

DAFTAR PUSTAKA

- Azmi, Z., & Dahria, M. (2013). Decision Tree Berbasis Algoritma Untuk Pengambilan Keputusan. *Saintikom*, 12, 157–164. [http://demo.pohonkeputusan.com/files/Decision Tree Berbasis Algoritma Untuk Pengambilan Keputusan.pdf?i=1](http://demo.pohonkeputusan.com/files/Decision%20Tree%20Berbasis%20Algoritma%20Untuk%20Pengambilan%20Keputusan.pdf?i=1)
- Bahri, S., & Lubis, A. (2020). Metode Klasifikasi Decision Tree Untuk Memprediksi Juara English Premier League. *Jurnal Sintaksis*, 2(1), 63–70.
- Eris, Viny Christanti M, J. P. (2017). *Penerapan Algoritma TextRank Untuk Automatic Summarization Pada Dokumen Berbahasa Indonesia*. 1(1), 71–78.
- Firdaus, M., & C, T. A. (2019). *Implementasi algoritma decision tree untuk klasifikasi pola serangan pada log file*. <http://repository.unmuhjember.ac.id/7150/1/JURNAL.pdf>
- Khontoro, C., Andjarwirawan, J., & Yulia. (2021). Penerapan Algoritma *TextRank* dan Dice Similarity untuk Verifikasi Berita Hoax. *Infra*, 9(1), 1–5.
- Nasrullah, A. H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2), 45–51. <https://doi.org/10.35329/jiik.v7i2.203>
- Riyadi, S., & Samiati, I. (2021). Pengenalan algoritma dalam model peringkasan teks untuk mempercepat pekerjaan akademik. *Jgmm*, 1(2), 19–26.
- Robianto, Sitorus, S., & Ristian, U. (2021). *Penerapan Metode Decision Tree Untuk Mengklasifikasikan Mutu Buah Jeruk Berdasarkan Fitur Warna Dan Ukuran*. 09(01), 76–86.
- Rojas-Simón, J., Ledeneva, Y., & García-Hernández, R. A. (2021). Evaluation of text summaries without human references based on the linear optimization of content metrics using a genetic algorithm. *Expert Systems with Applications*, 167(August 2020). <https://doi.org/10.1016/j.eswa.2020.113827>
- Yuliska, Y., & Syaliman, K. U. (2020). Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Teks Berbahasa Indonesia. *IT Journal Research and Development*, 5(1), 19–31. [https://doi.org/10.25299/itjrd.2020.vol5\(1\).4688](https://doi.org/10.25299/itjrd.2020.vol5(1).4688)
- Zamzam, M. A. (2020). Sistem Automatic Text Summarization Menggunakan Algoritma *TextRank*. *Matics*, 12(2), 111–116. <https://doi.org/10.18860/mat.v12i2.8372>

LAMPIRAN

A. Lampiran 1

1. Kode Program Metode *TextRank*

```
from flask import Flask, request, render_template
from wtforms import Form, TextAreaField, validators
from gensim.summarization.summarizer import summarize
from rouge import Rouge
import nltk

nltk.download('punkt')

app = Flask(__name__)

class InputForm(Form):
    text = TextAreaField(
        label='Input Text', default='',
        validators=[validators.InputRequired(),
validators.Length(100)]
    )

@app.route('/', methods=['GET', 'POST'])
def main():
    form = InputForm(request.form)
    if request.method == 'GET':
        return render_template('dashboard_input.html',
form=form)
    if request.method == 'POST' and form.validate():
        text = form.text.data
        ratio = 0.3
        sentences = nltk.sent_tokenize(text)
        if len(sentences) > 1:
            summary = TextRank_summarize(text, ratio=ratio)
        else:
            summary = text
        rouge_scores = evaluate_rouge(text, summary)
        if rouge_scores is not None:
            return render_template('dashboard_result.html',
text=text, summary=summary, rouge_scores=rouge_scores)
        else:
            return render_template('dashboard_result.html',
text=text, summary=summary)
    else:
        return render_template('dashboard_input.html',
form=form)

def TextRank_summarize(text, ratio=0.3):
    summarized_text = summarize(text, ratio=ratio)
    return summarized_text
```

```

def evaluate_rouge(reference, summary):
    rouge = Rouge()
    if summary:
        scores = rouge.get_scores(summary, reference)
        return scores[0]
    else:
        return None

if __name__ == "__main__":
    app.run(debug=True)

```

2. Penjelasan Cosine Similarity pada artikel ke 1

Tabel 4. 16 Similarity 1

Similarity (1, 2)	0.250
Similarity (1, 3)	0.381
Similarity (1, 4)	0.157
Similarity (1, 5)	0.126
Similarity (1, 6)	0.267
Similarity (1, 7)	0.267

Tabel 4. 17 Similarity 2

Similarity (2, 3)	0.546
Similarity (2, 4)	0.305
Similarity (2, 5)	0.245
Similarity (2, 6)	0.516
Similarity (2, 7)	0.516

Tabel 4. 18 Similarity 3

Similarity (3, 4)	0.452
Similarity (3, 5)	0.363
Similarity (3, 6)	0.763
Similarity (3, 7)	0.763

Tabel 4. 19 Similarity 4

Similarity (4, 5)	0.724
Similarity (4, 6)	0.243
Similarity (4, 7)	0.243

Tabel 4. 20 Similarity 5

Similarity (5, 6)	0.828
Similarity (5, 7)	0.828
Similarity (6, 7)	0.828

Keterangan:

Metode *cosine similarity* digunakan untuk mengukur sejauh mana dua vektor mendekati kesamaan atau keterkaitannya. Sehingga dapat dilihat pada tabel 4.3 sampai dengan tabel 4.7, dimana pada table 4.3 memiliki 7 data yang diberi nomor 1.2 hingga 1.7. Similarity antara dua dokumen dihitung dengan menggunakan *cosine similarity* dan nilai-nilai tersebut,

misalnya, dalam tabel 4.3, *Similarity* (1, 2) memiliki nilai 0.677. Ini berarti bahwa data nomor 1.2 memiliki tingkat kesamaan sebesar 0.677 menggunakan metode cosine similarity. Demikian pula, nilai-nilai *Similarity* yang diberikan dalam tabel-tabel lain menunjukkan tingkat kesamaan antara pasangan dokumen yang berbeda. Metode *cosine similarity* dihitung berdasarkan representasi vektor dari dokumen yang dianalisis. Vektor ini bisa berupa representasi *term frequency-inverse document frequency* (TF-IDF) dari setiap kata pada data tabel atau menggunakan metode representasi vektor lainnya.

Contoh kalimat 1 dan 2 pada artikel 3

1. "Seiring berkembangnya zaman"
2. "Tidak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis"

seiring	berkembangnya	zaman	tidak	terasa	kita	telah	hidup	di	Yang
penuh	dengan	kecanggihan	dan	semuanya	Serba	praktis			

1. "Seiring berkembangnya zaman"

1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0			

2. "Tidak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis"

0	0	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1			

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}},$$

$$(1*0) + (1*0) + (1*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1) + (0*1)$$

$$\sqrt{1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2} + \sqrt{0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2}$$

$$= \frac{1}{\sqrt{3} * \sqrt{15}}$$

$$= \frac{1}{4}$$

$$= 0,25$$

Jadi, nilai kecocokan antarkalimat di atas 16%

3. Penjelasan Rouge-1, Rouge-2, dan Rouge-L

1) ROUGE-1 dengan Unigram (perkata)

- *Precision*

Precision (P) = (Jumlah token dalam LCS) / (Jumlah token dalam hasil summary)

$$\begin{aligned} P &= \frac{LCS}{n} \\ &= \frac{31}{31} \\ &= 1.0 \end{aligned}$$

- *Recall*

Keterangan:

Recall = (Jumlah token dalam LCS) / (Jumlah token dalam kalimat asli)

$$\begin{aligned} R &= \frac{LCS}{m} \\ &= \frac{31}{105} \\ &= 0.295 \end{aligned}$$

- *F1 Score*

Keterangan:

F1 Score = $2 * (Precision * Recall) / (Precision + Recall)$

$$\begin{aligned} F_{lcs} &= \frac{2 * (P * R)}{(P+R)} \\ &= \frac{2 * (1.0 * 0.295)}{(1.0 + 0.295)} \\ &= \frac{2 * 0.295}{1.295} \\ &= \frac{0.59}{1.295} \\ &= 0.456 \end{aligned}$$

2) ROUGE-2 dengan *Bigram* (pasangan kata)

- *Precision*

Precision (P) = (Jumlah token dalam *LCS*) / (Jumlah token dalam hasil summary)

$$\begin{aligned} P &= \frac{LCS}{n} \\ &= \frac{31}{31} \\ &= 1.0 \end{aligned}$$

- *Recall*

Recall = (Jumlah token dalam *LCS*) / (Jumlah token dalam kalimat asli)

$$\begin{aligned} R &= \frac{LCS}{m} \\ &= \frac{31}{130} \\ &= 0.238 \end{aligned}$$

- *F1 Score*

F1 Score = $2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$

$$\begin{aligned} F_{lcs} &= \frac{2 * (P * R)}{(P+R)} \\ &= \frac{2 * (1.0 * 0.238)}{(1.0 + 0.238)} \\ &= \frac{2 * 0.238}{1.295} \\ &= \frac{0.476}{1.238} \\ &= 0.384 \end{aligned}$$

3) ROUGE-L

- *Precision*

Precision = (Jumlah unigram yang cocok / (jumlah unigram dalam teks ringkasan yang di hasilkan)

$$\begin{aligned} P &= \frac{LCS}{n} \\ &= \frac{31}{31} \\ &= 1.0 \end{aligned}$$

- *Recall*

Recall = (total unigram cocok) / total unigram dalam teks ringkasan referensi)

bigram hipotesis

$$\begin{aligned} R &= \frac{LCS}{m} \\ &= \frac{31}{105} \\ &= 0.295 \end{aligned}$$

- *F1 Score*

F1 Score = $2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$

$$\begin{aligned} F_{lcs} &= \frac{2 * (P * R)}{(P+R)} \\ &= \frac{2 * (1.0 * 0.295)}{(1.0 + 0.295)} \\ &= \frac{2 * 0.295}{1.295} \\ &= \frac{0.59}{1.295} \\ &= 0.455 \end{aligned}$$

LEMBAR PERBAIKAN SKRIPSI

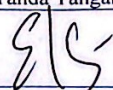
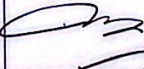


“Implementasi *TextRank* Pada Peringkasan Teks Otomatis Dokumen Berbahasa Indonesia”

OLEH:


A. Fiqram Abdillah Astri
D42116507

Skripsi ini telah dipertahankan pada Ujian Akhir Sarjana tanggal 17 Juli 2023.
Telah dilakukan perbaikan penulisan dan isi skripsi berdasarkan usulan dari penguji dan pembimbing skripsi.

Persetujuan perbaikan oleh tim penguji:

	Nama	Tanda Tangan
Ketua	Elly Warni, S.T., M.T.	
Sekretaris	Mukarramah Yusuf, B.Sc., M.Sc., Ph.D.	
Anggota	Dr. Eng. Ady Wahyudi Paundu, ST., M.T.	
	Muhammad Alief Fahdal Imran Oemar., ST., M.Sc.	

Persetujuan Perbaikan oleh pembimbing:

Pembimbing	Nama	Tanda Tangan
I	Elly Warni, S.T., M.T.	
II	Mukarramah Yusuf, B.Sc., M.Sc., Ph.D.	