

bahan_belajar

November 23, 2023

1 Text Mining and Web Scraping

Bagas Wibowo - 2023

1.1 Install Aplikasi Pendukung

- EmEditor (<https://www.emeditor.com/#download>) untuk big data text editor
- Anaconda (<https://www.anaconda.com/download/>) untuk platform Data Science berbasis Python
- Power BI Desktop (<https://powerbi.microsoft.com/en-us/downloads/>) untuk ETL Tools dan Visualisasi
- WebScrapet (<https://www.webscraper.io/>) untuk scrape data web menggunakan extension (plugin) di browser
- Visual Studio Code (<https://code.visualstudio.com/download>) untuk editor kode pemrograman

1.2 Scrape dengan menggunakan Web Scraper

#Sitemap artikel detik

```
{"_id":"artikel_detik","startUrl":["https://www.detik.com/edu/beasiswa/d-7010598/awardee-bongk
```

#Sitemap hasil pencarian detik

```
{"_id":"cari_info_detik","startUrl":["https://www.detik.com/search/searchall?query=lpdp&siteid
```

#Sitemap hasil pencarian google scholar

```
{"_id":"google_scholar","startUrl":["https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&as_y
```

2 Scrape dengan menggunakan Power BI Desktop

#URL Laporan Kinerja DJP 2022 <https://pajak.go.id/sites/default/files/2023-02/Laporan%20Kinerja>

```
#Kode Power Query let      Source = Pdf.Tables(Web.Contents("https://pajak.go.id/sites/default/
[Implementation="1.3"])),      Table057 = Source{[Id="Table057"]}[Data],
#"Promoted Headers" = Table.PromoteHeaders(Table057, [PromoteAllScalars=true]),
#"Filtered Rows" = Table.SelectRows("#Promoted Headers", each ([#"No."
<> null) and ([#"Kelompok#(lf)Pajak"] <> null)),      #"Changed Type" =
Table.TransformColumnTypes("#Filtered Rows",{{"Target#(lf)2022", type number},
{"2021", type number}, {"2022", type number}, {"%#(lf)Growth", type number},
```

```
{ "%#(lf)Growth_1", type number}, { "%#(lf)Penc.", type number}, { "%#(lf)Penc._2",
type number}}) in      #"Changed Type"
```

3 Scrape dengan menggunakan Python

```
import pandas
import os
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By

def sedotUmkh(url,setProvinsi,setKabupaten,setNama,pageSize,destinationDir):

    options=Options()
    options.add_argument("disable-infobars")
    options.add_argument("disable-extensions")
    options.add_argument("disable-blink-features=AutomationControlled")
    options.add_argument("start-maximized")

    browser=webdriver.Chrome(options=options)

    browser.get(url)
    provinsi=browser.find_element(By.NAME,"ctl00$MainContent$DropDownList1")
    provinsi.send_keys(setProvinsi)
    kabupaten=browser.find_element(By.NAME,"ctl00$MainContent$DropDownList2")
    kabupaten.send_keys(setKabupaten)
    nama=browser.find_element(By.NAME,"ctl00$MainContent$TextBox1")
    nama.send_keys(setNama)
    submit=browser.find_element(By.NAME,"ctl00$MainContent$Button1")
    submit.click()

    result=int(browser.find_element(By.XPATH,"//
↪*[@id='MainContent_DetailsView1']/tbody/tr/td[2]").text)

    pages=list(range(2,int(result / pageSize) + 2))

    def collectUrl(activeBrowser,targetUrls):
        links=activeBrowser.find_elements(By.TAG_NAME,"a")
        for link in links:
            hrefUrl=link.get_attribute("href")
            if(str(hrefUrl).startswith("https://umkm.depkop.go.id/Detail.aspx?
↪KoperasiId=")):
                targetUrls.append(hrefUrl)
        return targetUrls

    umkmUrls=[]
```

```

umkmUrls=collectUrl(browser,umkmUrls)

for page in pages:
    if(page%10 == 1):
        nextPage=browser.find_elements(By.LINK_TEXT,"...")[1]
    else:
        nextPage=browser.find_element(By.LINK_TEXT,str(page))
    nextPage.click()
    umkmUrls=collectUrl(browser,umkmUrls)
print(umkmUrls)

def storeTable(activeBrowser,destinationDir,fileName):
    table=activeBrowser.find_element(By.TAG_NAME,"table").
    ↪get_attribute("outerHTML")
    table=pandas.read_html(table,skiprows=0,header=0)[0]
    dfTable=pandas.DataFrame(table)

    if(destinationDir is None):
        print(dfTable)
    else:
        if(os.path.exists(destinationDir) is False):
            os.makedirs(destinationDir)
        targetName=fileName + ".tsv"
        dfTable.to_csv(destinationDir + os.sep +
    ↪targetName,index=False,sep="\t")

    for umkmUrl in umkmUrls:
        fileName=str(umkmUrl).split("=")[1]
        browser.get(umkmUrl)
        storeTable(browser,destinationDir,fileName)

url="https://umkm.depkop.go.id/"
provinsi="DKI Jakarta"
kota="Kota Jakarta Selatan"
kriteria="bengkel mobil"
pageSize=20
folder_simpan="C:\\repositories\\TextMining\\hasil_scraping\\python\\umkm"

sedotUmkm(url,provinsi,kota,kriteria,pageSize,folder_simpan)

```

3.1 TEXT MINING

3.2 Ekstrak data dari file CSV

```

import pandas as pd

#buka dataset dari file CSV dengan menggunakan Pandas

```

```
file_sumber="./hasil_scraping/webscraper/artikel_detik.csv"
csvDataFrame=pd.read_csv(file_sumber)
csvDataFrame
```

```
kontenDataFrame=csvDataFrame["konten"]
kontenDataFrame
```

```
konten=kontenDataFrame[0]
print(konten)
```

3.3 Ekstrak data dari file PDF

```
from pypdf import PdfReader

reader = PdfReader("./sumber_data/Awardee Bongkar 3 Rahasia Dapat Beasiswa LPDP.
↳pdf")
konten = ""
for page in reader.pages:
    konten += page.extract_text() + "\n"
print(konten)
```

Awardee Bongkar 3 Rahasia Dapat
Beasiswa LPDP, Intip Yuk!
Nikita Rosa - detikEdu
Senin, 30 Okt 2023 21:00 WIB

Jakarta - Margaret Ayu Caroline bertolak ke Inggris untuk melanjutkan studi di University College London (UCL) . Alumnus Fakultas Kedokteran Universitas Airlangga itu berangkat ke London dengan beasiswa Lembaga Pengelola Dana Pendidikan (LPDP) Kementerian Keuangan . Oline, panggilan akrabnya, menceritakan bahwa keinginan untuk melanjutkan kuliah berasal dari cita -citanya sebagai seorang dosen klinis . Baca juga: 3 Tips Persiapan Daftar S2 di Luar Negeri untuk Pekerja, Catat Nih!Baca juga: Sosok Dhani, Dokter Gigi Disabilitas Peraih Beasiswa LPDP ke Jerman "Aku punya cita - cita ingin jadi dosen klinis, karena aku tertarik menjadi seorang pengajar, " katanya dalam laman Unair dikutip Senin (30/10/2023) .

Ketertarikannya pada bidang ortopedi membuat Oline memilih Master of Science (MSc) Musculoskeletal Science and Medical Engineering sebagai jurusannya . Ia memilih UCL karena telah

bermitra dengan salah satu rumah sakit ortopedi terbaik dunia . Bagi mahasiswa yang berminat dengan beasiswa LPDP , Oline membongkar rahasianya . Yuk simak berikut ini .

Rahasia Dapat Beasiswa

LPDP : 1. Tentukan Alasan . Menurut Oline, segala proses akan berjalan dengan baik bila ada motivasi yang kuat . Bahkan Oline membutuhkan waktu yang cukup lama dalam menentukan alasan kuat miliknya . "Sebelum memutuskan mendaftar kuliah, aku harus menemukan alasan kuat dalam melanjutkan studi, ini berlangsung cukup lama . Kalau nggak ada alasan kuat, aku akan kesulitan dalam menjalankan segala proses, " jelasnya . 2. Punya Mentor Memiliki seorang mentor menjadi hal yang tak kalah penting . Jika ada mentor yang mendampingi, maka ia bisa memberi arahan ketika kamu mengalami kesulitan . "Mencari mentor artinya meminta bantuan dan bimbingan pendahulu yang sukses menimba pendidikan di luar negeri . Sehebat dan sepintar apapun, kita tidak bisa melakukan prosesnya sendiri, " papar Oline . 3. Cari Teman Seperjuangan . Menurut Oline, melanjutkan pendidikan bisa membuat seseorang merasa kesepian . Teman seperjuangan bisa memberikan dorongan motivasi serta tempat berkeluh kesah karena ada pada posisi yang sama .

Terlebih, proses yang dijalani tidak sebentar . Mulai dari mendaftar kampus hingga beasiswa . "Nggak ada yang namanya menyiapkan persyaratan masuk kampus hingga beasiswa hanya satu atau dua bulan . Harus ada timeline yang tersusun, " ujarnya . Baca juga: Kemenag Buka Pendaftaran Beasiswa

Penyelesaian Studi S2 -S3, Bernilai Rp 25 Juta

Oline berpesan bahwa dalam menjalani segala proses maka seseorang harus menikmatinya . Meski terasa sakit, proses harus terus berjalan . "Semakin berkembang maka rasa sakit yang terasa akan semakin besar . Memang prosesnya terasa nggak enak, rasanya menyakitkan . Tapi proses harus terus berjalan, " pungkasnya .

3.4 Cleansing Data dengan Regex

```
import re

#buang semua non alphanumeric, spasi dan titik
konten_clean=re.sub(r"[^A-Za-z0-9\s.]", "", konten)
```

```
#ubah semua multi-spasi menjadi single spasi
konten_clean=re.sub(r"\\s+", " ",konten_clean)

#ubah semua spasi sebelum titik
konten_clean=re.sub(r"\\s[.]", ".",konten_clean)
print(konten_clean)
```

Awardee Bongkar 3 Rahasia Dapat Beasiswa LPDP Intip Yuk Nikita Rosa detikEdu
 Senin 30 Okt 2023 2100 WIB Jakarta Margaret Ayu Caroline bertolak ke Inggris
 untuk melanjutkan studi di University College London UCL. Alumnus Fakultas
 Kedokteran Universitas Airlangga itu berangkat ke London dengan beasiswa Lembaga
 Pengelola Dana Pendidikan LPDP Kementerian Keuangan. Oline panggilan akrabnya
 menceritakan bahwa keinginan untuk melanjutkan kuliah berasal dari cita citanya
 sebagai seorang dosen klinis. Baca juga 3 Tips Persiapan Daftar S2 di Luar
 Negeri untuk Pekerja Catat NihBaca juga Sosok Dhani Dokter Gigi Disabilitas
 Peraih Beasiswa LPDP ke Jerman Aku punya cita cita ingin jadi dosen klinis
 karena aku tertarik menjadi seorang pengajar katanya dalam laman Unair dikutip
 Senin 30102023. Ketertarikannya pada bidang ortopedi membuat Oline memilih
 Master of Science MSc Musculoskeletal Science and Medical Engineering sebagai
 jurusannya. Ia memilih UCL karena telah bermitra dengan salah satu rumah sakit
 ortopedi terbaik dunia. Bagi mahasiswa yang berminat dengan beasiswa LPDP Oline
 membongkar rahasianya. Yuk simak berikut ini. Rahasia Dapat Beasiswa LPDP 1.
 Tentukan Alasan. Menurut Oline segala proses akan berjalan dengan baik bila ada
 motivasi yang kuat. Bahkan Oline membutuhkan waktu yang cukup lama dalam
 menentukan alasan kuat miliknya. Sebelum memutuskan mendaftar kuliah aku harus
 menemukan alasan kuat dalam melanjutkan studi ini berlangsung cukup lama. Kalau
 nggak ada alasan kuat aku akan kesulitan dalam menjalankan segala proses
 jelasnya. 2. Punya Mentor Memiliki seorang mentor menjadi hal yang tak kalah
 penting. Jika ada mentor yang mendampingi maka ia bisa memberi arahan ketika
 kamu mengalami kesulitan. Mencari mentor artinya meminta bantuan dan bimbingan
 pendahulu yang sukses menimba pendidikan di luar negeri. Sehebat dan sepintar
 apapun kita tidak bisa melakukan prosesnya sendiri papar Oline. 3. Cari Teman
 Seperjuangan. Menurut Oline melanjutkan pendidikan bisa membuat seseorang merasa
 kesepian. Teman seperjuangan bisa memberikan dorongan motivasi serta tempat
 berkeluh kesah karena ada pada posisi yang sama. Terlebih proses yang dijalani
 tidak sebentar. Mulai dari mendaftar kampus hingga beasiswa. Nggak ada yang
 namanya menyiapkan persyaratan masuk kampus hingga beasiswa hanya satu atau dua
 bulan. Harus ada timeline yang tersusun ujanya. Baca juga Kemenag Buka
 Pendaftaran Beasiswa Penyelesaian Studi S2 S3 Bernilai Rp 25 JutaOline berpesan
 bahwa dalam menjalani segala proses maka seseorang harus menikmatinya. Meski
 terasa sakit proses harus terus berjalan. Semakin berkembang maka rasa sakit
 yang terasa akan semakin besar. Memang prosesnya terasa nggak enak rasanya
 menyakitkan. Tapi proses harus terus berjalan pungkasnya.

3.5 Remove Stopwords

```
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
import re

def getStopWords(customList=None, listFilePath=None):
    #mendapatkan default stopwords bahasa Indonesia
    factory = StopWordRemoverFactory()
    stopwords = factory.get_stop_words()

    #menambahkan daftar stopwords custom
    if(customList is not None):
        my_additional_stop_words=customList
    else:
        my_additional_stop_words=[]
    if(listFilePath is not None):
        with open(listFilePath) as f:
            for line in f :
                my_additional_stop_words += [line[:-1]]
    all_stopwords = stopwords + my_additional_stop_words

    return all_stopwords

def removeStopWords(sentence, stop_words):
    words = sentence.split()
    return ' '.join([word for word in words if word.lower() not in stop_words])

konten_clean_from_stopwords=removeStopWords(konten_clean,getStopWords())
print(konten_clean_from_stopwords)
```

Awardee Bongkar 3 Rahasia Beasiswa LPDP Intip Yuk Nikita Rosa detikEdu Senin 30 Okt 2023 2100 WIB Jakarta Margaret Ayu Caroline bertolak Inggris melanjutkan studi University College London UCL. Alumnus Fakultas Kedokteran Universitas Airlangga berangkat London beasiswa Lembaga Pengelola Dana Pendidikan LPDP Kementerian Keuangan. Oline panggilan akrabnya menceritakan keinginan melanjutkan kuliah berasal cita citanya seorang dosen klinis. Baca 3 Tips Persiapan Daftar S2 Luar Negeri Pekerja Catat NihBaca Sosok Dhani Dokter Gigi Disabilitas Peraih Beasiswa LPDP Jerman Aku punya cita cita jadi dosen klinis aku tertarik menjadi seorang pengajar katanya laman Unair dikutip Senin 30102023. Ketertarikannya bidang ortopedi membuat Oline memilih Master of Science MSc Musculoskeletal Science and Medical Engineering jurusanannya. memilih UCL bermitra salah satu rumah sakit ortopedi terbaik dunia. mahasiswa berminat beasiswa LPDP Oline membongkar rahasianya. Yuk simak berikut ini. Rahasia Beasiswa LPDP 1. Tentukan Alasan. Oline segala proses berjalan baik bila motivasi kuat. Bahkan Oline membutuhkan waktu cukup lama menentukan alasan kuat miliknya. memutuskan mendaftar kuliah aku menemukan alasan kuat melanjutkan studi berlangsung cukup lama. Kalau alasan kuat aku kesulitan menjalankan segala

proses jelasnya. 2. Punya Mentor Memiliki seorang mentor menjadi tak kalah penting. mentor mendampingi memberi arahan kamu mengalami kesulitan. Mencari mentor artinya meminta bantuan bimbingan pendahulu sukses menimba pendidikan luar negeri. Sehebat sepintar apapun melakukan prosesnya sendiri papas Online. 3. Cari Teman Seperjuangan. Online melanjutkan pendidikan membuat seseorang merasa kesepian. Teman seperjuangan memberikan dorongan motivasi tempat berkeluh kesah posisi sama. Terlebih proses dijalani sebentar. Mulai mendaftar kampus hingga beasiswa. namanya menyiapkan persyaratan masuk kampus hingga beasiswa satu bulan. timeline tersusun ujanya. Baca Kemenag Buka Pendaftaran Beasiswa Penyelesaian Studi S2 S3 Bernilai Rp 25 Juta Online berpesanan menjalani segala proses seseorang menikmatinya. Meski terasa sakit proses terus berjalan. Semakin berkembang rasa sakit terasa semakin besar. Memang prosesnya terasa enak rasanya menyakitkan. proses terus berjalan pungkasnya.

3.6 Tokenizer dan Stemmer

```
modelName="indobenchmark/indobert-base-p1"

from transformers import AutoTokenizer,AutoModel,pipeline
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory

model=AutoModel.from_pretrained(modelName)
tokenizer=AutoTokenizer.from_pretrained(modelName)

def stemmer():
    # Inisialisasi stemmer
    factory=StemmerFactory()
    stemmerAgent=factory.create_stemmer()
    return stemmerAgent

stemmer=stemmer()
```

```
def stemFromText(stemmer,text):
    stemmedText=stemmer.stem(text)
    return stemmedText

stemText=stemFromText(stemmer,konten_clean_from_stopwords)
print(stemText)
```

awardee bongkar 3 rahasia beasiswa lpdp intip yuk nikita rosa detikedu senin 30 okt 2023 2100 wib jakarta margaret ayu caroline tolak inggris lanjut studi university college london ucl alumnus fakultas dokter universitas airlangga berangkat london beasiswa lembaga kelola dana didik lpdp menteri uang online panggil akrab cerita ingin lanjut kuliah asal cita cita orang dosen klinis baca 3 tips siap daftar s2 luar negeri kerja catat nihbaca sosok dhani dokter gigi disabilitas raih beasiswa lpdp jerman aku punya cita cita jadi dosen klinis aku tarik jadi orang ajar kata laman unair kutip senin 30102023 tari bidang ortopedi

buat online pilih master of science msc musculoskeletal science and medical engineering jurusan pilih ucl mitra salah satu rumah sakit ortopedi baik dunia mahasiswa minat beasiswa lpdp online bongkar rahasia yuk simak ikut ini rahasia beasiswa lpdp 1 tentu alasan online segala proses jalan baik bila motivasi kuat bahkan online butuh waktu cukup lama tentu alasan kuat milik putus daftar kuliah aku temui alasan kuat lanjut studi langsung cukup lama kalau alasan kuat aku sulit jalan segala proses jelas 2 punya mentor milik orang mentor jadi tak kalah penting mentor damping beri arah kamu alami sulit cari mentor arti minta bantu bimbing dahulu sukses timba didik luar negeri hebat pintar apa laku proses sendiri papir online 3 cari temanjuang online lanjut didik buat orang rasa sepi temanjuang beri dorongan motivasi tempat keluh kesah posisi sama lebih proses jalan sebentar mulai daftar kampus hingga beasiswa nama siap syarat masuk kampus hingga beasiswa satu bulan timeline susun ujar baca kemenag buka daftar beasiswa selesai studi s2 s3 nilai rp 25 juta online pesan jalan segala proses orang nikmat meski asa sakit proses terus jalan makin kembang rasa sakit asa makin besar memang proses asa enak rasa sakit proses terus jalan pungkas

```
def tokenFromText(pretrainedTokenizer,text):
    tokens=pretrainedTokenizer.tokenize(text)
    return tokens

tokens=tokenFromText(tokenizer,konten_clean_from_stopwords)
print(tokens)
```

```
['award', '##ee', 'bongkar', '3', 'rahasia', 'beasiswa', 'lp', '##dp', 'intip',
'yuk', 'nikita', 'rosa', 'detik', '##edu', 'senin', '30', 'okt', '202', '##3',
'2100', 'wib', 'jakarta', 'margar', '##et', 'ayu', 'carol', '##ine', 'bertolak',
'inggris', 'melanjutkan', 'studi', 'university', 'college', 'london', 'uc',
'##l', '.', 'alumnus', 'fakultas', 'kedokteran', 'universitas', 'airlangga',
'berangkat', 'london', 'beasiswa', 'lembaga', 'pengelola', 'dana', 'pendidikan',
'lp', '##dp', 'kementerian', 'keuangan', '.', 'oli', '##ne', 'panggilan',
'akrab', '##nya', 'menceritakan', 'keinginan', 'melanjutkan', 'kuliah',
'berasal', 'cita', 'cita', '##nya', 'seorang', 'dosen', 'klinis', '.', 'baca',
'3', 'tips', 'persiapan', 'daftar', 's', '##2', 'luar', 'negeri', 'pekerja',
'catat', 'nih', '##baca', 'sosok', 'dhani', 'dokter', 'gigi', 'disabilitas',
'peraih', 'beasiswa', 'lp', '##dp', 'jerman', 'aku', 'punya', 'cita', 'cita',
'jadi', 'dosen', 'klinis', 'aku', 'tertarik', 'menjadi', 'seorang', 'pengajar',
'katanya', 'laman', 'unair', 'dikutip', 'senin', '30', '##102', '##02', '##3',
.', 'ketertarikan', '##nya', 'bidang', 'ort', '##ope', '##di', 'membuat',
'oli', '##ne', 'memilih', 'master', 'of', 'science', 'msc', 'mus', '##cul',
'##os', '##kel', '##etal', 'science', 'and', 'medical', 'engineering',
'jurusannya', '.', 'memilih', 'uc', '##l', 'bermitra', 'salah', 'satu', 'rumah',
'sakit', 'ort', '##ope', '##di', 'terbaik', 'dunia', '.', 'mahasiswa',
'berminat', 'beasiswa', 'lp', '##dp', 'oli', '##ne', 'membongkar', 'rahasianya',
.', 'yuk', 'simak', 'berikut', 'ini', '.', 'rahasia', 'beasiswa', 'lp', '##dp',
'1', '.', 'tentukan', 'alasan', '.', 'oli', '##ne', 'segala', 'proses',
'berjalan', 'baik', 'bila', 'motivasi', 'kuat', '.', 'bahkan', 'oli', '##ne',
'membutuhkan', 'waktu', 'cukup', 'lama', 'menentukan', 'alasan', 'kuat',
```

'miliknya', '.', 'memutuskan', 'mendaftar', 'kuliah', 'aku', 'menemukan',
 'alasan', 'kuat', 'melanjutkan', 'studi', 'berlangsung', 'cukup', 'lama', '.',
 'kalau', 'alasan', 'kuat', 'aku', 'kesulitan', 'menjalankan', 'segala',
 'proses', 'jelasnya', '.', '2', '.', 'punya', 'mentor', 'memiliki', 'seorang',
 'mentor', 'menjadi', 'tak', 'kalah', 'penting', '.', 'mentor', 'mendampingi',
 'memberi', 'arahan', 'kamu', 'mengalami', 'kesulitan', '.', 'mencari', 'mentor',
 'artinya', 'meminta', 'bantuan', 'bimbingan', 'pendahulu', 'sukses', 'menimba',
 'pendidikan', 'luar', 'negeri', '.', 'seh', '##ebat', 'sepi', '##nt', '##ar',
 'apapun', 'melakukan', 'prosesnya', 'sendiri', 'papar', 'oli', '##ne', '.', '3',
 '.', 'cari', 'teman', 'seper', '##juan', '##gan', '.', 'oli', '##ne',
 'melanjutkan', 'pendidikan', 'membuat', 'seseorang', 'merasa', 'keseharian', '.',
 'teman', 'seper', '##juan', '##gan', 'memberikan', 'dorongan', 'motivasi',
 'tempat', 'berkel', '##uh', 'kesah', 'posisi', 'sama', '.', 'terlebih',
 'proses', 'dijalani', 'sebentar', '.', 'mulai', 'mendaftar', 'kampus', 'hingga',
 'beasiswa', '.', 'namanya', 'menyiapkan', 'persyaratan', 'masuk', 'kampus',
 'hingga', 'beasiswa', 'satu', 'bulan', '.', 'timeline', 'tersusun', 'ujarnya',
 '.', 'baca', 'kemenag', 'buka', 'pendaftaran', 'beasiswa', 'penyelesaian',
 'studi', 's', '##2', 's', '##3', 'bernilai', 'rp', '25', 'juta', '##olin',
 '##e', 'berpesan', 'menjalani', 'segala', 'proses', 'seseorang', 'menikmatinya',
 '.', 'meski', 'terasa', 'sakit', 'proses', 'terus', 'berjalan', '.', 'semakin',
 'berkembang', 'rasa', 'sakit', 'terasa', 'semakin', 'besar', '.', 'memang',
 'prosesnya', 'terasa', 'enak', 'rasanya', 'menyakitkan', '.', 'proses', 'terus',
 'berjalan', 'pungkasnya', '.']

3.7 Sentiment Analysis

```
def sentimentAnalyzer(modelName):
    nlp=pipeline(
        "sentiment-analysis",
        model=modelName,
        tokenizer=modelName
    )
    return nlp

def sentimentAnalysis(sentimentAnalyzer,text):
    result=sentimentAnalyzer(text)
    return result

sentimentAnalysisModelName="w11wo/indonesian-roberta-base-sentiment-classifier"
sentimentAnalyzer=sentimentAnalyzer(sentimentAnalysisModelName)

textPositif="Aku sangat senang sekali belajar hal baru seperti NLP"
sentimentAnaysisResult=sentimentAnalysis(sentimentAnalyzer,textPositif)
print(textPositif)
print(sentimentAnaysisResult)

textNegatif="Huh, susah banget sih belajar NLP"
```

```

sentimentAnaysisResult=sentimentAnalysis(sentimentAnalyzer,textNegatif)
print(textNegatif)
print(sentimentAnaysisResult)

sentimentAnaysisResult=sentimentAnalysis(sentimentAnalyzer,konten_clean_from_stopwords)
print(konten_clean_from_stopwords)
print(sentimentAnaysisResult)

```

Aku sangat senang sekali belajar hal baru seperti NLP

```
[{'label': 'positive', 'score': 0.9981694221496582}]
```

Huh, susah banget sih belajar NLP

```
[{'label': 'negative', 'score': 0.9986188411712646}]
```

Awardee Bongkar 3 Rahasia Beasiswa LPDP Intip Yuk Nikita Rosa detikEdu Senin 30 Okt 2023 2100 WIB Jakarta Margaret Ayu Caroline bertolak Inggris melanjutkan studi University College London UCL. Alumnus Fakultas Kedokteran Universitas Airlangga berangkat London beasiswa Lembaga Pengelola Dana Pendidikan LPDP Kementerian Keuangan. Oline panggilan akrabnya menceritakan keinginan melanjutkan kuliah berasal cita citanya seorang dosen klinis. Baca 3 Tips Persiapan Daftar S2 Luar Negeri Pekerja Catat NihBaca Sosok Dhani Dokter Gigi Disabilitas Peraih Beasiswa LPDP Jerman Aku punya cita cita jadi dosen klinis aku tertarik menjadi seorang pengajar katanya laman Unair dikutip Senin 30102023. Ketertarikannya bidang ortopedi membuat Oline memilih Master of Science MSc Musculoskeletal Science and Medical Engineering jurusanannya. memilih UCL bermitra salah satu rumah sakit ortopedi terbaik dunia. mahasiswa berminat beasiswa LPDP Oline membongkar rahasianya. Yuk simak berikut ini. Rahasia Beasiswa LPDP 1. Tentukan Alasan. Oline segala proses berjalan baik bila motivasi kuat. Bahkan Oline membutuhkan waktu cukup lama menentukan alasan kuat miliknya. memutuskan mendaftar kuliah aku menemukan alasan kuat melanjutkan studi berlangsung cukup lama. Kalau alasan kuat aku kesulitan menjalankan segala proses jelasnya. 2. Punya Mentor Memiliki seorang mentor menjadi tak kalah penting. mentor mendampingi memberi arahan kamu mengalami kesulitan. Mencari mentor artinya meminta bantuan bimbingan pendahulu sukses menimba pendidikan luar negeri. Sehebat sepintar apapun melakukan prosesnya sendiri papar Oline. 3. Cari Teman Seperjuangan. Oline melanjutkan pendidikan membuat seseorang merasa kesepian. Teman seperjuangan memberikan dorongan motivasi tempat berkeluh kesah posisi sama. Terlebih proses dijalani sebentar. Mulai mendaftar kampus hingga beasiswa. namanya menyiapkan persyaratan masuk kampus hingga beasiswa satu bulan. timeline tersusun ujanya. Baca Kemenag Buka Pendaftaran Beasiswa Penyelesaian Studi S2 S3 Bernilai Rp 25 JutaOline berpesan menjalani segala proses seseorang menikmatinya. Meski terasa sakit proses terus berjalan. Semakin berkembang rasa sakit terasa semakin besar. Memang prosesnya terasa enak rasanya menyakitkan. proses terus berjalan pungkasnya.

```
[{'label': 'negative', 'score': 0.8194615244865417}]
```

3.8 POS Analysis

```
import malaya

def posAnalyzer(modelName):
    model=malaya.pos.transformer(model=modelName)
    return model

def posAnalysis(posAnalyzer,text):
    result=posAnalyzer.predict(text)
    return result

modelName="bert"
posAnalyzer=posAnalyzer(modelName)
posAnalysis=posAnalysis(posAnalyzer,konten_clean_from_stopwords)

print(posAnalysis)
```

```
[('Awardee', 'PROPN'), ('Bongkar', 'PROPN'), ('3', 'NUM'), ('Rahasia', 'PROPN'),
('Beasiswa', 'PROPN'), ('LPDP', 'PROPN'), ('Intip', 'PROPN'), ('Yuk', 'PROPN'),
('Nikita', 'PROPN'), ('Rosa', 'PROPN'), ('detikEdu', 'PROPN'), ('Senin',
'PROPN'), ('30', 'NUM'), ('Okt', 'PROPN'), ('2023', 'NUM'), ('2100', 'NUM'),
('WIB', 'PROPN'), ('Jakarta', 'PROPN'), ('Margaret', 'PROPN'), ('Ayu', 'PROPN'),
('Caroline', 'PROPN'), ('bertolak', 'VERB'), ('Inggris', 'PROPN'),
('melanjutkan', 'VERB'), ('studi', 'NOUN'), ('University', 'PROPN'), ('College',
'PROPN'), ('London', 'PROPN'), ('UCL', 'PROPN'), ('.', 'PUNCT'), ('Alumnus',
'PROPN'), ('Fakultas', 'PROPN'), ('Kedokteran', 'PROPN'), ('Universitas',
'PROPN'), ('Airlangga', 'PROPN'), ('berangkat', 'VERB'), ('London', 'PROPN'),
('beasiswa', 'PROPN'), ('Lembaga', 'PROPN'), ('Pengelola', 'PROPN'), ('Dana',
'PROPN'), ('Pendidikan', 'PROPN'), ('LPDP', 'PROPN'), ('Kementerian', 'PROPN'),
('Keuangan', 'PROPN'), ('.', 'PUNCT'), ('Oline', 'PROPN'), ('panggilan',
'NOUN'), ('akrabnya', 'ADJ'), ('menceritakan', 'VERB'), ('keinginan', 'NOUN'),
('melanjutkan', 'VERB'), ('kuliah', 'NOUN'), ('berasal', 'VERB'), ('cita',
'NOUN'), ('citanya', 'NOUN'), ('seorang', 'DET'), ('dosen', 'NOUN'), ('klinik',
'ADJ'), ('.', 'PUNCT'), ('Baca', 'VERB'), ('3', 'NUM'), ('Tips', 'PROPN'),
('Persiapan', 'PROPN'), ('Daftar', 'PROPN'), ('S2', 'PROPN'), ('Luar', 'PROPN'),
('Negeri', 'PROPN'), ('Pekerja', 'PROPN'), ('Catat', 'VERB'), ('NihBaca',
'PROPN'), ('Sosok', 'PROPN'), ('Dhani', 'PROPN'), ('Dokter', 'PROPN'), ('Gigi',
'PROPN'), ('Disabilitas', 'PROPN'), ('Peraih', 'PROPN'), ('Beasiswa', 'PROPN'),
('LPDP', 'PROPN'), ('Jerman', 'PROPN'), ('Aku', 'PRON'), ('punya', 'VERB'),
('cita', 'NOUN'), ('cita', 'NOUN'), ('jadi', 'VERB'), ('dosen', 'NOUN'),
('klinik', 'ADJ'), ('aku', 'PRON'), ('tertarik', 'VERB'), ('menjadi', 'VERB'),
('seorang', 'DET'), ('pengajar', 'NOUN'), ('katanya', 'NOUN'), ('laman',
'NOUN'), ('Unair', 'PROPN'), ('dikutip', 'VERB'), ('Senin', 'PROPN'),
('30102023', 'NUM'), ('.', 'PUNCT'), ('Ketertarikannya', 'PROPN'), ('bidang',
'NOUN'), ('ortopedi', 'CCONJ'), ('membuat', 'VERB'), ('Oline', 'PROPN'),
('memilih', 'VERB'), ('Master', 'PROPN'), ('of', 'PROPN'), ('Science', 'PROPN'),
('MSc', 'PROPN'), ('Musculoskeletal', 'PROPN'), ('Science', 'PROPN'), ('and',
```

'CCONJ'), ('Medical', 'PROPN'), ('Engineering', 'PROPN'), ('jurusannya',
 'NOUN'), ('.', 'PUNCT'), ('memilih', 'VERB'), ('UCL', 'PROPN'), ('bermitra',
 'VERB'), ('salah', 'DET'), ('satu', 'DET'), ('rumah', 'NOUN'), ('sakit',
 'NOUN'), ('ortopedi', 'PROPN'), ('terbaik', 'ADJ'), ('dunia', 'NOUN'), ('.',
 'PUNCT'), ('mahasiswa', 'NOUN'), ('berminat', 'ADJ'), ('beasiswa', 'NOUN'),
 ('LPDP', 'PROPN'), ('Oline', 'PROPN'), ('membongkar', 'VERB'), ('rahasianya',
 'NOUN'), ('.', 'PUNCT'), ('Yuk', 'PROPN'), ('simak', 'VERB'), ('berikut',
 'VERB'), ('ini', 'DET'), ('.', 'PUNCT'), ('Rahasia', 'PROPN'), ('Beasiswa',
 'PROPN'), ('LPDP', 'PROPN'), ('1', 'NUM'), ('.', 'PUNCT'), ('Tentukan', 'ADV'),
 ('Alasan', 'NOUN'), ('.', 'PUNCT'), ('Oline', 'PROPN'), ('segala', 'DET'),
 ('proses', 'NOUN'), ('berjalan', 'VERB'), ('baik', 'ADJ'), ('bila', 'SCONJ'),
 ('motivasi', 'NOUN'), ('kuat', 'ADJ'), ('.', 'PUNCT'), ('Bahkan', 'SCONJ'),
 ('Oline', 'PROPN'), ('membutuhkan', 'VERB'), ('waktu', 'NOUN'), ('cukup',
 'ADV'), ('lama', 'ADJ'), ('menentukan', 'VERB'), ('alasan', 'NOUN'), ('kuat',
 'ADJ'), ('miliknya', 'NOUN'), ('.', 'PUNCT'), ('memutuskan', 'VERB'),
 ('mendaftar', 'VERB'), ('kuliah', 'NOUN'), ('aku', 'PRON'), ('menemukan',
 'VERB'), ('alasan', 'NOUN'), ('kuat', 'ADJ'), ('melanjutkan', 'VERB'), ('studi',
 'NOUN'), ('berlangsung', 'VERB'), ('cukup', 'ADV'), ('lama', 'ADJ'), ('.',
 'PUNCT'), ('Kalau', 'SCONJ'), ('alasan', 'NOUN'), ('kuat', 'ADJ'), ('aku',
 'PRON'), ('kesulitan', 'NOUN'), ('menjalankan', 'VERB'), ('segala', 'DET'),
 ('proses', 'NOUN'), ('jelasnya', 'ADJ'), ('.', 'PUNCT'), ('2', 'NUM'), ('.',
 'PUNCT'), ('Punya', 'PROPN'), ('Mentor', 'PROPN'), ('Memiliki', 'VERB'),
 ('seorang', 'DET'), ('mentor', 'NOUN'), ('menjadi', 'VERB'), ('tak', 'ADV'),
 ('kalah', 'VERB'), ('penting', 'ADJ'), ('.', 'PUNCT'), ('mentor', 'NOUN'),
 ('mendampingi', 'VERB'), ('memberi', 'VERB'), ('arahan', 'NOUN'), ('kamu',
 'PRON'), ('mengalami', 'VERB'), ('kesulitan', 'NOUN'), ('.', 'PUNCT'),
 ('Mencari', 'PROPN'), ('mentor', 'NOUN'), ('artinya', 'VERB'), ('meminta',
 'VERB'), ('bantuan', 'NOUN'), ('bimbingan', 'NOUN'), ('pendahulu', 'NOUN'),
 ('sukses', 'ADJ'), ('menimba', 'VERB'), ('pendidikan', 'NOUN'), ('luar', 'ADP'),
 ('negeri', 'NOUN'), ('.', 'PUNCT'), ('Sehebat', 'PROPN'), ('sepintar', 'NOUN'),
 ('apapun', 'DET'), ('melakukan', 'VERB'), ('prosesnya', 'NOUN'), ('sendiri',
 'ADJ'), ('papar', 'VERB'), ('Oline', 'PROPN'), ('.', 'PUNCT'), ('3', 'NUM'),
 ('.', 'PUNCT'), ('Cari', 'VERB'), ('Teman', 'PROPN'), ('Seperjuangan', 'PROPN'),
 ('.', 'PUNCT'), ('Oline', 'PROPN'), ('melanjutkan', 'VERB'), ('pendidikan',
 'NOUN'), ('membuat', 'VERB'), ('seseorang', 'NOUN'), ('merasa', 'VERB'),
 ('kesepian', 'PUNCT'), ('.', 'PUNCT'), ('Teman', 'PROPN'), ('seperjuangan',
 'ADJ'), ('memberikan', 'VERB'), ('dorongan', 'NOUN'), ('motivasi', 'NOUN'),
 ('tempat', 'NOUN'), ('berkeluh', 'VERB'), ('kesah', 'NOUN'), ('posisi', 'NOUN'),
 ('sama', 'ADJ'), ('.', 'PUNCT'), ('Terlebih', 'PROPN'), ('proses', 'NOUN'),
 ('dijalani', 'NOUN'), ('sebentar', 'ADV'), ('.', 'PUNCT'), ('Mulai', 'VERB'),
 ('mendaftar', 'VERB'), ('kampus', 'PROPN'), ('hingga', 'ADP'), ('beasiswa',
 'NOUN'), ('.', 'PUNCT'), ('namanya', 'NOUN'), ('menyiapkan', 'VERB'),
 ('persyaratan', 'NOUN'), ('masuk', 'VERB'), ('kampus', 'PROPN'), ('hingga',
 'ADP'), ('beasiswa', 'NOUN'), ('satu', 'NUM'), ('bulan', 'NOUN'), ('.',
 'PUNCT'), ('timeline', 'NOUN'), ('tersusun', 'NOUN'), ('ujarnya', 'NOUN'), ('.',
 'PUNCT'), ('Baca', 'VERB'), ('Kemenag', 'PROPN'), ('Buka', 'PROPN'),
 ('Pendaftaran', 'PROPN'), ('Beasiswa', 'PROPN'), ('Penyelesaian', 'PROPN'),
 ('Studi', 'PROPN'), ('S2', 'PROPN'), ('S3', 'PROPN'), ('Bernilai', 'PROPN'),

```
(('Rp', 'PROPN'), ('25', 'NUM'), ('JutaOline', 'PROPN'), ('berpesan', 'PROPN'),
('menjalani', 'VERB'), ('segala', 'DET'), ('proses', 'NOUN'), ('seseorang',
'NOUN'), ('menikmatinya', 'VERB'), ('.', 'PUNCT'), ('Meski', 'SCONJ'),
('terasa', 'VERB'), ('sakit', 'ADJ'), ('proses', 'NOUN'), ('terus', 'ADV'),
('berjalan', 'VERB'), ('.', 'PUNCT'), ('Semakin', 'ADV'), ('berkembang',
'NOUN'), ('rasa', 'NOUN'), ('sakit', 'ADJ'), ('terasa', 'VERB'), ('semakin',
'ADV'), ('besar', 'ADJ'), ('.', 'PUNCT'), ('Memang', 'PROPN'), ('prosesnya',
'NOUN'), ('terasa', 'VERB'), ('enak', 'ADJ'), ('rasanya', 'NOUN'),
('menyakitkan', 'VERB'), ('.', 'PUNCT'), ('proses', 'NOUN'), ('terus', 'ADV'),
('berjalan', 'VERB'), ('pungkasnya', 'PUNCT'), ('.', 'PUNCT'))]
```

3.9 NER Analysis

```
def nerAnalyzer(modelName):
    # Load model NER tagging
    model=malaya.entity.transformer(model=modelName)
    return model

def nerAnalysis(nerAnalyzer,text):
    # Prediksi NER
    result=nerAnalyzer.predict(text)
    return result

modelName="xlnet"
nerAnalyzer=nerAnalyzer(modelName)
result=nerAnalysis(nerAnalyzer,konten_clean_from_stopwords)
print(result)
```

```
[('Awardee', 'OTHER'), ('Bongkar', 'OTHER'), ('3', 'OTHER'), ('Rahasia',
'OTHER'), ('Beasiswa', 'OTHER'), ('LPDP', 'OTHER'), ('Intip', 'OTHER'), ('Yuk',
'person'), ('Nikita', 'person'), ('Rosa', 'person'), ('detikEdu', 'person'),
('Senin', 'time'), ('30', 'time'), ('Okt', 'time'), ('2023', 'time'), ('2100',
'time'), ('WIB', 'time'), ('Jakarta', 'location'), ('Margaret', 'OTHER'),
('Ayu', 'person'), ('Caroline', 'OTHER'), ('bertolak', 'OTHER'), ('Inggris',
'location'), ('melanjutkan', 'OTHER'), ('studi', 'OTHER'), ('University',
'OTHER'), ('College', 'OTHER'), ('London', 'location'), ('UCL', 'OTHER'), ('.',
'OTHER'), ('Alumnus', 'person'), ('Fakultas', 'OTHER'), ('Kedokteran', 'OTHER'),
('Universitas', 'OTHER'), ('Airlangga', 'OTHER'), ('berangkat', 'OTHER'),
('London', 'location'), ('beasiswa', 'OTHER'), ('Lembaga', 'OTHER'),
('Pengelola', 'OTHER'), ('Dana', 'OTHER'), ('Pendidikan', 'OTHER'), ('LPDP',
'OTHER'), ('Kementerian', 'OTHER'), ('Keuangan', 'OTHER'), ('.', 'OTHER'),
('Oline', 'person'), ('panggilan', 'OTHER'), ('akrabnya', 'OTHER'),
('menceritakan', 'OTHER'), ('keinginan', 'OTHER'), ('melanjutkan', 'OTHER'),
('kuliah', 'OTHER'), ('berasal', 'OTHER'), ('cita', 'OTHER'), ('citanya',
'OTHER'), ('seorang', 'OTHER'), ('dosen', 'OTHER'), ('klinis', 'OTHER'), ('.',
'OTHER'), ('Baca', 'OTHER'), ('3', 'OTHER'), ('Tips', 'OTHER'), ('Persiapan',
'OTHER'), ('Daftar', 'OTHER'), ('S2', 'OTHER'), ('Luar', 'OTHER'), ('Negeri',
'OTHER'), ('Pekerja', 'OTHER'), ('Catat', 'OTHER'), ('NihBaca', 'OTHER'),
```

('Sosok', 'OTHER'), ('Dhani', 'OTHER'), ('Dokter', 'OTHER'), ('Gigi', 'OTHER'),
 ('Disabilitas', 'OTHER'), ('Peraih', 'OTHER'), ('Beasiswa', 'OTHER'), ('LPDP',
 'OTHER'), ('Jerman', 'OTHER'), ('Aku', 'OTHER'), ('punya', 'OTHER'), ('cita',
 'OTHER'), ('cita', 'OTHER'), ('jadi', 'OTHER'), ('dosen', 'OTHER'), ('klinik',
 'OTHER'), ('aku', 'OTHER'), ('tertarik', 'OTHER'), ('menjadi', 'OTHER'),
 ('seorang', 'OTHER'), ('pengajar', 'OTHER'), ('katanya', 'OTHER'), ('laman',
 'OTHER'), ('Unair', 'OTHER'), ('dikutip', 'OTHER'), ('Senin', 'time'),
 ('30102023', 'time'), ('.', 'OTHER'), ('Ketertarikannya', 'OTHER'), ('bidang',
 'OTHER'), ('ortopedi', 'OTHER'), ('membuat', 'OTHER'), ('Oline', 'OTHER'),
 ('memilih', 'OTHER'), ('Master', 'OTHER'), ('of', 'OTHER'), ('Science',
 'OTHER'), ('MSc', 'OTHER'), ('Musculoskeletal', 'OTHER'), ('Science', 'OTHER'),
 ('and', 'OTHER'), ('Medical', 'OTHER'), ('Engineering', 'OTHER'), ('jurusannya',
 'OTHER'), ('.', 'OTHER'), ('memilih', 'OTHER'), ('UCL', 'OTHER'), ('bermitra',
 'OTHER'), ('salah', 'OTHER'), ('satu', 'OTHER'), ('rumah', 'OTHER'), ('sakit',
 'OTHER'), ('ortopedi', 'OTHER'), ('terbaik', 'OTHER'), ('dunia', 'OTHER'), ('.',
 'OTHER'), ('mahasiswa', 'OTHER'), ('berminat', 'OTHER'), ('beasiswa', 'OTHER'),
 ('LPDP', 'OTHER'), ('Oline', 'OTHER'), ('membongkar', 'OTHER'), ('rahasianya',
 'OTHER'), ('.', 'OTHER'), ('Yuk', 'OTHER'), ('simak', 'OTHER'), ('berikut',
 'OTHER'), ('ini', 'OTHER'), ('.', 'OTHER'), ('Rahasia', 'OTHER'), ('Beasiswa',
 'OTHER'), ('LPDP', 'OTHER'), ('1', 'OTHER'), ('.', 'OTHER'), ('Tentukan',
 'OTHER'), ('Alasan', 'OTHER'), ('.', 'OTHER'), ('Oline', 'OTHER'), ('segala',
 'OTHER'), ('proses', 'OTHER'), ('berjalan', 'OTHER'), ('baik', 'OTHER'),
 ('bila', 'OTHER'), ('motivasi', 'OTHER'), ('kuat', 'OTHER'), ('.', 'OTHER'),
 ('Bahkan', 'OTHER'), ('Oline', 'organization'), ('membutuhkan', 'OTHER'),
 ('waktu', 'OTHER'), ('cukup', 'OTHER'), ('lama', 'OTHER'), ('menentukan',
 'OTHER'), ('alasan', 'OTHER'), ('kuat', 'OTHER'), ('miliknya', 'OTHER'), ('.',
 'OTHER'), ('memutuskan', 'OTHER'), ('mendaftar', 'OTHER'), ('kuliah', 'OTHER'),
 ('aku', 'OTHER'), ('menemukan', 'OTHER'), ('alasan', 'OTHER'), ('kuat',
 'OTHER'), ('melanjutkan', 'OTHER'), ('studi', 'OTHER'), ('berlangsung',
 'OTHER'), ('cukup', 'OTHER'), ('lama', 'OTHER'), ('.', 'OTHER'), ('Kalau',
 'OTHER'), ('alasan', 'OTHER'), ('kuat', 'OTHER'), ('aku', 'OTHER'),
 ('kesulitan', 'OTHER'), ('menjalankan', 'OTHER'), ('segala', 'OTHER'),
 ('proses', 'OTHER'), ('jelasnya', 'OTHER'), ('.', 'OTHER'), ('2', 'OTHER'),
 ('.', 'OTHER'), ('Punya', 'OTHER'), ('Mentor', 'OTHER'), ('Memiliki', 'OTHER'),
 ('seorang', 'OTHER'), ('mentor', 'OTHER'), ('menjadi', 'OTHER'), ('tak',
 'OTHER'), ('kalah', 'OTHER'), ('penting', 'OTHER'), ('.', 'OTHER'), ('mentor',
 'OTHER'), ('mendampingi', 'OTHER'), ('memberi', 'OTHER'), ('arahan', 'OTHER'),
 ('kamu', 'OTHER'), ('mengalami', 'OTHER'), ('kesulitan', 'OTHER'), ('.',
 'OTHER'), ('Mencari', 'OTHER'), ('mentor', 'OTHER'), ('artinya', 'OTHER'),
 ('meminta', 'OTHER'), ('bantuan', 'OTHER'), ('bimbingan', 'OTHER'),
 ('pendahulu', 'OTHER'), ('sukses', 'OTHER'), ('menimba', 'OTHER'),
 ('pendidikan', 'OTHER'), ('luar', 'OTHER'), ('negeri', 'OTHER'), ('.', 'OTHER'),
 ('Sehebat', 'OTHER'), ('sepintar', 'OTHER'), ('apapun', 'OTHER'), ('melakukan',
 'OTHER'), ('prosesnya', 'OTHER'), ('sendiri', 'OTHER'), ('papar', 'OTHER'),
 ('Oline', 'person'), ('.', 'OTHER'), ('3', 'OTHER'), ('.', 'OTHER'), ('Cari',
 'OTHER'), ('Teman', 'OTHER'), ('Seperjuangan', 'OTHER'), ('.', 'OTHER'),
 ('Oline', 'person'), ('melanjutkan', 'OTHER'), ('pendidikan', 'OTHER'),
 ('membuat', 'OTHER'), ('seseorang', 'OTHER'), ('merasa', 'OTHER'), ('kesepian',

```
'OTHER'), ('.', 'OTHER'), ('Teman', 'OTHER'), ('seperjuangan', 'OTHER'),
('memberikan', 'OTHER'), ('dorongan', 'OTHER'), ('motivasi', 'OTHER'),
('tempat', 'OTHER'), ('berkeluh', 'OTHER'), ('kesah', 'OTHER'), ('posisi',
'OTHER'), ('sama', 'OTHER'), ('.', 'OTHER'), ('Terlebih', 'OTHER'), ('proses',
'OTHER'), ('dijalani', 'OTHER'), ('sebentar', 'OTHER'), ('.', 'OTHER'),
('Mulai', 'OTHER'), ('mendaftar', 'OTHER'), ('kampus', 'OTHER'), ('hingga',
'OTHER'), ('beasiswa', 'OTHER'), ('.', 'OTHER'), ('namanya', 'OTHER'),
('menyiapkan', 'OTHER'), ('persyaratan', 'OTHER'), ('masuk', 'OTHER'),
('kampus', 'OTHER'), ('hingga', 'OTHER'), ('beasiswa', 'OTHER'), ('satu',
'OTHER'), ('bulan', 'OTHER'), ('.', 'OTHER'), ('timeline', 'OTHER'),
('tersusun', 'OTHER'), ('ujarnya', 'OTHER'), ('.', 'OTHER'), ('Baca', 'OTHER'),
('Kemenag', 'OTHER'), ('Buka', 'OTHER'), ('Pendaftaran', 'OTHER'), ('Beasiswa',
'OTHER'), ('Penyelesaian', 'OTHER'), ('Studi', 'OTHER'), ('S2', 'OTHER'), ('S3',
'OTHER'), ('Bernilai', 'OTHER'), ('Rp', 'quantity'), ('25', 'quantity'),
('JutaOline', 'quantity'), ('berpesan', 'OTHER'), ('menjalani', 'OTHER'),
('segala', 'OTHER'), ('proses', 'OTHER'), ('seseorang', 'OTHER'),
('menikmatinya', 'OTHER'), ('.', 'OTHER'), ('Meski', 'OTHER'), ('terasa',
'OTHER'), ('sakit', 'OTHER'), ('proses', 'OTHER'), ('terus', 'OTHER'),
('berjalan', 'OTHER'), ('.', 'OTHER'), ('Semakin', 'OTHER'), ('berkembang',
'OTHER'), ('rasa', 'OTHER'), ('sakit', 'OTHER'), ('terasa', 'OTHER'),
('semakin', 'OTHER'), ('besar', 'OTHER'), ('.', 'OTHER'), ('Memang', 'OTHER'),
('prosesnya', 'OTHER'), ('terasa', 'OTHER'), ('enak', 'OTHER'), ('rasanya',
'OTHER'), ('menyakitkan', 'OTHER'), ('.', 'OTHER'), ('proses', 'OTHER'),
('terus', 'OTHER'), ('berjalan', 'OTHER'), ('pungkasnya', 'OTHER'), ('.',
'OTHER'))]
```

3.10 WordCloud Analysis

```
import numpy as np
from PIL import Image
from wordcloud import WordCloud
import matplotlib.pyplot as plt
from nltk.corpus import stopwords

wc = WordCloud(background_color='white', colormap = 'gist_rainbow',
               stopwords = ['meta'], width = 800, height = 500).
    generate(konten_clean_from_stopwords)
plt.axis("off")
plt.imshow(wc)
```

```
<matplotlib.image.AxesImage at 0x1eac00f97d0>
```


