yang lebih mudah untuk melakukan sesuatu"

(Robert A. Heinlein "Time Enough for Love")

"Seharusnya aku tidak melakukan hal merepotkan seperti ini, tapi lebih merepotkan lagi kalau tidak kulakukan."

(Shikamaru Nara "Naruto")



PERUNTUKAN

Dengan segala kerendahan hati dan rasa syukur, penulis ucapkan terima kasih kepada Allah SWT yang senantiasa memberikan segala keberanian, ketabahan, dan semangat yang tak kenal rasa menyerah dalam diri penulis. Atas segala izinnya selalu menjawab doa-doa baik yang penulis panjatkan dan memberikan takdir yang terbaik untuk penulis. Berkat rahmat dan anugerah-nya, akhirnya penulis bisa menyelesaikan tugas ini dengan baik.

Dengan segala kerendahan hati, karya skripsi ini penulis persembahkan kepada:

1. Allah SWT, Segala puji bagi Allah SWT atas limpahan rahmat, hidayah, dan kemudahan-Nya dalam setiap langkah perjalanan hingga skripsi ini dapat terselesaikan.
2. Kedua Orang Tua tercinta yang selalu memberikan doa, dan dukungan yang tiada henti. Semangat dan pengorbanan kalian menjadi sumber motivasi terbesar dalam menyelesaikan skripsi ini.
3. Dosen Pembimbing, terima kasih kepada Ibu Ivana Lucia Kharisma., M.Kom dan Ibu Gina Purnama Insany, S.ST, M.Kom atas bimbingan, arahan, serta ilmu yang diberikan dengan sabar dan penuh perhatian selama proses penyusunan skripsi ini.
4. Dosen-Dosen Program Studi Teknik Informatika, terima kasih yang sebesar-besarnya penulis sampaikan kepada seluruh keluarga besar Teknik Informatika, termasuk Kaprodi beserta jajarannya, serta semua dosen yang tidak dapat penulis sebutkan satu per satu. Ilmu dan pengalaman berharga yang kalian berikan selama masa perkuliahan telah menjadi fondasi yang kuat dan sangat bermanfaat dalam penelitian ini.
5. Keluarga yang selalu memberikan dukungan, motivasi, dan semangat, sehingga dapat menyelesaikan studi ini. Tak lupa penulis juga mengucapkan terima kasih kepada seluruh anggota keluarga atas cinta, doa, dan dukungan yang tiada henti. Kehadiran dan semangat kalian selalu menjadi kekuatan terbesar yang mendorong penulis untuk terus maju.

6. Untuk semua teman yang telah kebersamai dan menjadi motivasi, terutama rekan-rekan Teknik Informatika, Semoga kita dapat terus bersama dan saling mendukung ke depannya. Terima kasih atas kebersamaan dan dukungan selama masa perkuliahan ini.
7. Terakhir kepada pihak lainnya yang tidak bisa disebutkan satu persatu, untuk semua yang telah memberikan bantuan, dukungan, dan doa dalam berbagai bentuk selama masa studi dan penyusunan skripsi penulis, baik yang secara langsung maupun tidak langsung, penulis ucapkan terima kasih yang sebesar-besarnya.
8. Akhir kata, Penulis berharap semoga karya yang sederhana ini dapat memberikan manfaat bagi penulis sebagai penulis, pembaca, dan juga bagi perkembangan ilmu pengetahuan. Semoga segala kesalahan dan kekurangan yang ada dalam karya ini menjadi pelajaran berharga bagi penulis di masa mendatang. Aamiin.



ABSTRAK

Pembuatan meta deskripsi untuk *Search Engine Optimization* (SEO) secara manual merupakan proses yang tidak efisien, memakan waktu, dan sulit menjaga konsistensi, terutama untuk situs web dengan volume konten yang besar. Penelitian ini bertujuan untuk mengimplementasikan dan mengevaluasi model *text summarization* berbasis arsitektur Transformer, khususnya varian T5-small, guna mengotomatiskan pembuatan meta deskripsi SEO untuk konten berbahasa Indonesia. Metode penelitian yang digunakan adalah melakukan *fine-tuning* lanjutan terhadap model checkpoint panggi/t5-small-indonesian-summarization-cased dengan melatihnya kembali pada dataset berita skala besar fajrikoto/id_liputan6. Proses adaptasi model ini dijalankan di lingkungan *cloud computing* Kaggle Notebooks. Kinerja model diukur menggunakan metrik evaluasi standar NLP, yaitu ROUGE dan BLEU. Hasil evaluasi kuantitatif pada 10.972 data uji menunjukkan performa yang menjanjikan dengan pencapaian skor ROUGE-1 sebesar 34.09, ROUGE-2 sebesar 18.67, ROUGE-L sebesar 29.37, dan BLEU sebesar 10.41. Model yang telah di-*fine-tuning* kemudian berhasil diimplementasikan ke dalam sebuah aplikasi web fungsional menggunakan framework Streamlit sebagai bukti konsep penerapan praktis. Penelitian ini membuktikan bahwa model T5-small memiliki potensi sebagai solusi untuk meningkatkan efisiensi dan skalabilitas dalam produksi konten SEO, sekaligus memberikan kontribusi pada penerapan NLP untuk domain digital *marketing* di Indonesia.

Kata kunci: *text summarization*, Transformer T5-small, Meta Deskripsi SEO, *Search Engine Optimization* (SEO), T5-small, *Natural Language Processing* (NLP).

ABSTRACT

The manual creation of meta descriptions for Search Engine Optimization (SEO) is an inefficient, time-consuming process that struggles with consistency, especially for websites with a large volume of content. This research aims to implement and evaluate a Transformer-based text summarization model, specifically the T5-small variant, to automate the generation of SEO meta descriptions for Indonesian language content. The research method involves further fine-tuning the panggi/t5-small-indonesian-summarization-cased checkpoint model by retraining it on the large-scale news dataset fajrikoto/id_liputan6. This model adaptation process was conducted in the Kaggle Notebooks cloud computing environment. The model's performance was measured using standard NLP evaluation metrics, namely ROUGE and BLEU. Quantitative evaluation results on 10,972 test data samples show promising performance, achieving a ROUGE-1 score of 34.09, ROUGE-2 of 18.67, ROUGE-L of 29.37, and a BLEU score of 10.41. The fine-tuned model was then successfully implemented into a functional web application using the Streamlit framework as a practical proof of concept. This study demonstrates that the T5-small model has the potential to be a solution for improving efficiency and scalability in SEO content production. It also contributes to the application of NLP in the digital marketing domain in Indonesia.

Keywords: *text summarization, Transformer T5-small, Meta Description SEO, Search Engine Optimization (SEO), T5-small, Natural Language Processing (NLP).*

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya. Shalawat dan salam penulis haturkan kepada Nabi Muhammad SAW. Dengan penuh rasa syukur, penulis dapat menyelesaikan skripsi ini yang berjudul “IMPLEMENTASI MODEL TRANSFORMERS T5-small UNTUK *text summarization* DALAM PEMBUATAN META DESKRIPSI SEO” dengan baik.

Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna, baik dari segi penyajian, permasalahan, maupun pemecahan masalah. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang konstruktif untuk perbaikan di masa mendatang. Semoga skripsi ini dapat bermanfaat, khususnya bagi penulis dan umumnya bagi semua pihak yang membutuhkan.

Selama mengerjakan skripsi ini, penulis tidak lepas dari bantuan dan dorongan dari berbagai pihak, sehingga skripsi ini dapat terselesaikan. Pada kesempatan ini, penulis ingin mengucapkan terima kasih yang setulus-tulusnya kepada:

1. Bapak Dr. H. Kurniawan, S.T., MH, selaku Rektor Universitas Nusa Putra.
2. Bapak Ir.H. Paikum, S.T., IPM., Asean Eng, selaku Dekan Fakultas Universitas Nusa Putra.
3. Bapak Ir.Somantri, ST., M.Kom, selaku ketua program studi Teknik Informatika Universitas Nusa Putra.
4. Ibu Ivana Lucia Kharisma, M.Kom dan Ibu Gina Purnama Insany, S.ST, M.Kom, selaku dosen pembimbing skripsi atas segala bimbingan, arahan serta saran yang diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik.
5. Ibu Anggun Fergina, S.Kom, M.Kom, selaku pembimbing akademik, yang selalu memberikan bimbingan, arahan, dan dukungan selama proses perkuliahan.
6. Seluruh dosen/staff pengajar Fakultas Komputer Teknik dan desain program studi Teknik Informatika yang telah memberikan ilmu

pengetahuan yang tak ternilai selama penulis menempuh pendidikan di Universitas Nusa Putra.

7. Kepada seluruh keluarga, orang tua, bibi, adik, teman yang selalu memberikan dukungan penuh agar penulis bisa menyelesaikan skripsi ini. Dorongan dan motivasi dari mereka sangat berarti, dan menjadi sumber semangat yang tak ternilai selama proses penulisan.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu kritik dan saran yang membangun dari berbagai pihak sangat diharapkan demi perbaikan. Amin Yaa Rabbal 'Alamin.

Sukabumi, Agustus 2025

Penulis



Irvan Arif Agustian

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik UNIVERSITAS NUSA PUTRA, penulis yang bertanda tangan dibawah ini :

Nama : Irvan Arif Agustian
NIM 20210040145
Program Studi : Teknik Informatika
Jenis Karya : Tugas Akhir (Skripsi)

Demi pengembangan ilmu pengetahuan, penulis menyetujui untuk memberikan kepada Universitas Nusa Putra **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah yang berjudul:

“IMPLEMENTASI MODEL TRANSFORMERS T5-*small* UNTUK TEXT SUMMARIZATION DALAM PEMBUATAN META DESKRIPSI SEO”.

Dengan Hak Bebas Royalti Noneksklusif ini, Universitas Nusa Putra berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir (Skripsi) saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Sukabumi
Pada tanggal: Agustus 2025
Yang Menyatakan,

Irvan Arif Agustian

DAFTAR ISI

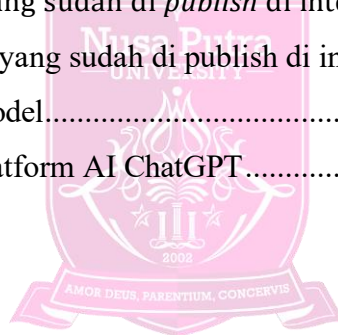
HALAMAN JUDUL	i
PERNYATAAN PENULIS	II
PENGESAHAN SKRIPSI	III
PERUNTUKAN	V
ABSTRAK	VII
KATA PENGANTAR	IX
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS SKRIPSI UNTUK KEPENTINGAN AKADEMIS	XI
DAFTAR ISI	XII
DAFTAR TABEL	XIV
DAFTAR GAMBAR	XV
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	6
1.3 Batasan Masalah	7
1.4 Tujuan Penelitian	8
1.5 Manfaat Penelitian	9
1.6 Sistematika Penelitian.....	10
BAB II	11
TINJAUAN PUSTAKA	11
2.1. Penelitian Terkait.....	11
2.2 Landasan Teori	17
2.2.1 <i>Natural Language Processing (NLP)</i>	17
2.2.2 <i>Text Summarization</i>	18
2.2.3 <i>T5-SMALL</i>	19
2.2.4 <i>Search Engine Optimization</i>	21
2.2.5 <i>Meta Deskripsi SEO</i>	22
2.3 Kerangka Berpikir	23
BAB III	25
METODOLOGI PENELITIAN	25
3.1. Metode Penelitian	25
3.2 Identifikasi Masalah.....	29
3.3 Studi Literatur dan Pemilihan Model	30

3.4 Pengumpulan Data.....	32
3.5 Data <i>Preprocessing</i>	33
3.6 <i>Fine-tuning</i> Model.....	34
3.7 Evaluasi Model.....	37
3.8 <i>Deployment</i>	40
BAB IV	46
HASIL DAN PEMBAHASAN.....	46
4.1 Persiapan Data dan Hasil <i>Preprocessing</i>	46
4.2 Hasil Pelatihan Model (<i>Fine-Tuning</i>).....	53
4.3 Hasil Evaluasi Kinerja Model	54
4.4 Pembahasan	59
4.5 Hasil Uji Website.....	63
BAB V	71
PENUTUP	71
5.1 Kesimpulan.....	71
5.2 Saran.....	72
DAFTAR PUSTAKA	73



DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	11
Tabel 3. 1 Konfigurasi Hiperparameter <i>Fine-Tuning</i>	35
Tabel 3. 2 Metrik Evaluasi yang Digunakan	37
Tabel 3. 3 Test Skenario.....	44
Tabel 4. 1 Data Artikel mentah.....	47
Tabel 4. 2 Data Artikel.....	48
Tabel 4. 3 Data Artikel dalam Bentuk Angka	50
Tabel 4. 4 Data Artikel Input Model.....	51
Tabel 4. 5 Data Artikel Target Model.....	52
Tabel 4. 6 Evaluasi	54
Tabel 4. 7 Data artikel test 1	55
Tabel 4. 8 Evaluasi satu data test.....	56
Tabel 4. 9 Data artikel yang sudah di <i>publish</i> di internet	56
Tabel 4. 10 Evaluasi data yang sudah di publish di internet	58
Tabel 4. 11 Hasil meta model.....	59
Tabel 4. 12 Hasil meta platform AI ChatGPT.....	62



DAFTAR GAMBAR

Gambar 2. 1 contoh meta deskripsi	23
Gambar 2. 2 contoh skripmeta deskripsi	23
Gambar 2. 3 Kerangka Berpikir	24
Gambar 3. 1 Tahapan Penelitian	26
Gambar 3. 2 Side bar.....	41
Gambar 3. 3 Halaman dasar beranda.....	42
Gambar 3. 4 Halaman dasar buat meta deskripsi	42
Gambar 3. 5 Halaman dasar panduan.....	43
Gambar 3. 6 Halaman dasar metodologi	43
Gambar 3. 7 Halaman dasar hasil.....	44
Gambar 4. 1 Halaman Beranda	64
Gambar 4. 2 Halaman buat meta deskripsi.....	64
Gambar 4. 3 Halaman panduan	65
Gambar 4. 4 Halaman metodologi.....	65
Gambar 4. 5 Halaman hasil dan pembahasan.....	66
Gambar 4. 6 Hasil Pengujian F1.....	66
Gambar 4. 7 Hasil Pengujian E1	67
Gambar 4. 8 Hasil Pengujian E2	68
Gambar 4. 9 Hasil Pengujian U1 Navigasi Panduan.....	68
Gambar 4. 10 Hasil Pengujian U2 Navigasi Metodologi	69
Gambar 4. 11 Hasil Pengujian U3 Navigasi Hasil	69

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital saat ini, kemajuan teknologi dan digitalisasi telah berdampak secara signifikan terhadap cara perusahaan berinteraksi dengan audiens mereka [1]. Dalam konteks persaingan yang semakin ketat dan tingginya arus informasi, perusahaan dituntut untuk mampu menyampaikan pesan secara relevan dan bernilai bagi konsumen. Di sinilah pemasaran konten (*content marketing*) memainkan peran penting [1]. Pendekatan ini tidak hanya menitikberatkan pada promosi langsung, melainkan berfokus pada penciptaan serta distribusi konten yang informatif, edukatif, atau menghibur guna menarik perhatian audiens dan membangun hubungan yang lebih personal. Konten yang ditargetkan dengan baik dapat membangun kepercayaan, memperkuat citra merek, serta membuka peluang untuk keterlibatan yang lebih dalam antara perusahaan dan konsumennya [1]. Melalui strategi pemasaran konten yang terencana, perusahaan mampu menjangkau audiens secara lebih strategis, meningkatkan loyalitas pelanggan, dan menciptakan nilai jangka panjang. Secara keseluruhan, pemasaran konten bukan hanya sebuah tren pemasaran, melainkan sebuah pendekatan penting untuk menjalankan strategi bisnis yang sukses di era yang semakin terhubung dan digital [1].

Search Engine Optimization (SEO) juga merupakan bagian penting dalam menciptakan kekuatan personal *branding* di situs web [2]. Penggunaan SEO dalam dunia digital *marketing* bertujuan untuk mengoptimalkan penggunaan mesin pencari. SEO memungkinkan sebuah situs web, merek, produk, atau layanan untuk lebih mudah diakses oleh pengunjung karena berada di halaman pertama hasil pencarian mesin pencari seperti Google [2]. Menerapkan praktik SEO terbaru dapat memperkuat visibilitas situs web, aksesibilitas, dan kesadaran

merek [2]. SEO membantu meningkatkan visibilitas online situs web atau halaman web di mesin pencari seperti Google, Bing, Baidu, Yandex, Yahoo, dan lainnya [2].

Meta Description adalah keterangan dari halaman berdasarkan kalimat artikel, agar *meta description* dapat menarik *traffic*, perlu dilakukan modifikasi [3]. *Keyword* akan ditempatkan pada 150 *character* pertama [dalam *meta description*] [3]. *Meta description* adalah deskripsi singkat yang muncul di hasil pencarian Google, dirancang untuk mencerminkan konten halaman dan harus mengandung kata kunci relevan [4]. Optimasi ini bertujuan meningkatkan visibilitas dan daya tarik halaman web, sehingga diharapkan dapat meningkatkan jumlah klik dan kunjungan ke situs web, serta berperan penting dalam strategi SEO *On-Page* [4].

Yoast SEO adalah salah satu *plugin* yang ada pada WordPress untuk mengoptimasi SEO pada sebuah *website*, dalam Yoast terdapat indikator SEO yang otomatis akan menganalisa kualitas konten pada website [5]. Untuk pengguna non-teknis, *plugin* ini menjadi solusi praktis dalam menerapkan strategi SEO tanpa harus memahami detail teknis yang kompleks. *Plugin* ini menyediakan indikator seperti *Focus Keyword*, *Meta Title*, *Meta Description*, *Permalink*, *Text Length*, *Internal Links*, dan lainnya, yang membantu pengguna menyusun elemen-elemen SEO sesuai standar [5]. Yoast SEO juga memiliki sistem standar penulisan artikel yang *SEO Friendly*, dan menampilkan indikator visual seperti lampu hijau jika elemen SEO terpenuhi [5].

Seiring berkembangnya teknologi, strategi SEO tidak lagi hanya mengandalkan pengaturan manual seperti *plugin*. Kini, integrasi kecerdasan buatan (AI) menjadi pendekatan baru dalam mengotomatiskan optimasi konten. Selain itu, teknologi AI dapat menganalisis data untuk kebutuhan SEO [6]. Dengan perangkat berbasis kecerdasan buatan, analisis metode pengoptimalan SEO dapat memutuskan apa yang perlu dilakukan baik pada struktur, kata kunci, konten, atau tautan untuk meningkatkan kehadiran internet mereka [6]. Masa depan SEO terletak di tangan kecerdasan buatan karena teknik tradisional telah kehilangan kekuatannya [6]. AI menghilangkan kebutuhan akan frasa yang tepat

sebagai kata kunci [6]. Alat-alat seperti ChatGPT memiliki potensi untuk merevolusi pengalaman pencarian, mengubah cara informasi dihasilkan dan disajikan, dan membangun titik akses baru untuk keterlibatan online [7]. Peringkasan dengan menggunakan model Transformer T5, dimana proses peringkasan dilakukan secara abstraktif dengan beberapa *scenario*, telah diimplementasikan dalam dokumen berbahasa Indonesia [8].

Meringkas teks menjadi ringkasan yang efektif merupakan tantangan utama dalam pengembangan konten digital. Salah satu pembahasan pada *Natural Language Processing* adalah *Text Summarization*. *Text Summarization* bertujuan untuk mengkaji tentang otomatis perangkuman teks dengan pilihan metode-metode/algorithm yang bermacam-macam [9]. Peringkasan teks pertama-tama memerlukan interpretasi dan penyajian teks sumber dan kemudian mengubah representasi sumber menjadi representasi ringkasan sebelum menghasilkan teks ringkasan [10]. *Automatic summarization* (otomatis perangkuman) adalah proses mengurangi dokumen teks menggunakan program komputer dalam rangka membuat rangkuman yang menyimpan titik-titik terpenting dari dokumen asli [9]. TS otomatis sudah ada sejak tahun 1958, ketika Luhn (1958) memperkenalkan konsep '*auto abstract*'. Sejak saat itu, banyak pendekatan dan teknik TS otomatis yang sudah dieksplorasi dan diusulkan [10]. Ringkasan harus dipadatkan, terintegrasi dengan konten, dan mudah dibaca [10]. Studi NLP baru-baru ini mulai berfokus pada perspektif pragmatis dari ringkasan teks [10]. Di zaman ini, teknologi *summarization* digunakan pada banyak sektor industri. Sebuah contoh adalah *search engine* seperti Google [9]. Integrasi antara pemodelan Transformer khususnya dalam tugas peringkasan teks dengan menggabungkan metode *text rank* telah mampu meringkas dokumen yang memiliki halaman yang cukup banyak sehingga dapat mengurangi waktu yang dibutuhkan jika ingin meringkas dokumen secara manual [11].

Salah satu pengembangan signifikan dari model Transformer adalah metode *Neural Transfer Reformulation* (NTR), yang dibangun dengan menyempurnakan model T5 untuk tugas perumusan ulang kueri, *Neural Transfer Reformulation*

(NTR), sebuah pendekatan yang dibangun dari karya sebelumnya untuk menyempurnakan *Text-to-Text Transfer Transformer* (T5) yang sudah terlatih [12]. Dibandingkan dengan hasil pemeringkatan penuh, semua metode perumusan ulang kueri menunjukkan peningkatan. Efektivitas pemeringkatan ulang NTR (T5) mendekati kueri manual [12]. Di antara semua metode CQR, NTR (T5) menghasilkan kualitas peringkat terbaik dan secara signifikan mengungguli metode CQR lainnya dalam hal NDGC@3 dan NDCG@1 [12]. Untuk mengatasi masalah *coreference* dan penghilangan dalam kueri percakapan, penelitian sebelumnya mengusulkan dua metode CQR, yaitu *Historical Query Expansion* (HQE) dan *Neural Transfer Reformulation* (NTR). NTR, yang dibangun di atas model urutan-ke-urutan yang telah dilatih sebelumnya, T5, secara langsung memformulasikan ulang kueri percakapan menjadi kueri yang alami bagi pemahaman manusia [12].

Serupa dengan penelitian sebelumnya oleh Fadhlillah dan Ikasari, varian T5-*small* dipilih karena efisiensi sumber daya [13], dan dilakukan di google colab. *Cloud Notebooks* dipilih karena menyediakan sumber daya yang diperlukan, karena keterbatasan yang dihadapi dalam penelitian [13]. Meskipun demikian, model bahasa yang telah disetel dengan baik masih dapat menghasilkan ringkasan abstrak yang layak dari sebuah artikel berita [13]. meskipun skor ROUGE yang dihasilkan relatif rendah, namun kualitas ringkasan tetap dianggap baik berdasarkan penilaian manual oleh responden, yang menilai ringkasan tetap koheren, relevan, dan sesuai konteks [14].

Penelitian ini menggunakan model T5-*small* dengan proses eksperimen yang dilakukan di Kaggle *Notebooks*. Dataset yang digunakan adalah fajrikoto/id_liputan6 yang tersedia di Hugging Face atau Github.

Penelitian ini bertujuan untuk mengembangkan model Transformer T5-*small* yang dapat menciptakan meta deskripsi SEO. Model ini diharapkan mampu meningkatkan efisiensi kerja praktisi digital *marketing*. Model ini diharapkan memiliki performa yang baik. Penelitian ini diharapkan berkontribusi dalam bidang NLP dan digital *marketing*, khususnya dalam penerapan *text*

summarization untuk optimasi mesin pencarian. Efektivitas model akan diukur menggunakan metrik evaluasi seperti ROUGE dan BLEU untuk memastikan ringkasan tetap ringkas tanpa kehilangan informasi penting.

Inovasi utama dari penelitian ini terletak pada integrasi antara bidang Pemrosesan Bahasa Alami (*Natural Language Processing* - NLP) dan *Search Engine Optimization* (SEO) untuk menghasilkan solusi praktis dan otomatis [26]. Penelitian ini bukan menciptakan model Transformer baru dari awal, melainkan menyajikan sebuah penerapan baru (*novel application*) dari teknologi peringkasan teks abstraktif untuk mengatasi tantangan spesifik dalam dunia pemasaran digital, yaitu pembuatan meta deskripsi SEO yang efektif dan efisien [17].

Bentuk inovasi dari penelitian ini dapat diuraikan menjadi tiga poin utama:

- a. Menjembatani Kesenjangan antara Peringkasan Teks Abstraktif dan Kebutuhan SEO

Penelitian sebelumnya telah mengeksplorasi penggunaan peringkasan teks untuk meta deskripsi, namun umumnya masih menggunakan metode ekstraktif tradisional yang hanya menyalin kalimat penting dari sumber [17]. Di sisi lain, penelitian yang telah menggunakan model Transformer seperti T5 untuk tugas peringkasan abstraktif pada artikel berita atau transkrip video berbahasa Indonesia, tetapi tujuannya belum spesifik untuk optimasi SEO.

Untuk menjembatani kesenjangan tersebut dengan secara khusus menerapkan model peringkasan abstraktif (T5-*small*) yang mampu menghasilkan kalimat-kalimat baru [25] [27], untuk tujuan praktis pembuatan meta deskripsi SEO [17].

- b. Otomatisasi Proses Manual dalam Strategi Konten *On-Page*

Pembuatan meta deskripsi adalah salah satu faktor penting dalam strategi SEO *On-Page* [21] [26]. Namun, proses ini sering kali dilakukan secara tradisional atau manual, yang terbukti tidak efisien, memakan waktu, dan sulit menjaga konsistensi kualitas, terutama untuk situs web dengan volume konten yang besar.

Untuk mengotomatiskan tugas manual ini dengan memanfaatkan kemampuan model T5-*small*. Dengan menyediakan alat bantu yang dapat menghasilkan draf meta deskripsi, memberikan solusi untuk meningkatkan efisiensi dan produktivitas bagi praktisi pemasaran digital, penulis konten, dan pengelola situs web.

c. Implementasi Model yang Efisien dan Dapat Diakses

Di saat banyak penelitian berlomba menggunakan model AI berskala sangat besar, penelitian ini terletak pada pemilihan strategis model T5-*small*. Model ini dipilih karena efisiensi sumber daya, yang memungkinkannya untuk dilatih dan diimplementasikan pada lingkungan komputasi dengan keterbatasan perangkat keras, seperti Kaggle *Notebooks* atau Google Colab [13] [23].

Ini menunjukkan teknologi dapat diadaptasi untuk menjadi solusi yang lebih aksesibel dan terjangkau. Hal ini membuktikan bahwa untuk tugas-tugas spesifik seperti pembuatan meta deskripsi, model yang lebih ringan pun dapat memberikan hasil baik, tanpa harus bergantung pada infrastruktur komputasi yang mahal.

1.2 Rumusan Masalah

Adapun rumusan masalah yang diidentifikasi adalah sebagai berikut :

1. Bagaimana cara menerapkan *text summarization* menggunakan model Transformer T5-*small* untuk menghasilkan meta deskripsi SEO yang relevan secara konteks dan memenuhi panjang karakter ?
2. Apa saja faktor yang mempengaruhi kualitas meta deskripsi SEO yang dihasilkan dari proses *text summarization*?
3. Bagaimana performa *text summarization* berbasis Transformer T5-*small* dalam menghasilkan meta deskripsi SEO?

1.3 Batasan Masalah

Pada penelitian ini, peneliti membatasi masalah antara lain :

1. Penelitian ini hanya menggunakan model Transformer varian T5-*small* untuk tugas peringkasan teks. Model lain seperti BART, PEGASUS, atau varian T5 yang lebih besar tidak digunakan sebagai pembanding langsung dalam implementasi.
2. Sumber data yang digunakan untuk proses *fine-tuning* terbatas pada dataset fajrikoto/id_liputan6, tanpa melakukan augmentasi atau penggabungan dengan korpus dari domain lain.
3. Fokus optimasi SEO hanya pada aspek konten *on-page*, yaitu pembuatan deskripsi meta. Aspek SEO teknis lainnya seperti kecepatan situs, struktur URL, *crawl budget*, maupun SEO *off-page* seperti strategi *backlink* tidak dibahas dalam penelitian ini.
4. Evaluasi kinerja model peringkasan dilakukan secara kuantitatif hanya menggunakan metrik NLP, yaitu ROUGE dan BLEU sebagai pendukung. Penelitian ini tidak mencakup evaluasi kualitatif dari para ahli SEO atau evaluasi performa bisnis seperti *Click-Through Rate* (CTR) di mesin pencari.
5. Penelitian ini tidak menganalisis dampak dari variasi linguistik (*linguistic variance*) secara mendalam, seperti varians konseptual, realisasi, atau nonkonformitas dalam teks yang dihasilkan, yang dapat memengaruhi kinerja sistem penilaian otomatis.
6. Efektivitas model yang diimplementasikan tidak diuji terhadap dinamika perubahan algoritma mesin pencari Google di masa mendatang, yang berpotensi memengaruhi relevansi dan peringkat meta deskripsi yang dihasilkan.
7. Proses *fine-tuning* dan eksperimen sepenuhnya dilakukan dalam lingkungan *cloud computing* Kaggle *Notebooks* dengan akselerator GPU T4. Penelitian ini tidak melakukan perbandingan performa atau waktu pelatihan jika dijalankan pada lingkungan komputasi atau dengan spesifikasi perangkat keras yang berbeda.

8. Penelitian ini tidak membahas *cloud computing* Kaggle Notebooks dan GPU T4.
9. Konfigurasi hiperparameter yang digunakan dalam proses *fine-tuning*, seperti *learning rate* dan *weight decay*, diatur berdasarkan praktik umum dari penelitian sebelumnya dan tidak dilakukan eksplorasi atau *tuning* ekstensif untuk menemukan kombinasi hiperparameter optimal untuk *dataset* ini.
10. Aplikasi web yang dikembangkan menggunakan Streamlit berfungsi sebagai contoh penerapan untuk mendemonstrasikan fungsionalitas model. Penelitian ini tidak melakukan pengujian *User Experience* (UX) atau studi kegunaan (*usability*) secara formal dengan pengguna akhir, Dan tidak membahas Streamlit.

1.4 Tujuan Penelitian

Terdapat beberapa tujuan dari penelitian ini adalah sebagai berikut :

1. Menerapkan *text summarization* berbasis model transformers T5-*small* untuk membuat atau menghasilkan Meta deskripsi SEO.
2. Menganalisis faktor-faktor yang mempengaruhi kualitas meta deskripsi SEO yang dihasilkan dari proses *text summarization*
3. Mengevaluasi performa model Transformer T5-*small* dalam menghasilkan meta deskripsi SEO berdasarkan metrik evaluasi ROUGE.

Berbagai penelitian telah mengimplementasikan model Transformer untuk peringkasan teks dalam bahasa Indonesia. Namun, sebagian besar studi belum mengarahkan hasil peringkasan untuk mendukung strategi optimasi SEO, khususnya dalam bentuk meta *description* yang terstruktur dan relevan dengan pencarian pengguna. Oleh karena itu, penelitian ini mencoba kekosongan tersebut melalui pendekatan *text summarization* berbasis T5 yang disesuaikan untuk kebutuhan konten SEO.

1.5 Manfaat Penelitian

Adapun untuk beberapa manfaat dari penelitian ini adalah sebagai berikut :

1. Manfaat Bagi Mahasiswa

- a. Menambah wawasan dan pemahaman mengenai penerapan *Natural Language Processing* (NLP) dalam dunia nyata, khususnya untuk keperluan SEO.
- b. Melatih keterampilan teknis dalam *preprocessing* data, *fine-tuning* model Transformer T5-*small*, dan evaluasi hasil dengan metrik NLP (ROUGE dan BLEU).
- c. Memberikan pengalaman langsung dalam membangun aplikasi berbasis AI sederhana sebagai sarana demonstrasi sistem.
- d. Menjadi bekal untuk penelitian lanjutan atau karier di bidang AI, Data Science, dan digital *marketing* berbasis teknologi.

2. Manfaat Bagi Peneliti

- a. Menjadi referensi bagi penelitian selanjutnya terkait penerapan AI dalam optimasi SEO dan digital *marketing*.
- b. Memberikan landasan bagi pengembangan sistem SEO-*aware* NLP atau meta SEO berbasis model ringan seperti T5-*small*.
- c. Menunjukkan penerapan model NLP pada perangkat komputasi terbatas, yang bermanfaat bagi penelitian dengan keterbatasan sumber daya.
- d. Menawarkan model yang dapat dikembangkan lebih lanjut untuk optimasi konten digital dan strategi pemasaran berbasis AI dan memperluas aplikasi NLP dalam SEO, digital *marketing*, dan analisis teks otomatis.

3. Manfaat Bagi Masyarakat

- a. Membantu pelaku bisnis digital, pemilik *website*, blogger, dan *content writer* dalam menghasilkan meta deskripsi SEO.
- b. Meningkatkan efisiensi waktu dalam pembuatan konten online terutama pembuatan meta deskripsi.
- c. Menawarkan solusi AI yang dapat diakses oleh masyarakat umum tanpa memerlukan pengetahuan teknis mendalam atau infrastruktur mahal.

1.6 Sistematika Penelitian

Memberikan gambaran secara garis besar, dalam hal ini dijelaskan isi dari masing masing bab dari tugas akhir ini. Sistematika penulisan dalam pembuatan laporan ini sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penelitian tentang IMPLEMENTASI MODEL TRANSFORMERS T5-*small* UNTUK *TEXT SUMMARIZATION* DALAM PEMBUATAN META DESKRIPSI SEO.

BAB II TINJAUAN PUSTAKA

Dalam bab ini dibahas dan diuraikan secara teoritis mengenai konsep-konsep yang dijadikan landasan teori masalah, penelitian terkait, dan juga kerangka pemikiran.

BAB III METODOLOGI PENELITIAN

Dalam bab ini membahas tentang metode-metode seperti metode penelitian, metode pengumpulan data, dan metode perancangan sistem.

BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini membahas tentang *output* yang dihasilkan dalam penelitian dan dibahas secara detail.

BAB V PENUTUP

Dalam bab ini membahas tentang kesimpulan dari isi semua penelitian skripsi ini dan diakhiri dengan penutup.

DAFTAR PUSTAKA

BAB II

TINJAUAN PUSTAKA

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini berhasil membuktikan bahwa penerapan *text summarization* menggunakan model Transformer T5-*small* untuk menghasilkan meta deskripsi SEO dapat diimplementasikan, ini membuktikan bahwa T5 berpotensi dalam bidang digital *marketing* terutama dalam SEO. Pencapaian ini diraih melalui proses adaptasi model yang kemudian diwujudkan menjadi sebuah aplikasi web fungsional menggunakan framework Streamlit sebagai solusi praktisnya.

Secara kuantitatif, performa model ini setelah dievaluasi pada 10.972 data uji menunjukkan pencapaian skor yang cukup baik, dengan skor ROUGE-1 sebesar 34.09, ROUGE-2 sebesar 18.67, dan ROUGE-L sebesar 29.37, serta didukung oleh skor BLEU sebesar 10.41. Meskipun angka-angka ini menunjukkan kapabilitas model, terdapat catatan penting terkait kualitas hasil akhirnya. Kualitas deskripsi yang dihasilkan model AI ini dipengaruhi oleh dua hal utama: pertama, seberapa jelas artikel sumber yang dimasukkan, dan kedua, cara model diatur untuk menghasilkan kalimat yang lebih kreatif dan bervariasi.

Meski hasil deskripsinya sudah terdapat kata kunci, tetapi faktor kedua inilah yang menjadi landasan dari keterbatasan utama model, yaitu inkonsistensi hasil yang terkadang memerlukan beberapa kali percobaan untuk mendapatkan keluaran yang optimal. Karena artikel *dataset* yang dilatih adalah artikel berita yang menggunakan bahasa baku dan formal serta menggunakan piramid terbalik, dan adanya kebutuhan pasca-pemrosesan untuk menyesuaikan panjang karakter sesuai standar SEO bersama dengan inkonsistensi tersebut menjadi catatan penting dalam penerapannya di dunia nyata. Dapat disimpulkan implementasi ini berhasil dan fungsional, namun dengan pemahaman bahwa kualitasnya tidak selalu stabil dan sering kali memerlukan iterasi untuk mencapai hasil yang memuaskan

5.2 Saran

Untuk pengembangan di masa depan, disarankan agar penelitian selanjutnya dapat mengeksplorasi arsitektur model yang lebih besar dan lebih lanjut, seperti T5-*base*, T5-*large*, atau bahkan model lain seperti BART dan PEGASUS, mengingat penelitian ini terbatas pada T5-*small*. Peningkatan kapabilitas model ini sebaiknya diimbangi dengan penggunaan *dataset* yang lebih spesifik untuk domain SEO, misalnya dengan membangun korpus dari artikel dan meta deskripsi yang terbukti memiliki performa tinggi, alih-alih menggunakan *dataset* berita umum seperti yang dilakukan saat ini. Dari sisi fungsionalitas, sebuah langkah maju yang signifikan adalah dengan mengembangkan fitur pembangkitan berbasis kata kunci (*keyword-driven generation*), yang memungkinkan pengguna memasukkan kata kunci target untuk mengarahkan proses peringkasan agar lebih sesuai dengan strategi SEO yang spesifik. Pada akhirnya, untuk memvalidasi semua penyempurnaan tersebut, metode evaluasi perlu diperluas melampaui metrik NLP kuantitatif.



DAFTAR PUSTAKA

- [1] S. Ulya Hasasi, M. Panca Rani, and A. Setiawan, “Bibliometric Analysis in Mapping Research with the Central Theme of Digital Content *marketing*,” International Journal of Engineering, Science and Information Technology, vol. 5, no. 2, pp. 85–96, Feb. 2025, doi: 10.52088/ijesty.v5i2.790.
- [2] A. Setiawan, Z. Harahap, D. Syamsuar, and Y. N. Kunang, “The Optimization of Website Visibility and Traffic by Implementing Search Engine Optimization (SEO) in Palembang Polytechnic of Tourism,” 2020. [Online]. Available: <http://poltekpar-palembang.ac.id/>
- [3] P. Septiani dan H. Kurniawan, “Analisa Penggunaan Keyword Untuk Implementasi Search Engine Optimization (SEO),” Jurnal Teknologi Informasi, vol. XV, no. 3, pp. 83–91, Nov. 2020, ISSN: 1907-2430.
- [4] A. Zahra Maulaa Habiibah, A. Hermawan, and R. Gelar Guntara, “Implementasi Teknik Seo Dengan Metode On Page dan Off Page Dalam Meningkatkan Peringkat Website Canopybandung.Com”.
- [5] N. Widiyasono, “IMPLEMENTASI YOAST SEO DAN OPTIMASI KEYWORD UNTUK MENINGKATKAN RANK PADA GOOGLE SEARCH ENGINE RESULT PAGE (SERP),” vol. 7, no. 2, p. 2021.
- [6] T. H. Muhammad Fawaati et al., “APPLICATION OF ARTIFICIAL INTELLIGENCE (AI) IN SEARCH ENGINE OPTIMIZATION (SEO),” 2022.
- [7] Evolution of Generative AI: A paradigm Shift in optimization of Search Engine Strategies (SEO),” INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, vol. 09, no. 02, pp. 1–9, Feb. 2025, doi: 10.55041/IJSREM41499.
- [8] U. Primakara, “IMPLEMENTASI PERINGKAS DOKUMEN BERBAHASA INDONESIA MENGGUNAKAN METODE TEXT TO TEXT TRANSFER TRANSFORMER (T5) I Nyoman Purnama 1) , Ni

Nengah Widya Utami 2) Program Studi Sistem Informasi 1) , Sistem Informasi Akutansi 2).”

- [9] M. A. Zamzam, “SISTEM AUTOMATIC *text summarization* MENGGUNAKAN ALGORITMA TEXTRANK,” MATICS, vol. 12, no. 2, pp. 111–116, Sep. 2020, doi: 10.18860/mat.v12i2.8372.
- [10] Y. Li, M. a. Thomas, and D. Liu, “From semantics to pragmatics: where IS can lead in *Natural Language Processing* (NLP) research,” European Journal of Information Systems, vol. 30, no. 5, pp. 569–590, 2021, doi: 10.1080/0960085X.2020.1816145.
- [11] I. L. Kharisma, Kamdan, A. Fergina, T. Hidayat, M. A. A. Hidayat, M. Muslih, and A. Erfina, "Integration of Transformer Model *text summarization* and Text-to-Speech in Helping Document Understanding in the Bukudio Application," IAIC International Conference Series, vol. 4, no. 1, pp. 178–186, 2023. doi: 10.34306/conferenceseries.v4i1.653.
- [12] S.-C. Lin, J.-H. Yang, R. Nogueira, M.-F. Tsai, C.-J. Wang, and J. Lin, “Multi-Stage Conversational Passage Retrieval: An Approach to Fusing Term Importance Estimation and Neural Query Rewriting,” May 2020, [Online]. Available: <http://arxiv.org/abs/2005.02230>
- [13] D. A. Fadhlillah and D. Ikasari, “Optimizing Online News Understanding: Abstractive Summarization Approach with T5 for Comprehend Content,” International Research Journal of Advanced Engineering and Science, vol. 8, no. 4, pp. 189–194, 2023.
- [14] A. A. Fattahila, A. Romadhony, and S. A. Faraby, “Peringkasan Artikel Berita Menggunakan Pendekatan Abstraktif Dengan Model Transformers,” e-Proceeding of Engineering, vol. 10, no. 5, pp. 4980–4986, 2023.
- [15] M. Faris et al., “INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION journal homepage : www.joiv.org/index.php/joiv INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

- Exploring Current Methods and Trends in *text summarization*: A Systematic Mapping Study.” [Online]. Available: <https://www.scopus.com/>
- [16] T. Zesch, A. Horbach, and F. Zehner, “To Score or Not to Score: Factors Influencing Performance and Feasibility of Automatic Content Scoring of Text Responses,” *Educational Measurement: Issues and Practice*, vol. 42, no. 1, pp. 44–58, Mar. 2023, doi: 10.1111/emip.12544.
- [17] G. Matošević, “Text summarization techniques for meta description generation in process of search engine optimization,” in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2019, pp. 165–173. doi: 10.1007/978-3-319-91189-2_17.
- [18] M. Yani, N. Siti Khodijah, and M. Mustamiin, “Aplikasi Peringkat Teks Bahasa Indonesia Menggunakan Model *text-to-text Transfer* Transformer (T5)”, doi: 10.37817/ikraith-informatika.v9i2.
- [19] C. Raffel et al., “Exploring the Limits of Transfer Learning with a Unified *text-to-text* Transformer,” Oct. 2019, [Online]. Available: <http://arxiv.org/abs/1910.10683>
- [20] A. Kosasih, L. Purnomo, O. Sidanta, and Sfenrianto, “Search Engine Optimization (SEO) in Promoting E-Commerce Start Aja. Com,” *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 6, pp. 577–580, Mar. 2020, doi: 10.35940/ijrte.F7396.038620.
- [21] T. Ganang Wicaksono, “Implementation of Search Engine Optimization (SEO) on Website of IKM Semanggi Harmony,” *International Journal of Progressive Sciences and Technologies (IJPSAT)*, vol. 19, no. 1, pp. 202–206, 2020, [Online]. Available: <http://ijpsat.ijsht-journals.org>
- [22] E. Woncharso, A. Muawwal, S. Informasi, and S. KHARISMA Makassar, “Penerapan Search Engine Optimization (SEO) untuk meningkatkan pengunjung pada website SCLEAN.” [Online]. Available: <https://tech.kharisma.ac.id>

- [23] M. F. Fadlilah, A. R. Atmadja, and M. D. Firdaus, “Pemanfaatan Transformer untuk Peringkasan Teks: Studi Kasus pada Transkripsi Video Pembelajaran,” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 2111–2119, Dec. 2024, doi: 10.47065/bits.v6i3.6342.
- [24] R. Saxena, H. Tang, and F. Keller, “End-to-End Long Document Summarization using Gradient Caching.”
- [25] M. Sie, R. Beek, M. Bots, S. Brinkkemper, and A. Gatt, “Summarizing long regulatory documents with a multi-step pipeline,” Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.09777>
- [26] V. M. S and M. C. Padma, “Insights into Search Engine Optimization using Natural Language Processing and Machine Learning,” 2023. [Online]. Available: www.ijacsa.thesai.org
- [27] A. G. Etemad, A. I. Abidi, and M. Chhabra, “Fine-Tuned T5 for Abstractive Summarization,” *International Journal of Performability Engineering*, vol. 17, no. 10, pp. 900–906, Oct. 2021, doi: 10.23940/ijpe.21.10.p8.900906.
- [28] F. Koto, J. H. Lau, and T. Baldwin, “Liputan6: A Large-scale Indonesian Dataset for Text Summarization,” in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 2020, pp. 598–608. [Online]. Available: https://huggingface.co/datasets/fajrikoto/id_liputan6 & https://github.com/fajri91/sum_liputan6
- [29] A. See, P. J. Liu, and C. D. Manning, “Get To The Point: Summarization with Pointer-Generator Networks,” *arXiv preprint arXiv:1704.04368*, 2017.
- [30] Google Cloud, “Mengevaluasi model AutoML,” *Google Cloud Translation Documentation*, 2024. [Online]. Tersedia: <https://cloud.google.com/translate/docs/advanced/automl-evaluate?hl=id>.

Lampiran :

Syntac File Kaggle Notebook

```
!pip install --upgrade -q transformers huggingface_hub accelerate
datasets sentencepiece

!pip install rouge_score evaluate

!pip install -q seacrowd jsonlines nltk

from huggingface_hub import notebook_login

import os

os.environ["HF_TOKEN"] = "hf token"

print("Token Hugging Face telah diatur secara manual untuk sesi
ini.")

from datasets import load_dataset

from transformers import AutoTokenizer, AutoModelForSeq2SeqLM

print("--- Memuat Aset dari Data Manual ---")

model_name = "panggi/t5-small-indonesian-summarization-cased"

dataset_script_path = "fajrikoto/id_liputan6"

dataset_config_name = "canonical"

data_path = "/kaggle/input/sum-liputan6-data/liputan6_data"

print(f"Memuat tokenizer untuk model: {model_name}...")

tokenizer = AutoTokenizer.from_pretrained(model_name)

print(f"Memuat model: {model_name}...")

model = AutoModelForSeq2SeqLM.from_pretrained(model_name)

print(f"Mencoba memuat data dari direktori: {data_path}...")

dataset = load_dataset(dataset_script_path, dataset_config_name,
data_dir=data_path, trust_remote_code=True)

print("\n--- Dataset dari data manual berhasil dimuat! ---")

print(dataset)

print("--- Memulai Preprocessing Data ---")

prefix = "ringkas: "
```

```

def preprocess_function(examples):
    """
    Fungsi untuk melakukan tokenisasi pada input dan output.
    """
    inputs = [prefix + doc for doc in examples["clean_article"]]
    model_inputs = tokenizer(inputs, max_length=512,
truncation=True)

    labels = tokenizer(text_target=examples["clean_summary"],
max_length=128, truncation=True)

    model_inputs["labels"] = labels["input_ids"]

    return model_inputs

tokenized_dataset = dataset.map(preprocess_function, batched=True)
print("\n--- Dataset Berhasil Diproses ---")

print("Contoh Kunci dari Dataset yang Sudah Diproses:",
tokenized_dataset['train'][0].keys())

from transformers import Seq2SeqTrainingArguments
nama_repo_baru = "Irvan14/t5-small-indonesian-summarization"
print(f"Model baru akan disimpan di repositori: {nama_repo_baru}")

training_args = Seq2SeqTrainingArguments(
    output_dir=nama_repo_baru,
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    fp16=True,
    eval_strategy="steps",
    eval_steps=5000,
    save_strategy="steps",
    save_steps=5000,
    save_total_limit=2,
    load_best_model_at_end=True,

```

```

        learning_rate=2e-5,
        weight_decay=0.01,
        num_train_epochs=3,
        predict_with_generate=True,
        push_to_hub=True,
    )
    print("\nKonfigurasi untuk pelatihan berhasil dibuat.")
    !pip install -q wandb
    import wandb
    import os
    wandb.login(key='api key')
    run = wandb.init(
        project="Fine-tune T5small indo summarization",
        job_type="Fine-tuning",
        save_code=True
    )
    print("Login dan inisialisasi Weights & Biases berhasil! Sesi
    pencatatan sudah aktif.")
    from transformers import DataCollatorForSeq2Seq, Seq2SeqTrainer
    data_collator = DataCollatorForSeq2Seq(tokenizer=tokenizer,
    model=model)
    trainer = Seq2SeqTrainer(
        model=model,
        args=training_args,
        train_dataset=tokenized_dataset["train"],
        eval_dataset=tokenized_dataset["validation"],
        tokenizer=tokenizer,
        data_collator=data_collator,
    )

```

```

print("\n--- MEMULAI PROSES PELATIHAN MODEL ---")
trainer.train()

print("\n--- MENGUNGGAH MODEL FINAL KE HUGGING FACE HUB ---")
trainer.push_to_hub()

print("\n--- SELAMAT! Pelatihan model telah selesai. ---")

import numpy as np
import nltk
import evaluate

from transformers import Seq2SeqTrainingArguments, Seq2SeqTrainer,
DataCollatorForSeq2Seq

try:
    model
    tokenizer
    tokenized_dataset

    print("--- Variabel dari sesi training sebelumnya berhasil
ditemukan. ---")
except NameError:
    print("ERROR: Variabel penting tidak ditemukan. Sesi Anda
mungkin sudah di-restart.")

    print("Harap jalankan kembali langkah-langkah sebelumnya.")

device = torch.device("cuda" if torch.cuda.is_available() else
"cpu")

model.to(device)

model.eval()

print("\n--- Menyiapkan fungsi evaluasi... ---")

rouge_metric = evaluate.load('rouge')
bleu_metric = evaluate.load('sacrebleu')

def compute_metrics(eval_preds):
    preds, labels = eval_preds

    if isinstance(preds, tuple):

```

```

        preds = preds[0]

        decoded_preds = tokenizer.batch_decode(preds,
skip_special_tokens=True)

        labels = np.where(labels != -100, labels,
tokenizer.pad_token_id)

        decoded_labels = tokenizer.batch_decode(labels,
skip_special_tokens=True)

        decoded_preds_rouge =
["\n".join(nltk.sent_tokenize(pred.strip())) for pred in
decoded_preds]

        decoded_labels_rouge =
["\n".join(nltk.sent_tokenize(label.strip())) for label in
decoded_labels]

        rouge_result =
rouge_metric.compute(predictions=decoded_preds_rouge,
references=decoded_labels_rouge)

        decoded_labels_bleu = [[label] for label in decoded_labels]

        bleu_result = bleu_metric.compute(predictions=decoded_preds,
references=decoded_labels_bleu)

        result = {key: value * 100 for key, value in
rouge_result.items()}

        result["bleu"] = bleu_result["score"]

        return {k: round(v, 4) for k, v in result.items()}

print("\n--- Menjalankan evaluasi pada DATA TES dengan optimasi
GPU... ---")

training_args = Seq2SeqTrainingArguments(

    output_dir="./eval_results_test",

    per_device_eval_batch_size=16,

    fp16=True,

    predict_with_generate=True,

)

data_collator = DataCollatorForSeq2Seq(tokenizer=tokenizer,
model=model)

```

```
eval_trainer = Seq2SeqTrainer(
    model=model,
    args=training_args,
    data_collator=data_collator,
    eval_dataset=tokenized_dataset["test"],
    compute_metrics=compute_metrics,
)
results = eval_trainer.evaluate()
print("\n=====")
print("    HASIL EVALUASI FINAL PADA DATA TES")
print("=====")
print(results)
```

Syntac File Streamlit

Buat meta deskripsi



```
import streamlit as st
import torch
from transformers import T5Tokenizer, T5ForConditionalGeneration
import re

st.set_page_config(
    page_title="Peringkasan Meta SEO",
    page_icon=" ",
    layout="wide"
)

@st.cache_resource
def load_components():
    """Memuat model dan tokenizer dari Hugging Face Hub."""
    repo_id = "Irvan14/t5-small-indonesian-summarization"
    try:
```

```

        tokenizer = T5Tokenizer.from_pretrained(repo_id)
        model = T5ForConditionalGeneration.from_pretrained(repo_id)
        return tokenizer, model

    except Exception as e:
        st.error(f"Gagal memuat komponen dari Hub. Kesalahan: {e}")
        return None, None

def smart_truncate(text, target_length=155):
    """Memotong teks secara cerdas di akhir kalimat terakhir sebelum
    target_length."""
    if len(text) <= target_length:
        return text

    truncated_text = text[:target_length]
    last_period_index = truncated_text.rfind('.')
    if last_period_index != -1:
        return truncated_text[:last_period_index + 1]
    else:
        last_space_index = truncated_text.rfind(' ')
        if last_space_index != -1:
            return truncated_text[:last_space_index] + "."
        else:
            return truncated_text

tokenizer, model = load_components()
if 'summary_result' not in st.session_state:
    st.session_state.summary_result = ""

col1, col2 = st.columns(2)
with col1:
    st.subheader("Masukkan Teks")

    article_title = st.text_input("Judul Artikel",
    placeholder="Masukkan judul artikel Anda di sini...")

```

```

    article_text = st.text_area("Isi Konten Artikel", height=300,
placeholder="Tempelkan isi konten artikel Anda di sini...")

    submit_button = st.button("Buat Deskripsi Meta", type="primary")
with col2:

    st.subheader("Deskripsi Meta SEO")

    output_text_value = st.session_state.summary_result

    if output_text_value:

        st.info(output_text_value)

    else:

        st.info("Deskripsi meta Anda akan muncul di sini...")

    char_count = len(output_text_value)

    st.caption(f"{char_count} karakter. Ideal: 130-155.")

    st.subheader("Pratinjau SEO")

    with st.container(border=True):

        preview_title = article_title or
st.session_state.get('prev_title', 'Contoh Judul Halaman')

        st.markdown(f"<h5 style='color:
#4A90E2;'>{preview_title}</h5>", unsafe_allow_html=True)

        st.markdown(f"<p style='color: #4B5563;'>{output_text_value
or 'Di sinilah deskripsi meta Anda akan muncul...'}</p>",
unsafe_allow_html=True)

    if submit_button and tokenizer and model:

        if not article_text or len(article_text) < 150:

            st.warning("Masukkan setidaknya 150 karakter artikel untuk
hasil terbaik.")

        else:

            with st.spinner("Membuat ringkasan dengan mode kreatif..."):

                try:

                    source_text_for_summary =
f"{article_title}\n\n{article_text}"

                    input_text = "ringkas: " + source_text_for_summary

```

```

        device = "cuda" if torch.cuda.is_available() else
"cpu"

        inputs = tokenizer(input_text, return_tensors="pt",
max_length=512, truncation=True).to(device)

        model.to(device)

        summary_ids = model.generate(
            inputs['input_ids'],
            max_length=80,
            min_length=30,
            do_sample=True,
            top_k=50,
            top_p=0.95,
            repetition_penalty=2.5,
            num_beams=1
        )

        raw_summary = tokenizer.decode(summary_ids[0],
skip_special_tokens=True)

        final_summary = smart_truncate(raw_summary)
        st.session_state.summary_result = final_summary
        st.session_state.prev_title = article_title
        st.rerun()

    except Exception as e:
        st.error(f"Terjadi kesalahan saat proses
peringkasan: {e}")

```