

1. PENDAHULUAN

1.1. Latar Belakang

Dalam era digital yang sangat dinamis, akses terhadap informasi kian berkembang. Sedemikian banyak platform penyedia berita dengan informasi yang lengkap mulai dari ekonomi, politik, berita internasional, hingga berita mengenai dunia entertainment. Namun, dengan volume informasi yang terus meningkat, para pembaca seringkali mengalami kesulitan dalam mengikuti perkembangan informasi terkini melalui membaca berita (Sitanggung, 2022). Terkadang judul berita clickbait yang dituliskan juga dibuat agar pembaca berita tertarik untuk membaca berita secara keseluruhan, meskipun judul dapat menyebabkan misleading (Binus University Communication, 2022).

Maka dari itu, dibutuhkan sebuah metode untuk mengakses informasi dengan jumlah besar untuk meningkatkan produktivitas pengguna berita (Ahmad, 2020). Salah satu metode untuk meringkas teks yang panjang menjadi ringkasan singkat yang sudah memuat intisari dari teks tersebut disebut Automatic Text Summarization (ATS). ATS terbagi menjadi 3 jenis, yaitu extractive, abstractive, dan hybrid. Sesuai namanya, ATS jenis extractive meringkas teks dengan cara mengambil potongan - potongan kalimat yang penting. Sedangkan jenis abstractive jauh lebih kompleks karena model tidak hanya mengambil bagian yang penting dari teks asli, tapi juga generate kata - kata untuk menggabungkan antar poin penting atau bahkan antar kalimat (Moirangthem, 2020). Hybrid approach merupakan jenis terbaru dari text summarization yang membentuk hasil ringkasan dengan menggabungkan extractive dan abstractive (Muniraj, 2023).

Pemanfaatan *hybrid* ATS membantu menangani masalah koherensi pada metode extractive (Widyassari, 2022), serta masalah kelemahan pemilihan kata OOV (Out of Vocabulary) pada metode abstractive. Salah satu metode hybrid yang terbukti dapat menangani masalah - masalah tersebut adalah Pointer-Generator Networks (See, 2017). Abigail See menyebutkan bahwa performa Pointer-Generator Networks atau PGN melampaui RNN pada tugas *Sequence to Sequence* seperti mesin translasi. PGN memanfaatkan teknik seperti n-gram blocking dan coverage mechanism untuk mengatasi repetisi kata dan model dapat menentukan kapan harus menghasilkan kata baru atau mengambil kata dari teks asli. Hal ini yang menjadikan PGN memiliki keunggulan dibandingkan arsitektur model lain dalam mengolah OOV dan kata repetitif.

Penelitian lain yang membahas *Pointer-Generator Networks* pada Bahasa Indonesia menggunakan dataset Liputan6, mendapatkan hasil yang kurang baik karena masalah repetisi yang tidak dapat diatasi (Gilang Akbar, 2021). Gilang menyebutkan bahwa *Pointer-Generator*

Networks mempercepat progres penurunan nilai *loss*, namun masalah kata repetitif diakibatkan karena tidak diterapkannya *coverage mechanism* pada penelitian tersebut. Hal ini membuktikan bahwa *coverage mechanism* merupakan unsur penting pada arsitektur Pointer-Generator Networks. Rini Wijayanti yang melakukan penelitian terhadap PGN pada Bahasa Indonesia juga menyebutkan bahwa *coverage mechanism* penting untuk mengatasi kata repetitif. Rini juga menyebutkan 1 masalah lain yang ada pada model PGN adalah adanya terselip hasil ringkasan *out of context* atau *hallucination* (hasil ringkasan yang tidak relevan dengan teks asli) pada bagian ringkasan yang dibuat secara *abstractive* oleh model (Wijayanti, 2021). Hal ini menjadi *drawback* bagi PGN yang belum dapat diperbaiki, terutama pada tugas *news summarization* karena dapat menimbulkan hoax dari sifat *abstractive* model.

Terdapat penelitian dari Rini Wijayanti (2021) yang melakukan *fine-tuning* BERTSum-Extractive menggunakan *pretrained* model IndoBERT untuk *text summarization* dan dihasilkan bahwa model tersebut mendapatkan nilai tertinggi dibandingkan model ekstraktif lainnya seperti Lead-3 dan NeuralSum pada dataset IndoSum, sebuah dataset khusus untuk *text summarization* task dalam Bahasa Indonesia yang merupakan hasil penelitian Kemal Kurniawan (2018). Model hasil BertSum-Ext-IndoBERT tidak terdeteksi adanya *hallucination* atau *out of context error* (Wijayanti, 2021).

BERTSum-Extractive merupakan tipe ringkasan *extractive* yang dapat memperhitungkan relevansi kalimat. Kalimat yang memiliki tingkat relevansi rendah cenderung tidak dimasukkan kedalam hasil ringkasan (Torres, 2021). Pada skripsi ini selain menggunakan PGN sebagai arsitektur model utama, ditambahkan juga model BERTSum-Extractive-IndoBERT, yaitu *pretrained model* BERT versi Bahasa Indonesia yang di-*fine tune* untuk tugas *summarization*.

Metrik pertama yang akan digunakan adalah ROUGE, yang telah digunakan pada penelitian - penelitian sebelumnya untuk membandingkan hasil dengan penelitian sebelumnya. Namun menurut penelitian Gilang Akbar, nilai evaluasi ROUGE tidak dapat dijadikan acuan baik atau buruknya performa model (Gilang Akbar, 2021). Metrik ROUGE dinilai kurang sesuai untuk diterapkan pada *abstractive text summarization* karena metrik ROUGE menghitung jumlah token yang sama persis antara hasil ringkasan dan *golden summary*, padahal terdapat proses *vocabulary generation* pada model *text summarization* PGN yang bersifat *abstractive*. Maka, pada Skripsi ini menggunakan metrik kedua adalah BERTScore, yang merupakan metrik khusus *text generation* karena menghitung bukan dari *exact match token*, tapi dari *contextual embedding* token (Zhang., 2020) yang lebih memperhatikan semantik kalimat (Koto, 2020).

Hasil yang diharapkan dari Skripsi ini adalah dapat menghasilkan sistem rangkuman otomatis yang lebih efektif dan akurat untuk berita Bahasa Indonesia, serta memperbaiki kelemahan out of context dan kata repetitif dari model PGN dengan menggabungkan pretrained model BERTSum-*Extractive*-IndoBERT. Skripsi ini juga akan membandingkan metrik ROUGE dengan penelitian sebelumnya oleh Gilang Akbar (2021) yang mendapatkan bahwa ROUGE hasil model penerapan PGN pada News Summarization Bahasa Indonesia kurang baik.

1.2. Rumusan Masalah

Berdasarkan dari latar belakang di atas, maka rumusan masalah penelitian dapat dijabarkan sebagai berikut:

1. Berapa angka metrik ROUGE dan BERTScore dari hasil ringkasan model PGN + BERTSum-*Extractive*-IndoBERT?
2. Apakah PGN + BERTSum-*Extractive*-IndoBERT dapat memperbaiki *out-of-context error* yang ada pada PGN?

1.3. Tujuan Penelitian

Tujuan penelitian dibagi menjadi 3 tujuan, yaitu :

1. Mengintegrasikan arsitektur *Pointer-Generator Networks* dengan *pretrained model* BERTSum-*Extractive*-IndoBERT sebagai kerangka sistem ATS.
2. Membandingkan performa model integrasi dengan model pada penelitian sebelumnya.
3. Mengevaluasi kualitas hasil text summarization dengan membandingkan metrik hasil penelitian sebelumnya.

1.4. Ruang Lingkup

Ruang lingkup dibatasi pada:

1. Data diambil dari dataset IndoSum yang terdiri dari 18.774 berita Bahasa Indonesia dengan 6 kategori :
 - a. *Entertainment*
 - b. *Inspiration*
 - c. *Sport*

- d. *Showbiz*
 - e. *Headline*, dan
 - f. *Technology*
2. Metode yang akan digunakan :
 - a. *Pointer-Generator Networks* (PGN)
 - b. *Bidirectional Encoder Representations from Transformers* (BERT)
 3. Sistem integrasi model yang diusulkan adalah dengan meringkas teks berita asli menggunakan PGN, lalu meringkas teks hasil ringkasan PGN menggunakan BERTSum-Extractive-IndoBERT.
 4. Bahasa pemrograman yang digunakan adalah Python.
 5. Program akan menggunakan :
 - a. *Platform* Google Colab (ipynb)
 - b. *Tensorflow*
 - c. *PyTorch*
 - d. *Transformer*
 6. Hasil yang diharapkan :
 - a. Model peringkasan berita :
 - i. *Pointer-Generator Networks*
 - ii. *Pointer-Generator Networks* + BERTSum-Extractive-IndoBERT
 - iii. BERTSum-Extractive-IndoBERT + *Pointer-Generator Networks*
 - b. Metrik Evaluasi ROUGE dan BERTScore
 - c. Teks hasil ringkasan dari berita
 7. Input Program berupa teks artikel berita dari dataset (poin 1)
 8. Proses program berupa :
 - a. *Data Preprocessing* : Membersihkan teks artikel berita dari dataset.
 - b. *Text Processing* : Melakukan *tokenization*, *word embedding* (BERT-base), dan *unique token insertion* untuk menyesuaikan format *input* yang diharapkan model PGN dan BERTSum-Extractive-IndoBERT
 - c. Token kata dimasukkan ke *Pointer Generator Networks*
 - d. Hasil *Pointer Generator Networks* akan dimasukkan ke BERTSum-Extractive-IndoBERT
 - e. Mendapatkan *output* program
 9. Output Program berupa teks ringkasan berita

10. Metrik evaluasi akan menggunakan ROUGE dan BERTScore terhadap *reference summary* dari *dataset*.
11. Penilaian model akan melibatkan beberapa responden yang adalah masyarakat umum dan bukan merupakan pakar Bahasa Indonesia.

1.5. Manfaat Penelitian

Hasil dari Skripsi ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem *Automatic Text Summarization* yang lebih efektif dalam Bahasa Indonesia. Kerangka model yang dihasilkan dapat digunakan untuk penelitian lebih lanjut tentang *Automatic Text Summarization*. Dengan adanya Skripsi ini, dapat melihat apakah BERTSum-Extractive-IndoBERT dapat membantu menghilangkan *hallucination* dari PGN. Sistem ATS yang dihasilkan dapat digunakan oleh pembaca dan analis berita untuk mendapatkan informasi dari berita dengan cepat. Skripsi ini juga akan menghasilkan kesimpulan hasil dari penggabungan metode PGN dan BERTSum-Extractive-IndoBERT pada *Text Summarization* pada berita Bahasa Indonesia.

1.6. Metodologi Penelitian

Langkah-langkah dalam pengerjaan skripsi:

1. Studi Literatur
 - a. Metode *Pointer-Generator Networks*
 - b. Metode BERTSum-Extractive-IndoBERT
 - c. Metrik evaluasi ROUGE dan BERTScore
2. Analisis Penelitian sebelumnya
3. *Data Gathering, Cleaning, and Preprocessing* :
 - a. Mengumpulkan artikel berita dari dataset IndoSum
 - b. Melakukan *text cleaning* dengan menghapus karakter tidak umum.
 - c. Melakukan *tokenization* dengan membagi teks artikel menjadi kumpulan *string* per kata.
 - d. Pembentukan *vocabulary* untuk pembelajaran model.
4. Pembuatan Model
 - a. Metode *Pointer-Generator Networks*
 - b. Metode BERTSum-Extractive-IndoBERT
5. Training Model dan Penggabungan Metode

- a. Membuat ringkasan dengan PGN
 - b. Membuat ringkasan dengan BERTSum-*Extractive*-IndoBERT pada hasil ringkasan PGN
 - c. Membuat kombinasi urutan metode PGN dan BERTSum-*Extractive*-IndoBERT :
 - i. *Pointer-Generator Networks*
 - ii. *Pointer-Generator Networks* + BERTSum-*Extractive*-IndoBERT
6. Pengujian & Evaluasi Model
- a. Membandingkan metrik ROUGE dengan penelitian sebelumnya.
 - b. Membandingkan metrik ROUGE dan BERTScore pada setiap kombinasi metode.
 - c. Penilaian hasil ringkasan oleh beberapa responden yang berpengalaman dalam membaca berita.
7. Pengambilan kesimpulan model ATS dengan metode teknis dan analisa manusia.

1.7. Sistematika Penulisan

Berikut adalah sistematika penulisan dari proyek akhir ini:

BAB I : PENDAHULUAN

Bab I berisi latar belakang, rumusan masalah, tujuan, ruang lingkup, manfaat penelitian, metodologi penelitian, serta sistematika penulisan pada proyek akhir ini.

BAB II : LANDASAN TEORI

Bab II berisi kumpulan teori dan metode yang akan digunakan selama pengerjaan proyek. Teori dan metode yang akan digunakan meliputi *Pointer-Generator Networks*, IndoBERT, BERTSum, ROUGE, BERTScore.

BAB III : ANALISIS DAN DESAIN SISTEM

Bab III berisi analisis dan desain sistem yang akan dibuat.

BAB IV : IMPLEMENTASI SISTEM

Bab IV berisi implementasi serta pengujian sistem dari desain sistem yang telah dibuat.

BAB V : EVALUASI DAN ANALISIS HASIL EVALUASI

Bab V berisi hasil pengujian sistem yang telah dibuat berdasarkan implementasi pada bab IV.

BAB VI : KESIMPULAN DAN SARAN

Bab VI berisikan kesimpulan yang didapat dari hasil evaluasi pengujian sistem, hasil yang dicapai, dan saran bermanfaat untuk pengembangan sistem selanjutnya.