

Aplikasi Peringkat Teks Bahasa Indonesia Menggunakan Model Text-to-Text Transfer Transformer (T5)

¹Mohammad Yani, ²Nur Siti Khodijah, ³Rendi, ⁴Muhamad Mustamiin
¹Rekayasa Perangkat Lunak, Politeknik Negeri Indramayu, Indramayu
²Teknik Informatika, Politeknik Negeri Indramayu, Indramayu
³Rekayasa Perangkat Lunak, Politeknik Negeri Indramayu, Indramayu
⁴Teknik Informatika, Politeknik Negeri Indramayu, Indramayu

E-mail: [1mohammad.yani@polindra.ac.id](mailto:mohammad.yani@polindra.ac.id), [2nursitikhodijah01@gmail.com](mailto:nursitikhodijah01@gmail.com),
[3rendi@polindra.ac.id](mailto:rendi@polindra.ac.id), [4mustamiin@polindra.ac.id](mailto:mustamiin@polindra.ac.id), ,

ABSTRAK

Di era digital ini, informasi dapat kita peroleh dengan mudah melalui beberapa media yang ada seperti mesin pencari, tempat penyimpanan (*repository*) karya ilmiah, media sosial, portal berita, dan laman web. Akan tetapi biasanya informasi yang tersedia masih dalam bentuk teks yang sangat panjang. Sebagai contoh informasi tentang “Joko Widodo” pada laman Wikipedia memiliki panjang teks setidaknya berisi 8.716 karakter. Sebuah teks yang cukup panjang untuk dapat dipahami intisarinya. Ada beberapa penelitian tentang pengembangan aplikasi peringkat teks. Akan tetapi aplikasi peringkat teks yang menerima masukan dari berbagai format belum tersedia. Pada penelitian ini penulis mengusulkan sebuah sistem peringkat teks yang dapat melakukan peringkat teks menggunakan masukan dalam berbagai format. Sistem peringkat teks ini menggunakan input informasi dalam bentuk teks, file dokumen, laman web, dan gambar yang kemudian akan diringkas menggunakan model Text-to-Text Transfer Transformer (T5) untuk menghasilkan sebuah luaran hasil peringkat teks yang dapat disimpan dalam beberapa format file. Hasil evaluasi menunjukkan bahwa peringkat teks dengan masukan file berjenis dokumen atau teks memiliki nilai ROUGE yang lebih baik dari masukan dalam bentuk non-teks mencapai rata-rata 0,87.

Kata kunci : *peringkatan teks, model T5, teks multi format, abstraktif*

ABSTRACT

In the digital era, information can be easily accessed through various available media such as search engines, academic repositories, social media, news portals, and websites. However, the information available is often presented in lengthy textual form. For instance, information about 'Joko Widodo' on the Wikipedia page contains at least 8,716 characters of text. Such a long text can be challenging to extract its main points efficiently. Several studies have been conducted on the development of text summarization applications. Nevertheless, a summarization application that accepts input from various formats has not yet been developed. In this research, the author proposes a text summarization system capable of summarizing text using inputs in multiple formats. This text summarization system processes information inputs in the form of text, document files, web pages, and images, which are subsequently summarized using the Text-to-Text Transfer Transformer (T5) model to generate a summarized output that can be stored in various file formats. Evaluation results indicate that text summarization with document or text file inputs yields higher ROUGE scores compared to non-text inputs, reaching an average score of 0.87.

Keyword : *text summarization, T5 model, multi format text, abstractif*

1. PENDAHULUAN

Akhir-akhir ini, informasi dapat diperoleh dengan mudah melalui berbagai macam media seperti mesin pencari, *repository* karya ilmiah, media sosial, portal berita, dan laman web. Akan tetapi, informasi yang tersedia dalam kanal tersebut biasanya berupa teks yang sangat panjang dan membuat sulit pembaca. Oleh karena itu, sistem peringkasan teks sangat dibutuhkan (Agustinus dan Edwin, 2023).

Peringkasan dokumen teks otomatis merupakan solusi efektif untuk mengekstrak inti informasi dari teks dengan cepat dan efisien. Proses peringkasan dapat dilakukan melalui berbagai pendekatan, termasuk ekstraktif dan abstraktif. Pendekatan ekstraktif mengambil kalimat-kalimat penting dari dokumen asli, sementara pendekatan abstraktif menghasilkan kalimat-kalimat baru untuk menyajikan intisari dari dokumen tersebut (Yuliska, 2020).

Pentingnya peringkasan dokumen tidak hanya untuk mempercepat pencarian dan pembacaan informasi, tetapi juga untuk menghindari duplikasi dan redundansi dalam dokumen. Teknologi peringkasan dokumen dapat diterapkan dalam berbagai konteks, pada artikel berita, teks sederhana untuk pencarian, dan efisiensi pencarian informasi medis (Yuliska, 2020).

Penelitian tentang peringkasan teks dalam bahasa Inggris cukup banyak. Penelitian yang dilakukan oleh N.A. Ranggianto (2023) menunjukkan bahwa metode ROUGE dapat menjadi alat yang efektif untuk mengukur kinerja peringkasan teks, baik secara ekstraktif maupun abstraktif. Studi ini menemukan bahwa metode BERT (J. Devlin et al., 2018) memberikan hasil terbaik dibandingkan metode lainnya dalam konteks bahasa Inggris. Penelitian oleh Mustamiin M., *et al* (2018) terkait peringkasan multi-dokumen dalam bahasa Indonesia sebagai salah satu contoh peringkasan sederhana yang masih menerapkan metode ekstraksi kalimat-

kalimat dari dokumen yang kemudian dikelompokkan menggunakan pengelompokan berbasis hirarki.

Selain itu, pada penelitian Nyoman dan Ni Nengah (2023) mendukung penggunaan model Text-to-Text Transfer Transformers (T5) dalam peringkasan teks berita, menunjukkan bahwa model ini unggul dalam menghasilkan ringkasan yang lebih bermakna.

Meskipun penelitian tentang peringkasan teks sudah banyak dilakukan, namun penelitian tentang peringkasan teks berbahasa Indonesia dengan multi-format input masih belum tersedia. Untuk itu pada artikel ini penulis akan membuat aplikasi peringkasan teks berbahasa Indonesia dengan menggunakan multi-format input.

Artikel ini disusun dengan susunan dimulai dengan penyampaian kajian literatur, metodologi yang digunakan, pembahasan hasil penelitian, dan diakhiri dengan kesimpulan.

2. KAJIAN LITERATUR

1. Sistem Peringkasan Teks

Sistem peringkasan teks adalah metode yang digunakan untuk membuat ringkasan singkat dan akurat dari suatu dokumen teks yang panjang. Ada dua jenis utama sistem peringkasan teks, yaitu peringkasan ekstraktif dan peringkasan abstraktif. Namun ada juga sebagian peneliti menggunakan pendekatan hibrida, yaitu gabungan antara ekstraktif dan abstraktif (R. Farah et al., 2023). Ada beberapa penelitian terkait dengan aplikasi peringkasan teks ini.

Ranggianto mengembangkan peringkasan teks menggunakan pendekatan ekstraktif dan abstraktif menggunakan model BERT (2023).

Mustamiin et al., mengembangkan aplikasi peringkasan teks sederhana Bahasa Indonesia dengan menggunakan pendekatan ekstraktif (2018).

Nyoman and Ni Nengah mengembangkan peringkasan teks menggunakan model T5 (2023).

2. Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan domain ilmu yang mempelajari tentang bagaimana mengolah bahasa manusia. NLP bertujuan untuk memungkinkan komputer memahami, menafsirkan, dan memproses bahasa alami dengan cara yang berguna. Beberapa teknik utama dalam NLP meliputi tokenisasi, *stemming*, *lemmatization*, *parsing*, dan *Named Entity Recognition (NER)*. NLP memiliki banyak aplikasi, seperti pencarian informasi, *chatbots*, analisis sentimen, dan terjemahan otomatis (Jurafsky & Martin, 2020), (I. Nitin and JD. Fred, 2010), (R. Kibble, 2013).

3. Text-to-Text Transfer Transformer (T5)

Text-to-Text Transfer Transformer (T5) adalah model pra *encoder* dan *decoder* yang dilatih menggunakan pendekatan *supervised* dan *unsupervised* untuk melakukan pengolahan berupa konversi dari teks ke teks (R. Colin et al., 2020).

4. ROUGE

ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) adalah serangkaian metrik untuk mengevaluasi kualitas ringkasan teks. ROUGE mengukur kesamaan antara ringkasan yang dihasilkan secara otomatis dengan ringkasan referensi manusia berdasarkan beberapa aspek seperti n-gram, LCS (*Longest Common Subsequence*), dan kehadiran frasa (Lin, 2004). Metrik ROUGE sering digunakan dalam penelitian NLP untuk menilai efektivitas model peringkasan teks (Halimah et al, 2022), (PI. Dewi and A. Fahri, 2019).

1. Black Box Testing

Black box testing merupakan metode pengujian yang paling banyak dan mudah digunakan. Metode ini hanya mengecek fungsionalitas dari sistem saja bukan pada detail kode secara rinci (B.B Agarwal et al., 2010).

3. METODOLOGI

1. Metodologi Pengembangan Perangkat Lunak

Metodologi pengembangan perangkat lunak yang digunakan adalah model *Waterfall* (S. Ian, 2001). Dimulai dari analisis kebutuhan pengguna, pembuatan rancangan perangkat lunak, koding, verifikasi dan validasi, perawatan sampai dengan dokumentasi.

a. Analisis Kebutuhan

Pada tahap ini semua kebutuhan baik kebutuhan pengguna maupun kebutuhan pengembang dianalisis dan disusun. Penyusunan kebutuhan tersebut dapat diperoleh dengan cara wawancara, survei, atau diskusi langsung.

b. Perancangan Perangkat Lunak

Tahap ini merupakan tahapan lanjutan setelah tahap analisis kebutuhan telah dilakukan dan lengkap.

c. Koding

Setelah sistem dirancang. Sistem akan diimplementasikan dengan melakukan penulisan-penulisan kode menggunakan bahasa pemrograman

d. Verifikasi dan Validasi

Pada tahap ini dilakukan verifikasi terhadap sistem dan pengujian terhadap sistem untuk mengetahui apakah sistem sudah memenuhi persyaratan. Kemudian hasilnya divalidasi oleh pengguna.

e. Perawatan

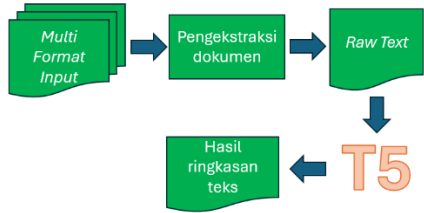
Tahap ini dilakukan setelah sistem dijalankan secara ril oleh pengguna. Tahap ini bertujuan untuk melakukan perawatan jika terdapat bagian dari sistem yang masih salah.

f. Dokumentasi

Setelah tahap pemeliharaan (*maintenance*), penting untuk melanjutkan dengan tahap Arsip dan Dokumentasi (*Archiving and Documentation*) untuk memastikan bahwa semua aspek pengembangan perangkat lunak terdokumentasi dengan baik dan siap untuk referensi di masa depan.

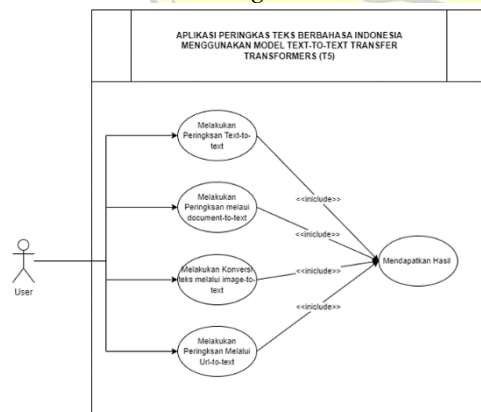
2. Rancangan Sistem

Rancangan sistem terdiri dari gambaran sistem secara umum dan rancangan aktivitas untuk empat fitur utama, yaitu: *text-to-text*, *document-to-text*, *image-to-text*, dan *url-to-text*. Gambaran sistem secara umum dapat dilihat pada Gambar 1.



Gambar 1. Gambaran Sistem Secara Umum

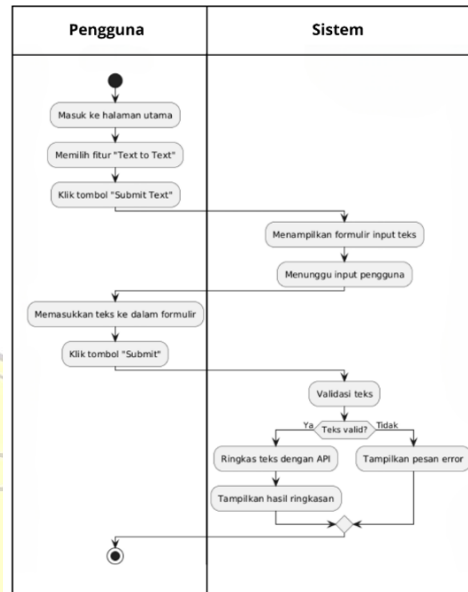
Dari Gambar 1 kita dapat lihat bahwa masukan dari aplikasi ini adalah dokumen dari beberapa format. Kemudian dokumen tersebut diekstraksi menjadi *raw text* yang kemudian menjadi input dari model T5. Model T5 kemudian melakukan peringkasan dengan pendekatan abstraktif untuk menghasilkan kalimat-kalimat baru untuk menyajikan intisari dari dokumen tersebut. Hasil abstraksi kemudian ditampilkan sebagai luaran. Gambar 2 menunjukkan gambaran sistem dilihat dari sisi *use case diagram*.



Gambar 2. Diagram Use Case Sistem

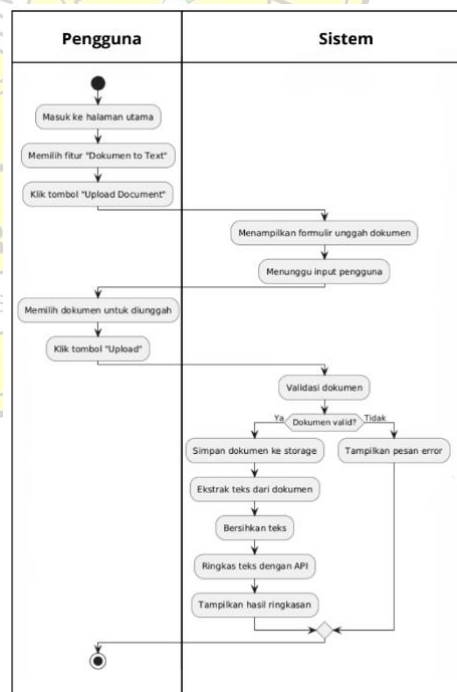
Gambar 2 menunjukkan bahwa aplikasi peringkasan teks yang dikembangkan terdiri dari empat fitur utama, yaitu: meringkas teks dari masukan teks murni, meringkas teks dari masukan dokumen, meringkas

teks dari masukan *image*, dan meringkas teks dari masukan URL.



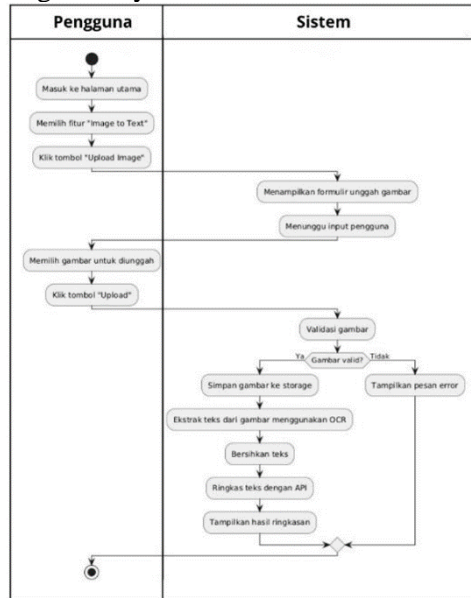
Gambar 3. Diagram Aktivitas untuk Fitur *Text-to-Text*

Gambar 3 menunjukkan bagaimana aplikasi ini melakukan peringkasan teks dari sebuah *raw-text*.



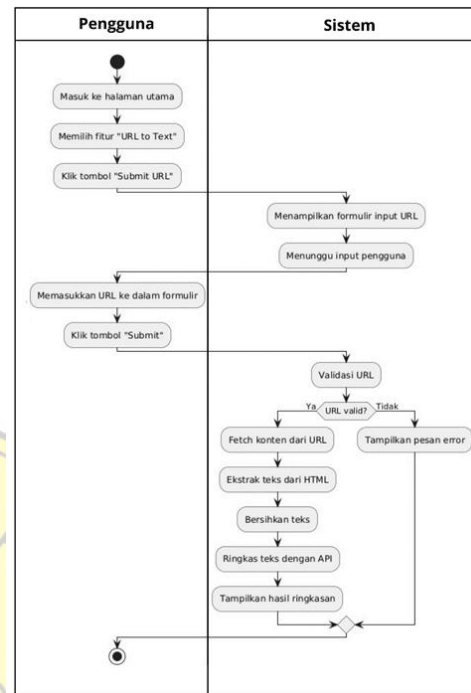
Gambar 4. Diagram Aktivitas untuk Fitur *Document-to-Text*

Gambar 4 menunjukkan bagaimana sebuah mauskun dalam format tertentu akan diekstrak menjadi *raw-text* sebelum digunakan oleh model T5 untuk dibuatkan ringkasannya.



Gambar 5. Diagram Aktivitas untuk Fitur *Image-to-Text*

Gambar 5 menunjukkan bagaimana sebuah masukan dalam format *image* diekstraksi ke dalam bentuk teks. Di sini penulis menggunakan *Optical Character Recognition* (OCR) untuk mengekstraksi teks yang ada pada *image*.



Gambar 6. Diagram Aktivitas untuk Fitur *Url-to-Text*

Gambar 6 menunjukkan bagaimana *input* dalam bentuk *Uniform Resource Locator* (URL) dapat diekstraksi menjadi *raw-text*. Ekstraksi diawali dengan mengambil konten dari HTML, kemudian konten HTML diekstrak menjadi sebuah teks, dan setelah itu proses pembersihan teks menjadi *raw-text*.

4. HASIL DAN PEMBAHASAN

Pada bagian ini penulis akan memberikan hasil *running example* dari aplikasi peringkasan teks yang penulis kembangkan dan hasil serta pembahasan dari penelitian ini.

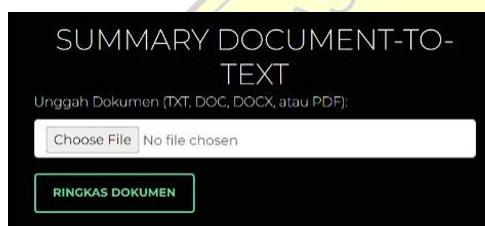
1. Tampilan *Running Example*

Pada bagian ini penulis akan memberikan hasil *running example* dari aplikasi peringkasan teks yang penulis kembangkan.



Gambar 7. Contoh dan Tampilan *Running Example* untuk Fitur *Text-to-Text*

Dalam contoh pada Gambar 7 ini, pengguna langsung memasukkan *raw-text* ke aplikasi peringkas teks. Kemudian pengguna dapat menggeneralisasi *raw-text* tersebut dengan menekan tombol “Ringkas”.



Gambar 8. Contoh dan Tampilan *Running Example* untuk Fitur *Document-to-Text*

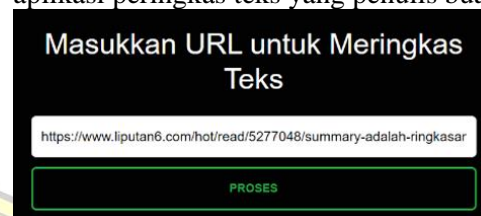
Pada fitur ini pada Gambar 8, pengguna dapat memberikan masukan ke dalam aplikasi dengan menggunakan format dokumen yang diinginkan dengan memilih file dokumen yang akan diringkas pada tombol “Choose File”. Kemudian untuk melihat hasil ringkasannya, pengguna dapat menekan tombol “Ringkas Dokumen”.



Gambar 9. Contoh dan Tampilan *Running Example* untuk Fitur *Image-to-Text*

Pada fitur sebagaimana pada Gambar 9 menunjukkan bagaimana pengguna dapat memberikan masukan ke aplikasi berupa file *image* yang informasinya akan

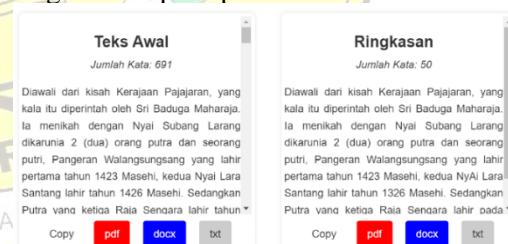
dilakukan peringkasan melalui tombol “Choose File”. Setelah file terpilih, pengguna dapat menggunakan tombol “Upload” untuk mengekstraksi informasi dalam *image* menjadi *raw-text* yang selanjutnya akan menjadi *input* bagi aplikasi peringkas teks yang penulis buat.



Gambar 10. Contoh dan Tampilan *Running Example* untuk Fitur *Url-to-Text*

Pada Gambar 10 menunjukkan bagaimana pengguna dapat memberikan masukan ke aplikasi dengan menggunakan alamat URL dari informasi yang akan diringkas. Untuk melihat hasil ringkasannya, pengguna dapat menggunakan tombol “Proses”.

Dari keempat fitur di atas pada akhirnya menampilkan sebuah tampilan hasil ringkasan seperti pada Gambar 11.



Gambar 11. Tampilan *Summary* Hasil Ringkasan Setelah Memilih Fitur Yang Digunakan

2. Hasil dan Pembahasan

Hasil ini menghasilkan Aplikasi Peringkas Teks Berbahasa Indonesia Menggunakan Model Text-To-Text-Transfer-Transformer(T5). Aplikasi ini dapat digunakan untuk meringkas teks melalui teks, *image*, dokumen dan url.

Pembahasan Pada bagian ini, dilakukan analisis mendalam terhadap hasil implementasi Aplikasi Peringkas Teks Berbahasa Indonesia menggunakan

model Text-To-Text Transfer Transformers (T5). Pembahasan ini mencakup beberapa aspek utama sebagai berikut:

a. Evaluasi Efektivitas Model T5

- Kinerja Model dalam Merangkum Teks: di sini, peneliti menganalisis bagaimana model T5 berfungsi dalam merangkum teks berbahasa Indonesia. Evaluasi dilakukan berdasarkan berbagai jenis teks yang diuji, termasuk teks dokumen, gambar yang diubah menjadi teks, dan teks yang diambil dari URL. Peneliti memeriksa apakah ringkasan yang dihasilkan oleh model tersebut relevan dan ringkas sesuai dengan konteks asli teks
- Kualitas Ringkasan diperiksa menggunakan metrik ROUGE. Hasil pengujian ini memberikan gambaran tentang seberapa baik model T5 dalam menciptakan ringkasan yang informatif dan relevan dibandingkan dengan ringkasan referensi yang dibuat oleh manusia.

b. Kesesuaian Desain Antarmuka

- Desain Antarmuka Pengguna (UI): Pembahasan mengenai bagaimana desain antarmuka pengguna aplikasi diimplementasikan untuk memfasilitasi interaksi yang intuitif dan mudah digunakan. Evaluasi dilakukan terhadap elemen-elemen desain, tata letak, dan navigasi aplikasi untuk memastikan bahwa aplikasi memenuhi kebutuhan pengguna dengan baik.
- Pengalaman Pengguna (UX): Peneliti mengevaluasi pengalaman pengguna dalam menggunakan aplikasi, termasuk kemudahan akses fitur, kejelasan instruksi, dan kenyamanan dalam proses interaksi dengan aplikasi. Umpan balik pengguna juga diperhatikan untuk mengevaluasi apakah aplikasi mudah dipahami dan digunakan.

c. Tantangan dalam Pengembangan dan Pengujian

- Masalah Teknis: identifikasi dan analisis tantangan teknis yang dihadapi selama pengembangan aplikasi, seperti integrasi dengan API Hugging Face, penanganan berbagai format file, dan pemrosesan teks yang kompleks. Peneliti juga mengidentifikasi kendala dalam pengujian dan evaluasi hasil ringkasan.
- Perbaikan yang Diperlukan: diskusi tentang area perbaikan yang mungkin diperlukan untuk meningkatkan kinerja aplikasi, baik dari segi teknis maupun desain. Ini bisa mencakup optimasi model T5, perbaikan antarmuka pengguna, atau peningkatan fungsionalitas aplikasi.

3. Pengujian Sistem

Pada tahap ini penulis mengevaluasi performa sistem dengan menggunakan metode ROUGE. Tabel 1 sampai Tabel 4 menunjukkan hasil evaluasi performa sistem untuk empat fitur utama. Pengukuran ROUGE dilakukan dengan tiga jenis, yaitu ROUGE 1 untuk pengukuran kesamaan kata secara unigram, ROUGE 2 untuk pengukuran kesamaan kata secara bigram, dan ROUGE L untuk pengukuran kesamaan dalam urutan kata secara *Longest Common Subsequence*.

Tabel 1. Hasil Pengujian Fitur *Text-to-Text*

| <i>Metric</i> | <i>Nilai</i> |
|---------------|--------------|
| ROUGE 1 | 0,5405 |
| ROUGE 2 | 0,3119 |
| ROUGE-L | 0,5405 |

Tabel 2. Hasil Pengujian Fitur *Document-to-Text*

| <i>Metric</i> | <i>Nilai</i> |
|---------------|--------------|
| ROUGE 1 | 0,8889 |
| ROUGE 2 | 0,8522 |
| ROUGE-L | 0,8889 |

Tabel 3. Hasil Pengujian Fitur *Image-to-Text*

| <i>Metric</i> | <i>Nilai</i> |
|---------------|--------------|
| ROUGE 1 | 0,3832 |
| ROUGE 2 | 0,3679 |
| ROUGE-L | 0,3832 |

Tabel 4. Hasil Pengujian Fitur *Url-to-Text*

| <i>Metric</i> | <i>Nilai</i> |
|---------------|--------------|
| ROUGE 1 | 0,6050 |
| ROUGE 2 | 0,5821 |
| ROUGE-L | 0,6050 |

Dari Tabel 1 sampai Tabel 4 jika dihitung reratanya dalam persen adalah ROUGE 1 mencapai 60,55%, ROUGE-2 52,85% dan ROUGE L 60,44%.

4. Analisis

Hasil evaluasi Image-to-Text dan Url-to-Text memiliki nilai ROUGE yang lebih rendah dibandingkan dengan Text-to-Text dan Document-to-Text karena terjadi beberapa kehilangan informasi pada saat proses ekstraksi image dan url ke dalam teks.

Hasil pengukuran Document-to-Text memiliki nilai paling tinggi dikarenakan ekstraksi ke teks lebih mudah dibandingkan dengan Image-to-Text dan Url-To-Text.

5. KESIMPULAN

Penelitian ini menghasilkan sebuah aplikasi peringkas teks menggunakan model T5 yang dapat mengatasi masalah multi format input yang mampu menerima input dari teks, input dalam bentuk dokumen dan laman web menghasilkan output yang mempertahankan substansi dari teks asli sebagai intisari dari output

Tingkat evaluasi menggunakan ROUGE menunjukkan bahwa input teks menggunakan dokumen atau secara langsung mencapai nilai ROUGE lebih baik dibandingkan dengan yang non-teks.

Penelitian ke depan adalah bagaimana mengembangkan sebuah ekstraktor dokumen *input* menjadi sebuah teks yang dapat menghasilkan hasil ekstraksi yang lebih bersih. Dengan demikian akan meningkatkan performa hasil ringkasan teks yang diberikan.

6. UCAPAN TERIMA KASIH

Ucapan terima kasih kepada Pusat Penelitian dan Pengabdian Kepada Masyarakat Politeknik Negeri Indramayu atas dana penelitian yang telah diberikan.

Penelitian ini adalah sebagai luaran wajib dari kegiatan penelitian yang dibiayai oleh Pusat Penelitian dan Pengabdian Kepada Masyarakat Politeknik Negeri Indramayu melalui hibah PUKTI SIPASTI dengan nomor kontrak : 0720/PL42.PL42.9/AL.04/2024.

DAFTAR PUSTAKA

- B.B Agarwal et al. (2010). Software Engineering and Testing, an Introduction. Jones and Bartlett Publishers .
- D. Jurafsky and J.H. Martin (2009). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2nd ed. London, England: Prentice Hall.
- Halimah et al. (2022). Peringkasan Teks Otomatis (automated text summarization) Pada Artikel Berbahasa Indonesia Menggunakan Algoritma Lexrank. Jurnal Computer Science and Information Technology (CoSciTech).
- I. Nitin and J.D. Fred (2010). Natural Language Processing Second Edition. CRC. Press.
- J. Devlin et al. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv: 1810.04805.
- Lin, C.-Y. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. Association for Computational Linguistics, 8, 74–81.

- Mustamiin, M., Ghozali, A. L., & Sifa, M. L. (2018). Peringkasan Multi-dokumen menggunakan Metode Pengelompokan berbasis Hirarki dengan Multi-level Divisive Coefficient. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 5(6).
- N.A. Ranggianto (2023). Peringkasan Teks dengan Pendekatan Abstraktif dan Ekstraktif Pada Ulasan Perjalanan Berbahasa Indonesia. Tesis di Faculty of Intelligent Electrical and Informatics Technology.
- P.I. Dewi and A. Fahri (2019). Peringkasan Ekstraktif Teks Bahasa Indonesia dengan Pendekatan Unsupervised Menggunakan Metode Clustering. *CYBERNETICS*. Vol. 3, no. 02.
- P. I Nyoman and W.U. Ni Nengah (2023). Implementasi Peringkasan Dokumen Berbahasa Indonesia Menggunakan Metode Text-to-Text Transfer Transformer (T5). *Jurnal Teknologi Informasi dan Komputer*.
- R. Colin et al. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research* 21.
- R. Kibble (2013). *Introduction to Natural Language Processing*. Goldsmith. University of London.
- R. Farah et al. (2023). Pendekatan Hybrid Pada Sistem Peringkasan Teks Artikel Berita Bahasa Inggris Menggunakan Natural Language Processing. *Telematika MKOM*.
- S. Ian (2001). *Software Engineering*. 9th Edition. Addison Wesley.
- Y.S.A.S. Agustinus and A. Edwin (2023). Pengujian Algoritma TextRank Dalam Merangkum Teks. Seminar Nasional AMIKOM Surakarta (SEMNAS).
- Yuliska., et al. (2020). Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Tek Berbahasa Indonesia, Tersedia di <https://journal.uir.ac.id/index.php/ITJRD/article/view/4688>. [Diakses pada tanggal 23 Juli 2024]