IMPLEMENTASI DAN EVALUASI MODEL *T5* (*TEXT-TO-TEXT TRANSFER TRANSFORMER)* UNTUK *TITLE GENERATOR* ARTIKEL ILMIAH BERDASARKAN ABSTRAK

SKRIPSI

Oleh :

HASAN ILHAM ZAUZAN
NIM. 210605110094

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

IMPLEMENTASI DAN EVALUASI MODEL *T5* (*TEXT-TO-TEXT TRANSFER TRANSFORMER)* UNTUK *TITLE GENERATOR* ARTIKEL ILMIAH BERDASARKAN ABSTRAK

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
HASAN ILHAM ZAUZAN
NIM. 210605110094

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2025

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

IMPLEMENTASI DAN EVALUASI MODEL *T5 (TEXT-TO-TEXT TRANSFER TRANSFORMER)* UNTUK *TITLE GENERATOR* ARTIKEL ILMIAH BERDASARKAN ABSTRAK

SKRIPSI

Oleh :
Hasan Ilham Zauzan
NIM. 210605110094

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 13 Juni 2025

Pembimbing I,	Pembimbing II,
Dr. Zainal Abidin, M.Kom	H. Fatchurrochman, M.Kom
NIP. 19760613 200501 1 004	NIP. 19700731 200501 1 002

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

iii

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

IMPLEMENTASI DAN EVALUASI MODEL *T5 (TEXT-TO-TEXT TRANSFER TRANSFORMER)* UNTUK *TITLE GENERATOR* ARTIKEL ILMIAH BERDASARKAN ABSTRAK

SKRIPSI

Oleh :
Hasan Ilham Zauzan
NIM. 210605110094

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 20 Juni 2025

Susunan Dewan Penguji

Ketua Penguji	: Nur Fitriyah Ayu Tunjung Sari, M.Cs NIP. 19911226 202012 2 001	( )
Anggota Penguji I	: Khadijah Fahmi Hayati Holle, M.Kom NIP. 19900626 202203 2 002	( )
Anggota Penguji II	: Dr. Zainal Abiddin, M.Kom NIP. 19760613 200501 1 004	( )
Anggota Penguji III	: Fatchurrohman, M.Kom NIP. 19700731 200501 1 002	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Ir. Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama	: HASAN ILHAM ZAUZAN
NIM	: 210605110094
Fakultas / Prodi	: Sains dan Teknologi / Teknik Informatika
Judul Skripsi	: Implementasi dan Evaluasi Model *T5* (*Text-To-Text Transfer Transformer*) untuk *Title Generator* Artikel Ilmiah Berdasarkan Abstrak

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 13 April 2025
Yang membuat pernyataan,

Hasan Ilham Zauzan
NIM. 210605110094

MOTTO

Balance in all things

HALAMAN PERSEMBAHAN

Puji syukur atas kehadirat Allah Subhanahu wa ta'ala, karena berkat rahmat dan petunjuk-Nya penulis dapat menyelesaikan skripsi ini dengan lancar.

Penulis persembahkan skripsi ini kepada kedua kedua orang tua, keluarga, dosen, guru, sahabat, dan semua pihak yang telah membantu secara aktif dalam menyelesaikan penelitian ini.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang senantiasa memberikan rahmat dan kesehatan, sehingga penulis mampu menyelesaikan skripsi ini dengan baik. Penulis menyampaikan ucapan Terimakasih kepada semua pihak yang pernah terlibat langsung maupun tidak langsung dalam menyelesaikan skripsi ini, bukan hanya karena usaha keras dari penulis sendiri, akan tetapi karena adanya dukungan dari berbagai pihak. Oleh karena itu penulis berterima kasih kepada:

1. Prof. Dr. M. Zainuddin, M.A., selaku rektor Universitas islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Hariani, M.Si., selaku dekan Fakutas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Ir. Fachrul Kurniawan, M.MT., IPU, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Zainal Abidin, M.Kom selaku Dosen Pembimbing I yang telah dengan sabar memberikan arahan baik dalam penulisan hingga program yang dibuat dalam menyelesaikan skripsi ini.
5. Fatchurrochman, M.Kom selaku Dosen Pembimbing II yang telah memberikan bimbingan, arahan serta bantuan dalam terwujudnya karya tulis skripsi ini.
6. Nur Fitriyah Ayu Tunjung Sari, M.Cs selaku ketua penguji, Khadijah Fahmi Hayati Holle, M.Kom, selaku penguji I yang telah meluangkan waktunya untuk menguji dan dengan sabar memberi arahan serta saran dalam menyelesaikan skripsi ini.

7. Ajib Hanani, M.T selaku Dosen Wali yang telah memberikan arahan selama proses perkuliahan.
8. Nia Faricha, S.Si selaku Admin Program Studi Teknik Informatika yang dengan sabar membantu, memberikan arahan informasi terkait perkuliahan.
9. Kedua orangtua penulis Bapak Tarianto dan Ibu Susanah yang selalu memberi dukungan dan doa serta selalu memberikan yang terbaik untuk kelancaran putrinya dalam pendidikan.
10. Seluruh kawan-kawanku Kontrakan Candi dan Discord UIN Maulana Maliki, yang telah membersamai penulis dan meramaikan harinya, terima kasih atas segala bentuk adu nasib, dan manifestasinya yang menjadi motivasi penulis dalam proses pengerjaan skripsi ini.
11. Zalfa Zahiyah yang telah menemani, memberikan semangat, serta membantu penulis selama proses pengerjaan skripsi ini.
12. Seluruh pihak yang terlibat namun tidak bisa penulis sebutkan satu persatu. Semoga Allah SWT membalas kebaikan dan mengampuni segala dosa kita semua.

Akhir kata, penulis mengakui bahwa penulisan pada skripsi ini masih banyak kekurangan. Penulis berharap semoga skripsi ini diterima sebagai amal ibadah yang tulus dan bermanfaat di sisi Allah Subhanahu Wa Ta'ala.

Malang, 17 Juni 2025

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN ..... **ii**
HALAMAN PERSETUJUAN ..... **iii**
HALAMAN PENGESAHAN ..... **iv**
PERNYATAAN KEASLIAN TULISAN ..... **v**
MOTTO ..... **vi**
HALAMAN PERSEMBAHAN ..... **vii**
KATA PENGANTAR ..... **viii**
DAFTAR ISI ..... **x**
DAFTAR GAMBAR ..... **xii**
DAFTAR TABEL ..... **xiii**
ABSTRAK ..... **xiv**
ABSTRACT ..... **xv**
مستخلص البحث ..... **xvi**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Tujuan Penelitian 5
1.4 Batasan Masalah 5
1.5 Manfaat Penelitian 5
BAB II STUDI PUSTAKA **6**
2.1 Penelitian Terkait 6
2.2 Text Generation 10
2.3 Transformer 11
2.4 Text-to-Text Transfer Transformer (*T5*) 17
2.4.1 Arsitektur Model *T5* 18
2.4.2 Perbedaan Parameter Model *T5* 21
2.4.3 *Preprocessing* dan Tokenisasi 23
BAB III METODE PENELITIAN **26**
3.1 Desain Penelitian 26
3.2 Dataset 26
3.3 *Preprocessing* 28
3.4 *Training* dan *Testing* Data 29
3.4.1 Model *T5* 29
3.4.2 *Training* Model *T5* 31
3.4.3 *Testing* Model *T5* 33
3.5 Evaluasi Matrik *ROUGE* 34
3.6 Deployment 36
3.7 Skenario Uji Coba 36
BAB IV HASIL DAN PEMBAHASAN **38**
4.1 Dataset 38

4.2 Proses *Training* & *Testing* Model 38
4.2.1 *Training* Model *T5* 39
4.2.2 *Testing* Model *T5* 40
4.3 Hasil Skenario Uji Coba 41
4.4 Hasil Evaluasi *ROUGE* 42
4.4.1 Analisis Skor *ROUGE* Tinggi 44
4.4.2 Analisis Skor ROUGE Rendah 46
4.5 Deployment 48
4.6 Pembahasan 49
4.7 Integrasi 54
BAB V KESIMPULAN DAN SARAN 57
5.1 Kesimpulan 57
5.2 Saran 58
DAFTAR PUSTAKA
LAMPIRAN

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Transformer (Vaswani et al., 2017) 12
Gambar 2. 2 Scaled Dot-Product Attention (Vaswani et al., 2017) 14
Gambar 2. 3 Perbedaan Arsitektur Model *T5* (Pal et al., n.d.) 20
Gambar 3. 1 Desain Penelitian 26
Gambar 3. 2 Dataset Artikel Ilmiah 27
Gambar 3. 3 Distribusi Token Abstract 27
Gambar 3. 4 Flowchart *Training* Model *T5* 32
Gambar 4. 1 Distribusi Panjang Token 38
Gambar 4. 2 Interface Title Generator 48

DAFTAR TABEL

Tabel 2. 1 Penelitian terkait 9
Tabel 2. 2 Perbandingan Varian Model *T5* 23
Tabel 3. 1 Perbandingan data setelah preprocessing..... 28
Tabel 3. 2 Jenis Matrik *ROUGE* 34
Tabel 3. 3 *Hyperparameter* Model *T5* 37
Tabel 4. 1 *Hyperparameter Tuning*..... 41
Tabel 4. 2 Konfigurasi Skenario Terbaik..... 43
Tabel 4. 3 Hasil Evaluasi *ROUGE*..... 43
Tabel 4. 4 Judul yang dihasilkan sistem dengan *ROUGE* tertinggi..... 44
Tabel 4. 5 Judul hasil sistem dengan *ROUGE* terendah 46
Tabel 4. 6 Sampel data yang diuji Validator..... 51

ABSTRAK

Zauzan, Hasan Ilham. 2025. **Implementasi dan Evaluasi Model *T5 (Text-To-Text Transfer Transformer)* Untuk *Title Generator* Artikel Ilmiah Berdasarkan Abstrak**. Skripsi, PROGRAM STUDI Teknik Informatika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (I) Dr. Zainal Abidin, M.Kom, (II) H. Fatchurrochman, M.Kom.

Kata Kunci: *Natural Language Processing, T5, Title Generation*

Penyusunan judul artikel ilmiah yang representatif dan menarik merupakan tantangan krusial karena seringkali memakan waktu dan bersifat subjektif, yang dapat menurunkan visibilitas karya ilmiah. Untuk mengatasinya, penelitian ini mengimplementasikan dan mengevaluasi model *T5-small* untuk menghasilkan judul secara otomatis berdasarkan abstrak, menggunakan 5.000 data artikel COVID-19 dari PubMed. Model dilatih dan dioptimalkan melalui *hyperparameter tuning* untuk menemukan konfigurasi terbaik. Hasil evaluasi dengan konfigurasi terbaik (*epoch* 6, *batch size* 4) menunjukkan pencapaian skor *F1-score ROUGE*-1 sebesar 0.413, *ROUGE*-2 0.217, dan *ROUGE*-L 0.345. Analisis kualitatif mengungkapkan bahwa skor ROUGE yang tinggi tidak selalu berkorelasi dengan kualitas gramatikal, di mana beberapa judul terpotong. Sebaliknya, judul yang baik secara semantik terkadang menerima skor rendah. Dari 50 sampel, 56% judul dinilai valid. Secara keseluruhan, penelitian ini menunjukkan model *T5-small* merupakan pendekatan yang efektif dan efisien untuk pembuatan judul otomatis yang mampu bersaing dengan arsitektur lebih besar.

ABSTRACT

Zauzan, Hasan Ilham. 2025. **Implementation and Evaluation of the *T5* (Text-To-Text Transfer Transformer) Model for Title Generator of Scientific Articles Based on Abstracts**. Thesis, Informatics Engineering Study Program. Faculty of Science and Technology. State Islamic University Maulana Malik Ibrahim Malang. Advisor: (I) Dr. Zainal Abidin, M.Kom, (II) H. Fatchurrochman, M.Kom.

Keywords: *Natural Language Processing, T5, Title Generation*

Crafting representative and engaging scientific article titles is a crucial challenge as it is often time-consuming and subjective, which can decrease the visibility of a scientific work. To address this, this research implements and evaluates the *T5*-small model to automatically generate titles from abstracts, using 5,000 COVID-19 article data from PubMed. The model was trained and optimized through hyperparameter tuning to find the best configuration. The evaluation with the best configuration (6 *epoch*s, batch size 4) achieved F1-scores of 0.413 for ROUGE-1, 0.217 for ROUGE-2, and 0.345 for ROUGE-L. Qualitative analysis reveals that high ROUGE scores do not always correlate with grammatical quality, with some titles being truncated. Conversely, semantically good titles sometimes received low scores. Out of 50 samples, 56% of the titles were deemed valid. Overall, this study demonstrates that the *T5*-small model is an effective and efficient approach for automatic title generation, capable of competing with larger architectures.

مستخلص البحث

زوزان ، .حسن إلهام ٢٠٢٥. **تنفيذ وتقييم نموذج *T5* (Text-To-Text Transfer Transformer) لمولد عناوين المقالات العلمية بناءً على الملخصات** ، أطروحة. قسم هندسة المعلوماتية. كلية العلوم والتكنولوجيا. جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (أولا) د. زين العابدين, الماجستير (ثانيا) فتح الرحمن, الماجستير

الكلمات المفتاحية: معالجة اللغة الطبيعية، *T5*، توليد العنوان

يعد توليد عناوين تمثيلية وجذابة للمقالات العلمية تحديًا حاسمًا لأنه غالبًا ما يستغرق وقتًا طويلاً وغير موضوعي، مما قد يقلل لتوليد العناوين تلقائيًا استنادًا إلى *T5*-small من ظهور الأوراق العلمية. ولمعالجة هذا الأمر، تقوم هذه الدراسة بتطبيق وتقييم نموذج وقد تم تدريب النموذج وتحسينه من خلال ضبط المعلمة .PubMed الملخصات، باستخدام 5000 مقالة عن كوفيد-19 من F1 ROUGE-1 الفائقة للعثور على أفضل تكوين. أظهرت نتائج التقييم بأفضل تكوين (الحلقة 6، حجم الدُفعة 4) تحقيق درجة البالغة 0.345. كشف التحليل النوعي أن درجات ROUGE-Lالبالغة 0.217، و ROUGE-2 البالغة 0.413، و المرتفعة لا ترتبط بالضرورة بالجودة النحوية، حيث كانت بعض العناوين مبتورة. وعلى العكس من ذلك، حصلت العناوين ROUGE الجيدة من الناحية الدلالية في بعض الأحيان على درجات منخفضة. من بين 50 عينة، تم تصنيف 56% من العناوين على أنها هو نهج فعال وكفء لتوليد العناوين التلقائي الذي يمكن أن ينافس *T5*-small صالحة. بشكل عام، تُظهر هذه الدراسة أن نموذج ..البنى الأكبر

BAB I

PENDAHULUAN

1.1 Latar Belakang

Judul adalah nama yang digunakan untuk memberikan suatu identitas pada sebuah karya tulis. Fungsi dari judul adalah sebagai gambaran dari seluruh isi teks yang dapat menarik perhatian pembaca. Judul juga merupakan bagian penting dalam artikel ilmiah karena mencerminkan isi dari dokumen tersebut. Judul yang baik harus mampu memberikan gambaran yang singkat namun jelas mengenai topik yang dibahas, sehingga pembaca dapat memahami konteks dari artikel tanpa harus membaca kesulurahan isinya (Cals & Kotz, 2013). Penulisan judul yang baik dapat meningkatkan visibilitas dari artikel di dalam database dan mesin pencarian seperti Google Scholar, sehingga memiliki peluang yang lebih besar untuk dikutip oleh peneliti.

Selain berperan dalam menarik perhatian pembaca, judul juga berfungsi sebagai sistem pengindeksan artikel ilmiah (Rossi & Brand, 2020). Dengan pemilihan kata kunci yang tepat, artikel akan lebih mudah ditemukan oleh peneliti atau penulis yang mencari referensi terkait. Sebaliknya, judul yang kurang spesifik atau terlalu umum dapat menyebabkan artikel sulit ditemukan dan kurang mendapatkan perhatian (Chen & Eger, 2022). Oleh karena itu, pemilihan judul harus dilakukan dengan cermat agar dapat mencerminkan isi penelitian serta memudahkan pencarian bagi pembaca maupun penulis.

Meskipun memiliki peran yang sangat penting, penyusunan judul sering kali menjadi tantangan bagi penulis. Proses ini bersifat subjektif serta bergantung pada pemahaman penulis terhadap bidang yang diteliti. Tidak adanya pedoman baku dalam pembuatan judul menyebabkan variasi dalam format dan gaya penulisannya, sehingga tidak semua judul yang dibuat dapat menggambarkan isi penelitian secara optimal (Fox & Burns, 2015). Dalam beberapa kasus, judul yang terlalu panjang atau terlalu pendek dapat menyebabkan kurangnya keterbacaan dan efektivitas dalam menyampaikan informasi utama dari penelitian (Jin & Hauptmann, 2001).

Selain itu, penulis yang tidak memiliki latar belakang bahasa akademik yang kuat sering kali mengalami kesulitan dalam menyusun judul yang sesuai dengan standar akademik. Permasalahan ini sering dialami oleh peneliti yang menggunakan bahasa tidak baku dalam publikasi mereka. Proses manual dalam menentukan judul juga sering kali memakan waktu yang cukup lama dan tidak selalu menghasilkan judul yang optimal (Jiang & Hyland, 2023). Oleh karena itu, diperlukan pendekatan yang lebih sistematis dan efisien untuk membantu penulis dalam menghasilkan judul yang lebih representatif dengan isi penelitian.

Kemajuan di bidang permrosesan bahasa alami telah membuka peluang baru dalam otomatisasi pemrosesan teks, termasuk pembuatan judul secara otomatis (Hanif, 2023). Salah satu teknologi terbaru yang terbukti efektif dalam tugas ini adalah model *Transformer*, seperti *T5 (Text-to-Text Transfer Transformer)*. Model *T5* bekerja dengan menggunakan mekanisme *self-attention*, yang memungkinkan sistem untuk memahami hubungan antara kata dalam sebuah teks dengan lebih akurat. Berbeda dengan pendekatan berbasis aturan linguistik atau metode statistik,

T5 mampu mempelajari pola bahasa dari dataset yang luas dan menghasilkan judul yang lebih natural juga sesuai dengan standar akademik (Rehman et al., 2025). Dengan teknologi ini, pembuatan judul tidak hanya menjadi lebih cepat, tetapi juga lebih konsisten dan sesuai dengan konteks ilmiah yang berlaku.

Model ini dirancang untuk menangkap konteks dengan lebih akurat dibandingkan model sebelumnya, seperti *Recurrent Neural Network (RNN)* atau *Long Short-Term Memory (LSTM),* sehingga dapat menghasilkan judul yang lebih sesuai dengan isi penelitian (Lodhwal & Choudhary, 2023). Model *T5* juga memiliki keunggulan lain seperti fleksibilitas dan performa yang lebih tinggi di berbagai tugas *NLP* karena pendekatan *text-to-text*. Sering kali model ini dapat melebihi Model *Transformer* lainnya seperti *BERT* dan *GPT* dalam tugas yang membutuhkan pemahaman dalam generasi teks, serta sangat mudah di-fine-tune untuk berbagai aplikasi tanpa perubahan arsitektur.

Berdasarkan permasalahan yang telah diidentifikasi, penelitian ini bertujuan untuk mengeksplorasi model berbasis *Transformer* yang dapat menghasilkan judul artikel ilmiah secara otomatis berdasarkan abstrak. Asbtrak merupakan ringkasan singkat yang mencakup informasi inti dari sebuah karya tulis ilmiah. Hal ini membuat abstrak menjadi sangat cocok sebagai input untuk *title generator* karena mengandung informasi yang relevan dengan isi karya tulis ilmiah. Dengan adanya sistem ini, diharap dapat membantu akademisi, peneliti, dan mahasiswa agar dapat lebih mudah membuat judul yang lebih akurat, relevan, serta efisien. Manfaat ini sangat sesuai dengan firman Allah SWT. yang diterangkan pada surat Al-Insyirah ayat 5-6 yang berbunyi:

فَإِنَّ مَعَ ٱلۡعُسۡرِ يُسۡرًا (٥) إِنَّ مَعَ ٱلۡعُسۡرِ يُسۡرًا (٦)

"*Maka sesungguhnya beserta kesulitan ada kemudahan. Sesungguhnya beserta kesulitan ada kemudahan.*"*(QS. Al-Insyirah: 6)*

Ayat diatas sangat sesuai dengan manfaat penelitian karena memudahkan akademisi dan penulis karya tulis ilmiah dalam menyusun judul untuk karya mereka. Selain membantu dalam pembuatan judul otomatis, penelitian ini juga membuka eksplorasi baru di bidang pemrosesan bahasa alami. Pendekatan otomatis ini juga diharapkan dapat mengurangi subjektivitas dalam pembuatan judul serta meningkatkan kualitas standar dalam publikasi akademik.

Selain itu, penelitian ini diharapkan dapat menjadi referensi bagi pengembangan lebih lanjut dalam pemanfaatan *NLP* untuk berbagai tugas akademik lainnya, seperti peringkasan teks dan klasifikasi dokumen ilmiah. Penelitian ini juga diharap dapat berkontribusi terhadap perkembangan teknologi kecerdasan buatan di bidang akademik dan penerbitan ilmiah. Karena semakin banyaknya artikel yang dipublikasikan setiap tahun, penyusunan judul secara otomatis diharap dapat membantu mengurangi beban kerja editor jurnal serta mempercepat proses publikasi.

1.2 Rumusan Masalah

Seberapa tinggi *ROUGE* yang dihasilkan oleh model *T5* dan implementasinya untuk menghasilkan judul menggunakan abstrak dari artikel ilmiah?

1.3 Tujuan Penelitian

Mengetahui Seberapa tinggi *ROUGE* yang dihasilkan oleh model *T5* dan implementasinya untuk menghasilkan judul menggunakan abstrak dari artikel ilmiah.

1.4 Batasan Masalah

a. Dataset yang digunakan merupakan pasangan abstrak dan judul artikel ilmiah yang menggunakan Bahasa Inggris.
b. Sistem yang dikembangkan pada penelitian hanya sampai proses deployment dan tidak sampai diujikan pada user.
c. Deployment yang dibangun hanya dapat menerima abstrak dari artikel berbahasa inggris.

1.5 Manfaat Penelitian

a. Memudahkan peneliti dalam menyusun serta menentukan judul untuk artikel ilmiah dengan lebih cepat dan efisien.
b. Mengeksplorasi model *Transformer* untuk pembuatan judul secara otomatis.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terkait

Penelitian yang dilakukan oleh (Lodhwal & Choudhary, 2023) membahas pendekatan dalam generasi judul otomatis yang berfokus pada implementasi metode *Recurrent Neural Network (RNN)* dan *Transformer* yang sudah dilatih sebelumnya. Penelitian ini berfokus pada evaluasi dan perbandingan kinerja model *NLP* tersebut dalam tugas generasi judul otomatis terutama dalam konteks akademik. Hasil penelitian menunjukkan model berbasis *Transfomer* seperti *GPT-2* dan *GPT-Neo* menghasilkan judul yang lebih baik dalam hal relevansi dan kesesuaian dengan abstrak dibandingkan dengan model *LSTM*.

(Putra & Khodra, 2017) dalam penelitiannya yang berjudul *"Automatic Title Generation in Scientific Articles for Authorship Assistance"* membahas tentang bagaimana mengembangkan sebuah sistem yang mampu menghasilkan judul secara otomatis menggunakan metode *Adaptive K-Nearest Neighbor (AKNN).* Metode tersebut digunakan untuk mengklasifikasi kalimat berdasarkan *rhetorical categories,* hal ini *Bertujuan* untuk membantu sistem memilih bagian abstrak yang paling relevan dalam proses pembuatan judul otomatis. Hasil penelitian menunjukkan metode *AKNN* dapat menghasilkan judul yang cukup baik. Dari hasil evaluasi menggunakan metrik F1-score, hasil yang diperoleh menunjukkan bahwa model lebih efektif dalam menghasilkan judul untuk artikel dalam bidang linguistik

dibandingkan bidang kimia, dengan skor F1 yang bervariasi antara 0,109 hingga 0,255.

Selanjutnya penelitian yang dilakukan oleh (Deshmukh et al., n.d.) membahas tentang bagaimana mengembangkan model *LSTM* untuk melakukan ringkasan otomatis *(text summarization)* sekaligus menghasilkan judul yang sesuai dengan isi dokumen. Penelitian ini menggunakan pendekatan *sequence-to-sequence,* di mana model dilatih untuk memetakan input berupa teks yang panjang ke bentuk output berupa teks yang lebih ringkas. Hasil penelitian ini yaitu model mampu menghasilkan ringkasan dan judul yang lebih informatif dan singkat dengan akurasi model 95,2 lebih tinggi dibandinkan dengan metode yang lain seperti *SVM, E-RNN, CNN-RNN,* dll.

Penelitian selanjutnya dilakukan oleh (Mishra et al., 2021) yang membahas tentang bagaimana mengembangkan sistem untuk menghasilkan judul otomatis menggunakan model *GPT-2* dengan beberapa tahapan pendekatan yaitu, yang pertama model *GPT-2* digunakan untuk menghasilkan beberapa kandidat judul berdasarkan abstrak, selanjutnya dilakukan seleksi judul dengan membandingkan kesamaan makna antara judul yang dihasilkan dan abstrak serta mengidentifikasi kemiripan dalam penggunaan frasa kata benda antara keduanya, terakhir judul yang terpilih diperbaiki menggunakan model *GPT-2* agar lebih berhubungan dan sesuai dengan standar akademik. Hasil dari penelitian adalah dilakukannya evaluasi menggunakan metrik *ROUGE* dan *BLEU* yang menunjukkan model *Transformer* memiliki performa yang lebih baik dibanding metode sebelumnya seperti *TextRank, LexRank,* dan *Bi-GRU* dengan *Attention* dengan akurasi model 0,376 untuk

ROUGE-1, 0.188 untuk *ROUGE-2,* 0.336 untuk *ROUGE-L,* dan 0.124 untuk *BLEU-4*.

Penelitian dengan judul *"Research Paper Summarization Using Text-To-Text Transfer Transformer (T5) Model"* yang dilakukan oleh (Hanif, 2023) mengangkat topik utama terkait efisensi dari memahami isi dokumen secara singkat yang memerlukan sebuah sistem untuk meringkas dokumen. Penelitian ini membahas terkait evaluasi tingkat efektivitas dari model *T5* dalam merangkum sebuah dokumen artikel ilmiah. Hasil penelitian menunjukkan bahwa model tersebut mampu menghasilkan ringkasan yang informatif serta tetap mempertahankan isi dari artikel ilmiah. Evaluasi dilakukan menggunakan metrik *ROUGE* dan *BLEU* yang menunjukkan model *T5* lebih unggul dibanding model lain seperti *BERT, GPT-2,* dan *BART* dengan akurasi *ROUGE-1* 83%, *ROUGE-2* 82%, *ROUGE-L* 83%, dan *BLEU* 47%.

Penelitian yang dilakukan oleh (Rashmi & Maheswari, 2023)) menunjukkan bahwa penggunaan *LSTM* dengan *Beam search* dapat meningkatkan kualitas judul yang dihasilkan dibandingkan dengan metode seperti *Vanila RNN* dan *LSTM* tanpa *Beam search.* Metode yang menggunakan *Beam search* mampu menghasilkan judul yang lebih informatif, relevan, dan sesuai dengan konteks. Teknik *Beam search* memungkinkan eksplorasi lebih banyak judul sebelum memilih yang paling sesuai, sehingga dapat mengurangi kemungkinan menghasilkan judul yang kurang representatif terhadap dokumen aslinya. Hasil evaluasi penelitian menggunakan 4 model dan 4 dataset menunjukkan bahwa model yang menggunakan *Beam search* memiliki rata- rata akurasi sebesar 96,25.

Berdasarkan penelitian sebelumnya, metode yang digunakan memiliki keterbatasan dalam menangkap teks yang panjang seperti halnya penelitian yang dilakukan (Rashmi & Maheswari, 2023) yaitu menggunakan *LSTM* dan *Vanilla RNN*. Penelitian ini menggunakan model yang dapat menyelesaikan permasalahan tersebut, dengan menerapkan arsitektur *Transformer*. Keunggulan dari model *Transformer* yaitu dalam mekanisme self-*attention*, dimana memungkinkan model untuk memperhatikan seluruh bagian input secara bersamaan, tanpa harus memprosesnya secara berurutan seperti pada *LSTM* atau *RNN*. Hal ini dapat meningkatkan efisiensi komputasi, serta memungkinkan model dalam memahami konteks global dari sebuah teks, termasuk informasi yang tersebar di seluruh bagian abstrak.

Tabel 2. 1 Penelitian terkait

No	Nama	Judul	Dataset	Metode	Hasil
1.	(Lodhwal & Choudhary, 2023)	Survey Paper: Automatic Title Generation for Text with RNN and Pre-trained *Transformer* Language Model	CNN, Gigaword, NEWSROOM	*GPT*-2, *GPT*-Neo, Chat*GPT*, dan LSTM	Hasil dari penelitian adalah model yang dihasilkan oleh *Transformer* lebih baik dibandingkan dengan model yang dihasilkan oleh LSTM.
2.	(Putra & Khodra, 2017)	Automatic Title Generation in Scientific Articles for Authorship Assistance	Dataset bidang Computational Linguistic, Datset bidang Kimia	AKNN	Hasil evaluasi menggunakan matrik F1-score menunjukkan bahwa model lebih efektif dalam menghasilkan judul untuk bidang linguistik. Untuk score F1 diantara 0,109 – 0,255
3.	(Deshmukh et al., n.d.)	Automatic Text Summarization and Title Generation Using LSTM	Abstrak artikel ilmiah	LSTM & sequence-to-sequence	Hasil penelitian adalah model LSTM dapat menghasilkan ringkasan judul yang lebih singkat dan informatif dengan skor akurasi model 95,2

No	Nama	Judul	Dataset	Metode	Hasil
4.	(Mishra et al., 2021)	Automatic Title Generation for Text with Pre-trained *Transformer* Language Model	Dataset berasal dari jurnal Arxiv	*GPT*-2	Hasil penelitian adalah model yang digunakan dapat menghasilkan judul yang lebih akurat dibandingkan dengan metode baseline seperti TextRank
5.	(Hanif, 2023)	Research Paper Summarization Using Text-To-Text Transfer *Transformer* (*T5*) Model	Abstrak dan ringkasan artikel ilmiah	Text-to-Text Transfer *Transformer* (*T5*)	Hasil penelitian menunjukkan Model *T5* lebih unggul dibanding model lain seperti *BERT, GPT-2*, dan *BART* dengan akurasi *ROUGE-1* 83%, *ROUGE-2* 82%, *ROUGE-L* 83%, dan *BLEU* 47%.
6.	(Rashmi & Maheswari, 2023)	Title Generation Using Lstm With *Beam search* For Enhancing The Contextually Relevant Headlines	Artikel berita dan judul	Vanilla RNN, LSTM, LSTM *Beam search*	Model LSTM dengan *Beam search* mampu menghasilkan judul yang lebih akurat dibanding model lain dengan akurasi rata-rata 96,25

2.2 Text Generation

Text Generation merupakan proses menghasilkan teks secara otomatis oleh sistem komputer berdasarkan input tertentu, seperti teks, kalimat, atau bahkan hanya kata kunci. Tujuan utama dari *text generation* adalah menciptakan output teks yang relevan dan menyerupai bahasa alami manusia (Yu et al., 2022). Proses ini banyak digunakan dalam berbagai aplikasi seperti dialog sistem, pembuatan konten otomatis, penulisan ringkasan, serta *title generation* atau pembuatan judul artikel ilmiah. Secara umum pendekatan dalam *text generation* dapat dibagi menjadi 2 jenis yaitu menggunakan aturan linguistik (*rule-based generation*) untuk menyusun kalimat dan neural text generation yang memanfaatkan model

pembelajaran mesin, khusunya *deep learning* untuk mempelajari pola bahasa dari data dan menghasilkan teks baru yang lebih alami.

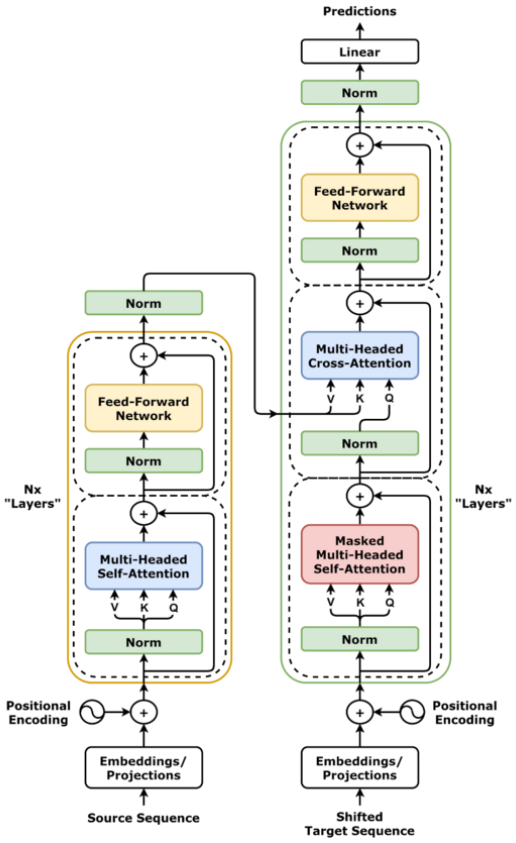
Dalam pengembangan *text generation* yang modern, pendekatan berbasis *neural network* seperti *RNN*, *LSTM*, hingga *Transformer* menjadi metode yang dominan digunakan. *Model Transformer* telah menunjukkan performa yang sangat baik dalam menghasilkan teks berkualitas tinggi yang sesuai dengan konteks dari input. Dalam konteks ilmiah, *text generation* dapat dimanfaatkan untuk berbagai keperluan, seperti pembuatan judul artikel otomatis, pembuatan abstrak, ataupun peringkasan teks otomatis. Dengan meningkatnya kemampuan model dalam memahami dan memproduksi teks secara otomatis, *text generation* tidak hanya mempercepat proses produksi konten ilmiah, tetapi juga membuka peluang dalam mengembangkan sistem yang dapat membantu penulis dalam Menyusun elemen penting dalam publikasi akademik secara otomatis dan efisien.

2.3 Transformer

Transformer merupakan model deep learning berbasis *self-attention* yang dikembangkan oleh (Vaswani et al., 2017) untuk mengatasi keterbatasan dari model berbasis *RNN* dan *CNN* dalam pemrosesan bahasa alami. Berbeda dengan model sebelumnya yang menggunakan pemrosesan secara sekuensial (satu per satu), *Transformer* memungkinkan pemrosesan *parallel* (secara bersamaan) yang lebih efisien dan skalabel. Sebelum adanya *Transformer*, model bahasa alami banyak menggunakan pendekatan berbasis *RNN*, *LSTM*, dan *GRU* dimana memiliki keterbatasan dalam menangani urutan panjang suatu kalimat atau teks yang yang

terdiri dari banyak kata disebabkan oleh *vanishing gradient*, serta sulitnya menangkap hubungan antar kata yang berjauhan dalam teks.

Transformer memperkenalkan pendekatan baru dengan menggunakan *self-attention* sebagai mekanisme utama (Edelman et al., 2022). Model *Transformer* menggunakan arsitektur *encoder-decoder*, dimana *encoder* mengubah input teks menjadi representasi vektor dan *decoder* menghasilkan output berdasarkan representasi tersebut. Layer utama encoder meliputi *Self-Attention Layer*, *Feed-forward Neural Network*, dan Layer *Normalization & Residual Connection*. Sedangkan Layer utama dari decoder adalah *Masked Self-Attention*, *Encoder-Decoder Attention*, dan *Feed-forward Neural Network*.



Gambar 2. 1 Arsitektur *Transformer* (Vaswani et al., 2017)

Gambar 2. 1 memperlihatkan arsitektur *Transformer* yang tersusun atas 2 blok yaitu *encoder* disebelah kiri dan *decoder* disebelah kanan.

a. *Input Embedding Layer*

Merupakan layer yang merubah input berupa kata maupun sub-kata menjadi representasi vektor numerik . Proses ini bertujuan untuk menangkap makna semantik dari setiap token yang ada di dalam teks input.

b. *Positional Encoding*

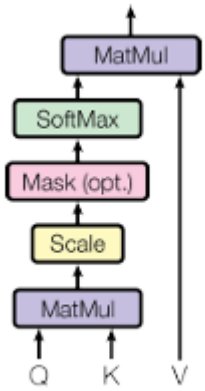
Dalam arsitektur *Transformer* termasuk model *T5*, karena tidak memiliki mekanisme rekurensi seperti pada *RNN* maka digunakan *Positional Encoding* yang ditambahkan ke dalam *input embedding* untuk memberikan informasi mengenai posisi setiap token dalam sebuah urutan, sehingga dapat memungkinkan model untuk memahami posisi dari masing-masing kata dalam suatu kalimat. *Transformer* menggunakan fungsi sinus dan cosinus pada *Positional Encoding* untuk menciptakan pola yang berbeda pada setiap posisi dalam urutan kata. Urutan kata yang telah diberi *Positional Encoding* digunakan sebagai input ke bagian *encoder Transformer*. Setiap blok dari arsitektur *Transformer* memiliki *Positional Encoding* setelah layer *embedding* mereka.

c. *Attention* Mechanism

Dalam *Transformer*, mekanisme *self-attention* memungkinkan model untuk menghubungkan setiap kata dalam input dengan kata-kata lainnya dalam satu urutan. Sebagai contoh, dalam sebuah kalimat, model dapat belajar untuk mengaitkan kata “health” dengan “care”. Sehingga

mekanisme ini dapat memungkinkan *encoder* untuk memusatkan perhatian pada bagian-bagian tertentu dari urutan input saat memproses setiap token. Token yang sudah diubah kemudian direpresentasikan dalam setiap vektor input berupa *key* (K), *value* (V), dan *query* (Q). Vektor Q adalah vektor yang mewakili token tertentu dari urutan input, sedangkan vektor K merupakan vektor yang sesuai dengan setiap token dalam urutan input. Setiap *value* dikaitkan dengan *key* dan digunakan untuk menyusun output dari lapisan *attention*. Skor *attention* bernilai tinggi apabila, *query* dan *key* memiliki kesesuaian.

d. *Scaled Dot-Product Attention*



Gambar 2. 2 *Scaled Dot-Product Attention* (Vaswani et al., 2017)

Pada Gambar 2. *2 scaled dot-product attention* mengubah input berupa *query, key,* dan *value* untuk menghasilkan representasi yang lebih banyak. Perhitungan antara *query* dan *key* menghasilkan matriks yang merepresentasikan setiap *query* dan *key*-nya, kemudian hasil di skala dan dibagi dengan dimensi vektor *key* untuk mengatasi gradien yang terjadi dalam fungsi softmax. Fungsi softmax diterapkan pada setiap baris matriks

untuk mendapatkan distribusi probabilitas, yang memastikan bahwa semua elemen dalam setiap baris dijumlah menjadi satu lalu mengubah *value* menjadi bobot probabilitas. Rumus yang digunakan adalah sebagai berikut

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_K}})V \quad (2.1)$$

Keterangan :

Q adalah *query*, K adalah *key*, dan V adalah *value*.
d_k adalah ukuran dimensi dari vektor *key*
V adalah ukuran dimensi dari vektor *value*

e. *Multi-Headed Attention*

Dapat diartikan sebagai fungsi *attention* untuk menghasilkan nilai vektor secara *parallel*. Nilai vektor yang dihasilkan kemudian dijumlah untuk menghasilkan nilai akhir.

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^O \quad (2.2)$$

$$where\ head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$$

Keterangan :

$MultiHead(Q, K, V)$ adalah gabugan 3 input (*query*, *key*, *value*) menjadi 1 representasi vektor.
$Concat(head_1, \dots, head_h)$ adalah setiap *head attention* yang dihasilkan oleh fungsi *attention* digabung menjadi satu vektor panjang.
W^O adalah matriks proyeksi output.
$head_i$ adalah perhitungan *head* yang sudah diproyeksikan dengan matriks output.
$Attention(QW_i^Q, KW_i^K, VW_i^V)$ adalah penerapan operasi *attention* pada *query*, *key*, *value* yang sudah diproyeksikan linear untuk menghasilkan vektor *attention*.

f. *Feed-forward Network*

Fungsi *FFN* diterapkan pada masing masing blok *encoder* dan *decoder* dengan *ReLU* didalamnya. Jaringan ini berperan dalam melakukan penyempurnaan tambahan terhadap representasi yang ada.

$$FFN(x) = \max(0, xW_1 + b_1)\, W_2 + b_2 \quad (2.3)$$

x dikalikan dengan matriks bobot W_1 dan ditambah dengan bias b_1. Hasil dari operasi tersebut kemudian dimasukkan ke dalam fungsi *ReLU* untuk menambahkan non-linearitas. Setelah itu, output dari *ReLU* tersebut dikali dengan matriks bobot W_2 dan ditambah dengan bias b_2 untuk menghasilkan output akhir.

g. *Normalization dan Residual Connection*

Pada tahap normalisasi, setiap sub-layer dalam lapisan *encoder* diikuti oleh langkah normalisasi yang berfungsi untuk mengatur distribusi nilai-nilai input agar stabil selama proses pelatihan model. Pada tahap *residual connection*, setiap output dari sub-layer ditambahkan kembali ke inputnya sendiri sebelum dilakukan operasi selanjutnya untuk membantu dalam mengatasi *vanishing gradient*. Proses *residual connection* ini juga terjadi setelah sub-layer *FFN*. Setelah *FFN* mengolah input, outputnya juga ditambah kembali ke input sebelum normalisasi. Tujuannya untuk memastikan bahwa setiap tahap dalam lapisan *encoder* tidak hanya meningkatkan representasi input, tetapi juga mempertahankan konsistensi dan stabilitas informasi dalam proses transformasi data.

h. *Linear dan Softmax*

Tahapan akhir dari arsitektur *Transformer* adalah saat melewati lapisan *linear* akhir, yang berfungsi sebagai klasifikator. Ukuran dari klasifikator ini sesuai dengan jumlah kelas yang terlibat, yaitu jumlah seluruh kata unik (*vocabulary size*) yang dikenal oleh model. Misalkan jika

terdapat 1000 kelas yang merepresentasikan 1000 kata berbeda, maka output dari lapisan klasifikator akan berupa array berdimensi 1000 elemen. Output ini kemudian akan diproses lebih lanjut menggunakan fungsi aktivasi *softmax*, yang akan mengubah nilai-nilai tersebut menjadi skor probabilitas dalam rentang antara 0 hingga 1. Dari hasil ini, skor probabilitas tertinggi akan dipilih sebagai hasil akhir, dan indeks dari nilai tertinggi tersebut menunjuk lansung ke kata yang diprediksi sebagai token selanjutnya dalam urutan kalimat. Dalam konteks penelitian yang sedang dilakukan, Langkah ini merupakan tahapan dimana judul dihasilkan satu kata demi kata. Model akan memilih kata dengan probabilitas tertinggi pada setiap langkah, hingga membentuk satu judul yang lengkap.

2.4 Text-to-Text Transfer Transformer (*T5*)

Model *T5* (*Text-to-Text Transfer Transformer*) adalah arsitektur berbasis *Transformer* yang dikembangkan oleh *Google Research* sebagai solusi untuk berbagai tugas *NLP* dalam format *text-to-text* (Kale & Rastogi, 2020). Pendekatan ini memungkinkan semua tugas *NLP* seperti penerjemahan, meringkas teks, analisis sentimen, dan tanya jawab dapat direpresentasikan dalam bentuk teks baik itu input maupun outputnya. Salah satu aspek penting dalam model *T5* adalah penggunaan *prefix* (awalan teks) yang berfungsi sebagai petunjuk model untuk memahami konteks tugas yang diberikan. Dengan adanya *prefix*, model dapat menangani berbagai tugas hanya dengan menyesuaikan instruksi dalam input.

Berbeda dengan model lain seperti *BERT*, *GPT*, dan *BART*, model *T5* memiliki beberapa keunggulan (Zaki, 2024). Yang pertama, *T5* tidak hanya berfokus pada

pemahaman teks seperti *BERT* atau generasi teks seperti *GPT*, model *T5* dapat menangani kedua aspek tersebut dengan lebih fleksibel. Kedua, *T5* dilatih menggunakan dataset *C4* (*Colossal Clean Crawled Corpus*) dalam skala yang besar, sehingga menjadikan model tersebut lebih kuat dalam transfer learning, di mana model yang telah dilatih dapat dengan mudah diadaptasi untuk berbagai tugas *NLP* tanpa banyak penyesuaian ulang. Keunggulan lain dari *T5* adalah skalabilitasnya, dengan berbagai ukuran model mulai dari *T5-Small* dengan 60 juta parameter hingga *T5-XXL* dengan 11 miliar parameter.

2.4.1 Arsitektur Model *T5*

Secara arsitektural, *T5* dibangun menggunakan model dari arsitektur *encoder-decoder* Transformer asli yang awalnya diperkenalkan untuk tugas penerjemahan mesin. Struktur ini terdiri dari dua komponen utama:

1. ***Encoder***

   Encoder arsitektur model *T5* terdiri dari beberapa lapisan, di mana setiap lapisan memiliki dua sub-lapisan utama yaitu mekanisme *multi-head self-attention* dan jaringan *feed-forward* yang terhubung ke setiap lapisan berikutnya (Raileanu et al., 2020). Dalam mekanisme *self-attention*, setiap token dalam urutan input dapat memperhatikan (*attend to*) semua token lainnya dalam urutan yang sama, baik yang mendahuluinya maupun yang mengikutinya. Kemampuan untuk melihat konteks "ke depan" dan "ke belakang" ini memungkinkan *encoder* untuk membangun pemahaman yang mendalam dari seluruh teks input. Misalnya, untuk memahami makna kata "bank" dalam kalimat "The man went to the bank to get some cash," model perlu

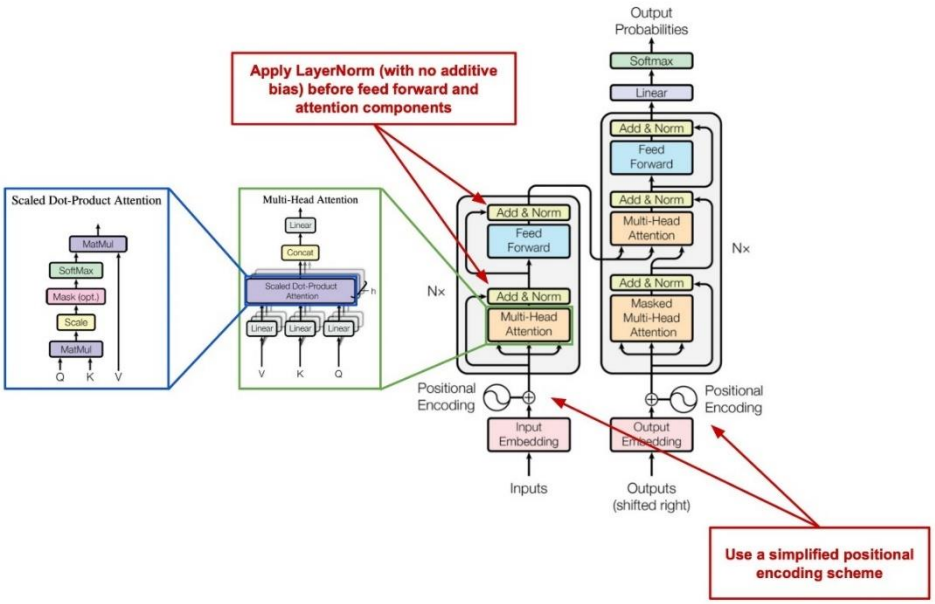
melihat kata-kata di sekitarnya seperti "cash" untuk membedakannya dari "river bank". *Encoder T5* menangkap hubungan dan dependensi jangka panjang *self-attention*, menghasilkan representasi input yang sangat informatif untuk digunakan oleh *decoder* (Shaw et al., 2018).

2. ***Decoder***

Decoder T5 memiliki struktur yang lebih kompleks karena harus melakukan dua tugas sekaligus: menghasilkan teks yang koheren dan memastikan teks tersebut relevan dengan input. Untuk mencapai ini, setiap lapisan *decoder* memiliki tiga sub-lapisan, yang mencakup dua jenis mekanisme *attention* yang berbeda :

1) ***Masked Multi-Head Self-Attention***: mekanisme yang mirip dengan yang digunakan oleh model GPT. Saat menghasilkan token pada posisi ke-i, *decoder* hanya diizinkan untuk memperhatikan (*attention*) token-token yang telah dihasilkan sebelumnya (dari posisi 1 hingga i−1). Perhatian (*attention*) pada token-token di masa depan (posisi i+1 dan seterusnya) "ditutupi" (*masked*). Hal ini memastikan bahwa proses generasi bersifat autoregresif sehingga kata berikutnya dihasilkan berdasarkan kata-kata sebelumnya yang penting untuk menghasilkan teks yang mengalir secara alami dan logis (Yin et al., 2024).
2) ***Multi-Head Cross-Attention***: merupakan mekanisme penghubung antara *decoder* dan *encoder*. Dalam sub-lapisan ini, *decoder* memperhatikan (*attention*) output dari *encoder* yaitu representasi dari *seluruh urutan input* (Libovický et al., 2018). Pada setiap langkah generasi, *cross-attention*

memungkinkan *decoder* untuk memfokuskan perhatiannya (*attention*) pada bagian-bagian yang paling relevan dari teks input untuk menginformasikan token apa yang harus dihasilkan selanjutnya. Misalnya, dalam tugas peringkasan, saat *decoder* menulis kalimat ringkasan, selanjutnya menggunakan *cross-attention* untuk "melihat kembali" ke kalimat-kalimat kunci dalam artikel asli yang disediakan oleh *encoder*.



Gambar 2. 3 Perbedaan Arsitektur Model *T5* (Pal et al., n.d.)

Pada Gambar 2. 3 arsitektur model *T5* menyederhanakan beberapa aspek dari arsitektur asli *Transformer* seperti :

a. *Positional Encoding*

Dalam *T5* skema yang digunakan lebih sederhana dari model *Transformer* asli, dimana *T5* menambahkan vektor posisi, scalar ke logit terkait untuk menghitung bobot *attention*. Pendekatan ini menyederhanakan perhitungan dan mengurangi kompleksitas.

b. *Layer Normalization*

T5 menerapkan normalisasi lapisan sebelum setiap sub-layer. Jadi, input ke setiap sub-layer *attention* dan *feed-forward* sudah dinormalisasi telebih dahulu. Perubahan urutan ini. Dapat mempengaruhi stabilitas pelatihan dan kinerja model.

c. *Additive Bias*

T5 tidak menggunakan bias tambahan dalam normalisasi layer, yang berarti setalah normalisasi tidak ada penyesuaian tambahan yang dilakukan. Pendekatan ini dapat mengurangi jumlah parameter yang harus dipelajari model, yang bisa membuat model lebih sederhana dan lebih cepat dalam pelatihan.

d. Mekanisme *Dropout*

T5 memperluas penggunaan dropout ke seluruh jaringan, termasuk pada bobot *attention*, lapisan *feed-forward*, dan lainnya. dengan menerapkan *dropout* secara lebih luas, *T5* berusaha untuk meningkatkan generalisasi model dan mencegah *overfitting* pada berbagai komponen layer.

2.4.2 Perbedaan Parameter Model *T5*

Perbedaan parameter menjadi kunci yang membedakan kemampuan, ukuran, dan kebutuhan komputasi dari setiap varian dari model *T5*. Secara mendasar, semakin besar jumlah parameter, semakin kompleks model tersebut dan berpotensi memiliki performa yang lebih baik dalam memahami dan menghasilkan teks, namun dengan konsekuensi kebutuhan sumber daya yang lebih besar.

Parameter dalam model *T5*, seperti pada model *deep learning* lainnya, adalah nilai-nilai yang dipelajari oleh model selama proses pelatihan. Nilai-nilai inilah yang menentukan bagaimana model akan memproses input dan menghasilkan output. Pada arsitektur *T5* yang berbasis *encoder-decoder*, parameter-parameter ini tersebar di berbagai komponennya. Perbedaan jumlah parameter antar varian *T5* terutama berasal dari penyesuaian pada beberapa komponen arsitektur utamanya:

a. `d_model` (Ukuran Model/Dimensi Hidden) adalah dimensi dari *embedding* dan *hidden state* di seluruh lapisan model. Nilai `d_model` yang lebih besar berarti representasi vektor dari teks menjadi lebih kaya dan detail, sehingga memungkinkan model untuk menangkap nuansa yang lebih halus.
b. `num_layers` (Jumlah Lapisan) Model *T5* memiliki lapisan *encoder* dan *decoder* yang identik. `num_layers` merujuk pada jumlah lapisan pada masing-masing komponen tersebut. Menambah jumlah lapisan akan meningkatkan kedalaman model, memungkinkannya untuk mempelajari pola yang lebih kompleks dalam data.
c. `num_heads` (Jumlah Kepala Atensi)**:** Dalam mekanisme *multi-head attention*, `num_heads` menentukan berapa banyak "kepala" atensi yang bekerja secara paralel. Setiap kepala fokus pada aspek yang berbeda dari hubungan antar kata dalam sebuah kalimat. Jumlah kepala yang lebih banyak memungkinkan model untuk menangkap berbagai jenis dependensi kontekstual.

d. `d_ff` (Dimensi Lapisan *Feed-Forward*) Setiap lapisan *encoder* dan *decoder* memiliki jaringan *feed-forward*. `d_ff` adalah ukuran dari lapisan tersembunyi di dalam jaringan ini. Biasanya, `d_ff` memiliki nilai yang lebih besar dari `d_model`.

e. `d_kv` (Dimensi *Key/Value*)**:** Dalam mekanisme atensi, ini adalah dimensi dari *key* dan *value* untuk setiap kepala atensi. Biasanya dihitung sebagai `d_model` / `num_heads`.

Berikut adalah perbandingan beberapa varian populer dari model *T5*, yang menggambarkan bagaimana perubahan pada komponen di atas berdampak pada jumlah total parameter:

Tabel 2. 2 Perbandingan Varian Model *T5*

Varian	Jumlah Total Parameter	d_model	num_layers	num_heads	d_ff
T5-Small	~60 Juta	512	6	8	2048
T5-Base	~220 Juta	768	12	12	3072
T5-Large	~770 Juta	1024	24	16	4096
T5-3B	~3 Miliar	1024	24	32	16384
T5-11B	~11 Miliar	1024	24	64	65536

2.4.3 *Preprocessing* dan Tokenisasi

T5 memiliki *pipeline preprocessing* yang mencakup tokenisasi dan format input prefiks tugas. Beberapa tahapan untuk preprocessing *T5* sebelum masuk ke *input embedding* adalah sebagai berikut:

a. Melakukan tokenisasi dengan memecah *string* dari teks mentah menjadi urutan unit yang lebih kecil yang disebut token. Token ini bisa berupa kata atau sub-kata.

b. Setiap token kemudian dipetakan ke ID *integer* sesuai dengan kosakata (*vocabulary*) yang telah ditentukan sebelumnya.
c. Karena model *Transformer* bekerja dengan input berukuran tetap, urutan ID token yang lebih panjang dari panjang maksimum yang telah ditentukan akan dipotong (*truncated*), dan yang lebih pendek akan ditambahkan dengan token *padding* khusus hingga mencapai panjang yang seragam.
d. Token khusus, seperti halnya token akhir urutan (*end-of-sequence*, EOS), akan ditambahkan ke urutan ID. Untuk model *T5*, token `</s>` umumnya digunakan sebagai penanda EOS.

Untuk langkah tokenisasi, *T5* secara khusus menggunakan *tokenizer SentencePiece*. *Tokenizer SentencePiece* adalah sebuah *library* tokenisasi yang diimplementasikan menggunakan C++ untuk mengolah data dengan lebih cepat (Dang et al., 2024). Salah satu keunggulan utama dari *tokenizer SentencePiece* adalah kemampuannya untuk melatih model tokenisasi langsung dari teks mentah, tanpa memerlukan pra-tokenisasi yang bergantung pada bahasa (seperti memisahkan kata berdasarkan spasi).

Secara spesifik, *T5* menggunakan implementasi model *unigram* dari *SentencePiece*, yang berbeda dari pendekatan *Byte-Pair Encoding* (*BPE*) yang digunakan oleh model seperti *GPT* dan *BERT* (Dang et al., 2024). Kosakata standar dari *T5* memiliki ukuran tetap yaitu 32.000 token, yang telah dilatih pada campuran data dari berbagai bahasa untuk mendukung kemampuan *multilingual* dasar.

Berbeda dengan banyak *tokenizer* lain yang menggunakan spasi sebagai pemisah kata, *SentencePiece* memperlakukan spasi sebagai simbol biasa.

Umumnya, spasi diganti dengan meta-simbol khusus seperti (U+2581) sebelum tokenisasi. Ini akan memastikan bahwa proses tokenisasi sepenuhnya dapat dibalik (*lossless* dan *reversible*). Urutan token yang dihasilkan mengandung semua informasi yang diperlukan untuk merekonstruksi teks asli dengan sempurna, termasuk spasi.

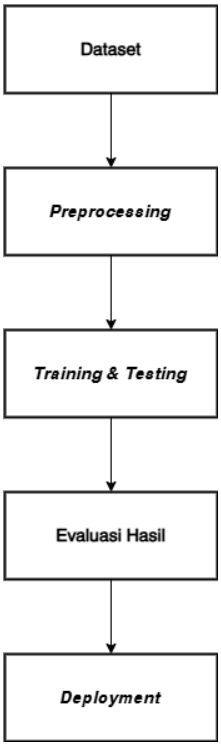
Adapun elemen paling inovatif dalam *pipeline* input *T5* adalah penggunaan prefiks tugas (*task-specific prefixes*). Merupakan *string* teks pendek yang ditambahkan di awal setiap urutan input untuk memberi tahu model tugas apa yang harus dilakukan (Atitienei & Marginean, 2024). Prefiks berfungsi sebagai instruksi misalnya, prefiks "`summarize`:" mengindikasikan bahwa model harus menghasilkan ringkasan dari teks yang mengikutinya, bukan menerjemahkannya atau menjawab pertanyaan tentangnya. Mekanisme sederhana ini memungkinkan satu model tunggal untuk melakukan multitasking secara efektif tanpa memerlukan perubahan arsitektur apa pun. Model belajar untuk mengasosiasikan setiap prefiks dengan perilaku output yang sesuai selama fase *pre-training* campuran atau *fine-tuning*.

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian merupakan serangkaian alur penelitian yang digunakan oleh peneliti untuk mengarahkan proses penelitian, dimulai dari proses pengumpulan data hingga fase *deployment* sistem. Desain penelitian dapat dilihat pada Gambar 3. 1 yang memberikan gambaran singkat mengenai alur penelitian.



Gambar 3. 1 Desain Penelitian

3.2 Dataset

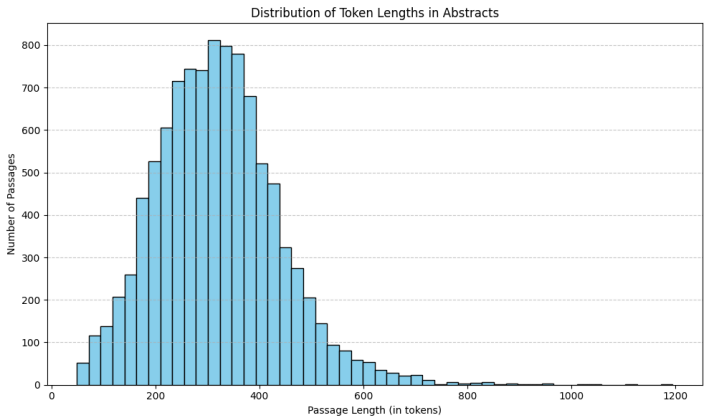
Penelitian ini menggunakan dataset yang berasal dari platform online bernama Kaggle. Dataset yang digunakan berasal dari jurnal publikasi PubMed

yang merupakan pusat publikasi artikel ilmiah di bidang kesehatan. Jumlah dataset adalah 10.000 data, untuk setiap artikel jurnal yang berisi judul artikel, abstrak, dan url dari setiap artikel jurnal. Penelitian ini menggunakan 2 kolom dari dataset yaitu kolom *title* yang digunakan sebagai target output yang akan diprediksi oleh model, dan kolom *abstract* sebagai input dalam model *T5*. Dataset dapat diakses pada tautan https://www.kaggle.com/datasets/anandhuh/covid-abstracts

title Title of research paper	abstract Abstract of the research paper	url URL of the research paper
10000 unique values	10000 unique values	10000 unique values
Real-World Experience with COVID-19 Including Direct COVID-19 Antigen Testing and Monoclonal-Antibo...	This article summarizes the experiences of COVID-19 patients diagnosed and treated at Faulkton Area ...	https://pubmed.ncbi.nlm.nih.gov/35008137
Successful outcome of pre-engraftment COVID-19 in an HCT patient impact of targeted therapies and c...	Coronavirus disease 2019 COVID-19 caused by severe acute respiratory syndrome coronavirus 2 SARS...	https://pubmed.ncbi.nlm.nih.gov/35008104

Gambar 3. 2 Dataset Artikel Ilmiah

Distribusi panjang token untuk setiap abstrak dapat dilihat pada Gambar 3. 3. Dari 10.000 data dapat dilihat bahwa rata-rata token pada data berjumlah 200 sampai 400 token.



Gambar 3. 3 Distribusi Token Abstract

3.3 *Preprocessing*

Setelah didapatkan dataset proses implementasi sistem dilanjutkan dengan preprocessing data dimana merupakan tahapan yang dilakukan untuk mempersiapkan data yang akan dilatih oleh model. Tahapan yang dilakukan pada preprocessing kali ini adalah *lowercasing* yaitu merubah sebuah kata yang memilki awalan huruf besar menjadi huruf kecil semua. Pendekatan ini dipilih karena model *T5* telah di*training* menggunakan tokenizer dari model itu sendiri yang mampu menangani variasi bentuk kata, karakter khusus, serta struktur kalimat tanpa perlu dilakukan *stemming*, *stopword removal*, atau pembersihan simbol secara manual (HaCohen-Kerner et al., 2020).

Melakukan *stopword removal* dan *stemming* justru dapat menghilangkan informasi penting yang dibutuhkan dalam tugas generasi judul. *Lowercasing* dianggap sebagai tahap preprocessing yang paling sesuai dan aman untuk memastikan konsistensi input tanpa mengurangi kualitas informasi yang diberikan kepada model.

Tabel 3. 1 Perbandingan data setelah preprocessing

Sebelum *lowercasing*	Setelah *lowercasing*
The indispensable role of plastic products in our daily life is highlighted by the COVID-19 pandemic again Disposable face masks made of polymer materials as effective and cheap personal protective equipment PPE have been extensively used by the public to slow down the viral transmission The repercussions of this have generated million tons of plastic waste being littered into the environment because of the improper disposal and mismanagement amid And plastic waste can release microplastics MPs with the help of physical chemical and biological processes which is placing a huge MPs contamination burden on the ecosystem In this work the knowledge regarding to the combined effects of MPs and pollutants from the release of face masks and	the indispensable role of plastic products in our daily life is highlighted by the covid-19 pandemic again disposable face masks made of polymer materials as effective and cheap personal protective equipment ppe have been extensively used by the public to slow down the viral transmission the repercussions of this have generated million tons of plastic waste being littered into the environment because of the improper disposal and mismanagement amid and plastic waste can release microplastics mps with the help of physical chemical and biological processes which is placing a huge mps contamination burden on the ecosystem in this work the knowledge regarding to the combined effects of mps and pollutants from the release of face masks and

the impacts of wasted face masks and MPs on the environment terrestrial and aquatic ecosystem was systematically discussed In view of these some green technologies were put forward to reduce the amounts of discarded face masks in the environment therefore minimizing MPs pollution at its source Moreover some recommendations for future research directions were proposed based on the remaining knowledge gaps In a word MPs pollution linked to face masks should be a focus worldwide	the impacts of wasted face masks and mps on the environment terrestrial and aquatic ecosystem was systematically discussed in view of these some green technologies were put forward to reduce the amounts of discarded face masks in the environment therefore minimizing mps pollution at its source moreover some recommendations for future research directions were proposed based on the remaining knowledge gaps in a word mps pollution linked to face masks should be a focus worldwide

3.4 *Training* dan *Testing* Data

Dataset yang telah *preprocessing* kemudian disimpan yang nantinya akan digunakan pada proses *training* & *testing* data. Pada tahap ini data yang telah dipreprocessing kemudian dibagi menjadi 2 yaitu sebagai data *training* dan *testing*. Model nantinya akan dilatih menggunakan data *training* untuk mengenali pola dan hubungan antara abstrak dan judul. Kemudian setelah model di*training* selanjutnya dilakukan *testing* menggunakan data *testing* untuk melihat seberapa baik model dapat menghasilkan judul dari abstrak yang belum dikenali sebelumnya.

3.4.1 Model *T5*

Penelitian yang dilakukan menggunakan *T5* varian *small* dengan 60 juta parameter. Penggunaan *T5-small* didasari karena kebutuhan sumber daya komputasi yang secara signifikan lebih sedikit daripada *T5-Base* yang memiliki 220 juta parameter atau *T5-Large* 770 juta parameter. Selain itu dikarenakan penggunaan parameter yang lebih sedikit, proses pelatihan dapat dilakukan lebih cepat dibandingkan dengan model *T5* yang lain

Arsitektur model *T5* terdiri dari 2 komponen utama yaitu *encoder* dan *decoder*. Pada tahap encoder model menerima input berupa teks abstrak yang

selanjutnya diproses melalui *embedding layer* dan diubah ke dalam representasi vektor, kemudian ditambahkan dengan *positional encoding* untuk memberikan informasi urutan kata. Hasil ini diproses oleh encoder melalui beberapa lapisan yang masing-masing terdiri dari multi-head self-attention, feed-forward network, serta residual connection dan normalization.

Setelah melalui tahap encoder, proses generasi output pada decoder dimulai dengan token awal, kemudian model secara bertahap menghasilkan satu token (kata) pada satu waktu. Token yang baru dihasilkan akan digunakan sebagai masukan untuk memprediksi token berikutnya. Proses ini akan terus berulang hingga membentuk satu kalimat judul yang utuh. Pendekatan ini disebut sebagai autoregressive decoding, di mana setiap prediksi token bergantung pada token token sebelumnya.

Decoder menggunakan 2 jenis attention yaitu masked self-attention dimana layer ini mencegah model untuk "melihat" token selanjutnya saat menghasilkan kata demi kata. Selanjutnya ada cross-attention layer yang berfungsi untuk menghubungkan informasi dari input (abstrak) dengan output yang dihasilkan. Berdasarkan rumus pada Persamaan 2.1, langkah-langkah untuk menghitung fungsi softmax dapat dilihat pada pseudocode dibawah ini:

Pseudocode 3. 1 Langkah menghitung fungsi *Softmax*

```
Tokenize the text into words:
  tokens = ["physical", "activity", "as", "a", "human", "right"]
Convert each token into an embedding vector:
   embeddings = embed(tokens) # shape: (n_tokens, embedding_dim)
Initialize weight matrices for Q, K, V:
   Wq, Wk, Wv = initialize_matrices(embedding_dim, d_k)
Compute Query (Q), Key (K), and Value (V) matrices:
   Q = embeddings × Wq  # shape: (n_tokens, d_k)
   K = embeddings × Wk  # shape: (n_tokens, d_k)
   V = embeddings × Wv  # shape: (n_tokens, d_v)
Compute the attention scores (dot product between Q and K^T):
```

```
   scores = Q × transpose(K)  # shape: (n_tokens, n_tokens)
Scale the scores by sqrt(d_k):
   scaled_scores = scores / sqrt(d_k)
Apply the softmax to each row (to get attention weights):
   attention_weights = softmax(scaled_scores, axis=1)
Multiply attention weights with V to get the output:
   attention_output = attention_weights × V  # shape: (n_tokens,
d_v)
Return attention_output
```

Selanjutnya langkah untuk menghitung *Feed Forward Network* menggunakan Persamaan 2.3 dapat dilihat pada pseudocode dibawah ini

Pseudocode 3. 2 Langkah menghitung *Feed Forward Network*

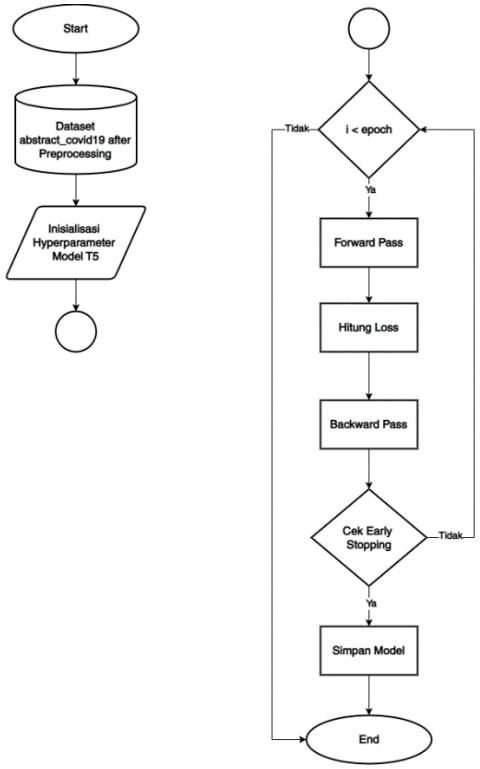
```
Define layer dimensions:
   d_model = input/output dimension (e.g., 512)
   d_ff = hidden layer size (e.g., 2048)
Initialize parameters:
   W1 = weight matrix of shape (d_model, d_ff)
   b1 = bias vector of shape (d_ff,)
   W2 = weight matrix of shape (d_ff, d_model)
   b2 = bias vector of shape (d_model,)
Compute hidden layer with ReLU activation:
   hidden = max(0, x × W1 + b1)  # shape: (seq_len, d_ff)
Compute output layer:
   ffn_output = hidden × W2 + b2  # shape: (seq_len, d_model)
Return ffn_output
```

3.4.2 *Training* Model *T5*

Proses *training* dilakukan menggunakan model *T5* yang didapat dari website *simpletransformers*, sebuah platform online yang menyediakan model *transformers* ringan yang sudah di pre-trained dan siap digunakan. Penelitian ini menggunakan varian *T5-small*, salah satu versi paling ringan dari arsitektur *T5* yang telah di *pre-trained*. Alasan menggunakan *T5-small* karena model yang digunakan lebih mudah di-*fine-tune* dari pada varian di atasnya. Meskipun model yang lebih besar seperti *T5*-Base umumnya memiliki performa lebih unggul, varian *T5-Small* menunjukkan hasil yang kompetitif dengan waktu pelatihan yang lebih cepat, menjadikannya pilihan yang tepat untuk implementasi di lingkungan dengan

sumber daya terbatas (Setiyarini et al., 2024). Model yang digunakan dapat diakses pada laman https://simpletransformers.ai/docs/*T5*-model/ untuk dokumentasi lebih lanjutnya

Laman tersebut menyediakan model *T5* dalam bentuk siap pakai melalui *library simpletransformer*, yang memungkinkan integrasi langsung tanpa perlu melakukan *training* dari awal. Proses *training* model *T5* pada penelitian ini dapat dilihat pada flowchart Gambar 3. 4



Gambar 3. 4 Flowchart *Training* Model *T5*

Proses *training* di mulai dengan pemuatan dataset yang telah dipreprocessing, lalu dilakukan inisialisasi parameter untuk model *T5* seperti jumlah *epoch*, ukuran batch, panjang maksimum token, dan lain-lain. Dilanjut mulai proses iterasi berdasarkan jumlah *epoch* yang telah ditentukan. Di bagian

Forward Pass abstrak dataset yang telah ditokenisasi diinput pada bagian *encoder* model. Bagian encoder selanjutnya memproses input menjadi representasi vektor yang kemudian digunakan oleh bagian *decoder* model untuk menghasilkan judul artikel secara bertahap dalam batch per batch.

Di bagian menghitung *loss*, output prediksi dari *decoder* lalu dibandingkan dengan judul asli. *Loss* digunakan untuk mengukur performa model dalam menghasilkan output yang mendekati judul asli dari artikel yang tersedia di dataset. Nilai *loss* yang telah diukur digunakan untuk menghitung gradient dari parameter model. Setelah satu *epoch*, performa model akan dievaluasi. Jika tidak ada peningkatan performa, maka *early stopping* akan menghentikan proses pelatihan secara otomatis untuk mencegah *overfitting*. Apabila performa yang ditunjukkan model bagus, maka model akan lanjut untuk disimpan dan siap digunakan pada implementasi sistem.

3.4.3 *Testing* Model *T5*

Testing adalah proses dimana model yang telah di*training* digunakan untuk memprediksi data yang belum pernah dilihat sebelumnya (Naseeba et al., 2024). Data *testing* yang digunakan pada proses *testing* berjumlah 20% dari total dataset, yang dipisahkan secara acak menggunakan *random samples* dari *library pandas*. Dalam penelitian ini, proses *testing* dilakukan menggunakan *function eval* yang sudah ada dalam *library simpletransformers* . Proses ini bertujuan untuk melihat seberapa baik model dalam menghasilkan judul baru dari abstrak yang tidak termasuk dalam data *training*.

3.5 Evaluasi Matrik *ROUGE*

Setelah model dilatih untuk menghasilkan judul baru dari abstrak yang telah diberikan, peneliti mengevaluasi seberapa baik model yang digunakan dalam menghasilkan judul baru. Evalusi dilakukan dengan menggunakan metrik *ROUGE*. Metrik ini bekerja dengan mengukur kecocokan jumlah *n-gram* antara teks yang menjadi kandidat dan refrensi. *N-gram* dapat diibaratkan sebagai sekelompok token yang dimulai dari unigram (1-gram) atau satu buah kata. Dilanjut dengan bigram (2-gram) yang terdiri dari dua buah kata hingga seterusnya. Jenis *ROUGE* yang digunakan dalam evaluasi model dapat dilihat pada Tabel 3. *2*.

Tabel 3. 2 Jenis Matrik *ROUGE*

Metrik	Deskripsi
ROUGE-1	Mengukur kesamaan berdasarkan unigram antara hasil dari model dengan judul asli.
ROUGE-2	Mengukur kesamaan berdasarkan bigram.
ROUGE-L	Mengukur kesamaan struktur kalimat antara judul dari hasil model dengan judul asli

Rumus yang digunakan untuk menghitung *ROUGE*-N ada didalam persamaan 3.1 sampai dengan 3.3

$$\text{Recall} = \frac{Jumlah\ N{-}gram\ yang\ sesuai\ dalam\ refrensi\ dan\ prediksi}{Jumlah\ total\ N{-}gram\ dalam\ refrensi} \quad (3.1)$$

$$\text{Precission} = \frac{Jumlah\ N{-}gram\ yang\ sesuai\ dalam\ refrensi\ dan\ prediksi}{Jumlah\ total\ N{-}gram\ dalam\ prediksi} \quad (3.2)$$

$$\text{F1} = \frac{2{\times}Precission{\times}Recall}{(Precission{+}Recall)} \quad (3.3)$$

Keterangan :

$N - gram$ adalah urutan kata dengan panjang N.

Selanjutnya untuk menghitung *ROUGE*-L digunakan persamaan 3.4 sampai dengan 3.6. Matriks ini nantinya akan digunakan untuk mengukur kesamaan

struktur kalimat dari judul yang dihasilkan model dengan judul asli. Semakin tinggi nilai yang dihasilkan maka semakin baik model dalam menghasilkan urutan kata yang sesuai.

$$\mathrm{R}ecall_{lcs} = \frac{LCS\,(X,Y)}{m} \tag{3.4}$$

$$\mathrm{Precission}_{lcs} = \frac{LCS\,(X,Y)}{n} \tag{3.5}$$

$$\mathrm{F1}_{lcs} = \frac{(1+\beta^2)\mathrm{R}_{lcs}\mathrm{P}_{lcs}}{\mathrm{R}_{lcs}+\beta^2\mathrm{P}_{lcs}} \tag{3.6}$$

Keterangan :

$LCS(X,Y)$ adalah panjang dari *Longest Common Subsquence* antara teks prediksi (X) dan refrensi (Y).
m dan n adalah ukuran kalimat refrensi dan kandidat.
β adalah bias terhadap recall atau precission yang bernilai 1, artinya recall dan precission dianggap sama pentingnya.

Model dianggap "baik" apabila skor *ROUGE* yang dihasilkan model cenderung tinggi, yang menunjukkan bahwa model berhasil menghasilkan teks yang lebih dekat dengan referensi yang diinginkan. Dalam penelitian yang dilakukan oleh (Hanif, 2023), skor *ROUGE-1* dan *ROUGE-2* sering digunakan untuk menilai sejauh mana model dapat menangkap elemen penting dalam teks. Skor *ROUGE-1* mengukur kesamaan unigram antara hasil yang dihasilkan oleh model dan referensi, yang memberikan gambaran tentang sejauh mana kata-kata penting dari teks asli ada di dalam ringkasan yang dihasilkan. Sedangkan *ROUGE-2* lebih berfokus pada kesamaan *bigram*, yang menunjukkan kemampuan model dalam mempertahankan hubungan antar kata yang lebih kompleks dan konteks yang lebih dalam dari teks sumber. Semakin tinggi nilai *ROUGE-1* dan *ROUGE-2*,

semakin baik model dalam menghasilkan ringkasan yang relevan dan sesuai dengan informasi utama yang ada dalam teks asli.

3.6 Deployment

Fase *deployment* merupakan tahap implementasi model yang telah di *training* ke dalam bentuk aplikasi yang dapat digunakan oleh user secara langsung. Pada penelitian ini, tahap *deployment* hanya dilakukan untuk mengimplementasikan model dalam menghasilkan judul artikel ilmiah dari input berupa abstrak. Platform online *Hugging Face* dipilih sebagai integrasi untuk model karena mendukung *framework Gradio* serta desain antarmuka yang memudahkan penggunaan user.

3.7 Skenario Uji Coba

Pengujian *hyperparameter* dilakukan pada penelitian yang bertujuan untuk melakukan *tuning* dengan menemukan setiap kombinasi *hyperparameter* terbaik. *Hyperparameter* yang digunakan untuk proses pengujian adalah *batch size* dan *epoch. Batch size* adalah jumlah data yang diproses dalam satu iterasi. Misalkan *batch size* model adalah 16, maka model akan membaca 16 data sekaligus. *Epoch* adalah satu iterasi penuh dimana seluruh dataset *training* diproses sekali oleh model *machine learning.*

Pada penelitian yang dilakukan oleh (Rehman et al., 2025), proses tuning dilakukan pada pelatihan dataset CSPubSum, dimana *epoch* dan *batch size* disesuaikan untuk menghasilkan performa terbaik pada pelatihan model. Setiap parameter memiliki beberapa nilai yang menjadi konfigurasi. Hyperparameter yang digunakan pada penelitian ini ditampilkan pada Tabel 3. 1 berikut.

Tabel 3. 3 Hyperparameter Model *T5*

Hyperparameter	**Konfigurasi**
reprocess_input_data	True
overwrite_output_dir	True
max_seq_length	512
train_batch_size	4, 8
num_train_*epoch*s	4, 5, 6
num_beams	5
early_stopping	True
gradient_accumulation_steps	4
save_model_every_*epoch*	False

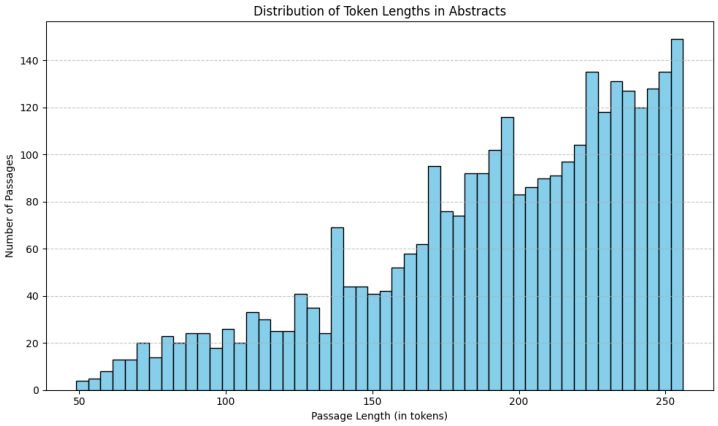
Tujuan dari skenario pengujian adalah untuk menemukan kombinasi parameter yang optimal, yang dapat menghasilkan model dengan performa terbaik. Selanjutnya hasil dari masing – masing konfigurasi parameter akan dibandingkan dan dianalisis untuk menetukan parameter terbaik yang dapat digunakan pada fase *deployment*. Pembagian dataset dilakukan dengan perbandingan 8:2 pada proses *training* dan *testing*. Dimana data yang di*training* adalah 80% dari dataset, sedangkan data yang di*testing* adalah 20%-nya

BAB IV

HASIL DAN PEMBAHASAN

4.1 Dataset

Karena keterbatasan model *T5-Small* dalam memproses data dengan panjang lebih dari 512 token (kata) tanpa terjadinya *truncation* otomatis (Koupaee & Wang, 2018), (Pal et al., n.d.) dalam penelitiannya yang berjudul "*Text Summarization using BERT and T5*" melakukan filtering terhadap data yang memiliki jumlah token melebihi batas tersebut, karena keterbatasan memori yang digunakan pada penelitiannya. Pada penelitian ini, panjang maksimal dari input abstrak dibatasi menjadi 512 token untuk *hyperparameter tuning* di bagian *max sequence length*. Data *training dan testing* yang digunakan adalah 50% dari dataset yaitu berjumlah 5000 data yang merupakan merupakan pasangan dari abstrak (input) dan judul (output) artikel ilmiah.



Gambar 4. 1 Distribusi Panjang Token

4.2 Proses *Training* & *Testing* Model

4.2.1 *Training* Model *T5*

Proses *training* model *T5* dalam penelitian bertujuan untuk melakukan *fine-tuning* pada model agar mampu menghasilkan judul artikel ilmiah yang optimal dengan skor *ROUGE* yang tinggi berdasarkan abstrak. Berikut adalah beberapa tahapan mengenai alur *training* model:

4.2.1.1 Persiapan Dataset untuk Pelatihan

Langkah pertama adalah persiapan dataset. Dataset yang digunakan adalah dataset yang telah di-*preprocessing* lalu disimpan dan di muat dengan nama `dataset_lowercased.csv`. Penelitian menggunakan 50% dari total data yang diambil secara acak (`frac=0.5, random_state=50`) untuk digunakan dalam penelitian. Melakukan inisialisasi *abstract* dan *title* menjadi `input_text` dan `target_text` untuk memenuhi format standar yang dibutuhkan oleh *library* `simpletransformers`.

Selanjutnya, dataset dibagi menjadi dua bagian dengan perbandingan 80:20. Sebesar 80% dari data digunakan sebagai set pelatihan (`train_df`), sementara 20% sisanya dipisahkan sebagai set evaluasi (`eval_df`) yang akan digunakan pada tahap pengujian. Model *T5* menggunakan `prefix` yaitu sebuah intruksi mengenai tugas yang akan dilakukan dengan nilai `"summarize"` pada kedua set data tersebut.

4.2.1.2 Konfigurasi Model

Penelitian menggunakan model *T5* varian small yang diinisialisasi melalui *library* `simpletransformers` dengan `model_type="T5"` dan `model_name="T5-small"`. Proses pelatihan dilakukan menggunakan *hyperparameter* yang telah

diidentifikasikan. Panjang token input maksimal dibatasi pada 512 token (`max_seq_length`), yang mengindikasikan setiap abstrak yang lebih panjang dari batas ini akan dipotong secara otomatis.

. Untuk menerapkan ukuran *batch* yang lebih besar tanpa membebani memori, digunakan teknik *gradient accumulation* sebanyak 4 langkah (`gradient_accumulation_steps`), sehingga menghasilkan ukuran *batch* yang efektif sebesar 32 (8 x 4). Hyperparameter lain seperti *early stopping* diterapkan untuk mencegah *overfitting* pada data yang di *training* serta penggunaan `num_beams` untuk meningkatkan kualitas teks yang dihasilkan pada saat evaluasi.

4.2.1.3 Proses *Training*

Setelah melakukan konfigurasi *hyperparameter*, proses pelatihan dimulai dengan memanggil *function* `model.train_model(train_df)` dari *library simpletransformers*. Selama proses *training*, model *T5*-small menyesuaikan bobot internalnya untuk mempelajari pola dan hubungan antara `input_text` (abstrak) dan `target_text` (judul) dari set pelatihan. Tujuannya adalah untuk meminimalkan *loss* atau kesalahan prediksi. Setelah proses pelatihan selama *epoch* yang ditentukan selesai, model yang telah di-*fine-tune* beserta *tokenizer*-nya disimpan ke dalam direktori `saved_model512` untuk digunakan pada tahap pengujian selanjutnya.

4.2.2 *Testing* Model *T5*

Tahap pengujian bertujuan untuk menguji kemampuan generalisasi model pada data yang belum pernah dilihat sebelumnya.

4.2.2.1 Proses *Testing* Model *T5*

Langkah awal yang dilakukan adalah memuat kembali model *T5* dan *tokenizer* yang telah di-*fine-tune* dan disimpan dari direktori `saved_model512`. *Testing* dilakukan dengan menggunakan fungsi `model.eval_model(eval_df)`. Fungsi ini akan meminta model untuk menghasilkan judul (`target_text`) untuk setiap abstrak (`input_text`) yang ada di dalam set data *testing* (`eval_df`).

Karena `eval_df` berisi data yang tidak digunakan selama pelatihan, proses ini mengukur seberapa baik model dalam mengaplikasikan pengetahuan yang telah dipelajarinya pada data baru. Hasil dari proses adalah nilai *loss* dari *testing*, yang memberikan gambaran mengenai performa model. Hasil tersebut kemudian menjadi dasar untuk analisis lebih lanjut menggunakan *ROUGE*.

4.3 Hasil Skenario Uji Coba

Berdasarkan Tabel 3. 3, dilakukan skenario uji coba untuk setiap hyperparameter pada proses training data untuk melihat seberapa baik model belajar. Skenario dari konfigurasi hyperparameter terbaik didapatkan apabila model menghasilkan training loss yang relatif rendah pada proses training yang mengindikasikan model belajar dengan baik pada saat proses training.

Tabel 4. 1 *Hyperparameter Tuning*

No	*train_batch_size*	*num_train_epochs*	*train_loss*
1	4	4	1.830
2	4	5	1.746
3	4	6	1.665
4	8	4	1.850
5	8	5	1.772
6	8	6	1.691

Dari Tabel 4. *1*, terlihat bahwa peningkatan jumlah *epoch* secara umum berkontribusi terhadap penurunan nilai loss pada *training*, yang menandakan bahwa model dapat belajar semakin baik terhadap data *training* seiring bertambahnya iterasi. Dapat dilihat pada konfigurasi dengan batch size 4, nilai loss *training* menurun secara signifikan dari 1.830 pada konfigurasi dengan *epoch* berjumlah 4 menjadi 1.746 pada konfigurasi dengan jumlah *epoch* 5 dan akhirnya turun lagi menjadi 1.665 dengan jumlah *epoch* 6.

Penurunan nilai *train loss* yang signifikan menunjukkan bahwa model mengalami peningkatan dalam proses *training* terhadap data seiring bertambahnya *epoch*/iterasi. *Batch size* yang lebih kecil juga membuat model dapat melakukan update terhadap bobot data lebih sering, sehingga model dapat menangkap pola data jauh lebih detail.

Pada skenario dengan jumlah *batch size* 4 model menghasilkan *training loss* yang lebih rendah dibandingkan dengan *batch size* 8, hal ini karena bobot model diperbarui lebih sering saat iterasi *training*, sehingga model dapat lebih responsif terhadap variasi data untuk menghindari *overfitting*. Penggunaan *batch size* yang lebih kecil selama pelatihan *neural network* sering kali menghasilkan generalisasi yang lebih baik. *Batch size* yang kecil memberikan perhitungan gradien yang lebih mutakhir, sehingga menghasilkan pelatihan yang lebih stabil dan andal (Masters & Luschi, 2018).

4.4 Hasil Evaluasi *ROUGE*

Dari hasil pengujian skenario, didapatkan parameter terbaik adalah dengan menggunakan kombinasi konfigurasi yang ditampilkan pada Tabel 4. *2*

Tabel 4. 2 Konfigurasi Skenario Terbaik

No	*train_batch_size*	*num_train_epochs*	*train_loss*
1	4	6	1.665

Model dengan konfigurasi tersebut kemudian dievaluasi menggunakan matriks *ROUGE* untuk mengukur seberapa tinggi akurasi yang dihasilkan. Tujuan melakukan evaluasi ini adalah untuk melihat seberapa mirip judul yang dihasilkan oleh model yang telah di*training* dengan judul asli dari dataset. Hasil evaluasi tersebut menghasilkan rata-rata sebagai berikut:

Tabel 4. 3 Hasil Evaluasi *ROUGE*

ROUGE	*Recall*	*Precision*	*F1-Score*
ROUGE-1	0.3770	0.4982	0.4135
ROUGE-2	0.1973	0.2684	0.2179
ROUGE-L	0.3145	0.4181	0.3459

Dari Tabel 4. 3 yang ditampilkan, hasil evaluasi *ROUGE*-1 untuk *score F1* sebesar 0.413 menunjukkan kinerja model yang mampu menangkap kata yang sesuai dari judul asli dengan cukup baik. Sekitar 41% dari kata yang ada pada judul asli dapat dihasilkan/diprediksi oleh model. Dengan nilai *precision* yang lebih tinggi dibandingkan *recall* menandakan kata yang dihasilkan oleh model lebih relevan dengan judul asli, meskipun tidak setiap kata yang ada pada judul asli dapat ditangkap oleh model.

Untuk hasil evaluasi *ROUGE*-2 kurang baik dibandingkan dengan *ROUGE*-1 dengan score 0.217 atau hanya sekitar 21% kata-kata yang ada pada judul asli yang dapat dihasilkan oleh model. Hal ini sangat wajar untuk tugas *text generation* terutama dengan konteks yang lebih spesifik untuk teks pendek seperti judul, karena dua kata yang berurutan lebih sulit untuk ditangkap/diprediksi secara tepat.

Dari hasil evaluasi *ROUGE*-L didapatkan hasil 0.345 yang mengindikasikan bahwa model sudah dapat mempertahankan urutan kata penting dalam judul sekitar 34%. Hasil evaluasi *ROUGE*-L memberikan gambaran bahwa meskipun tidak semua kata yang dihasilkan model tepat/sesuai dengan judul refrensi, struktur kalimat atau keterbacaan judul masih terjaga.

4.4.1 Analisis Skor *ROUGE* Tinggi

Tabel 4. 4 menyajikan beberapa hasil dari generated title yaitu judul yang telah dihasilkan oleh sistem dan juga hasil dari evaluasi ROUGE beserta kualitas dari judul yang dihasilkan.

Tabel 4. 4 Judul yang dihasilkan sistem dengan *ROUGE* tertinggi

judul asli	**judul baru**	**Rouge-1**	**Rouge-2**	**Rouge-L**	**Penilaian Kualitatif**	**Kualitas**
online medical teaching in china during the covid-19 pandemic tools modalities and challenges	the challenges and experiences of conducting online medical teaching practices during the covid-19 pandemic in china	0.774	0.483	0.516	Sangat Baik	Judul yang dihasilkan relevan, gramatikal, dan menangkap konsep inti dari referensi dengan tumpang tindih leksikal yang tinggi.
a population-based analysis of diabetes-related care measures foot complications and amputation during the covid-19 pandemic in ontario canada	association of the covid-19 pandemic with diabetes-related care measures foot complications and a	0.722	0.529	0.444	Kurang Baik	Judul secara gramatikal tidak lengkap (terpotong). Skor tinggi disebabkan oleh tumpang tindih kata kunci yang signifikan, mengabaikan kegagalan sintaksis.
impact of sars-cov-2 infection during pregnancy on infant neurobehavioral development a	sars-cov-2 infection and infant neurobehavioral development a case-	0.692	0.583	0.692	Kurang Baik	Skor tinggi disebabkan oleh ekstraksi frasa yang kemungkinan ada di teks sumber. Namun, judul yang dihasilkan

judul asli	judul baru	Rouge-1	Rouge-2	Rouge-L	Penilaian Kualitatif	Kualitas
case-control study						tidak lengkap dan kurang abstraksi.
impact of coronavirus disease on the management of lower urinary tract symptoms and voiding dysfunction	impact of covid-19 on lower urinary tract symptoms and management	0.692	0.417	0.615	Baik	Skor yang cukup tinggi dapat dibenarkan. Judul ini adalah parafrasa yang baik dan ringkas dari referensi, mempertahankan kata kunci penting. Sedikit abstraksi dengan hasil yang baik.

Skor ROUGE yang tinggi tidak menjamin kesesuaian gramatikal. Model dapat menghasilkan urutan kata kunci yang memiliki kecocokan tinggi dengan referensi, tetapi gagal membentuk kalimat yang koheren. Contoh yang jelas berdasarkan tabel adalah pada data dengan kolom judul asli: "a population-based analysis of diabetes-related care measures foot complications and amputation during the covid-19 pandemic in ontario canada". Judul baru: "association of the covid-19 pandemic with diabetes-related care measures foot complications and a". Judul ini mencapai skor ROUGE-1 sebesar 0.722, yang tergolong tinggi. Namun, judul tersebut secara gramatikal tidak lengkap dan terpotong di akhir kata "a". Model jelas gagal menyelesaikan kalimatnya, namun dikarena sebagian besar kata kunci yang relevan hadir, skor ROUGE-1 tetap tinggi. Ini menunjukkan kelemahan ROUGE-1: ia menghitung kecocokan "kantong kata" (*bag-of-words*) tanpa mempertimbangkan struktur sintaksis.

4.4.2 Analisis Skor ROUGE Rendah

Dalam banyak kasus, skor ROUGE yang rendah secara akurat mencerminkan kualitas ringkasan yang buruk. Umumnya hal tersebut terjadi ketika judul yang dihasilkan sama sekali tidak relevan, tidak koheren secara gramatikal di mana model menghasilkan teks yang tidak didukung oleh sumber. Beberapa contoh dari judul yang dihasilkan dari sistem serta evaluasi ROUGE dari judul yang dihasilkan dapat dilihat pada tabel berikut:

Tabel 4. 5 Judul hasil sistem dengan ROUGE terendah

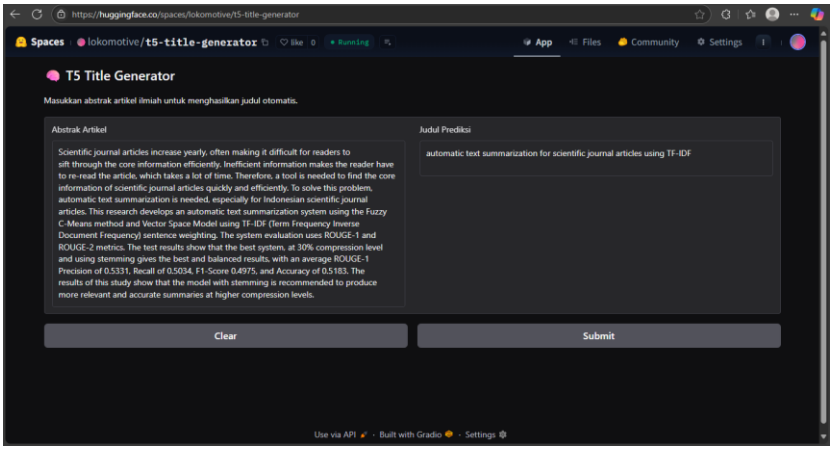
judul asli	**judul baru**	**Rouge-1**	**Rouge-2**	**Rouge-L**	**Penilaian Kualitatif**	**Kualitas**
discrepancies of sars-cov-2 *testing* results among patients with total laryngectomy	nasopharyngeal swabs in trachea and	0.000	0.000	0.000	Sangat Buruk	Skor rendah dapat dibenarkan karena judul yang dihasilkan tidak lengkap dan tidak relevan secara kontekstual. ROUGE secara akurat mengidentifikasi kegagalan dari generasi judul.
a storytelling robot managing persuasive and ethical stances via act-r an exploratory study	a needs-driven model of a storytelling robot controlled via act-r cognitive architecture	0.428	0.308	0.428	Sangat Baik	Judul yang dihasilkan secara kualitatif sangat baik dan deskriptif, skor ROUGE yang diperoleh cenderung rendah. Hal ini disebabkan oleh kemampuan model dalam melakukan parafrasa, di mana perbedaan leksikal judul tidak dapat ditangkap dengan baik oleh metrik ROUGE yang hanya mengukur kecocokan kata.

judul asli	judul baru	Rouge-1	Rouge-2	Rouge-L	Penilaian Kualitatif	Kualitas
going viral researching safely on social media	using social media to recruit participants for a study on abortion stigma in australia	0.285	0.105	0.190	Sangat Baik	Judul yang dihasilkan jauh lebih informatif dan relevan dengan teks sumber. rendahnya skor ROUGE disebabkan oleh minimnya tumpang tindih leksikal akibat kualitas judul referensi yang kurang baik.
covid-19 insights into virus-receptor interactions	a deep understanding of the high transmissibility of sars-cov-2 and other	0.000	0.000	0.000	Sangat Buruk	Skor rendah dapat dibenarkan. Judul yang dihasilkan terpotong dan tidak menyampaikan ide yang relevan dengan referensi.

Contoh dalam tabel adalah data yang menghasilkan skor ROUGE bernilai nol di semua metrik. Misalnya, untuk data judul asli: "a storytelling robot managing persuasive and ethical stances via act-r an exploratory study". Judul yang dihasilkan adalah: "a needs-driven model of a storytelling robot controlled via act-r cognitive architecture". Judul ini menerima skor ROUGE-1 F1 yang tergolong biasa saja yaitu sebesar 0.428. Namun, secara kualitatif, judul yang dihasilkan sangat baik. Ia secara akurat menggambarkan isi makalah, yang berfokus pada "model yang didorong oleh kebutuhan" (needs-driven model) yang mengontrol robot. Judul ini lebih deskriptif dan informatif daripada referensi. Namun, karena menggunakan frasa "needs-driven model" dan "controlled via act-r cognitive architecture" alih-alih "managing persuasive and ethical stances", kecocokan leksikal yang diberikan terbatas, sehingga menghasilkan skor yang rendah.

4.5 Deployment

Model dengan konfigurasi *parameter* terbaik kemudian diimplementasikan dalam sistem melalui website. Sebelum model dideploy ke sistem, hasil *training* model dengan konfigurasi terbaik disimpan terlebih dahulu. Kemudian hasil *training* tersebut diunggah ke laman *huggingface* untuk dilakukan *deploy* pada *space* sehingga dapat menghasilkan sistem dalam tampilan website. untuk tampilan dari model yang sudah diimplementasikan pada sistem dapat dilihat pada Gambar 4. 2.



Gambar 4. 2 *Interface Title Generator*

Hasil implementasi model dapat diakses pada laman https://huggingface.co/spaces/lokomotive/*T5*-title-generator. Untuk menggunakan sistem, input abstrak ke kolom abstrak artikel. Klik tombol submit untuk menghasilkan judul baru di kolom judul. Tombol clear digunakan untuk membersihkan kedua kolom.

4.6 Pembahasan

Penelitian yang dilakukan bertujuan untuk mengimplementasikan model *T5* guna menghasilkan judul artikel ilmiah secara otomatis serta melihat seberapa tinggi ROUGE yang dihasilkan oleh model dari evaluasi yang dilakukan terhadap data *testing* dengan judul asli dari dataset artikel ilmiah. Model *T5* yang akan digunakan di*training* menggunakan dataset artikel ilmiah yang telah preprocessing menggunakan *lowercase*. Data yang digunakan berjumlah 5000 data dari total 10.000 data yang ada pada dataset. Proses pembagian data *testing* dan *training* dilakukan secara random, menggunakan *random samples* dari *library pandas* dengan perbandingan 8:2.

Model di*training* dengan menggunakan beberapa skenario uji coba *hyperparameter tuning*. *Hyperparameter* yang digunakan untuk uji coba diantaranya adalah *epoch* dan jumlah *batch size*. *Hyperparameter* tersebut disesuaikan dengan konfigurasi yang telah ditentukan pada Tabel 3. 3. Dari hasil pengujian menggunakan parameter tersebut, performa terbaik didapat dari skenario uji coba *hyperparameter epoch* yang berjumlah 6 dengan ukuran *batch size* 4.

Hasil pengujian terlihat bahwa pemilihan kombinasi *parameter* dapat memberikan pengaruh yang signifikan terhadap performa model. *Batch size* yang kecil, menunjukkan bahwa model dengan konfigurasi yang lebih sederhana justru dapat memberikan hasil yang lebih optimal pada data yang digunakan. Di sisi lain, jumlah data *training* serta *epoch* yang lebih banyak membuat model dapat lebih baik bekerja karena model lebih banyak belajar terhadap data.

Model dengan performa terbaik selanjutnya disimpan lalu di*testing* untuk menguji seberapa baik kemampuan model bekerja pada data yang belum dilihat sebelumnya. Dari hasil *testing* selanjutnya dilakukan evaluasi model menggunakan matrik *ROUGE*. Dari hasil evaluasi tersebut didapatkan skor untuk *ROUGE*-1 sebesar 0.413, *ROUGE*-2 sebesar 0.217, dan *ROUGE*-L sebesar 0.345 berdasarkan rata-rata dari masing-masing nilai *F1-score* matrik *ROUGE*.

Nilai *F1-score* menunjukkan bahwa keseimbangan antara recall dan precission yang dihasilkan oleh model dalam mempertahankan informasi yang relevan kurang baik. Akurasi *F1-score* untuk *ROUGE*-1 mencapai nilai 0.413 menunjukkan bahwa generasi judul oleh sistem memiliki 41% kesamaan kata yang cukup relevan dengan refrensi, hal ini menandakan model berhasil menangkap cukup banyak kata utama dari teks asli. *ROUGE*-2 yang lebih rendah yaitu 0.217 menandakan bahwa pasangan kata yang relevan dalam judul yang dihasilkan perlu adanya perbaikan agar jauh lebih mirip dengan refrensi. *ROUGE*-L sebesar 0.345 menunjukkan bahwa sistem mampu menangkap 34% dari struktur kalimat dan ide utama teks asli, meskipun tidak selalu sama untuk setiap susunan frasanya, hal ini mengindikasikan bahwa model cukup baik dalam mempertahankan kata yang relevan, meskipun masih ada perbedaan kata dengan judul refrensi.

Dilakukan uji validasi dengan validator eksternal dengan mengujikan 50 sampel data untuk dinilai valid tidaknya judul yang dihasilkan oleh sistem. Berdasarkan analisis dari 50 sampel data yang dipilih secara random dari data uji juga telah dilakukan uji validasi oleh validator eksternal pada Tabel 4. 6, ditemukan hubungan antara skor ROUGE dan validitas judul yang dihasilkan oleh model.

Tabel 4. 6 Sampel data yang diuji Validator

Abstrak	**Judul yang Dihasilkan**	**Rouge-1**	**Rouge-2**	**Rouge-L**	**Validitas Judul**
background post-covid19 syndrome also referred as long covid describes persisting symptoms after sars-cov2 infection including myalgia fatigue respiratory or neurological symptoms objective symptoms are often lacking thus resembling a fibromyalgia-like syndrome...	fibromyalgia-like syndrome a new mobile health application for patients with co	500	267	313	Tidak Valid (Terpotong)
background the novel coronavirus disease 2019 covid-19 pandemic has tested the ability of universities to provide a high-quality safe educational experience for students due to campuses shutting down as a result online learning could shift from a traditional classroom teaching mode...	the challenges and experiences of conducting online medical teaching practices during the covid-19 pandemic in china	774	483	516	Valid
we present a storytelling robot controlled via the act-r cognitive architecture able to adopt different persuasive techniques and ethical stances while conversing about some topics concerning covid-19 the main contribution of the paper consists in the proposal of a needs-driven model...	a needs-driven model of a storytelling robot controlled via act-r cognitive architecture	429	308	429	Valid
cardiac rehabilitation cr is a well-known intervention for the secondary prevention of cardiovascular diseases however in japan the outpatient cr participation rate is estimated to be very low cardiac telerehabilitation ctr can be defined as a remote cr program using digital health technology...	cardiac rehabilitation in japan evidence of telerehabilitation using digital health technology	400	111	300	Valid
objectives assess differences in movement behaviours within the 24-hour cycle including light intensity physical activity lpa	differences in movement behaviours during the covid-19 pandemic a	545	323	545	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
moderate-to-vigorous physical activity mvpa sedentary time and sleep before and during the covid-19 pandemic and assess these differences stratified by several relevant factors...	longitudinal cohort study				
background during the covid-19 pandemic nurses might experience added emotional stress this study examined the relationship between gratitude and psychological stress to explore effective psychological support among nurses methods a cross-sectional survey assessed the level of psychological distress...	the relationship between gratitude and psychological stress to nurses during the covid-19 pandemic a cross	452	69	452	Tidak Valid (Terpotong)
background cardiogenic shock cs is a life-threatening condition characterized by circulatory insufficiency caused by an acute dysfunction of the heart pump the pathophysiological approach to cs has recently been enriched by the tissue consequences of low flow including inflammation...	cardiogenic shock reversal in cardiogenic shock cs a multicentered multicentered	296	80	296	Tidak Valid (Canggung & Repetitif)
…	…	…	…	…	…
aims the study aimed to understand the role and the core values of pharmacists and the professional expectations of medical staff for pharmacists in treating covid-19 patients from the perspectives of the frontline medical staff the findings help to understand and provide a reference for the career growth path...	pharmacists role and the core values of pharmacists in treating covid-19 patients perspectives of frontline	611	353	500	Valid
successful implementation of virtual healthcare depends immensely on patients perceptions and satisfaction this cross-sectional study assessed patients perceptions of and factors associated with poor and average satisfaction with the	patients perceptions of and factors associated with poor and average satisfaction with outpatient telemedicine clinics	457	242	457	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
outpatient telemedicine clinics in the kingdom of saudi arabia ksa...					
the american academy of pediatric dentistry recommends regular preventive dental examinations and cleanings for children starting from the time their first tooth appears or by age 1 year 1 in 2020 dental practices adjusted their services in response to the covid-19 pandemic and access to dental care was disrupted...	the american academy of pediatric dentistry recommends regular preventive dental examinations and cleanings for children aged	267	71	200	Valid
purpose to evaluate the impact of the covid-19 pandemic on canadian intensive care unit icu workers methods between june and august 2020 we distributed a cross-sectional online survey of icu workers evaluating the impact of the pandemic coping strategies symptoms of post-traumatic stress disorder ptsd...	impact of the covid-19 pandemic on icu workers in canada a cross-	643	462	643	Tidak Valid (Terpotong)

Dari 50 sampel yang dianalisis terdapat 28 judul (56%) yang dinilai Valid. Judul-judul tersebut memiliki tata bahasa yang benar, koheren, dan relevan dengan konten abstraknya. Sedangkan 22 judul (44%) dinilai Tidak Valid. Sebagian besar ketidak validan dari judul disebabkan oleh judul yang terpotong di tengah kalimat. Penyebab lainnya termasuk pengulangan kata yang canggung dan ketidak koherenan.

Dari analisis uji validitas yang dilakukan, skor ROUGE tinggi tidak menjamin validitas. Berdasarkan uji validitas skor ROUGE yang tinggi tidak selalu berarti judul yang dihasilkan berkualitas baik. Contoh yang paling menonjol adalah pada data dengan judul: "*sars-cov-2 infection and infant neurobehavioral*

development a case-“ dengan Skor ROUGE-1 0.692 ROUGE-L 0.692. Meskipun skor yang dihasilkan tinggi, judul ini jelas tidak valid karena terpotong. Skor yang tinggi ini disebabkan oleh tumpang tindih frasa kunci yang signifikan dengan teks target, yang menunjukkan bahwa ROUGE lebih mengukur dalam kesamaan leksikal daripada keutuhan gramatikal. Model berhasil mengekstrak bagian penting, tetapi gagal menyelesaikannya menjadi kalimat secara utuh.

Adapun skor ROUGE yang rendah tidak secara otomatis menandakan judul yang tidak valid atau tidak dapat dibaca. Beberapa judul yang valid secara gramatikal dan relevan secara semantik justru mendapatkan skor ROUGE yang rendah. Data dengan judul “*using social media to recruit participants for a study on abortion stigma in australia*” menghasilkan skor ROUGE-1 0.286, ROUGE-L 0.190. Judul yang dihasilkan valid, sangat informatif, dan merangkum isi abstrak dengan baik. Namun, skor ROUGE-nya sangat rendah karena judul referensi (target_text) adalah "*going viral researching safely on social media*", yang sangat umum. Model *T5* menghasilkan judul yang lebih spesifik dan deskriptif (sebuah bentuk peringkasan abstraktif yang baik), tetapi karena pilihan katanya sangat berbeda dari referensi, ia tidak masuk perhitungan oleh metrik ROUGE.

4.7 Integrasi

Judul merupakan penghubung antara pembaca dengan sebuah karya tulis yang berperan penting menarik perhatian bagi pembaca. Sehingga apabila judul dari suatu karya tulis ilmiah menarik maka banyak yang akan tertarik untuk membaca sebuah karya tulis ilmiah tersebut. Hal ini sesuai dengan ayat Al-Qur’an surah Al-‘Alaq ayat 1-5 yang menganjurkan setiap ummat untuk membaca dan menulis.

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ خَلَقَ الْإِنسَانَ مِنْ عَلَقٍ اقْرَأْ وَرَبُّكَ الْأَكْرَمُ الَّذِي عَلَّمَ بِالْقَلَمِ عَلَّمَ الْإِنسَانَ مَا لَمْ يَعْلَمْ

Artinya:

Bacalah, (wahai Nabi Muhammad saw., wahyu Ilahi yang beberapa saat lagi akan engkau terima; dan bacalah juga alam dan masyarakatmu) dengan (atau demi) nama Tuhan Pemeliharanmu yang menciptakan (semua makhluk)! (Dia adalah Tuhan) yang telah menciptakan manusia dari 'alaq (sesuatu yang berdempet di dinding rahim). Bacalah (berulang-ulang)dan Tuhan Pemeliharamu Maha Pemurah. Yang mengajar dengan pena (yakni dengan usaha dan sarana mereka), (dan Dia juga yang) mengajar manusia (tanpa alat dan usaha mereka) apa yang belum diketahui-(nya).(*QS. Al-'Alaq ayat 1-5*)(Shihab, 2020)

Sejalan dengan ayat tersebut, penelitian ini berfokus pada pemanfaatan kalam dalam bentuk modern, yaitu data tekstual digital. Penelitian yang dilakukan menggunakan abstrak dari artikel ilmiah sebagai sumber data. Abstrak merupakan ringkasan dari keseluruhan isi artikel ilmiah yang disajikan dalam bentuk tulisan. Dengan mengembangkan sistem yang dapat menghasilkan judul baru berdasarkan abstrak, penelitian yang dilakukan dapat mencerminkan proses pemanfaatan ilmu berbasis kalam digital.

Menurut Tafsir al-Misbah Kata القلم terambil dari kata kerja قلم yang berarti memotong ujung sesuatu. Kata qalam disini dapat berarti hasil dari pnggunaan ayat tersebut, yakni tulisan, karena bahasa sering kali menggunakan kata yang berarti "alat" atau penyebab untuk menunjuk akibat atau hasil dari penyebab atau penggunaan alat tersebut (Lubis & Nasution, 2018). Proses ini sangat relevan dengan tafsir tersebut yang menekankan pentingnya pena sebagai medium pengajaran dan penyebaran ilmu. Model *T5* yang diimplementasikan dalam penelitian berfungsi sebagai "pena digital" yang menghasilkan tulisan (judul) dari tulisan lain (abstrak).

Sistem yang digunakan untuk menghasilkan judul secara otomatis dapat menjadi sebuah sarana untuk membantu dalam mencapai pemahaman tentang isi dari sebuah artikel, hal ini sesuai dengan sabda Rasulullah,

أحبُّ الناسِ إلى اللهِ تعالى أنفعُهم للناسِ

"Orang yang paling dicintai Allah adalah yang paling bermanfaat untuk orang lain." (HR At-Thabrani)

Hukum yang mubah, dapat menjadi sunnah jika sesuatu yang bermanfaat dapat membantu seseorang menjadi lebih dekat dengan Allah SWT. seperti yang diterangkan pada hadist tersebut bahwa sistem generasi judul otomatis yang dibuat, dapat dimanfaatkan oleh manusia untuk sarana mencari ilmu dan akan mendapat pahala melalui ilmu yang didapatkannya. Penelitian yang dilakukan dapat membantu penulis serta peneliti dalam menentukan judul untuk karya tulis ilmiah mereka dengan lebih mudah hanya berdasarkan abstrak dari tulisan yang sudah mereka buat.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian yang dilakukan berhasil mengimplementasikan model *T5 small* untuk menghasilkan judul artikel ilmiah secara otomatis berdasarkan abstrak serta mengevaluasi tinggi skor *ROUGE* yang dihasilkan oleh model. Model dilatih menggunakan 5.000 data artikel ilmiah dari dataset PubMed yang telah melalui *preprocessing*. Pengujian dilakukan dengan berbagai kombinasi *hyperparameter*, hasil terbaik diperoleh pada konfigurasi *epoch* 6 dan *batch size* 4. Evaluasi performa menggunakan metrik *ROUGE* menghasilkan skor F1 masing-masing sebesar *ROUGE*-1 0.413, *ROUGE*-2 0.217, dan *ROUGE*-L 0.345, yang menunjukkan bahwa model mampu menghasilkan judul yang cukup relevan, meskipun belum sepenuhnya komprehensif dalam merepresentasikan seluruh kata dari judul referensi.

Analisis kualitatif terhadap 50 sampel data menunjukkan bahwa skor ROUGE yang tinggi tidak selalu menjamin validitas judul secara gramatikal. Ditemukan kasus di mana judul dengan skor ROUGE tinggi ternyata terpotong atau tidak koheren. Sebaliknya, beberapa judul yang secara kualitatif baik dan informatif justru mendapatkan skor ROUGE yang rendah karena menggunakan parafrasa yang berbeda secara leksikal dari judul aslinya, menyoroti keterbatasan metrik ROUGE dalam menilai kualitas semantik secara menyeluruh. Dari 50 sampel, 56% (28 judul) dinilai valid, sementara 44% (22 judul) tidak valid, sebagian besar karena terpotong. Secara keseluruhan, penelitian ini menunjukkan bahwa

model *T5*-small efisien untuk tugas generasi judul artikel ilmiah karena dengan penggunaan memori yang lebih sedikit jika dibandingkan dengan model varian *T5* yang lain, model mampu menghasilkan judul yang relevan secara leksikal dan struktural dengan cukup baik.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, terdapat beberapa saran yang perlu dipertimbangkan untuk penelitian selanjutnya sebagai berikut:

a. Penggunaan model yang lebih besar seperti *T5*-Base atau *T5*-Large disarankan untuk penelitian selanjutnya, karena model dengan kapasitas parameter lebih besar berpotensi menghasilkan performa yang lebih baik dalam memahami konteks dan struktur kalimat yang kompleks.
b. Disarankan untuk melakukan eksperimen dengan teknik fine-tuning tambahan, seperti pengaturan learning rate atau strategi early stopping yang lebih adaptif.
c. Perlu dilakukan pengujian sistem terhadap pengguna langsung (user *testing*) untuk mengetahui efektivitas sistem secara praktis, serta mendapatkan feedback mengenai kualitas dan kegunaan judul yang dihasilkan.
d. Perluasan cakupan dataset ke bidang ilmu lain di luar kesehatan juga penting untuk menguji generalisasi sistem dalam menghasilkan judul dari abstrak di berbagai domain ilmiah.

DAFTAR PUSTAKA

Atitienei, S. C., & Marginean, A. (2024). Exploring the Limits of *T5*: Multitask Fine-Tuning for Reasoning, Summarization, Sentiment Analysis and Figurative Language Understanding. 01–08.

Cals, J. W., & Kotz, D. (2013). Effective writing and publishing scientific papers, part II: title and abstract. Journal of Clinical Epidemiology, 66(6), 585.

Chen, Y., & Eger, S. (2022). Transformers go for the LOLs: Generating (humourous) titles from scientific abstracts end-to-end. arXiv Preprint arXiv:2212.10522.

Dang, T. A., Raviv, L., & Galke, L. (2024). Tokenization and Morphology in Multilingual Language Models: A Comparative Analysis of m*T5* and By*T5*. arXiv Preprint arXiv:2410.11627.

Deshmukh, M., Sharma, H., Sunar, A., Somavanshi, S., & Salunke, D. (n.d.). AUTOMATIC TEXT SUMMARIZATION AND TITLE GENERATION USING LSTM.

Fox, C. W., & Burns, C. S. (2015). The relationship between manuscript title structure and success: Editorial decisions and citation performance for an ecological journal. Ecology and Evolution, 5(10), 1970–1980.

HaCohen-Kerner, Y., Miller, D., & Yigal, Y. (2020). The influence of preprocessing on text classification using a bag-of-words representation. PloS One, 15(5), e0232525.

Hanif, U. (2023). Research Paper Summarization Using Text-To-Text Transfer Transformer (*T5*) Model.

Jiang, F. K., & Hyland, K. (2023). Titles in research articles: Changes across time and discipline. Learned Publishing, 36(2), 239–248.

Jin, R., & Hauptmann, A. G. (2001). Title generation for machine-translated documents. 1229–1234.

Kale, M., & Rastogi, A. (2020). Text-to-text pre-*training* for data-to-text tasks. arXiv Preprint arXiv:2005.10433.

Koupaee, M., & Wang, W. Y. (2018). Wikihow: A large scale text summarization dataset. arXiv Preprint arXiv:1810.09305.

Libovický, J., Helcl, J., & Mareček, D. (2018). Input combination strategies for multi-source transformer decoder. arXiv Preprint arXiv:1811.04716.

Lodhwal, V., & Choudhary, G. (2023). Survey paper: Automatic title generation for text with RNN and pre-trained transformer language model. International Journal for Research in Applied Science and Engineering Technology, 11, 2117–2124.

Lubis, S., & Nasution, M. R. (2018). Nilai pendidikan pada surah al-alaq ayat 1-5 menurut quraish shihab. Jurnal Ilmiah Al-Hadi, 4(2), 919–941.

Masters, D., & Luschi, C. (2018). Revisiting small batch *training* for deep neural networks. arXiv Preprint arXiv:1804.07612.

Mishra, P., Diwan, C., Srinivasa, S., & Srinivasaraghavan, G. (2021). Automatic title generation for text with pre-trained transformer language model. 17–24.

Naseeba, B., Nikhil, S. B., Arjun, P., Nair, N. S., Sivathmika, T., & Challa, N. P. (2024). Generating Headlines from Article Summaries Using Transformer Models. 154–162.

Pal, A., Fan, L., & Igodifo, V. (n.d.). Text Summarization using BERT and *T5*.

Putra, J. W. G., & Khodra, M. L. (2017). Automatic title generation in scientific articles for authorship assistance: A summarization approach. Journal of ICT Research and Applications, 11(3), 253–267.

Raileanu, R., Goldstein, M., Szlam, A., & Fergus, R. (2020). Fast adaptation to new environments via policy-dynamics value functions. 7920–7931.

Rashmi, G. D., & Maheswari, O. (2023). Generation Using Lstm With Beam Search For Enhancing The Contextually Relevant Headlines. Journal of Namibian Studies, 33.

Rehman, T., Sanyal, D. K., & Chattopadhyay, S. (2025). Can pre-trained language models generate titles for research papers? 154–170.

Rossi, M. J., & Brand, J. C. (2020). Journal article titles impact their citation rates. Arthroscopy: The Journal of Arthroscopic & Related Surgery, 36(7), 2025–2029.

Setiyarini, D., Adji, T. B., & Hidayah, I. (2024). Evaluating Performance of Transformer Models for Dialogue Summarization: A Comparison of *T5*-Base, *T5*-Small, and BART-Base. 184–190.

Shaw, P., Uszkoreit, J., & Vaswani, A. (2018). Self-attention with relative position representations. arXiv Preprint arXiv:1803.02155.

Shihab, M. Q. (2020). Al-Quran dan Maknanya. Lentera Hati.

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.

Yin, Q., He, X., Zhuang, X., Zhao, Y., Yao, J., Shen, X., & Zhang, Q. (2024). Stablemask: Refining causal masking in decoder-only transformer. arXiv Preprint arXiv:2402.04779.

Yu, W., Zhu, C., Li, Z., Hu, Z., Wang, Q., Ji, H., & Jiang, M. (2022). A survey of knowledge-enhanced text generation. ACM Computing Surveys, 54(11s), 1–38.

Zaki, M. Z. (2024). Revolutionising Translation Technology: A Comparative Study of Variant Transformer Models–BERT, GPT and *T5*. Computer Science and Engineering–An International Journal, 14(3), 15–27.

LAMPIRAN

Lampiran 1

Profil validator eksternal untuk memvalidasi judul yang sesuai dengan kaidah penulisan dalam artikel ilmiah, serta dalam hal keterbacaan

Nama : Mohammad Daffa' Alfarizi

Riwayat Pendidikan

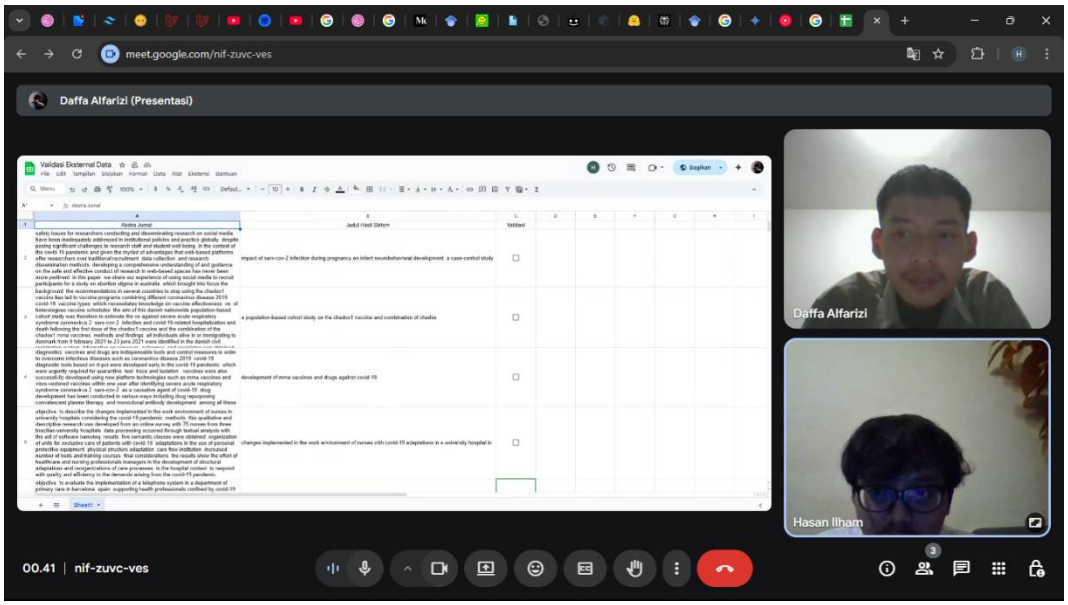
1. S-1 Bahasa dan Sastra Inggris - UIN Maulana Malik Ibrahim Malang

Riwayat Profesi

1. Penulis Media Online Radar Malang

2. Penulis dan Editor Media Online Jawapos

Dokumentasi dengan validator



Lampiran 2

Sampel data yang diuji oleh validator

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
background post-covid19 syndrome also referred as long covid describes persisting symptoms after sars-cov2 infection including myalgia fatigue respiratory or neurological symptoms objective symptoms are often lacking thus resembling a fibromyalgia-like syndrome...	fibromyalgia-like syndrome a new mobile health application for patients with co	500	267	313	Tidak Valid (Terpotong)
background the novel coronavirus disease 2019 covid-19 pandemic has tested the ability of universities to provide a high-quality safe educational experience for students due to campuses shutting down as a result online learning could shift from a traditional classroom teaching mode...	the challenges and experiences of conducting online medical teaching practices during the covid-19 pandemic in china	774	483	516	Valid
we present a storytelling robot controlled via the act-r cognitive architecture able to adopt different persuasive techniques and ethical stances while conversing about some topics concerning covid-19 the main contribution of the paper consists in the proposal of a needs-driven model...	a needs-driven model of a storytelling robot controlled via act-r cognitive architecture	429	308	429	Valid
cardiac rehabilitation cr is a well-known intervention for the secondary prevention of cardiovascular diseases however in japan the outpatient cr participation rate is estimated to be very low cardiac telerehabilitation ctr can be defined as a remote cr program using digital health technology...	cardiac rehabilitation in japan evidence of telerehabilitation using digital health technology	400	111	300	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
objectives assess differences in movement behaviours within the 24-hour cycle including light intensity physical activity lpa moderate-to-vigorous physical activity mvpa sedentary time and sleep before and during the covid-19 pandemic and assess these differences stratified by several relevant factors...	differences in movement behaviours during the covid-19 pandemic a longitudinal cohort study	545	323	545	Valid
background during the covid-19 pandemic nurses might experience added emotional stress this study examined the relationship between gratitude and psychological stress to explore effective psychological support among nurses methods a cross-sectional survey assessed the level of psychological distress...	the relationship between gratitude and psychological stress to nurses during the covid-19 pandemic a cross	452	69	452	Tidak Valid (Terpotong)
background cardiogenic shock cs is a life-threatening condition characterized by circulatory insufficiency caused by an acute dysfunction of the heart pump the pathophysiological approach to cs has recently been enriched by the tissue consequences of low flow including inflammation...	cardiogenic shock reversal in cardiogenic shock cs a multicentered multicentered	296	80	296	Tidak Valid (Canggung & Repetitif)
background as the municipality council area in colombo cmc experienced the highest number of cases until the end of january 2021 in sri lanka we carried out a serosurvey prior to initiation of the vaccination program to understand the extent of the sars-cov-2 outbreak...	sars-cov-2 seropositivity and seroprevalence in	500	222	400	Tidak Valid (Terpotong)
safety issues for researchers conducting and disseminating research on social media have been	using social media to recruit participants for a study on	286	105	190	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
inadequately addressed in institutional policies and practice globally despite posing significant challenges to research staff and student well-being in the context of the covid-19 pandemic...	abortion stigma in australia				
background previous studies on the pneumonia outbreak caused by the severe acute respiratory syndrome coronavirus 2 sars-cov-2 have focused on the general population and pregnant women while little is known about the effects of sars-cov-2 on retardation during and after pregnancy...	sars-cov-2 infection and infant neurobehavioral development a case-	692	583	692	Tidak Valid (Terpotong)
objective to discuss the consideration of performing radical cystectomy for rhabdomyosarcoma in children during the covid-19 pandemic in the urology department of adam malik hospital medan introduction rhabdomyosarcoma is a rare malignancy that develops from primitive mesenchymal stem cells...	radical cystectomy for rhabdomyosarcoma in children during the covid	387	138	387	Tidak Valid (Terpotong)
purpose of review coronavirus disease covid-19 has caused a crisis in the entire healthcare system since its emergence the urgency and priority of various diseases have impacted the medical and surgical treatment in this period we aim to review the impact of covid-19 on lower urinary tract symptoms...	impact of covid-19 on lower urinary tract symptoms and management	692	417	615	Valid
severe acute respiratory syndrome coronavirus 2 sars-cov-2 is a causative virus in the development of coronavirus disease 2019 covid-19 pandemic respiratory manifestations of sars-cov-2 infection such	immune response to the sars-cov-2 invasion may lead to uncontrolled release of pro	143	0	71	Tidak Valid (Terpotong & Tidak Relevan)

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
as acute lung injury ali and acute respiratory distress syndrome ards leads to hypoxia...					
background tracheostomy is performed in patients expected to require prolonged mechanical ventilation but to date optimal timing of tracheostomy has not been established the evidence concerning tracheostomy in covid-19 patients is particularly scarce we aimed to describe the relationship between early tracheostomy...	early tracheostomy in covid-19 patients a prospective cohort study	452	138	387	Valid
purpose prevention of severe acute respiratory syndrome coronavirus 2 sars-cov-2 infection is obtained with nasopharyngeal swabs by the way there is no consensus regarding sampling in totally laryngectomized subjects who thus breathe directly by the tracheostomy and theoretically may be infected in the trachea...	nasopharyngeal swabs in trachea and	0	0	0	Tidak Valid (Terpotong & Tidak Koheren)
the african union bureau of heads of state and government endorsed the covid-19 vaccine development and access strategy to vaccinate at least 60 of each country s population with a safe and efficacious vaccine by 2022 to achieve the population-level immunity needed to bring the pandemic under control...	covid-19 vaccination trends in west african union bureau and state estimates	467	214	400	Valid
disclaimer in an effort to expedite the publication of articles related to the covid-19 pandemic ajhp is posting these manuscripts online as soon as possible after acceptance accepted manuscripts have been	ajhp a review of pharmacy specialties certification for advanced pharmacy areas	333	0	250	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
peer-reviewed and copyedited but are posted online before technical formatting and author proofing...					
the recent outbreak of coronavirus disease 2019 covid-19 calls for rapid mobilization of scientists to probe and explore solutions to this deadly disease a limited understanding of the high transmissibility of sars-cov-2 severe acute respiratory syndrome coronavirus 2 relative to other coronavirus strains...	a deep understanding of the high transmissibility of sars-cov-2 and other	0	0	0	Tidak Valid (Terpotong & Tidak Koheren)
importance deferred diabetic foot screening and delays in timely care of acute foot complications during the covid-19 pandemic may have contributed to an increase in limb loss objective to evaluate the association of the covid-19 pandemic with diabetes-related care measures foot complications and amputation...	association of the covid-19 pandemic with diabetes-related care measures foot complications and a	722	529	444	Tidak Valid (Terpotong)
multisystem inflammatory syndrome in children mis-c is a delayed-onset covid-19-related hyperinflammatory systemic illness characterized by sars-cov-2 antigenemia cytokine storm and immune dysregulation however the role of the neutrophil has yet to be defined in adults with severe covid-19...	neutrophil activation in children with covid-19 and acute covid-19	522	286	348	Valid
background the aim of this study was to determine the relationship of inflammation with mortality in covid-19 hospitalized patients and to assess if the relationship differed by strata of type 2 diabetes status we hypothesized that the association of	the association of inflammation with mortality in covid-19 hospitalized patients a case-control died-	552	370	483	Tidak Valid (Terpotong)

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
inflammation with mortality was different by type 2 diabetes status...					
background the necessity to identify and isolate covid-19 patients to avoid intrahospital cross infections is particularly felt as a challenge clinically occult sars-cov-2 infection among patients admitted to the hospital is always considered a risk during the pandemic...	ct scan to reveal unexpected covid-19 in patients needing hospital admission	485	194	424	Valid
background vaccines against covid-19 are a powerful tool to control the current sars-cov-2 pandemic a thorough description of their immunogenicity among people living with hiv plwhiv is necessary we aimed to assess the immunogenicity of the mrna-1273 vaccine among plwhiv...	immunogenicity of mrna-1273 vaccine among people living with hiv plwhi	389	294	222	Tidak Valid (Terpotong)
introduction the covid-19 healthcare personnel study chps was designed to assess adverse short-term and long-term physical and mental health impacts of the coronavirus disease-2019 covid-19 pandemic on new york s physicians nurse practitioners and physician assistants...	covid-19 healthcare personnel study in new york s physicians nurse practitioners and physician assistants a	429	308	429	Tidak Valid (Terpotong)
awareness of infections which are transmitted between animals and humans have been given prominence following the covid-19 pandemic the orf infection in humans is rare recognition of orf lesions avoids misdiagnosing and incorrect treatment we present a case of a scottish-farmer with pain and discomfort...	scottish-farmer pain and discomfort from a lesion on her finger during the	242	0	182	Valid
severe acute respiratory syndrome coronavirus-2	sars-cov-2 a new breakthrough in the	345	74	345	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
sars-cov-2 has become a global crisis due to strong infectivity and fast...source	diagnostic and treatment of covid-19				
introduction coronavirus disease 19 covid-19 is an infectious disease caused by the newly discovered severe acute respiratory syndrome coronavirus 2 sars-cov-2 we have plenty of data about the clinical features of the disease s acute phase while little is known about the long-term consequences on survivors...	clinical and functional consequences of covid-19 pneumonia months after hospital discharge evidence synthesis	320	174	320	Valid
background covid-19 disproportionally affects pregnant women and their newborn yet little is known about the variables that modulate the maternal-fetal immune response to infection methods we prospectively studied socioeconomic biologic and clinical factors affecting humoral immunity...	maternal immunity to infection in unvaccinated pregnant women admitted to hospital for covid-19	258	69	258	Valid
introduction patients with acute respiratory distress syndrome caused by coronavirus disease 2019 covid-19 are at risk for superadded infections especially infections caused by multidrug resistant mdr pathogens before the covid-19 pandemic the prevalence of mdr infections...	a systematic review of the antimicrobial resistance pattern and biofilm formation in k pneumoniae	421	333	421	Valid
background the recommendations in several countries to stop using the chadox1 vaccine has led to vaccine programs combining different coronavirus disease 2019 covid-19 vaccine types which necessitates knowledge on vaccine effectiveness ve of heterologous vaccine schedules...	a population-based cohort study on the chadox1 vaccine and combination of chadox	419	195	233	Tidak Valid (Terpotong)

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
responding to the coronavirus disease 2019 covid-19 pandemic has been an unexpected and unprecedented global challenge for humanity in this century during this crisis specialists from the laboratories and frontline clinical personnel have made great efforts to prevent and treat covid-19...	sars-cov-2 sars-cov-2 a novel novel novel novel	0	0	0	Tidak Valid (Tidak Koheren & Repetitif)
in the age of digitalization social media has played a significant role in quickly spreading the news about current affairs from december 2019 to now coronavirus disease 2019 covid-19 with its several mutated shapes has more transmissible potential catastrophe and has become a severe phenomenon issue worldwide...	impact of social media on its users in china	167	91	167	Valid
background tracheostomy has been proposed as an option to help organize the healthcare system to face the unprecedented number of patients hospitalized for a covid-19-related acute respiratory distress syndrome ards in intensive care units icu it is however considered a particularly high-risk procedure for contamination...	tracheostomy in covid-19 critically ill patients a retrospective analysis	519	240	519	Valid
background the impact of qualitative olfactory disorders is underestimated parosmia the distorted perception of familiar odors and phantosmia the experience of odors in the absence of a stimulus can arise following postinfectious anosmia and the incidences of both have increased substantially...	postinfectious olfactory dysfunction syndrome a threaded and unstructured	200	111	200	Valid
background the united states food and drug administration fda monitors	a review of health products for covid-19	429	231	357	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
inspects and enforces the promotion of products by companies that claim to mitigate prevent treat diagnose or cure covid-19 the introduction of covid-19-related diagnostics and therapeutics during the pandemic...	infection a systematic review				
background capturing the complexity of an individual s uniqueness and life experiences is difficult new technology approaches are needed to simplify the process of knowing the person s story to improve care and communication method lifebio is currently developing a cross-platform application...	lifebio memory a novel machine learning-based application for people living with alzheimer	222	80	222	Valid
rationale over a year after the outbreak of sars-cov-2 and the ensuing covid-19 pandemic with its lockdowns and social distancing requirements being together with others again seems possible against this backdrop important questions arise about how to safely manage gatherings of large numbers of unrelated people...	covid-19 pandemic risks and risks a cross-sectional study at one of	0	0	0	Tidak Valid (Tidak Koheren)
abstractexperiencing empathy for others has been linked to worsening others feelings against their wishes these paternalistic empathic goals have been theorised to happen at the dyad level when an agent aims to worsen a target s emotional state they may also operate at a broader level...	agents s emotional state and covid-19 s ntotal vulnerability study 1 age studies	160	87	160	Tidak Valid (Tidak Koheren)
the outbreak of a new coronavirus sars-cov-2 was first identified in wuhan people s republic of china in 2019 which has led to a severe life-threatening form of pneumonia covid-19 research scientists all	conserved macrodomain adp ribose phosphatase phosphat	214	154	214	Tidak Valid (Terpotong)

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
around the world have been trying to find small molecule drugs to treat covid-19...					
mrna covid-19 vaccines have hegemonized the world market and their administration to the population promises to stop the pandemic however the waning of the humoral immune response which does not seem to last so many months after the completion of the vaccination program...	immunological mechanisms of waning immunity in the case of mrna covid-19	250	182	250	Valid
purpose this study detects sars-cov-2 in the ocular surface through one-step reverse-transcription droplet digital pcr one-step rt-ddpcr and evaluates the possibility of the ocular surface as a possible transmission route methods a single-center prospective observational study was designed...	sars-cov-2 in the ocular surface through one-step reverse-tran	485	323	485	Tidak Valid (Terpotong)
severe acute respiratory syndrome coronavirus 2 sars-cov-2 which has dramatically changed the world is a highly contagious virus the timely and accurate diagnosis of sars-cov-2 infections is vital for disease control and prevention here in this work a fluorescence immunoassay was developed...	fluorescence immunoassay to detect 2019 novel coronavirus antibodies 2019-ncov	364	129	303	Valid
the rapid surge of covid-19 cases worldwide drew attention to covid-19 infection as a new source of risk in transport the virus introduced a need for viral transmission mitigation as a major priority when selecting a mode of travel and caused a significant drop in public transport use...	risk perception worry and priority of covid-19 transmission in public transport	500	267	438	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
introduction this article explores one mental health company s urgent response to the global covid-19 pandemic and the multifaceted implications of quickly transitioning to telehealth services objectives the purpose of this article is to share information with interdisciplinary professionals...	transitioning to telehealth services in mental health company s urgent response to covid-19 pan	571	308	429	Tidak Valid (Terpotong)
to investigate the effect of covid-19 lockdown on match-play metrics in professional football referees during official matches of the spanish professional leagues forty-two professional football referees from the first n 20 and second division n 22 were monitored during 564 official games...	covid-19 lockdown and match-play metrics in professional football referees during official matches	621	370	552	Valid
a 63-year-old male with no significant past history and not on any regular medications previously had mild respiratory symptoms post the first dose of the astrazeneca cambridge england coronavirus disease 2019 covid-19 vaccine which were self-limiting following the second dose of the vaccine...	a 63-year-old male with no history of respiratory symptoms following the first astra	0	0	0	Tidak Valid (Terpotong)
background chronic obstructive pulmonary disease copd has become one of the most critical public health problems worldwide because many copd patients are using video-based social media to search for health information there is an urgent need to assess the information quality of copd videos on social media...	information quality of copd videos on social media a cross sectional study	323	69	258	Valid
aims the study aimed to understand the role and the core values of pharmacists	pharmacists role and the core values of pharmacists in	611	353	500	Valid

Abstrak	Judul yang Dihasilkan	Rouge-1	Rouge-2	Rouge-L	Validitas Judul
and the professional expectations of medical staff for pharmacists in treating covid-19 patients from the perspectives of the frontline medical staff the findings help to understand and provide a reference for the career growth path...	treating covid-19 patients perspectives of frontline				
successful implementation of virtual healthcare depends immensely on patients perceptions and satisfaction this cross-sectional study assessed patients perceptions of and factors associated with poor and average satisfaction with the outpatient telemedicine clinics in the kingdom of saudi arabia ksa...	patients perceptions of and factors associated with poor and average satisfaction with outpatient telemedicine clinics	457	242	457	Valid
mrna covid-19 vaccines have hegemonized the world market and their administration to the population promises to stop the pandemic however the waning of the humoral immune response which does not seem to last so many months after the completion of the vaccination program...	immunological mechanisms of waning immunity in the case of mrna covid-19	250	182	250	Valid