

PERINGKASAN ABSTRAKTIF PADA DIALOG MENGGUNAKAN  
*FINE-TUNING PRE-TRAINED LANGUAGE MODEL BART*

Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Jurusan Teknik Informatika



Oleh :

HANIF SYAHRI RAMADHANI

NIM : 09021282126042

**Jurusan Teknik Informatika**

**FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**

**2025**

# HALAMAN PENGESAHAN

## SKRIPSI

### PERINGKASAN ABSTRAKTIF PADA DIALOG MENGGUNAKAN FINE-TUNING PRE-TRAINED LANGUAGE MODEL BART

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Teknik Informatika

Oleh:

**HANIF SYAHRI RAMADHANI**

**09021282126042**

**Pembimbing 1 : Desty Rodiah, S.Kom., M.T.**  
**NIP. 198912212020122011**

**Mengetahui**

**Ketua Jurusan Teknik Informatika**



**Hadipurnawan Satria, Ph.D**

**198004182020121001**

## TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Jumat tanggal 24 Januari 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Hanif Syahri Ramadhani  
NIM : 09021282126042  
Judul : Peringkasan Abstraktif pada dialog menggunakan *Fine-Tuning Pre-Trained Language Model BART*

dan dinyatakan LULUS

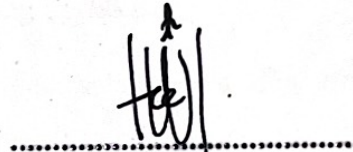
1. Ketua Penguji

Prof. Dr. Ermatita M.Kom  
NIP. 196709132006042001



2. Penguji

Hadipurnawan Satria, Ph.D.  
NIP. 198004182020121001



3. Pembimbing I

Desty Rediah, S.Kom., M.T.  
NIP. 198912212020122011



Mengetahui,

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D.  
NIP. 198004182020121001

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Hanif Syahri Ramadhani  
NIM : 09021282126042  
Program Studi : Teknik Informatika  
Judul Skripsi : Peringkasan Abstraktif pada dialog menggunakan *Fine Tuning Pre-Trained Language Model BART*

Hasil pengecekan *Software iThenticate/Turnitin* : 8%

Menyatakan bahwa laporan penelitian saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan penelitian ini, maka saya bersedia menerima sanksi akademik Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Palembang, 30 Januari 2025



Hanif Syahri Ramadhani  
NIM 09021282126042

## MOTTO DAN PERSEMBAHAN

*“When you want something, all the universe conspires in helping you to achieve it.”*

— **Paulo Coelho, *The Alchemist***

*“Sesungguhnya Allah menyukai orang-orang yang berbuat baik.”*

— **QS. Al-Baqarah: 195**

Kupersembahkan Karya Tulis ini kepada:

- Allah SWT
- Kedua Orang Tua dan Saudara
- Keluarga Besar
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

## ABSTRACT

*The massive increase in communication over digital platforms generates conversation data that contains a variety of important information and requires effective analysis methods to filter the information, one of which is abstractive summarization. This technique uses a natural language processing (NLP) approach to reorganize dialogue information into a more concise and coherent form. This research aims to develop an abstractive dialogue summarization system using pre-trained BART (Bidirectional and Auto-Regressive Transformers) models that are fine-tuned with a sequence-to-sequence approach on the DialogSum dataset. The dataset consists of 14,460 English dialogues, split for training and testing with 7:3 ratio. Evaluation using the ROUGE metrics shows that the average values on ROUGE-1 are precision 53.16%, recall 49.90% and f1-score 50.44%. On ROUGE-2, precision was 25.66%, recall 24.95% and f1-score 24.76%. The results show that the system is able to generate dialogue summaries with the highest performance on the ROUGE-1 metric which measures the number of individual words absorbed, in line with the focus of abstractive summarization which performs new sentence formation through paraphrasing.*

**Keywords:** *Abstractive Summarization, Dialogue Summarization, BART, Natural Language Processing, sequence-to-sequence, ROUGE Score.*

## ABSTRAK

Komunikasi melalui *platform* digital yang meningkat secara masif menghasilkan data percakapan yang memuat berbagai informasi penting dan memerlukan metode analisis efektif untuk menyaring informasinya, salah satu metodenya adalah peringkasan abstraktif. Teknik ini menggunakan pendekatan *natural language processing* (NLP) untuk menyusun kembali informasi dialog menjadi lebih ringkas dan koheren. Penelitian ini bertujuan mengembangkan sistem peringkasan dialog abstraktif menggunakan model *pre-trained* BART (*Bidirectional and Auto-Regressive Transformers*) yang di-*fine-tune* dengan pendekatan *sequence-to-sequence* pada *dataset* DialogSum. *Dataset* terdiri dari 14.460 dialog berbahasa Inggris, yang dibagi untuk pelatihan dan pengujian dengan rasio 7:3. Evaluasi menggunakan metrik ROUGE menunjukkan rata-rata nilai pada ROUGE-1 sebesar *precision* 53,16%, *recall* 49,90%, dan *f1-score* 50,44%. Pada ROUGE-2, diperoleh *precision* 25,66%, *recall* 24,95%, dan *f1-score* 24,76%. Hasil menunjukkan bahwa sistem mampu menghasilkan ringkasan dialog dengan performa tertinggi pada metrik ROUGE-1 yang mengukur jumlah kata tunggal yang diserap, sejalan dengan fokus peringkasan abstraktif yang melakukan pembentukan kalimat baru melalui parafrase.

**Kata Kunci:** Peringkasan Abstraktif, *Dialogue Summarization*, BART, *Natural Language Processing*, *sequence-to-sequence*, ROUGE Score.

## KATA PENGANTAR

Puji dan syukur ke hadirat Allah SWT atas segala nikmat, rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan Skripsi yang berjudul “Peringkasan Abstraktif pada Dialog menggunakan *Fine-Tuning Pre-Trained Language Model BART*”. Skripsi ini disusun sebagai salah satu syarat kelulusan program Strata-1 pada Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam menyelesaikan skripsi ini, penulis menerima bantuan, bimbingan dan dukungan dari banyak pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT atas rahmat dan nikmat-Nya penulis dapat menyelesaikan skripsi ini dengan baik.
2. Bapak Suryanto, S.M. Ibu Nurlelawati, Adik Rizky Muhammad Tsani, serta Keluarga Besar Suwarno dan Jumadi yang telah mendoakan, selalu memberi semangat, motivasi, dan nasihat untuk menyelesaikan skripsi ini.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Hadipurnawan Satria, Ph.D. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Desty Rodiah, S.Kom., M.T. selaku dosen pembimbing skripsi yang telah membimbing dan memberikan arahan kepada penulis dalam pengerjaan skripsi.
6. Bapak Dr. M. Fachrurrozi, S.Si., M.T. selaku dosen pembimbing akademik yang telah memberikan bimbingan selama proses kegiatan perkuliahan.
7. Seluruh Dosen, Admin, dan Staf Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.



8. Teman seperjuangan “Project Penting” Ahmad Azhari, Dellin Irawan, Risky Armansyah, Vanya Terra Ardani, dan Yolendri Anisyahfitri yang telah menemani, memberikan motivasi, dan dukungan kepada penulis.
9. Jihaan Jasmine Jahroo, Cindy Angreini, dan Yasmin Nuha Khoirunnisa selaku teman penulis yang selalu menjadi tempat bertukar pendapat, memberikan semangat dan motivasi.
10. Teman – teman Senja 2.0, client OnlyNet., serta teman – teman Teknik Informatika angkatan 2021 yang telah memberikan motivasi, dan dukungan selama perkuliahan.
11. Pihak – pihak lain yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak sekali kekurangan dikarenakan kurangnya pengalaman dan pengetahuan penulis. Oleh karena itu penulis mengharapkan kritik dan saran yang membangun guna kemajuan penelitian selanjutnya. Semoga tugas akhir ini dapat bermanfaat. Terima kasih.

Indralaya, 30 Januari 2025

Penulis

Hanif Syahri Ramadhani

## DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN KOMPREHENSIF .....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT .....	vi
ABSTRAK .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI .....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xiv
BAB I PENDAHULUAN .....	I-1
1.1 Pendahuluan .....	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah .....	I-3
1.4 Tujuan Penelitian.....	I-3
1.5 Manfaat Penelitian .....	I-4
1.6 Batasan Masalah.....	I-4
1.7 Sistematika Penulisan .....	I-4
1.8 Kesimpulan .....	I-6
BAB II KAJIAN LITERATUR .....	II-1
2.1 Pendahuluan .....	II-1
2.2 Landasan Teori .....	II-1
2.2.1 <i>Text Summarization</i> .....	II-1
2.2.2 Pra-Pemrosesan Teks.....	II-3
2.2.3 <i>Pre-Trained Language Model BART</i> .....	II-4
2.2.4 <i>Sequence-to-Sequence</i> .....	II-5
2.2.5 <i>ROUGE Score</i> .....	II-7

2.3	<i>Rational Unified Process</i> .....	II-9
2.4	Penelitian Lain yang Relevan.....	II-11
2.5	Kesimpulan .....	II-12
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan .....	III-1
3.2	Pengumpulan Data .....	III-1
3.2.1	Jenis dan Sumber Data .....	III-1
3.2.2	Metode Pengumpulan Data .....	III-2
3.3	Tahapan Penelitian .....	III-2
3.3.1	Kerangka Kerja .....	III-3
3.3.2	Kriteria Pengujian .....	III-4
3.3.3	Format Data Pengujian.....	III-4
3.3.4	Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-5
3.3.5	Pengujian Penelitian.....	III-6
3.3.6	Analisis Hasil Pengujian dan Membuat Kesimpulan.....	III-6
3.4	Metode Pengembangan Perangkat Lunak.....	III-7
3.4.1	Fase Insepsi .....	III-7
3.4.2	Fase Elaborasi .....	III-7
3.4.3	Fase Konstruksi.....	III-8
3.4.4	Fase Transisi.....	III-8
3.5	Manajemen Proyek Penelitian .....	III-9
3.6	Kesimpulan .....	III-12
BAB IV PENGEMBANGAN PERANGKAT LUNAK .....		IV-1
4.1	Pendahuluan .....	IV-1
4.2	Fase Insepsi .....	IV-1
4.2.1	Permodelan Bisnis.....	IV-1
4.2.2	Kebutuhan Sistem .....	IV-2
4.2.3	Analisis dan Desain.....	IV-3
4.3	Fase Elaborasi .....	IV-13
4.3.1	Permodelan Bisnis.....	IV-13

4.3.2	Perancangan Data.....	IV-13
4.3.3	Kebutuhan Sistem .....	IV-13
4.3.4	Diagram Aktivitas ( <i>Activity Diagram</i> ).....	IV-14
4.3.5	Diagram Urutan ( <i>Sequence Diagram</i> ) .....	IV-17
4.3.6	Perancangan Antarmuka.....	IV-21
4.4	Fase Konstruksi.....	IV-23
4.4.1	Diagram Kelas.....	IV-23
4.4.2	Implementasi .....	IV-25
4.4.2.1	Implementasi Kelas .....	IV-25
4.4.2.2	Implementasi Antarmuka .....	IV-26
4.5	Fase Transisi.....	IV-28
4.5.1	Permodelan Bisnis.....	IV-28
4.5.2	Rencana Pengujian .....	IV-28
4.5.3	Implementasi .....	IV-28
4.6	Kesimpulan .....	IV-30
BAB V	HASIL DAN ANALISIS PENELITIAN .....	V-1
5.1	Pendahuluan .....	V-1
5.2	Hasil Percobaan Penelitian.....	V-1
5.2.1	Konfigurasi Percobaan .....	V-1
5.2.2	Data Hasil Konfigurasi.....	V-2
5.3	Analisis Hasil Penelitian .....	V-4
5.4	Kesimpulan .....	V-5
BAB VI	KESIMPULAN DAN SARAN.....	VI-1
6.1	Pendahuluan .....	VI-1
6.2	Kesimpulan .....	VI-1
6.3	Saran.....	VI-2
DAFTAR PUSTAKA .....		xv
LAMPIRAN.....		xviii

## DAFTAR TABEL

<b>Tabel II-1.</b> Contoh Text Cleaning.....	II-3
<b>Tabel II-2.</b> Contoh Tokenizing .....	II-4
<b>Tabel III-1.</b> Rancangan Data Masukan .....	III-5
<b>Tabel III-2.</b> Alat Bantu Penelitian .....	III-5
<b>Tabel III-3.</b> Evaluasi Data Hasil Pengujian.....	III-6
<b>Tabel III-4.</b> Evaluasi Hasil Pengujian .....	III-6
<b>Tabel III-5.</b> Perencanaan Aktivitas Penelitian dalam Bentuk WBS.....	III-10
<b>Tabel IV-1.</b> Kebutuhan Fungsional .....	IV-2
<b>Tabel IV-2.</b> Kebutuhan Non Fungsional.....	IV-2
<b>Tabel IV-3.</b> Contoh Data .....	IV-4
<b>Tabel IV-4.</b> Contoh Pembersihan Teks .....	IV-5
<b>Tabel IV-5.</b> Contoh Tokenisasi .....	IV-5
<b>Tabel IV-6.</b> Contoh Vektorisasi BART .....	IV-6
<b>Tabel IV-7.</b> Definisi Use Case .....	IV-8
<b>Tabel IV-8.</b> Skenario Use Case Meringkas Dialog.....	IV-9
<b>Tabel IV-9.</b> Skenario Use Case Menampilkan Riwayat Ringkasan. ....	IV-10
<b>Tabel IV-10.</b> Skenario Use Case Menampilkan Contoh Dataset.....	IV-11
<b>Tabel IV-11.</b> Skenario Use Case Menampilkan Performa Model. ....	IV-12
<b>Tabel IV-12.</b> Kebutuhan Sistem .....	IV-14
<b>Tabel IV-13.</b> Deskripsi Implementasi Kelas.....	IV-25
<b>Tabel IV-14.</b> Rencana Pengujian Use Case .....	IV-28
<b>Tabel IV-15.</b> Hasil Pengujian Use Case .....	IV-29
<b>Tabel V-1.</b> Hasil Ringkasan Sistem .....	V-2
<b>Tabel V-2.</b> Hasil Evaluasi ROUGE Score .....	V-3
<b>Tabel V-3.</b> Rata-rata ROUGE Score .....	V-4

## DAFTAR GAMBAR

<b>Gambar II-1.</b> Arsitektur Sistem Extractive Summarization .....	II-2
<b>Gambar II-2.</b> Arsitektur Sistem Abstractive Summarization .....	II-3
<b>Gambar II-3.</b> Arsitektur Pre-Trained Language Model BART .....	II-5
<b>Gambar II-4.</b> Arsitektur Sequence-to-sequence .....	II-6
<b>Gambar II-5.</b> Arsitektur Rational Unified Process (Anwar, 2014).....	II-9
<b>Gambar III-1.</b> Tahapan Penelitian .....	III-2
<b>Gambar III-2.</b> Kerangka Kerja Penelitian .....	III-3
<b>Gambar IV-1.</b> Use Case Diagram .....	IV-7
<b>Gambar IV-2.</b> Activity Diagram Meringkas Dialog.....	IV-15
<b>Gambar IV-3.</b> Activity Diagram Menampilkan Riwayat Ringkasan .....	IV-16
<b>Gambar IV-4.</b> Activity Diagram Menampilkan Contoh Dataset.....	IV-16
<b>Gambar IV-5.</b> Activity Diagram Menampilkan Performa Model .....	IV-17
<b>Gambar IV-6.</b> Sequence Diagram Meringkas Dialog.....	IV-18
<b>Gambar IV-7.</b> Sequence Diagram Menampilkan Riwayat Ringkasan .....	IV-19
<b>Gambar IV-8.</b> Sequence Diagram Menampilkan Contoh Dataset.....	IV-20
<b>Gambar IV-9.</b> Sequence Diagram Menampilkan Performa Model .....	IV-20
<b>Gambar IV-10.</b> Antarmuka Peringkat Dialog .....	IV-21
<b>Gambar IV-11.</b> Antarmuka Riwayat Ringkasan .....	IV-21
<b>Gambar IV-12.</b> Antarmuka Contoh Dataset .....	IV-22
<b>Gambar IV-13.</b> Antarmuka Performa Model.....	IV-22
<b>Gambar IV-14.</b> Class Diagram .....	IV-24
<b>Gambar IV-15.</b> Antarmuka Peringkat Dialog .....	IV-26
<b>Gambar IV-16.</b> Antarmuka Riwayat Ringkasan.....	IV-26
<b>Gambar IV-17.</b> Antarmuka Contoh Dataset .....	IV-27
<b>Gambar IV-18.</b> Antarmuka Performa Model - Rouge Score.....	IV-27
<b>Gambar IV-19.</b> Antarmuka Performa Model - Evaluation Metrics.....	IV-27

# BAB I

## PENDAHULUAN

### 1.1 Pendahuluan

Pada bab ini akan menguraikan latar belakang penelitian, rumusan masalah penelitian, tujuan dan manfaat penelitian serta batasan masalah penelitian. Bagian ini akan berisi penjelasan umum dan memberikan gambaran keseluruhan terkait pembahasan penelitian pada bab-bab berikutnya.

### 1.2 Latar Belakang Masalah

Saat ini komunikasi melalui *platform* digital sudah menjadi salah satu kebutuhan dalam kehidupan sehari-hari. Data percakapan yang dihasilkan melalui berbagai *platform* komunikasi seperti pesan instan, forum diskusi, dan media sosial meningkat secara masif. Data ini menyimpan informasi yang sangat bernilai yang dapat dimanfaatkan untuk berbagai tujuan, termasuk analisis isi percakapan dan penyusunan ringkasan dialog.

Terdapat tantangan besar yang dihadapi dalam analisis data percakapan, yaitu kompleksitas dan volume informasi yang terkandung di dalamnya. Penyajian informasi yang terlalu banyak dan gaya bahasa yang beragam, sering kali menyulitkan pengguna untuk mendapatkan esensi dari percakapan tersebut (Feng et al., 2021). Peringkasan menjadi salah satu solusi yang tepat untuk masalah ini.

Salah satu metode peringkasan yaitu *abstractive summarization* (peringkasan abstraktif). Teknik ini merupakan bagian dari *Natural Language Processing* (NLP) yang dirancang untuk menghasilkan ringkasan teks yang tidak hanya memotong

dan menempelkan bagian-bagian tertentu, tetapi juga menyusun kembali informasi penting dalam bentuk yang lebih ringkas dan koheren (Gupta & Gupta, 2019).

Dalam *abstractive dialogue summarization*, penggunaan *pre-trained language model* sangat penting. Pada penelitian ini, jenis *pre-trained language model* yang akan digunakan ialah BART (*Bidirectional and Auto-Regressive Transformers*). Menurut Macharla et al. (2023) berdasarkan hasil penelitian dengan data *radiology reports*, BART adalah salah satu model *deep learning* terbaru yang telah menunjukkan performa mengesankan dalam tugas-tugas *Natural Language Processing* (NLP) yang dievaluasi menggunakan metrik Recall-Oriented Understudy for Gisting Evaluation (ROUGE) mencapai hasil skor ROUGE-1 sebesar 32.33%. Pada penelitian *abstractive dialogue summarization* yang dilakukan Chen & Yang (2020), model BART menghasilkan performa dengan skor ROUGE-1 dan 2 sebesar 48.2% dan 24.5%, dan pada penelitian *Dialogue Summarization using BART* oleh Lundberg et al. (2022), model BART menghasilkan skor ROUGE-1 dan 2 sebesar 47.29% dan 21.65%.

Namun demikian, masih belum banyak penelitian model BART dalam konteks ringkasan dialog, khususnya pada *dataset* yang mengandung percakapan sehari-hari seperti DialogSum. Salah satu penelitian oleh Chauhan et al. (2022), data DialogSum dengan model Topic\_FT\_PEGASUS mendapatkan skor ROUGE-1 dan 2 sebesar 49.42% dan 21.81%. Oleh karena itu, penelitian ini difokuskan pada upaya untuk mengoptimalkan model BART dalam menghasilkan ringkasan abstraktif yang akurat dan informatif dari *dataset* DialogSum melalui proses *fine-tuning* dan akan di evaluasi kinerja menggunakan metrik ROUGE Score.



Berdasarkan penjelasan di atas, peringkasan abstraktif menggunakan model BART menjadi topik yang menarik untuk dikembangkan. Penelitian ini akan membangun sistem peringkasan abstraktif pada teks dialog menggunakan *pre-trained language model* BART yang di *fine-tuning* dengan *seq2seq*.

### 1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Bagaimana mengembangkan sistem peringkasan abstraktif pada dialog menggunakan *pre-trained language model* BART yang di *fine-tuning* dengan *seq2seq*?
2. Bagaimana kinerja sistem peringkasan abstraktif pada dialog menggunakan *pre-trained language model* BART yang di *fine-tuning* dengan *seq2seq* berdasarkan ROUGE Score?

### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menghasilkan sebuah sistem yang dapat melakukan peringkasan abstraktif pada dialog dengan menggunakan *pre-trained language model* BART yang di *fine-tuning* dengan *seq2seq*.
2. Mengetahui kinerja sistem peringkasan abstraktif pada dialog menggunakan *pre-trained language model* BART yang di *fine-tuning* dengan *seq2seq* berdasarkan ROUGE Score.

### 1.5 Manfaat Penelitian

Manfaat yang didapat dari penelitian ini adalah sebagai berikut:

1. Sistem dapat melakukan peringkasan abstraktif pada data dialog.
2. Masyarakat dapat menggunakan sistem untuk meringkas data dialog.
3. Hasil penelitian dapat menjadi rujukan pada penelitian terkait.

### 1.6 Batasan Masalah

Agar permasalahan tidak menyimpang dari batasan yang telah ditetapkan, maka adapun batasan dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah *dataset* yang berisi 14460 percakapan berbahasa inggris dari DialogSum<sup>1)</sup>.
2. *Pre-trained Language Model* BART yang digunakan merupakan *pretrained model* bart-large-xsum oleh Facebook.
3. Model di *fine-tuning* menggunakan pendekatan *sequence-to-sequence*.

### 1.7 Sistematika Penulisan

Struktur penulisan dalam penelitian ini mengikuti pedoman penulisan proposal skripsi yang ditetapkan oleh Fakultas Ilmu Komputer Universitas Sriwijaya, yang mencakup:

---

<sup>1)</sup> <https://aclanthology.org/2021.findings-acl.449>

## **BAB I. PENDAHULUAN**

Pada bab ini, akan diuraikan mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah. Pokok bahasan pada bab ini akan menjadi landasan pada bab selanjutnya.

## **BAB II. KAJIAN LITERATUR**

Bab ini akan menjelaskan landasan teori yang digunakan dalam rangka menunjang penelitian, seperti penjelasan mengenai Peringkasan Abstraktif, Pra-Pemrosesan Teks, *Pre-trained Language Model* BART, *Sequence-to-Sequence*, *ROUGE Score*, serta literatur-literatur yang relevan terhadap penelitian ini.

## **BAB III. METODOLOGI PENELITIAN**

Pada bab ini, akan dibahas mengenai metodologi penelitian yang menguraikan rencana tahapan penelitian serta manajemen proyek penelitian. Rencana tahapan penelitian akan digunakan sebagai landasan pada pengembangan perangkat lunak.

## **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Bab ini akan menguraikan proses yang dilakukan dalam mengimplementasikan *pre-trained language model* BART yang telah di *fine-tuning* dengan pendekatan seq2seq dalam pengembangan perangkat lunak.

## **BAB V. HASIL DAN ANALISIS PENELITIAN**

Bab ini akan memaparkan hasil pengujian dan pengembangan perangkat lunak yang telah dilakukan sesuai dengan format pengujian yang ditetapkan. Hasil dari analisis penelitian ini akan menjadi dasar dalam penarikan kesimpulan.

## **BAB VI. KESIMPULAN DAN SARAN**

Bab ini akan memaparkan kesimpulan yang didapat berdasarkan penelitian dan analisis yang telah dilakukan sebelumnya. Bab ini juga akan memberikan saran yang nantinya diharapkan dapat menjadi masukan untuk penelitian berikutnya yang mengimplementasikan pendekatan atau model yang sama.

### **1.8 Kesimpulan**

Pada bab ini, telah diuraikan tentang latar belakang penelitian, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penulisan yang akan dijadikan sebagai landasan dan pokok penelitian.

## DAFTAR PUSTAKA

- Barbella, M., & Tortora, G. (2022). Rouge Metric Evaluation for Text Summarization Techniques. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4120317>
- Chauhan, V., Roy, P., Dey, L., & Goel, T. (2022). *TCS\_WITM\_2022 @ DialogSum : Topic oriented Summarization using Transformer based Encoder Decoder Model*. <https://huggingface.co/google/>
- Chen, J., & Yang, D. (2020). *Multi-View Sequence-to-Sequence Models with Conversational Structure for Abstractive Dialogue Summarization*. <http://arxiv.org/abs/2010.01672>
- Cui, L., Wu, Y., Liu, S., Zhang, Y., & Zhou, M. (2020). *MuTual: A Dataset for Multi-Turn Dialogue Reasoning*. <https://doi.org/10.48550/arXiv.2004.04494>
- El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic text summarization: A comprehensive survey. In *Expert Systems with Applications* (Vol. 165). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2020.113679>
- Feng, X., Feng, X., & Qin, B. (2021). *A Survey on Dialogue Summarization: Recent Advances and New Frontiers*. <http://arxiv.org/abs/2107.03175>
- Garg, N., & Sharma, K. (2022). Text pre-processing of multilingual for sentiment analysis based on social network data. *International Journal of Electrical and Computer Engineering*, 12(1), 776–784. <https://doi.org/10.11591/ijece.v12i1.pp776-784>
- Gulli, A., & Pal, S. (2017). *Deep Learning with Keras*. Packt Publishing. [https://books.google.co.id/books?id=u\\_fVswEACAAJ](https://books.google.co.id/books?id=u_fVswEACAAJ)
- Gupta, S., & Gupta, S. K. (2019). Abstractive summarization: An overview of the state of the art. In *Expert Systems with Applications* (Vol. 121, pp. 49–65). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2018.12.011>
- Halimah, Surya Agustian, & Siti Ramadhani. (2022). Peringkasan teks otomatis (automated text summarization) pada artikel berbahasa indonesia menggunakan algoritma lexrank. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(3), 371–381. <https://doi.org/10.37859/coscitech.v3i3.4300>
- Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, 25(1), 114–146. <https://doi.org/10.1177/1094428120971683>
- Kim, J., Yun, J. H., & Kim, H. C. (2020). Anomaly detection for industrial control systems using sequence-to-sequence neural networks. *Lecture*

*Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11980 LNCS, 3–18. [https://doi.org/10.1007/978-3-030-42048-2\\_1](https://doi.org/10.1007/978-3-030-42048-2_1)

- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). *BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension*. <http://arxiv.org/abs/1910.13461>
- Li, Y., Su, H., Shen, X., Li, W., Cao, Z., & Niu, S. (2017). *DailyDialog: A Manually Labelled Multi-turn Dialogue Dataset*. <https://arxiv.org/abs/1710.03957>
- Lin, C.-Y. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. *Text Summarization Branches Out*, 74–81.
- Lin, H., & Ng, V. (2019). Abstractive Summarization: A Survey of the State of the Art. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9815–9822. <https://doi.org/10.1609/aaai.v33i01.33019815>
- Lundberg, C., Sánchez Viñuela, L., & Biales, S. (2022). Dialogue Summarization using BART. In S. Shaikh, T. Ferreira, & A. Stent (Eds.), *Proceedings of the 15th International Conference on Natural Language Generation: Generation Challenges* (pp. 121–125). Association for Computational Linguistics. <https://aclanthology.org/2022.inlg-genchal.17>
- Macharla, S., Madamanchi, A., & Kancharla, N. (2023). nav-nlp at RadSum23: Abstractive Summarization of Radiology Reports using BART Finetuning. *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks*, 541–544. <https://doi.org/10.18653/v1/2023.bionlp-1.55>
- Ngurah, G., Nata, M., Yudiastara, P., Stmik, ), Bali, S., & Raya Puputan, J. (2017). *Konferensi Nasional Sistem & Informatika*.
- Shafiee, S., Wautelet, Y., Hvam, L., Sandrin, E., & Forza, C. (2020). Scrum versus Rational Unified Process in facing the main challenges of product configuration systems development. *Journal of Systems and Software*, 170. <https://doi.org/10.1016/j.jss.2020.110732>
- Sun, K., Yu, D., Chen, J., Yu, D., Choi, Y., & Cardie, C. (2019). DREAM: A Challenge Data Set and Models for Dialogue-Based Reading Comprehension. *Transactions of the Association for Computational Linguistics*, 7, 217–231. [https://doi.org/10.1162/tacl\\_a\\_00264](https://doi.org/10.1162/tacl_a_00264)
- Verma, P., & Verma, A. (2020). A Review on Text Summarization Techniques. *Journal of Scientific Research*, 64(01), 251–257. <https://doi.org/10.37398/jsr.2020.640148>

Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., Affandy, A., & Setiadi, D. R. I. M. (2022). Review of automatic text summarization techniques & methods. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 34, Issue 4, pp. 1029–1046). King Saud bin Abdulaziz University. <https://doi.org/10.1016/j.jksuci.2020.05.006>