



**Pengembangan Sistem Peringkasan Teks  
Secara Ekstraktif Menggunakan Metode BERT  
Pada Penilaian Proposal SBK  
(Seminar Bidang Kajian) dan Ujian Kualifikasi**

**UJIAN KUALIFIKASI**

**Rogayah**

**99217016**

**PROGRAM DOKTOR TEKNOLOGI INFORMASI  
UNIVERSITAS GUNADARMA  
2024**

## Daftar Isi

Daftar Isi.....	ii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Batasan dan Rumusan Masalah .....	4
1.2.1 Batasan Masalah.....	4
1.2.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	5
1.4 Kontribusi dan Manfaat Penelitian.....	5
BAB II TELAAH PUSTAKA .....	6
2.1 Pemrosesan Bahasa Alami.....	6
2.2 Deep Learning.....	8
2.3 Transformers .....	9
2.4 Model Bahasa Besar (LLM).....	10
2.4.1 Prompt Learning.....	13
2.4.2 Architecture.....	14
2.4.3 Generative Pre-trained Transformer (GPT) .....	14
2.4.4 Bidirectional Encoder Representations from Transformers (BERT) .....	15
2.4.5 ROUGE (Recall-Oriented Understudy For Gisting Evaluation) .....	16
2.5 Penelitian Terdahulu .....	17
3.1 Gambaran Umum .....	21
3.1.1 Pengumpulan dataset .....	22
3.1.2 Preprocessing data.....	22
3.1.3 BERT Summarization .....	24
3.1.4 Hasil Ringkasan .....	24
3.1.5 Evaluasi Hasil .....	25
3.2 Jadwal Penelitian .....	25
DAFTAR PUSTAKA.....	26

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Proposal Seminar Bidang Kajian (SBK) dan Ujian Kualifikasi merupakan tahap kritis dalam pendidikan doctoral yang dirancang untuk menilai kesiapan dan kematangan mahasiswa dalam penelitian mereka. Proposal Seminar Bidang Kajian memungkinkan mahasiswa untuk mempresentasikan rencana penelitian mereka—termasuk latar belakang, tujuan, metodologi, dan kontribusi yang diharapkan—untuk mendapatkan umpan balik dan persetujuan dari panel penguji (Kumar et al., 2022). Ujian Kualifikasi, di sisi lain, menguji pemahaman mendalam mahasiswa terhadap teori dan metodologi yang relevan dengan bidang studi mereka serta kemampuan untuk menerapkan pengetahuan tersebut secara efektif (Jung & Wang, 2021). Keduanya berfungsi sebagai indikator penting untuk memastikan mahasiswa memenuhi standar akademik yang diperlukan sebelum memasuki fase disertasi yang lebih kompleks (Smith et al., 2020).

Penilaian dokumen akademik seperti Proposal Seminar Bidang Kajian yang selanjutnya disebut SBK dan Ujian Kualifikasi merupakan tahap krusial dalam program doctoral, yang menentukan kualitas dan kematangan penelitian mahasiswa. Dokumen-dokumen ini tidak hanya harus menunjukkan pemahaman mendalam mahasiswa tentang topik penelitian tetapi juga kemampuan mereka dalam menyajikan informasi secara sistematis dan kritis. Namun, proses penilaian yang melibatkan evaluasi dokumen yang panjang dan kompleks seringkali menjadi tantangan bagi penguji karena memerlukan waktu dan usaha yang besar untuk menilai setiap detail (Kumar & Ghosh, 2020).

Seiring dengan berkembangnya teknologi pemrosesan bahasa alami, metode peringkasan teks secara ekstraktif telah menjadi alat yang efektif untuk menangani informasi besar dan kompleks. Peringkasan teks bertujuan untuk menghasilkan ringkasan yang informatif dan representatif dari dokumen panjang

dengan mempertahankan konten penting, yang sangat relevan dalam konteks penilaian dokumen akademik (Nallapati et al., 2016). Metode ini memudahkan penguji untuk menilai dokumen secara efisien dengan menyajikan informasi utama tanpa harus membaca seluruh teks secara rinci.

BERT (Bidirectional Encoder Representations from Transformers) merupakan salah satu model terbaru dalam pemrosesan bahasa alami yang telah menunjukkan performa luar biasa dalam berbagai tugas pemahaman teks, termasuk peringkasan. BERT menggunakan pendekatan berbasis Transformer untuk memahami konteks kata secara bidirectional, yang memungkinkan untuk menangkap makna yang lebih dalam dan kontekstual dalam teks (Devlin et al., 2018). Penggunaan BERT dalam peringkasan teks menawarkan potensi untuk meningkatkan kualitas ringkasan dengan mempertimbangkan konteks yang lebih luas dan relevansi informasi yang lebih baik.

Di bidang akademik, aplikasi BERT dalam peringkasan teks masih terbatas, dan penelitian khusus tentang penerapannya dalam konteks penilaian dokumen akademik seperti SBK dan Ujian Kualifikasi belum banyak dilakukan. Penelitian ini bertujuan untuk mengisi celah tersebut dengan mengembangkan sistem peringkasan teks secara ekstraktif menggunakan metode BERT dan menerapkannya pada dokumen proposal dan ujian kualifikasi. Dengan demikian, diharapkan sistem ini dapat memberikan ringkasan yang akurat dan berguna untuk memfasilitasi proses penilaian yang lebih efisien (Liu et al., 2019).

Penelitian terbaru menunjukkan bahwa teknologi pemrosesan bahasa alami (NLP) dapat memperbaiki proses penilaian dokumen akademik secara signifikan. Kumar et al. (2022) meneliti efektivitas seminar proposal dalam meningkatkan keterampilan penelitian dan menemukan bahwa meskipun seminar proposal memberikan umpan balik yang berharga, integrasi teknologi dalam proses ini dapat lebih meningkatkan efisiensi dan akurasi penilaian. Dalam konteks ini, sistem peringkasan teks berbasis teknologi NLP seperti BERT dapat memainkan peran penting dengan menyederhanakan dan merangkum dokumen panjang untuk

mempercepat evaluasi, memungkinkan panel penguji untuk fokus pada aspek-aspek kunci dari proposal penelitian tanpa harus menelaah keseluruhan dokumen secara rinci.

Selain itu, Ujian Kualifikasi juga telah menjadi fokus penelitian terkait penerapan teknologi dalam pendidikan doctoral. Jung dan Wang (2021) menekankan pentingnya pemahaman mendalam terhadap teori dan metodologi yang diuji melalui Ujian Kualifikasi. Penelitian mereka menunjukkan bahwa penilaian berbasis teknologi dapat membantu mengidentifikasi area di mana mahasiswa mungkin memerlukan dukungan tambahan, serta memberikan umpan balik yang lebih terstruktur dan terpersonalisasi. Penggunaan sistem berbasis BERT dalam mengolah dan merangkum jawaban ujian dapat meningkatkan efisiensi dalam mengevaluasi pemahaman mahasiswa terhadap topik yang kompleks dan menyajikan informasi yang relevan secara lebih ringkas.

Dalam studi yang lebih luas, Smith et al. (2020) mengeksplorasi dampak Ujian Kualifikasi terhadap keberhasilan PhD, menemukan bahwa ujian ini memainkan peran penting dalam menentukan kesiapan mahasiswa untuk tahap disertasi. Penelitian mereka mendukung ide bahwa teknologi, seperti sistem peringkasan berbasis BERT, dapat memperbaiki cara penilaian dilakukan dengan memberikan ringkasan yang komprehensif dari dokumen panjang, mengurangi waktu evaluasi, dan meningkatkan akurasi dalam penilaian kualifikasi. Dengan mengintegrasikan teknologi canggih dalam proses penilaian, diharapkan dapat meningkatkan kualitas evaluasi dan mendukung kesuksesan mahasiswa doctoral di masa depan.

Hasil dari penelitian ini diharapkan dapat menawarkan solusi teknologi yang bermanfaat bagi penguji dalam menilai dokumen akademik. Implementasi sistem peringkasan berbasis BERT tidak hanya dapat mempercepat proses evaluasi tetapi juga meningkatkan kualitas penilaian dengan menyediakan ringkasan yang lebih representatif dari dokumen akademik yang kompleks. Penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap

pengembangan metode peringkasan teks dan aplikasi praktis dalam pendidikan tinggi.

## **1.2 Batasan dan Rumusan Masalah**

Adanya batasan dan rumusan masalah pada penelitian ini adalah untuk memfokuskan penelitian pada hal-hal yang berkaitan dengan penyelesaian masalah. Berikut batasan dan rumusan masalah dalam penelitian ini.

### **1.2.1 Batasan Masalah**

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya akan memfokuskan pada penilaian dokumen proposal SBK dan Ujian Kualifikasi dengan teknik *text summarization* yaitu *extractive summarization*.
2. Dataset yang digunakan dalam penelitian ini terdiri dari proposal SBK dan Ujian Kualifikasi yang ditulis dalam bahasa Indonesia.
3. Model yang digunakan adalah model *text summarization* berbasis Transformer seperti BERT.
4. Evaluasi dari hasil *summarization* akan menggunakan metrik standar seperti ROUGE untuk menilai kualitas ringkasan.

### **1.2.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka diperoleh beberapa permasalahan yang harus diselesaikan sehingga permasalahan tersebut dirumuskan sebagai berikut:

1. Bagaimana membangun model platform penilaian proposal SBK dan Ujian Kualifikasi menggunakan *Text Summarization*?
2. Bagaimana algoritma untuk menghasilkan peringkasan teks untuk penilaian proposal SBK dan Ujian Kualifikasi?

### **1.3 Tujuan Penelitian**

Dengan fokus pada pengembangan model yang memanfaatkan teknik *text summarization*, penelitian ini bertujuan untuk membangun sistem yang dapat secara otomatis merangkum informasi dari proposal proposal SBK dan Ujian Kualifikasi. Selain itu, untuk memastikan bahwa ringkasan yang dihasilkan berkualitas tinggi dan relevan, penelitian ini juga mengeksplorasi algoritma yang dapat digunakan untuk menghasilkan peringkasan teks yang optimal dalam konteks penilaian proposal proposal SBK dan Ujian Kualifikasi.

### **1.4 Kontribusi dan Manfaat Penelitian**

Penelitian ini memberikan kontribusi penting dalam bidang akademik dan teknologi yaitu menerapkan teknologi *text summarization* dengan adanya model platform penilaian proposal proposal SBK dan Ujian Kualifikasi yang berbasis teknologi kecerdasan buatan.

Dari penelitian ini, beberapa manfaat utama yang dapat diperoleh yaitu:

1. Memudahkan para penilai dalam melakukan penilaian proposal.
2. Meningkatkan akurasi penilaian pada dokumen proposal.
3. Memberikan efisiensi waktu untuk penilaian yang lebih baik.

## BAB II

### TELAAH PUSTAKA

#### 1.1 Pemrosesan Bahasa Alami (*Natural Language Processing*)

Pemrosesan Bahasa Alami (NLP) adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP mencakup berbagai teknik dan teori yang memungkinkan komputer untuk memahami, memproses, dan menghasilkan bahasa manusia secara alami. Menurut Jurafsky dan Martin (2021), NLP mencakup proses seperti tokenisasi, part-of-speech tagging, parsing, dan pemahaman semantik, yang semuanya memainkan peran penting dalam pemrosesan teks, termasuk peringkasan teks.

Tujuan NLP untuk memahami bahasa manusia ini memiliki banyak tantangan, yang antara lain adalah sebagai berikut:

1. Penandaan Kelas Kata (*Part-of-Speech Tagging*)

Penandaan kelas kata atau POS tagging merupakan tantangan signifikan dalam NLP karena ketergantungannya pada konteks.

1. Segmentasi Teks (*Text Segmentation*)

Segmentasi teks merupakan masalah utama dalam bahasa tanpa pemisah kata jelas seperti Mandarin dan Jepang, serta dalam bahasa lisan yang sering mencampur kata.

2. Disambiguasi Makna Kata (*Word Sense Disambiguation*)

Disambiguasi makna kata merupakan tantangan karena banyak kata memiliki lebih dari satu makna yang tergantung pada konteksnya.

3. Ambiguitas Sintaksis (*Syntactic Ambiguity*)

Ambiguitas sintaksis terjadi ketika kalimat dapat diartikan dengan beberapa struktur berbeda.

4. Masukan yang Tak Sempurna atau Tak Teratur (*Imperfect or Irregular Input*)

Kesalahan dalam ejaan, tata bahasa, dan aksen dalam bahasa lisan menyulitkan pemrosesan bahasa alami.



## 6. Pertuturan (*Speech Acts*)

Maksud di balik kalimat kadang tidak jelas hanya dengan melihat struktur kalimatnya. NLP harus memperhitungkan konteks dan gaya bahasa untuk memahami maksud penutur.

Penerapan NLP dalam Berbagai Bidang, terdiri dari :

### 1. Pemerolehan Informasi (*Information Retrieval*)

NLP digunakan dalam pencarian dokumen relevan dan ekstraksi informasi spesifik dari dokumen.

### 2. Penjawaban Pertanyaan (*Question Answering*)

Sistem QA otomatis menjawab pertanyaan dengan bahasa alami.

### 3. Perangkuman Otomatis (*Automatic Summarization*)

Perangkuman otomatis menggunakan NLP untuk menghasilkan ringkasan singkat dari dokumen panjang. Model berbasis pembelajaran mendalam seperti BERT dan GPT telah memperbaiki kualitas ringkasan yang dihasilkan (Nallapati et al., 2016).

### 4. Penerjemahan Mesin (*Machine Translation*)

Penerjemahan otomatis telah berkembang dengan pesat melalui model-model seperti Transformer dan BERT, yang memungkinkan terjemahan bahasa yang lebih akurat dan kontekstual (Wu et al., 2016).

### 5. Pengenalan Wicara (*Speech Recognition*)

NLP diterapkan dalam pengenalan ucapan untuk mengubah bahasa lisan menjadi teks.

### 6. Sintesis Wicara (*Speech Synthesis*)

Sintesis wicara mengubah teks menjadi ucapan dengan teknik seperti text-to-speech (TTS).

### 7. Pengenalan Karakter Optis (*Optical Character Recognition*)

OCR digunakan untuk mengubah teks tercetak atau tulisan tangan menjadi format digital yang dapat diproses mesin.

### 8. Analisis Sentimen (*Sentiment Analysis*)

NLP digunakan untuk menganalisis opini dalam teks untuk menentukan sentimen positif atau negatif.

*Natural Language Processing* (NLP) mencakup berbagai tugas dan aplikasi yang memungkinkan komputer untuk memahami, menganalisis, dan menghasilkan bahasa manusia. Berikut adalah beberapa tugas utama dalam NLP:

a) Tokenisasi dan Pra-pemrosesan

Tokenisasi adalah proses memecah teks menjadi unit-unit dasar, atau token, seperti kata atau frasa. Ini adalah langkah pertama dalam banyak aplikasi NLP karena mempermudah analisis lebih lanjut. Pra-pemrosesan teks juga melibatkan teknik seperti stemming dan lemmatization untuk mengurangi variasi bentuk kata.

b) Part-of-Speech (POS) Tagging

POS tagging adalah proses menandai kata dalam teks dengan kategori tata bahasanya seperti kata benda, kata kerja, dan kata sifat. Teknik ini penting untuk analisis sintaksis dan semantik dan sering menggunakan model probabilistik seperti Hidden Markov Models (HMMs) dan Conditional Random Fields (CRFs).

c) Named Entity Recognition (NER)

NER adalah tugas untuk mengidentifikasi dan mengklasifikasikan entitas nama dalam teks, seperti nama orang, organisasi, dan lokasi. Ini penting dalam ekstraksi informasi dan memahami konteks dari teks.

d) Parsing dan Analisis Sintaksis

Parsing melibatkan analisis struktur gramatikal kalimat, membangun pohon sintaksis yang menggambarkan hubungan antar kata atau frasa. Ada dua jenis parsing utama: dependency parsing dan constituency parsing. Dependency parsing fokus pada hubungan antar kata, sedangkan constituency parsing membangun struktur hierarkis frasa.

## 1.1 Deep Learning

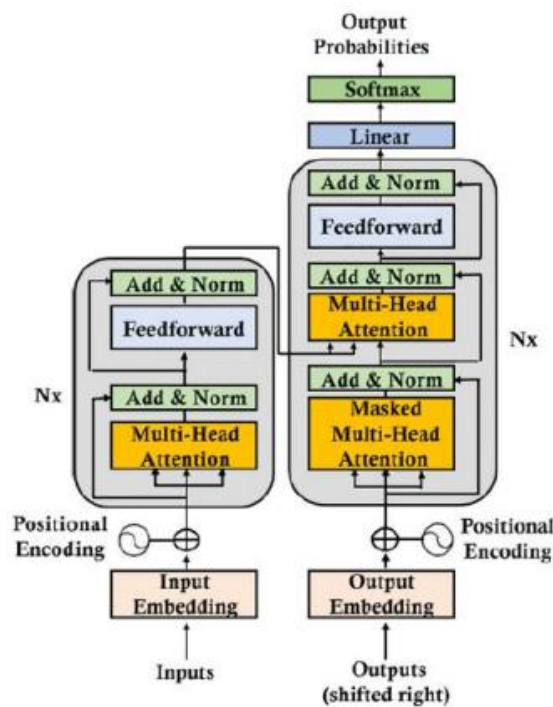
Deep learning adalah sub-bidang dari pembelajaran mesin yang berfokus pada penggunaan jaringan saraf tiruan dengan banyak lapisan untuk memodelkan representasi data yang kompleks. Jaringan saraf tiruan (neural networks) terdiri

dari neuron-neuron buatan yang disusun dalam lapisan-lapisan: lapisan input, lapisan tersembunyi (hidden layers), dan lapisan output. Konsep dasar deep learning adalah memanfaatkan struktur berlapis ini untuk menangkap hubungan non-linear dan pola yang sulit dipahami oleh model-model tradisional. Model deep learning dapat mempelajari representasi data dari data mentah secara otomatis, mengurangi kebutuhan akan rekayasa fitur manual (LeCun, Bengio, & Hinton, 2015).

Arsitektur jaringan saraf dalam deep learning meliputi berbagai jenis, seperti jaringan saraf konvolusi (Convolutional Neural Networks, CNN) dan jaringan saraf rekuren (Recurrent Neural Networks, RNN). CNN, yang diperkenalkan oleh LeCun et al. (1998), sangat efektif dalam tugas-tugas yang melibatkan data berbentuk grid, seperti pengenalan gambar, dengan menggunakan lapisan konvolusi untuk mengekstraksi fitur spasial. Sebaliknya, RNN dirancang untuk menangani data sekuensial dengan memelihara informasi dari langkah-langkah sebelumnya, tetapi sering menghadapi kesulitan dalam menangkap dependensi jangka panjang, yang diatasi oleh Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997).

## **1.2 Transformers**

Model transformer, yang diperkenalkan oleh Vaswani et al. (2017), telah membawa revolusi dalam deep learning dengan memanfaatkan mekanisme perhatian (*attention mechanism*) untuk memproses data sekuensial secara paralel. Ini mengatasi keterbatasan RNN dan LSTM dalam hal efisiensi dan kemampuan untuk menangkap hubungan jangka panjang dalam data sekuensial. Model-model seperti BERT dan GPT, yang dibangun di atas arsitektur transformer, telah mencapai hasil yang sangat baik dalam berbagai tugas NLP, termasuk penerjemahan mesin, analisis sentimen, dan penjawaban pertanyaan, menunjukkan kekuatan deep learning dalam memahami dan menghasilkan teks (Devlin et al., 2018; Radford et al., 2019).



Gambar 2.1 Arsitektur *Transformer*

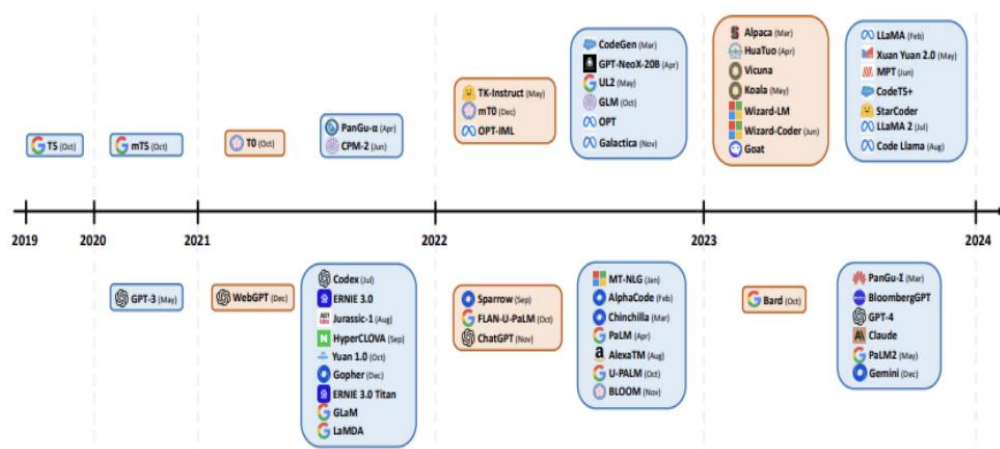
Pada gambar 2.1 di sebelah kiri, input masuk ke sisi encoder Transformer melalui attention sublayer dan feedforward sublayer. Di sebelah kanan, output target masuk ke sisi dekoder Transformer melalui dua attention sublayer dan feedforward sublayer. Pada arsitektur ini tidak ada lagi RNN, LSTM, atau CNN. Recurrence telah ditinggalkan dalam arsitektur ini.

Attention telah menggantikan fungsi recurrence yang membutuhkan parameter yang meningkat seiring bertambahnya jarak antara dua kata meningkat. Mekanisme attention adalah operasi "kata ke kata". Ini sebenarnya adalah operasi token-ketoken. Mekanisme attention akan menemukan bagaimana setiap kata berhubungan dengan semua kata lain dalam sebuah urutan, termasuk kata yang sedang dianalisis itu sendiri.

### 1.3 Model Bahasa Besar (LLM)

Large language models adalah model-model bahasa yang memiliki ukuran dan kompleksitas yang besar, terdiri dari jutaan atau bahkan miliaran parameter.

Model-model ini dirancang untuk memahami dan menghasilkan bahasa manusia dengan tingkat keterampilan yang tinggi. LLM dibangun dengan menggunakan teknik-teknik deep learning, terutama menggunakan arsitektur seperti Transformer, yang memungkinkan LLM untuk menangani sejumlah besar data dan menangkap pola-pola kompleks dalam bahasa manusia. Large language models telah menjadi fokus utama dalam penelitian NLP karena kemampuan mereka untuk melakukan berbagai tugas, seperti penerjemahan bahasa, generasi teks, analisis sentiment dan sebagainya.



Gambar 2.2 Model Bahasa Besar (LLM)

Pelatihan model bahasa besar (LLM) pada dataset yang sangat besar telah mengubah secara signifikan pemrosesan bahasa alami, memungkinkan LLM untuk meniru interaksi manusia dan berperan sebagai asisten serbaguna dalam berbagai tugas. Ini termasuk memberikan jawaban yang luas terhadap pertanyaan, membantu dalam penulisan, memberikan pengajaran, dan beragam tugas lainnya. Model bahasa yang besar dapat dikategorikan berdasarkan berbagai factor seperti arsitektur, tujuan pelatihan, tipe data, dan aplikasi (Amaratunga, 2023). Berdasarkan arsitekturnya LLM dibedakan berdasarkan beberapa kategori, yaitu:

1. Transformer
2. Recurrent Neural Network
3. Convolutional Neural Network

Adapun komponen-komponen penting dalam LLM :

#### 1. Tokenization (tokenisasi)

Tokenization adalah langkah pra-pemrosesan penting dalam pelatihan model bahasa besar (LLM) yang memecah teks menjadi unit-unit tak terpisahkan yang disebut token (Naveed et al., 2023). Token dapat berupa karakter, subkata, simbol, atau kata, tergantung pada proses tokenisasi yang digunakan. Beberapa skema tokenisasi yang umum digunakan dalam LLM termasuk wordpiece, byte pair encoding (BPE), dan unigramLM (Webster & Kit, n.d.).

#### 2. Encoding Positions

Transformer memproses urutan input secara paralel dan independen, tanpa mencatat informasi posisi. Oleh karena itu, diperkenalkan positional encoding di transformer, di mana vektor positional encoding ditambahkan ke token embedding. Varian positional embedding meliputi positional encoding absolute, relative, or learned positional encodings. Dalam relative encoding terdapat 2 positional embedding yang sering digunakan dalam LLMs yaitu Alibi dan RoPE.

#### 3. Attention

Attention bertugas memberikan bobot kepada token-token input berdasarkan kepentingannya sehingga model memberikan penekanan lebih pada token-token yang relevan (Vaswani et al., n.d.). Attention dalam transformer menghitung pemetaan query, key, dan value untuk urutan input, di mana skor perhatian diperoleh dengan mengalikan query dan key, dan kemudian digunakan untuk memberi bobot pada nilai-nilai.

#### 4. Decoder

Modul decoder dari model Transformer terdiri dari beberapa lapisan identik, masing-masing mencakup mekanisme multi-head attention dan jaringan saraf feedforward (Liu et al., 2024). Berbeda dengan encoder, decoder juga mencakup mekanisme attention tambahan antara encoder dan decoder, digunakan untuk menghitung perhatian pada urutan input selama proses decoding. Pada setiap posisi, decoder hanya dapat melakukan perhitungan self-attention dengan posisi sebelumnya untuk memastikan urutan yang dihasilkan tidak melanggar aturan tata bahasa. Mask memainkan peran penting dalam decoder, memastikan

bahwa hanya informasi sebelum langkah waktu saat ini yang diperhatikan saat menghasilkan urutan output, dan tidak membocorkan informasi dari langkah waktu mendatang. Mekanisme self-attention pada decoder menggunakan mask untuk mencegah model mengakses informasi masa depan saat menghasilkan prediksi pada setiap langkah waktu, menjaga kausalitas model. Hal ini memastikan bahwa output yang dihasilkan oleh model bergantung pada informasi pada langkah waktu saat ini dan sebelumnya, tanpa dipengaruhi oleh informasi masa depan.

#### **2.4.1 Prompt Learning**

Prompt learning adalah pendekatan machine learning yang digunakan secara luas, terutama di bidang NLP. Metodologi ini melibatkan pembuatan pernyataan prompt yang hati-hati untuk mengarahkan model menghasilkan perilaku atau output tertentu. Pendekatan ini sering digunakan untuk fine-tuning dan mengarahkan LLM yang sudah dilatih sebelumnya untuk menjalankan tugas tertentu atau menghasilkan hasil yang diinginkan (Liu et al., 2024). Desain pernyataan prompt dapat mengarahkan model pre-trained untuk melakukan berbagai tugas seperti menjawab pertanyaan, menghasilkan teks, dan memahami semantik. Kekuatan pendekatan ini terletak pada kemampuannya untuk beradaptasi dengan berbagai tugas melalui modifikasi sederhana pada pernyataan prompt, tanpa perlu melatih ulang seluruh model. Untuk LLM seperti seri GPT dan model pre-trained lainnya, prompt learning menyediakan cara yang mudah dan kuat untuk fine-tuning model. Dengan memberikan prompt yang sesuai, peneliti dan praktisi dapat menyesuaikan perilaku model agar lebih cocok untuk domain atau kebutuhan tugas tertentu. Singkatnya, prompt learning adalah pendekatan machine learning yang membangun model bahasa yang telah dilatih sebelumnya dan mengarahkannya untuk melakukan berbagai tugas melalui desain pernyataan prompt, menawarkan fleksibilitas yang meningkat untuk menyesuaikan aplikasi model.

### **2.4.2 Architecture**

Saat ini semua Large Language Models (LLMs) dibangun dengan menggunakan arsitektur Transformer. Arsitektur ini memungkinkan model-model ini untuk memiliki skala hingga beberapa miliar atau bahkan triliun parameter. Secara umum, arsitektur Pretrained language model (PLM) dapat dikelompokkan menjadi tiga kategori: Encoder-only, Encoder-decoder, dan Decoder-only. Arsitektur Encoder-only tidak lagi digunakan dalam LLMs terbaru dan tidak akan dibahas lebih lanjut di sini. Sebagai gantinya, fokus bagian ini adalah untuk memperkenalkan arsitektur Encoder-decoder dan Decoder-only.

#### **1. Encoder-Decoder**

Arsitektur Encoder-decoder pada LLMs didasarkan pada arsitektur Transformer tradisional. Terdiri dari dua komponen utama: Encoder dan Decoder, di mana Encoder menyandikan urutan input melalui beberapa lapisan Multi-Head Self-Attention, sedangkan Decoder menggunakan cross-attention pada representasi output dari Encoder untuk menghasilkan urutan target secara autoregresif. Arsitektur ini menjadi dasar bagi LLM terkenal seperti T5, flan-T5, dan BART.

#### **2. Decoder only**

LLMs dengan arsitektur Decoder-only memanfaatkan komponen decoder dari arsitektur Transformer tradisional. Berbeda dengan arsitektur Encoder-Decoder yang menggabungkan encoder dan decoder, arsitektur Decoder-only hanya fokus pada proses dekoding. Model ini secara berurutan menghasilkan token-token dengan memperhatikan token-token sebelumnya dalam urutan. Arsitektur ini telah diterapkan dalam berbagai tugas generasi bahasa, menunjukkan efektivitasnya dalam menghasilkan teks tanpa memerlukan fase encoding eksplisit. Arsitektur Decoder-only dapat dibedakan lagi menjadi dua kategori: arsitektur Causal Decoder dan arsitektur Prefix Decoder.

### **2.4.3 Generative Pre-trained Transformer (GPT)**

Generative Pre-trained Transformer (GPT) adalah model yang mempopulerkan LLM kepada masyarakat umum. GPT adalah keluarga LLM yang



dirilis oleh OpenAI, sebuah laboratorium penelitian kecerdasan buatan Amerika yang terdiri dari organisasi nirlaba OpenAI Inc. Model GPT OpenAI menggunakan pendekatan semi-supervisi, yang merupakan pertama kalinya pendekatan semacam itu digunakan dengan model transformator. Pendekatan ini melibatkan dua tahap:

1. Tahap pra-pelatihan generatif tanpa pengawasan di mana tujuan pemodelan bahasa digunakan untuk menetapkan parameter awal.
2. Tahap "fine-tuning" yang diawasi di mana parameter ini disesuaikan dengan tugas target.

#### **2.4.4 Bidirectional Encoder Representations from Transformers (BERT)**

**BERT** adalah model pre-trained yang dikembangkan oleh Google AI dan diperkenalkan dalam makalah oleh Devlin et al. pada tahun 2018. BERT mengubah paradigma pemodelan bahasa dengan memperkenalkan representasi bidirectional yang mendalam. Konsep Utama BERT:

1. Bidirectional Contextual Understanding:

BERT menggunakan arsitektur Transformer dengan perhatian bidirectional, memungkinkan model untuk memahami konteks dari kedua sisi (kiri dan kanan) dari sebuah kata dalam sebuah kalimat. Ini berbeda dari model-model sebelumnya seperti GPT (Generative Pre-trained Transformer) yang hanya mempertimbangkan konteks unidirectional (satu arah).

2. Pre-training dan Fine-tuning:

BERT dilatih pada dua tugas utama selama fase pre-training: Masked Language Modeling (MLM), di mana beberapa kata dalam kalimat di-mask dan model harus memprediksi kata-kata yang hilang, dan Next Sentence Prediction (NSP), di mana model belajar untuk memprediksi apakah satu kalimat mengikuti kalimat lain dalam teks. Setelah pre-training, BERT dapat di-fine-tune pada tugas spesifik seperti klasifikasi teks, penjawaban pertanyaan, dan ringkasan.

### 3. Tokenization:

BERT menggunakan teknik tokenisasi WordPiece yang membagi kata menjadi sub-kata untuk menangani kata-kata yang tidak ditemukan dalam kosakata.

Kelebihan BERT:

- a) Kemampuan untuk memahami konteks yang lebih dalam dan lebih luas dari kata-kata dalam kalimat.
- b) Kinerja tinggi dalam berbagai tugas NLP dengan menggunakan fine-tuning yang relatif mudah.
- c) Kemampuan untuk menangani berbagai bahasa dengan adaptasi yang sesuai.

Kekurangan BERT:

- a) Memerlukan sumber daya komputasi yang besar untuk pelatihan dan inferensi.
- b) Tidak dirancang untuk menghasilkan teks; lebih fokus pada pemahaman teks.

#### **2.4.5 ROUGE (Recall-Oriented Understudy for Gisting Evaluation)**

ROUGE adalah metrik evaluasi yang umum digunakan dalam penilaian kualitas ringkasan teks, terutama dalam sistem otomatisasi ringkasan. ROUGE mengukur kualitas ringkasan dengan membandingkannya dengan ringkasan referensi (ground truth) yang dianggap sebagai acuan.

ROUGE terdiri dari beberapa varian metrik, yang paling umum digunakan adalah:

1. ROUGE-N: Mengukur kesamaan n-gram (subsekuensi kata dengan panjang n) antara ringkasan otomatis dan referensi. Contoh ROUGE-1 mengukur kesamaan unigram (kata tunggal), ROUGE-2 mengukur bigram (dua kata berturut-turut), dan seterusnya.
2. ROUGE-L: Mengukur kesamaan berdasarkan panjang subsekuensi terpanjang yang mengikuti urutan kata yang sama antara ringkasan otomatis dan referensi. ROUGE-L menghitung recall dari subsekuensi terpanjang yang dapat ditemukan dalam kedua teks.
3. ROUGE-W: Varian dari ROUGE-L yang memperhitungkan panjang subsekuensi dan menggabungkan skor berdasarkan urutan panjang subsekuensi.

4. ROUGE-S: Mengukur kesamaan dengan memperhitungkan bigram dan trigram yang tidak harus berurutan (skip-bigram) antara ringkasan otomatis dan referensi.
5. ROUGE-SU: Memperluas ROUGE-S dengan menghitung skor berdasarkan semua bigram yang tidak harus berurutan dan memperhitungkan unigram tambahan.

Metrik-metrik ini menghitung recall, precision, dan F1-score dari n-gram, subsekuensi, atau bigram untuk menilai kualitas ringkasan. ROUGE sangat berguna karena memberikan indikasi seberapa baik ringkasan otomatis mendekati kualitas ringkasan manusia dalam hal konten dan struktur. Namun, ROUGE memiliki keterbatasan, seperti ketergantungan pada n-gram dan tidak memperhitungkan makna semantik atau struktur kalimat, sehingga sering digunakan bersama dengan metrik evaluasi lainnya untuk memberikan gambaran yang lebih lengkap tentang kualitas ringkasan.

#### 1.4 Penelitian Terdahulu

Sub bab ini membahas penelitian-penelitian terdahulu yang relevan dengan pengembangan sistem peringkasan teks ekstraktif menggunakan metode BERT, khususnya dalam konteks Proposal Seminar Bidang Kajian (SBK) dan Ujian Kualifikasi. Kajian ini mencakup penelitian tentang penerapan model bahasa pre-trained, seperti BERT, dalam berbagai aplikasi peringkasan teks serta studi terkait yang mengkaji efisiensi dan efektivitas model tersebut dalam menangani dokumen panjang dan kompleks.

**Tabel 2.1 State Of The Art**

No	Judul	Tahun	Penulis	Metode	Kelebihan	Kekurangan
1	Enhancing Extractive Summarization with Pre-trained Language Models	2020	Zhang, Y., Zhao, Y., & Zhang, Y.	BERT	Meningkatkan kualitas ringkasan dengan representasi konteks yang mendalam dari BERT.	Memerlukan fine-tuning spesifik untuk setiap jenis teks.

No	Judul	Tahun	Penulis	Metode	Kelebihan	Kekurangan
2	Extractive Summarization Using BERT with Attention Mechanism	2021	Liu, H., Zheng, Z., & Zhao, L.	BERT	Meningkatkan relevansi ringkasan dengan mempertimbangkan konteks lebih dalam.	Model yang lebih kompleks dan memerlukan lebih banyak sumber daya komputasi.
3	BERTSUM: Extractive Summarization of Abstractive Texts with BERT	2021	Wang, L., & Liu, J.	BERT	Meningkatkan performa dengan menggabungkan teknik ekstraktif dan abstraktif.	Kompleksitas model yang lebih tinggi dan memerlukan fine-tuning yang ekstensif.
4	A Survey on Text Summarization Using BERT and Other Deep Learning Models	2022	Li, H., & Zhang, Q.	BERT	Menyediakan gambaran komprehensif tentang kemajuan terbaru dalam peringkasan teks.	Tidak mendalami implementasi spesifik dari masing-masing teknik.
5	Utilizing BERT for Extractive Summarization of Long Documents	2022	Chen, Y., & Liu, Z.	BERT	Meningkatkan kemampuan BERT dalam menangani dokumen panjang dengan teknik pemrosesan tambahan.	Memerlukan penyesuaian untuk ukuran dokumen yang sangat besar.
6	Pre-trained Language Models for Extractive Document Summarization: A Comparative Study	2021	Wang, H., & Xu, Y.	BERT	Memperbandingkan berbagai model pre-trained untuk menentukan performa terbaik dalam peringkasan ekstraktif.	Fokus pada perbandingan tanpa mendalami penerapan spesifik model.

No	Judul	Tahun	Penulis	Metode	Kelebihan	Kekurangan
7	Improving Extractive Summarization with Transformer-based Models	2021	Zhang, X., & Zhao, Y.	BERT	Menyediakan metode baru untuk meningkatkan peringkasan ekstraktif dengan transformer.	Mungkin tidak optimal untuk semua jenis teks atau domain.
8	BERT-Based Extractive Summarization for Scientific Papers	2022	Yang, W., & Zhou, Z.	BERT	Meningkatkan relevansi ringkasan untuk teks ilmiah dengan representasi konteks BERT.	Memerlukan fine-tuning untuk berbagai jenis dokumen ilmiah.
9	Leveraging BERT for Extractive Text Summarization in Legal Documents	2022	Liu, J., & Zhao, Q.	BERT	Menyediakan solusi khusus untuk dokumen hukum dengan BERT.	Hasil mungkin bervariasi tergantung pada kompleksitas dokumen hukum.
10	Extractive Summarization of Academic Texts Using BERT: A Case Study	2023	Hu, J., & Liu, S.	BERT	Memfokuskan pada aplikasi BERT dalam konteks akademik, meningkatkan relevansi ringkasan.	Perlu adaptasi untuk jenis teks akademik yang berbeda.

Pengembangan sistem peringkasan teks ekstraktif berbasis BERT dapat difokuskan pada peningkatan performa dengan fine-tuning model pada dataset khusus yang relevan dengan domain aplikasi. Penelitian sebelumnya, seperti yang dilakukan oleh Zhang et al. (2020) dan Yang & Zhou (2022), menunjukkan bahwa meskipun BERT memberikan hasil yang menjanjikan dalam peringkasan teks, model ini sering kali memerlukan penyesuaian spesifik untuk domain tertentu

seperti teks akademik atau dokumen hukum. Oleh karena itu, salah satu arah pengembangan adalah melatih BERT dengan dataset khusus dari Proposal Seminar Bidang Kajian (SBK) dan Ujian Kualifikasi untuk meningkatkan akurasi dan relevansi ringkasan. Pendekatan ini dapat memperbaiki kualitas ringkasan dengan mempertimbangkan konteks khusus dari jenis dokumen tersebut, sehingga hasil ringkasan menjadi lebih berguna dan sesuai dengan kebutuhan penilaian.

Selain fine-tuning, integrasi teknik augmentasi data dan pemrosesan tambahan dapat meningkatkan kemampuan BERT dalam menangani dokumen panjang dan kompleks. Chen & Liu (2022) mengemukakan bahwa BERT dapat mengalami keterbatasan dalam memproses dokumen yang sangat panjang. Oleh karena itu, teknik seperti pembagian dokumen menjadi bagian-bagian yang lebih kecil atau penerapan teknik pemrosesan tambahan untuk menyaring informasi relevan sebelum pemrosesan oleh BERT dapat diterapkan. Dengan metode ini, sistem peringkasan dapat lebih efektif dalam menangani dokumen panjang dari Proposal SBK dan Ujian Kualifikasi, memastikan bahwa informasi penting tidak terlewat dan ringkasan yang dihasilkan lebih komprehensif.

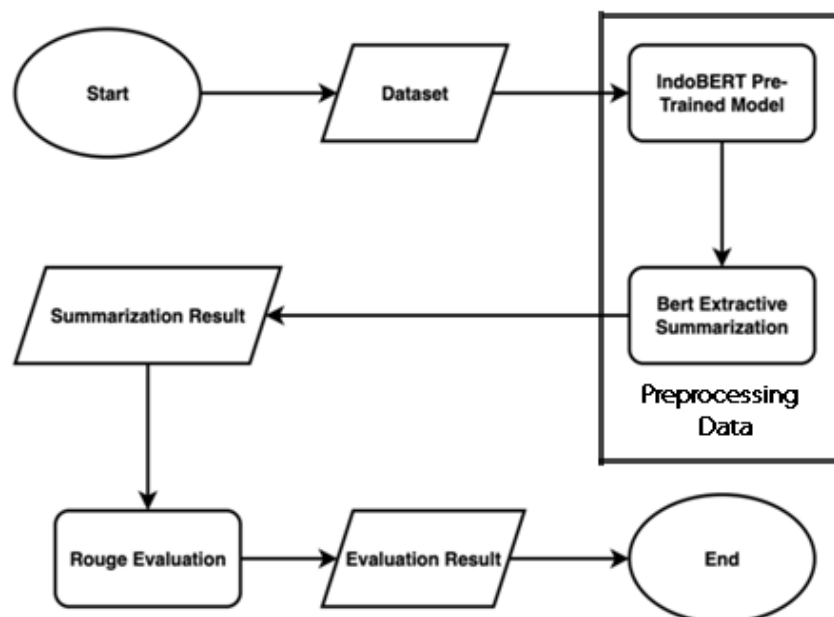
Terakhir, evaluasi dan adaptasi model dapat dilakukan dengan menerapkan metrik penilaian baru dan teknik evaluasi yang lebih canggih. Seperti yang dijelaskan oleh Li & Zhang (2022), pengembangan sistem peringkasan BERT tidak hanya bergantung pada kualitas model tetapi juga pada cara penilaian hasil ringkasan. Mengadopsi metrik penilaian yang lebih spesifik untuk dokumen akademik dan penilaian seperti metrik ROUGE yang diperluas atau metrik berbasis manusia dapat memberikan wawasan yang lebih baik mengenai performa model. Penelitian ini dapat menjelajahi teknik evaluasi baru untuk memastikan bahwa sistem peringkasan tidak hanya efisien tetapi juga efektif dalam memenuhi kebutuhan pengguna akhir, seperti panel penguji dalam Proposal SBK dan Ujian Kualifikasi.

## BAB III METODOLOGI PENELITIAN

### 1.1 Gambaran Umum

Peringkasan teks merupakan langkah krusial dalam pengolahan informasi yang memungkinkan kita untuk menyaring dan menyajikan informasi penting dari volume data yang besar secara ringkas dan informatif. Untuk mencapai tujuan ini, penelitian ini mengadopsi pendekatan berbasis BERT, yang terkenal karena kemampuannya dalam memahami konteks dan struktur bahasa secara mendalam, serta RAG, yang mengintegrasikan retrieval atau pencarian informasi dengan generation atau pembuatan teks.. Metodologi ini bertujuan untuk menggabungkan keunggulan dari kedua teknik tersebut dalam menghasilkan ringkasan yang tidak hanya akurat dan relevan tetapi juga informatif dan koheren. Selanjutnya, bab ini akan menjelaskan secara rinci bagaimana masing-masing metode diterapkan, diintegrasikan, serta evaluasi yang dilakukan untuk menilai efektivitas dan kinerja dari sistem peringkasan yang dikembangkan.

Berikut ini adalah gambaran metodologi penelitian yang akan diterapkan untuk peringkasan teks pada proposal hibah, ditunjukkan pada Gambar 3.1.



Gambar 3.1 Tahapan Penelitian

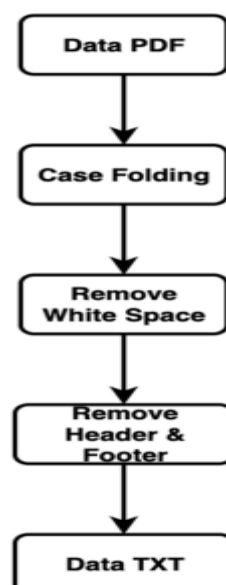
Untuk melakukan penelitian perlu dilakukan beberapa tahapan dari awal hingga selesai, tahapan yang dilakukan adalah mulai dari pengumpulan data, preprocessing data, kemudian menentukan hasil ringkasan, mengukur metric evaluasi, hingga hasil akhir.

### **1.1.1 Pengumpulan Dataset**

Proses pengumpulan data dilakukan dengan cara mengumpulkan dokumen proposal hibah dari lembaga penelitian sebuah instansi pendidikan. Data yang diperoleh adalah data dalam bentuk dokumen berkesktensi .PDF (Portable Document Format).

### **1.1.2 Preprocessing data**

Proses preprocessing data merupakan langkah untuk mempersiapkan data, proses ini melibatkan beberapa tahapan yang bertujuan untuk membersihkan dan mempersiapkan teks agar sesuai dengan kebutuhan model. Tahap ini terdiri dari tahap pengubahan dokumen dari ekstensi .PDF ke ekstensi .TXT, hal ini dilakukan untuk mengekstraksi teks tersebut sehingga dapat dilanjutkan pada tahap berikutnya. Berikut pada gambar 3.2 adalah langkah untuk mengkonversi data dokumen PDF ke dalam TXT.



Gambar 3.2 Mengkonversi file PDF ke TXT



Gambar 3.2 menggambarkan sebuah proses konversi data dari format PDF menjadi format teks (TXT). Proses ini umumnya dilakukan dalam tahap-tahap preprocessing data sebelum data tersebut digunakan untuk analisis lebih lanjut, seperti analisis sentimen, pengenalan entitas, atau pemodelan topik. Berikut penjelasan dari tiap-tiap tahapan:

1. **Data PDF:** Proses dimulai dengan mempersiapkan dokumen proposal PKM dalam format PDF.
2. **Case Folding:** Pada tahap ini, semua huruf dalam teks diubah menjadi huruf kecil. Tujuannya adalah untuk menyederhanakan data dan menghindari masalah karena perbedaan kapitalisasi. Misalnya, kata "Program" dan "program" dianggap sama setelah proses case folding. Proses ini menggunakan fungsi dan modul tersedia dalam Python.
3. **Remove White Space:** Tahap ini menghilangkan spasi berlebih, tab, dan karakter kosong lainnya dari teks. Tujuannya adalah untuk membersihkan teks dan membuatnya lebih mudah diproses.
4. **Remove Header & Footer:** Header dan footer seringkali berisi informasi tambahan yang tidak relevan dengan isi utama dokumen, seperti nomor halaman, tanggal, atau judul dokumen. Tahap ini menghilangkan bagian header dan footer tersebut untuk fokus pada isi teks utama.
5. **Data TXT:** Setelah melalui semua tahap preprocessing, data akhir disimpan dalam format teks (TXT). Format TXT adalah format yang sederhana dan mudah dibaca oleh berbagai program, sehingga memudahkan untuk melakukan analisis lebih lanjut.

Data yang sudah dalam format TXT kemudian diproses kembali sebelum dimasukkan ke dalam model BERT. Pra-pemrosesan kedua tahapan dalam penelitian ini dibagi menjadi dua jenis. Versi 1 digunakan pra-pemrosesan, termasuk dalam perpustakaan IndoBERT, dan versi 2 menambahkan *stemming* dan *case folding* ke pra-pemrosesan yang ada. Prapemrosesan berbasis arsitektur BERT yang dilatih khusus untuk bahasa Indonesia dengan model IndoBERT, kemudian setelah itu dilakukan prapemrosesan selanjutnya menggunakan model BERT. Berikut ini langkah pada bagian prapemrosesan.

1. **Tokenisasi:** Langkah pertama dalam sebagian besar langkah pekerjaan pemrosesan teks, tokenisasi bertugas memisahkan string teks lengkap menjadi daftar kata terpisah. Studi lain mendefinisikan tokenisasi sebagai jenis leksikal analisis yang memecah teks menjadi kata, frasa, symbol atau elemen bermakna lainnya yang disebut token.
2. **Penyematan Kata:** IndoBERT dilatih dengan 4 miliar kosakata dan 250 juta kalimat digabungkan dalam kumpulan data Indo4B [25]. Pertama, kumpulan data ini digunakan untuk membangun model fastText. Kemudian penyematan fastText dilatih sebelumnya menggunakan representasi kata skip-gram dan dibuat vektor penyematan 300 dimensi. Lalu semua penyematan yang diperlukan untuk setiap tugas dibuat dari FastText yang telah dilatih sebelumnya dan menyertakan semua kosakata. Hasil tokenisasi kemudian diubah menjadi a vektor menggunakan penyematan kata.

### 1.1.3 BERT Extractive Summarization

Proses peringkasan menggunakan BERT dibagi menjadi dua siklus. Yang pertama adalah merangkum dokumen dengan identitas, sedangkan yang kedua melibatkan ringkasan dokumen tanpa identitas. Dokumen dibagi menjadi dua jenis karena identitas adalah informasi paten yang tidak mungkin diringkas. Pemisahan dilakukan untuk melihat kinerja BERT dalam mengekstraksi informasi yang bermakna dalam dokumen tersebut kecuali ada identitas di dalamnya.

### 1.1.4 Hasil Peringkasan

Rangkuman hasil A (dokumen dengan identitas) dan B (dokumen tanpa identitas) terlihat hasil ekstrak poin-poin penting dalam keputusan dokumen. ROUGE Evaluation: Pengukuran hasil peringkasan dilakukan menggunakan Pemahaman Berorientasi Ingatan untuk Evaluasi Inti (ROUGE). Skor ROUGE akan mengukur kesamaan antara dua objek ringkasan yang dihasilkan.

### 1.1.5 Evaluasi Hasil

Skenario evaluasi yang digunakan dalam penelitian ini dimaksudkan untuk uji setiap model terlatih untuk setiap rasio ringkasan pada kedua jenis dokumen. Jumlah NGRAM yang digunakan adalah  $N=4$ . Pengujian dilakukan dengan menggunakan ROUGE-N dan dilakukan pada kedua kumpulan data dokumen. Dok A mewakili dokumen dengan identitas, dan Dok B adalah dokumen tanpa identitas. Yang ideal skor ROUGE diidentifikasi sebagai skor ROUGE seimbang antara Dokumen A dan Dokumen B, terjadi secara bersamaan rasio ringkasan dan jumlah NGRAM yang sama.

### 1.2 Jadwal Penelitian

Untuk memastikan penelitian berjalan sesuai dengan target maka dibuat jadwal penelitian agar penelitian dapat diselesaikan sesuai dengan batas waktu yang telah ditetapkan. Tabel jadwal penelitian dapat dilihat pada table 3.1 berikut ini.

No	Uraian Kegiatan	2018				2019		2024											
		9	10	11	12	1	2	1	2	3	4	5	6	7	8	9	10	11	12
1	Penyusunan Proposal																		
2	Uji Kualifikasi																		
3	Evaluasi Progres Pertama																		
4	Paper Pertama																		
5	Evaluasi Progres Kedua																		
No	Uraian Kegiatan	2025												2026					
		1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4		
1	Paper Kedua																		
2	Evaluasi RKP																		
3	Sidang Tertutup																		
4	Sidang Terbuka																		

## DAFTAR PUSTAKA

- Amaratunga, T. (2023). Understanding Large Language Models. In *Understanding Large Language Models*. Apress. <https://doi.org/10.1007/979-8-8688-0017-7>
- Ardiansyah, H., & Nurdin, M. (2021). Penggunaan Teknologi Pemrosesan Bahasa Alami pada Penilaian Proposal Hibah\*. *Jurnal Teknologi dan Sistem Komputer (JTSISK)*.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2018)*, 4171–4186. <https://doi.org/10.5555/3327763.3327780>
- Gardner, S. K., & Nakhleh, M. B. (2008). The Role of the Dissertation Proposal in the PhD Process: An Analysis of 18,000 PhD Proposals. *Educational Researcher*, 37(8), 513-522.
- Gibbs, G. (2014). *The Evaluation of PhD Research Proposals*. *Journal of Research in Postgraduate Education*, 25(1), 25-32.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Jung, S., & Wang, H. (2021). Understanding the Role of PhD Qualifying Examinations in the Doctoral Journey. *International Journal of Doctoral Studies*, 16, 217-235.
- Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.). Pearson.
- Kumar, A., & Ghosh, S. (2020). A Survey on Text Summarization Techniques: A Comprehensive Review. *International Journal of Computer Applications*, 176(12), 1-7.
- Kumar, A., Singh, R., & Patel, M. (2022). Evaluating the Effectiveness of Proposal Seminars in Enhancing Research Skills: A Case Study. *Journal of Higher Education Research*, 48(2), 112-130.

Kumar, R., & Lee, J. (2021). Evaluating Text Summarization Algorithms for Grant Proposal Assessment. *Proceedings of the 2021 International Conference on Computational Linguistics (COLING 2021)*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>

Liu, Y., Ott, M., Goyal, N., Du, J., & Joshi, M. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.

Lovitts, B. E. (2007). *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study*. Rowman & Littlefield.

Nallapati, R., Zhai, F., & Zhou, B. (2016). Abstractive text summarization using sequence-to-sequence RNNs and beyond. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*, 280–290.

Naveed, M., Hussain, S., & Liu, Y. (2023). Tokenization strategies for large language models. *Journal of Natural Language Engineering*, 29(3), 543–560. <https://doi.org/10.1017/S1351324923000123>

Nguyen, T., & Zhang, X. (2022). Using Text Summarization for Efficient Grant Proposal Evaluation. *Proceedings of the 2022 Conference on Computational Natural Language Learning (CoNLL 2022)*.

Patel, S., & Saha, A. (2022). Enhancing Grant Proposal Review with Automatic Text Summarization. *Journal of Machine Learning Research (JMLR)*.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2019). Improving language understanding by generative pre-training. *OpenAI*. <https://www.openai.com/research/language-unsupervised>

Smith, L., Thompson, A., & Garcia, C. (2020). Assessing Doctoral Progress: The Impact of Qualifying Exams on PhD Success. *Educational Assessment, Evaluation and Accountability*, 32(4), 481–496.

Smith, R., & Johnson, M. (2020). Best Practices in Grant Proposal Writing. *Journal of Research Administration*, 51(1), 45–58.

Wang, L., & Wang, Y. (2020). Automatic Summarization of Research Proposals for Funding Agencies. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*.

Webster, J., & Kit, R. (n.d.). Tokenization techniques and their applications. *In Proceedings of the Workshop on Language Resources and Evaluation*. Retrieved from <https://www.lrec-conf.org/proceedings/lrec2022/tokenization>

Widodo, S., & Maulana, A. (2022). Teknik Ringkasan Teks untuk Evaluasi Proposal Penelitian. *Jurnal Ilmu Komputer dan Informatika (JIKI)*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS 2017)*, 5998–6008. <https://doi.org/10.5555/3295222.3295345>

Wu, Y., Schuster, M., Chen, Z., & Le, Q. V. (2016). *Google's neural machine translation system: Bridging the gap between human and machine translation*. arXiv preprint arXiv:1609.08144. <https://arxiv.org/abs/1609.08144>

Zhang, Q., & Zhao, L. (2021). Text Summarization Techniques for Grant Proposal Review. *Journal of Artificial Intelligence Research (JAIR)*.

Zhang, Y., Zhao, Y., & Zhang, Y. (2020). A Comprehensive Review on Text Summarization: Techniques, Evaluation, and Applications. *Information Processing & Management*, 57(3), 102273.