

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi *Natural Language Processing* (NLP) menjadi salah satu fokus riset penting dalam kecerdasan buatan, terutama untuk memproses dan memahami teks secara otomatis. Peringkasan teks (*text summarization*) merupakan salah satu tugas krusial dalam NLP yang bertujuan menghasilkan ringkasan dari dokumen panjang dengan tetap mempertahankan informasi utama. Pendekatan *abstractive summarization* menggunakan model pembelajaran mendalam untuk menghasilkan kalimat ringkasan baru yang menyerupai ringkasan manusia, sehingga memiliki nilai praktis lebih tinggi dibandingkan pendekatan *extractive* yang hanya memilih kalimat penting secara langsung dari teks sumber. Penelitian tentang *abstractive summarization* terus berkembang karena kemampuannya menangani informasi berjumlah besar dengan lebih efisien dalam berbagai aplikasi digital [1].

Teknik peringkasan teks modern terbagi menjadi dua pendekatan utama, yaitu *extractive summarization* dan *abstractive summarization*. Peringkasan *abstractive* adalah pendekatan yang meringkas dengan menafsirkan dan menyusun ulang isi asli, bertujuan untuk menangkap esensi dokumen seperti yang dilakukan manusia [2]. Sedangkan pendekatan *extractive* merupakan teknik ringkasan teks otomatis yang memilih kalimat paling penting dari dokumen asli, tanpa mengubah struktur bahasa sehingga ringkasan tetap berasal langsung dari sumber teks [3]. Temuan serupa disampaikan penelitian tahun 2023, yang menegaskan bahwa pendekatan ekstraktif sering gagal menangkap hubungan semantik yang lebih dalam sehingga kualitas ringkasannya berada di bawah pendekatan abstraktif [4]. Pendekatan abstraktif sendiri banyak digunakan dalam industri media digital karena dapat menghasilkan ringkasan yang menyerupai penulisan manusia, sehingga memberi nilai tambah bagi media online serta perusahaan berita yang ingin menyajikan informasi lebih efektif kepada pembaca mereka [5]. *Abstractive summarization* terbukti lebih efektif daripada *extractive* dalam menghasilkan ringkasan yang koheren dan mudah dipahami pembaca [6].

Perkembangan *Natural Language Processing (NLP)* mengalami perubahan besar sejak hadirnya arsitektur Transformer, yang memanfaatkan mekanisme *self-attention* untuk memahami konteks secara lebih luas dan mendalam. Model berbasis Transformer seperti *BERT*, *BART*, dan *mBART* banyak diadopsi dalam penelitian *summarization* modern karena kemampuannya menangkap fitur semantik secara lebih efektif dibandingkan metode tradisional [7].

Bidirectional Encoder Representations from Transformers (BERT) adalah sebuah pendekatan *bidirectional* yang mempunyai kemampuan memahami konteks dari kedua arah (sebelum dan sesudah kata pada kalimat). *BERT* memiliki kekuatan dalam memahami makna dan struktur teks yang membuatnya efektif untuk tugas *NLP* salah satunya adalah peringkasan teks [1]. Berbeda dengan *BERT*, *Bidirectional and Auto-Regressive Transformers (BART)* dirancang sebagai model encoder-decoder yang menggabungkan pemahaman *bidirectional* dengan kemampuan generatif *auto-regressive* pada bagian *decoder*. Arsitektur ini menjadikan *BART* sangat cocok untuk tugas generatif seperti peringkasan teks abstraktif. Penelitian yang mengevaluasi *BART* pada peringkasan teks berbahasa Indonesia menunjukkan bahwa *BART* mampu menghasilkan ringkasan dengan metrik *ROUGE* yang menjanjikan, menunjukkan keunggulan pendekatan *encoder-decoder* pada tugas generasi teks dibandingkan pendekatan hanya encoder [2].

Selain model *monolingual*, perkembangan dalam *NLP* juga mencakup model multibahasa seperti *Multilingual Bidirectional and Auto-Regressive Transformers (mBART)*, yang dilatih pada banyak bahasa sekaligus. *mBART* memberikan kemampuan *transfer* lintas bahasa yang berguna untuk bahasa dengan sumber daya terbatas, termasuk Bahasa Indonesia. Model ini menunjukkan performa yang kompetitif dalam tugas *abstractive summarization* ketika dievaluasi dengan metrik standar, meskipun tantangan tetap ada terkait dataset dan kualitas ringkasan [8].

Penelitian lain juga mengevaluasi model *Transformer* untuk tugas *abstractive summarization* dalam berbagai variasi model seperti *IndoT5* dan *mBART*, menunjukkan bagaimana pendekatan *Transformer* berbeda berkontribusi pada kemampuan ringkasan teks abstraktif. Pada penelitian tersebut, model-model *Transformer* dievaluasi menggunakan metrik *ROUGE* dan *BERT-Score* untuk

menilai kesejajaran antara ringkasan yang dihasilkan dengan ringkasan referensi manusia, yang semakin memperkuat relevansi *Transformer* dalam peringkasan abstraktif terutama untuk teks berbahasa Indonesia [8].

Dalam dunia yang semakin dipenuhi oleh informasi, kemampuan untuk meringkas dokumen secara efisien dan akurat menjadi semakin krusial. Oleh karena itu, penelitian ini tidak hanya memberikan kontribusi pada pengembangan teknologi *NLP*, tetapi juga berperan dalam meningkatkan efisiensi dan efektivitas dalam pengelolaan informasi [1].

Meskipun terdapat banyak penelitian yang mengevaluasi model *Transformer* untuk peringkasan teks (termasuk *BERT*, *BART*, dan model generatif lain seperti *T5*) [9]. Tetapi, kajian yang secara komprehensif membandingkan performa model berdasarkan pendekatan arsitektur (*bidirectional encoder vs bidirectional & auto-regressive vs multilingual auto-regressive*) masih terbatas khususnya untuk peringkasan teks abstraktif berbahasa Indonesia.

Keterbatasan kajian komparatif ini menjadi tantangan bagi peneliti dan pengembang sistem *NLP* dalam menentukan model *Transformer* mana yang paling sesuai untuk tugas peringkasan teks abstraktif berbahasa Indonesia, terutama ketika mempertimbangkan performa efektifitas, efisiensi dan kecocokan dengan karakteristik bahasa Indonesia. Efektivitas model dalam meringkas dokumen diukur dari kemampuannya menghasilkan ringkasan yang mencakup informasi penting dengan akurat. Efisiensi model dilihat dari kecepatan dan penggunaan sumber daya yang minimal dalam menghasilkan ringkasan tersebut [1]. Oleh karena itu, perlu dilakukan penelitian yang membandingkan performa model *Transformer* berdasarkan perbedaan pendekatan arsitektur, seperti *bidirectional encoder*, *bidirectional and auto-regressive*, serta *multilingual bidirectional and auto-regressive*, untuk memberikan wawasan yang lebih komprehensif mengenai keunggulan dan keterbatasan masing-masing pendekatan [1].

Permasalahan yang diangkat dalam penelitian ini bersifat praktis. Pengembang sistem peringkasan berita berbahasa Indonesia dihadapkan pada banyak pilihan model *Transformer*, tetapi belum tersedia perbandingan yang dijalankan pada data dan skema evaluasi yang sama untuk model dengan arsitektur berbeda, yaitu *encoder-only* (*BERT*), *encoder-decoder* (*BART*), dan *encoder-decoder*

multibahasa (mBART). Akibatnya, pemilihan model sering didasarkan pada asumsi atau hasil penelitian pada bahasa dan dataset lain. Untuk memastikan masalah ini benar dialami calon pengguna, dilakukan survei pendahuluan melalui Google Form terhadap 84 responden yang dipilih secara purposive, yaitu pengembang (51,2%) dan mahasiswa (41,7%) yang berpotensi bersinggungan dengan pengembangan sistem peringkasan teks. Hasilnya menunjukkan bahwa 88,1% responden masih kesulitan menentukan model yang paling efektif untuk peringkasan berita berbahasa Indonesia (Gambar 4.5), dan 97,6% responden (81 dari 83 yang menjawab) menyatakan bahwa penelitian komparatif BERT, BART, dan mBART penting untuk mendukung keputusan implementasi sistem NLP (Gambar 4.6). Survei ini merupakan penjajakan kebutuhan (needs assessment) dan digunakan sebagai pendukung urgensi, bukan sebagai dasar statistik yang mewakili seluruh populasi pengembang. Rancangan survei dijelaskan pada Subbab 3.1.1 dan hasilnya diuraikan pada Subbab 4.1.1.2.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk membandingkan performa *Transformer Model Bidirectional Encoder Representations of Transformers, Bidirectional and Auto-Regressive Transformers*, serta *Multilingual Bidirectional and Auto-Regressive Transformers* dalam peringkasan teks abstraktif berbahasa Indonesia. Evaluasi dilakukan menggunakan metrik yang relevan seperti *ROUGE* dan *BERTScore* untuk menghasilkan gambaran objektif terkait performa masing-masing model dalam tugas *abstractive summarization* ini.

1.2 Rumusan Masalah

1. Bagaimana perbandingan performa BERT, BART, dan mBART dalam meringkas teks berita berbahasa Indonesia berdasarkan metrik ROUGE-1, ROUGE-2, ROUGE-L, dan BERTScore?
2. Model manakah yang menunjukkan performa paling baik pada kondisi eksperimen penelitian ini?

1.3 Tujuan Penelitian

1. Mengukur dan membandingkan performa BERT (ekstraktif), BART, dan mBART (abstraktif) pada 200 data uji berita berbahasa Indonesia menggunakan ROUGE dan BERTScore.

2. Menentukan model dengan performa paling seimbang sebagai acuan bagi pengembang dalam memilih model peringkasan berita berbahasa Indonesia.

1.4 Manfaat Penelitian

Penelitian tugas akhir ini diharapkan memberikan beberapa manfaat, antara lain:

1.4.1 Manfaat Bagi Peneliti

Penelitian ini dapat memberikan wawasan yang lebih mendalam mengenai performa model *BERT*, *BART*, dan *MBART* dalam peringkasan teks berita bahasa Indonesia. Dengan membandingkan ketiga model tersebut, peneliti dapat memahami kelebihan dan kekurangan masing-masing model dalam menghasilkan ringkasan yang relevan dan informatif. Selain itu, penelitian ini membantu peneliti mengenali faktor-faktor yang memengaruhi kualitas ringkasan otomatis, seperti ukuran dataset, metode pelatihan, dan metrik evaluasi (*ROUGE* dan *BERTScore*), sehingga dapat menjadi dasar bagi pengembangan penelitian selanjutnya di bidang *Natural Language Processing (NLP)*.

1.4.2 Manfaat Bagi Masyarakat

Bagi masyarakat, hasil penelitian ini diharapkan memberikan kemudahan bagi masyarakat dalam memperoleh informasi secara cepat dan efisien. Misalnya, pembaca berita dapat mengakses ringkasan artikel yang padat tanpa harus membaca seluruh isi berita, sehingga meningkatkan efisiensi waktu dan pemahaman informasi. Selain itu, pihak media atau jurnalis juga dapat memanfaatkan penelitian ini untuk menghasilkan ringkasan berita yang lebih cepat dan akurat, yang akan mendukung penyebaran informasi yang tepat dan terpercaya kepada publik. Sebagai gambaran kasus nyata, model *BART* yang terbukti paling unggul pada penelitian ini dapat diadaptasi menjadi fitur “ringkasan otomatis” atau “baca cepat” pada portal berita seperti CNN Indonesia, Liputan6, maupun Kompas — sumber data yang sama dengan yang digunakan dalam penelitian ini — sehingga pembaca dapat menangkap inti sebuah berita hanya dalam beberapa kalimat sebelum memutuskan untuk membaca artikel secara lengkap. Bagi redaksi media, kemampuan ini juga dapat mempercepat proses pembuatan ringkasan untuk keperluan notifikasi aplikasi, caption media sosial, maupun breaking news, tanpa mengurangi keakuratan informasi yang disampaikan kepada publik.

1.4.3 Manfaat Bagi Akademik

Penelitian ini dapat menjadi referensi yang berguna bagi pengembangan studi akademik selanjutnya dalam bidang *NLP* dan *text summarization*. Hasil penelitian ini diharapkan dapat digunakan oleh akademisi atau mahasiswa lain sebagai dasar untuk mengembangkan model peringkasan teks yang lebih efektif, baik untuk bahasa Indonesia maupun bahasa lain. Selain itu, penelitian ini juga dapat menunjang pengembangan aplikasi atau sistem peringkasan teks otomatis untuk keperluan pendidikan, penelitian, maupun industri, sehingga memperluas pemanfaatan teknologi *NLP* di berbagai bidang.

1.5 Batasan Masalah

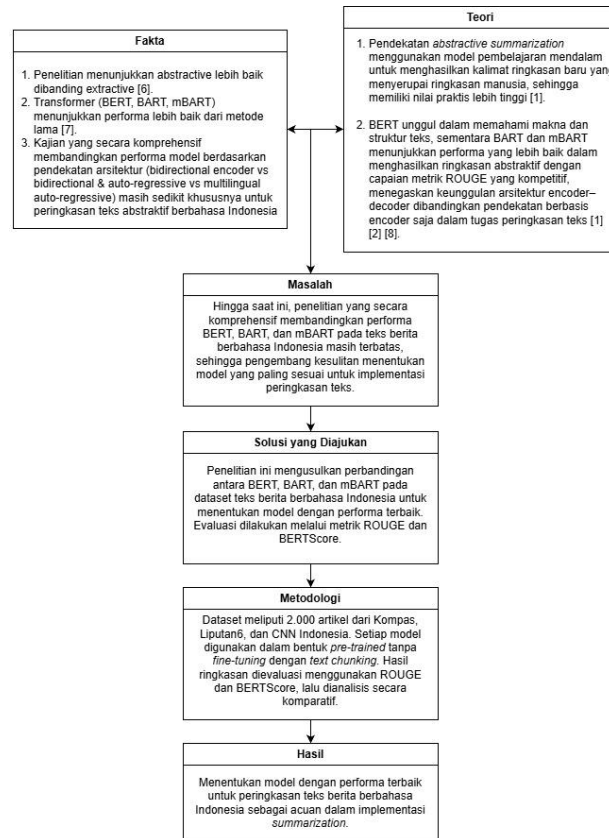
Terdapat beberapa Batasan masalah yang menjadi batas pada penelitian kali ini yaitu sebagai berikut:

1. Penelitian ini hanya menggunakan teks berita berbahasa Indonesia dari CNN Indonesia, Liputan6, Kompas sebagai dataset. Artikel dari *domain* lain, seperti jurnal ilmiah, opini, atau cerita fiksi, tidak termasuk dalam penelitian ini.
2. Penelitian ini membandingkan tiga *model pre-trained language model*, yaitu *BERT (encoder-only)*, *BART (encoder-decoder)*, dan *MBART (multilingual encoder-decoder)*. Model lain di luar ketiga ini tidak diteliti.
3. Penelitian ini fokus pada *abstract summarization*, yaitu menghasilkan ringkasan yang bersifat parafrase dan ringkas, bukan sekadar menyalin kalimat dari teks asli (*extractive summarization*).
4. Kinerja model akan dievaluasi menggunakan metrik *ROUGE* dan *BERTScore*. Metriks lain seperti *METEOR* atau *BLEU*, tidak digunakan dalam penelitian ini.
5. Penelitian ini terbatas pada analisis performa model dalam konteks ringkasan berita. Penerapan model di bidang lain atau integrasi ke aplikasi nyata berskala besar tidak menjadi fokus penelitian karena penelitian ini bertujuan hanya untuk membandingkan performa modelnya.

1.6 Kerangka Pemikiran

Kerangka pemikiran penelitian ini ditunjukkan pada Gambar 1.1. Kerangka tersebut menggambarkan hubungan antara fakta, teori, permasalahan, solusi yang

diusulkan, metodologi, hingga hasil akhir yang diharapkan. Dengan memahami alur ini, pembaca dapat melihat secara jelas bagaimana pendekatan penelitian diusulkan untuk menjawab permasalahan.



Gambar 1.1 Kerangka Pemikiran

Berdasarkan Gambar 1.1, kerangka berpikir penelitian ini disusun berdasarkan perkembangan riset peringkasan teks abstraktif yang menunjukkan kecenderungan penggunaan model pembelajaran mendalam berbasis *Transformer*. Sejumlah penelitian menunjukkan bahwa pendekatan *abstractive summarization* menghasilkan ringkasan yang lebih baik dibandingkan pendekatan *extractive* karena mampu membentuk kalimat baru yang menyerupai ringkasan manusia [6]. Selain itu, model *Transformer* seperti BERT, BART, dan mBART terbukti menunjukkan performa yang lebih baik dibandingkan metode peringkasan tradisional, khususnya dalam memahami konteks dan hubungan antar kata dalam teks panjang [7]. Meskipun demikian, kajian yang secara komprehensif membandingkan performa model berdasarkan pendekatan arsitektur, yaitu *bidirectional encoder* (BERT), *bidirectional and auto-regressive encoder-decoder*

(*BART*), serta *multilingual auto-regressive (mBART)*, masih terbatas, terutama pada konteks peringkasan teks abstraktif berbahasa Indonesia.

Fenomena tersebut kemudian dikaitkan dengan teori-teori utama dalam bidang *Natural Language Processing*. Pendekatan *abstractive summarization* memanfaatkan model pembelajaran mendalam untuk menghasilkan ringkasan baru yang lebih natural dan memiliki nilai praktis lebih tinggi dibandingkan sekadar mengekstraksi kalimat dari teks asli[1]. *BERT*, sebagai model *bidirectional encoder*, memiliki kekuatan dalam memahami makna dan struktur teks secara kontekstual sehingga efektif untuk berbagai tugas NLP, termasuk peringkasan teks[1]. Sementara itu, *BART* mengadopsi arsitektur encoder-decoder yang memungkinkan proses pemahaman dan generasi teks dilakukan secara terintegrasi, serta menunjukkan hasil evaluasi *ROUGE* yang menjanjikan pada tugas peringkasan dibandingkan pendekatan yang hanya menggunakan *encoder* [2]. Di sisi lain, *mBART* dirancang sebagai model multibahasa dan menunjukkan performa yang kompetitif dalam tugas *abstractive summarization* ketika dievaluasi menggunakan metrik standar, sehingga relevan untuk bahasa dengan sumber daya terbatas seperti bahasa Indonesia [8].

Keterkaitan antara fakta empiris dan landasan teori tersebut menegaskan adanya kesenjangan penelitian, yaitu belum adanya studi yang secara sistematis membandingkan kinerja *BERT*, *BART*, dan *mBART* berdasarkan karakteristik arsitekturnya dalam tugas peringkasan teks abstraktif berbahasa Indonesia. Oleh karena itu, penelitian ini mengusulkan pendekatan *benchmark* komparatif untuk mengevaluasi ketiga model tersebut menggunakan dataset berita berbahasa Indonesia. Evaluasi dilakukan melalui metrik *ROUGE* dan analisis kualitas ringkasan untuk menentukan model yang paling optimal. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam pemetaan performa model *Transformer* untuk *abstractive summarization* serta memberikan panduan praktis bagi pengembang sistem, media daring, dan akademisi dalam memilih model yang sesuai untuk implementasi peringkasan teks berbahasa Indonesia

1.7 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang proses penulisan yang dilakukan untuk menyelesaikan penelitian ini. Sistematika penulisan penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penelitian, perumusan masalah, tujuan penelitian, serta batasan yang ditetapkan agar penelitian tetap fokus. Selain itu, bab ini juga menjelaskan kerangka pemikiran yang menjadi dasar penelitian, metodologi yang digunakan, serta gambaran sistematika penulisan sebagai panduan dalam membaca penelitian ini.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori-teori dan konsep dasar yang relevan dengan topik penelitian. Kajian pustaka ini bertujuan untuk memberikan pemahaman mendalam serta mendukung analisis dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan langkah-langkah metodologi yang digunakan dalam penelitian, mulai dari tahap awal hingga akhir. Pembahasan mencakup analisis produk serta perencanaan eksekusi aplikasi agar penelitian berjalan secara sistematis dan terarah.

BAB IV HASIL DAN PEMBAHASAN

Bab ini terdiri dari dua hal utama, pertama pemaparan tentang temuan atau hasil penelitian berdasarkan tahapan penelitian yang dilakukan. Pemaparan hasil penelitian disesuaikan dengan urutan masalah penelitian. Kedua pembahasan hasil atau temuan penelitian untuk menjawab rumusan penelitian.

BAB V PENUTUP

Bab ini terdiri dari dua hal utama, pertama pemaparan tentang temuan atau hasil penelitian berdasarkan tahapan penelitian yang dilakukan. Pemaparan hasil

penelitian disesuaikan dengan urutan masalah penelitian. Kedua pembahasan hasil atau temuan penelitian untuk menjawab rumusan penelitian.

