



## Penerapan Natural Language Processing untuk Klasifikasi Bidang Minat berdasarkan Judul Tugas Akhir

Ahmad Jimly Hanif<sup>1</sup>, Mifta Nur Farid<sup>2✉</sup>, Barokatun Hasanah<sup>3</sup>

<sup>1,2,3</sup>Institut Teknologi Kalimantan, Indonesia

[miftanurfarid@lecturer.itk.ac.id](mailto:miftanurfarid@lecturer.itk.ac.id)

### Abstrak

Teknik Elektro ITK memiliki 5 bidang minat sehingga terdapat berbagai pilihan judul tugas akhir yang dapat diambil oleh mahasiswa tingkat akhir. Beberapa kondisi yang dapat mempengaruhi mahasiswa dalam memilih judul tugas akhir, salah satunya yaitu berdasarkan bidang minat yang diambil oleh mahasiswa tersebut. Pelaksanaan penelitian ini dilaksanakan dengan menggunakan Python software versi 3.9.13, Spyder versi 5.2.2, dan library NLTK. Output pada perancangan sistem ini adalah hasil rekomendasi bidang minat dan dosen pembimbing berdasarkan dari judul tugas akhir. Setelah dilakukan simulasi, diperoleh hasil pengambilan data dari Dataset berisi 100 judul yang digunakan sebagai parameter data testing berdasarkan 5 bidang minat teknik elektro itk. Adapun hasil dari proses pengambilan data yaitu judul tugas akhir bidang minat sistem tenaga didapatkan hasil rekomendasi berupa bidang minat sistem tenaga, teknik komputer, telekomunikasi, sistem pengaturan, dan elektronika. Pada bidang minat sistem pengaturan, didapatkan hasil rekomendasi berupa bidang minat sistem pengaturan dan sistem tenaga. Pada bidang minat elektronika, didapatkan hasil rekomendasi berupa bidang minat elektronika, sistem tenaga, teknik komputer, telekomunikasi, dan sistem pengaturan. Pada hasil bidang minat teknik komputer, didapatkan hasil rekomendasi berupa bidang minat teknik komputer, elektronika, sistem tenaga, telekomunikasi, dan sistem pengaturan. Dan Pada hasil bidang minat telekomunikasi, didapatkan hasil rekomendasi berupa bidang telekomunikasi, sistem tenaga, teknik komputer, sistem pengaturan, dan elektronika. Adapun tingkat akurasi pada aplikasi ini sebesar 86, 8%.

**Kata kunci:** NLP, Sentiment Analysis, Text Mining, Tugas Akhir, Elektro.

*JSISFOTEK is licensed under a Creative Commons 4.0 International License.*



### 1. Pendahuluan

Menentukan topik tugas akhir merupakan masalah awal yang dialami oleh mahasiswa tingkat akhir. Mahasiswa biasanya merasa kesulitan, ragu-ragu dan bahkan tidak tahu bagaimana memutuskan suatu topik sehingga mereka membutuhkan lebih banyak waktu untuk memulai tugas akhirnya. Salah satu fasilitas yang dapat digunakan oleh mahasiswa untuk mencari judul tugas akhir yang telah tersedia adalah melalui web perpustakaan kampus, Namun informasi yang ditampilkan hanya berupa judul tugas akhir, penulis, program studi, dan lokasi kampus mahasiswa tugas akhir tersebut. Akan tetapi sistem web perpustakaan kampus belum dapat membatasi judul tugas akhir berdasarkan topik bidang minat dan tahun dari judul tugas akhir tersebut [12].

Pada tahap awal penulisan tugas akhir, mahasiswa pada umumnya sering terkena kendala untuk menentukan judul tugas akhir dan menentukan dosen pembimbing dan pada umumnya mahasiswa masih belum mengetahui langkah dalam penentuan tema atau judul tugas akhir yang akan diajukan sehingga dapat memperlambat proses pengajuan judul tugas akhir tersebut. Jadi langkah pertama dalam menentukan judul tugas akhir adalah menentukan topik mengenai penulisan tugas akhir terdahulu seperti mahasiswa harus menentukan tentang tugas akhir apa yang mau diteliti. Salah satu cara untuk membantu menemukan solusi penentuan judul tugas akhir dan dosen pembimbing berdasarkan bidang minat adalah dengan memanfaatkan sistem teknologi informasi. Oleh karena itu, penulis membuat sistem rekomendasi yang dapat membantu mahasiswa dalam mengambil judul tugas akhir dan menentukan dosen pembimbing yang tepat berdasarkan bidang minat yang terdapat dari judul yang diajukan oleh mahasiswa tersebut [6].

Pada umumnya, Terdapat banyak metode yang dapat membantu mahasiswa untuk mengetahui termasuk kedalam jenis bidang mana judul tugas akhir yang telah diajukan oleh mahasiswa tersebut, Namun sebagian besar juga masih banyak terdapat mahasiswa yang belum mengetahui bidang minat dari judul tugas akhir yang telah diajukan sehingga dilakukan penelitian ini dengan memanfaatkan data judul tugas akhir yang digunakan sebagai data informasi untuk menganalisis data yang disimpan pada sistem yang bertujuan mencari informasi yang bermanfaat untuk menyelesaikan permasalahan kesulitan dalam menentukan bidang minat dari judul tugas akhir yang telah diajukan dan juga menentukan dosen pembimbing yang tepat sesuai dengan bidang minat yang telah ditentukan

dari judul tugas akhir yang telah diajukan oleh mahasiswa tersebut. Untuk proses pengambilan data dapat menggunakan software sistem berbentuk basis data yang dapat dilakukan dengan menggunakan Natural Language Processing (NLP) dengan menggunakan metode bahasa alami [2].

Berdasarkan ringkasan permasalahan di atas, terdapat ide untuk membuat suatu sistem software yang dapat memberikan output atau keluaran mengenai penentuan bidang minat dan dosen pembimbing yang sesuai dengan judul tugas akhir yang diajukan oleh mahasiswa. Penelitian ini menggunakan software Python yang telah ter-install dengan library tools pendukung yang bernama NLTK (Natural Language Toolkit) yang bertujuan untuk membaca pemrosesan metode bahasa alami yang akan menghasilkan output keluaran berupa bidang minat dari judul tugas akhir serta dosen pembimbing yang tepat sesuai dengan tema tugas akhir yang diajukan oleh mahasiswa.

## 2. Metode

Metode penelitian yang digunakan pada penelitian ini terdiri dari beberapa langkah. Langkah pertama adalah pembuatan *dataset* yang terdiri dari judul tugas akhir, bidang minat dari judul tugas akhir tersebut, dan nama dosen pembimbing dari masing-masing bidang minat. Langkah kedua adalah identifikasi kebutuhan *hardware* maupun software yang digunakan untuk merancang sistem. Langkah ketiga adalah perancangan sistem. Langkah keempat adalah pengujian terhadap sistem yang telah dibuat.

### 2.1. Pembuatan Dataset

Dataset dibuat dalam bentuk dua *file csv*. *File csv* pertama berisikan informasi tentang judul tugas akhir pada kolom pertama dan bidang minatnya pada kolom kedua sebagaimana yang diilustrasikan oleh Gambar 1(a). *File csv* kedua berisikan informasi tentang nama-nama Dosen Program Studi Teknik Elektro - ITK pada kolom pertama dan bidang keahliannya pada kolom kedua sebagaimana yang diilustrasikan oleh Gambar 1(b). Judul tugas akhir didapatkan baik dari *repository* Institut Teknologi Kalimantan (ITK) maupun kampus di selain ITK. Jumlah judul yang didapatkan sebanyak 200 judul.

	Judul Tugas Akhir	Bidang Minat
1	Analisis Aspek Ekonomi dari Keandalan Sistem Distribusi 20 kV pada PT. PLN (Persero) Area Balikpapan Penyulang I1 dan I10	Sistem Tenaga
2	Analisis Perbandingan Penggunaan Transformator Sisi dan Uprating Transformator dalam Menanggulangi Drop Tegangan pada Gardu Distribusi GHI 1011 Penyulang J5	Sistem Tenaga
3	Studi Perencanaan Pembangkit Listrik Tenaga Sampah pada TPA Sambutan Kota Samarinda	Sistem Tenaga
4	Optimalisasi Penempatan Capacitor Bank untuk Memperbaiki Kualitas Daya Listrik pada Sistem Kelistrikan PDAM Tirta Kencana Kota Samarinda	Sistem Tenaga
5	Multi-Objective Optimal Power Flow pada Sistem Mahakam 150 kV Menggunakan Chaotic Bat Algorithm	Sistem Tenaga
6	RANCANG BANGUN ALAT PEREKAM KECEPATAN DAN KEMIRINGAN JALAN KENDARAAN BERMOTOR BERBASIS INTERNET OF THINGS (IOT)	Elektronika
7	Rancang Bangun Robot Pemancing Ikan menggunakan Sensor ACXL 345 dengan Monitoring Android	Elektronika
8	Rancang Bangun Alat Wiper Pintar Pada Mobil Berbasis Raspberry Pi 3	Elektronika
9	Rancang Bangun Pengaturan Suhu Air pada Valve dengan Controller Ni Myrio-1900	Elektronika
10	Perancangan Alat Pengendali Suhu dan Kelembapan Pada Budidaya Cacing Lumbricus Rubellus dengan Menggunakan Metode Anfis (Adaptive Neuro Fuzzy Inference System)	Elektronika
11	Analisis Perbandingan dan Implementasi Boost Converter dan Quadratic Boost Converter	Sistem Pengaturan
12	Perancangan Fuzzy-PID Controller untuk Kestabilan Hover pada Hexacopter	Sistem Pengaturan
13	Analisis Penerapan Kontrol Proportional Integral Derivative (PID) pada Pengendalian Kecepatan Motor Sinkron Magnet Permanen Dengan Metode Indirect Field Oriented Control	Sistem Pengaturan
14	Analisis Penerapan Kontroler PID pada Pengendalian Kecepatan Motor Induksi dengan Metode Direct Field Oriented Control (DFOC)	Sistem Pengaturan
15	Penerapan Kendali PID (Proporsional Integral Derivative) pada cuk Converter	Sistem Pengaturan
16	Sistem Informasi Suhu, Cucia, dan Polusi Udara Menggunakan Metode Logika Fuzzy Sugeno	Teknik Komputer
17	Perancangan Sistem Deteksi Kantuk Menggunakan Metode Artificial Intelligent (AI) Berbasis Neurosky Mindwave dan SMS Gateway	Teknik Komputer
18	Perancangan Sistem Robot Line Follower untuk Pembelajaran di Taman Rekreasi Sengkalang dengan Identifikasi RFID disertai Wireless Monitoring (Software)	Teknik Komputer
19	Sistem Parkir Otomatis Prabayar Menggunakan Near Field Communication (NFC) dan Identifikasi Plat Nomor Kendaraan (Software)	Teknik Komputer
20	Implementasi Near Field Communication (NFC) dan Light Emitting Diode (Led) Touchscreen untuk Mengontrol Absensi Karyawan Berbasis Web	Teknik Komputer
21	Rancang Bangun Antena Mikrostrip Menggunakan Material Dielektrik Artifisial pada Frekuensi 2,4 GHz	Telekomunikasi
22	ALOKASI RESOURCE BERBASIS ALGORITMA GREEDY DAN ROUND ROBIN UNTUK PENINGKATAN PERFORMANSI PADA SISTEM 4G LTE	Telekomunikasi
23	Perancangan Dan Implementasi Antena Open-Ended Rectangular Waveguide Untuk Uji Non-Destructif Menggunakan Vector Network Analyzer	Telekomunikasi
24	Analisis Kinerja MIMO-OFDM Dengan Menggunakan Mode Spatial	Telekomunikasi
25	Perancangan Antena 3d Serta Pengaruh Jumlah Seed (Patch) Terhadap Performansi Antena Pada Frekuensi Kerja 2.4 Ghz	Telekomunikasi

(a)

	Nama Dosen	Bidang Minat
1	Amalia Rizqi Utami, S.T., M.T.	Telekomunikasi
2	Adi Mahmud Jaya Marindra, S.T., M.Eng.,	Telekomunikasi
3	Andhika Giyantara, S.T., M.T.	Sistem Pengaturan
4	Barokatun Hasanah, S.T., M.T.	Telekomunikasi
5	Happy Aprillia, S.ST., M.T., M.Eng., Ph.D.	Sistem Tenaga
6	Firilia Filiana, S.T., M.T.	Sistem Tenaga
7	Risty Jayanti Yuniar, S.T., M.T.	Sistem Pengaturan
8	Himawan Wicaksono, S.ST., M.T.	Teknik Komputer
9	Kharis Sugianto, SST., M.T.	Elektronika
10	Mifta Nur Farid, S.T., M.T.	Teknik Komputer
11	Sena Sukmananda Suprpto, S.T., M.T.	Elektronika
12	Thorikul Huda, S.T., M.T.	Sistem Tenaga
13	Vicky Andria Kusuma, S.ST., M.T.	Sistem Tenaga
14	Yun Tonce Kusuma Priyanto, S.T., M.T.	Sistem Tenaga
15	Muhammad Agung Nursyeha, S.T., M.T.	Elektronika
16	Muhammad Ridho Dewanto, S.T., M.T.	Teknik Komputer
17	Riza Hadi Saputra, S.T., M.T.	Sistem Pengaturan

(b)

Gambar 1. Dataset berupa (a) judul tugas akhir beserta bidang minatnya, serta (b) nama dosen beserta bidang keahliannya.

Judul harus jelas dan singkat. Nama penulis dan afiliasinya seperti yang tertulis diatas. Nama penulis ditulis secara jelas tanpa gelar. Penomoran heading dengan sistem Arabic dengan sub-heading maksimal hingga 3 tingkat.

## 2.2. Identifikasi Kebutuhan Sistem

Dilakukan identifikasi kebutuhan sistem, baik kebutuhan terhadap *hardware* maupun *software*. Kebutuhan *hardware* merupakan spesifikasi *hardware* yang digunakan dalam perancangan sistem NLP. Spesifikasi *hardware* yang digunakan ditunjukkan oleh Tabel 1. Sedangkan kebutuhan terhadap *software* adalah sebagai berikut ini:

1. Python versi 3.9.13
2. Spyder versi 5.2.2
3. *Python Library*
  - a. *Natural Language Tool Kit (NLTK)*: *library* yang digunakan untuk proses pengolahan teks yaitu *tokenization*, *stemming*, *case folding*, *stopword (filtering)*, dan penghapusan karakter spesial.
  - b. *Pandas*: *library* yang digunakan untuk membaca dataset berbentuk file csv.
  - c. *Numpy*: *library* yang digunakan untuk melakukan perhitungan saintifik matematika seperti perhitungan matriks, aljabar, statistik, dll.
  - d. *Sastrawi*: *library* yang dipakai adalah *StemmerFactory* sebagai pengubah kata berimbuhan bahasa Indonesia menjadi kata baku.
  - e. *Random*: *library* yang digunakan untuk menghasilkan angka acak.
  - f. *Regex*: *library* yang digunakan untuk mendeteksi angka, menghilangkan tanda baca dan angka.
  - g. *Tkinter*: *library* yang digunakan untuk membuat GUI/

Tabel 1. Spesifikasi *hardware* yang digunakan dalam perancangan sistem NLP

Komponen	Spesifikasi
Sistem Operasi	Windows 7 64 bit
CPU	Intel Core i5 7250U CPU @ 3.4 GHz
Memori	12 GB
Harddisk	500 GB

Keseluruhan *library* yang dibutuhkan akan di-*import* sebagaimana yang ditunjukkan oleh Algoritma 1.

### Algoritma 1. Import python library yang dibutuhkan

```

import pandas as pd
import re
import nltk
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
import numpy as np
import random
import tkinter as tk
nltk.download('punkt')
nltk.download('stopwords')

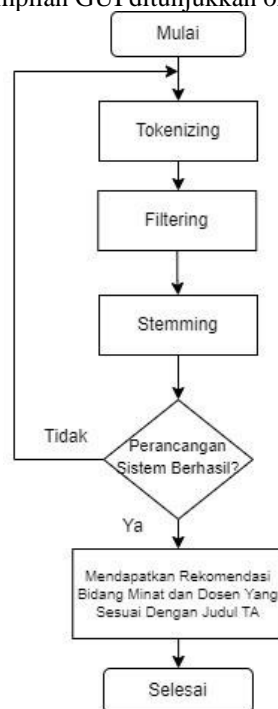
```

## 2.3. Perancangan Sistem

Sistem NLP yang dirancang mengacu pada diagram alir yang ditunjukkan oleh Gambar 2. Sistem dimulai dari perancangan proses *Tokenizing*, *Filtering*, dan juga *Stemming*. Proses *Tokenizing* adalah proses memecah kalimat dipecah menjadi kata. Proses *Tokenizing* ini dimulai dengan merubah keseluruhan kata di dalam kalimat menjadi huruf kecil, merubah selain huruf dan angka menjadi spasi, dan memecah kalimat menjadi kata. Deklarasi fungsi *Tokenizing* untuk proses training ditunjukkan oleh Algoritma 2 sedangkan proses testing ditunjukkan oleh Algoritma 3. Proses selanjutnya adalah *Filtering*. Proses ini adalah proses penyaringan kata berbahasa Indonesia. Penyaringan yang dimaksud adalah kata hasil *Tokenizing* yang tidak ada di dalam daftar kata *Corpus* Bahasa Indonesia akan dibuang. Deklarasi fungsi *Filtering* untuk proses training ditunjukkan oleh Algoritma 4 sedangkan proses testing ditunjukkan oleh Algoritma 5. Selanjutnya adalah proses *Stemming*. Proses ini adalah proses dimana kata hasil *Filtering* diubah menjadi kata baku tanpa awalan dan akhiran [1]. Deklarasi fungsi *Stemming* untuk proses training ditunjukkan oleh Algoritma 6 sedangkan proses testing ditunjukkan oleh Algoritma 7. Ketiga proses ini, yaitu *Tokenizing*, *Filtering*, dan *Stemming*, dideklarasikan dalam satu fungsi *preprocessing* sebagaimana yang ditunjukkan oleh Algoritma 8.

Setelah ketiga proses berhasil dirancang, selanjutnya dilakukan perancangan proses *training* dan *testing* yang disertai dengan pemanggilan dataset. Pada proses training, digunakan dilakukan perhitungan frekuensi distribusi dari kata-kata yang digunakan dalam judul tugas akhir terhadap masing-masing bidang minatnya. Judul tugas akhir yang digunakan sebagai *training* berjumlah 100 judul sedangkan 100 judul sisanya digunakan untuk *testing*.

Dirancang juga GUI untuk mempermudah pengguna dalam melakukan penginputan judul tugas akhir sebagai proses testing. Tampilan dari GUI yang dirancang ditunjukkan oleh Gambar 2. Algoritma dari proses training terhadap 100 dataset, proses testing, dan tampilan GUI ditunjukkan oleh Algoritma 9.



Gambar 1. Diagram alir perancangan sistem NLP

---

#### Algoritma 2. Proses Tokenizing untuk training

---

```
def tokenizing(doc):
    hasil_token = []
    for i in doc:
        ubah_huruf = i.lower()
        cleansing = re.sub('[^A-Za-z0-9]+', '', ubah_huruf)
        token = nltk.tokenize.word_tokenize(cleansing)
        hasil_token.append(token)
    return hasil_token
```

---

---

#### Algoritma 3. Proses Tokenizing untuk testing

---

```
def tokenizingQuery(query):
    ubah_huruf = query.lower()
    cleansing = re.sub('[^A-Za-z]+', '', ubah_huruf)
    token = nltk.tokenize.word_tokenize(cleansing)
    return token
return hasil_token
```

---

---

#### Algoritma 4. Proses Filtering untuk Training

---

```
def filtering(doc):
    hasil_filter = []
    stop_word = nltk.corpus.stopwords.words('indonesian')
    for i in doc:
        a = []
        for j in i:
            if(j not in stop_word):
                a.append(j)
        hasil_filter.append(a)
    return hasil_filter
```

---

---

Algoritma 5. Proses Filtering untuk Testing

---

```
def filteringQuery(query):
    stop_word = nltk.corpus.stopwords.words('indonesian')
    hasil_filter = []
    for i in query:
        if(i not in stop_word):
            hasil_filter.append(i)
    return hasil_filter
```

---

---

Algoritma 6. Proses Stemming untuk Training

---

```
def stemming(doc):
    hasil_stem = []
    stemmer = StemmerFactory().create_stemmer()
    for i in doc:
        a = []
        for j in i:
            a.append(stemmer.stem(j))
        hasil_stem.append(a)
    return hasil_stem
```

---

---

Algoritma 7. Proses Stemming untuk Testing

---

```
def stemmingQuery(query):
    hasil_stem = []
    stemmer = StemmerFactory().create_stemmer()
    for i in query:
        hasil_stem.append(stemmer.stem(i))
    return hasil_stem
```

---

---

Algoritma 7. Proses Preprocessing

---

```
def preprocessingDoc(doc):
    tokenizingDoc = tokenizing(doc)
    filteringDoc = filtering(tokenizingDoc)
    stemmingDoc = stemming(filteringDoc)
    return stemmingDoc
```

---

---

Algoritma 9. Proses Training, Testing, dan desain GUI

---

```
def hitungTF(doc):
    hasil_tf = []
    for i in doc:
        hasil_tf.append(nltk.FreqDist(i).most_common())
    return hasil_tf

def program():
    #TRAINING
    dataset = {"dataset": pd.read_csv("dataset fix.csv", delimiter=";")}
    datadosen = {"dosen": pd.read_csv("dosen.csv", delimiter=";"),
                 "dosen2": pd.read_csv("dosen2.csv", delimiter=";"),}
    judul = dataset["dataset"]["Judul Tugas Akhir"]
    bidangminat = dataset["dataset"]["Bidang Minat"]
    opsibidangminat = dataset["dataset"]["Kemungkinan Bidang Minat"]
    bidangdosen = datadosen["dosen"]["Bidang Minat"]
    dosen = datadosen["dosen"]["Nama Dosen"]
    dosen2 = datadosen["dosen2"]["Nama Dosen"]
    dokumen = hitungTF(preprocessingDoc(judul))

    #TESTING
    query = preprocessingQuery(input_judul.get())
    hasil = []
    for i in dokumen:
        a=0
```

```
for j in i:
    for k in query:
        if j[0] == k:
            a+=j[1]
hasil.append(a)
sort = np.flip(np.argsort(hasil))
hasil_bidang_minat = bidangminat[sort[0]]
kemungkinan_bidang_minat = opsibidangminat[sort[0]]

temp = []
for idx, i in enumerate(bidangdosen):
    if i == bidangminat[sort[0]]:
        temp.append(idx)

temp2 = []
for idx, i in enumerate(dosen):
    if idx in temp:
        temp2.append(i)
hasil_rekomendasi_dosen = random.sample(temp2,2)
while hasil_rekomendasi_dosen[0] in list(dosen2):
    hasil_rekomendasi_dosen = random.sample(temp2,2)

if kemungkinan_bidang_minat == '-':
    var_bidang.set(hasil_bidang_minat)
else:
    var_bidang.set(hasil_bidang_minat' / '+kemungkinan_bidang_minat)
var_dosen1.set(hasil_rekomendasi_dosen[0])
var_dosen2.set(hasil_rekomendasi_dosen[1])

if __name__ == '__main__':

    r = tk.Tk()
    r.geometry("600x300")
    r.title('PENENTUAN JUDUL TUGAS AKHIR')

    gambar = tk.PhotoImage(file="logoitk.png")
    w1 = tk.Label(r, image=gambar).place(x = 0, y = 0)
    gambar2 = tk.PhotoImage(file="logoelektro.png")
    w2 = tk.Label(r, image=gambar2).place(x = 50, y = 5)

    l = tk.Label(r, text = "PENENTUAN JUDUL TUGAS AKHIR",font=("Arial Bold", 15)).place(x = 140,y = 40)

    cari = tk.Button(r, text="CARI", command=program).place(x = 500,y = 82)

    judul = tk.Label(r, text = "JUDUL TUGAS AKHIR",font=("Segoe UI", 10)).place(x = 30, y = 80)
    input_judul = tk.Entry(r)
    input_judul.place(x = 180, y = 82, width=300, height=25)

    bidang_minat = tk.Label(r, text = "REKOMENDASI BIDANG MINAT",font=("Segoe UI", 10)).place(x = 30, y = 120)
    var_bidang = tk.StringVar()
    hasil_bidang = tk.Entry(r, state='readonly',textvariable=var_bidang)
    hasil_bidang.place(x = 230, y = 122, width=250, height=25)

    dosen1 = tk.Label(r, text = "REKOMENDASI PEMBIMBING 1",font=("Segoe UI", 10)).place(x = 30, y = 160)
    var_dosen1 = tk.StringVar()
    hasil_dosen1 = tk.Entry(r, state='readonly',textvariable=var_dosen1)
    hasil_dosen1.place(x = 230, y = 162, width=250, height=25)

    dosen2 = tk.Label(r, text = "REKOMENDASI PEMBIMBING 2",font=("Segoe UI", 10)).place(x = 30, y = 200)
    var_dosen2 = tk.StringVar()
    hasil_dosen2 = tk.Entry(r, state='readonly',textvariable=var_dosen2)
    hasil_dosen2.place(x = 230, y = 202, width=250, height=25)
    r.resizable(False, False)
    r.mainloop()
```

(a)

(b)

Gambar 2. Tampilan GUI aplikasi yang telah dibuat (a) saat belum diberikan *input* dan (b) saat diberikan *input* serta *output* yang dihasilkan.

### 3. Hasil dan Pembahasan

Dilakukan analisis akurasi, presisi, *recall* dan *f1-score* terhadap hasil dari sistem NLP yang telah dirancang. Hasil analisis dilakukan terhadap masing-masing bidang minat. Nilai akurasi, presisi, *recall*, dan *f1-score* didapatkan dari *confusion matrix* masing-masing bidang minat. *Confusion matrix* dari masing-masing bidang minat ditunjukkan oleh Tabel 2 untuk bidang Sistem Tenaga, Tabel 3 untuk bidang Elektronika, Tabel 4 untuk bidang Teknik Komputer, Tabel 5 untuk bidang Sistem Pengaturan, dan Tabel 6 untuk bidang Telekomunikasi. Dari Tabel 2 hingga Tabel 6 didapatkan nilai akurasi, presisi, *recall*, dan *f1-score* sebagaimana yang ditunjukkan oleh Tabel 7.

Berdasarkan Tabel 7, dapat diperhatikan bahwa bidang minat sistem pengaturan memiliki nilai akurasi, presisi, *recall* dan *f1-score* yang lebih tinggi dibandingkan dengan bidang minat yang lain. Sedangkan nilai terendah

dimiliki oleh bidang minat teknik komputer. Selanjutnya, nilai akurasi, presisi, *recall*, dan *f1-score* rata-rata dari keseluruhan bidang minat adalah 86.8 %, 66.2 %, 54.14 %, dan 64.775 %.

Tabel 2. Confusion matrix bidang sistem tenaga

Prediksi	Aktual	
	Sistem Tenaga	Selain Sistem Tenaga
Sistem Tenaga	13	7
Selain Sistem Tenaga	7	73

Tabel 3. Confusion matrix bidang elektronika

Prediksi	Aktual	
	Elektronika	Selain Elektronika
Elektronika	14	6
Selain Elektronika	7	72

Tabel 4. Confusion matrix bidang teknik komputer

Prediksi	Aktual	
	Teknik Komputer	Selain Teknik Komputer
Teknik Komputer	11	9
Selain Teknik Komputer	7	73

Tabel 5. Confusion matrix bidang sistem pengaturan

Prediksi	Aktual	
	Sistem Pengaturan	Selain Sistem Pengaturan
Sistem Pengaturan	16	4
Selain Sistem Sistem Pengaturan	6	74

Tabel 6. Confusion matrix bidang telekomunikasi

Prediksi	Aktual	
	Telekomunikasi	Selain Telekomunikasi
Telekomunikasi	14	6
Selain Telekomunikasi	6	74

Tabel 7. Nilai akurasi, presisi, recall, dan f1-score dari masing-masing bidang minat

Bidang Minat	Akurasi	Presisi	Recall	F1-Score
Sistem Tenaga	86 %	65 %	65 %	65 %
Elektronika	86 %	63 %	70 %	66.3 %
Teknik Komputer	84 %	61 %	55 %	57.8 %
Sistem Pengaturan	<b>90 %</b>	<b>72 %</b>	<b>80 %</b>	<b>75.7 %</b>
Telekomunikasi	88 %	70 %	70 %	70 %
<b>Rata-Rata</b>	<b>86.8 %</b>	<b>66.2 %</b>	<b>54.14 %</b>	<b>64.775 %</b>

#### 4. Kesimpulan

Berdasarkan hasil yang didapatkan, dapat disimpulkan bahwa kata-kata yang digunakan dalam judul tugas akhir memiliki pengaruh terhadap tingkat akurasi prediksi bidang minatnya.

#### Daftar Rujukan

- [1] Agastya, I. M. A. (2018). Pengaruh stemmer bahasa Indonesia terhadap performa analisis sentimen terjemahan ulasan film. *Jurnal Tekno Kompak*, 12(1), 18-23.
- [2] Panggali, R. R., Marisa, F., & Purnomo, D. (2017). Sistem Pendukung Keputusan Untuk Menentukan Judul Skripsi Jurusan Teknik Informatika Menggunakan Metode Topsis. *JIMP (Jurnal Informatika Merdeka Pasuruan)*, 2(1).
- [3] Fitriyah, N., Warsito, B., & Di Asih, I. M. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM). *Jurnal Gaussian*, 9(3), 376-390.
- [4] Herwin, H., & Andesa, K. (2019). Super Agent Chatbot “3S” Sebagai Media Informasi Menggunakan Metoda Natural Language Processing (NLP). *Jurnal Teknologi Dan Open Source*, 2(1), 53-64.
- [5] Indrayuni, E. (2019). Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Jurnal Khatulistiwa Informatika*, 7(1).
- [6] Khazari, A. S., Marisa, F., & Wijaya, I. D. (2017). Sistem Rekomendasi Penentuan Judul Skripsi Menggunakan Algoritma Decision Tree. *Jurnal Teknologi dan Manajemen Informatika*, 3(1).
- [7] Kurniawati, I. D., & Kusumawardhani, A. (2017). Implementasi Algoritma Canny Dalam Pengenalan Wajah Menggunakan Antarmuka GUI Matlab. *ResearchGate, Surabaya*.



- [8] Lutz, M. (2001). *Programming python*. " O'Reilly Media, Inc."
- [9] Novialdy, N., Darussalam, U., & Rahman, B. (2020). Expert System of Text Mining to Analyze Student Interaction in FTKI UNAS Online Lectures: Expert System of Text Mining to Analyze Student Interaction in FTKI UNAS Online Lectures. *Jurnal Mantik*, 3(4), 268-277.
- [10] Setiawan, G., Palit, H. N., & Setyati, E. (2019). Aspect Based Sentiment Analysis pada Layanan Umpan Balik Universitas dengan Menggunakan Metode Naïve Bayes dan Latent Semantic Analysis. *Jurnal Infra*, 7(1), 170-174.
- [11] Simarangkir, M. S. H. (2017). Studi Perbandingan Algoritma-Algoritma Stemming Untuk Dokumen Teks Bahasa Indonesia. *Jurnal Inkofar*, 1(1).
- [12] Soyusiawaty, D., & Jones, A. H. S. (2018). Penerapan Antar Muka Bahasa Alami dalam Pencarian Informasi Skripsi pada suatu Program Studi. *SEMNASTEKNOMEDIA ONLINE*, 6(1), 1-10.
- [13] Tommy, L., Kirana, C., & Lindawati, V. (2019). Recommender System dengan Kombinasi Apriori dan Content-Based Filtering pada Aplikasi Pemesanan Produk. *Jurnal Teknoinfo*, 13(2), 84-95.