BAB II
KAJIAN LITERATUR

2.1 Portal Berita Digital

Portal berita digital merupakan media yang memanfaatkan platform internet untuk menyebarkan informasi jurnalistik secara cepat, luas, dan tanpa batasan waktu bagi pengguna [8]. Kehadiran media digital telah merevolusi praktik jurnalisme melalui distribusi konten yang lebih interaktif, sehingga informasi dapat terus diperbarui sesuai perkembangan peristiwa terbaru [9].

Dalam studi ini, portal berita digital dianggap sebagai sumber data utama yang bersifat dinamis. Kecepatan penyebaran informasi di portal ini menjadi alasan utama kebutuhan akan sistem ringkasan otomatis agar pengguna dapat memperoleh informasi terkini tanpa ketinggalan oleh laju pembaruan berita.

2.1.1 Karakteristik Portal Berita Digital

Ciri-ciri portal berita digital mencakup kemampuan menyajikan konten secara langsung, penggabungan berbagai format multimedia, serta adanya interaksi timbal balik antara penyedia konten dan audiens [8]. Tidak seperti media tradisional, media digital memungkinkan keterlibatan aktif audiens dalam proses konsumsi berita yang didukung oleh infrastruktur teknologi informasi untuk penyimpanan data secara besar-besaran [9].

Karakteristik tersebut memerlukan suatu sistem yang dapat mengekstrak informasi utama di antara banyaknya elemen pendukung dalam berita digital. Sistem yang dikembangkan dalam penelitian ini berorientasi pada elemen teks untuk disederhanakan menggunakan teknik *abstractive summarization* agar interaksi pengguna menjadi lebih efektif.

2.1.2 Penyebaran Informasi Berita Secara Digital

Perkembangan teknologi internet telah mendorong perubahan pola konsumsi informasi masyarakat menuju media digital yang lebih fleksibel dan mudah dijangkau. Pengguna cenderung memilih platform digital karena memungkinkan akses informasi yang cepat, praktis, serta dapat diperoleh kapan saja sesuai kebutuhan aktivitas sehari-hari [8],[9].

Dengan tingginya jumlah informasi yang tersebar di media digital, pengguna berpotensi mengalami kesulitan dalam memilah dan memahami informasi yang penting dengan efisien. Keadaan ini dapat menambah beban berpikir pengguna ketika harus memproses informasi yang banyak dalam waktu singkat. Dengan demikian, teknologi ringkasan teks dapat digunakan sebagai solusi untuk membantu pengguna mendapatkan informasi utama dengan lebih cepat dan terorganisir melalui aplikasi berbasis web.

2.1.3 Permasalahan dalam Konsumsi Berita Digital

Kemudahan dalam mengakses informasi melalui media digital juga membawa tantangan baru berupa kelebihan informasi atau *information overload*, yaitu situasi saat individu menerima informasi dalam jumlah yang banyak sehingga kemampuan untuk menyortir, memahami, dan memproses informasi menjadi terhambat. Kondisi ini dapat menurunkan daya tangkap pengguna terhadap informasi yang diterima, khususnya pada konten berita yang memiliki struktur panjang dan kompleks. Akibatnya, pengguna cenderung membaca dengan cepat (*skimming*) sehingga dapat melewatkan informasi penting yang ada dalam berita [9].

Masalah tersebut menjadi salah satu alasan utama penerapan metode *abstractive summarization* dalam penelitian ini. Teknologi ringkasan teks memungkinkan pengguna untuk mendapatkan informasi utama secara lebih singkat tanpa perlu membaca seluruh isi berita, sehingga pemahaman informasi menjadi lebih efektif dan mudah diterima.

2.1.4 Berita Lokal

Berita lokal adalah informasi yang berisi peristiwa atau isu yang berlangsung di suatu daerah tertentu dan berkaitan langsung dengan kebutuhan serta kepentingan masyarakat setempat. Kehadiran berita lokal memiliki peran krusial dalam mendukung distribusi informasi mengenai situasi sosial, kebijakan daerah, layanan publik, serta kemajuan lingkungan sekitar yang mempengaruhi kehidupan warga [9].

Penelitian ini berfokus pada berita lokal Kota Medan karena informasi pada tingkat daerah biasanya memiliki ciri bahasa dan konteks yang lebih khusus dibandingkan berita nasional. Di samping itu, tingginya permintaan masyarakat akan informasi lokal mendorong kebutuhan akan sistem yang dapat menyajikan informasi inti dengan cara yang lebih singkat dan efisien. Sehingga, pengembangan sistem ringkasan teks pada berita lokal Medan

diharapkan mampu memudahkan pengguna dalam memahami informasi inti dengan lebih cepat tanpa harus membaca keseluruhan isi berita yang cenderung panjang.

Berdasarkan tinjauan teori mengenai portal berita digital dan karakteristik berita lokal di atas, dapat disimpulkan bahwa terdapat kesenjangan antara volume informasi yang tersedia dengan kemampuan audiens dalam mengonsumsi berita secara efektif. Permasalahan *information overload* dan perilaku *skimming* pembaca menjadi urgensi utama dilakukannya penelitian ini. Relevansi portal berita digital dalam skripsi ini adalah sebagai sumber data primer, sementara fokus pada berita lokal Medan diambil untuk mengisi celah kebutuhan informasi yang spesifik dan kontekstual bagi masyarakat setempat. Oleh karena itu, penelitian ini memposisikan perancangan *website* peringkasan teks sebagai solusi teknologi yang relevan untuk menjembatani melimpahnya arus berita lokal Medan dengan kebutuhan pembaca akan informasi yang ringkas, akurat, dan mudah dipahami.

2.2 Peringkasan Teks Otomatis (*Text Summarization*)

Peringkasan teks otomatis (*text summarization)* merupakan proses yang menghasilkan ringkasan dari suatu teks dengan tetap mempertahankan informasi penting yang ada dalam teks tersebut. Tujuan utama dari peringkasan teks adalah untuk menyajikan informasi utama yang lebih singkat dari sebuah teks sehingga dapat memudahkan pengguna dalam memahami isi teks tanpa harus membaca keseluruhan isi teks. Perkembangan teknik peringkasan teks dipengaruhi oleh kemajuan di bidang NLP dan *machine learning*. Saat ini terdapat dua pendekatan utama dalam *text summarization* yaitu :

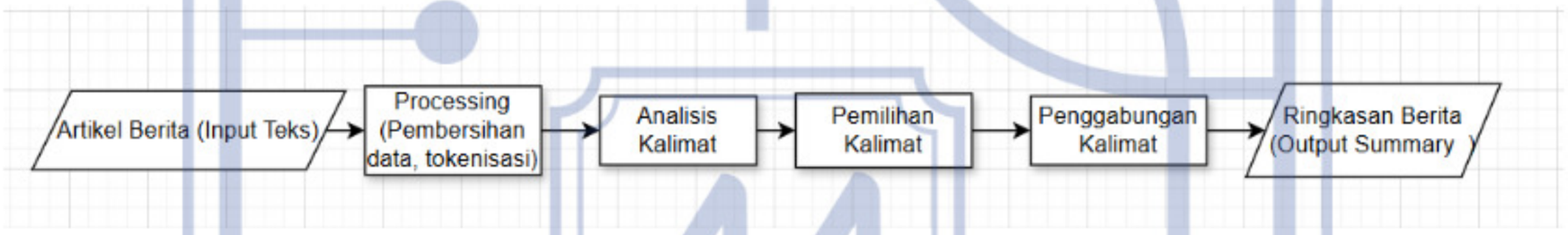
2.2.1 *Extractive Summarization*

Extractive summarization merupakan metode ringkasan teks yang dilakukan dengan memilih kalimat atau bagian penting secara langsung dari teks sumber tanpa mengubah struktur kalimat aslinya. Metode ini bekerja dengan mengidentifikasi kalimat yang memiliki nilai penting tinggi melalui analisis tertentu, lalu menyusunnya menjadi ringkasan yang komprehensif. Karena tidak melakukan penggantian kalimat atau penyusunan ulang, metode ini dapat menjaga informasi asli dokumen sehingga kemungkinan perubahan arti dapat diminimalkan [10], [11].

Proses *extractive summarization* biasanya dimulai dengan tahap pra-pemrosesan *(pre-processing)* yang mencakup pembersihan teks, normalisasi, dan tokenisasi. Tahap ini bertujuan untuk mempersiapkan data agar dapat diproses dengan lebih efisien pada tahap

selanjutnya. Kemudian, sistem menganalisis setiap kalimat guna menilai sejauh mana relevansinya dengan keseluruhan isi dokumen. Evaluasi relevansi dapat dilakukan dengan metode pembobotan seperti TF-IDF atau pendekatan berbasis pembelajaran mendalam yang dapat mendeteksi informasi penting dalam dokumen dengan lebih tepat [12], [13].

Selain itu, teknik yang berbasis graf seperti *TextRank* juga kerap digunakan dalam proses ringkasan teks. Metode ini berfungsi dengan merepresentasikan kalimat sebagai simpul dan hubungan antar kalimat sebagai tepi dalam sebuah grafik. Tingkat relevansi sebuah kalimat kemudian ditentukan oleh hubungannya dengan kalimat lain sehingga kalimat yang paling mewakili dapat dipilih sebagai elemen dari ringkasan. Metode ini populer karena mudah dan efisien, meskipun ringkasan yang dihasilkan cenderung tidak terlalu alami jika dibandingkan dengan metode *abstractive summarization* [14].



Gambar 2.1 *Flowchart Extractive Summarization*

Berdasarkan Gambar 2.1, proses *extractive summarization* dimulai dengan artikel berita sebagai masukan (*input text*) yang akan disederhanakan oleh sistem. Teks selanjutnya masuk ke tahap *pre-processing* yang mencakup pembersihan data, normalisasi, dan tokenisasi. Tahap ini bertujuan untuk menghapus karakter yang tidak relevan serta menyiapkan teks agar dapat dianalisis secara lebih efisien dalam proses selanjutnya [10], [11].

Setelah tahap *pre-processing*, sistem menganalisis kalimat untuk memahami makna dari setiap kalimat dalam dokumen. Pada fase ini, setiap kalimat dinilai tingkat relevansinya dengan menggunakan metode tertentu, baik yang berbasis pembobotan maupun yang menggunakan pendekatan deep learning. Proses pemilihan kalimat selanjutnya dilakukan dengan memilih kalimat yang mendapatkan skor tertinggi dan dianggap paling mencerminkan isi dokumen. Metode berbentuk graf seperti *TextRank* kerap digunakan pada fase ini untuk menilai relevansi kalimat berdasarkan koneksi antar kalimat dalam dokumen [12], [13], [14].

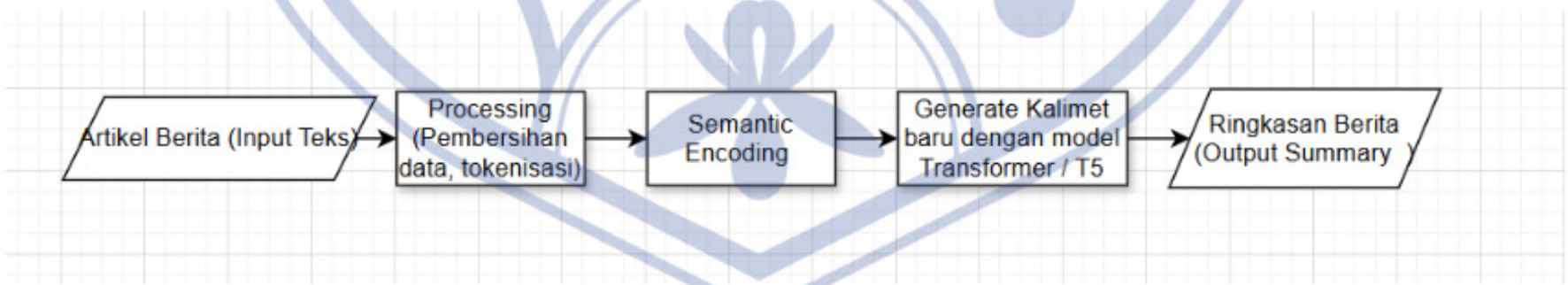
Langkah selanjutnya adalah penggabungan kalimat (*sentence extraction*), yakni merangkai kembali kalimat-kalimat yang telah dipilih menjadi sebuah ringkasan yang utuh. Pada tahap ini, susunan kalimat asli tetap dijaga tanpa perubahan sehingga informasi yang

disampaikan tetap sesuai dengan dokumen yang asli. Proses diakhiri dengan menciptakan ringkasan berita (*output summary*) sebagai hasil akhir sistem. Walaupun dapat menjaga ketepatan informasi, hasil ringkasan *extractive* sering kali terasa kurang alami dan kurang kohesif dibandingkan dengan metode *abstractive* karena hanya mengambil kalimat dari dokumen asli tanpa melakukan parafrase atau menciptakan kalimat baru [10], [15].

2.2.2 *Abstractive Summarization*

Abstractive summarization adalah metode ringkasan teks yang menghasilkan ringkasan melalui pemahaman arti dan konteks dari teks asli, lalu menyusun ulang dalam bentuk kalimat baru yang lebih singkat dan informatif. Tidak seperti *extractive summarization* yang hanya mengambil kalimat dari dokumen aslinya, metode abstraktif dapat melakukan parafrase serta menciptakan kalimat baru, sehingga hasil ringkasan lebih alami dan mirip dengan cara manusia merangkum informasi. Metode ini mengedepankan penyampaian informasi utama tanpa perlu mempertahankan susunan kalimat dari dokumen sumber yang asli [10], [11].

Metode ini biasanya menggunakan teknik *Natural Language Processing* (NLP) yang didasarkan pada *deep learning*, terutama model *sequence-to-sequence* yang terdiri dari *encoder* dan *decoder*. *Encoder* berfungsi untuk memahami dan merepresentasikan informasi dari teks input ke dalam bentuk vektor atau representasi angka, sementara *decoder* menciptakan ringkasan dalam bentuk kalimat baru berdasarkan representasi itu. Metode ini memungkinkan sistem untuk menghasilkan ringkasan yang lebih kontekstual, fleksibel, dan tidak tergantung pada urutan kalimat dari dokumen asli [12], [13].



Gambar 2.2 *Flowchart Abstractive Summarization*

Berdasarkan Gambar 2.2, proses *abstractive summarization* dimulai dengan artikel berita sebagai masukan (*input text*) yang akan disederhanakan oleh sistem. Teks selanjutnya masuk ke tahap *pre-processing* yang mencakup pembersihan data, normalisasi, dan tokenisasi. Tahap ini bertujuan untuk menghapus karakter yang tidak relevan serta menyiapkan teks agar dapat dianalisis secara lebih efisien dalam proses selanjutnya [10], [11].

Setelah proses *pre-processing*, teks kemudian menjalani tahap *semantic encoding*. Pada fase ini, model mengonversi teks menjadi representasi angka yang dapat menangkap arti, hubungan antar kata, serta konteks umum dokumen. Representasi itu menjadi landasan bagi model dalam memahami informasi signifikan yang terdapat dalam teks sebelum membuat ringkasan. Kemampuan untuk memahami konteks secara keseluruhan adalah salah satu kelebihan utama dari pendekatan abstraktif dibandingkan pendekatan ekstraktif [12], [13].

Langkah selanjutnya adalah tahap pembuatan ringkasan dengan memanfaatkan model *Transformer*, seperti *T5-base-indonesian-summarization-cased*. Pada fase ini, *decoder* menciptakan kalimat baru berdasarkan representasi yang telah dipahami oleh *encoder*. Berbeda dengan metode ekstraktif yang sekadar mengambil kalimat dari teks asli, metode abstraktif mampu menyusun informasi secara baru, melakukan parafrase, dan menciptakan kalimat yang lebih singkat serta lebih mudah dimengerti. Kemampuan ini membuat ringkasan yang dihasilkan menjadi lebih alami dan terarah.

Langkah terakhir adalah menciptakan ringkasan berita (*output summary*) sebagai hasil akhir dari proses peringkasan. Ringkasan yang dibuat diharapkan dapat menyampaikan informasi inti dari dokumen sumber dengan singkat, jelas, dan tetap menjaga makna yang ada di dalamnya. Dengan cara ini, pembaca dapat mendapatkan pokok informasi tanpa harus menelusuri seluruh dokumen.

Metode ringkasan abstraktif memiliki kelebihan dalam menciptakan ringkasan yang lebih alami dan mirip dengan cara manusia menguraikan informasi karena dapat membentuk kalimat baru berdasarkan pemahaman konteks dari dokumen. Meskipun begitu, metode ini juga mempunyai beberapa batasan, seperti kemungkinan terjadinya modifikasi atau kehilangan informasi jika model tidak dapat memahami konteks dengan benar. Di samping itu, model abstraktif umumnya memerlukan lebih banyak sumber daya komputasi dibandingkan metode ekstraktif karena melibatkan proses pemahaman dan pembuatan bahasa yang kompleks [16], [15].

Berdasarkan perbandingan antara metode *extractive summarization* dan *abstractive summarization* di atas, peneliti menyimpulkan bahwa pendekatan *abstractive* memiliki relevansi yang lebih tinggi untuk diterapkan dalam sistem peringkas berita lokal Medan. Meskipun metode *extractive* lebih mudah diimplementasikan, metode *abstractive* dipandang lebih unggul dalam menghasilkan ringkasan yang bersifat "*human-like*" karena mampu melakukan parafrase dan menyusun kalimat baru yang lebih padat. Hal ini sangat penting untuk pengolahan berita lokal Medan yang sering kali memiliki gaya bahasa jurnalistik yang

panjang dan repetitif. Dengan memahami mekanisme kedua metode ini, peneliti memfokuskan penggunaan model *abstractive* (seperti *T5-base-indonesian-summarization-cased*) guna memastikan bahwa *output* ringkasan pada *website* nantinya tidak hanya sekadar potongan kalimat dari teks asli, melainkan sebuah sari pati berita yang mengalir secara alami, koheren, dan efisien bagi pembaca di kota Medan.

2.3 *Natural Language Processing* (NLP)

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan yang menitikberatkan pada kemampuan mesin untuk mengenali, mengolah, dan menghasilkan bahasa manusia secara otomatis. NLP mengintegrasikan ilmu bahasa, kecerdasan buatan, dan pembelajaran mesin agar komputer dapat memahami arti, konteks, dan struktur bahasa yang dipakai manusia. Teknologi NLP saat ini banyak digunakan dalam berbagai aplikasi seperti *chatbot*, analisis emosi, terjemahan bahasa, sistem tanya jawab, dan ringkasan teks. Dengan menggunakan NLP, komputer dapat menganalisis susunan kalimat, memahami arti suatu teks, serta menghasilkan informasi baru yang lebih padat tanpa menghilangkan pokok informasi dari teks asli [17].

2.3.1 Tahapan dalam *Natural Language Processing*

Pengolahan Bahasa Alami (*NLP*) melibatkan berbagai langkah dalam penanganan teks agar dapat dipahami oleh komputer. Tahapan tersebut biasanya terdiri dari *preprocessing*, analisis sintaksis, dan analisis semantik. *Pre-processing* digunakan untuk membersihkan teks, melakukan normalisasi, dan memecah teks menjadi *token*, sehingga data siap untuk dianalisis pada tahapan selanjutnya. Analisis sintaksis menekankan pada susunan kalimat dan keterkaitan antar kata, sedangkan analisis semantik berfokus pada pemahaman makna teks secara lebih mendalam [18].

2.3.2 Penerapan NLP dalam Peringkasan Teks

Salah satu penggunaan utama NLP adalah peringkasan teks (*text summarization*), yang merupakan proses untuk membuat ringkasan dari sebuah dokumen sambil mempertahankan informasi penting yang ada di dalamnya. NLP memungkinkan sistem untuk mengenali informasi penting, menentukan bagian yang paling berkaitan, dan menyusun ringkasan yang lebih ringkas serta mudah dimengerti. Evolusi model bahasa

terkini telah meningkatkan mutu ringkasan otomatis, memungkinkan untuk menciptakan ringkasan yang lebih informatif dan terstruktur [10], [15].

2.3.3 NLP dalam *Abstractive* dan *Extractive Summarization*

Dalam peringkasan teks, NLP diterapkan melalui dua metode utama, yaitu *extractive summarization* dan *abstractive summarization*. *Extractive summarization* menciptakan ringkasan dengan memilih kalimat-kalimat penting secara langsung dari dokumen asli, sementara *abstractive summarization* menciptakan ringkasan dengan memahami konten teks terlebih dahulu, lalu mengatur kembali informasi tersebut ke dalam kalimat baru yang lebih ringkas. Kemajuan model *Transformer* dan model bahasa yang telah dilatih sebelumnya telah memperkuat kemampuan sistem dalam menciptakan ringkasan abstraktif yang lebih alami, konsisten, dan mendekati cara manusia merangkum informasi [12], [13].

2.3.4 Tantangan dalam NLP

Walaupun NLP telah berkembang pesat dalam beberapa tahun terakhir, masih ada berbagai rintangan dalam memahami kompleksitas bahasa manusia. Ketidakjelasan makna kata, perbedaan struktur kalimat, pemakaian bahasa nonformal, dan ketergantungan pada konteks merupakan elemen yang dapat memengaruhi mutu hasil pengolahan teks. Dalam tugas merangkum teks, tantangan utama adalah menjaga informasi penting tetap ada tanpa mengubah arti utama dari dokumen asli. Di samping itu, model yang didasarkan pada *Transformer* juga membutuhkan sumber daya komputasi yang cukup besar agar dapat memberikan kinerja yang maksimal [18], [19].

Berdasarkan tinjauan teori mengenai tahapan dan tantangan dalam *Natural Language Processing* di atas, peneliti menyimpulkan bahwa teknologi NLP merupakan fondasi operasional paling kritis dalam keberhasilan perancangan *website* ini. Relevansi utama NLP dalam penelitian ini terletak pada kemampuannya untuk mengonversi data teks berita lokal Medan yang bersifat tidak terstruktur menjadi representasi vektor yang dapat dipahami oleh mesin. Mengingat berita lokal Medan sering kali memiliki karakteristik bahasa jurnalistik yang khas dan repetitif, penerapan NLP melalui pendekatan *abstractive summarization* menjadi sangat relevan untuk mengatasi tantangan ambiguitas makna dan menjaga koherensi hasil ringkasan. Dengan demikian, pemahaman mendalam mengenai

teknik NLP ini memungkinkan peneliti untuk mengoptimalkan performa model dalam menyajikan informasi yang padat dan akurat bagi pembaca berita di kota Medan.

2.4 Arsitektur *Transformer* dan Model *T5 Base Indonesian Summarization Cased*

Dalam perkembangan teknologi *Natural Language Processing* (NLP), model yang didasarkan pada *Transformer* menjadi salah satu metode yang paling umum diterapkan dalam berbagai tugas pemrosesan bahasa alami, termasuk rangkuman teks. *Transformer* dapat menangani informasi dalam sebuah teks secara bersamaan dan lebih efektif dalam memahami relasi antar kata melalui mekanisme perhatian. Tidak seperti model tradisional seperti *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), dan *Gated Recurrent Unit* (GRU) yang memproses data satu per satu, *Transformer* mampu menangani ketergantungan jarak jauh antar kata dengan lebih efektif sehingga lebih efisien dalam mengelola dokumen yang panjang [17], [19].

2.4.1 Arsitektur *Transformer*

Model *Transformer* adalah arsitektur yang terdiri dari dua bagian utama yaitu, *encoder* dan *decoder*. *Encoder* berfungsi untuk memahami dan merepresentasikan teks yang masuk, sementara *decoder* digunakan untuk menghasilkan *output* berdasarkan representasi yang dihasilkan oleh *encoder*. Kedua komponen tersebut terdiri dari beberapa lapisan yang disusun secara bertingkat (*stacked layers*) sehingga dapat menangkap informasi dan konteks Bahasa dengan lebih mendalam [19].

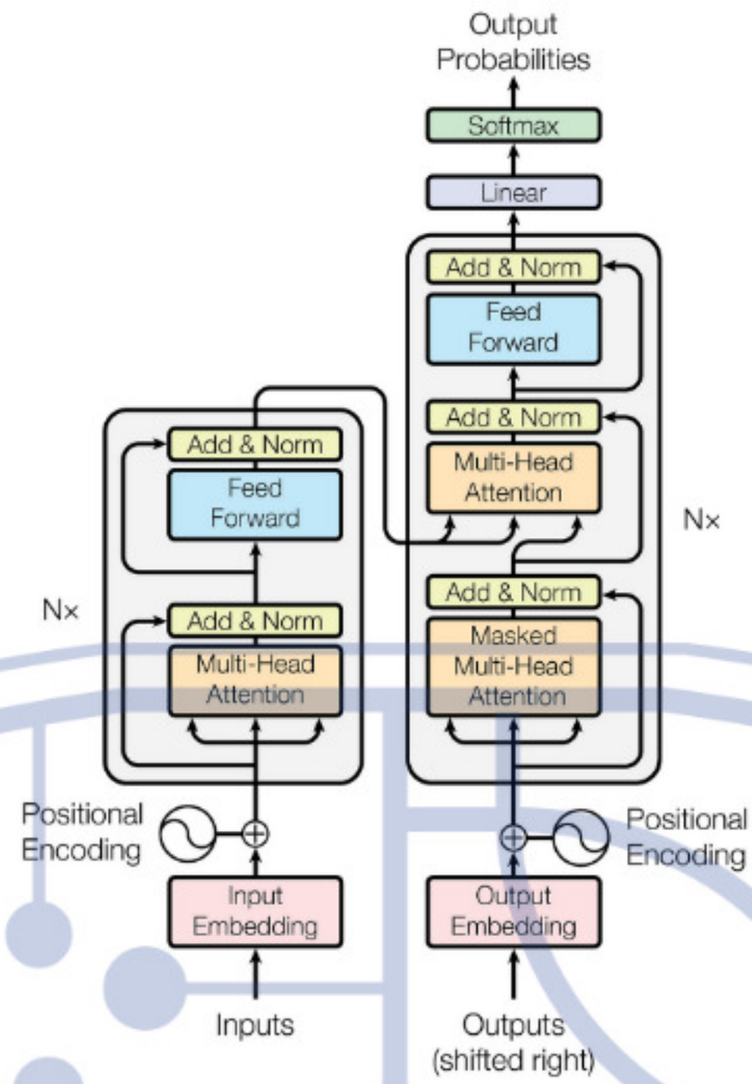


Figure 1: The Transformer - model architecture.

Gambar 2.3 Arsitektur Transformer

Gambar diatas menunjukkan arsitektur *Transformer encoder-decoder*. Pada bagian proses *encoder*, teks input akan diubah menjadi *vector* melalui *embedding* dan *positional encdoding* yang kemudian akan diproses menggunakan *self attention* untuk memahami hubungan antar kata dan menangkap konteks secara keseluruhan.

Selanjutnya, *decoder* bertugas untuk menghasilkan ringkasan secara bertahap (*token per token*). Sebagai contoh, jika diberikan *input* "*Cuaca hari ini sangat panas sehingga banyak orang memilih untuk tetap dirumah",* maka *decoder* akan mulai membentuk kalimat secara berurutan. Dalam proses ini, decoder akan menggunakan *masked-self attention* agar model hanya mempertimbangkan kalimat yang telah dihasilkan sebelumnya sehingga tidak dapat mengakses informasi di masa depan. *Decoder* kemudian akan menghasilkan *output* secara bertahap, contohnya dimulai dari kalimat "*Cuaca panas*" dan berkembang menjadi *"Cuaca panas membuat orang tetap dirumah"*. Proses ini yang membuktikan bagaimana model menyusun kembali informasi penting dari *input* yang dimasukkan menjadi kalimat baru yang lebih ringkas namun tetap mempertahankan makna utama.

Kelebihan utama arsitektur *Transformer* terletak pada kemampuannya untuk memproses data secara bersamaan melalui mekanisme perhatian. Kemampuan ini mempercepat proses pelatihan dan inferensi dibandingkan dengan model berbasis RNN dan LSTM yang memproses data dengan cara berurutan. Di samping itu, *Transformer* juga lebih

efisien dalam memahami keterkaitan antar kata yang terpisah jauh, sehingga sangat cocok digunakan untuk tugas-tugas NLP yang melibatkan dokumen panjang, seperti *abstractive summarization* [17], [19].

2.4.2 *Self-Attention Mechanism*

Salah satu komponen penting dalam arsitektur *Transformer* adalah mekanisme *self-attention*, yang memungkinkan model untuk mengenali hubungan antar kata dalam sebuah kalimat atau tulisan. Dengan mekanisme ini, setiap kata dapat memperhatikan kata-kata lain dalam urutan teks, sehingga model dapat menangkap informasi kontekstual dengan lebih efisien. Tidak seperti pendekatan sekuensial yang digunakan oleh RNN dan LSTM, *self-attention* memungkinkan pemrosesan semua kata secara simultan sehingga hubungan antara kata, termasuk yang posisinya berjauhan dalam kalimat, dapat dipahami dengan lebih baik [19].

Mekanisme *self-attention* berfungsi dengan memberikan bobot perhatian (*attention weight*) kepada setiap kata relatif terhadap kata-kata lain dalam kalimat. Bobot itu menggambarkan sejauh mana pentingnya sebuah kata dalam mendukung pemahaman makna kata yang sedang dianalisis. Semakin tinggi nilai bobot perhatian yang diberikan, semakin besar dampak kata tersebut pada representasi akhir yang dihasilkan oleh model. Melalui metode ini, *Transformer* mampu mengidentifikasi bagian mana dari kalimat yang paling penting untuk diperhatikan dalam konteks tertentu [19].

Dalam tugas peringkasan teks, *self-attention* sangat berperan penting karena memungkinkan model untuk memahami keterkaitan antara informasi yang tersebar di berbagai bagian dokumen. Kemampuan ini memfasilitasi model untuk mengenali informasi penting dan menjaga konteks saat menghasilkan ringkasan. Oleh sebab itu, mekanisme *self-attention* menjadi salah satu elemen kunci yang memungkinkan model yang didasarkan pada *Transformer*, termasuk T5 untuk menghasilkan ringkasan yang lebih tepat, koheren, dan informatif dibandingkan metode konvensional [17], [19].

2.4.3 *Encoder* dan *Decoder*

Encoder adalah bagian dari *Transformer* yang berfungsi untuk mengubah teks yang masuk menjadi representasi vektor yang padat dengan informasi kontekstual. Tahapan ini dimulai dengan mengubah setiap token menjadi bentuk numerik melalui *embedding*, setelah itu ditambahkan *positional encoding* untuk memberikan data tentang posisi setiap token dalam urutan teks. Setelah itu, representasi tersebut diolah melalui mekanisme *self-attention*

dan *feed-forward network* agar model dapat mengenali keterkaitan antara kata serta memahami konteks keseluruhan kalimat dengan baik [17], [19].

Di sisi lain, *decoder* memiliki peran untuk menghasilkan *output* berdasarkan representasi yang didapatkan dari *encoder*. *Decoder* menciptakan teks secara bertahap (*token-by-token*) dengan menggunakan mekanisme *masked self-attention*, yaitu mekanisme yang hanya memungkinkan model untuk memperhatikan *token-token* yang sudah dihasilkan sebelumnya. Di samping itu, *decoder* juga memanfaatkan *encoder-decoder attention* untuk mendapatkan informasi krusial dari representasi yang dihasilkan oleh *encoder*. Kombinasi kedua sistem ini memungkinkan model untuk menghasilkan teks yang sesuai, konsisten, dan tetap menjaga konteks dari teks yang diberikan [17], [19].

Dalam tugas merangkum teks, *encoder* berfungsi untuk memahami keseluruhan isi dokumen, sementara *decoder* bertugas untuk menyusun informasi penting tersebut menjadi ringkasan yang lebih ringkas dan mudah dimengerti. Kolaborasi antara *encoder* dan *decoder* inilah yang menjadi dasar keberhasilan model berbasis *Transformer* dalam menciptakan ringkasan abstraktif yang informatif dan alami.

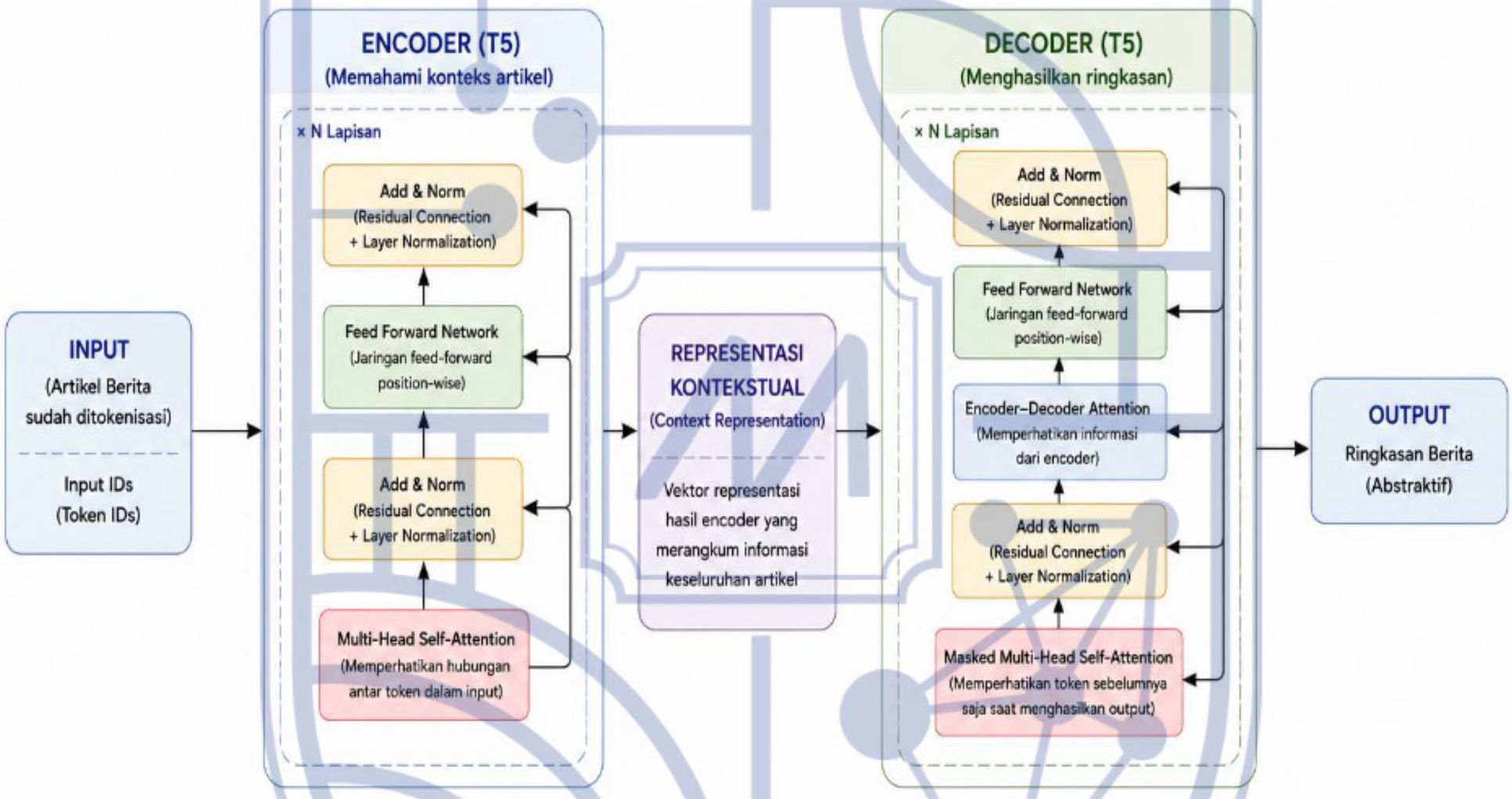
2.4.4 Model *T5 Base Indonesian Summarization Cased*

Model T5 (*Text-to-Text Transfer Transformer*) adalah model berbasis arsitektur *Transformer* yang dikembangkan oleh *Google Research* untuk membantu berbagai tugas *Natural Language Processing* (NLP) dengan menggunkaan pendekatan *text-to-text*. Pada pendekatan ini, seluruh permasalahan NLP, seperti penerjemahan Bahasa, klasifikasi teks, tanya jawab, hinga peringkasan teks, diperlakukan sebagai proses mengubah teks masukan (*input text*) menjadi teks keluaran (*output text*) [20].Pendekatan tersebut memungkinkan satu model digunakan untuk berbagai tugas tanpa memerlukan perubahan pada arsitektur utama, sehingga proses pengembangan model menjadi lebih sederhana dan fleksibel.

Tidak seperti beberapa model *Transformer* yang dibuat untuk tugas tertentu, T5 mengandalkan satu kerangka kerja yang konsisten untuk semua tugas NLP. Model hanya perlu penyesuaian *format input* dan proses *fine-tuning* sesuai dengan dataset yang digunakan. Dengan cara ini, keterampilan yang didapat selama proses *pre-training* dapat digunakan kembali untuk berbagai tugas lainnya melalui metode *transfer learning*, sehingga durasi pelatihan menjadi lebih efisien dan kebutuhan data pelatihan dapat diminimalkan [20].

Penelitian ini menggunakan model T5 yang dapat ditemukan di platform *Hugging Face* dengan nama *cahya/t5-base-indonesian-summarization-cased*. Model ini adalah penerapan T5 *Base* yang telah disesuaikan untuk tugas *abstractive summarization* menggunakan dataset berita berbahasa Indonesia. Dibandingkan dengan model T5 yang bersifat umum, model ini telah diadaptasi sesuai dengan karakteristik bahasa Indonesia sehingga dapat menghasilkan ringkasan yang lebih relevan, koheren, dan mudah dimengerti [21].

A. Arsitektur *T5 Base Indonesian Summarization Cased*



Gambar 2.4 Arsitektur *t5-base-indonesian-summarization-cased*

Model *T5-base-indonesian-summarization-cased* memanfaatkan arsitektur *encoder-decoder* yang berbasis *Transformer* dan ditujukan untuk menghasilkan ringkasan dalam bentuk abstraktif. Arsitektur ini terdiri dari dua komponen utama, yaitu *encoder* yang berfungsi memahami konteks keseluruhan berita dan *decoder* yang menghasilkan ringkasan berdasarkan informasi yang diperoleh dari *encoder*. Kedua elemen tersebut berfungsi secara berurutan sehingga model dapat menghasilkan ringkasan yang tetap menjaga informasi penting dari dokumen asli [20].

Berdasarkan Gambar 2.4, proses ringkasan dimulai dari *Input*, yaitu berita yang telah melalui tahap tokenisasi sehingga diubah menjadi *Input IDs* atau representasi numerik yang dapat dipahami oleh model. Representasi itu berfungsi sebagai *input* untuk *encoder* agar dapat memahami konteks dan keterkaitan antar token dalam berita.

Pada bagian *encoder*, setiap *token* melewati beberapa lapisan *Transformer* yang disusun berulang sebanyak *N layer*. Setiap lapisan terdiri dari *Multi-Head Self-Attention*, *Feed Forward Network*, dan *Add & Norm*. *Multi-Head Self-Attention* berfungsi untuk memahami hubungan antar *token* dalam teks agar model bisa menangkap konteks secara menyeluruh. *Output* dari mekanisme perhatian tersebut selanjutnya diproses oleh *Feed Forward Network* untuk meningkatkan representasi fitur yang telah didapatkan. Selanjutnya, *Add & Norm*, yang merupakan gabungan dari *Residual Connection* dan *Layer Normalization*, dipakai untuk menjaga kestabilan proses pelatihan sekaligus mempertahankan informasi yang telah dipelajari di setiap lapisan [20].

Setelah semua lapisan *encoder* selesai diproses, model menghasilkan Representasi Kontekstual (*Context Representation*). Representasi ini berupa vektor yang menyajikan ringkasan informasi penting dari seluruh artikel berita, sehingga dapat merepresentasikan makna dan konteks dokumen secara keseluruhan. Representasi kontekstual itu selanjutnya dikirim ke decoder sebagai landasan dalam proses pembuatan ringkasan.

Pada bagian *decoder*, representasi kontekstual dari *encoder* diproses ulang melalui beberapa lapisan *Transformer* yang juga disusun sebanyak *N layer*. Tidak seperti *encoder*, *decoder* menggunakan *Masked Multi-Head Self-Attention* untuk memastikan bahwa prediksi hanya memanfaatkan *token-token* yang sudah dihasilkan sebelumnya. Selain itu, *decoder* menggunakan *Attention Encoder-Decoder* untuk mengekstrak informasi penting dari representasi kontekstual yang dihasilkan oleh *encoder* agar ringkasan yang dibuat tetap relevan dengan isi artikel. Setelah itu, hasil pemrosesan diperkuat oleh *Feed Forward Network* dan distabilkan dengan *Add & Norm* sebelum diteruskan ke lapisan selanjutnya [20].

Setelah semua tahapan pada *decoder* selesai, model menghasilkan *Output*, yaitu ringkasan berita yang bersifat abstraktif. Ringkasan tersebut disusun dengan kalimat yang berbeda berdasarkan pemahaman model terhadap seluruh konten artikel, bukan hanya menyalin kalimat dari dokumen asli. Melalui penggunaan mekanisme perhatian (*attention mechanism*) di *encoder* dan *decoder*, *T5-base-indonesian-summarization-cased* dapat menciptakan ringkasan yang lebih koheren, informatif, dan tetap menjaga makna inti dari berita [20].

B. Proses *Pre-training* dan *Fine-tuning*

Model *Text-to-Text Transfer Transformer* (T5) diciptakan dengan metode *transfer learning*, di mana model ini terlebih dahulu menjalani tahap *pre-training* pada korpus teks yang besar sebelum disesuaikan untuk tugas tertentu lewat proses *fine-tuning* [20]. Metode ini memungkinkan model untuk mempelajari representasi bahasa secara keseluruhan, sehingga pengetahuan yang diraih selama proses pelatihan awal dapat digunakan kembali untuk berbagai tugas Pemrosesan Bahasa Alami (NLP), termasuk merangkum teks.

Pada tahap *pre-training*, model T5 dilatih dengan metode *text-to-text*, di mana seluruh data masukan dan keluaran direpresentasikan dalam format teks. Dengan metode tersebut, berbagai tugas NLP, seperti terjemahan bahasa, klasifikasi teks, tanya jawab, dan ringkasan teks, dipelajari menggunakan kerangka kerja yang serupa. Dalam proses ini, model mempelajari struktur bahasa, keterkaitan antara kata, serta konteks antar kalimat sehingga dapat membangun representasi bahasa yang umum dan dapat digunakan untuk berbagai tugas setelah penyesuaian dilakukan [20].

Setelah proses *pre-training*, model melanjutkan ke tahap *fine-tuning*, yaitu penyesuaian *parameter* model dengan menggunakan dataset yang relevan untuk tugas yang akan dikerjakan. Dalam penelitian ini, dataset yang diterapkan terdiri dari kumpulan berita yang telah di kumpulkan dari tahun 2021-2026. Selama pelatihan, parameter model disesuaikan untuk dapat menghasilkan ringkasan yang memiliki tingkat kesesuaian yang tinggi dengan ringkasan acuan [20].

Dalam penelitian ini digunakan model *cahya/t5-base-indonesian-summarization-cased*, yaitu model T5 yang telah diadaptasi untuk tugas *abstractive summarization* berbahasa Indonesia. Model ini dikembangkan melalui proses *fine-tuning* dengan dataset berita berbahasa Indonesia, sehingga dapat memahami struktur kalimat, kosakata, dan karakteristik bahasa Indonesia lebih baik dibandingkan model T5 yang masih umum [21].

Walaupun telah tersedia sebagai model pralatih (*pre-trained model*), *cahya/t5-base-indonesian-summarization-cased* masih bisa menjalani *fine-tuning* kembali dengan dataset penelitian. Dalam penelitian ini, model dilatih dengan menggunakan dataset berita lokal dari Kota Medan yang berisikan pasangan artikel berita dan ringkasan referensi. Proses ini bertujuan agar model dapat menyesuaikan *parameter* sesuai karakteristik data penelitian, sehingga menghasilkan ringkasan yang lebih relevan dan tetap menjaga informasi utama dari artikel.

Pada proses *fine-tuning*, artikel berita terlebih dahulu diproses dengan *tokenizer default* T5 untuk menghasilkan *Input IDs* yang kemudian dikirim ke *encoder*. Selanjutnya, *decoder* menciptakan ringkasan berdasarkan representasi kontekstual yang dibuat oleh *encoder*. Ringkasan yang dihasilkan oleh model dibandingkan dengan ringkasan acuan untuk menghitung nilai kerugian. Angka tersebut dijadikan acuan dalam memperbaharui *parameter* model melalui proses optimasi sampai model mencapai performa yang diinginkan.

Setelah proses *fine-tuning* rampung, model digunakan untuk membuat ringkasan pada data uji. Kualitas ringkasan selanjutnya dinilai dengan menggunakan metrik ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) dan BERTScore. ROUGE yang terdiri dari **ROUGE-1, ROUGE-2, dan ROUGE-L** digunakan untuk menilai sejauh mana kesamaan antara ringkasan yang diprediksi dengan ringkasan acuan. Sementara itu, BERTScore digunakan untuk mengukur kesamaan semantik antara kedua ringkasan berdasarkan representasi *embedding* dari model BERT sehingga mampu menilai kesamaan makna meskipun pemilihan kata dan kalimat berbeda. Penilaian tersebut dimanfaatkan untuk mengukur kemampuan model dalam menghasilkan ringkasan berita yang tepat, informatif, dan sejalan dengan konten artikel [21].

C. Model yang Digunakan dalam Penelitian

Dalam penelitian ini, model yang digunakan adalah *cahya/t5-base-indonesian-summarization-cased* yang diakses melalui pustaka *Hugging Face Transformers*. Model ini adalah implementasi T5 *Base* yang telah di *fine-tune* untuk tugas *abstractive summarization* dalam bahasa Indonesia, sehingga dapat menghasilkan ringkasan yang lebih sesuai dengan karakteristik bahasa tersebut. Pemilihan model ini dilakukan karena telah dimodifikasi untuk menghasilkan ringkasan berita dalam bahasa Indonesia, yang menjadikannya sebagai model awal (*pre-trained model*) dalam penelitian ini.

Model *cahya/t5-base-indonesian-summarization-cased* dirancang dengan menggunakan arsitektur *Transformer encoder-decoder* yang memanfaatkan mekanisme *multi-head self-attention* untuk memahami hubungan antar token dalam dokumen. Artikel berita yang menjadi *input* terlebih dahulu diproses menggunakan *tokenizer* bawaan T5 sehingga teks diubah menjadi *token-token* yang dapat dimengerti oleh model. *Token* tersebut selanjutnya diproses oleh *encoder* untuk menghasilkan representasi kontekstual dari berita, sementara

decoder secara bertahap menghasilkan ringkasan berdasarkan representasi yang dihasilkan oleh *encoder* [20].

Dalam penelitian ini, model tidak hanya berfungsi sebagai model pralatih, tetapi juga mengalami proses *fine-tuning* dengan memanfaatkan dataset berita Bahasa Indonesia yang telah disiapkan. Proses ini bertujuan agar *parameter* model mampu menyesuaikan dengan karakteristik berita lokal Kota Medan, sehingga menghasilkan ringkasan yang lebih relevan dan informatif. Setelah proses pelatihan selesai, model diterapkan untuk menghasilkan ringkasan pada data pengujian yang kemudian dievaluasi dengan menggunakan metrik ROUGE*Score* dan BERT*Score*.

D. Alasan Pemilihan *T5 Base Indonesian Summarization Cased*

Pemilihan model *cahya/t5-base-indonesian-summarization-cased* pada penelitian ini didasarkan pada beberapa pertimbangan.

Pertama, Model ini mengimplementasikan arsitektur *Text-to-Text Transfer Transformer* (T5) yang dapat menghasilkan ringkasan secara abstraktif dengan pendekatan *text-to-text*, sehingga informasi penting dalam artikel dapat disampaikan kembali dalam kalimat baru tanpa hanya menyalin konten dari dokumen sumber [20].

Kedua, model *cahya/t5-base-indonesian-summarization-cased* telah dirancang secara khusus untuk tugas *abstractive summarization* dalam bahasa Indonesia, sehingga lebih tepat digunakan pada penelitian yang memproses artikel berita berbahasa Indonesia. Kemampuan ini memungkinkan model untuk lebih memahami struktur kalimat, kosakata, dan ciri-ciri bahasa Indonesia dibandingkan dengan model yang hanya dilatih dengan korpus bahasa asing.

Ketiga, model ini memungkinkan proses *fine-tuning* agar dapat disesuaikan kembali dengan menggunakan dataset penelitian. Kemampuan itu memungkinkan model untuk memahami ciri-ciri berita lokal Kota Medan sehingga kualitas ringkasan yang dihasilkan lebih sesuai dengan kebutuhan penelitian.

Keempat, model *cahya/t5-base-indonesian-summarization-cased* dapat diakses melalui pustaka *Hugging Face Transformers*, sehingga memudahkan integrasi ke dalam sistem yang dikembangkan. Selain menawarkan dokumentasi yang menyeluruh, pustaka tersebut juga mendukung proses pelatihan dan inferensi dengan efisien, sehingga mempermudah penerapan model di aplikasi berbasis web.

2.4.5 Kelebihan *Transformer* dibanding RNN/LSTM

Transformer memiliki sejumlah kelebihan dibandingkan model sebelumnya seperti RNN, LSTM. dan GRU, antara lain:

1. Dapat mengolah data secara paralel sehingga lebih cepat
2. Dapat memahami hubungan jangka panjang antar kata
3. Lebih efektif dalam memahami konteks keseluruhan
4. Memiliki performa yang lebih baik pada berbagai tugas NLP

Meski demikian, *Transformer* juga memiliki kekurangan, yakni membutuhkan banyak sumber daya komputasi dan data pelatihan yang cukup agar dapat berfungsi dengan baik.

Integrasi model *cahya/t5-base-indonesian-summarization-cased* dalam studi ini menjadi elemen krusial karena berita lokal memiliki ciri bahasa yang beragam, penggunaan istilah unik daerah, serta penyampaian informasi yang ringkas. Dengan memanfaatkan arsitektur *encoder-decoder* yang didasarkan pada *Transformer*, model ini dapat memahami konteks keseluruhan artikel sebelum menciptakan ringkasan yang bersifat abstraktif, sehingga informasi utama dari berita tetap terjaga [20]. Selain itu, penerapan model *cahya/t5-base-indonesian-summarization-cased* yang telah dirancang untuk tugas merangkum teks berbahasa Indonesia menawarkan kemampuan yang lebih baik dalam memahami struktur kalimat serta karakteristik bahasa Indonesia. Dengan melakukan *fine-tuning* menggunakan dataset penelitian, diharapkan model dapat menghasilkan ringkasan berita lokal Kota Medan yang lebih presisi, koheren, dan sesuai dengan isi artikel, sehingga dapat membantu sistem peringkas berita dalam menyajikan informasi yang singkat dan mudah dimengerti oleh pengguna [21].

2.5 Penelitian Terdahulu

Penelitian terkait peringkasan teks otomatis (*text summarization*) telah banyak dilakukan, khususnya yang memanfaatkan pendekatan abstraktif berbasis model *deep learning* dan *Transformer*. Untuk mengindentifikasi celah penelitian (*research gap*) dan memposisikan penelitian ini diantara penelitian-penelitian sebelumnya, berikut adalah rangkuman beberapa penelitian terdahulu yang relevan:

Tabel 2.1 Rangkuman Penelitian Terdahulu

No	Peneliti & Tahun	Judul	Metode	Dataset	Hasil	Kekurangan
1	Aini Nur Khasanah & Mardhiya Hayaty (2023) [22]	ABSTRACTIVE-BASED AUTOMATIC TEXT SUMMARIZATION ON INDONESIAN NEWS USING GPT-2	*Abstractive Summarization* menggunakan model Transformer **GPT-2** (*Generative Pre-Trained 2*)	Dataset **Indosum** (berita berbahasa Indonesia) sebanyak 9.387 data	Menghasilkan ringkasan dengan rata-rata *recall* ROUGE-1: 0.61, ROUGE-2: 0.51, dan ROUGE-L: 0.57	Ringkasan sudah mampu memparafrase, tetapi masih banyak mengulang kata asli dari teks karena kurangnya data latih
2	Antonius Sakti Wiradinata & Viny Christanti Mawardi (2024) [2]	Abstractive Text Summarization Berita Bahasa Indonesia Menggunakan Retrieval-Augmented Generation	*Retrieval-Augmented Generation* (RAG) menggunakan model generatif **Pegasus**	1.000 tautan artikel berita hasil *web scraping* dari **CNN Indonesia dan CNBC Indonesia**	Model mampu menghasilkan ringkasan yang koheren dengan nilai kepresisian ROUGE-1 sebesar 0.7432 dan ROUGE-2 sebesar 0.6174	Nilai *Recall* ROUGE tergolong sangat rendah (0.0561), yang mengindikasikan banyak informasi penting tidak tertangkap sepenuhnya ke dalam ringkasan

3	Mohammad Wahyu Bagus Dwi Satya, dkk. (2025) [23]	Comparative Analysis of T5 Model Performance for Indonesian Abstractive Text Summarization	Analisis komparatif varian model **T5** (T5-Base, FLAN-T5, mT5)	Dataset **INDOSUM** (19.000 pasang artikel berita dan ringkasan)	T5-Base memberikan performa terbaik dengan ROUGE-1 73.51%, ROUGE-2 64.49%, ROUGE-L 69.54%	T5-Base memiliki kelemahan dalam redundansi kata dan butuh komputasi serta waktu latih yang lama, sedangkan FLAN-T5 sering terpotong (*truncation issues*)
4	Muhammad Furqon Fadlilah, dkk. (2024) [24]	Pemanfaatan Transformer untuk Peringkasan Teks: Studi Kasus pada Transkripsi Video Pembelajaran	Integrasi model **Whisper Turbo** (transkripsi) dan *Fine-tuning* model **T5** (peringkasan abstraktif)	Transkrip video pembelajaran dan Dataset berita **Liputan6**	Menghasilkan evaluasi F1-Score ROUGE-1 39.23, ROUGE-2 13.17, dan ROUGE-L 23.84 tanpa menggunakan praproses teks	Terdapat keterbatasan kapasitas input maksimal (*512 token*) sehingga teks sangat panjang harus dipotong menjadi beberapa segmen, berisiko

						menghilangkan konteks
5	Gaurav Sahu, dkk. (2025) [25]	A Guide To Effectively Leveraging LLMs for Low-Resource Text Summarization: Data Augmentation and Semi-supervised Approaches	Pendekatan Semi-Supervised (PPSL) & *Data Augmentation* menggunakan LLaMA-3 untuk melatih model **DistilBART CNN** (*distilbart-12-6-cnn backbone*)	Dataset TweetSumm, WikiHow, dan ArXiv/PubMed	Pengetahuan dari LLM besar berhasil didistilasi ke dalam model DistilBART (306M parameter) dengan sangat efektif dan efisien meskipun data latih minim	Masih berfokus pada teks berbahasa Inggris dan pengolahan dokumen panjang memakan waktu karena metode *chunk-and-summarize*
6	Rahma Hayuning Astuti, dkk. (2024) [26]	Indonesian News Text Summarization Using MBART Algorithm	*Abstractive Summarization* dengan proses *fine-tuning* menggunakan model pra-terlatih **mBART50**	Dataset **XL-Sum** yang difilter khusus teks berbahasa Indonesia dari BBC News (47.800 artikel)	Model memberikan performa stabil mengungguli model mT5 *base* dengan evaluasi ROUGE-1 35.94, ROUGE-2 16.43, dan	keterbatasan komputasi (pelatihan hanya dilakukan 1 *epoch*) dan sangat bergantung pada proses prapemrosesan teks

					ROUGE-L 29.91	

Berdasarkan penelitan terdahulu yang tertera pada tabel di atas, disimpulkan bahwa mayoritas penelitian di bidang peringkasan teks menekankan pada metode *abstractive summarization* dengan menggunakan model berbasis *Transformer*, seperti GPT-2, T5, mBART, dan Pegasus. Dataset yang digunakan biasanya diambil dari dataset standar, seperti IndoSum, XL-SUM, atau koleksi berita nasional dari media besar, seperti CNN Indonesia dan CNBC Indonesia. Hal ini mengindikasikan bahwa penelitian sebelumnya lebih sering menggunakan data yang bersifat umum dan berskala nasional, sementara penelitian yang secara spesifik menyoroti peringkasan berita lokal masih tergolong sedikit.

Namun, terdapat sejumlah kekurangan dalam penelitian yang bisa ditingkatkan. Sebagian besar studi sebelumnya belum membahas penerapan model *T5-base-indonesian-summarization-cased* pada sistem berbasis web yang secara otomatis mengolah berita lokal. Di samping itu, penelitian yang secara khusus menggunakan artikel berita lokal Kota Medan sebagai sumber data juga masih sangat sedikit. Beberapa penelitian sebelumnya menunjukkan adanya kesulitan dalam membuat ringkasan yang dapat mempertahankan informasi penting, memahami konteks dokumen secara keseluruhan, serta menyesuaikan hasil ringkasan dengan ciri bahasa Indonesia.

Dengan mempertimbangkan celah penelitian (*research gap*) tersebut, penelitian ini bertujuan untuk merancang sistem peringkas berita lokal di Kota Medan yang dilengkapi fitur *abstractive text summarization* dengan menggunakan model *cahya/t5-base-indonesian-summarization-cased.* Model ini dipilih karena telah dirancang untuk tugas ringkasan teks dalam bahasa Indonesia dan masih bisa melalui proses *fine-tuning* dengan dataset penelitian, sehingga dapat menyesuaikan karakteristik berita lokal yang diterapkan. Dengan memanfaatkan keunggulan arsitektur *Text-to-Text Transfer Transformer* (T5) dalam memahami konteks dokumen dan menghasilkan ringkasan secara abstrak, penelitian ini diharapkan dapat menghasilkan ringkasan berita yang lebih tepat, koheren, dan relevan, sekaligus memberikan sumbangan dalam pengembangan sistem peringkas berita lokal Medan berbasis web yang mampu menyajikan informasi secara ringkas dan mudah dimengerti oleh pengguna.