

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital yang ditandai dengan kecepatan arus informasi, masyarakat menghadapi tantangan dalam mengelola banyaknya informasi terutama dari *platform* berita daring. Menurut laporan Reuters Institute (2023), sekitar 88% masyarakat Indonesia mengakses berita melalui media daring, yang menunjukkan tingginya konsumsi berita digital. Namun keterbatasan waktu dan kompleksitas isi artikel membuat pembaca sering kali hanya mampu menangkap sebagian kecil dari keseluruhan informasi yang disajikan. Tantangan ini menjadi semakin besar pada berita daring berbahasa Indonesia karena adanya perbedaan struktur kalimat, variasi kosakata, serta penggunaan ungkapan (idiom) yang memerlukan pemahaman mendalam.

Tantangan ini mendorong kebutuhan akan teknologi yang mampu membantu menyaring dan menyajikan informasi secara efisien. Salah satu solusi yang ditawarkan oleh bidang *Natural Language Processing* (NLP) adalah peringkasan teks otomatis, yaitu proses menyajikan inti informasi dari sebuah dokumen dalam bentuk yang lebih ringkas. Studi Gupta *et al* (2022) membuktikan bahwa ringkasan efektif dalam pengambilan informasi penting dengan cepat dari sejumlah besar teks, sehingga mempercepat pemahaman tanpa kehilangan konteks.

Dalam proses peringkasan teks dikenal dua pendekatan utama, yaitu ekstraktif dan abstraktif. Pendekatan ekstraktif meski akurat, terbatas dalam mengungkap hubungan antarperistiwa secara menyeluruh. Selain itu, metode ini juga terbatas oleh jumlah token yang digunakan, yang dapat memengaruhi kelengkapan informasi yang disampaikan (Shukla *et al.*, 2022). Sebaliknya, pendekatan abstraktif lebih sesuai karena dapat menghasilkan ringkasan yang kontekstual, koheren, dan ringkas (Tsonkov *et al.*, 2021). Pendekatan abstraktif memungkinkan pemaduan informasi dari berbagai bagian dokumen secara lebih alami, sehingga dapat dengan mudah dipahami (Kumar *et al.*, 2024).

Seiring berkembangnya model berbasis *deep learning*, muncul berbagai model peringkasan abstraktif seperti T5, PEGASUS, dan BART. Di antara model-model tersebut, *Bidirectional and Auto-Regressive Transformer* (BART) menjadi salah satu yang paling banyak digunakan karena menggabungkan keunggulan *encoder bidirectional* untuk memahami konteks dan *decoder auto-regressive* untuk menghasilkan teks baru yang koheren (Lewis *et al.*, 2020). Studi Hartawan *et al* (2024) menunjukkan keefektifan BART dalam meringkas berita secara abstraktif dengan *score* ROUGE-L 33.85. Penelitian Deokar & Shah (2021) melaporkan BART unggul atas T5 dalam tugas peringkasan, dengan skor rata-rata ROUGE dan F1 sebesar 33%, dibandingkan T5 yang hanya mencapai 26%. BART juga menunjukkan nilai *recall* dan *precision* yang lebih baik, terutama untuk artikel berita berukuran sedang. Penelitian Dharrao *et al*(2024) menunjukkan bahwa BART unggul dibandingkan PEGASUS

dalam peringkasan berita bisnis, dengan skor ROUGE-1 sebesar 0,308 dan METEOR 0,28, sementara PEGASUS hanya mencapai ROUGE-1 0,245 dan METEOR 0,25.

Namun, meskipun model seperti BART telah dilatih secara umum pada data besar, performanya tetap perlu ditingkatkan ketika digunakan pada domain atau bahasa yang lebih spesifik. Hal ini disebabkan karena model pra-latih (*pre-trained*) cenderung menggeneralisasi bahasa berdasarkan data pelatihan awal yang sebagian besar bersifat global atau berbasis bahasa Inggris. Struktur sintaksis dan karakteristik linguistik bahasa Indonesia dapat menjadi salah satu faktor yang memengaruhi kualitas ringkasan, sehingga hasil yang diperoleh masih memiliki keterbatasan dalam aspek penyampaian makna dan kesesuaian informasi. Oleh karena itu, dibutuhkan proses *fine tuning* agar model dapat beradaptasi dengan karakteristik bahasa dan domain teks tertentu.

Selama ini, sebagian besar proses *fine tuning* dilakukan dengan menggunakan *Cross-Entropy Loss*, yaitu fungsi *loss* yang mengukur perbedaan antara prediksi dan label referensi berdasarkan kesesuaian token secara berurutan. Pendekatan ini cenderung mengoptimalkan kesesuaian sintaksis tanpa memperhitungkan kesetaraan makna antar kalimat. Secara semantik, kualitas ringkasan seharusnya dinilai dari ketercakupan informasi, relevansi konten, dan koherensi antar ide, tidak hanya dari kesamaan urutan kata. Dengan kata lain, dua ringkasan dapat memiliki makna yang identik meskipun menggunakan susunan kata yang berbeda. Ilustrasi skematisnya dapat digambarkan sebagai dua lingkaran yang merepresentasi himpunan informasi dari ringkasan referensi dan ringkasan prediksi. Pada *Cross-Entropy Loss* fokus berada

pada tumpang tindih kata, sedangkan pada penilaian berbasis semantik fokus evaluasi berada pada tumpang tindih makna atau informasi diantara kedua lingkaran tersebut.

Untuk mengatasi keterbatasan penilaian yang hanya berfokus pada kesesuaian sintaksis, dikembangkan metode *fine-tuning* berbasis *Ranking Loss* yaitu *Bringing Order to Abstractive Summarization* (BRIO) yang mempertimbangkan kualitas semantik ringkasan. BRIO memanfaatkan metrik evaluasi seperti ROUGE sebagai bagian dari fungsi *loss* dan melatih model untuk mengurutkan hasil ringkasan berdasarkan kualitas. Pendekatan ini memungkinkan model untuk lebih fokus pada kualitas ringkasan, bukan hanya kesesuaian literal terhadap referensi. Penelitian Liu *et al* (2022) menunjukkan bahwa BRIO secara signifikan meningkatkan performa model BART pada dataset CNN/DM. Selain itu, penelitian oleh Lam *et al* (2023) di Vietnam membuktikan bahwa BRIO juga berhasil meningkatkan kualitas ringkasan pada bahasa non-Inggris, seperti bahasa Vietnam, dengan skor ROUGE yang lebih tinggi dibandingkan *fine tuning* biasa.

Meskipun pendekatan BRIO telah menunjukkan hasil yang menjanjikan dalam berbagai penelitian, hingga saat ini belum ditemukan kajian yang secara khusus menerapkannya pada teks berbahasa Indonesia. Sebagian besar penelitian sebelumnya terbatas pada teks bahasa Inggris, sehingga efektivitas BRIO dalam bahasa Indonesia yang memiliki struktur, tata bahasa, dan pilihan kosakata berbeda belum pernah dievaluasi. Dalam penelitian ini, BRIO dibandingkan dengan pendekatan *fine tuning* menggunakan *Cross-Entropy Loss*, yang merupakan fungsi *loss* standar dan paling umum digunakan pada model peringkasan. Kedua pendekatan ini berada pada lingkup

teknologi yang setara, karena BRIO dibangun dengan menambahkan *ranking loss* pada mekanisme *generation loss* yang konsepnya serupa dengan *Cross-Entropy Loss*. Perbandingan ini penting untuk mengetahui sejauh mana BRIO dapat memberikan peningkatan kinerja dibandingkan metode konvensional pada konteks bahasa Indonesia. Oleh karena itu, penelitian ini mempraktikkan dan membandingkan *fine tuning* menggunakan *Ranking Loss* BRIO dengan *Cross-Entropy Loss* pada model *transformer*, untuk memberikan bukti empiris serta rekomendasi teknis bagi pengembangan sistem peringkasan otomatis yang lebih sesuai dengan karakteristik bahasa Indonesia.

1.2 Identifikasi Masalah

Berdasarkan uraian latar belakang, berikut merupakan identifikasi masalah pada penelitian ini.

1. Banyaknya informasi berita di *platform* daring menyebabkan pembaca kesulitan dalam menyerap seluruh isi berita secara efisien.
2. Metode *fine tuning* menggunakan *Cross-Entropy Loss* belum secara langsung langsung mengoptimalkan kualitas ringkasan berdasarkan metrik ROUGE.
3. Metode BRIO dengan *Ranking Loss* menawarkan peningkatan kualitas ringkasan, tetapi evaluasi terhadap metode ini pada bahasa Indonesia masih terbatas.

4. Studi yang membandingkan *Cross-Entropy* dan *Ranking Loss* BRIO pada peringkasan berita berbahasa Indonesia masih sangat terbatas, sehingga efektivitas masing-masing pendekatan belum teruji secara menyeluruh.

1.3 Batasan Masalah

Adapun batasan masalah mengenai penelitian, diantaranya:

1. Penelitian ini menggunakan model *Bidirectional and Auto-Regressive Transformer* (mBART) sebagai arsitektur utama dalam peringkasan teks secara abstraktif.
2. Metode *fine-tuning* yang dibandingkan terbatas pada dua pendekatan, yaitu *Cross-Entropy Loss* dan *Ranking-Loss* BRIO.
3. Dataset yang digunakan dalam penelitian ini adalah Liputan6 oleh Koto *et al* (2020), yaitu kumpulan berita berbahasa Indonesia yang terdiri dari sekitar 215 ribu pasang artikel dan ringkasan berita bahasa Indonesia. Dari keseluruhan dataset tersebut, penelitian ini menggunakan sebagian dataset yang berasal dari data *train*, yaitu sekitar 26 ribu pasang artikel dan ringkasan yang telah melalui tahap *preprocessing*.
4. Evaluasi performa model dilakukan dengan dua pendekatan utama, yaitu:
 - a. Evaluasi otomatis dengan metrik ROUGE (ROUGE-1, ROUGE-2, ROUGE-L) untuk mengukur kesamaan *n-gram* antara ringkasan hasil model dan ringkasan referensi serta METEOR, untuk menilai kualitas ringkasan berdasarkan aspek kesamaan semantik dan sinonim.

- b. Evaluasi manual, dilakukan dengan tiga ahli di bidang bahasa yang akan menilai kualitas ringkasan berdasarkan indikator *fluency* (kefasihan), *coherence* (koherensi), *consistency* (konsistensi), dan *relevance* (relevansi). Evaluasi ini bertujuan untuk memberikan perspektif manusia terhadap kualitas ringkasan yang dihasilkan model.
5. Penelitian ini hanya difokuskan pada teks berita umum berbahasa Indonesia, dan tidak mencakup domain teks lain seperti opini, fiksi, atau teks ilmiah.
6. Skenario peringkasan yang diterapkan dalam penelitian ini merupakan skenario yang umum digunakan dalam penelitian di bidang *Natural Language Processing* (NLP).

1.4 Rumusan Masalah

Merujuk pada latar belakang yang telah diuraikan, adapun rumusan masalah dalam penelitian ini, diantaranya:

1. Bagaimana pemodelan BART menggunakan *fine tuning Cross-Entropy* dan *Ranking-Loss* BRIO dalam peringkasan abstraktif berita bahasa Indonesia?
2. Bagaimana hasil analisis perbandingan tingkat akurasi model BART dengan dua metode *fine tuning*, yaitu *Cross-Entropy* dan *Ranking-Loss* BRIO dalam peringkasan abstraktif berita bahasa Indonesia?

1.5 Tujuan Penelitian

Adapun tujuan yang diharapkan dalam penelitian ini, diantaranya:

1. Untuk mengetahui pemodelan BART menggunakan *fine tuning Cross-Entropy* dan *Ranking-Loss* BRIO dalam peringkasan abstraktif berita bahasa Indonesia.
2. Untuk mengetahui hasil analisis perbandingan tingkat akurasi model BART dengan dua metode *fine tuning*, yaitu *Cross-Entropy* dan *Ranking-Loss* BRIO dalam peringkasan abstraktif berita bahasa Indonesia.

1.6 Manfaat Penelitian

Terdapat dua manfaat pada penelitian ini yaitu manfaat teoretis dan manfaat praktis. Berikut adalah manfaat yang diharapkan melalui hasil penelitian ini.

1.6.1 Manfaat Teoretis

1. Penelitian ini diharapkan dapat memberikan kontribusi sebagai sumber referensi bagi peneliti lain di masa mendatang.
2. Meningkatkan pemahaman mengenai *Natural Language Processing* (NLP) khususnya dalam bidang peringkasan teks secara abstraktif.
3. Memberikan pemahaman mengenai efektivitas model berbasis Transformer, khususnya BART dalam peringkasan teks abstraktif pada berita bahasa Indonesia.
4. Memberikan pemahaman mengenai pengaruh metode *fine tuning* terhadap performa model bahasa, khususnya melalui perbandingan *Cross-Entropy* dan *Ranking-Loss* BRIO.

1.6.2 Manfaat Praktis

1. Bagi masyarakat umum, penelitian ini dapat membantu dalam memahami informasi inti berita secara cepat dan efisien. Sehingga mengurangi risiko misinterpretasi pembaca akibat berita yang terlalu panjang.
2. Bagi akademisi, penelitian ini diharapkan dapat memberikan wawasan dalam memilih dan menerapkan metode *fine-tuning* yang sesuai untuk tugas peringkasan teks, serta menjadi referensi dalam pengembangan penelitian lanjutan.
3. Bagi pengembang teknologi, diharapkan hasil penelitian ini bisa dijadikan acuan dalam membangun aplikasi atau sistem yang mampu meringkas teks secara otomatis, khususnya dalam Bahasa Indonesia.

