

Sehingga dari hasil *summary* yang dikeluarkan oleh hasil proses menggunakan rasio 0.3 sebagai batasan prediksi dengan rujukan kalimat asli sebanyak 120 kata 5 kalimat, sehingga menghasilkan hasil kesimpulan/*summary* yaitu 21 kata dan 1 kalimat. Setelah itu dilakukan proses dilakukan evaluasi menggunakan data hasil *summary*. Hasil evaluasi dapat berupa metrik seperti ROUGE-1, ROUGE-2, dan ROUGE-L untuk mengevaluasi keakuratan hasil *summary* yang dihasilkan oleh sistem sebagai berikut:

Tabel 4. 10 Hasil Perhitungan *Rouge TextRank* Artikel Kelima

ROUGE	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
ROUGE-1	1.0	0.215909091	0.355140184
ROUGE-2	1.0	0.172413793	0.294117645
ROUGE-L	1.0	0.215909091	0.355140184

Berdasarkan hasil ROUGE-1, ROUGE-2, dan ROUGE-L memberikan gambaran tentang sejauh mana ringkasan mencerminkan teks asli. Pada nilai *Precision*, ketiga ROUGE mendapatkan nilai 100%, menunjukkan bahwa ringkasan memiliki kesesuaian yang tinggi dengan teks asli dalam hal kata-kata individu.

Namun, pada nilai *Recall*, ROUGE-1 dan ROUGE-L memperoleh nilai 21.59%, sedangkan ROUGE-2 mendapatkan nilai 17.24%. Hal ini menunjukkan bahwa tidak semua informasi penting dalam teks asli tercakup dalam ringkasan. Sebagian informasi dalam teks asli mungkin tidak diikutsertakan dalam ringkasan tersebut.

Dari segi *F1 Score*, ROUGE-1 dan ROUGE-L memiliki nilai yang sama yaitu 35.51%, sedangkan ROUGE-2 memiliki nilai 29.41%. Nilai *F1 Score* yang relatif rendah menunjukkan bahwa ringkasan belum sepenuhnya mencerminkan teks asli secara keseluruhan dengan baik.

## 4.2 Pembahasan

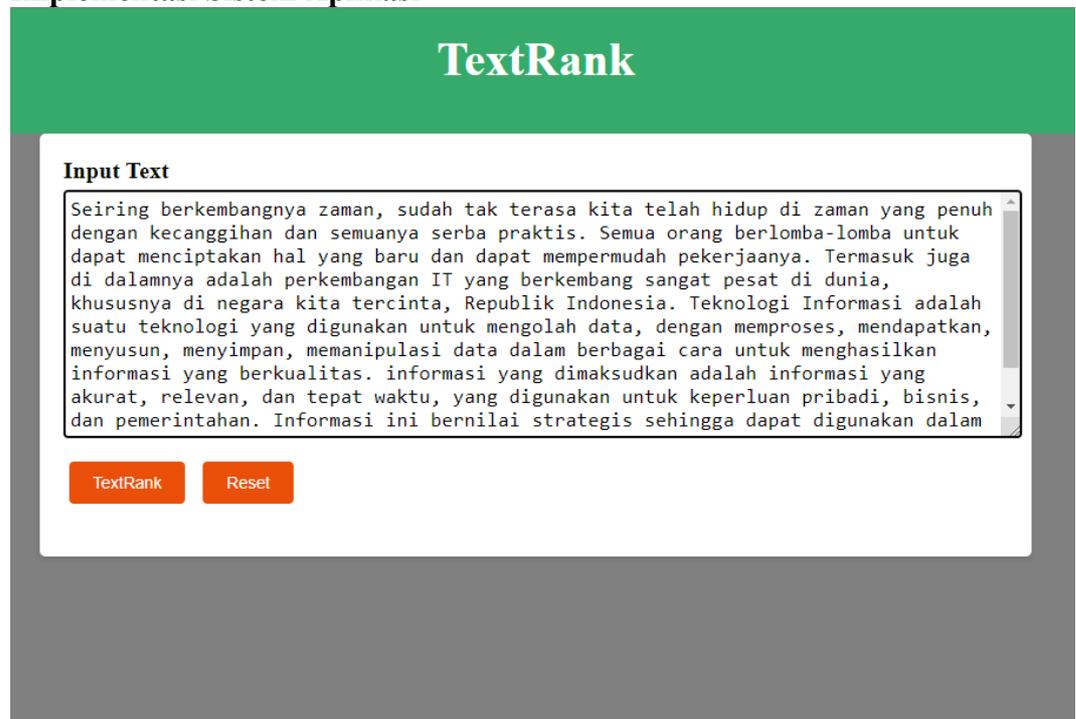
### A. Skenario Uji Coba

Langkah pertama dalam skenario uji coba adalah dengan mengumpulkan kumpulan data teks artikel teknologi informasi bahasa Indonesia pada internet berupa *Blogger*, *WordPress* dan berita *online* yang membahas tentang perkembangan teknologi informasi, dilakukan pengujian kedalam sistem aplikasi peringkas otomatis bahasa Indonesia menggunakan metode *TextRank*, sehingga dapat menghasilkan kesimpulan hasil ringkasan yang lebih *simple* dan efektif sehingga dapat memudahkan dalam penyimpulan yang lebih efisien dan optimal.

Pengujian ini akan dilakukan untuk mengetahui apakah sistem aplikasi sudah berjalan sesuai dengan metode *TextRank* yang telah dibuat. Pengujian dilakukan secara langsung dan teliti, agar proses pengujian berjalan dengan sesuai harapan. Adapun skenario sebelum melakukan pengujian terhadap sistem sebagai berikut:

1. Menjalankan aplikasi *Visual Studio Code* dan melakukan *running script* aplikasi peringkas otomatis bahasa Indonesia menggunakan metode *TextRank* yang telah dibuat menggunakan bahasa pemrograman *python* dengan cara mengklik *run python file* pada aplikasi tersebut.
2. Menjalankan sistem aplikasi melalui salah satu aplikasi *web browser* yaitu *google chrome*.
3. Setelah sistem terbuka selanjutnya masukan *text* artikel bahasa Indonesia kedalam *form input text* yang berfungsi sebagai *form* masukan yang akan di proses dalam pengujian pada sistem tersebut.

## B. Implementasi Sistem Aplikasi



Gambar 4. 6 Skema Tampilan Aplikasi

1. Form input *text* berfungsi untuk menambahkan *text* artikel bahasa Indonesia kedalam sistem aplikasi.
2. Button *TextRank* berfungsi sebagai tombol proses sistem aplikasi dari *text* asli menjadi hasil ringkasan/*summary* kesimpulan *text* artikel bahasa Indonesia sesuai dengan kaidah metode *TextRank*.

# TextRank

## Original Text

Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas. Informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.

## Hasil Summary

Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas.

## ROUGE Scores

ROUGE	Precision	Recall	F1 Score
ROUGE-1	1.0	0.2911392405063291	0.45098038866397544
ROUGE-2	1.0	0.23529411764705882	0.38095237786848074
ROUGE-L	1.0	0.2911392405063291	0.45098038866397544

Kembali

3. Tampilan setelah teks disubmit menampilkan hasil ringkasan teks dengan tampilan original teks dan *summary* teksnya dan menampilkan hasil nilai *Rouge-1*, *Rouge-2*, dan *Rouge-L*.

## C. Pembahasan *Source Code* Program

### 1. Impor *library* yang diperlukan:

```
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')
```

Keterangan:

- a) “**flask**”: Digunakan untuk membuat aplikasi web menggunakan *framework Flask*.
- b) “**Request**”: Digunakan untuk mengakses data yang dikirimkan oleh klien ke server.
- c) “**Form**” dan “**TextAreaField**” dari “**wtforms**”: Digunakan untuk membuat formulir dengan Flask-WTF.
- d) “**render\_template**”: Digunakan untuk merender template *HTML*.
- e) “**wtforms**”: Digunakan untuk membuat *form HTML* dan melakukan validasi input.
- f) “**summarize**” dari “**gensim.summarization.summarizer**”: Digunakan untuk menjalankan algoritma *TextRank* dan menghasilkan ringkasan teks.
- g) “**Rouge**”: Digunakan untuk menghitung skor ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) untuk mengevaluasi ringkasan teks.
- h) “**nltk**”: Digunakan untuk pemrosesan teks, termasuk tokenisasi kalimat.

### 2. Inialisasi Aplikasi Flask

```
app = Flask(__name__)
```

Keterangan:

- a) “**app = Flask(\_\_name\_\_)**”: Digunakan untuk membuat objek aplikasi *Flask* dengan menggunakan “**Flask(\_\_name\_\_)**”.

### 3. Definisi *Form Input*

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

Keterangan:

- a) “**InputForm**” adalah kelas turunan dari Form yang didefinisikan menggunakan Flask-WTF. Di dalamnya terdapat atribut “**text**” yang merupakan “**TextAreaField**” untuk memasukkan teks. Validasi juga ditambahkan untuk memastikan bahwa teks diinputkan dan memiliki panjang minimal 100 karakter.
4. Mendefinisikan rute utama dan logika aplikasi:

```
@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)
```

Keterangan:

- a) Metode GET: Ketika klien melakukan permintaan GET ke rute utama, halaman “**dashboard\_input.html**” dirender dengan formulir yang telah dibuat sebelumnya.
- b) Metode POST: Ketika klien melakukan permintaan POST ke rute utama dan formulir valid, teks yang diinputkan oleh pengguna akan diambil. Variabel “**ratio**” diberikan nilai 0.3 sebagai rasio panjang ringkasan terhadap panjang teks asli. Teks tersebut akan dipecah menjadi kalimat-kalimat menggunakan “**nltk.sent\_tokenize**”. Jika terdapat lebih dari satu kalimat, fungsi “**TextRank\_summarize**” akan digunakan untuk melakukan ringkasan teks menggunakan metode *TextRank*. Jika hanya terdiri dari satu kalimat, teks asli akan dikembalikan. Selanjutnya, skor ROUGE akan dihitung menggunakan fungsi “**evaluate\_rouge**”. Jika skor ROUGE tersedia,

halaman “**dashboard\_result.html**” akan dirender dengan teks, ringkasan, dan skor ROUGE. Jika skor ROUGE tidak tersedia, halaman “**dashboard\_result.html**” akan dirender dengan teks dan ringkasan saja. Jika formulir tidak valid, halaman “**dashboard\_input.html**” akan dirender kembali dengan formulir yang telah diisi sebelumnya.

5. Fungsi “**TextRank\_summarize (text, ratio=0.3)**”:

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

Keterangan:

- a) “**TextRank\_summarize**” adalah fungsi yang digunakan untuk melakukan ringkasan teks menggunakan metode *TextRank*. Fungsi ini mengambil teks sebagai input dan menghasilkan ringkasan teks dengan menggunakan fungsi *summarize* dari modul *gensim.summarization.summarizer*. Argumen opsional *ratio* digunakan untuk menentukan rasio panjang ringkasan terhadap panjang teks asli. Nilai *default* rasio adalah menggunakan rasio 0.3.

6. Fungsi “**evaluate\_rouge()**”:

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

Keterangan:

- b) Menerima teks referensi (asli) dan teks ringkasan sebagai argumen.  
c) Membuat instance dari *Rouge* dan menggunakan “**get\_scores()**” untuk menghitung skor ROUGE antara teks referensi dan teks ringkasan.  
d) Jika teks ringkasan ada, fungsi ini mengembalikan skor ROUGE. Jika tidak ada, maka mengembalikan “**None**”.

7. Bagian utama:

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

Bagian ini menjalankan aplikasi *flask* saat file ini dieksekusi langsung, dengan mode *debug* aktif. Sehingga pada bagian yang melakukan proses *TextRank* adalah:

```
ratio = 0.3
sentences = nltk.sent_tokenize(text)
if len(sentences) > 1:
    summary = summarize(text, ratio=ratio)
```

```
else:  
    summary = text
```

Jadi, proses *TextRank* terjadi pada bagian yang memanggil fungsi “*summarize()*” dan memeriksa jumlah kalimat dalam teks.

#### D. Pembahasan Analisis Metode *TextRank*

Pada tahap ini dilakukan analisis data text masukan artikel bahasa Indonesia yang telah di proses oleh sistem dengan mengimplementasikan metode *TextRank*. Adapun pembahasan dilakukan dengan melihat hasil perhitungan ROUGE-1, ROUGE-2 dan ROUGE-L pada metode *TextRank* menggunakan rasio 0.3 pada pengambilan keputusan prediksi, sehingga memperoleh hasil kesimpulan/summary yang lebih efisien dan efektif serta informatif.

##### c) Cara Kerja Kalimat Di Ringkas

###### 1) Input Teks

Adapun teks yang akan di input:

“Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas. informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.”

###### 2) Validasi minimal karakter

Teks yang di validasi untuk menghitung minimal karakter:  
“Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas. informasi yang dimaksudkan adalah informasi yang akurat,

relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.”

Total Karakter: Kalimat: (5), Kata:(106) dan Karakter(823)

### 3) Preprocessing

- Tokenisasi

“Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas. informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.”

Setelah di tokenisasi:

“Seiring”, “berkembangnya”, “zaman”, “,””, “sudah”, “tak”, “terasa”, “kita”, “telah”, “hidup”, “di”, “zaman”, “yang”, “penuh”, “dengan”, “kecanggihan”, “dan”, “semuanya”, “serba”, “praktis”, “.””, “Semua”, “orang”, “berlomba”, “-“, “lomba”, “untuk”, “dapat”, “menciptakan”, “hal”, “yang”, “baru”, “dan”, “dapat”, “mempermudah”, “pekerjaannya”, “.””, “Termasuk”, “juga”, “di”, “dalamnya”, “adalah”, “perkembangan”, “IT”, “yang”, “berkembang”, “sangat”, “pesat”, “di”, “dunia”, “,””, “khususnya”, “di”, “negara”, “kita”, “tercinta”, “,””, “Republik”, “Indonesia”, “.””, “Teknologi”, “Informasi”, “adalah”, “suatu”, “teknologi”, “yang”, “digunakan”, “untuk”, “mengolah data”, “,””, “dengan”, “memproses”, “,””, “mendapatkan”, “,””, “menyusun”, “,””, “menyimpan”, “,””, “manipulasi”, “data”, “dalam”, “berbagai”, “cara”, “untuk”, “menghasilkan”, “informasi”, “yang”, “berkualitas”, “.””, “informasi” “yang”, “dimaksudkan”, “adalah”, “informasi”, “yang”, “akurat”, “,””, “relevan”, “,””, “dan”, “tepat”, “waktu”, “,””, “yang”, “digunakan”, “untuk”, “keperluan”, “pribadi”, “,””, “bisnis”, “,””, “dan”, “pemerintahan”, “.””, “Informasi”,

“ini”, “bernilai”, “strategis”, “sehingga”, “dapat”, “digunakan”, “dalam”, “pengambilan”, “keputusan”, “.”,

#### 4) Ratio

Teks yang akan di ratio 0,3:

“Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas. informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.”

Setelah di ratio:

"Seiring berkembangnya zaman, hidup kita penuh dengan kecanggihan dan praktis. Orang berlomba menciptakan hal baru untuk mempermudah pekerjaan. Termasuk perkembangan IT yang pesat di dunia, terutama di Indonesia. Teknologi Informasi mengolah data, memproses, menyimpan, memanipulasi data untuk informasi berkualitas. Informasi akurat, relevan, tepat waktu digunakan dalam keperluan pribadi, bisnis, dan pemerintahan. Informasi bernilai strategis untuk pengambilan keputusan."

Teks di atas telah dipendekkan dengan mempertahankan informasi penting dan mengurangi beberapa kalimat.

#### 5) *TextRank*

- Reprerentasi vektor

Kalimat yang akan di reprerentasi vektor:

“Seiring berkembangnya zaman, sudah tak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis. Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya. Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia. Teknologi Informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk

menghasilkan informasi yang berkualitas. informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan. Informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan.”

Setelah di representasi vektor:

1. "seiring berkembangnya zaman"
2. "tidak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis"
3. "semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya"
4. "termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia"
5. "teknologi informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas"
6. "informasi yang dimaksudkan adalah informasi yang akurat, relevan, dan tepat waktu, yang digunakan untuk keperluan pribadi, bisnis, dan pemerintahan"
7. "informasi ini bernilai strategis sehingga dapat digunakan dalam pengambilan keputusan"

Dalam representasi vektor *TextRank*, setiap unit teks (kalimat atau frasa) diwakili oleh simpul dalam graf.

- Similarity matrix (cosine Similarity)

Kalimat yang akan di process:

1. "Seiring berkembangnya zaman"
2. "Tidak terasa kita telah hidup di zaman yang penuh dengan kecanggihan dan semuanya serba praktis"
3. "Semua orang berlomba-lomba untuk dapat menciptakan hal yang baru dan dapat mempermudah pekerjaannya"
4. "Termasuk juga di dalamnya adalah perkembangan IT yang berkembang sangat pesat di dunia, khususnya di negara kita tercinta, Republik Indonesia"
5. "Teknologi informasi adalah suatu teknologi yang digunakan untuk mengolah data, dengan memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas"