

Peringkasan *Hybrid* Teks Berita Bahasa Indonesia Berbasis *Improved TextRank* dan *Transformer*

Syadza Anggraini^{1,*}, Ali Impron², Linda Sutriani³, Kurniawan Tri Putra⁴

Program Studi Informatika, Universitas Muhammadiyah Sampit, Indonesia

¹anggrainisyadza@gmail.com; ²ali.impron@gmail.com; ³lindasutriani19@gmail.com; ⁴kurniawan_tri_putra@umsa.ac.id

*penulis korespondensi

INFO ARTIKEL

Sejarah Artikel

Diterima: 30 Oktober 2025
Direvisi: 23 Desember 2025
Diterbitkan: 31 Desember 2025

Kata Kunci

Improved TextRank
mT5-small
Peringkasan *Hybrid*
Peringkasan Teks
Sentence Embedding

ABSTRAK

Peringkasan teks otomatis berbahasa Indonesia masih menghadapi tantangan dalam menghasilkan ringkasan yang informatif namun tetap koheren secara semantik. Sebagian besar penelitian sebelumnya hanya menggunakan metode ekstraktif seperti *TextRank* atau metode abstraktif seperti *mT5-small* tanpa mengoptimalkan hubungan semantik antar kalimat. Terdapat masalah di antaranya metode ekstraktif cenderung kaku dan tidak mengubah susunan kata dalam kalimat, sedangkan metode abstraktif bisa menyebabkan risiko kesalahan fakta ataupun *output* yang kurang relevan jika teks terlalu panjang. Untuk mengatasi masalah tersebut, penelitian ini mengusulkan metode peringkasan teks *hybrid* yang menggabungkan *Improved TextRank* dengan *mT5-small*. Pada tahap awal, dilakukan *praproses* dan ekstraksi kalimat dengan representasi semantik berbasis *embedding*. Hasil ekstraksi dimasukkan sebagai input di model *mT5-small* untuk menghasilkan ringkasan secara *abstractive* melalui proses parafrasa dan penyusunan ulang kalimat. Penelitian dilakukan terhadap 1000 dokumen berita dataset IndoSum dengan metrik evaluasi ROUGE. Hasil evaluasi menunjukkan bahwa metode usulan mencapai nilai ROUGE sebesar 0.687, 0.451, dan 0.634, melampaui performa *TextRank* klasik 0.472, 0.307, 0.441 dan *mT5-Small* 0.553, 0.362, 0.508 untuk hasil evaluasi ROUGE 1, 2 dan L secara berturut-turut. Hasil ini membuktikan bahwa integrasi *sentence embedding* dan pendekatan *hybrid* efektif meningkatkan kualitas ringkasan dari segi relevansi semantik. Sehingga pendekatan ini berpotensi menjadi dasar pengembangan model peringkasan teks Bahasa Indonesia yang lebih *robust* dan semantik.

PENDAHULUAN

Dalam era digital yang ditandai oleh ledakan informasi, jumlah berita yang dipublikasikan secara daring meningkat secara eksponensial setiap harinya sekitar 400-500 artikel berita per situs (<https://detik.com>) dengan jumlah lebih dari 43.000 situs (<https://nasional.kompas.com>) di Indonesia. Pembaca dihadapkan pada volume teks yang sangat besar sehingga sulit untuk menyaring informasi penting secara cepat dan efisien [1][2]. Selain itu, karakteristik berita daring yang terus diperbarui secara *real-time* menyebabkan informasi menjadi berlimpah namun terfragmentasi, sehingga pengguna kesulitan memperoleh pemahaman utuh terhadap suatu topik. Kondisi ini menimbulkan tantangan dalam peringkasan teks, khususnya dalam menjaga koherensi dan relevansi semantik ketika merangkum informasi dari berbagai sumber yang saling tumpang tindih [3][4]. Lebih lanjut, hal tersebut mendorong kebutuhan akan peringkasan teks otomatis sebagai solusi untuk menghasilkan ringkasan singkat dari dokumen berita tanpa menghilangkan makna utama [5].

Peringkasan otomatis dapat mempercepat proses pemahaman informasi sekaligus mendukung sistem temu kembali informasi dan *question answering*. Selain itu, analisis opini publik (*sentiment analysis*) yang membutuhkan representasi ringkas dari dokumen teks dan analisis berita dalam skala besar [4]. Selain itu, dalam konteks analisis berita, ringkasan otomatis memudahkan proses klasifikasi topik, deteksi isu *trending*, serta analisis perbandingan antar-sumber berita [6]. Namun demikian, efektifitas peringkasan teks bergantung pada pendekatan yang digunakan dalam menghadapi tantangan kelengkapan informasi dan koherensi semantik.

Peringkasan dokumen telah dikembangkan dari berbagai pendekatan, mulai dari metode ekstraktif yang memilih potongan kalimat paling informatif dari dokumen sumber tanpa menghasilkan teks baru, hingga metode abstraktif yaitu memungkinkan pembuatan kalimat baru, merangkum konsep secara lebih natural [7][8]. Peringkasan berbasis ekstraktif di antaranya menggunakan *Maximum Marginal Relevance* (MMR) yang memanfaatkan *Cosine Similarity* untuk menyeimbangkan relevansi dan redundansi antar kalimat, sehingga ringkasan yang dihasilkan lebih ringkas namun tetap informatif [9] dan gabungan metode dalam pencarian kata kunci dengan *Latent Dirichlet Allocation* (LDA) yang digunakan untuk memodelkan topik dokumen, sehingga pemilihan kalimat didasarkan pada representasi topik utama dalam teks berbahasa Indonesia [10]. Adapun peringkasan multi-dokumen berita berdasarkan relevansi kalimat menggunakan fitur *POS Tagging* untuk mengidentifikasi struktur kalimat yang dianggap lebih informatif dan relevan [11], serta penggunaan struktur graf dalam algoritma *Ant* [12] dan algoritma *Lexrank* yang memanfaatkan *centrality* dalam graf kalimat dipadukan dengan YAKE [13] untuk memperkuat aspek kata kunci penting dalam menghasilkan ringkasan ekstraktif berita Bahasa Indonesia.

Dalam peringkasan abstraktif, sejumlah penelitian terkini di antaranya menggunakan algoritma berbasis *transformer* yang memungkinkan pemodelan konteks global dan pembelajaran perhatian (*attention*) lintas keseluruhan input yang lebih koheren dan kontekstual dibandingkan metode ekstraktif tradisional [14]. Model-model seperti T5 [15], *Bart* [16], mT5 [15], LSTM [17] dan GRU [18] menjadikan peringkasan sebagai tugas *text-to-text*, sehingga memungkinkan ringkasan menghasilkan kalimat baru yang tidak hanya memilih kembali kalimat dari sumber, tetapi juga menyusun ulang informasi secara manusiawi. Selain itu, Miracle Aurelia, dkk [19] menegaskan bahwa penerapan model *abstractive summarization* tidak hanya memperpendek teks, tetapi juga dapat meningkatkan aksesibilitas informasi, terutama pada sektor analitik media, jurnalis *big data*, dan sistem rekomendasi berita. Namun, terdapat sejumlah tantangan di antaranya, metode ekstraktif cenderung kaku, tidak mengubah susunan kata dalam kalimat, berisi kalimat panjang yang berulang, hingga antar kalimat yang tidak koheren karena dipilih tanpa konteks dalam menghasilkan hasil ringkasan [20]. Selain itu, dalam metode abstraktif bisa menyebabkan risiko kesalahan fakta, output yang kurang relevan jika teks terlalu panjang, kehilangan detail spesifik hingga kompleksitas komputasi yang tinggi karena input yang besar [21][22].

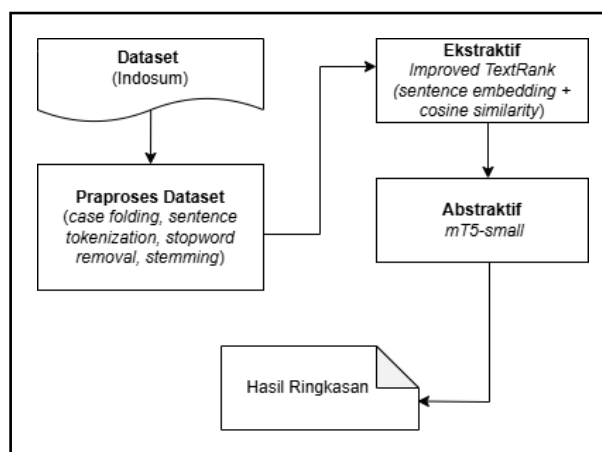
Berdasarkan kelemahan pendekatan tersebut, penelitian ini mengusulkan metode *hybrid* sebagai kontribusi yang menggabungkan *TextRank* berbasis *embedding* sebagai tahap ekstraktif untuk menyeleksi kalimat penting dari dokumen berita, dan *mT5-small* sebagai tahap abstraktif untuk menyusun ringkasan baru yang lebih padat dan koheren. Tidak hanya itu, metode *TextRank* dilakukan peningkatan dengan mengganti perhitungan frekuensi kata (TF-IDF) dengan *Sentence Embedding*. Hal ini dikarenakan, pada *TextRank* klasik berbasis *bag of words* (frekuensi kata) yang tidak bisa mengenali sinonim dan konteks serta berdasarkan kesamaan kata. Sedangkan, pada *sentence embedding* dapat mengenali hubungan semantik dan sintaksis yang berdasarkan kesamaan makna serta, proses ekstraksi kalimat dapat menghasilkan hasil yang lebih koheren antar kalimatnya. Lebih lanjut, dataset

menggunakan berita Bahasa Indonesia Indosum dengan batasan penelitian sebanyak 1000 subset yang akan dievaluasi dengan metrik ROUGE.

Oleh karena itu, penelitian ini merumuskan masalah bagaimana mengimplementasikan dan hasil evaluasi terhadap metode peringkasan *hybrid improved TextRank* dan *mT5-small* pada dataset berita Bahasa Indonesia. Selain itu, tujuan penelitian ini diharapkan mampu menghasilkan ringkasan berita berbahasa Indonesia yang tidak hanya ringkas, tetapi juga relevan secara semantik. Lebih lanjut manfaat secara teoretis dapat memberikan kontribusi terhadap pengembangan peringkasan teks menggunakan integrasi metode ekstraktif berbasis embedding dan abstraktif di mana representasi semantik embedding mampu mengatasi keterbatasan model berbasis *bag-of-words* dalam menangkap hubungan makna antar kalimat. Secara praktis, metode dapat dijadikan acuan pengembangan sistem peringkasan otomatis pada domain lain berbahasa Indonesia.

METODE

Penelitian dirancang menggunakan metode *hybrid* yang menggabungkan keunggulan metode ekstraktif dan abstraktif dengan algoritma *Improved TextRank* dan *mT5-small* berbasis *Transformer*, sehingga hasil ringkasan tidak hanya merepresentasikan informasi penting dari teks sumber, tetapi juga memiliki struktur kalimat yang lebih alami dan koheren secara semantik. *Improved TextRank* digunakan pada tahap awal untuk melakukan seleksi kalimat atau frasa paling relevan serta, untuk meningkatkan akurasi dan relevansi hasil ringkasan, dilakukan perbaikan pada tahap ekstraksi dengan menambahkan komponen *sentence embedding* berbasis *cosine similarity*, yang berfungsi sebagai penyaring semantik (*semantic-enhanced pre-filtering*). Hasil seleksi tersebut kemudian menjadi masukan (*input refined corpus*) bagi model *mT5-small*, yang berperan dalam menghasilkan ringkasan akhir secara *abstraktif* dengan mempertimbangkan konteks global antar kalimat. Selain itu, *mT5-small* terbukti efektif pada bahasa *low-resource* seperti Bahasa Indonesia, dengan kemampuan generalisasi lintas domain yang lebih baik dibandingkan pendekatan tunggal [23][24].



Gambar 1. Bagan Tahapan Penelitian

Tahapan Penelitian

Tahap eksperimen penelitian dirancang melalui 3 (tiga) tahap utama yang ditunjukkan pada Gambar 1, yaitu: (1) Praproses dataset, (2) peringkasan ekstraktif dengan *Improved TextRank* dalam pembentukan representasi kalimat menggunakan *sentence embedding* berbasis *cosine similarity* dan pemeringkatan kalimat, serta (3) penyusunan ringkasan akhir

menggunakan model *mT5-small*. Praproses merupakan tahapan persiapan dataset set sebelum diolah lebih lanjut. Kemudian, tahap peringkasan ekstraktif dataset hasil praproses dilakukan peringkasan menggunakan *Improved TextRank*. Lebih lanjut, hasil peringkasan ekstraktif menjadi input dalam peringkasan abstraktif menggunakan *mT5-small*.

Dataset

Penelitian menggunakan dataset Indosum [25] yang merupakan *benchmark* untuk peringkasan teks berita Bahasa Indonesia. Dataset Indosum berisi artikel berita dari sejumlah portal berita *online* (Detik, Kompas, Tempo, Antara, Tribun, Liputan6) dengan rata-rata panjang 350-450 kata per artikel dari berbagai kategori dan distribusi per kategori di antaranya olahraga 4768 artikel, teknologi 2303 artikel, berita utama 7192, *showbiz* 2578 artikel, hiburan 1803 artikel dan inspirasi 130 artikel. Namun, penelitian ini menggunakan subset 1000 dataset yang dipilih secara acak dan representatif, di mana terbagi ke dalam 800 data latih, 100 data validasi dan 100 data tes. Pemilihan subset ini bertujuan untuk memfokuskan evaluasi pada efektivitas metode peringkasan *hybrid* yang diusulkan, bukan pada pengaruh skala data. Selain itu, penggunaan subset efisien secara komputasi dan merepresentasikan kondisi *low-resource* yang biasa ada pada penelitian NLP berbahasa Indonesia.

Praproses Data

Tahap yang pertama adalah praproses dataset yang terdiri dari *case folding* yang merubah dan menyeragamkan semua huruf dan teks menjadi kecil dan mengurangi variasi kata. Kemudian, *sentence tokenization* yang merupakan pemecahan teks berita menjadi unit kalimat, dan unit kalimat menjadi kata (*token*) dengan menggunakan *library Natural Language Toolkit* (NLTK) [26] dan *spaCy*. Tidak hanya itu, praproses dataset juga dilakukan *stopword removal* yang merupakan penghapusan kata-kata umum atau tidak penting seperti kata penghubung “dan”, “atau”, “yang”, agar tidak mempengaruhi perhitungan bobot semantik. Serta, *stemming* yang merubah suatu kata ke dalam bentuk dasar. Tahap praproses ini dilakukan untuk mengurangi *noise* dan memastikan setiap kalimat memiliki representasi yang konsisten untuk proses *embedding* berikutnya.

Peringkasan Ekstraktif (*Improved TextRank*)

Tahap peringkasan ekstraktif menggunakan *Improved TextRank*. Adapun proses pemilihan representasi kalimat dengan *sentence embedding* yang menggunakan representasi semantik kalimat berbasis *embedding* vektor berdimensi tinggi. Proses *embedding* dilakukan menggunakan model *Sentence-BERT* (SBERT) [27][28] yang mampu menangkap makna semantik antar kalimat dalam Bahasa Indonesia. Setiap kalimat dalam teks diubah menjadi vektor berdimensi yang merepresentasikan makna keseluruhan kalimat. Representasi ini memungkinkan sistem mengenali kalimat yang semantik serupa meskipun tidak identik secara kata. Selanjutnya, bobot keterkaitan antar kalimat dihitung menggunakan *cosine similarity* yang ditunjukkan pada Persamaan (1).

$$\text{CosineSimilarity}(V_j, V_i) = \frac{V_j \cdot V_i}{\|V_j\| \|V_i\|} \quad (1)$$

di mana V_j dan V_i adalah vektor *embedding* dari dua kalimat berbeda. Nilai *cosine similarity* berada pada rentang antara 0 hingga 1. Semakin mendekati nilai 1 maka kemiripan semantik lebih kuat. Matriks kesamaan hasil perhitungan ini kemudian menjadi bobot antar *node* dalam graf *TextRank*, menggantikan perhitungan berbasis *co-occurrence* tradisional. Dengan demikian, setiap kalimat tidak hanya diperingkat berdasarkan kedekatan kata, tetapi

berdasarkan *kedekatan makna*. Setelah didapatkan representasi kalimat menggunakan *sentence embedding*, selanjutnya dilakukan pemeringkatan kalimat dengan membentuk graf di mana setiap kalimat menjadi simpul (*node*), dan setiap hubungan antar kalimat direpresentasikan oleh bobot cosine *similarity*. Nilai pentingnya setiap kalimat dihitung menggunakan algoritma *PageRank* dengan rumus iteratif yang ditunjukkan pada Persamaan (2). Kalimat dengan nilai skor tertinggi dipilih sebagai *candidate summary*, yang kemudian menjadi input bagi model abstraktif *mT5-small*.

$$S(V_i) = (1 - d) + d \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} S(V_j) \quad (2)$$

di mana, $S(V_i)$ = skor kepentingan kalimat ke-i, d = faktor peredaman (0.85), w_{ji} = bobot *cosine similarity* antar kalimat V_j dan V_i .

Peringkasan Abstraktif (*mT5-small*)

Tahap abstraktif adalah penyusunan ringkasan akhir menggunakan *mT5-small*. Model *mT5-small* merupakan varian *multilingual Text-to-Text Transfer Transformer* dengan arsitektur *Transformer Encoder-Decoder*. *Encoder* mengubah teks masukan menjadi representasi vektor kontekstual melalui mekanisme *self-attention* serta, memahami hubungan semantik antar kata dan antar kalimat dalam konteks global. Kemudian, *decoder* menghasilkan token teks keluaran secara autoregresif, yaitu satu token demi satu dengan memperhatikan hasil keluaran sebelumnya yang menggunakan dua mekanisme atensi yaitu: *Masked Self Attention* dan *Cross Attention*. Selain itu, bagian dari arsitektur *mT5-small* adalah *Feed-Forward Network* (FFN) yang berfungsi memperdalam representasi dan non-linearitas antar lapisan, serta *softmax layer* berfungsi mengonversi hasil keluaran menjadi distribusi probabilitas kata yang paling mungkin muncul berikutnya. Selanjutnya, mekanisme *self-attention* tersebut bekerja dengan menghitung bobot perhatian antar token yang ditunjukkan pada Persamaan (3).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

di mana Q, K, V merupakan matriks *Query*, *Key*, dan *Value* hasil proyeksi dari *input embedding*. Sedangkan d_k merupakan dimensi vektor *Key*, dan *softmax* digunakan untuk memastikan bobot perhatian berada pada rentang 0 sampai 1. Nilai *attention* menentukan seberapa besar perhatian token tertentu terhadap token lainnya dalam satu konteks. Dengan demikian, *mT5-small* mampu menghasilkan ringkasan yang memperhatikan keterkaitan semantik antar kalimat, bukan sekadar urutan kata. Penelitian ini menggunakan konfigurasi model *mT5-small* yang disesuaikan dengan ukuran subset dataset dan keterbatasan komputasi. *Learning rate* sebesar $2e-5$ dipilih untuk menjaga stabilitas *fine-tuning* pada model pra-latih multibahasa. Lebih lanjut, jumlah *epoch* dibatasi hingga 5 (lima) untuk mencegah *overfitting* pada subset berukuran kecil, serta *batch size* kecil digunakan agar eksperimen dapat dijalankan secara efisien pada GPU gratis *Google Colab*.

Evaluasi Kinerja

Proses evaluasi menggunakan 3 (tiga) metrik ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) [29] yaitu ROUGE-1, ROUGE-2, dan ROUGE-L, untuk mengukur kesesuaian isi ringkasan yang dihasilkan terhadap ringkasan referensi. ROUGE-1 mengukur kesesuaian unigram antara ringkasan sistem dan ringkasan referensi, sehingga merefleksikan tingkat ketercakupannya informasi utama (*content coverage*). Sedangkan ROUGE-2, mengevaluasi kesamaan bigram terhadap keterhubungan antar kata dan struktur kalimat, sehingga relevan untuk menilai koherensi semantik. Serta, ROUGE-L, metrik berbasis *Longest Common Subsequence* mengukur kesesuaian urutan kata secara global, yang mencerminkan ringkasan secara keseluruhan. Ketiga metrik evaluasi tersebut memberikan evaluasi yang komprehensif terhadap ringkasan, dari sisi kelengkapan informasi, koherensi, dan keselarasan struktur. Selain itu, untuk menilai efektivitas metode peringkasan *hybrid* yang menggabungkan pendekatan ekstraktif dan abstraktif.

HASIL DAN PEMBAHASAN

Metode usulan penelitian menggunakan 3 (tiga) metrik evaluasi ROUGE yaitu: ROUGE-1, ROUGE-2, dan ROUGE-L. Metrik evaluasi tersebut dilakukan uji pada empat model termasuk metode usulan terhadap subset sebanyak 1.000 data untuk proses pelatihan dan pengujian. Hasil evaluasi terhadap empat model peringkasan teks pada dataset IndoSum menunjukkan adanya variasi performa yang cukup signifikan meningkat terutama pada metode usulan. Hasil ROUGE terhadap empat model ditunjukkan pada Tabel 1.

Tabel 1. Hasil Evaluasi ROUGE

No	Model	ROUGE-1	ROUGE-2	ROUGE-L
1	<i>TextRank</i> klasik	0.472	0.307	0.441
2	<i>mT5-small</i>	0.553	0.362	0.508
3	<i>Improved TextRank</i>	0.612	0.403	0.561
4	<i>Hybrid: Improved TextRank + mT5-small</i> (metode usulan)	0.687	0.451	0.634

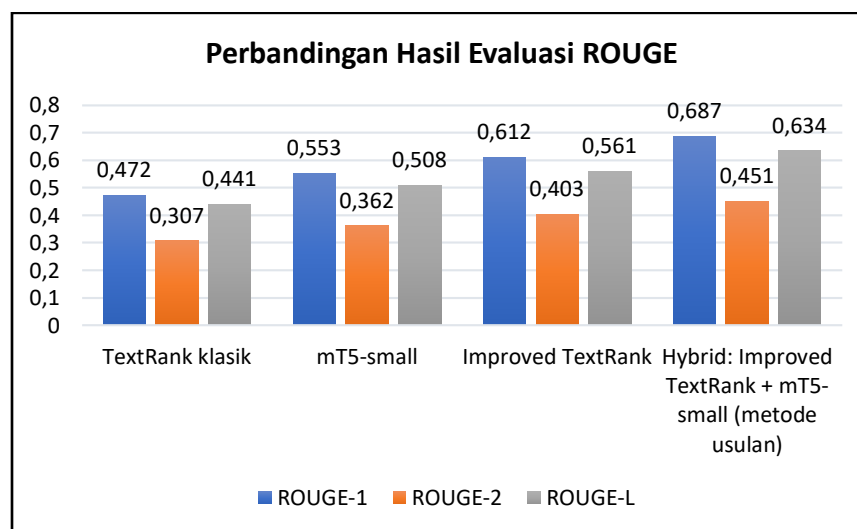
Model *TextRank* klasik menghasilkan ringkasan dengan tingkat kesesuaian leksikal yang cukup baik namun belum optimal dalam koherensi antar kalimat. Nilai ROUGE-1 sebesar 0.472 menunjukkan bahwa sebagian besar kata penting dari teks sumber masih berhasil dipertahankan dalam hasil ringkasan. Namun, skor ROUGE-2 sebesar 0.307 dan ROUGE-L sebesar 0.441 yang relatif lebih rendah memperlihatkan bahwa kesinambungan antarfrasa dan struktur kalimat masih lemah. Kelemahan ini muncul karena *TextRank* klasik mengandalkan kemiripan berbasis frekuensi kata (TF-IDF) dan tidak memperhitungkan hubungan semantik antar kalimat. Dengan kata lain, ringkasan yang dihasilkan masih cenderung panjang, kaku, dan bersifat *word-level similarity*. Akibatnya, model cenderung memilih kalimat dengan bobot kata tinggi tanpa memahami makna keseluruhan teks.

Kemudian, untuk model *mT5-small* sebagai pendekatan abstraktif menghasilkan ringkasan yang lebih alami, di mana model membangun ulang kalimat berdasarkan representasi semantik teks sumber. Model ini dilatih dengan arsitektur *encoder-decoder transformer*, di mana *encoder* memahami makna konteks penuh dari dokumen, dan *decoder* menghasilkan kalimat baru yang lebih padat dan alami. Meskipun demikian, nilai ROUGE-1 sedikit lebih tinggi 0.553 dibandingkan dengan *TextRank* klasik karena model cenderung melakukan generalisasi yang kadang menghilangkan detail spesifik dari teks sumber. Selain itu, *mT5-small* memiliki kelemahan yang terkadang melakukan parafrase berlebihan atau menghilangkan detail minor seperti waktu kejadian atau lokasi sekunder.

Selanjutnya, untuk model *Improved TextRank* berbasis *sentence embedding* mengganti perhitungan kemiripan antar kalimat menggunakan *sentence embedding* dan *cosine*

similarity dan memberikan peningkatan performa yang cukup signifikan, dengan skor ROUGE-1 naik menjadi 0.612, begitu juga dengan ROUGE-2 dan ROUGE-L. Representasi vektor dari setiap kalimat dihasilkan menggunakan model *transformer* yang mampu menangkap konteks semantik dari keseluruhan kalimat, bukan sekadar frekuensi kata. Penggunaan *embedding* semantik memungkinkan sistem memahami hubungan antar kalimat secara kontekstual, misalnya mengenali sinonim dan parafrasa. Selain itu, *embedding* memungkinkan pemilihan kalimat penting berdasarkan kedekatan makna, bukan sekadar kesamaan kata, sehingga konteks yang diberikan kepada *mT5-small* menjadi lebih semantik, ringkas, dan bebas redundansi. Kondisi ini dapat mengurangi beban model abstraktif dalam menyaring informasi relevan dari teks panjang, serta mengurangi risiko informasi yang kurang relevan. Sehingga secara kualitatif, ringkasan yang dihasilkan lebih padat dan fokus, meskipun masih bersifat ekstraktif.

Metode usulan yaitu *hybrid Improved TextRank* dan *mT5-small* menunjukkan hasil terbaik, dengan skor ROUGE-1 sebesar 0.687, ROUGE-2 sebesar 0.451, dan ROUGE-L sebesar 0.63. Secara persentase dibandingkan *TextRank* klasik, peningkatan mencapai lebih dari 45%, sementara terhadap *mT5-small* dan *Improved TextRank* masing-masing meningkat sekitar 24% dan 12% yang dapat dilihat grafik perbandingannya pada Gambar 2. Hasil ini mengindikasikan bahwa kombinasi metode ekstraktif berbasis semantik (*Improved TextRank* + *sentence embedding*) dengan model abstraktif multilingual (*mT5-small*) mampu menghasilkan ringkasan yang tidak hanya ringkas dan informatif, tetapi juga semantik dan koheren secara alami. Peningkatan performa terjadi karena proses pra-seleksi kalimat penting oleh *TextRank* berbasis kesamaan semantik (*sentence embedding* + *cosine similarity*) berhasil memfilter informasi inti sebelum diproses oleh model abstraktif.



Gambar 2. Grafik Perbandingan Hasil Evaluasi ROUGE

Secara linguistik, model *hybrid* menghasilkan struktur kalimat baru yang berbeda dari teks sumber, namun tetap mempertahankan pesan utama. Perbandingan dengan model *mT5-small* menunjukkan bahwa pra ekstraksi mampu mengurangi redundansi dan mengarahkan perhatian model generatif ke informasi inti. Hal tersebut dipengaruhi oleh teks Bahasa Indonesia yang termasuk kategori *low-resource language*, artinya jumlah data pelatihan model besar relatif terbatas dibanding bahasa seperti Inggris. Model *mT5-small* yang multibahasa dirancang untuk mampu melakukan *cross-lingual transfer*, yakni memanfaatkan representasi semantik dari bahasa lain yang memiliki sumber daya lebih besar. Namun, ketika *mT5-small* digunakan secara mandiri, model tersebut masih

menghadapi kesulitan dalam mempertahankan struktur faktual berita berbahasa Indonesia karena variasi morfologi dan gaya sintaksis. Oleh karena itu, tahap *Improved TextRank* membantu dengan menyaring kalimat berdasarkan makna semantik lokal, sistem *hybrid* mampu memberikan konteks yang lebih spesifik terhadap bahasa Indonesia.

Dari hasil perbandingan ROUGE pada Tabel 1 di atas relatif lebih rendah dibandingkan penelitian serupa dengan menggunakan dataset yang sama. Hal ini dikarenakan penelitian ini menggunakan subset yang representatif dari dataset Indosum, bukan keseluruhan data pelatihan, dengan tujuan untuk memfokuskan evaluasi pada efektifitas metode peringkasan *hybrid* yang diusulkan. Penggunaan subset memungkinkan pelaksanaan eksperimen yang lebih terkontrol, mengurangi beban komputasi, serta merepresentasi kondisi *low resource* dalam penelitian NLP berbahasa Indonesia. Selain itu, metode ekstraktif bekerja secara *unsupervised* dan metode abstraktif memanfaatkan pengetahuan linguistik lintas bahasa dari *pretraining* multibahasa yang tidak memerlukan data dalam jumlah besar untuk beradaptasi dengan domain berita Bahasa Indonesia. Penggunaan subset ini dapat menimbulkan bias dataset dan pada metode abstraktif memiliki potensi risiko *hallucination*. Namun, bias dataset tersebut tidak mempengaruhi validitas kinerja antar metode, karena seluruh metode yang dibandingkan dievaluasi menggunakan subset yang sama dan pengaturan eksperimen yang identik, serta membatasi masukan (*input*) pada model abstraktif untuk mengatasi *hallucination* agar tetap relevan secara semantik dan berlandaskan fakta.

Dengan demikian, eksperimen subset dataset tidak menyampingkan fakta bahwa metode usulan memiliki nilai ROUGE yang cukup signifikan tinggi dibandingkan model lainnya. Sehingga, dengan eksperimen dataset secara penuh ataupun *subset* yang lebih kecil tetap meningkatkan nilai ROUGE meskipun masih ada keterbatasan. Selain itu, berdasarkan penelitian ini telah menjawab rumusan masalah bahwa terjadi peningkatan hasil evaluasi ROUGE berdasarkan implementasi metode *hybrid* dengan mengintegrasikan *TextRank* berbasis *embedding* dan *mT5-small*. Tabel 2 menunjukkan contoh perbandingan contoh hasil ringkasan metode usulan terhadap metode pembanding lainnya. Hasil ringkasan tersebut mengubah panjang teks menjadi 30-46 kata dari teks asli yang memiliki panjang sebanyak 145 kata. Metode usulan memiliki ringkasan yang lebih panjang daripada metode lainnya karena memuat informasi lebih lengkap yang ditandai dengan informasi seperti “*Peristiwa ini mengingatkan pada gempa 2009 di L’Aquila yang menewaskan ratusan orang.*” yang tidak dimuat pada hasil ringkasan model lain. Hal ini menunjukkan bahwa metode lain belum mampu menangkap informasi penting dari teks asli. Selain itu, metode usulan tidak hanya meningkatkan hasil evaluasi berdasarkan metrik ROUGE tetapi juga terlihat dari perbandingan hasil ringkasan yang memuat informasi penting dan disajikan secara komprehensif. Sehingga, penelitian ini dapat dikembangkan lebih lanjut dengan mengeksplorasi penggunaan model *embedding* dan model generatif yang lebih mutakhir guna meningkatkan pemahaman konteks dan konsistensi fakta, serta memperluas evaluasi pada berbagai domain teks dan metrik berbasis semantik.

Tabel 2. Perbandingan Contoh Hasil Ringkasan

Teks Asli	Model	Hasil Ringkasan
Gempa berkekuatan 6,1 skala richter mengguncang Italia. Sepuluh orang dikabarkan tewas akibat tertimpa bangunan saat tertidur. Gempa terjadi sekira pukul 3.30 dini hari waktu setempat. Pusat Seismologi Mediterania-Eropa menyebut gempa berkekuatan 6,1 SR, sementara Badan Geologi AS 6,2 SR. Episentrum gempa terjadi di wilayah Norcia, sekitar 105	<i>TextRank</i> Klasik	Gempa 6,1 SR mengguncang Italia dan menewaskan sepuluh orang. Episentrum terjadi di Norcia, 105 mil tenggara Roma. Kota Amatrice, Accumoli, dan Norcia paling parah terdampak. Pemerintah mengirimkan bantuan dan alat berat.
	<i>mT5-small</i>	Gempa kuat 6,1 SR mengguncang wilayah tengah Italia pada dini hari, menewaskan sedikitnya sepuluh orang dan merusak kota Amatrice. Pemerintah

mil dari tenggara Roma dengan kedalaman 6 mil. Kota yang paling parah terdampak gempa adalah Amatrice, Accumoli dan Norcia. Ribuan orang berhamburan keluar rumah menyelamatkan diri. Wali Kota Amatrice Sergio Pirozzi mengatakan ada sejumlah korban yang masih terkubur puing-puing reruntuhan bangunan. Kota gelap gulita dan butuh alat berat untuk menyingkirkan reruntuhan dari jalan. "Kota ini sudah tidak ada lagi," kata Sergio. Matteo Renzi, staf Perdana Menteri Italia mengatakan pihaknya telah mengirimkan bantuan dan alat berat ke wilayah yang terpapar gempa. Ini bukan kali pertama Italia diguncang gempa. Tahun 2009, 300 orang tewas akibat gempa 6,3 SR di L'Aquila.		Italia segera mengirim bantuan dan alat berat ke lokasi terdampak.
	<i>Improved TextRank</i>	Gempa 6,1 SR mengguncang Italia tengah dan menewaskan sepuluh orang di wilayah Norcia, Amatrice, dan Accumoli. Gempa terjadi dini hari dengan pusat 105 mil tenggara Roma. Pemerintah mengirim alat berat untuk mengevakuasi korban yang tertimbun reruntuhan.
	<i>Hybrid: Improved TextRank + mT5-small</i> (metode usulan)	Gempa berkekuatan 6,1 SR mengguncang wilayah Norcia di Italia tengah pada dini hari, menewaskan sepuluh orang dan merusak kota Amatrice serta Accumoli. Pemerintah Italia mengirim bantuan dan alat berat untuk mengevakuasi korban dari reruntuhan. Peristiwa ini mengingatkan pada gempa 2009 di L'Aquila yang menewaskan ratusan orang.

KESIMPULAN

Berdasarkan hasil pengujian terhadap dataset IndoSum, metode usulan yang menggabungkan pendekatan *Improved TextRank* dengan *mT5-small* menunjukkan performa yang paling unggul dibandingkan tiga metode pembanding lainnya. Evaluasi dilakukan menggunakan metrik ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*), yang secara luas digunakan untuk menilai kualitas sistem peringkasan otomatis dengan membandingkan keluaran model terhadap ringkasan referensi manusia. Dari hasil analisis perbandingan metode usulan terhadap 3 (tiga) metode pembanding lainnya dapat disimpulkan bahwa kombinasi strategi ekstraktif dan abstraktif memberikan kontribusi teoretis signifikan terhadap peningkatan kualitas ringkasan, terutama pada aspek koherensi dan kelengkapan informasi. Selain itu, peningkatan performa metode hybrid juga mengindikasikan bahwa penggunaan *sentence embedding* sebagai representasi semantik pada tahap ekstraktif menjadi faktor utama yang memperbaiki relevansi dan konteks antar kalimat sebelum proses abstraktif dilakukan oleh *mT5-small*. Dengan kata lain, *embedding* berperan sebagai penyaring semantik yang memastikan hanya kalimat yang paling representatif yang diteruskan ke model generatif, sehingga beban sintaksis dan semantik pada *mT5-small* menjadi lebih ringan dan efisien. Selain itu, secara teknis metode *hybrid* juga terbukti lebih stabil pada *subset* data berukuran kecil, menunjukkan ketahanan model terhadap keterbatasan data pelatihan (*data efficiency*). Temuan penelitian ini memperkuat pemahaman bahwa kombinasi metode berbasis pembelajaran semantik dan generatif dapat meningkatkan kualitas peringkasan teks Bahasa Indonesia, khususnya pada konteks berita yang memiliki variasi struktur kalimat tinggi. Sebagai arah penelitian selanjutnya, pendekatan *embedding* bisa dilakukan dengan integrasi model *multilingual* generasi terbaru untuk meningkatkan pemahaman relasi antar kalimat, serta mengkaji strategi *fine-tuning* *mT5* dengan data yang lebih beragam pada domain yang lebih spesifik.

REFERENSI

- [1] C. P. R. Ardyanti, Y. Wibisono, and R. Megasari, "Peringkasan Teks Berita Berbahasa Indonesia Menggunakan LSTM dan Transformer," *Indonesian Journal of Applied Informatics*, vol. 8, no. 2, p. 122, 2024.
- [2] N. Hayatin, G. I. Marthasari, and S. Anggraini, "Improvement of cluster importance algorithm with sentence position for news summarization," *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, vol. 2018-October, pp. 483–488, 2018.
- [3] L. Shahrzadi, A. Mansouri, M. Alavi, and A. Shabani, "Causes, consequences, and strategies to deal with information overload: A scoping review," *International Journal of Information Management Data Insights*, vol. 4, no. 2, 2024.
- [4] Y. Zhang, H. Jin, D. Meng, J. Wang, and J. Tan, "A Comprehensive Survey on Process-Oriented Automatic Text Summarization with Exploration of LLM-Based Methods," *Knowledge Based System*, 2025.
- [5] A. F. Muharam, Y. A. Gerhana, D. S. Maylawati, M. A. Ramdhani, and T. K. A. Rahman, "Enhancing Abstractive Multi-Document Summarization with Bert2Bert Model for Indonesian Language," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 10, no. 1, pp. 110–121, 2025.
- [6] I. N. Mahardika, I. G. S., Arya Sasmita, G. M., & Piarsa, "Abstractive Text Summarization of Indonesian News Articles Using Long Short-Term Memory (LSTM)," *JITTER: Jurnal Ilmiah Teknologi dan Komputer*, vol. 6, no. 1, pp. 2401–2410, 2025.
- [7] M. D. B. Laksana, A. E. Karyawati, L. A. A. R. Putri, I. W. Santiyasa, N. A. Sanjaya ER, and I. G. A. G. A. Kadnyanan, "Text Summarization terhadap Berita Bahasa Indonesia menggunakan Dual Encoding," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 11, no. 2, p. 339, 2022.
- [8] K. N. Aisyah, S. Anggraini, and A. Z. Arifin, "Peringkasan Teks Multi-Dokumen Berdasarkan Metode Sentence Extraction Dan Word Sense Disambiguation," *NJCA (Nusantara Journal of Computers and Its Applications)*, vol. 4, no. 1, pp. 1–8, 2019.
- [9] Z. Idhafi, S. Agustian, F. Yanto, and N. S. H., "Marginal relevance Automatic Text Summarization in Indonesian Articles Using the Maximum Marginal Relevance Method," vol. 4, no. 3, pp. 609–618, 2023.
- [10] B. H. Mawaridi, M. Faisal, and H. Nurhayati, "Peringkasan Teks Berbahasa Indonesia dengan Latent Dirichlet Allocation dan Maximum Marginal Relevance," *Techno.Com*, vol. 23, no. 3, pp. 623–632, 2024.
- [11] M. Z. Abdullah and C. Fatichah, "Feature-based POS tagging and sentence relevance for news multi-document summarization in Bahasa Indonesia," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 1, pp. 541–549, 2022.
- [12] A. S. Girsang and F. J. Amadeus, "Extractive Text Summarization for Indonesian News Article Using Ant System Algorithm," *Journal of Advances in Information Technology*, vol. 14, no. 2, pp. 295–301, 2023.
- [13] J. Wijaya and A. S. Girsang, "Indonesian News Extractive Summarization using Lexrank and YAKE Algorithm," *Statistics, Optimization and Information Computing*, vol. 12, no. 6, pp. 1973–1983, 2024.
- [14] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers," *AI Open*, vol. 3, no. September, pp. 111–132, 2022.
- [15] M. W. Bagus Dwi Satya, A. Luthfiarta, and M. N. Althoff, "Comparative Analysis of T5 Model Performance for Indonesian Abstractive Text Summarization," *Sistemasi*, vol. 14, no. 3, p. 1092, 2025.
- [16] G. Hartawan, D. Sa'adillah Maylawati, and W. Uriawan, "Bidirectional and Auto-Regressive Transformer (BART) for Indonesian Abstractive Text Summarization," *JIP (Jurnal Informatika Polinema)*, pp. 535–542, 2024.
- [17] S. Song, H. Huang, and T. Ruan, "Abstractive text summarization using LSTM-CNN based deep learning," *Multimedia Tools and Applications*, vol. 78, no. 1, pp. 857–875, 2019.
- [18] T. Rehman, S. Das, D. K. Sanyal, and S. Chattopadhyay, "Abstractive Text Summarization Using Attentive GRU Based Encoder-Decoder BT - Applications of Artificial Intelligence and Machine Learning," 2022, pp. 687–695.
- [19] M. Aurelia, S. Monica, and A. S. Girsang, "Transformer-based abstractive indonesian text summarization," *International Journal of Informatics and Communication Technology*, vol. 13, no. 3, pp. 388–399, 2024.
- [20] N. Giarelis, C. Mastrokostas, and N. Karacapilidis, "Abstractive vs. Extractive Summarization: An Experimental Review," *Applied Sciences (Switzerland)*, vol. 13, no. 13, 2023.
- [21] V. Balachandran, H. Hajjishirzi, W. W. Cohen, and Y. Tsvetkov, "Correcting Diverse Factual Errors in Abstractive Summarization via Post-Editing and Language Model Infilling," *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022*, pp. 9818–9830, 2022.

- [22] Z. Gao, P. Li, F. Jiang, X. Chu, and Q. Zhu, "Factual Relation Discrimination for Factuality-oriented Abstractive Summarization," *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 977–986, 2023.
- [23] T. Hasan *et al.*, "XL-Sum: Large-Scale Multilingual Abstractive Summarization for 44 Languages," *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 4693–4703, 2021.
- [24] Y. Bai, Y. Gao, and H. Huang, "Cross-lingual abstractive summarization with limited parallel resources," *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, vol. 1, pp. 6910–6924, 2021.
- [25] K. Kurniawan and S. Louvan, "IndoSum: A New Benchmark Dataset for Indonesian Text Summarization," *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, pp. 215–220, 2018.
- [26] E. J. Rifano, A. C. Fauzan, A. Makhi, E. Nadya, Z. Nasikin, and F. N. Putra, "Text Summarization Menggunakan Library Natural Language Toolkit (NLTK) Berbasis Pemrograman Python," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 2, no. 1, pp. 8–17, 2020.
- [27] K. D. R. Diana and M. L. Khodra, "IndoSBERT: Enhancing Indonesian Sentence Embeddings with Siamese Networks Fine-tuning," in *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 2023, pp. 1–6.
- [28] N. Chamidah, E. Yulianti, and I. Budi, "Evaluating the Impact of Sentence Tokenization on Indonesian Automated Essay Scoring Using Pretrained Sentence Embeddings," *Revue d'Intelligence Artificielle*, vol. 37, no. 5, pp. 1101–1108, 2023.
- [29] C. Y. Lin, "Rouge: A Package for Automatic Evaluation of Summaries," *Proceedings of Workshop on Text Summarization Brances Out Association for Computational Linguistics Barcelona*, no. 1, pp. 25–26, 2004.