

**SISTEM PEMBUAT GAMBAR SAMPUL OTOMATIS DENGAN MODEL  
*STABLE DIFFUSION* PADA RINGKASAN TEKS MENGGUNAKAN  
MODEL BERT (*BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM  
TRANSFORMERS*)**

**SKRIPSI**

**NIKO DHARMAWAN**

**191401102**



**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS SUMATERA UTARA  
MEDAN  
2023**

SISTEM PEMBUAT GAMBAR SAMPUL OTOMATIS DENGAN MODEL  
*STABLE DIFFUSION* PADA RINGKASAN TEKS MENGGUNAKAN  
MODEL BERT (*BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM  
TRANSFORMERS*)

SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah  
Sarjana Ilmu Komputer

NIKO DHARMAWAN

191401102



PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS SUMATERA UTARA  
MEDAN  
2023

# PERSETUJUAN

Judul : SISTEM PEMBUAT GAMBAR SAMPUL OTOMATIS DENGAN MODEL STABLE DIFFUSION PADA RINGKASAN TEKS MENGGUNAKAN MODEL BERT (BIDIRECTIONAL REPRESENTATIONS FROM TRANSFORMERS)

Kategori : SKRIPSI

Nama : NIKO DHARMAWAN

Nomor Induk Mahasiswa : 191401102

Program Studi : SARJANA (S1) ILMU KOMPUTER

Fakultas : ILMU KOMPUTER DAN TEKNOLOGI INFORMASI UNIVERSITAS SUMATERA UTARA

Telah diuji dan dinyatakan lulus di Medan, 6 Desember 2023

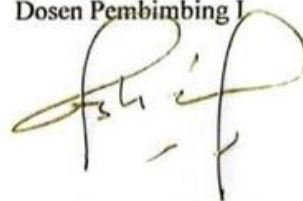
Dosen Pembimbing II



Hayatunnufus S.Kom, M.Cs

NIP. 199207192020012001

Dosen Pembimbing I



Amer Sharif S.Si, M.Kom.

NIP. 196910212021011001

Diketahui/disetujui oleh

Program Studi S1 Ilmu Komputer



Dr. Amalia, S.T., M.T.

NIP. 197812212014042001

**PERNYATAAN**

**SISTEM PEMBUAT GAMBAR SAMPUL OTOMATIS DENGAN MODEL  
STABLE DIFFUSION PADA RINGKASAN TEKS MENGGUNAKAN  
MODEL BERT (BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM  
TRANSFORMERS)**

**SKRIPSI**

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan,



Niko Dharmawan

191401102`

## UCAPAN TERIMA KASIH

Puji syukur kehadiran Tuhan Yang Maha Esa atas anugerah, berkat dan karunia yang telah dilimpahkan kepada penulis sehingga penulis dapat menyelesaikan skripsi ini yang merupakan salah satu persyaratan untuk memperoleh gelar Sarjana Komputer dari program studi S-1 Ilmu Komputer Universitas Sumatera Utara.

Pada kesempatan ini penulis ingin memberikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah mendukung dan membantu pengerjaan skripsi ini. Penulis mengucapkan terima kasih kepada:

1. Tuhan Yang Maha Esa yang senantiasa menyertai dan memberikan berkat Kesehatan kepada penulis baik dalam pembuatan skripsi ini.
2. Bapak Dr. Muryanto Amin, S.Sos., M.Si. selaku Rektor Universitas Sumatera Utara
3. Ibu Dr. Maya Silvi Lydia B.Sc., M.Sc. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.
4. Ibu Dr. Amalia S.T., M.T. selaku Ketua Program Studi S-1 Ilmu Komputer Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.
5. Ibu Dewi Sartika Br Ginting S.Kom., M.Kom selaku Dosen Pembimbing Akademik penulis di Program Studi S-1 Ilmu Komputer.
6. Bapak Amer Sharif S.Si, M.Kom. selaku Dosen Pembimbing I yang telah membimbing penulis dalam penyusunan skripsi ini.
7. Ibu Hayatunnufus S.Kom, M.Cs selaku Dosen Pembimbing II yang telah membimbing dan memberikan saran dan masukan kepada penulis dalam penyusunan skripsi ini.
8. Bapak/Ibu staff pengajar serta seluruh pegawai Program Studi S-1 Ilmu Komputer Fasilkom-TI USU.
9. Orang tua dan keluarga yang selalu mendoakan dan memberikan dukungan baik dari segi moral maupun material.

10. Sahabat-sahabat saya Ivan, Sinu, Ricky, Filbert, Harris, Prisko, Vinny, Christine, Isaac, Andre, Steven, Javin, Ericko, Billy, Kevin, Vanessa, Wynona, dan Wynola yang telah menemani penulis dalam proses penulisan skripsi ini.

11. Serta seluruh pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat disebutkan secara rinci.

Demikianlah skripsi ini dibuat, semoga bermanfaat bagi pembaca pada umumnya dan kepada penulis khususnya. Penulis menyadari bahwa masih banyak terdapat kekurangan yang ada di dalam skripsi ini. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari semua pihak yang membaca skripsi ini dan untuk itu penulis mengucapkan terima kasih.

Medan,



Penulis

## ABSTRAK

Konteks penelitian ini terinspirasi oleh pergeseran paradigma dalam konsumsi berita, di mana masyarakat beralih dari sumber berita tradisional ke artikel *online* dan blog, namun sering kali menghadapi kendala ketika artikel terlalu panjang untuk dibaca secara menyeluruh. Kondisi literasi yang rendah di Indonesia menjadi latar belakang tambahan yang memperumit pengiriman informasi yang efektif. Dalam upaya mengatasi masalah ini, sistem peringkasan teks dianggap sebagai solusi potensial untuk menyajikan informasi secara ringkas namun tetap mempertahankan esensi teks. Selain itu, pembuatan sampul artikel juga menjadi aspek penting dalam menarik perhatian pembaca, serta memperkenalkan topik artikel dengan efektif. Sampul artikel yang sesuai harus mencerminkan isi artikel secara akurat. Dalam penelitian ini dikembangkan sebuah sistem peringkasan teks dan pembuat gambar berbasis model BERT dan *Stable Diffusion*. Sistem yang dihasilkan dapat memanfaatkan model BERT untuk meringkas teks tanpa kehilangan makna utama teks dengan hasil ringkasan teks yang mencapai rata-rata  $r = 0.29$ ,  $p = 0.93$ , dan  $f = 0.31$  yang kemudian diterjemahkan menjadi bahasa Inggris melalui *GoogleTranslate*. Lalu, hasil terjemahan digunakan sebagai panduan model *Stable Diffusion* untuk membuat sebuah gambar yang akan digunakan sebagai sampul. Sistem ini memiliki beberapa batasan, seperti jumlah token teks yang terbatas, ringkasan yang kekurangan konteks, dan gambar yang terkadang mengandung tulisan atau huruf-huruf acak yang tidak memiliki makna. Dengan demikian, penelitian ini memberikan kontribusi dalam menghadirkan solusi yang dapat membantu pembaca dalam mengakses informasi melalui ringkasan teks, sekaligus memberikan manfaat bagi para penulis dan jurnalis *online* yang membutuhkan gambar sampul artikel yang relevan dengan topik yang mereka bahas.

**Kata Kunci:** *Deep Learning, Transformer, Bidirectional Encoder Representations from Transformers, self-attention, ringkasan, ekstraktif, Stable Diffusion, text-to-image*

*AUTOMATIC COVER IMAGE GENERATION SYSTEM WITH  
STABLE DIFFUSION ON TEXT SUMMARY USING  
BERT (BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM  
TRANSFORMERS)*

**ABSTRACT**

*The context of this research is inspired by the paradigm shift in news consumption, where people are shifting from traditional news sources to online articles and blogs, but often face obstacles when articles are too long to read in full. The low literacy conditions in Indonesia are an additional background that complicates the effective delivery of information. In an effort to overcome this problem, text summarization systems are considered a potential solution to present information in a concise but still maintain the essence of the text. In addition, article cover making is also an important aspect in attracting reader attention, as well as introducing article topics effectively. Appropriate article covers should accurately reflect the content of the article. In this research, a text summarization and image generation system based on the BERT and Stable Diffusion models was developed. The resulting system can utilize the BERT model to summarize text without losing the main meaning of the text with text summary results reaching an average  $r = 0.29$ ,  $p = 0.93$ , and  $f = 0.31$  which will then be translated into English through GoogleTranslate. The translation results are then used as a guide for the Stable Diffusion model to create an image that will be used as a cover. This system has some limitations, such as the limited number of text tokens, summaries that lack context, and images that sometimes contain text or random letters that have no meaning. Thus, this research contributes to presenting a solution that can help readers access information through text summaries, while also providing benefits for online writers and journalists who need article cover images that are relevant to the topics they discuss.*

**Keywords:** *Deep Learning, Transformer, Bidirectional Encoder Representations from Transformers, self-attention, extractive text summarization, Stable Diffusion, text-to-image*



## DAFTAR ISI

<b>PERSETUJUAN.....</b>	<b>Error! Bookmark not defined.</b>
<b>PERNYATAAN.....</b>	<b>iii</b>
<b>UCAPAN TERIMA KASIH .....</b>	<b>iv</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>viii</b>
<b>DAFTAR GAMBAR.....</b>	<b>x</b>
<b>DAFTAR TABEL .....</b>	<b>xi</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian .....	4
1.5. Manfaat Penelitian .....	4
1.6. Metodologi Penelitian .....	5
1.7. Sistematika Penulisan .....	6
<b>BAB II LANDASAN TEORI .....</b>	<b>6</b>
2.1. <i>Deep Learning</i> .....	7
2.1.1. <i>Transformer</i> .....	7
2.1.2. <i>Bidirectional Encoder Representation for Transformer (BERT)</i> .....	8
2.1.3. <i>Diffusion Models</i> .....	10
2.2. <i>Text Summarization</i> .....	13
2.2.1. <i>ROUGE</i> .....	14
<b>BAB III ANALISIS DAN PERANCANGAN SISTEM.....</b>	<b>15</b>
3.1. Penjelasan Sistem.....	15
3.2. Sumber <i>Dataset</i> .....	17
3.3. Pra-proses Gambar .....	17
3.4. Pelatihan Model <i>Stable Diffusion</i> dengan <i>Dreambooth</i> .....	17
3.5. Ringkasan Ekstraktif dengan BERT .....	17
3.6. Evaluasi .....	18
<b>BAB IV IMPLEMENTASI DAN PENGUJIAN .....</b>	<b>22</b>

4.1.	Perangkat yang Digunakan .....	22
4.1.1.	Perangkat Keras .....	22
4.1.2.	Perangkat Lunak.....	22
4.1.3.	Library.....	22
4.2.	Penerapan Tahap Pengumpulan <i>Dataset</i> .....	23
4.3.	Penerapan Tahap Pra-proses <i>Dataset</i> .....	24
4.4.	Penerapan Tahap Pelatihan <i>Stable Diffusion</i> dengan <i>Dreambooth</i> .....	25
4.5.	Penerapan Tahap Peringkasan Ekstraktif.....	28
4.6.	Penerapan Tahap Evaluasi .....	38
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>43</b>
5.1.	Kesimpulan .....	43
5.2.	Saran.....	44
<b>DAFTAR PUSTAKA .....</b>		<b>45</b>

X

**DAFTAR GAMBAR**

<b>Gambar 2.1</b> Arsitektur sebuah model difusi.....	11
<b>Gambar 3.1</b> Arsitektur Umum Sistem. ....	16
<b>Gambar 3.2</b> Flowchart Sistem. ....	18

**DAFTAR TABEL**

<b>Tabel 4.1</b> Dataset gambar sebelum di proses. ....	23
<b>Tabel 4.2</b> Contoh gambar-gambar yang telah diproses. ....	24
<b>Tabel 4.3</b> Tabel ringkasan. ....	29
<b>Tabel 4.4</b> Tabel translasi ringkasan. ....	36
<b>Tabel 4.5</b> Tabel inference. ....	40

## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Pada akhir tahun 2019 hingga saat ini, kertas koran yang merupakan sumber berita konvensional mulai ditinggalkan dikarenakan pandemi. Sebagian besar masyarakat sekarang lebih memilih untuk membaca berita melalui artikel *online* atau *blog*. Namun, masalah muncul ketika beberapa artikel atau blog yang ditawarkan terlalu panjang untuk dibaca. Selain itu, menurut penelitian yang dilakukan oleh UNESCO, literasi di negara Indonesia hanya mencapai 0,001%. Berarti hanya 1 dari 1000 orang yang memiliki keinginan untuk membaca. Oleh karena itu, ringkasan menjadi solusi yang efektif dalam menyajikan karangan atau teks yang panjang dalam bentuk yang singkat, tanpa menghilangkan ide dan pendekatan asli pengarang (Halim, F, et al., 2022).

Di sisi lain, sampul artikel pada sebuah berita juga menjadi ketertarikan sendiri bagi pembaca. Pembuatan sampul artikel juga menjadi hal penting bagi penulis dalam menarik perhatian pembaca dan memperkenalkan topik artikel. Selain itu, konten dalam artikel harus relevan dan informatif dengan topik yang dibahas dalam sampul artikel tersebut. Sebuah artikel yang berkualitas dengan sampul artikel yang baik akan memperlihatkan kualitas dan profesionalitas penulis serta dapat meningkatkan minat pembaca untuk membaca seluruh artikel.

*Automatic text summarization* (peringkas teks otomatis) menghasilkan ringkasan yang berisi kalimat yang mencakup semua informasi relevan dari dokumen asli misalnya artikel, berita, jurnal, dan lain-lain agar informasi yang dibutuhkan dapat diterima atau dimengerti dengan cepat tanpa kehilangan makna utama (Allahyari, M, et al., 2017).

Dalam meringkas teks terdapat dua metode utama yaitu abstraktif dan ekstraktif. Ringkasan abstraktif menghasilkan ringkasan dengan kalimat atau kata yang ringkas dan koheren yang tidak hanya mengekstrak kalimat dari dokumen asli (Erkan, G, et al., 2004), sedangkan ringkasan ekstraktif mengekstrak kalimat yang secara kolektif mewakili informasi yang paling penting atau relevan dalam konten

asli (Kyoomarsi, F, et al., 2008). Pada penelitian ini penulis memutuskan untuk menggunakan metode ekstraktif.

Salah satu model *deep learning* yang terbukti efektif dalam menangani tugas-tugas *Natural Language Processing* (NLP) adalah *Bidirectional Encoder Representations from Transformers* (BERT). BERT memiliki 6 lapisan *Transformer* yang terdiri dari *encoder* dan *decoder*, yang memungkinkannya untuk menangani proses pelatihan yang kompleks dan konfigurasi yang tinggi. Namun, kompleksitas pelatihan BERT menghasilkan biaya yang tinggi dan waktu yang banyak. Proses BERT dimulai dengan representasi *embedding* kata pada lapisan *embedding*. Setiap lapisan BERT melakukan perhitungan *multi-headed attention* pada representasi kata dari lapisan sebelumnya untuk membuat representasi perantara baru (Devlin, J, et al., 2018).

Dalam BERT, perhitungan *multi-headed attention* memainkan peran penting dalam memperoleh representasi yang akurat dari teks yang kompleks. Setiap lapisan BERT melakukan beberapa perhitungan *multi-headed attention* pada representasi kata dari lapisan sebelumnya untuk menghasilkan representasi perantara yang lebih baik. Hal ini memungkinkan BERT untuk memperoleh pemahaman yang lebih dalam tentang konteks dan makna di balik kata-kata dalam teks.

Penelitian yang dilakukan oleh Althar Mohammed (2021) menunjukkan bahwa kemampuan utama BERT adalah *self-attention*. BERT dapat membaca teks dari kedua arah, sehingga model ini dapat memahami teks berdasarkan konteks kalimat sebelum dan sesudahnya. Selain itu, model BERT juga dapat mempelajari kata yang tidak dikenali dengan mengubahnya menjadi sub-kata. Hal ini telah terbukti efektif dalam mengatasi keterbatasan *Recurrent Neural Network* dalam memproses kata demi kata secara berurutan, di mana pada metode BERT dapat memproses seluruh input kata secara bersamaan (Yang, L, et al., 2019). Berdasarkan penelitian yang telah dilakukan oleh Yang Gu dan Yanke Hu (2019), teks hasil ringkasan dari model BERT lebih mudah dibaca dan lebih relevan dengan topik artikel.

*Stable Diffusion* adalah model difusi *text-to-image* laten yang telah dilatih menggunakan gambar-gambar berukuran 512x512 dari database LAION-5B (Rombach, R, et al., 2022). Model *Stable Diffusion* biasanya digunakan untuk menghasilkan gambar yang dikondisikan pada deskripsi teks (*prompt*). Namun, model ini juga dapat diterapkan untuk tugas lain seperti *inpainting*, *outpainting*, dan *image-to-image* yang dipandu oleh *prompt* (Rombach, R, et al., 2022).

Sistem yang akan dibuat oleh penulis dapat membantu jurnalis dan pembaca dengan mengotomatisasi proses merangkum dan membuat gambar yang dapat digunakan sebagai sampul. User memasukkan isi artikel ke dalam sistem dan sistem melakukan peringkasan. Selanjutnya, hasil ringkasan akan dikirim dalam bentuk *prompt* pada model *Stable Diffusion* yang akan menghasilkan gambar sampul untuk artikel tersebut.

Kombinasi antara *text summarization* pada model BERT dan *text-to-image* pada model *Stable Diffusion* diharapkan dapat menghasilkan gambar sampul yang menarik dan sesuai dengan topik artikel.

## **1.2. Rumusan Masalah**

Pembaca teks artikel *online* sering sekali merasa bosan dengan sajian informasi yang terlalu panjang dikarenakan informasi yang diberikan sering kali bertele-tele. Untuk menyelesaikan permasalahan ini dibutuhkan suatu pembuatan ringkasan terhadap informasi-informasi yang tersedia secara *online*. Teks artikel *online* juga membutuhkan gambar yang relevan untuk mendukung informasi yang disampaikan, sering kali jurnalis terkendala dengan, oleh karena itu diperlukan sebuah model *Artificial Intelligence* dalam membuat sebuah gambar dari teks yang sudah disusun oleh jurnalis.

### 1.3. Batasan Masalah

Dalam penelitian ini ada beberapa batasan, yaitu:

1. Teks hanya bisa dimasukkan secara manual.
2. Sistem dibangun menggunakan bahasa pemrograman *Python*.
3. Metode yang digunakan untuk meringkas teks menggunakan metode ekstraktif
4. Teks yang dimasukkan harus dibatasi sebanyak 512 token.
5. Teks deskriptif atau *prompt* yang dimasukkan melewati *Stable Diffusion* harus dalam bahasa Inggris.
6. Sistem membutuhkan minimal 8GB VRAM ketika menghasilkan gambar.

### 1.4. Tujuan Penelitian

BERT merupakan salah satu model *deep learning* yang mengadopsi konsep *self-attention* dengan pendekatan *Transformer*. Konsep ini membantu BERT untuk memahami konteks sebuah kalimat. Dengan menggunakan model BERT diharapkan informasi-informasi yang tersedia secara *online* dapat diringkas dengan baik tanpa kehilangan isi teks sehingga waktu membaca dapat disingkat. *Stable Diffusion* adalah model difusi *text-to-image* laten yang bisa membuat sebuah gambar berdasarkan teks yang diinputkan atau dikenal dengan *prompt*. Dengan menggunakan kedua model ini diharapkan proses meringkas dan pembuatan sampul artikel dapat dipercepat dan dipermudah.

### 1.5. Manfaat Penelitian

Penelitian ini dapat membantu para pembaca dalam menerima informasi melalui hasil ringkasan dan jurnalis, terutama *online journalists* dan *blogger* yang biasanya memerlukan gambar untuk digunakan sebagai sampul yang tidak hanya berhubungan dengan isi teks tetapi juga dapat menarik perhatian banyak pembaca terhadap teks mereka.



## 1.6. Metodologi Penelitian

Metodologi penelitian yang digunakan adalah sebagai berikut:

### 1. Studi Pustaka

Pada tahap ini penelitian dimulai dengan mencari referensi dari berbagai sumber terpercaya dan melakukan peninjauan pustaka melalui buku-buku, *e-book*, jurnal, artikel ilmiah yang berhubungan dengan *text-summarization*, BERT (*Bidirectional Encoder Representation for Transformer*), dan juga Stable Diffusion.

### 2. Analisa dan Perancangan

Berdasarkan ruang lingkup penelitian, penulis melakukan analisa terhadap apa saja yang dibutuhkan dalam penelitian untuk segera dirancang dalam sebuah diagram alir (*flowchart*).

### 3. Implementasi

Pada tahap ini, membuat sebuah sistem dengan menggunakan bahasa pemrograman python sesuai dengan diagram alir yang telah dirancang.

### 4. Pengujian

Pada tahap ini, sistem yang telah dirancang akan diuji coba untuk mengetahui apakah sistem berjalan sesuai tujuan.

### 5. Dokumentasi

Pada tahap ini dilakukanlah dokumentasi mulai dari tahap analisa sistem sampai kepada pengujian sistem dalam format penulisan berbentuk skripsi.

## **1.7. Sistematika Penulisan**

Sistematika penulisan skripsi ini dibagi dalam beberapa bab, yaitu:

### **BAB 1 PENDAHULUAN**

Bab pendahuluan membahas mengenai latar belakang penelitian, rumusan masalah, batasan masalah penelitian, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan skripsi.

### **BAB 2 LANDASAN TEORI**

Bab landasan teori membahas tentang penjelasan mengenai *Deep Learning*, *Transformer*, *BERT*, *Diffusion Models*, *Stable Diffusion*, *Text Summarization*, dan penelitian-penelitian yang relevan.

### **BAB 3 ANALISIS DAN PERANCANGAN SISTEM**

Bab analisis dan perancangan sistem membahas tentang analisis dari masalah-masalah penelitian dan solusi yang diberikan untuk permasalahan tersebut dengan sistem perancangan sistem yang akan dibangun.

### **BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM**

Bab implementasi dan pengujian sistem membahas tentang implementasi dari sistem yang telah dirancang dan hasil dari pengujian sistem yang berfungsi untuk memeriksa kesesuaian dan keberhasilan sistem.

### **BAB 5 KESIMPULAN DAN SARAN**

Bab kesimpulan dan saran berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran dari hasil pengujian sehingga dapat diterapkan menjadi masukan dalam penelitian selanjutnya.

## **BAB II**

### **LANDASAN TEORI**

#### ***2.1. Deep Learning***

*Deep learning* adalah salah satu algoritma dalam perkembangan *machine learning*. Berbeda dengan *machine learning*, *deep learning* memiliki lapisan unit pemrosesan nonlinier yang terdiri dari beberapa *layer*, yang bisa digunakan untuk transformasi dan ekstraksi fitur dari data. Setiap lapisan pada *deep learning* menggunakan *output* dari lapisan sebelumnya sebagai *input*, sehingga *deep learning* mampu mempelajari representasi tingkat abstraksi yang lebih kompleks. (Zhang, et al., 2018)

Algoritma *deep learning* memiliki aplikasi yang luas dalam berbagai bidang, seperti dalam pengenalan wajah, klasifikasi citra, dan bahkan dalam prediksi harga saham. Salah satu keunggulan dari algoritma *deep learning* adalah kemampuannya untuk mengerti hubungan antar variabel serta pengetahuan yang terkait dengan hubungan tersebut. Dalam studi Deng, et al., (2014), *deep learning* mampu mempelajari tingkat representasi yang sesuai dengan tingkat abstraksinya. Sebagai contoh, pada tugas pengenalan gambar, *deep learning* dapat mempelajari representasi visual dari gambar, yang berbeda dari tugas *machine learning* konvensional yang lebih bergantung pada fitur yang diambil secara manual. Dalam hal ini, *deep learning* membantu dalam pemodelan dan analisis data yang lebih kompleks dan lebih akurat, yang tidak mungkin dilakukan dengan pendekatan konvensional.

##### ***2.1.1. Transformer***

Dalam studi Vaswani, A., et al. (2017), *Transformer* merupakan salah satu model *Deep learning* yang telah banyak digunakan di berbagai bidang seperti *Computer Vision* (CV), *Natural Language Processing* (NLP), dan *speech processing*. Model ini memiliki kemampuan untuk mencapai kinerja state-of-the-art dalam berbagai tugas NLP, sehingga menjadi arsitektur yang sering digunakan dalam bidang ini.

Salah satu keunggulan dari arsitektur *Transformer* adalah adanya *multi-head self-attention* dalam arsitekturnya. *Self-attention* ini memungkinkan *Transformer* untuk memahami kalimat secara keseluruhan, bukan hanya satu kata atau frase dalam kalimat tersebut. Hal ini membuat *Transformer* lebih mampu memahami konteks

dan hubungan antar kata dalam suatu kalimat. Saat ini, *Transformer* banyak dikembangkan dalam berbagai model, seperti BERT yang merupakan salah satu model *Transformer* terbaru yang dianggap sangat sukses dalam memecahkan berbagai masalah NLP. BERT telah terbukti memberikan hasil yang sangat baik dalam tugas-tugas seperti pemahaman bahasa alami, penerjemahan bahasa, dan banyak lagi.

### **2.1.2. *Bidirectional Encoder Representation for Transformer (BERT)***

Dalam studi Athar Mohammed & Ali H. Ali (2021) BERT atau *Bidirectional Encoder Representations for Transformers* merupakan salah satu pengembangan dari model *Transformer* dalam bidang *Natural Language Processing (NLP)*. Berbeda dengan *Transformer* yang menggunakan arsitektur *encoder* dan *decoder*, BERT hanya menggunakan arsitektur *encoder* untuk memproses data teks. *Encoder* tersebut terdiri atas 2 bagian *self-attention* dan *feed-forward neural network*.

BERT memiliki 2 konsep utama yaitu :

#### **1. *Mask Language Model (MLM)***

Untuk melatih model ini, BERT melakukan penggantian pada 15% kata dalam setiap kalimat dengan menggunakan token *[MASK]* yang disebut dengan *masking*, kemudian BERT akan berusaha memprediksi token yang sebenarnya berdasarkan konteks kalimat tersebut.

#### **2. *Next Sentence Prediction (NSP)***

BERT berupaya memahami hubungan antara dua kalimat. Dalam proses tersebut, input akan melalui dua tahap, yaitu pengubahan kalimat menjadi *sentence embedding* dan *embedding* posisi.

Pada penelitian yang dilakukan oleh Yang Liu dan Mirella Lapata (2019), diperlihatkan bahwa BERT yang telah dilatih sebelumnya dapat diterapkan dalam *text-summarization*. Hasil eksperimen pada tiga *dataset* menunjukkan bahwa model BERT mencapai hasil terbaik atau *state-of-the-art* di seluruh bidang dalam protokol evaluasi otomatis dan berbasis manusia.

Dalam mengolah teks, biasanya BERT akan mengubah teks menjadi *input embeddings* terlebih dahulu. Proses perubahan teks input menjadi *input embeddings* dalam model BERT melibatkan beberapa tahap yang penting seperti berikut:

1. Tokenisasi Teks Input:

- Tahap pertama adalah tokenisasi, di mana teks input dibagi menjadi *token-token* (unit-unit terkecil) seperti kata-kata. Setiap token diberi indeks numerik sesuai dengan kamus yang digunakan oleh model BERT.

2. *Word Embeddings*:

- Setelah tokenisasi, setiap token diubah menjadi representasi vektor berdasarkan *word embeddings* yang telah dipelajari sebelumnya. Model BERT biasanya menggunakan *word embeddings* yang telah dilatih dalam tahap *pre-training*. Representasi vektor ini mencerminkan makna kata atau *token* tersebut.
- Misalnya, kata "makan" akan diubah menjadi vektor numerik yang mencerminkan makna kata "makan."

3. *Interval Segment Embeddings* (Jika Ada):

- Jika Anda memiliki dua segmen atau dua kalimat yang berbeda dalam dokumen yang sama (misalnya, pada tugas pertanyaan-jawaban atau terjemahan), *segment embeddings* digunakan untuk membedakan antara segmen-segmen ini. Biasanya, *token* dalam segmen pertama diberi label satu, dan *token* dalam segmen kedua diberi label lain, misalnya, nol.
- *Segment embedding* mengidentifikasi apakah suatu token termasuk dalam segmen pertama atau kedua dalam pasangan teks

4. *Position Embeddings*:

- *Position embeddings* penting untuk memahami urutan kata dalam teks input karena model BERT adalah arsitektur *transformer* yang tidak memiliki pemahaman tentang urutan sekuensial secara alami. *Position embeddings* memberikan informasi tentang posisi absolut setiap token dalam kalimat.

- Misalnya, model BERT perlu tahu apakah token "makan" berada di awal atau di akhir kalimat.

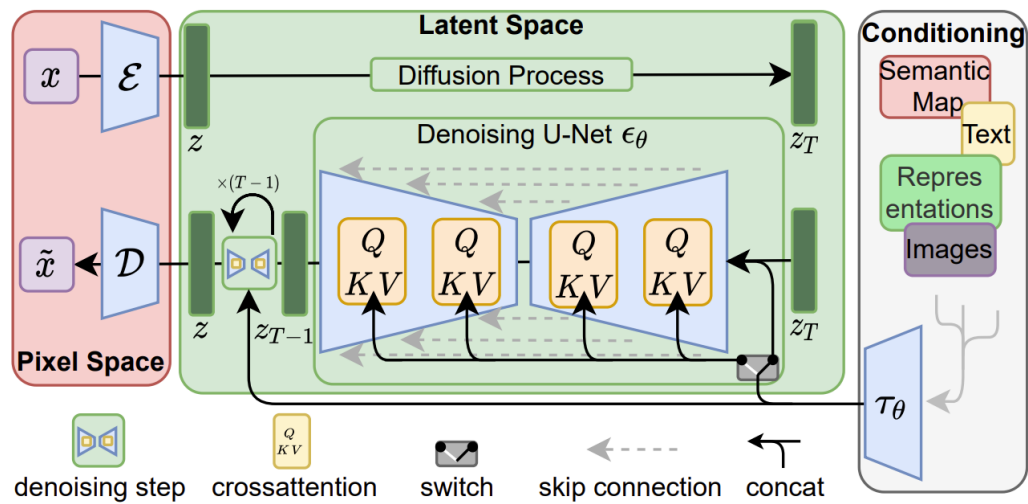
*Output* dari lapisan *encoder* BERT adalah urutan *embedding* kontekstual, di mana setiap *embedding* mewakili makna token dengan konteksnya di dalam suatu kalimat. *Embedding* ini menangkap hubungan semantik antar kata dan memberikan representasi yang komprehensif dari makna kalimat (Miller, D., 2019).

### 2.1.3. Diffusion Models

Model Difusi (DM) merupakan penambahan terbaru pada keluarga model generatif dalam *deep learning* dan telah menjadi pendekatan *state-of-the-art*. Model ini dapat mengakhiri dominasi jangka panjang dari *Generative Adversarial Network* (GAN) yang dibuat oleh Goodfellow dalam tugas yang menantang yaitu sintesis gambar (Dhariwal, P, et al., 2021). Dengan mendekomposisi proses pembentukan gambar menjadi pengaplikasian *denoising autoencoders* secara berurutan, model difusi (DM) dapat mencapai hasil sintesis terbaik pada data gambar dan juga yang lebih dari itu. Untuk sintesis gambar, model-model paling sukses bergantung pada varian yang diberi bobot dari batas bawah varian pada  $p(x)$ , yang mencerminkan skor penyamaran bising. Model-model ini dapat diinterpretasikan sebagai urutan *denoising autoencoders*  $\in \theta(x_t, t); t = 1 \dots T$ , yang dilatih untuk memprediksi varian yang telah dibersihkan dari input mereka  $x_t$ , di mana  $x_t$  adalah versi berisik dari input  $x$  (Rombach, et al., 2022).

$$L_{DM} = E_{x, \epsilon \sim N(0,1), t} \left[ \|\epsilon - \epsilon_{\theta}(x_t, t)\|_2^2 \right],$$

dengan  $t$  diambil secara acak seragam dari  $\{1, \dots, T\}$ .



**Gambar 2.1** Arsitektur sebuah model difusi.

Pada sebuah model difusi laten terdapat 3 komponen utama yaitu:

1. *Autoencoder (VAE)*

Model VAE memiliki dua bagian, yaitu *encoder* dan *decoder*. Selama pelatihan model difusi, *encoder* mengonversi gambar  $512 \times 512 \times 3$  menjadi representasi laten berdimensi rendah dari gambar berukuran misalnya  $64 \times 64 \times 4$  untuk proses *forward diffusion*. Versi terenkripsi kecil dari gambar-gambar ini disebut sebagai laten. Pada setiap langkah pelatihan, lebih banyak *noise* akan diterapkan pada laten ini. Representasi laten terenkripsi dari gambar ini berfungsi sebagai input untuk model *U-Net*.

*Decoder* mengubah representasi laten kembali menjadi gambar dengan cara mengubah representasi laten yang telah di *denoise* yang dihasilkan oleh proses difusi terbalik menjadi gambar menggunakan decoder VAE.

Selama inferensi (proses menggambar), kita hanya memerlukan decoder VAE untuk mengonversi gambar yang telah di *denoise* menjadi gambar-gambar aktual.

2. *U-Net*

*U-Net* memprediksi representasi gambar yang telah di *denoise* dari laten yang berisikan *noise*. Di sini, laten yang berisikan *noise* berperan sebagai masukan untuk *U-Net* dan keluaran dari *U-Net* adalah *noise* pada laten

tersebut. Dengan begitu, dapat diperoleh laten sebenarnya dengan mengurangi *noise* dari laten yang berisikan *noise*.

*Unet* ini menerima data laten yang berisikan *noise* ( $x$ ) dan memprediksi *noise*. Dengan menggunakan model kondisional yang juga menerima langkah waktu ( $t$ ) dan *text-embedding* yang telah di *input* dalam bentuk *prompt* sebagai panduan.

### 3. *Text-encoder*

*Text-encoder* mengubah *prompt* yang telah di *input* menjadi *embedding space* yang menjadi masukan untuk *U-Net*. Ini berfungsi sebagai panduan bagi laten yang memiliki *noise* saat kita melatih *Unet* untuk proses denoising-nya. *Text-encoder* biasanya adalah *encoder* berbasis *transformer* sederhana yang memetakan rangkaian token masukan menjadi rangkaian sematan teks laten. Pada *Stable Diffusion* bukan *text-encoder* baru yang dilatih, tetapi menggunakan *text-encoder* yang sudah dilatih sebelumnya, yaitu CLIP. *Text-encoder* ini menciptakan *embedding* yang sesuai dengan teks masukan.

#### 2.1.3.1. *Stable Diffusion*

*Stable Diffusion* adalah model difusi *text-to-image* laten yang telah dilatih dengan menggunakan gambar-gambar berukuran 512x512 dari sebagian *database* LAION-5B (Rombach, et al., 2022). *Stable Diffusion* ini pada umumnya digunakan untuk menghasilkan gambar yang dikondisikan pada deskripsi teks (*prompt*), *stable diffusion* juga dapat diterapkan untuk tugas lain seperti *inpainting*, *outpainting*, dan menghasilkan *image-to-image* yang dipandu oleh *prompt* (Rombach, et al., 2022). *Stable Diffusion* bekerja dengan menambahkan *noise* ke sebuah gambar. Model kemudian menjalankan proses pengangkatan *noise* atau *denoising* secara bertahap dan meningkatkan kualitas gambar hingga tidak ada *noise* yang tersisa, sehingga menghasilkan sebuah gambar realistis yang sesuai dengan *prompt* teks.



### 2.1.3.2. *Dreambooth*

*DreamBooth* adalah metode untuk mempersonalisasi model text-to-image seperti *Stable Diffusion* hanya dengan beberapa (3-5) gambar subjek. Ini memungkinkan model untuk menghasilkan gambar-gambar yang terkontekstualisasi dari subjek tersebut dalam berbagai adegan, pose, dan pandangan (Ruiz, et al., 2023).

## 2.2. *Text Summarization*

Sekarang peringkas teks otomatis telah menjadi alat penting dan relevan yang dapat membantu dalam memahami dan menginterpretasi data teks (Jezek , et al., 2008). Meskipun ada banyak informasi teks yang tersedia di internet, tidak semuanya dapat dibaca oleh manusia sehingga muncul masalah dalam membaca dan meresapi semua informasi relevan yang tersedia. Tujuan utama dari *text-summarization* adalah meringkas teks sumber menjadi versi yang lebih singkat tetapi tanpa kehilangan informasi dan makna utama dari teks tersebut.

Metode *text-summarization* dapat dikelompokkan menjadi dua jenis yaitu:

- Ekstraktif

Metode ekstraktif terdiri dari pemilihan kalimat atau paragraf penting dari dokumen asli, lalu menggabungkannya menjadi bentuk yang lebih singkat. Pentingnya kalimat tersebut ditentukan berdasarkan fitur statistik dan linguistik dari kalimat (Kyoomarsi, F, et al., 2008).

- Abstraktif

Ringkasan abstraktif berusaha untuk memahami konsep utama dalam sebuah dokumen dan mengekspresikan konsep-konsep tersebut dalam bahasa alami yang jelas. Metode ini menggunakan metode linguistik untuk memeriksa dan menafsirkan teks kemudian menemukan konsep dan ekspresi baru untuk menjelaskan teks tersebut dengan menghasilkan teks baru yang lebih pendek yang menyampaikan informasi penting dari dokumen aslinya (Erkan, G, et al., 2004).

### **2.2.1. ROUGE**

Library ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) adalah sebuah toolkit yang digunakan untuk mengevaluasi kualitas ringkasan teks yang dihasilkan oleh model mengukur seberapa banyak kata yang sama yang terdapat dalam ringkasan yang dihasilkan dan ringkasan yang dibuat oleh manusia (Lin., 2004). Tetapi, ROUGE dapat juga digunakan untuk membandingkan kesamaan antara teks hasil ringkasan dengan teks orisinal dengan menghitung beberapa metrik yang digunakan untuk mengukur kesamaan antara dua teks.

## BAB III

### ANALISIS DAN PERANCANGAN SISTEM

#### 3.1. Penjelasan Sistem

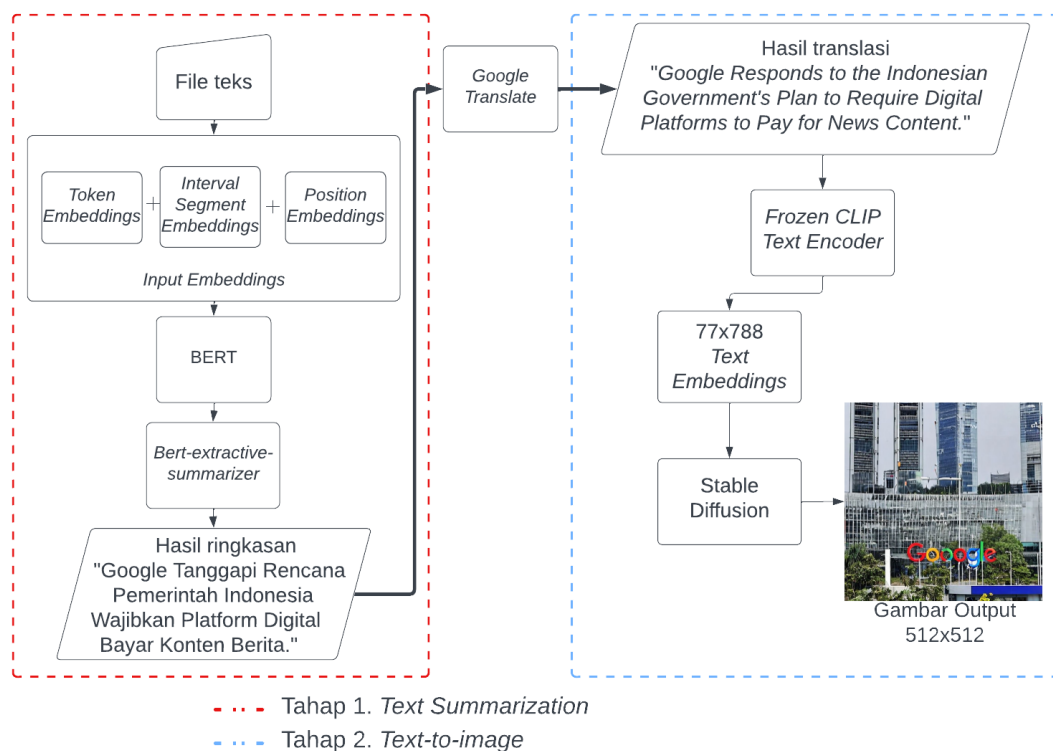
Pada penelitian ini, penulis mengembangkan sebuah sistem untuk mengotomatisasi proses meringkas dan membuat gambar sampul. Sistem dikembangkan dengan menggunakan *Bidirectional Encoder Representation for Transformer* (BERT) untuk meringkas sebuah teks secara ekstraktif, dan *Stable Diffusion* untuk membuat gambar sampul berukuran 512x512 pixel. Model *Stable Diffusion* akan dilatih lagi dengan gambar-gambar yang telah dilakukan pra proses dan berasal dari beberapa website berita atau artikel *online*, untuk model BERT sendiri penulis memutuskan untuk menggunakan sebuah library yang memungkinkan model BERT tersebut melakukan *text summarization* secara ekstraktif.

Penelitian dimulai dengan mengumpulkan data-data yang berupa gambar dari beberapa *website* berita berbahasa Indonesia yaitu [kompas.com](https://kompas.com), [detik.com](https://detik.com), dan [cnnindonesia.com](https://cnnindonesia.com). Gambar-gambar tersebut di-*crop* atau *resize* menjadi gambar berukuran 512x512 yang kemudian disimpan di dalam *google drive* dan diakses menggunakan *google colab*. Dengan menggunakan *dataset* tersebut tahap pelatihan model *Stable Diffusion* dengan pendekatan *Dreambooth* oleh *Google research team* yang memungkinkan pelatihan model *Stable Diffusion* hanya dengan beberapa gambar saja. Pelatihan model tersebut dilakukan dalam dua langkah yaitu :

- Melakukan *fine-tuning* pada model dengan memasukkan gambar-gambar *dataset* sebagai input dan dipasangkan dengan sebuah *text prompt* unik, pada penelitian ini penulis menggunakan “berita” sebagai *text prompt* unik. Secara paralel, diterapkan *class-specific prior preservation loss*, yang memanfaatkan prioritas semantik yang dimiliki model terhadap kelas tersebut dan mendorongnya untuk menghasilkan contoh gambar yang beragam yang termasuk dalam kelas subjek menggunakan nama kelas dalam *text prompt*.

- Melakukan fine-tuning pada komponen *super-resolution* pada model dengan beberapa pasang gambar *low-resolution* dan *high-resolution* yang diambil dari gambar-gambar pada *dataset*, yang memungkinkan kita mempertahankan keakuratan tinggi terhadap detail-detail kecil dari subjek.

Setelah proses pelatihan model *Stable Diffusion*, dimasukkan input berupa teks yang merupakan hasil terjemahan dari teks berita online yang telah diringkas oleh model BERT (*Bidirectional Encoder Representation for Transformer*) menggunakan pendekatan library *Bert Extractive Summarizer* yang memanfaatkan library dari *HuggingFace Pytorch transformers* untuk melakukan ringkasan ekstraktif. Proses ini dimulai dengan menyematkan kalimat-kalimat (*text embedding*), lalu menjalankan algoritma pengelompokan (*clustering*), menemukan kalimat-kalimat yang paling dekat dengan pusat-pusat kluster. Hasil ringkasan tersebut kemudian diterjemahkan ke Bahasa Inggris dengan menggunakan library *GoogleTranslate* yang memanfaatkan API dari *Google Translate*. Dengan teks input tersebut, model *Stable Diffusion* yang telah dilatih dapat menghasilkan gambar sesuai dengan konteks yang berasal dari teks input tersebut.



**Gambar 3.1** Arsitektur Umum Sistem.

### 3.2. Sumber Dataset

*Dataset* yang digunakan merupakan gambar-gambar berita yang berhubungan dengan Negara Indonesia yang diambil pada beberapa website berita berbahasa Indonesia *online*, ukuran minimal piksel gambar *dataset* yang digunakan untuk melatih model adalah 512x512 piksel agar mempertahankan kualitas gambar yang dihasilkan oleh model *Stable Diffusion*.

### 3.3. Pra-proses Gambar

Selanjutnya gambar-gambar pada *dataset* di pra-proses dengan melakukan *cropping* atau *resizing* agar semua gambar pada *dataset* memiliki dimensi gambar yang sama yaitu 512x512 piksel.

### 3.4. Pelatihan Model *Stable Diffusion* dengan *Dreambooth*

Proses pelatihan model *Stable Diffusion* dimulai dengan memasang gambar-gambar yang telah di pra-proses dengan sebuah *class* (misalnya untuk melatih model dengan gambar-gambar “anjing shiba” maka *class* nya adalah “anjing”), pada penelitian ini *class* yang digunakan adalah “berita”. Kemudian pelatihan dilanjutkan dengan melatih *text encoder* yang terdapat pada model dengan *class* dan gambar-gambar yang telah dipasangkan tersebut sebanyak *n* kali, lalu dilanjutkan lagi dengan melatih *U-Net* dengan cara menambahkan noise pada gambar-gambar *dataset* lalu mencoba memprediksi noise tersebut sebanyak *n* kali. Dengan menggunakan *Dreambooth* untuk melakukan pelatihan model akurasi dan detail gambar tetap bisa dijaga meskipun hanya dengan beberapa gambar saja. Proses pelatihan ini menggunakan *resource* yang sangat besar dari perangkat yang digunakan dan juga membutuhkan waktu yang cukup lama.

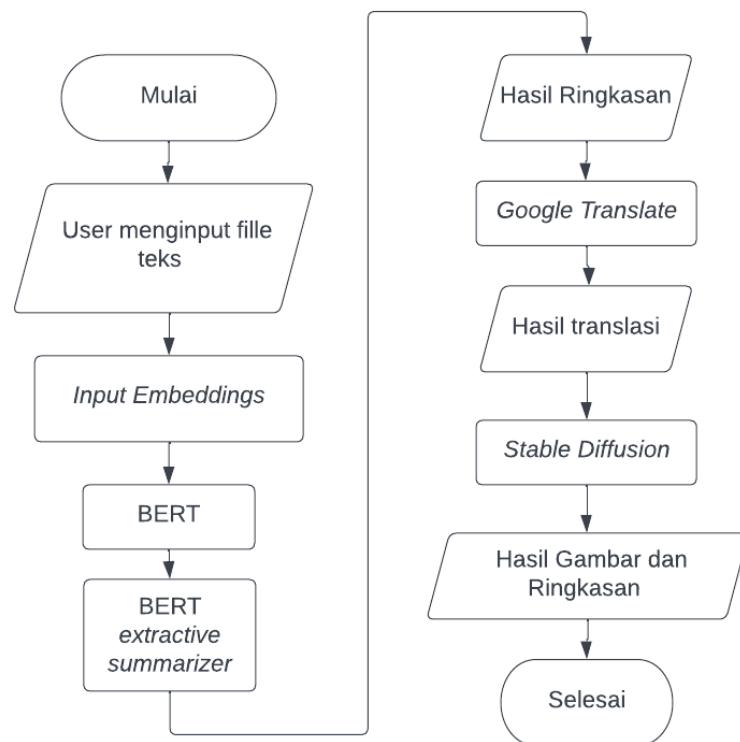
### 3.5. Ringkasan Ekstraktif dengan BERT

Teks input yang ingin diringkas dimasukkan melalui model BERT untuk diringkas. Proses meringkas menggunakan *library* BERT *Extractive Summarizer* yang memanfaatkan *library* dari *HuggingFace Pytorch transformers* untuk melakukan ringkasan ekstraktif dengan cara mengubah teks menjadi *text embedding*, kemudian menjalankan algoritma pengelompokan (*clustering*) yang mengelompokkan kata-

kata dan kalimat-kalimat yang dianggap penting oleh model BERT, lalu menemukan kalimat-kalimat yang paling dekat dengan pusat-pusat kluster. Kalimat-kalimat tersebut kemudian disusun kembali menjadi teks hasil ringkasan.

### 3.6. Evaluasi

Untuk mengevaluasi sistem, hasil dari kedua tahap yang dijalankan oleh sistem diperlihatkan secara jelas. Pada tahap pertama, sistem menggunakan model BERT dan *library* BERT *extractive summarizer* untuk meringkas teks, hasil ringkasan tersebut kemudian diterjemahkan menjadi bahasa Inggris dengan bantuan *library* *GoogleTranslate*. Hasil terjemahan dari ringkasan tersebut kemudian dijadikan *input* untuk model *Stable Diffusion* sebagai prompt yang menghasilkan gambar akhir.



**Gambar 3.2** Flowchart Sistem.

Pada penelitian ini digunakan model BERT *pre-trained* yaitu model IndoBERT yang telah dilatih pada kumpulan data teks bahasa Indonesia memperlihatkan keunggulan performanya dibanding dengan model lainnya atau model *Multilingual*

BERT (M-BERT) (Lintang, S., 2020). Sebelum melakukan peringkasan, teks input harus diubah menjadi *input embeddings* terlebih dahulu. Pada penelitian Franky Halim, et al (2022), peringkasan pada model BERT menggunakan *encoder* untuk memahami hubungan antar kalimat dalam sebuah artikel berita.

Setelah itu, dengan memanfaatkan *library BERT extractive summarizer* yang bekerja dengan menggunakan *clustering K-means* untuk mengelompokkan *embedding* kalimat, yang mewakili makna setiap kalimat dalam sebuah teks. Algoritma ini dikonfigurasi untuk menghasilkan sejumlah *cluster* yang sesuai dengan panjang ringkasan yang diinginkan.

Proses *Clustering library BERT extractive summarizer* pada penelitian Derek Miller (2019) adalah sebagai berikut:

1. Inisialisasi *Cluster*: *K centroid* awal dipilih secara acak dari kumpulan *embedding* kalimat. *Centroid* ini mewakili posisi awal *cluster*.
2. Penugasan Titik Data: Setiap *embedding* kalimat ditetapkan ke *cluster* yang *centroidnya* paling dekat dengannya. Penugasan ini didasarkan pada kesamaan kosinus antara *embedding* kalimat dan *centroid cluster*.
3. Perhitungan Ulang *Centroid*: Setelah semua *embedding* kalimat ditetapkan, *centroid* dari setiap *cluster* dihitung ulang dengan menghitung rata-rata *embedding* dari kalimat yang termasuk dalam *cluster* itu.
4. Iterasi dan Konvergensi: Langkah 3 dan 4 diulang secara berulang hingga *centroid* mencapai posisi yang stabil. Setelah *cluster* mencapai titik stabil maka proses pengelompokan selesai.

Setelah algoritma *clustering K-means* mencapai konvergensi, kalimat dalam *centroid* dari setiap cluster dipilih untuk membentuk ringkasan akhir. Kalimat ini dianggap paling representatif dan informatif dalam cluster masing-masing, memastikan ringkasan kuliah yang ringkas dan komprehensif.

Hasil ringkasan kemudian dievaluasi dengan menggunakan *library* ROUGE. Beberapa metrik yang paling umum digunakan pada ROUGE adalah:

- ROUGE-N: Menghitung jumlah *n-gram* yang sama yang terdapat dalam dua teks.
- ROUGE-L: Menghitung jumlah kata yang sama yang terdapat dalam dua teks, tanpa memperhatikan urutan kata.
- ROUGE-W: Menghitung jumlah kata yang sama yang terdapat dalam dua teks, dengan mempertimbangkan bobot kata berdasarkan frekuensinya dalam *corpus*.

Untuk menghitung nilai ROUGE, *library* ROUGE menggunakan algoritma yang disebut *Longest Common Subsequence* (LCS). LCS adalah algoritma yang digunakan untuk menemukan urutan subkata yang terpanjang yang sama dalam dua teks.

Untuk menghitung nilai ROUGE-N, *library* ROUGE menghitung jumlah *n-gram* yang sama yang terdapat dalam dua teks. Misalnya, untuk menghitung nilai ROUGE-1, *library* ROUGE menghitung jumlah kata yang sama yang terdapat dalam dua teks.

Untuk menghitung nilai ROUGE-L, *library* ROUGE menghitung jumlah kata yang sama yang terdapat dalam dua teks, tanpa memperhatikan urutan kata. Misalnya, untuk menghitung nilai ROUGE-L, *library* ROUGE menghitung jumlah kata "saya", "pergi", dan "ke" dalam dua teks, meskipun kata-kata tersebut tidak terdapat dalam urutan yang sama.

Untuk menghitung nilai ROUGE-W, *library* ROUGE menghitung jumlah kata yang sama yang terdapat dalam dua teks, dengan mempertimbangkan bobot kata berdasarkan frekuensinya dalam *corpus*. Misalnya, untuk menghitung nilai ROUGE-W, *library* ROUGE memberikan bobot yang lebih tinggi untuk kata "saya" daripada kata "ke", karena kata "saya" lebih sering muncul dalam *corpus*.

Nilai ROUGE yang lebih tinggi menunjukkan bahwa ringkasan yang dihasilkan lebih mirip dengan ringkasan yang asli.



Pada tahap kedua, sistem menggunakan model *Stable Diffusion* untuk melakukan proses *inference* (membuat gambar) berdasarkan teks ringkasan yang dihasilkan pada tahap pertama. Tetapi sebelum itu, ringkasan tersebut harus diterjemahkan menjadi bahasa Inggris terlebih dahulu dikarenakan batasan dari model *Stable Diffusion* yang digunakan penulis hanya dapat memahami teks dalam bahasa Inggris. Untuk melakukan penerjemahan ini, sistem menggunakan *library Google Translate*. Kemudian, dengan memanfaatkan *library huggingface* untuk mengambil dan menjalankan model *Stable Diffusion* yang telah dilatih. Hasil terjemahan tersebut dimasukkan sebagai *prompt* untuk membimbing model *Stable Diffusion*.

## **BAB IV**

### **IMPLEMENTASI DAN PENGUJIAN**

#### **4.1. Perangkat yang Digunakan**

Perangkat yang digunakan untuk membuat sistem peringkas teks dan pembuat sampul otomatis dengan pendekatan model transformer BERT (*Bidirectional Encoder Representations from Transformers*) dan *Stable Diffusion* adalah sebagai berikut:

##### **4.1.1. Perangkat Keras**

Detail perangkat keras yang digunakan untuk membuat sistem adalah sebagai berikut:

1. 1 x *single core hyper threaded Xeon Processors @2.3Ghz*
2. 13.0 GB RAM
3. 78.2 GB *Disk space*
4. NVIDIA Tesla T4 16GB VRAM

##### **4.1.2. Perangkat Lunak**

Detail perangkat lunak yang digunakan untuk membuat sistem adalah sebagai berikut:

1. *Google Colaboratory*
2. *Google Drive*
3. Sistem Operasi Windows 11

##### **4.1.3. Library**




Detail *library* yang digunakan untuk membuat sistem adalah sebagai berikut:

1. *HuggingFace*
2. *Transformers*
3. *bert-extractive-summarizer*
4. *GoogleTranslate*
5. *Diffusers*
6. *Accelerate*

#### 4.2. Penerapan Tahap Pengumpulan *Dataset*

*Dataset* yang digunakan untuk melatih dan mengevaluasi model didapat dari beberapa website berita berbahasa Indonesia yaitu [kompas.com](http://kompas.com), [detik.com](http://detik.com), dan [cnnindonesia.com](http://cnnindonesia.com). *Dataset* yang telah dikumpulkan kemudian di *crop* atau *resize* agar semua gambar pada *dataset* memiliki ukuran piksel sebesar 512 x 512 piksel. Jumlah total *dataset* ada sebanyak 31 gambar yang terdiri dari:

**Tabel 4.1** *Dataset* gambar sebelum di proses.

Sumber	Jumlah Gambar (Total = 31)	Contoh Gambar
<a href="http://kompas.com">kompas.com</a>	17	
<a href="http://detik.com">detik.com</a>	6	
<a href="http://cnnindonesia.com">cnnindonesia.com</a>	8	

### 4.3. Penerapan Tahap Pra-proses *Dataset*

Untuk melakukan *cropping* dan *resizing*, penulis memutuskan untuk menggunakan website [birme.net](http://birme.net) yang dimana user dapat melakukan *batch crop* atau *resize*.

**Tabel 4.2** Contoh gambar-gambar yang telah diproses.

Gambar Asli	Gambar yang telah diproses
	
	
	

#### 4.4. Penerapan Tahap Pelatihan *Stable Diffusion* dengan *Dreambooth*

Tahap ini dimulai dengan menginstal dan mengimport semua *dependencies* yang dibutuhkan oleh *Dreambooth* dan *Stable Diffusion*.

```
os.chdir('/content')
!pip install -qq --no-deps accelerate==0.12.0
!wget -q -i https://raw.githubusercontent.com/TheLastBen/fast-
stable-diffusion/main/Dependencies/dbdeps.txt
!dpkg -i *.deb
!tar -C / --zstd -xf gcolabdeps.tar.zst
!rm *.deb | rm *.zst | rm *.txt
!git clone -q --depth 1 --branch main
https://github.com/TheLastBen/diffusers
!pip install gradio==3.16.2 --no-deps -qq
```

*Code Snippet* ini berisikan *library-library* dan *software-software* yang dibutuhkan untuk menjalankan serta melatih sebuah model *Stable Diffusion*, contohnya *pytorch*, *transformer*, *gradio*, dll. Kemudian, model *Stable Diffusion* v1.5 di *download* dan *load* dari *huggingface*.

```
%cd /content/
clear_output()
!mkdir /content/stable-diffusion-v1-5
%cd /content/stable-diffusion-v1-5
!git init
!git lfs install --system --skip-repo
!git remote add -f
origin "https://huggingface.co/runwayml/stable-diffusion-v1-
5"
!git config core.sparsecheckout true
!echo -e
"scheduler\ntext_encoder\ntokenizer\nunet\nvae\nmodel_index.js
on\n!vae/diffusion_pytorch_model.bin\n!*safetensors\n!*fp16.
bin" > .git/info/sparse-checkout
!git pull origin main
```

Setelah itu, gambar-gambar pada *dataset* di *load* ke dalam sebuah *folder google drive* yang dinamai dengan class yang dipasangkan dengan gambar-gambar *dataset* (penulis menggunakan “berita” sebagai *class*), *folder* tersebut nantinya juga digunakan untuk menyimpan hasil model yang telah dilatih.

```

Session_Name = "berita"
INSTANCE_NAME=Session_Name
OUTPUT_DIR="/content/models/"+Session_Name
SESSION_DIR=WORKSPACE+'/Sessions/'+Session_Name
INSTANCE_DIR=SESSION_DIR+'/instance_images'
CAPTIONS_DIR=SESSION_DIR+'/captions'
MDLPTH=str(SESSION_DIR+"/"+Session_Name+'.ckpt')

IMAGES_FOLDER="/content/gdrive/MyDrive/cover_berita"

Smart_Crop_images= True
Crop_size = 512

```

Setelah *dataset* di *load*, proses pelatihan model dilanjutkan dengan melatih *text encoder* dengan *class* dan gambar-gambar yang telah dipasangkan yang kemudian dilanjutkan dengan pelatihan *U-Net* dengan gambar-gambar pada *dataset*. Kedua proses ini menggunakan library *Diffusers* dan *Dreambooth*. Kemudian, pelatihan *text encoder* dilakukan dengan melatih model untuk memahami hubungan antara gambar dataset dengan class yang sudah ditentukan sebanyak *n* steps dengan *n* tidak kurang dari 200 dan tidak lebih dari 500 untuk menghindari underfitting yang terjadi ketika model tidak dilatih dengan cukup lama dan overfitting yang terjadi ketika model dilatih dengan berlebihan.

```

MODEL_NAME=MODEL_NAME

UNet_Training_Steps=1500
UNet_Learning_Rate = 2e-6
untnlr=UNet_Learning_Rate

Text_Encoder_Training_Steps=350

Text_Encoder_Learning_Rate = 1e-6
txtnlr=Text_Encoder_Learning_Rate

!accelerate launch
/content/diffusers/examples/dreambooth/train_dreambooth.py
\
$trnlnltxt \
$extrnlcptn \
$ofstnse \
--image_captions_filename \

```

```

--train_text_encoder \
--dump_only_text_encoder \
--pretrained_model_name_or_path="$MODEL_NAME" \
--instance_data_dir="$INSTANCE_DIR" \
--output_dir="$OUTPUT_DIR" \
--captions_dir="$CAPTIONS_DIR" \
--instance_prompt="$PT" \
--seed=$Seed \
--resolution=$TexRes \
--mixed_precision=$precision \
--train_batch_size=1 \
--gradient_accumulation_steps=1 --gradient_checkpointing \
--use_8bit_adam \
--learning_rate=$txlr \
--lr_scheduler="linear" \
--lr_warmup_steps=0 \
--max_train_steps=$Training_Steps

```

Lalu, proses pelatihan *U-Net* pada kode berikut ini dilakukan dengan menambahkan *noise* pada gambar *dataset*, kemudian model memprediksi hasil gambar dari *noise* tersebut.

```

!accelerate launch
/content/diffusers/examples/dreambooth/train_dreambooth.p
y \
$extrnlcptn \
$ofstnse \
--image_captions_filename \
--train_only_unet \
--save_starting_step=$stpsv \
--save_n_steps=$stp \
--Session_dir=$SESSION_DIR \
--pretrained_model_name_or_path="$MODEL_NAME" \
--instance_data_dir="$INSTANCE_DIR" \
--output_dir="$OUTPUT_DIR" \
--captions_dir="$CAPTIONS_DIR" \
--instance_prompt="$PT" \
--seed=$Seed \
--resolution=$Res \
--mixed_precision=$precision \
--train_batch_size=1 \
--gradient_accumulation_steps=1 $GCUNET \
--use_8bit_adam \
--learning_rate=$untlr \
--lr_scheduler="linear" \
--lr_warmup_steps=0 \
--max_train_steps=$Training_Steps

```

Pelatihan ini dilakukan sebanyak 1500 *steps* dengan alasan yang sama dengan pelatihan *text encoder*, setelah proses pelatihan selesai model disimpan ke dalam *folder* sebagai sebuah *file* dengan nama “berita.ckpt” dimana .ckpt (*checkpoint*) merupakan ekstensi file model agar bisa digunakan untuk melakukan *inference*.

#### 4.5. Penerapan Tahap Peringkasan Ekstraktif

Tahap ini diawali dengan mendownload dan load model IndoBERT dari *huggingface*. Kemudian teks berita diubah menjadi *sentence embedding*, lalu dengan menggunakan *library bert-extractive-summary* diaplikasikan sebuah algoritma *clustering* yang memungkinkan pencarian kalimat-kalimat yang paling dekat dengan pusat *cluster*.

```
#BERT & Extractive Summarizer
custom_config =
AutoConfig.from_pretrained(pretrained_model_name_or_path="sara
hlintang/IndoBERT")
custom_config.output_hidden_states = True
custom_tokenizer =
AutoTokenizer.from_pretrained(pretrained_model_name_or_path="s
arahlintang/IndoBERT")
custom_model =
AutoModel.from_pretrained("sarahlintang/IndoBERT",
config=custom_config)

model = Summarizer(custom_model=custom_model,
custom_tokenizer=custom_tokenizer)
clear_output()
```

Disini model IndoBERT dan *library bert-extractive-summarizer* dipersiapkan untuk melakukan peringkasan. Untuk memulai peringkasan, fungsi model dapat dipanggil dengan memasukkan parameter berupa teks yang ingin diringkas seperti pada *code snippet* berikut ini.



```

#@markdown
teks = "" #@param {type:"string"}

if teks != "" :
    summary = model(teks, num_sentences=4)
    clear_output()

else :
    teks = input('')
    summary = model(teks, num_sentences=4)
    clear_output()

```

Kemudian 3 contoh teks diambil dari beberapa *website* artikel yang berhubungan dengan berita nasional dan diringkas, hasil ringkasan dari contoh-contoh tersebut dapat dilihat pada **Tabel 4.3**.

**Tabel 4.3** Tabel ringkasan.

<p>Contoh 1</p> <p>Sumber : <a href="https://kemenparekraf.go.id">https://kemenparekraf.go.id</a></p> <p>637 Kata / 1411 <i>tokens</i> (teks awal) → 77 Kata / 189 <i>tokens</i> (ringkasan)</p>	
Teks Awal	<p>Kementerian Pariwisata dan Ekonomi Kreatif/Badan Pariwisata dan Ekonomi Kreatif mendukung penyelenggaraan kejuaraan nasional e-sports "Ksatriya Mahardhika Student Championship 2023" yang akan mempertandingkan Lokapala, gim MOBA (Multiplayer Online Battle Arena) pertama buatan Indonesia. Wakil Menteri Pariwisata dan Ekonomi Kreatif/Wakil Kepala Badan Pariwisata dan Ekonomi Kreatif (Wamenparekraf/Wakabaparekraf) Angela Tanoesoedibjo dalam sambutannya di acara kick off "Ksatriya Mahardhika Student Championship 2023", Senin (14/8/2023), mengatakan kejuaraan nasional esports yang ditargetkan diikuti lebih dari 5.000 peserta dari berbagai daerah di Indonesia ini akan memperkuat pengembangan IP (intelektual property) gim lokal. "Kemenparekraf akan terus mendukung perkembangan IP gim lokal, karena ini adalah salah satu bentuk dari subsektor</p>

	<p>ekonomi kreatif dengan potensi yang besar," kata Wamenparekraf Angela. Berdasarkan data, pada 2022 transaksi gim di seluruh dunia mencapai Rp3.500 triliun. Sementara di Indonesia jumlahnya mencapai Rp30 triliun. Namun sayangnya dari jumlah tersebut, mayoritas gim yang dimainkan oleh lebih dari 175 juta pemain gim online di Indonesia merupakan gim dari pengembang asing. Pemerintah, kata Angela, menaruh perhatian yang besar dalam pengembangan industri gim tanah air. Salah satunya adalah sedang dilakukannya finalisasi pembentukan Peraturan Presiden (Perpres) tentang percepatan industri gim tanah air. Perpres ini akan menjadi payung hukum atau dasar hukum yang kuat dalam membangun ekosistem gim di Indonesia yang lebih inklusif, relevan, dan lebih berpihak kepada gim lokal dan berkelanjutan. "Kita sedang memfinalisasi perpres gim ini bersama-sama dengan semua kementerian dan lembaga untuk kita bisa menguatkan ekosistem gim di Indonesia dan mempunyai keberpihakan terhadap pemain-pemain lokal, ini penting sekali," ujar Angela. Karenanya ia berharap penyelenggaraan "Ksatriya Mahardhika Student Championship 2023" dapat memperkuat pengembangan industri gim tanah air, selain tentunya sebagai bentuk promosi atau diplomasi budaya tanah air. "Saya selalu percaya bahwa perpaduan dari ekonomi kreatif dan budaya, kalau dikelola dengan baik bisa memberikan nilai tambah. Tidak hanya nilai tambah ekonomi tetapi juga bisa nilai tambah persatuan kebangsaan dan juga nilai-nilai keberagaman," ujar Angela. Selain itu, Angela mengatakan, kejuaraan nasional "Ksatriya Mahardhika Student Championship 2023" juga akan menjadi ruang ekspresi yang baik bagi para konten kreator juga generasi muda yang ingin terjun ke dalam dunia esports. Lokapala sendiri adalah gim yang dipertandingkan dalam eksibisi esport di PON XX Papua</p>
--	---

	<p>2021 dan akan menjadi gim lokal yang akan dipertandingkan di PON 2024. "Kawan-kawan para pemain Lokapala bisa berkesempatan untuk menjadi atlet. Tapi tentunya kita harus terus menanamkan nilai-nilai kedisiplinan, tanggung jawab, karena ketika nanti bisa berkarier di industri gim, nilai-nilai tentang kedisiplinan dan tanggung jawab tetap diperlukan," ujar Angela. CEO dari Anantarupa Studios sekaligus penyelenggara "Ksatriya Mahardhika Student Championship 2023", Ivan Chen, mengatakan kejuaraan ini perwujudan nyata semangat kompetitif yang tumbuh di antara generasi muda tanah air, sambil mempromosikan budaya esports yang sehat dan inklusif. "Dengan melibatkan para pemuda yang berkompetisi dengan semangat juang dan sportivitas dalam kejuaraan ini, kami ingin mendorong intisari dari Astabrata ini ke dalam pengembangan kepemimpinan, kerjasama tim, dan disiplin yang akan membawa dampak positif bagi masyarakat dan bangsa," kata Ivan Chen. Peserta dari berbagai kalangan pelajar dari delapan daerah dari seluruh Indonesia akan bersaing dengan semangat juang yang tinggi untuk memperebutkan gelar juara "Ksatriya Mahardhika Student Championship 2023". Delapan daerah itu mencakup DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Sulawesi, Sumatera, Bali dan NTB, serta Daerah Istimewa Yogyakarta. Acara ini juga akan dimeriahkan berbagai kegiatan pendukung seperti berbagai macam sayembara budaya, lomba tari, ajang bagi content creator, roadshow ke sekolah-sekolah, pasar rakyat UMKM, pelestarian budaya, workshop edukatif, maupun coaching game, serta beragam kuis dan giveaway menarik bagi para penggemar setia Lokapala. "Lokapala bisa menjadi medium untuk diplomasi budaya, sarana pendidikan bagi anak-anak karena banyak sekali pemain lokal yang bisa belajar sejarah dan budaya lewat Lokapala," ujar Ivan Chen.</p>
--	--

	<p>Turut hadir mendampingi Wamenparekraf Angela dalam kesempatan tersebut, Deputi Bidang Ekonomi Digital dan Produk Kreatif Kemenparekraf/Baparekraf, Muhammad Neil El Himam; Staf Ahli Bidang Inovasi dan Kreativitas Kemenparekraf/Baparekraf, Josua Simanjuntak; serta Direktur Utama Badan Otorita Borobudur, Agustin Peranginangin.</p>
Ringkasan	<p>Kementerian Pariwisata dan Ekonomi Kreatif/Badan Pariwisata dan Ekonomi Kreatif mendukung penyelenggaraan kejuaraan nasional e-sports "Ksatriya Mahardhika Student Championship 2023" yang akan mempertandingkan Lokapala, gim MOBA (Multiplayer Online Battle Arena) pertama buatan Indonesia. Berdasarkan data, pada 2022 transaksi gim di seluruh dunia mencapai Rp3.500 triliun. Selain itu, Angela mengatakan, kejuaraan nasional "Ksatriya Mahardhika Student Championship 2023" juga akan menjadi ruang ekspresi yang baik bagi para konten kreator juga generasi muda yang ingin terjun ke dalam dunia esports.</p>

## Contoh 2

Sumber : <https://regional.kompas.com>276 Kata / 648 *tokens* (teks awal) → 65 Kata / 161 *tokens* (ringkasan)

Teks Awal	<p>Dua orang anggota polisi lalu lintas (Polantas) mengambil alih kemudi ambulans di ruas tol Pekanbaru-Dumai demi mengantar pasien darurat yang mengalami patah tulang. Salah satu polisi, Bripka Fernandoz menceritakan peristiwa saat dirinya menjadi sopir ambulans dadakan. Bripka Fernandoz mengungkapkan, peristiwa tersebut terjadi pada Rabu (19/9/2023) dini hari di Tol Pekanbaru-Dumai. Saat itu, polisi sedang melakukan patroli dengan mobil operasional. Bripka Fernandoz mengatakan dirinya sedang bersama rekannya Aiptu Jansen Sihombing. "Kami ketika itu sedang melaksanakan patroli antisipasi microsleep (tertidur sesaat) pengemudi. Kita patroli mulai jam 01.00 WIB," kata Bripka Fernandoz, seperti dilansir dari Tribun Pekanbaru, Sabtu (23/9/2023). Mereka lalu melihat ambulans berhenti di bahu jalan tol, tepatnya di jalur B kilometer 50 dari arah Dumai ke Pekanbaru. Ternyata saat didekati, ambulans tersebut berisi pasien darurat. Sedangkan sang sopir mengantuk dan tertidur. "Kami kira awalnya entah pecah ban atau gimana, maka saya dan senior saya mendekat. Ternyata sopir ambulans, maka saya dan senior saya mendekat, sopir ambulans saya lihat posisinya sedang memejamkan mata," ungkap Bripka Fernandoz. Saat ditanya, sopir ambulans mengaku tidur karena sangat mengantuk. "Dia kelelahan dan mau tidur sebentar karena mengantuk sekali, lantaran sudah jalan beberapa trip," ungkapnya. Di dalam ambulans, katanya, terdapat seorang pasien laki-laki dengan kondisi serius, beserta istri dan dua anak. "Sopir ini bilang dia lagi membawa pasien patah pinggang rujukan Rumah Sakit Bagansiapi-api, Rokah Hilir yang mau dioperasi pada pagi hari," ujarnya. Tak berpikir</p>
-----------	--

	<p>lama, Brika Fernandoz kemudian mengambil alih kemudi ambulans. "Saya langsung ambil alih kemudi, sementara senior saya mengawal dengan mobil patroli. Karena ini kondisi genting, pasien harus sampai tepat waktu," paparnya. Sementara Direktur Lalu Lintas (Dirlantas) Polda Riau Kombes Dwi Nur Setiawan membenarkan kejadian tersebut.</p>
Ringkasan	<p>Dua orang anggota polisi lalu lintas (Polantas) mengambil alih kemudi ambulans di ruas tol Pekanbaru-Dumai demi mengantar pasien darurat yang mengalami patah tulang. "Ternyata sopir ambulans, maka saya dan senior saya mendekat, sopir ambulans saya lihat posisinya sedang memejamkan mata," ungkap Brika Fernandoz. "Sopir ini bilang dia lagi membawa pasien patah pinggang rujukan Rumah Sakit Bagansiapi-api, Rokah Hilir yang mau dioperasi pada pagi hari," ujarnya.</p>
<p style="text-align: center;">Contoh 3</p> <p style="text-align: center;">Sumber : <a href="https://www.cnnindonesia.com">https://www.cnnindonesia.com</a></p> <p style="text-align: center;">153 Kata / 368 <i>tokens</i> (teks awal) → 57 Kata / 146 <i>tokens</i> (ringkasan)</p>	
Teks Awal	<p>Penjabat (Pj) Bupati Flores Timur, Nusa Tenggara Timur (NTT), Doris Alexander Rihi, mengimbau warga tidak makan nasi tiap Jumat. Salah satu alasannya, mengantisipasi krisis pangan dan ancaman kekeringan. Surat yang dikeluarkan pada Kamis (21/9) tersebut ditujukan kepada pimpinan Organisasi Perangkat Daerah (OPD), camat, kepala desa/lurah, serta pimpinan BUMN dan BUMD di seluruh Kabupaten Flores Timur. Penjabat (Pj) Bupati Flores Timur, Nusa Tenggara Timur (NTT), Doris Alexander Rihi, mengimbau warga tidak makan nasi tiap Jumat. Salah satu alasannya, mengantisipasi krisis pangan dan ancaman kekeringan. Surat yang dikeluarkan pada Kamis (21/9) tersebut ditujukan kepada pimpinan</p>

	Organisasi Perangkat Daerah (OPD), camat, kepala desa/lurah, serta pimpinan BUMN dan BUMD di seluruh Kabupaten Flores Timur. Ia juga mengatakan gerakan ini untuk mewujudkan sumber daya manusia yang sehat, aktif, dan produktif melalui kecukupan pangan yang beragam, bergizi seimbang, dan aman. "Nona Sari Setia juga untuk mengendalikan inflasi yang terjadi akibat kenaikan harga beras," kata dia.
Ringkasan	Penjabat (Pj) Bupati Flores Timur, Nusa Tenggara Timur (NTT), Doris Alexander Rihi, mengimbau warga tidak makan nasi tiap Jumat. Salah satu alasannya, mengantisipasi krisis pangan dan ancaman kekeringan. Surat yang dikeluarkan pada Kamis (21/9) tersebut ditujukan kepada pimpinan Organisasi Perangkat Daerah (OPD), camat, kepala desa/lurah, serta pimpinan BUMN dan BUMD di seluruh Kabupaten Flores Timur.

Hasil ringkasan tersebut kemudian diterjemahkan menjadi bahasa Inggris dengan menggunakan *library GoogleTranslate*.

```
#translate Indonesian to English
translator = Translator()
translations = translator.translate(summary)
en_prompt = translations.text
```

Dengan menggunakan *library GoogleTranslate* yang di impor sebagai *Translator*, ringkasan dapat di translasi menjadi bahasa Inggris agar dapat dijadikan *input* dan dipahami oleh model *Stable Diffusion*. Hasil terjemahan dari ketiga contoh dapat dilihat pada **Tabel 4.4**.

**Tabel 4.4** Tabel hasil penerjemahan ringkasan.

Contoh	Ringkasan Indonesia	Hasil terjemahan (English)
1	<p>Kementerian Pariwisata dan Ekonomi Kreatif/Badan Pariwisata dan Ekonomi Kreatif mendukung penyelenggaraan kejuaraan nasional e-sports "Ksatriya Mahardhika Student Championship 2023" yang akan mempertandingkan Lokapala, gim MOBA (Multiplayer Online Battle Arena) pertama buatan Indonesia. Berdasarkan data, pada 2022 transaksi gim di seluruh dunia mencapai Rp3.500 triliun. Selain itu, Angela mengatakan, kejuaraan nasional "Ksatriya Mahardhika Student Championship 2023" juga akan menjadi ruang ekspresi yang baik bagi para konten kreator juga generasi muda yang ingin terjun ke dalam dunia esports.</p>	<p><i>The Ministry of Tourism and Creative Economy/Tourism and Creative Economy Agency supports the holding of the national e-sports championship "Ksatriya Mahardhika Student Championship 2023" which will compete in Lokapala, the first MOBA (Multiplayer Online Battle Arena) game made in Indonesia. Based on data, in 2022 game transactions around the world will reach IDR 3,500 trillion. Apart from that, Angela said, the national championship "Ksatriya Mahardhika Student Championship 2023" will also be a good space for expression for content creators as well as the younger generation who want to get involved in the world of esports.</i></p>



2	<p>Dua orang anggota polisi lalu lintas (Polantas) mengambil alih kemudi ambulans di ruas tol Pekanbaru-Dumai demi mengantar pasien darurat yang mengalami patah tulang. "Ternyata sopir ambulans, maka saya dan senior saya mendekat, sopir ambulans saya lihat posisinya sedang memejamkan mata," ungkap Bripka Fernandoz. "Sopir ini bilang dia lagi membawa pasien patah pinggang rujukan Rumah Sakit Bagansiapi-api, Rokah Hilir yang mau dioperasi pada pagi hari," ujarnya.</p>	<p><i>Two members of the traffic police (Polantas) took over the wheel of an ambulance on the Pekanbaru-Dumai toll road to transport an emergency patient who had broken bones. "It turned out to be an ambulance driver, so my senior and I approached, I saw the ambulance driver was closing his eyes," said Bripka Fernandoz. "This driver said he was carrying a patient with a broken hip referred to Bagansiapi-api Hospital, Rokah Hilir who wanted surgery in the morning," he said.</i></p>
3	<p>Penjabat (Pj) Bupati Flores Timur, Nusa Tenggara Timur (NTT), Doris Alexander Rihi, mengimbau warga tidak makan nasi tiap Jumat. Salah satu alasannya, mengantisipasi krisis pangan dan ancaman kekeringan. Surat yang dikeluarkan pada Kamis (21/9) tersebut ditujukan kepada pimpinan Organisasi Perangkat Daerah (OPD), camat, kepala desa/lurah, serta pimpinan BUMN dan BUMD di seluruh Kabupaten Flores Timur.</p>	<p><i>Acting Regent of East Flores, East Nusa Tenggara (NTT), Doris Alexander Rihi, appealed to residents not to eat rice every Friday. One of the reasons is anticipating a food crisis and the threat of drought. The letter issued on Thursday (21/9) was addressed to leaders of Regional Apparatus Organizations (OPD), sub-district heads, village/sub-district heads, as well as leaders of BUMN and BUMD throughout East Flores Regency.</i></p>

#### 4.6. Penerapan Tahap Evaluasi

Tahap evaluasi dimulai dengan *Lalu* untuk mengukur kesamaan hasil ringkasan dengan teks asli, digunakan *ROUGE score* untuk mencari *score* setiap contoh untuk melihat apakah hasil ringkasan mirip dengan teks awal.

```
from rouge import Rouge
import pandas as pd

rouge = Rouge()
rouge_scores_bert = rouge.get_scores(summary, teks, avg=True)
pd.DataFrame(rouge_scores_bert)
```

Hasil rata-rata *ROUGE score* yang didapatkan adalah  $r = 0.29$ ,  $p = 0.93$ ,  $f = 0.31$ . Dimana  $r$ ,  $p$ , dan  $f$  dalam *ROUGE* adalah singkatan dari *Recall* (Akurasi), *Precision* (Presisi), dan *F-1 score* (Nilai F-1). Ketiganya merupakan ukuran kualitas terjemahan mesin.

- Akurasi adalah proporsi terjemahan yang benar yang benar-benar teridentifikasi. Dengan kata lain, ini adalah persentase kata dalam terjemahan manusia yang juga ditemukan dalam terjemahan mesin.
- Presisi adalah proporsi terjemahan yang teridentifikasi yang benar-benar benar. Dengan kata lain, ini adalah persentase kata dalam terjemahan mesin yang juga ditemukan dalam terjemahan manusia.
- Nilai F-1 adalah rata-rata harmonis dari akurasi dan presisi. Ini adalah ukuran yang menyeimbangkan pentingnya akurasi dan presisi.

Lalu, mempersiapkan model *Stable Diffusion* dengan mengimport *library-library* yang diperlukan untuk menjalankan model *Stable Diffusion* yang telah dilatih.

```
from diffusers import StableDiffusionPipeline
from IPython.display import clear_output
```

*Library-library* yang ada pada *code snippet* tersebut digunakan untuk men *download* dan *load* model *Stable Diffusion* yang telah dilatih dari *huggingface*.

```
#Stable Diffusion
model_id = "Nikodhrmwn/sampul-berita-indonesia"
pipe = StableDiffusionPipeline.from_pretrained(model_id,
torch_dtype=torch.float16,
use_auth_token="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")
```

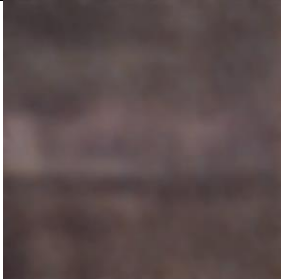
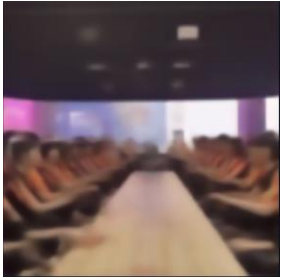

Kemudian, untuk menghasilkan gambar dari model yang sudah diambil dari *huggingface*. Variabel *pipe* yang digunakan untuk merepresentasikan sebuah *pipeline* dipindahkan menuju GPU/VGA untuk mempercepat proses *inference*.

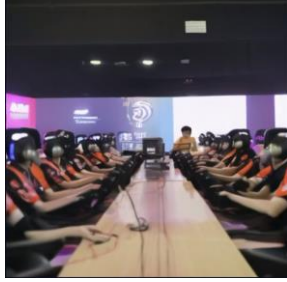



```
pipe = pipe.to("cuda") # device used to make inference
images = pipe("en_prompt + ', berita'",
num_inference_steps=20)
image = pipe(en_prompt, negative_prompt='texts, brand,
messages, error messages').images[0]
image.save("inference.png")
```

Disini model diberikan *parameter* berupa hasil terjemahan *en\_prompt* dan tambahan *string* “, berita” yang merupakan *class* dari gambar-gambar *dataset* yang digunakan dalam pelatihan model. Gambar dihasilkan dengan melakukan *denoising* pada sebuah *gaussian noise* yang dihasilkan secara random sambil dituntun oleh *en\_prompt* secara berulang-ulang sebanyak 20 *inference steps*. Hasil gambar tersebut kemudian disimpan dengan nama *inference.png*.

**Tabel 4.5** menunjukan hasil gambar dari hasil penerjemahan ringkasan ketiga contoh.

**Tabel 4.5** Tabel *inference*.

Contoh	Ringkasan (Bahasa Inggris)	Gambar (Setiap 5 steps)
1	<p>The Ministry of Tourism and Creative Economy/Tourism and Creative Economy Agency supports the holding of the national e-sports championship "Ksatriya Mahardhika Student Championship 2023" which will compete in Lokapala, the first MOBA (Multiplayer Online Battle Arena) game made in Indonesia. Based on data, in 2022 game transactions around the world will reach IDR 3,500 trillion. Apart from that, Angela said, the national championship "Ksatriya Mahardhika Student Championship 2023" will also be a good space for expression for content creators as well as the younger generation who want to get involved in the world of esports.</p>	<ol style="list-style-type: none"> <li>  <p>Pada iterasi 5, proses denoising sudah mulai memperlihatkan <i>outline</i> atau bayangan gambar akhir nantinya.</p> </li> <li>  <p>Pada iterasi 10, gambar sudah mulai terlihat lebih jelas.</p> </li> <li>  <p>Pada iterasi 15, gambar mulai terlihat lebih jelas lagi tetapi masih buram secara keseluruhan.</p> </li> </ol>

		 <p>4. Pada iterasi 20, gambar sudah terlihat jelas. Waktu yang dibutuhkan untuk menjalankan 20 <i>steps</i> jika menggunakan google colab (<i>free</i>) yang menggunakan GPU Nvidia Tesla T4 sekitar 7-8s atau detik.</p>
2	<p>Two members of the traffic police (Polantas) took over the wheel of an ambulance on the Pekanbaru-Dumai toll road to transport an emergency patient who had broken bones. "It turned out to be an ambulance driver, so my senior and I approached, I saw the ambulance driver was closing his eyes," said Bripka Fernandoz. "This driver said he was carrying a patient with a broken hip referred to Bagansiapi-api Hospital, Rokah Hilir who wanted surgery in the morning," he said.</p>	 <p>1.</p>  <p>2.</p>  <p>3.</p>

		<div data-bbox="820 465 847 504" data-label="Text">4.</div> <div data-bbox="868 230 1155 495" data-label="Image"> </div>
3	<p>Acting Regent of East Flores, East Nusa Tenggara (NTT), Doris Alexander Rihi, appealed to residents not to eat rice every Friday. One of the reasons is anticipating a food crisis and the threat of drought. The letter issued on Thursday (21/9) was addressed to leaders of Regional Apparatus Organizations (OPD), sub-district heads, village/sub-district heads, as well as leaders of BUMN and BUMD throughout East Flores Regency.</p>	<div data-bbox="820 792 847 831" data-label="Text">1.</div> <div data-bbox="868 539 1155 824" data-label="Image"> </div> <div data-bbox="820 1122 847 1160" data-label="Text">2.</div> <div data-bbox="868 864 1155 1151" data-label="Image"> </div> <div data-bbox="820 1451 847 1489" data-label="Text">3.</div> <div data-bbox="868 1193 1155 1480" data-label="Image"> </div> <div data-bbox="820 1765 847 1803" data-label="Text">4.</div> <div data-bbox="868 1520 1155 1794" data-label="Image"> </div>

## BAB V KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Kesimpulan yang dapat diambil dalam penelitian ini adalah sebagai berikut:

1. Peringkasan teks menggunakan *library bert-extractive-summarizer* yang memanfaatkan model BERT dapat meringkas teks tanpa kehilangan makna utama dari sebuah teks.
2. Dikarenakan batasan dari 512 *tokens* pada model BERT, hasil ringkasan dari sebuah teks yang terdiri atas lebih dari 512 *tokens* bukanlah hasil ringkasan teks secara keseluruhan.
3. Dikarenakan *library bert-extractive-summarizer* hanya memilih 1 kalimat paling dekat dengan *centroid*, ringkasan yang dihasilkan tidak bagus dan masih belum bisa merepresentasikan teks secara keseluruhan.
4. Gambar yang dihasilkan model *Stable Diffusion* sesuai dengan teks ringkasan dari model BERT yang dimasukkan sebagai *input*
5. *Fine-tuning* dengan menggunakan metode *dreambooth* juga berhasil mempengaruhi *output* dari model yang dapat dilihat pada *output* gambar dari teks contoh 2 dan 3.
6. Pada hasil gambar contoh 2 dari model *Stable Diffusion* dapat diamati bahwa model ini belum mengenali tulisan/huruf/bahasa sehingga terkadang *output* gambar mengandung tulisan atau huruf-huruf acak yang tidak memiliki makna.
7. Sistem berhasil mengotomatisasi dan mempercepat proses peringkasan, penerjemahan, dan menghasilkan gambar yang sesuai dengan ringkasan yang telah dihasilkan dengan memanfaatkan model BERT, *library GoogleTranslate*, dan model *Stable Diffusion*.

## 5.2. Saran

Saran yang dapat diberikan untuk membantu pengembangan sistem selanjutnya adalah sebagai berikut:

1. Diharapkan pada penelitian selanjutnya dapat menggunakan model baru yang lebih mampu, lebih ringan, dan lebih mudah untuk dipersiapkan.
2. Diharapkan pada penelitian selanjutnya dapat menggunakan model BERT yang di *fine-tune* untuk meringkas atau menggunakan model lain yang memiliki batasan *token* lebih dari 512 *tokens* dan dapat menghasilkan kualitas meringkas lebih baik.



## DAFTAR PUSTAKA

- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin (2017). “*Attention is all you need*”. [arXiv:1706.03762](https://arxiv.org/abs/1706.03762) [cs.CL]
- Athar Mohammed & Ali H. Ali (2021). “*Survey of BERT (Bidirectional Encoder Representation Transformer) types*“. in *Journal of Physics: Conference Series*. 1963. 012173.10.1088/1742- 6596/1963/1/012173.
- Chin-Yew Lin. (2004). “*ROUGE: A Package for Automatic Evaluation of Summaries*”. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.
- Derek Miller. (2019). “*Leveraging BERT for Extractive Text Summarization on Lectures*”. [arXiv:1906.04165](https://arxiv.org/abs/1906.04165) [cs.CL]
- Farshad Kyoomarsi, Hamid Khosravi, Esfandiar Eslami and Pooya Khosravayan Dehkordy, “*Optimizing Text Summarization Based on Fuzzy Logic*”. In *proceedings of Seventh IEEE/ACIS International Conference on Computer and Information Science, IEEE, University of Shahid Bahonar Kerman, UK*, 347-352, 2008.
- Feliks Victor Parningotan Samosir, Hapnes Toba, Mewati Ayub (2022).”*BESKlus : BERT Extractive Summarization with K-Means Clustering in Scientific Paper*”. *Jurnal Teknik Informatika dan Sistem Informasi* Volume 8 Nomor 1 April 2022.
- Franky Halim, Liliana, Kartika Gunadi (2022). “*Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT*”.
- Gunes Erkan and Dragomir R. Radev (2004). “*LexRank: Graph-based Lexical Centrality as Salience in Text Summarization*”. *Journal of Artificial Intelligence Research* 22 (2004) 457-479.

- Ian J. Goodfellow, Jean Pouget-Abadie , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair , Aaron Courville, Yoshua Bengio (2014). “*Generative Adversarial Networks*”. [arXiv:1406.2661](https://arxiv.org/abs/1406.2661) [stat.ML]
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. “*BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*”. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assef, Saeid Safaei, Elizabeth D. Trippe, Juan B. Gutierrez, Krys Kochut (2017). “*Text Summarization Techniques: A Brief Survey*”. [arXiv:1707.02268](https://arxiv.org/abs/1707.02268) [cs.CL]
- Nataniel Ruiz, Yael Pritch, Yuanzhen Li, Michael Rubinstein, Varun Jampani, Kfir Aberman. (2023). “*DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation*”. [arXiv:2208.12242](https://arxiv.org/abs/2208.12242) [cs.CV]
- Prafulla Dhariwal & Alex Nichol (2021). “*Diffusion Models Beat GANs on Image Synthesis*”. [arXiv:2105.05233](https://arxiv.org/abs/2105.05233) [cs.LG]
- Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Bjorn Ommer (2022). “*Stable Diffusion*”. Munich : Ludwig Maximilian University of Munich & IWR. <https://github.com/CompVis/stable-diffusion>.
- Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Bjorn Ommer (2022). “*High-Resolution Image Synthesis with Latent Diffusion Models*”. Munich : Ludwig Maximilian University of Munich & IWR.
- Robin Zbinden (2022). “*Implementing and Experimenting with Diffusion Models for Text-to-Image Generation*”. [arXiv:2209.10948](https://arxiv.org/abs/2209.10948) [cs.CV]

- Sarah Lintang Sariwening (2020). “IndoBERT: *Transformer-based Model for Indonesian Language Understanding*”. Yogyakarta : Universitas Gadjah Mada.
- Thomas Indrias Biniam & Adam Morén (2021). “*Extractive Text Summarization of Norwegian News Articles Using BERT*”. LiU-ITN-TEK-A--21/016-SE.
- Wen Jun Zhang & Guosheng Yang & Yingzi Lin & Chunli Ji & Madan Gupta. (2018). “*On Definition of Deep Learning*”. 1-5. 10.23919/WAC.2018.8430387.
- Yang Gu dan Yanke Hu. (2019). “*Extractive Summarization with Very Deep Pretrained Language Model*”. *International Journal of Artificial Intelligence and Applications (IJAIA)*, Vol.10, No.2 .
- Yang Liu dan Mirella Lapata (2019). “*Text Summarization with Pretrained Encoders*”. *University of Edinburgh*.
- Zahra Khaerunnisa, Dr.techn. Ahmad Ashari, M.I.Kom;Yunita Sari, S.Kom., M.Sc., Ph.D. (2022). ”Perbandingan Penggunaan *Word Embedding Word2Vec, GloVe*, dan *FastText* dalam Peringkasan Teks Ekstraktif Berita Berbahasa Indonesia Menggunakan RNN”. Universitas Gadjah Mada.