Muhammad Adam Zaky Jiddyansah

Pencarian dan Penambangan Web

To blah, blah, and blah.

Table of contents

Selamat Datang di Jupyter Book Saya

Selamat Datang di Quarto Book Saya Website ini menyajikan beragam konten terkait Penambangan Data atau Data Mining. Mulai dari tugas, materi, hingga artikel-artikel menarik yang mencakup dua tema utama: Pencarian Penambangan informasi melalui Web dan teknik-teknik penambangan data yang relevan. Jelajahi menu di samping untuk menemukan lebih banyak wawasan dan praktik terkait, serta kunjungi Website Personal(https://adamzakys.github.io/) untuk mengetahui biodata saya secara lebih rinci. Selamat menikmati pembelajaran dan penjelajahan!

1

tokenisasi ulang

##Import dan Install modul yang diperlukan

```
!pip install nltk
!pip install Sastrawi
!pip install gensim
import pandas as pd
import re
import nltk
import numpy as np
from gensim.models import Word2Vec
from gensim.models import LdaModel
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.tokenize import RegexpTokenizer
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer,CountVectorizer
from sklearn.preprocessing import OneHotEncoder
from gensim import corpora, models
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.3.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.1)
Collecting Sastrawi
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (209 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 209.7/209.7 kB 2.9 MB/s eta 0:00:00
Installing collected packages: Sastrawi
Successfully installed Sastrawi-1.0.1
Requirement already satisfied: gensim in /usr/local/lib/python3.10/dist-packages (4.3.2)
Requirement already satisfied: numpy>=1.18.5 in /usr/local/lib/python3.10/dist-packages (from gensim)
Requirement already satisfied: scipy>=1.7.0 in /usr/local/lib/python3.10/dist-packages (from gensim)
Requirement already satisfied: smart-open>=1.8.1 in /usr/local/lib/python3.10/dist-packages (from ge
```

##Konfigurasi nltk

```
nltk.download("punkt")
nltk.download("stopwords")
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
```

```
True
```

##mengambil data pta yang sudah dicrawl di program sebelumnya

```
from google.colab import drive
drive.mount('/content/drive')

file_path = "/content/drive/My Drive/Penambangan WEB/Data/full_210411100234_data_pgsd.csv"

df = pd.read_csv(file_path)
df
```

Mengubah teks menjadi lower

```
df=df.astype(str)
df["Abstrak"] = df["Abstrak"].apply(lambda x: x.lower())

abstrak_column = df["Abstrak"]
```

##Tokenisasi menggunakan word tokenize pemisahan kata dalam suatu kalimat dengan tujuan untuk proses analisis teks lebih lanjut

```
def process_tokenize(text):
    text = text.split()
    return text

processed_abstrak = abstrak_column.apply(process_tokenize)

df["processed_abstrak"] = processed_abstrak

data = pd.DataFrame(df, columns=['processed_abstrak'])
data
```

##Punctuation Menghiilangkan tanda baca

```
def process_punctuation(tokens):
    cleaned_tokens = [re.sub(r'[.,():-]', '', token) for token in tokens]
    cleaned_tokens = [re.sub(r'\d+', '', token) for token in cleaned_tokens]

    return cleaned_tokens

df['processed_abstrak'] = df['processed_abstrak'].apply(process_punctuation)

data = pd.DataFrame(df, columns=['processed_abstrak'])
data
```

##mensaring kata2 dengan Stopword

```
def process_stopword_token(tokens):
    stop_words = set(stopwords.words("indonesian"))
    filtered_tokens = [token for token in tokens if token.lower() not in stop_words]
    return " ".join(filtered_tokens)

df['processed_abstrak'] = df['processed_abstrak'].apply(process_stopword_token)

data = pd.DataFrame(df, columns=['processed_abstrak'])
```

##Proses menghilangkan imbuhan baik yang berada di awal kata ataupun yang berada di akhir(Steeming) contoh : kata “running” dan “runner” dapat di-stem menjadi “run”.

##Feature Extraction melakukan perhitungan dan perbandingan yang bisa digunakan untuk mengklasifikasikan ciri-ciri yang dimiliki

mengonversi koleksi dokumen teks menjadi representasi numerik

```
countvectorizer = CountVectorizer(analyzer= 'word', stop_words='english')
count_wm = countvectorizer.fit_transform(df['processed_abstrak'])
count_tokens = countvectorizer.get_feature_names_out()
df_countvect = pd.DataFrame(data = count_wm.toarray(),columns = count_tokens)
print('Count Vectorizer\n')
df_countvect
```

mengonversi teks menjadi representasi numerik yang lebih informatif

##SKIP GRAM Word2Vec

```
#Menambahkan kode untuk melatih model Word2Vec (Skip-gram)
```

```
sentences = df['processed_abstrak'].apply(lambda x: x.split()).tolist()
model = Word2Vec(sentences, vector_size=100, window=5, min_count=1, sg=1)
```

```
#contoh penggunaan skip gram pada dataset abstrak yang sudah dinormalisasi
print("Model Word2Vec (Skip-gram) berhasil dilatih.")
print("Contoh penggunaan:")
word_vector = model.wv['akurat']
print("Vektor kata 'kata_contoh':", word_vector)
similar_words = model.wv.most_similar('akurat')
print("Kata-kata yang mirip dengan 'akurat':", similar_words)
```

##LDA

```
documents = df['processed_abstrak'].apply(lambda x: x.split())

# Membuat representasi teks dalam bentuk "bag of words" /  konsep yang diambil dari analisis te
dictionary = corpora.Dictionary(documents)
corpus = [dictionary.doc2bow(doc) for doc in documents]

# create and training model LDA
lda_model = LdaModel(corpus, num_topics=3, id2word=dictionary, passes=15)

# Menghitung proporsi kata dalam setiap topic
topic_word_proposals = lda_model.get_topics()

# Membuat DataFrame untuk menyimpan proporsi kata dalam topik
topic_word_proposals_df = pd.DataFrame(topic_word_proposals, columns=[dictionary[i] for i in ra

# Menyimpan DataFrame proporsi kata dalam topik ke dalam file CSV
topic_word_proposals_df.to_csv("proporsi_kata_dalam_topik.csv", index=False)

# Menghitung proporsi topik dalam dokumen
document_topic_proposals = [lda_model.get_document_topics(doc) for doc in corpus]

# Membuat DataFrame untuk menyimpan proporsi topik dalam dokumen
document_topic_proposals_df = pd.DataFrame(columns=["Judul"] + [f"Topic {i+1}" for i in range(l

for i, doc_topic_proposals in enumerate(document_topic_proposals):
    row_data = {"Judul": df['Judul'].iloc[i]}  # Menggunakan kolom "Judul" dari DataFrame awal
    for topic, prop in doc_topic_proposals:
        row_data[f"Topic {topic + 1}"] = prop
    # Menambahkan baris data ke DataFrame
    document_topic_proposals_df = pd.concat([document_topic_proposals_df, pd.DataFrame([row_dat
```

```
# ubah nilai NaN dengan 0
document_topic_proposals_df = document_topic_proposals_df.fillna(0)

document_topic_proposals_df
```

##TF-IDF

```
# Melakukan pra-pemrosesan teks untuk TF-IDF
documents_tfidf = df['processed_abstrak']  # Menggunakan processed_abstrak2 yang sudah diproses

# Melatih model TF-IDF Vectorizer
tfidfvectorizer = TfidfVectorizer(analyzer='word', stop_words='english')
tfidf_wm = tfidfvectorizer.fit_transform(documents_tfidf)
tfidf_tokens = tfidfvectorizer.get_feature_names_out()

# Mengubah matriks TF-IDF menjadi DataFrame
df_tfidfvect = pd.DataFrame(data=tfidf_wm.toarray(), columns=tfidf_tokens)

# Menambahkan kolom dokumen ke DataFrame TF-IDF
df_tfidfvect.insert(0, 'Judul', df['Judul'])  # Menggunakan kolom 'Judul' sebagai dokumen

print('TF-IDF Vectorizer\n')
df_tfidfvect
```

##Perbandingan silhouette score antara LDA dan TF-IDF setelah Cluster

```
from sklearn.metrics import silhouette_score

# Jumlah kluster yang diinginkan
n_clusters = 3

# Melakukan clustering pada data hasil LDA
kmeans_lda = KMeans(n_clusters=n_clusters, random_state=0, n_init=10)
kmeans_lda.fit(document_topic_proposals_df.iloc[:, 1:])  # Menggunakan proporsi topik sebagai f

# Melakukan clustering pada data TF-IDF
kmeans_tfidf = KMeans(n_clusters=n_clusters, random_state=0, n_init=10)
kmeans_tfidf.fit(tfidf_wm)

# Menghitung Silhouette Score untuk LDA
silhouette_lda = silhouette_score(document_topic_proposals_df.iloc[:, 1:], kmeans_lda.labels_)

# Menghitung Silhouette Score untuk TF-IDF
```

```
silhouette_tfidf = silhouette_score(tfidf_wm, kmeans_tfidf.labels_)


# Menambahkan label kluster ke DataFrame hasil LDA
document_topic_proposals_df['Cluster_LDA'] = kmeans_lda.labels_

# Menambahkan label kluster ke DataFrame hasil TF-IDF
df_tfidfvect['Cluster_TFIDF'] = kmeans_tfidf.labels_

# Menampilkan hasil clustering LDA dan TF-IDF
print("Hasil Clustering LDA:")
print(document_topic_proposals_df[['Judul', 'Cluster_LDA']])
print("\nHasil Clustering TF-IDF:")
print(df_tfidfvect[['Judul','Cluster_TFIDF']])
print("============================================================")
# Menampilkan hasil perbandingan
print(f"Silhouette Score untuk LDA: {silhouette_lda}")
print(f"Silhouette Score untuk TF-IDF: {silhouette_tfidf}")
```

#Simpan Hasil Preprocessing Data

```
import pandas as pd

svdf = pd.DataFrame(df)
svdf.to_csv("210411100234_data_pgsd_with_preprocessing.csv",index=False)
document_topic_proposals_df.to_csv("210411100234_ModelingTopic.csv", index=False)
```

```
from google.colab import drive
import shutil
```

```
drive.mount('/content/drive')
```

```
source_path = "210411100234_data_pgsd_with_preprocessing.csv"
destination_path = "/content/drive/My Drive/Penambangan WEB/Data/210411100234_data_pgsd_with_pr

source_path = "210411100234_ModelingTopic.csv"
destination_path = "/content/drive/My Drive/Penambangan WEB/Data/210411100234_ModelingTopic.csv

shutil.copy(source_path, destination_path)
```

2

Crawl Berita Kompas

Impor Modul:

requests: Digunakan untuk membuat permintaan HTTP ke situs web. BeautifulSoup dari bs4: Sebuah pustaka untuk mengekstrak data dari file HTML dan XML.

pandas: Sebuah pustaka untuk manipulasi data dalam analisis data.

Inisialisasi:

Skrip menginisialisasi kamus kosong bernama csv dengan kunci "Judul", "Berita", dan "Kategori" untuk menyimpan data yang diambil. Daftar categories berisi kategori-kategori untuk artikel berita yang akan diambil. Loop Pencarian Berita:

Skrip melakukan iterasi melalui setiap kategori dan nomor halaman untuk membentuk URL halaman indeks berita di situs Kompas. Kemudian, skrip mengirim permintaan HTTP ke URL yang dibentuk dan menguraikan konten HTML menggunakan BeautifulSoup.

Ekstraksi Artikel:

Untuk setiap halaman, skrip mencari semua elemen div dengan kelas "latest–indeks mt2 clearfix" yang berisi informasi tentang artikel. Selanjutnya, skrip mengekstrak tautan artikel dan mengambil konten masing-masing artikel dengan mengirim permintaan HTTP lainnya.

Ekstraksi Data:

Skrip mengekstrak judul dan isi dari setiap artikel menggunakan BeautifulSoup berdasarkan struktur HTML situs web Kompas. Data yang diekstrak kemudian ditambahkan ke daftar yang sesuai dalam kamus csv. Membuat DataFrame:

Setelah mengambil semua artikel, skrip mengonversi data yang dikumpulkan ke dalam DataFrame pandas yang dinamai data.

Menyimpan ke CSV:

DataFrame kemudian disimpan ke file CSV bernama

“Data_Berita_All_Kategori.csv” pada path yang ditentukan (“/content/drive/My Drive/Penambangan WEB/Data/”).

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/dr
```

```
import requests
from bs4 import BeautifulSoup as soup
import pandas as pd

csv = {"Judul": [], "Berita": [], "Kategori": []}

categories = ["bola", "lifestyle", "umkm", "tekno"]

for kategori in categories:
    for i in range(1, 100):
        url = f"https://indeks.kompas.com/?site={kategori}&page={i}"
        print(f"Crawling: {url}")
        client = requests.get(url)
        page_html = client.content
        page_soup = soup(page_html, "html.parser")
        berita = page_soup.findAll("div", {"class": "latest--indeks mt2 clearfix"})

        print(f"Found {len(berita)} articles on this page.")

        for h in berita:
            link_berita = h.select_one('a.article__link')
            if link_berita:
                link_berita = link_berita['href']
                r = requests.get(link_berita)
                page = soup(r.content, "html.parser")
                halaman_isi = page.select_one("div", {"class": "col-bs10-10"})

                judul_berita = halaman_isi.select("h1", {"class": "read__title"})
                judul = judul_berita[0].text

                isi_berita = halaman_isi.findAll('p')
                isi = '\n'.join([p.get_text() for p in isi_berita])

                csv["Judul"].append(judul)
                csv["Berita"].append(isi)
                csv["Kategori"].append(kategori)
```

```
data = pd.DataFrame(csv)
csv_path_drive = '/content/drive/My Drive/Penambangan WEB/Data/Data_Berita_All_Kategori.csv'
data.to_csv(csv_path_drive, index=False)
```

```
Crawling: https://indeks.kompas.com/?site=bola&page=1
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=2
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=3
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=4
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=5
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=6
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=7
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=8
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=9
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=10
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=11
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=12
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=13
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=14
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=15
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=16
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=17
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=18
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=19
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=20
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=bola&page=21
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=22
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=23
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=24
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=25
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=26
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=27
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=28
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=29
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=30
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=31
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=32
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=33
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=34
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=35
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=36
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=37
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=38
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=39
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=40
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=41
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=42
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=43
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=44
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=45
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=46
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=47
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=48
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=49
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=50
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=51
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=52
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=53
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=54
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=55
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=56
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=57
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=58
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=59
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=60
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=61
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=62
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=63
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=64
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=65
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=bola&page=66
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=67
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=68
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=69
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=70
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=71
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=72
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=73
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=74
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=75
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=76
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=77
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=78
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=79
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=80
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=81
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=82
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=83
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=84
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=85
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=86
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=87
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=88
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=89
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=90
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=91
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=92
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=93
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=94
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=95
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=96
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=97
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=98
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=bola&page=99
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=1
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=2
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=3
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=4
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=5
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=6
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=7
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=8
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=9
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=10
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=11
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=lifestyle&page=12
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=13
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=14
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=15
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=16
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=17
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=18
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=19
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=20
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=21
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=22
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=23
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=24
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=25
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=26
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=27
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=28
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=29
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=30
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=31
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=32
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=33
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=34
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=35
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=36
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=37
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=38
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=39
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=40
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=41
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=42
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=43
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=44
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=45
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=46
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=47
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=48
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=49
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=50
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=51
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=52
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=53
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=54
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=55
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=56
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=lifestyle&page=57
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=58
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=59
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=60
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=61
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=62
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=63
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=64
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=65
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=66
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=67
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=68
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=69
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=70
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=71
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=72
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=73
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=74
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=75
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=76
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=77
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=78
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=79
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=80
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=81
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=82
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=83
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=84
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=85
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=86
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=87
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=88
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=89
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=90
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=91
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=92
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=93
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=94
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=95
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=96
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=97
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=98
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=lifestyle&page=99
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=1
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=2
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=umkm&page=3
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=4
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=5
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=6
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=7
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=8
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=9
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=10
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=11
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=12
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=13
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=14
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=15
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=16
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=17
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=18
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=19
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=20
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=21
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=22
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=23
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=24
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=25
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=26
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=27
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=28
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=29
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=30
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=31
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=32
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=33
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=34
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=35
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=36
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=37
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=38
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=39
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=40
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=41
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=42
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=43
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=44
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=45
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=46
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=47
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=umkm&page=48
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=49
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=50
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=51
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=52
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=53
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=54
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=55
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=56
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=57
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=58
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=59
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=60
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=61
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=62
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=63
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=64
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=65
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=66
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=67
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=68
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=69
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=70
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=71
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=72
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=73
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=74
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=75
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=76
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=77
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=78
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=79
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=80
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=81
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=82
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=83
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=84
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=85
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=86
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=87
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=88
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=89
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=90
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=91
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=92
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=umkm&page=93
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=94
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=95
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=96
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=97
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=98
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=umkm&page=99
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=1
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=2
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=3
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=4
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=5
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=6
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=7
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=8
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=9
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=10
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=11
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=12
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=13
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=14
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=15
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=16
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=17
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=18
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=19
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=20
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=21
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=22
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=23
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=24
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=25
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=26
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=27
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=28
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=29
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=30
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=31
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=32
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=33
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=34
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=35
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=36
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=37
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=38
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=tekno&page=39
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=40
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=41
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=42
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=43
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=44
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=45
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=46
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=47
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=48
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=49
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=50
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=51
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=52
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=53
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=54
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=55
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=56
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=57
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=58
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=59
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=60
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=61
```

```
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=62
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=63
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=64
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=65
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=66
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=67
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=68
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=69
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=70
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=71
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=72
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=73
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=74
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=75
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=76
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=77
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=78
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=79
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=80
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=81
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=82
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=83
Found 1 articles on this page.
```

```
Crawling: https://indeks.kompas.com/?site=tekno&page=84
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=85
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=86
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=87
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=88
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=89
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=90
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=91
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=92
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=93
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=94
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=95
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=96
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=97
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=98
Found 1 articles on this page.
Crawling: https://indeks.kompas.com/?site=tekno&page=99
Found 1 articles on this page.
```

```
drive.mount('/content/drive')
file_path = "/content/drive/My Drive/Penambangan WEB/Data/Data_Berita_All_Kategori.csv"
df = pd.read_csv(file_path)
df
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/dr
```

	Judul	Berita
0	Hasil Persib Vs PSM Makassar 0-0: Gagal Menang...	Hasil Persib Vs PSM Makassar 0-0: Gagal M
1	Manchester City Pantau Bintang Argentina yang ...	Manchester City Pantau Bintang Argentina
2	Sepak Terjang Paris Brunner, Sang Pemain Terba...	Sepak Terjang Paris Brunner, Sang Pemain
3	Musim Gemilang Astra Honda Racing Team, Domina...	Musim Gemilang Astra Honda Racing Team

	Judul	Berita
4	Pembelajaran dari Piala Dunia U17 2023 demi Se...	Pembelajaran dari Piala Dunia U17 2023 de
...	...	...
391	Fitur di Smartwatch Huawei Watch GT 4 Bisa Pan...	Fitur di Smartwatch Huawei Watch GT 4 B
392	Hands-on Gelang Pintar Xiaomi Smart Band 8, Ga...	Hands-on Gelang Pintar Xiaomi Smart Ban
393	Google Rayakan Ulang Tahun Ke-25, Ini Sejarah ...	Google Rayakan Ulang Tahun Ke-25, Ini Se
394	Social Commerce dan E-commerce, Apa Bedanya? B...	Social Commerce dan E-commerce, Apa Be
395	Samsung Sudah Siapkan Galaxy S25, Ini Buktinya	Samsung Sudah Siapkan Galaxy S25, Ini Bu

3

Tugas Crawling_berita_Kompas dan ekstrak kata kunci

- Nama : Muhammad Adam Zaky Jiddyansah
- NIM : 210411100234
- Kelas : Penambangan dan Pencarian Web B

3.1 Menginstall dan mengimport library yang dibutuhkan

Libraries yang Digunakan * BeautifulSoup (bs4): Library untuk melakukan web scraping dan ekstraksi data HTML. * Requests: Library untuk membuat permintaan HTTP ke situs web dan mendapatkan konten halaman. * Pandas (pd): Library untuk mengelola dan menyimpan data dalam format CSV.

```
from bs4 import BeautifulSoup as soup
import requests
import pandas as pd
import nltk
from nltk.tokenize import sent_tokenize
import re
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer,CountVectorizer, TfidfTransformer
from sklearn.metrics.pairwise import cosine_similarity
from nltk.tokenize import word_tokenize
from collections import Counter
import networkx as nx
import matplotlib.pyplot as plt
```

```
nltk.download("punkt")
nltk.download("stopwords")
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
```

```
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
```

```
True
```

3.2 Web Scraping Berita Kompas dan Penyimpanan ke Google Drive

Kode ini menggunakan teknik web scraping untuk mengumpulkan judul dan isi berita dari 200 halaman indeks situs Kompas.com. Berikut adalah penjelasan komponen utama dari kode tersebut:

Proses Web Scraping dan menyimpan data * Iterasi sebanyak 5 (karna yang akan dipreprocessing kali ini cuman 1 berita) halaman indeks(tergantung indeks pada website) di situs Kompas.com. * Untuk setiap halaman, melakukan permintaan HTTP dan mendapatkan konten HTML. * Menggunakan BeautifulSoup untuk mengekstrak daftar berita dari halaman tersebut. * Untuk setiap berita, mengakses halaman individu untuk mengambil judul dan isi berita. * Data judul dan isi berita ditambahkan ke dalam dictionary csv. Penyimpanan Data * Setelah proses scraping selesai, data disimpan dalam file CSV dengan nama “Data_BeritaKompas.csv” terlebih dahulu di environment Colab. * File CSV tersebut disalin ke dalam Google Drive pada path “/content/drive/My Drive/Penambangan WEB/Data/”.

```
from google.colab import drive
drive.mount('/content/drive')

csv = {"Judul": [], "Berita": []}

for i in range(1, 20):
    url = "https://indeks.kompas.com/?page={}".format(i)
    client = requests.get(url)
    page_html = client.content
    page_soup = soup(page_html, "html.parser")
    berita = page_soup.findAll("div", {"class": "latest--indeks mt2 clearfix"})

    for h in berita:
        r = requests.get(h.select_one('a.article__link')['href'])
        page = soup(r.content, "html.parser")
        halaman_isi = page.select_one("div", {"class": "col-bs10-10"})
```

```
        judul_berita = halaman_isi.select("h1", {"class": "read__title"})
        judul = judul_berita[0].text

        isi_berita = halaman_isi.findAll('p')
        isi = '\n'.join([p.get_text() for p in isi_berita])

        csv["Judul"].append(judul)
        csv["Berita"].append(isi)

        data = pd.DataFrame(csv)
        csv_path_drive = '/content/drive/My Drive/Penambangan WEB/Data/Data_BeritaKompas.csv'
        data.to_csv(csv_path_drive, index=False)
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/dr
```

3.3 Import dan Membaca Data CSV

Menggunakan Pandas untuk membaca file CSV yang telah disimpan sebelumnya. File CSV tersebut berisi data judul dan isi berita dari Kompas.com yang telah di-web scrape sebelumnya, lalu menampilkan data yang telah dibaca dari file CSV dalam bentuk DataFrame untuk dapat dianalisis lebih lanjut.

```
from google.colab import drive
drive.mount('/content/drive')
file_path = "/content/drive/My Drive/Penambangan WEB/Data/Data_BeritaKompas.csv"
df = pd.read_csv(file_path)
df
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/dr
```

	Judul	Berita
0	Lirik dan Makna Lagu Oh Indang Oh Apang, Lagu ...	Lirik dan Makna Lagu Oh Indang Oh Apa
1	Berapa Banyak Kalori yang Dibakar dengan Berja...	Berapa Banyak Kalori yang Dibakar denga
2	Pendaftaran Petugas KPPS Pemilu 2024 Dibuka, B...	Pendaftaran Petugas KPPS Pemilu 2024 I
3	Produser Ungkap Tiket Advance Siksa Neraka Sol...	Produser Ungkap Tiket Advance Siksa Ne
4	Peringati Hari Korban 40 Ribu Jiwa, Pj Gubernu...	Peringati Hari Korban 40 Ribu Jiwa, Pj G
5	Lirik dan Chord Lagu Salam Kenal - Vidi Aldiano	Lirik dan Chord Lagu Salam Kenal - Vidi
6	Pikap Tabrak Truk Tronton di Kota Malang, Satu...	Pikap Tabrak Truk Tronton di Kota Mala
7	Perbandingan Yamaha XMAX 250 Tech Max dan Kymc...	Perbandingan Yamaha XMAX 250 Tech M
8	Hasil Mediasi Keluarga Bayi HNM: RS Hermina Po...	Hasil Mediasi Keluarga Bayi HNM: RS He
9	Tanggapi Hasil Survei Litbang ”Kompas”, Mahfud...	Tanggapi Hasil Survei Litbang ”Kompas”,

	Judul	Berita
10	Kepsek SMA 17 Makassar Dimutasi Dampak Aksi De...	Kepsek SMA 17 Makassar Dimutasi Damp
11	Status WhatsApp Bakal Bisa Dibagikan Langsung ...	Status WhatsApp Bakal Bisa Dibagikan L
12	Alat yang Bekerja Berdasarkan Prinsip Tuas	Alat yang Bekerja Berdasarkan Prinsip Tu
13	Pelaku Perusakan Belasan Mobil Dinas di Semara...	Pelaku Perusakan Belasan Mobil Dinas di
14	Rafael Alun Dihukum Bayar Uang Pengganti 18,9 ...	Rafael Alun Dihukum Bayar Uang Pengga
15	Begini Peran Teknik Industri dalam Pemanfaatan...	Begini Peran Teknik Industri dalam Pema
16	Lexus Siapkan RZ550e, RZ Berperforma Tinggi	Lexus Siapkan RZ550e, RZ Berperforma T
17	Pengendara Motor Tewas di Lokasi Kejadian Sete...	Pengendara Motor Tewas di Lokasi Kejadi
18	Newcastle Vs Milan, Kabar Baik untuk Rossoneri...	Newcastle Vs Milan, Kabar Baik untuk Ro

3.4 Mengambil salah satu berita untuk menjadi sampel data yang akan dianalisis

```
berita = df['Berita'].iloc[15]
berita
```

```
'Begini Peran Teknik Industri dalam Pemanfaatan AI di Era Industri 4.0\n\nKOMPAS.com -
Implementasi kecerdasan buatan atau artificial intelligence (AI) dalam industri menjadi salah satu poi
```

3.5 Pra-pemrosesan Teks / Preprocessing

Mengambil teks sebagai input dan melakukan beberapa langkah pra-pemrosesan pada teks dalam bahasa Indonesia. Langkah-langkah tersebut mencakup penghapusan angka, pembersihan karakter khusus, konversi teks ke huruf kecil, penghapusan kata-kata umum (stopwords), dan penggabungan kata-kata yang telah melalui proses sebelumnya. Tujuannya adalah membersihkan dan mempersiapkan teks agar dapat digunakan lebih lanjut dalam analisis atau pemrosesan lanjutan.

```
def preprocessing(text):
    text = re.sub(r'\d+', '', text)
    text = re.sub(r'[^\w\s.]', '', text)
    text = text.lower()

    stop_words = set(stopwords.words('indonesian'))
```

```
    words = text.split()
    filtered_words = [word for word in words if word.lower() not in stop_words]

    preprocessing_text = ' '.join(filtered_words)

    return preprocessing_text
```

```
berita = preprocessing(berita)
print(berita)
```

```
peran teknik industri pemanfaatan ai era industri . kompas.com implementasi kecerdasan buatan artifici
```

##Memisahakan kalimat dengan word tokenize

Tokenisasi kalimat adalah proses memecah sebuah teks menjadi kalimat-kalimat yang lebih kecil atau unit yang disebut token, di mana token dapat berupa kata, frasa, atau karakter. Word tokenization, atau tokenisasi kata, adalah jenis tokenisasi yang fokus pada memecah teks menjadi unit kata.

```
kalimat = nltk.sent_tokenize(berita)
kalimat = [sentence.replace('.', '') for sentence in kalimat]
print(kalimat)
```

```
['peran teknik industri pemanfaatan ai era industri ', 'kompascom implementasi kecerdasan buatan artif
```

```
kata = word_tokenize(berita)
kata = [k.lower() for k in kata if k != '.']
kata = list(set(kata))
print(kata)
```

```
['dihadapi', 'as', 'merangkul', 'pemborosan', 'printing', 'solusi', 'perguruan', 'signifikan', 'konsu
```

```
# Inisialisasi DataFrame
matriks_kata = pd.DataFrame(0, index=kata, columns=kata)
```

```
for sentence in kalimat:
    words = sentence.split()
    for i in range(len(words)-1):
        current_word = words[i]
        next_word = words[i+1]

        # Pastikan kedua kata ada dalam indeks matrikskata sebelum mengupdate nilai
```

```
        if current_word in matriks_kata.index and next_word in matriks_kata.columns:
            matriks_kata.at[current_word, next_word] += 1
```

```
matriks_kata
```

	dihadapi	as	merangkul	pemborosan	printing	solusi	perguruan	signifikan	konsums
dihadapi	0	0	0	0	0	0	0	0	0
as	0	0	0	0	0	0	0	0	0
merangkul	0	0	0	0	0	0	0	0	0
pemborosan	0	0	0	0	0	0	0	0	0
printing	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...
operasional	0	0	0	0	0	0	0	0	0
industrial	0	0	0	0	0	0	0	0	0
panji	0	0	0	0	0	0	0	0	0
menekankan	0	0	0	0	0	0	0	0	0
computer	0	0	0	0	0	0	0	0	0

3.6 Perhitungan Kemiripan Kosinus

Menghitung kemiripan kosinus antara semua pasangan kalimat menggunakan matriks TF-IDF yang telah dihasilkan sebelumnya. Kemudian, hasil kemiripan kosinus tersebut disimpan dalam DataFrame untuk kemudahan analisis.

Rumus Perhitungan Kemiripan Kosinus: * Untuk dua vektor A dan B, kemiripan kosinus dihitung dengan rumus

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

- Di sini, adalah operasi perkalian dot (dot product), dan adalah norma Euclidean. Dalam konteks ini, vektor A dan B adalah vektor representasi TF-IDF dari dua kalimat. Hasilnya adalah skalar yang menunjukkan seberapa mirip dua kalimat tersebut, dengan nilai 1 menunjukkan kemiripan sempurna.

```
cosine = cosine_similarity(matriks_kata, matriks_kata)
```

```
similarity = pd.DataFrame(cosine, columns=matriks_kata.index, index=matriks_kata.index)
similarity
```

	dihadapi	as	merangkul	pemborosan	printing	solusi	perguruan	signifikan	konsum
dihadapi	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
as	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
merangkul	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
pemborosan	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
printing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...
operasional	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
industrial	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
panji	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
menekankan	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
computer	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

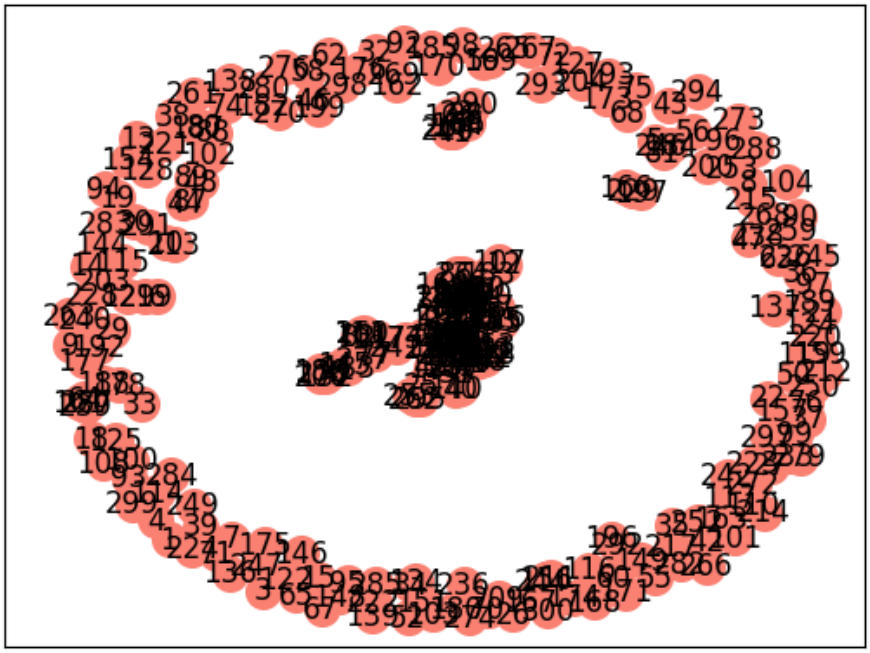
```
Warning: Total number of columns (301) exceeds max_columns (20) limiting to first (20) columns.
```

3.7 Visualisasi Hubungan Antar Kalimat Menggunakan Graf (Graph) berdasarkan Cosine Similarity

Membuat graf yang merepresentasikan hubungan antara kalimat-kalimat berdasarkan nilai cosine similarity. Setiap kalimat direpresentasikan sebagai node, dan garis yang menghubungkan dua node menunjukkan tingkat kemiripan antara kalimat tersebut. Semakin panjang dan tebal garis, semakin tinggi tingkat kemiripan antara dua kalimat. Graf ini membantu memvisualisasikan struktur hubungan antar kalimat dalam teks.

```
G = nx.DiGraph()
for i in range(len(cosine)):
    G.add_node(i)

for i in range(len(cosine)):
    for j in range(len(cosine)):
        similarity = cosine[i][j]
        if similarity > 0.1 and i != j:
            G.add_edge(i, j)

pos = nx.spring_layout(G)
nx.draw_networkx_nodes(G, pos, node_size=200, node_color='salmon')
nx.draw_networkx_edges(G, pos, edge_color='red', arrows=True)
nx.draw_networkx_labels(G, pos)
```

```
plt.show()
```



#Menemukan Kata kunci

##Page Rank **PageRank** adalah algoritma yang dikembangkan oleh Larry Page dan Sergey Brin, pendiri Google. Algoritma ini digunakan oleh mesin pencari Google untuk menentukan peringkat atau relevansi halaman web dalam hasil pencarian. Ide utama di balik PageRank adalah bahwa halaman web yang banyak di-link oleh halaman web lain dianggap lebih penting dan memiliki peringkat yang lebih tinggi.

Konsep Dasar:

Prinsip Dasar: PageRank bekerja berdasarkan prinsip bahwa sebuah tautan dari halaman A ke halaman B dapat dianggap sebagai suara suara atau dukungan dari halaman A untuk halaman B. Semakin banyak tautan yang menuju ke sebuah halaman, semakin besar “nilai suara” atau “nilai dukungan” yang diberikan oleh halaman tersebut.

Tautan Kembali (Backlink): Pentingnya suatu halaman tidak hanya bergantung pada jumlah tautan masuk, tetapi juga pada kualitas tautan tersebut. Tautan dari halaman dengan peringkat tinggi lebih berharga daripada tautan dari halaman dengan peringkat rendah. Rumus dasar PageRank dapat

direpresentasikan sebagai sistem persamaan linear. Misalkan ada N halaman web, dan PR(A) adalah PageRank halaman A Persamaan PageRank untuk halaman A adalah :

$[PR(A) = (1-d)+dleft(fracPR(B)L(B)+fracPR(C)L(C)+ldots+fracPR(N)L(N)right)$

Dengan :

(d) adalah faktor damping (biasanya diatur sekitar 0.85).”

(PR(B), PR(C), \ldots, PR(N)) adalah PageRank dari halaman-halaman yang terhubung ke halaman (A).”

(L(B), L(C), \ldots, L(N))$$ adalah jumlah tautan keluar dari halaman (B, C, \ldots, N).”

```
pagerank = nx.pagerank(G)

sorted_pagerank= sorted(pagerank.items(), key=lambda x: x[1], reverse=True)
print("Page Rank :")
for node, pagerank in sorted_pagerank:
    print(f"Node {node}: {pagerank:.4f}")
```

```
Page Rank :
Node 194: 0.0188
Node 208: 0.0167
Node 251: 0.0163
Node 164: 0.0156
Node 231: 0.0129
Node 79: 0.0126
Node 69: 0.0126
Node 198: 0.0125
Node 232: 0.0112
Node 61: 0.0112
Node 235: 0.0106
Node 254: 0.0103
Node 11: 0.0096
Node 227: 0.0096
Node 0: 0.0091
Node 2: 0.0091
Node 16: 0.0091
Node 57: 0.0091
Node 141: 0.0091
Node 158: 0.0091
Node 210: 0.0091
Node 256: 0.0091
```

```
Node 281: 0.0091
Node 289: 0.0091
Node 91: 0.0090
Node 20: 0.0089
Node 195: 0.0089
Node 246: 0.0084
Node 25: 0.0084
Node 80: 0.0084
Node 229: 0.0084
Node 243: 0.0084
Node 24: 0.0076
Node 54: 0.0076
Node 83: 0.0076
Node 109: 0.0076
Node 113: 0.0076
Node 150: 0.0076
Node 179: 0.0076
Node 275: 0.0076
Node 147: 0.0073
Node 31: 0.0067
Node 277: 0.0067
Node 41: 0.0064
Node 216: 0.0064
Node 264: 0.0064
Node 241: 0.0059
Node 6: 0.0057
Node 10: 0.0057
Node 30: 0.0057
Node 34: 0.0057
Node 44: 0.0057
Node 46: 0.0057
Node 47: 0.0057
Node 48: 0.0057
Node 59: 0.0057
Node 63: 0.0057
Node 64: 0.0057
Node 87: 0.0057
Node 89: 0.0057
Node 134: 0.0057
Node 161: 0.0057
Node 162: 0.0057
Node 169: 0.0057
Node 196: 0.0057
Node 199: 0.0057
Node 211: 0.0057
```

```
Node 213: 0.0057
Node 219: 0.0057
Node 226: 0.0057
Node 238: 0.0057
Node 239: 0.0057
Node 244: 0.0057
Node 269: 0.0057
Node 291: 0.0057
Node 292: 0.0057
Node 81: 0.0056
Node 184: 0.0056
Node 181: 0.0055
Node 28: 0.0055
Node 132: 0.0054
Node 186: 0.0054
Node 206: 0.0054
Node 278: 0.0054
Node 82: 0.0052
Node 111: 0.0052
Node 131: 0.0052
Node 260: 0.0052
Node 207: 0.0045
Node 183: 0.0045
Node 88: 0.0044
Node 166: 0.0044
Node 187: 0.0044
Node 197: 0.0044
Node 133: 0.0043
Node 145: 0.0043
Node 155: 0.0043
Node 258: 0.0043
Node 53: 0.0042
Node 126: 0.0042
Node 135: 0.0042
Node 205: 0.0042
Node 191: 0.0041
Node 237: 0.0041
Node 257: 0.0041
Node 262: 0.0040
Node 295: 0.0040
Node 45: 0.0039
Node 123: 0.0039
Node 12: 0.0037
Node 107: 0.0037
Node 73: 0.0035
```

```
Node 172: 0.0033
Node 201: 0.0033
Node 296: 0.0033
Node 5: 0.0032
Node 142: 0.0032
Node 286: 0.0032
Node 290: 0.0032
Node 40: 0.0031
Node 140: 0.0031
Node 77: 0.0030
Node 105: 0.0029
Node 23: 0.0028
Node 66: 0.0028
Node 106: 0.0028
Node 78: 0.0026
Node 121: 0.0026
Node 152: 0.0026
Node 153: 0.0026
Node 190: 0.0026
Node 248: 0.0026
Node 86: 0.0024
Node 234: 0.0024
Node 84: 0.0024
Node 130: 0.0024
Node 255: 0.0023
Node 85: 0.0022
Node 143: 0.0020
Node 112: 0.0020
Node 225: 0.0018
Node 17: 0.0015
Node 118: 0.0015
Node 165: 0.0015
Node 202: 0.0015
Node 49: 0.0014
Node 120: 0.0014
Node 218: 0.0014
Node 271: 0.0014
Node 160: 0.0014
Node 51: 0.0014
Node 230: 0.0014
Node 1: 0.0009
Node 3: 0.0009
Node 4: 0.0009
Node 7: 0.0009
Node 8: 0.0009
```

```
Node 9: 0.0009
Node 13: 0.0009
Node 14: 0.0009
Node 15: 0.0009
Node 18: 0.0009
Node 19: 0.0009
Node 21: 0.0009
Node 22: 0.0009
Node 26: 0.0009
Node 27: 0.0009
Node 29: 0.0009
Node 32: 0.0009
Node 33: 0.0009
Node 35: 0.0009
Node 36: 0.0009
Node 37: 0.0009
Node 38: 0.0009
Node 39: 0.0009
Node 42: 0.0009
Node 43: 0.0009
Node 50: 0.0009
Node 52: 0.0009
Node 55: 0.0009
Node 56: 0.0009
Node 58: 0.0009
Node 60: 0.0009
Node 62: 0.0009
Node 65: 0.0009
Node 67: 0.0009
Node 68: 0.0009
Node 70: 0.0009
Node 71: 0.0009
Node 72: 0.0009
Node 74: 0.0009
Node 75: 0.0009
Node 76: 0.0009
Node 90: 0.0009
Node 92: 0.0009
Node 93: 0.0009
Node 94: 0.0009
Node 95: 0.0009
Node 96: 0.0009
Node 97: 0.0009
Node 98: 0.0009
Node 99: 0.0009
```

```
Node 100: 0.0009
Node 101: 0.0009
Node 102: 0.0009
Node 103: 0.0009
Node 104: 0.0009
Node 108: 0.0009
Node 110: 0.0009
Node 114: 0.0009
Node 115: 0.0009
Node 116: 0.0009
Node 117: 0.0009
Node 119: 0.0009
Node 122: 0.0009
Node 124: 0.0009
Node 125: 0.0009
Node 127: 0.0009
Node 128: 0.0009
Node 129: 0.0009
Node 136: 0.0009
Node 137: 0.0009
Node 138: 0.0009
Node 139: 0.0009
Node 144: 0.0009
Node 146: 0.0009
Node 148: 0.0009
Node 149: 0.0009
Node 151: 0.0009
Node 154: 0.0009
Node 156: 0.0009
Node 157: 0.0009
Node 159: 0.0009
Node 163: 0.0009
Node 167: 0.0009
Node 168: 0.0009
Node 170: 0.0009
Node 171: 0.0009
Node 173: 0.0009
Node 174: 0.0009
Node 175: 0.0009
Node 176: 0.0009
Node 177: 0.0009
Node 178: 0.0009
Node 180: 0.0009
Node 182: 0.0009
Node 185: 0.0009
```

```
Node 188: 0.0009
Node 189: 0.0009
Node 192: 0.0009
Node 193: 0.0009
Node 200: 0.0009
Node 203: 0.0009
Node 204: 0.0009
Node 209: 0.0009
Node 212: 0.0009
Node 214: 0.0009
Node 215: 0.0009
Node 217: 0.0009
Node 220: 0.0009
Node 221: 0.0009
Node 222: 0.0009
Node 223: 0.0009
Node 224: 0.0009
Node 228: 0.0009
Node 233: 0.0009
Node 236: 0.0009
Node 240: 0.0009
Node 242: 0.0009
Node 245: 0.0009
Node 247: 0.0009
Node 249: 0.0009
Node 250: 0.0009
Node 252: 0.0009
Node 253: 0.0009
Node 259: 0.0009
Node 261: 0.0009
Node 263: 0.0009
Node 265: 0.0009
Node 266: 0.0009
Node 267: 0.0009
Node 268: 0.0009
Node 270: 0.0009
Node 272: 0.0009
Node 273: 0.0009
Node 274: 0.0009
Node 276: 0.0009
Node 279: 0.0009
Node 280: 0.0009
Node 282: 0.0009
Node 283: 0.0009
Node 284: 0.0009
```

```
Node 285: 0.0009
Node 287: 0.0009
Node 288: 0.0009
Node 293: 0.0009
Node 294: 0.0009
Node 297: 0.0009
Node 298: 0.0009
Node 299: 0.0009
Node 300: 0.0009
```

```
print("3 Node Tertinggi Page Rank :")
sentence = ""
for node, pagerank in sorted_pagerank[:3]:
  top_sentence = kata[node]
  sentence += top_sentence + ", "
  print(f"Node {node}: Page Rank = {pagerank:.4f}")
  print(f"Kalimat: {top_sentence}")
```

```
3 Node Tertinggi Page Rank :
Node 194: Page Rank = 0.0188
Kalimat: ai
Node 208: Page Rank = 0.0167
Kalimat: proses
Node 251: Page Rank = 0.0163
Kalimat: industri
```

```
news = df['Berita'].iloc[15]
print('Berita yang digunakan : ')
news
```

```
Berita yang digunakan :

'Begini Peran Teknik Industri dalam Pemanfaatan AI di Era Industri 4.0\n\nKOMPAS.com -
Implementasi kecerdasan buatan atau artificial intelligence (AI) dalam industri menjadi salah satu poi
```

```
print('Kata Kunci :', sentence)
```

```
Kata Kunci : ai, proses, industri,
```

4

Meringkas Berita

Nama : Muhammad Adam Zaky Jiddyansah

Kelas : Pencarian dan Penambangan Web

5

Menginstall dan importing modul / libray yang akan digunakan

```
!pip install seaborn matplotlib
import pandas as pd
import nltk
from nltk.tokenize import sent_tokenize
import re
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer,CountVectorizer, TfidfTransformer
from sklearn.metrics.pairwise import cosine_similarity
from nltk.tokenize import word_tokenize
from collections import Counter
import networkx as nx
import matplotlib.pyplot as plt
import seaborn as sns
```

```
Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-packages (0.12.2)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: numpy!=1.24.0,>=1.17 in /usr/local/lib/python3.10/dist-packages (from
Requirement already satisfied: pandas>=0.25 in /usr/local/lib/python3.10/dist-packages (from seaborn)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matp
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotl
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from mat
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from mat
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matpl
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplot
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matp
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateu
```

```
nltk.download("punkt")
nltk.download("stopwords")
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
```

```
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!

True
```

6

Membaca Data Berita

```
from google.colab import drive
drive.mount('/content/drive')
file_path = "/content/drive/My Drive/Penambangan WEB/Data/Data_BeritaKompas.csv"
df = pd.read_csv(file_path)
df
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/dr
```

	Judul	Berita
0	Lirik dan Makna Lagu Oh Indang Oh Apang, Lagu ...	Lirik dan Makna Lagu Oh Indang Oh Apa
1	Berapa Banyak Kalori yang Dibakar dengan Berja...	Berapa Banyak Kalori yang Dibakar denga
2	Pendaftaran Petugas KPPS Pemilu 2024 Dibuka, B...	Pendaftaran Petugas KPPS Pemilu 2024 I
3	Produser Ungkap Tiket Advance Siksa Neraka Sol...	Produser Ungkap Tiket Advance Siksa Ne
4	Peringati Hari Korban 40 Ribu Jiwa, Pj Gubernu...	Peringati Hari Korban 40 Ribu Jiwa, Pj G
5	Lirik dan Chord Lagu Salam Kenal - Vidi Aldiano	Lirik dan Chord Lagu Salam Kenal - Vidi
6	Pikap Tabrak Truk Tronton di Kota Malang, Satu...	Pikap Tabrak Truk Tronton di Kota Mala
7	Perbandingan Yamaha XMAX 250 Tech Max dan Kymc...	Perbandingan Yamaha XMAX 250 Tech M
8	Hasil Mediasi Keluarga Bayi HNM: RS Hermina Po...	Hasil Mediasi Keluarga Bayi HNM: RS He
9	Tanggapi Hasil Survei Litbang ”Kompas”, Mahfud...	Tanggapi Hasil Survei Litbang ”Kompas”,
10	Kepsek SMA 17 Makassar Dimutasi Dampak Aksi De...	Kepsek SMA 17 Makassar Dimutasi Damp
11	Status WhatsApp Bakal Bisa Dibagikan Langsung ...	Status WhatsApp Bakal Bisa Dibagikan L
12	Alat yang Bekerja Berdasarkan Prinsip Tuas	Alat yang Bekerja Berdasarkan Prinsip Tu
13	Pelaku Perusakan Belasan Mobil Dinas di Semara...	Pelaku Perusakan Belasan Mobil Dinas di
14	Rafael Alun Dihukum Bayar Uang Pengganti 18,9 ...	Rafael Alun Dihukum Bayar Uang Pengga
15	Begini Peran Teknik Industri dalam Pemanfaatan...	Begini Peran Teknik Industri dalam Pema
16	Lexus Siapkan RZ550e, RZ Berperforma Tinggi	Lexus Siapkan RZ550e, RZ Berperforma T
17	Pengendara Motor Tewas di Lokasi Kejadian Sete...	Pengendara Motor Tewas di Lokasi Kejadi
18	Newcastle Vs Milan, Kabar Baik untuk Rossoneri...	Newcastle Vs Milan, Kabar Baik untuk Ro

6.1 Mengambil salah satu *berita*

```
berita = df['Berita'].iloc[15]
berita
```

```
'Begini Peran Teknik Industri dalam Pemanfaatan AI di Era Industri 4.0\n\nKOMPAS.com -
Implementasi kecerdasan buatan atau artificial intelligence (AI) dalam industri menjadi salah satu poi
```

7

Pra-pemrosesan Teks / Preprocessing

Mengambil teks sebagai input dan melakukan beberapa langkah pra-pemrosesan pada teks dalam bahasa Indonesia. Langkah-langkah tersebut mencakup penghapusan angka, pembersihan karakter khusus, konversi teks ke huruf kecil, penghapusan kata-kata umum (stopwords), dan penggabungan kata-kata yang telah melalui proses sebelumnya. Tujuannya adalah membersihkan dan mempersiapkan teks agar dapat digunakan lebih lanjut dalam analisis atau pemrosesan lanjutan.

```
def preprocessing(text):
    text = re.sub(r'\d+', '', text)
    text = re.sub(r'[^\w\s.]', '', text)
    text = text.lower()

    stop_words = set(stopwords.words('indonesian'))
    words = text.split()
    filtered_words = [word for word in words if word.lower() not in stop_words]

    preprocessing_text = ' '.join(filtered_words)

    return preprocessing_text
```

```
berita = preprocessing(berita)
print(berita)
```

```
peran teknik industri pemanfaatan ai era industri . kompas.com implementasi kecerdasan buatan artifici
```

```
kalimat = nltk.sent_tokenize(berita)
kalimat = [sentence.replace('.', '') for sentence in kalimat]
print(kalimat)
```

```
['peran teknik industri pemanfaatan ai era industri ', 'kompascom implementasi kecerdasan buatan artif
```

8

Ekstraksi kalimat

Ekstraksi kalimat adalah suatu teknik atau proses yang bertujuan untuk mengidentifikasi dan mengekstrak kalimat-kalimat penting atau relevan dari sebuah dokumen teks. Tujuan utamanya adalah untuk merangkum informasi yang terkandung dalam teks tanpa menghilangkan inti dari pesan yang ingin disampaikan. Metode ekstraksi kalimat berfokus pada pemilihan kalimat-kalimat yang dianggap memiliki nilai informasi tinggi atau representatif dari keseluruhan teks.

8.1 Metode pengukuran TF-IDF

Mengonversi teks yang sudah melewati tahap pra-pemrosesan menjadi representasi matriks numerik menggunakan skema TF-IDF. Matriks ini mencerminkan bobot TF-IDF dari setiap kata dalam setiap kalimat. Setelah transformasi, DataFrame pandas dibuat untuk menyajikan matriks tersebut dengan kata-kata sebagai kolom dan setiap baris mewakili satu kalimat, menunjukkan bobot TF-IDF dari kata-kata tersebut dalam kalimat tersebut.

```
tfidf_vectorizer = TfidfVectorizer()
tfidf_preprocessing = tfidf_vectorizer.fit_transform(kalimat)

terms = tfidf_vectorizer.get_feature_names_out()
tfidf_preprocessing = pd.DataFrame(data=tfidf_preprocessing.toarray(), columns=terms)

tfidf_preprocessing
```

	ai	akses	akurat	alat	all	analisis	aplikasi	ar	artificial	as
0	0.242502	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.142902	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.301414	0.000000
2	0.187425	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.333926	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000

	ai	akses	akurat	alat	all	analisis	aplikasi	ar	artificial	as
6	0.235964	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
7	0.127922	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.269819
8	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.117251	0.000000	0.0000	0.000000	0.00000	0.223311	0.000000	0.000000	0.000000	0.000000
10	0.126915	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.151523	0.000000	0.3196	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
13	0.266678	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
15	0.185576	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
16	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
17	0.150751	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
18	0.119230	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
19	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
20	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
21	0.120177	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
22	0.259671	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
23	0.123091	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
24	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
25	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
26	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
27	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
28	0.000000	0.000000	0.0000	0.000000	0.00000	0.233250	0.000000	0.258319	0.000000	0.000000
29	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
30	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
31	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
32	0.140822	0.297029	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
33	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
34	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
35	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
36	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
37	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
38	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
39	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
40	0.000000	0.000000	0.0000	0.362831	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
41	0.000000	0.000000	0.0000	0.259730	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
42	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.371988	0.000000	0.000000	0.000000
43	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
44	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
45	0.000000	0.000000	0.0000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000
46	0.000000	0.000000	0.0000	0.000000	0.57735	0.000000	0.000000	0.000000	0.000000	0.000000

9

Membuat Graph

Membuat graf yang merepresentasikan hubungan antara kalimat-kalimat berdasarkan nilai cosine similarity. Setiap kalimat direpresentasikan sebagai node, dan garis yang menghubungkan dua node menunjukkan tingkat kemiripan antara kalimat tersebut. Semakin panjang dan tebal garis, semakin tinggi tingkat kemiripan antara dua kalimat. Graf ini membantu memvisualisasikan struktur hubungan antar kalimat dalam teks.

9.1 Perhitungan Kemiripan Kosinus

Menghitung kemiripan kosinus antara semua pasangan kalimat menggunakan matriks TF-IDF yang telah dihasilkan sebelumnya. Kemudian, hasil kemiripan kosinus tersebut disimpan dalam DataFrame untuk kemudahan analisis.

Rumus Perhitungan Kemiripan Kosinus: * Untuk dua vektor A dan B, kemiripan kosinus dihitung dengan rumus

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

- Di sini, adalah operasi perkalian dot (dot product), dan adalah norma Euclidean. Dalam konteks ini, vektor A dan B adalah vektor representasi TF-IDF dari dua kalimat. Hasilnya adalah skalar yang menunjukkan seberapa mirip dua kalimat tersebut, dengan nilai 1 menunjukkan kemiripan sempurna.

```
cosine_preprocessing = cosine_similarity(tfidf_preprocessing, tfidf_preprocessing)
```

```
similarity_preprocessing = pd.DataFrame(cosine_preprocessing, columns=range(len(kalimat)), inde
similarity_preprocessing
```

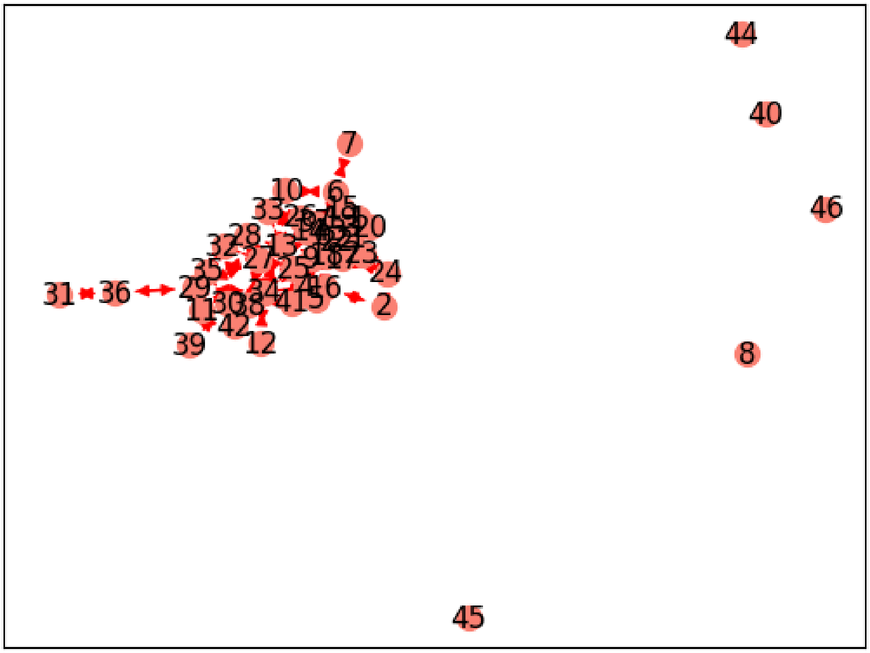
	0	1	2	3	4	5	6	7	8	9
0	1.000000	0.184595	0.045451	0.431353	0.061610	0.000000	0.149652	0.031021	0.0	0.166220
1	0.184595	1.000000	0.026783	0.215648	0.018153	0.000000	0.060953	0.018280	0.0	0.057353

	0	1	2	3	4	5	6	7	8	9
2	0.045451	0.026783	1.000000	0.062586	0.000000	0.000000	0.044226	0.023976	0.0	0.021976
3	0.431353	0.215648	0.062586	1.000000	0.042418	0.000000	0.142433	0.042717	0.0	0.134020
4	0.061610	0.018153	0.000000	0.042418	1.000000	0.119776	0.029974	0.000000	0.0	0.125885
5	0.000000	0.000000	0.000000	0.000000	0.119776	1.000000	0.000000	0.000000	0.0	0.208740
6	0.149652	0.060953	0.044226	0.142433	0.029974	0.000000	1.000000	0.139675	0.0	0.094703
7	0.031021	0.018280	0.023976	0.042717	0.000000	0.000000	0.139675	1.000000	0.0	0.014999
8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.0	0.000000
9	0.166220	0.057353	0.021976	0.134020	0.125885	0.208740	0.094703	0.014999	0.0	1.000000
10	0.030777	0.018136	0.023787	0.042380	0.000000	0.096650	0.138575	0.066486	0.0	0.055231
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
12	0.036745	0.021653	0.028399	0.050598	0.057598	0.000000	0.035754	0.079378	0.0	0.060705
13	0.064670	0.038109	0.049982	0.089051	0.000000	0.000000	0.062926	0.034114	0.0	0.031268
14	0.000000	0.103994	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
15	0.117695	0.298119	0.034782	0.112018	0.023574	0.000000	0.079156	0.023739	0.0	0.074480
16	0.000000	0.000000	0.121047	0.000000	0.131560	0.135592	0.000000	0.000000	0.0	0.041469
17	0.036558	0.079956	0.028255	0.050340	0.000000	0.000000	0.035572	0.019284	0.0	0.017676
18	0.283892	0.111057	0.076355	0.136282	0.174434	0.090798	0.096301	0.015252	0.0	0.194076
19	0.178885	0.129801	0.000000	0.055237	0.026017	0.000000	0.039032	0.000000	0.0	0.058185
20	0.040296	0.118520	0.046599	0.027744	0.071758	0.000000	0.019605	0.000000	0.0	0.029225
21	0.134107	0.093337	0.076962	0.072541	0.052812	0.000000	0.051260	0.015373	0.0	0.076222
22	0.341437	0.067077	0.048669	0.156743	0.032986	0.000000	0.110760	0.033218	0.0	0.104218
23	0.266514	0.093698	0.023070	0.218949	0.069729	0.000000	0.075961	0.015746	0.0	0.165391
24	0.063022	0.018569	0.109380	0.043391	0.070703	0.000000	0.030662	0.000000	0.0	0.083179
25	0.100582	0.029636	0.000000	0.069251	0.032618	0.195544	0.048935	0.000000	0.0	0.154584
26	0.067099	0.086143	0.000000	0.046198	0.021759	0.000000	0.032645	0.000000	0.0	0.048664
27	0.100006	0.029466	0.000000	0.068854	0.032431	0.000000	0.048655	0.000000	0.0	0.072530
28	0.047973	0.014135	0.000000	0.033030	0.062111	0.000000	0.023340	0.000000	0.0	0.173673
29	0.000000	0.000000	0.000000	0.000000	0.041297	0.000000	0.000000	0.000000	0.0	0.000000
30	0.000000	0.000000	0.000000	0.000000	0.087183	0.000000	0.000000	0.000000	0.0	0.000000
31	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
32	0.034150	0.020124	0.026394	0.047024	0.053530	0.000000	0.033229	0.018014	0.0	0.056418
33	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
34	0.094452	0.027829	0.000000	0.065031	0.113342	0.000000	0.045953	0.000000	0.0	0.068502
35	0.042871	0.055039	0.000000	0.029517	0.013903	0.000000	0.020858	0.000000	0.0	0.077640
36	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
37	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
38	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
39	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
40	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
41	0.000000	0.000000	0.000000	0.000000	0.000000	0.138930	0.000000	0.000000	0.0	0.000000
42	0.000000	0.000000	0.079888	0.000000	0.060496	0.000000	0.000000	0.000000	0.0	0.000000
43	0.160423	0.021199	0.000000	0.049536	0.023332	0.000000	0.035004	0.000000	0.0	0.052180
44	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000

	0	1	2	3	4	5	6	7	8	9
45	0.000000	0.062608	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
46	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000

9.2 Graph

```
G_preprocessing = nx.DiGraph()

# Add nodes to the graph
for i in range(len(cosine_preprocessing)):
    G_preprocessing.add_node(i)

# Add edges based on similarity
for i in range(len(cosine_preprocessing)):
    for j in range(len(cosine_preprocessing)):
        similarity = cosine_preprocessing[i][j]
        if similarity > 0.1 and i != j:
            G_preprocessing.add_edge(i, j)

# Specify the layout (spring_layout for example)
pos = nx.spring_layout(G_preprocessing)

# Draw the graph
nx.draw_networkx_nodes(G_preprocessing, pos, node_size=100, node_color='salmon')
nx.draw_networkx_edges(G_preprocessing, pos, edge_color='red', arrows=True)
nx.draw_networkx_labels(G_preprocessing, pos)

plt.show()
```


44
40
7
10
46
31
36
29
39
12
2
24
8
45

10

Presentasikan Ukuran dengan Matriks Sentralitas

10.1 Closeness

Menghitung dan menganalisis Closeness Centrality pada graf G_preprocessing

```
closeness_preprocessing = nx.closeness_centrality(G_preprocessing)

sorted_closeness_preprocessing = sorted(closeness_preprocessing.items(), key=lambda x: x[1], re
print("Closeness Centrality:")
for node, closeness in sorted_closeness_preprocessing:
    print(f"Node {node}: {closeness:.4f}")
```

```
Closeness Centrality:
Node 18: 0.4746
Node 34: 0.4746
Node 0: 0.4626
Node 25: 0.4568
Node 27: 0.4568
Node 22: 0.4153
Node 13: 0.4106
Node 16: 0.4106
Node 21: 0.4060
Node 43: 0.3972
Node 3: 0.3929
Node 17: 0.3888
Node 23: 0.3888
Node 30: 0.3888
Node 4: 0.3847
Node 9: 0.3847
Node 19: 0.3847
Node 1: 0.3807
Node 26: 0.3807
```

```
Node 41: 0.3767
Node 29: 0.3729
Node 37: 0.3691
Node 15: 0.3654
Node 14: 0.3583
Node 5: 0.3548
Node 32: 0.3548
Node 35: 0.3447
Node 20: 0.3384
Node 6: 0.3353
Node 38: 0.3322
Node 33: 0.3292
Node 42: 0.3263
Node 11: 0.3234
Node 24: 0.3206
Node 10: 0.3178
Node 12: 0.3123
Node 28: 0.3097
Node 2: 0.3020
Node 39: 0.2727
Node 36: 0.2687
Node 7: 0.2453
Node 31: 0.2076
Node 8: 0.0000
Node 40: 0.0000
Node 44: 0.0000
Node 45: 0.0000
Node 46: 0.0000
```

11

Hasil Ringkasan Berita Dengan Closeness Centrality

```
ringkasan_closeness_preprocessing = ""
print("Tiga Node Tertinggi Closeness Centrality Menggunakan Preprocessing:")
for node, closeness_preprocessing in sorted_closeness_preprocessing[:3]:
    top_sentence = kalimat[node]
    ringkasan_closeness_preprocessing += top_sentence + " "
    print(f"Node {node}: Closeness Centrality = {closeness_preprocessing:.4f}")
    print(f"Kalimat: {top_sentence}\n")
```

```
Tiga Node Tertinggi Closeness Centrality Menggunakan Preprocessing:
Node 18: Closeness Centrality = 0.4746
Kalimat: teknik industri solusi menghadapi kompleksitas transformasi industri didorong ai ilmu teknik

Node 34: Closeness Centrality = 0.4746
Kalimat: mahasiswa memahami menganalisis proses industri manufaktur menyeluruh

Node 0: Closeness Centrality = 0.4626
Kalimat: peran teknik industri pemanfaatan ai era industri

Ringkasan Berita Menggunakan Closeness Centrality:
teknik industri solusi menghadapi kompleksitas transformasi industri didorong ai ilmu teknik industri
```

```
print("Ringkasan Berita Menggunakan Closeness Centrality:")
print(ringkasan_closeness_preprocessing)
```

```
Ringkasan Berita Menggunakan Closeness Centrality:
teknik industri solusi menghadapi kompleksitas transformasi industri didorong ai ilmu teknik industri
```

11.1 Menghitung peringkat halaman dengan Pagerank

Rumus Perhitungan Manual PageRank:

- PageRank mengukur pentingnya suatu node dalam sebuah graf berdasarkan seberapa banyak node lain yang mengarah ke node tersebut. Rumusnya adalah sebagai berikut:
$$PR(x) = (1 - d) + d \left(\frac{L(y)}{PR(y)} \right)$$

- PR(x) adalah PageRank dari node x.
- d adalah faktor damping, biasanya diatur sebagai 0,85.
- PR(y) adalah PageRank dari node yang memiliki tautan ke node y.
- L(y) adalah jumlah tautan keluar dari node y.
- Iterasi dilakukan hingga konvergensi atau sejumlah iterasi tertentu. Semakin tinggi nilai PageRank, semakin "penting" node tersebut dalam graf.

```
pagerank_preprocessing = nx.pagerank(G_preprocessing)

sorted_pagerank_preprocessing= sorted(pagerank_preprocessing.items(), key=lambda x: x[1], rever
print("Page Rank :")
for node, pagerank_preprocessing in sorted_pagerank_preprocessing:
    print(f"Node {node}: {pagerank_preprocessing:.4f}")
```

```
Page Rank :
Node 34: 0.0512
Node 18: 0.0480
Node 0: 0.0442
Node 21: 0.0377
Node 27: 0.0361
Node 30: 0.0336
Node 25: 0.0334
Node 9: 0.0331
Node 22: 0.0320
Node 29: 0.0309
Node 13: 0.0303
Node 3: 0.0287
Node 43: 0.0276
Node 19: 0.0275
Node 26: 0.0259
Node 16: 0.0252
Node 23: 0.0251
Node 1: 0.0250
Node 6: 0.0236
Node 17: 0.0236
Node 37: 0.0219
Node 20: 0.0218
Node 5: 0.0205
Node 14: 0.0204
Node 4: 0.0202
```

```
Node 35: 0.0195
Node 32: 0.0193
Node 38: 0.0191
Node 15: 0.0185
Node 41: 0.0176
Node 24: 0.0174
Node 36: 0.0170
Node 42: 0.0153
Node 33: 0.0138
Node 11: 0.0115
Node 10: 0.0112
Node 2: 0.0108
Node 31: 0.0107
Node 28: 0.0107
Node 39: 0.0076
Node 7: 0.0075
Node 12: 0.0071
Node 8: 0.0035
Node 40: 0.0035
Node 44: 0.0035
Node 45: 0.0035
Node 46: 0.0035
```

12

Hasil Ringkasan dengan Page Rank

```
ringkasan_pagerank_preprocessing = ""
print("Tiga Node Tertinggi Page Rank Menggunakan Preprocessing:")
for node, pagerank_preprocessing in sorted_pagerank_preprocessing[:3]:
    top_sentence = kalimat[node]
    ringkasan_pagerank_preprocessing += top_sentence + " "
    print(f"Node {node}: Page Rank = {pagerank_preprocessing:.4f}")
    print(f"Kalimat: {top_sentence}\n")
```

```
Tiga Node Tertinggi Page Rank Menggunakan Preprocessing:
Node 34: Page Rank = 0.0512
Kalimat: mahasiswa memahami menganalisis proses industri manufaktur menyeluruh

Node 18: Page Rank = 0.0480
Kalimat: teknik industri solusi menghadapi kompleksitas transformasi industri didorong ai ilmu teknik

Node 0: Page Rank = 0.0442
Kalimat: peran teknik industri pemanfaatan ai era industri
```

```
print("Ringkasan Berita Menggunakan Page Rank:")
print(ringkasan_pagerank_preprocessing)
```

```
Ringkasan Berita Menggunakan Page Rank:
mahasiswa memahami menganalisis proses industri manufaktur menyeluruh teknik industri solusi menghadap
```

12.1 menggunakan metode eigenvector centrality untuk menghitung sentralitas eigenvector pada graf

Rumus Perhitungan Manual Eigenvector Centrality: * Eigenvector Centrality mengukur pentingnya suatu node dalam suatu graf berdasarkan pada seberapa banyak node lain yang terhubung ke node tersebut. Rumusnya adalah

sebagai berikut:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j$$

- xi adalah eigenvector centrality dari node i
- Aij adalah elemen matriks ketetanggaan yang menunjukkan apakah ada tautan antara node i dan j
- adalah nilai eigen (eigenvalue) yang sesuai dengan eigenvector yang dicari.
- Iterasi dilakukan hingga konvergensi atau sejumlah iterasi tertentu. Semakin tinggi nilai eigenvector centrality, semakin "penting" node tersebut dalam graf.

```
eigenvector_preprocessing = nx.eigenvector_centrality(G_preprocessing)

sorted_eigenvector_preprocessing= sorted(eigenvector_preprocessing.items(), key=lambda x: x[1],
print("Eigen Vector :")
for node, eigenvector_preprocessing in sorted_eigenvector_preprocessing:
    print(f"Node {node}: {eigenvector_preprocessing:.4f}")
```

```
Eigen Vector :
Node 0: 0.3588
Node 18: 0.3579
Node 21: 0.2843
Node 22: 0.2628
Node 3: 0.2366
Node 43: 0.2350
Node 9: 0.2343
Node 19: 0.2339
Node 23: 0.2249
Node 25: 0.2062
Node 1: 0.1965
Node 15: 0.1618
Node 34: 0.1588
Node 26: 0.1552
Node 37: 0.1496
Node 20: 0.1494
Node 27: 0.1486
Node 4: 0.1170
Node 6: 0.1109
Node 17: 0.1055
Node 16: 0.1011
Node 5: 0.0917
Node 14: 0.0859
Node 41: 0.0805
```

```
Node 24: 0.0790
Node 13: 0.0679
Node 30: 0.0670
Node 29: 0.0553
Node 38: 0.0408
Node 33: 0.0383
Node 35: 0.0380
Node 32: 0.0357
Node 28: 0.0335
Node 42: 0.0331
Node 11: 0.0266
Node 2: 0.0223
Node 10: 0.0222
Node 12: 0.0197
Node 7: 0.0137
Node 39: 0.0083
Node 36: 0.0070
Node 31: 0.0009
Node 8: 0.0000
Node 40: 0.0000
Node 44: 0.0000
Node 45: 0.0000
Node 46: 0.0000
```

13

Hasil Ringkasan menggunakan Eigen Vektor

```
ringkasan_eigenvector_preprocessing = ""
print("Tiga Node Tertinggi Eigen Vector Menggunakan Preprocessing:")
for node, eigenvector_preprocessing in sorted_eigenvector_preprocessing[:3]:
    top_sentence = kalimat[node]
    ringkasan_eigenvector_preprocessing += top_sentence + " "
    print(f"Node {node}: Page Rank = {eigenvector_preprocessing:.4f}")
    print(f"Kalimat: {top_sentence}\n")
```

```
Tiga Node Tertinggi Eigen Vector Menggunakan Preprocessing:
Node 0: Page Rank = 0.3588
Kalimat: peran teknik industri pemanfaatan ai era industri

Node 18: Page Rank = 0.3579
Kalimat: teknik industri solusi menghadapi kompleksitas transformasi industri didorong ai ilmu teknik

Node 21: Page Rank = 0.2843
Kalimat: program studi teknik industri binus university menyajikan kurikulum terkini relevan menjembat
```

```
print("Ringkasan Berita Menggunakan Eigenvector Centrality:")
print(ringkasan_eigenvector_preprocessing)
```

```
Ringkasan Berita Menggunakan Eigenvector Centrality:
peran teknik industri pemanfaatan ai era industri  teknik industri solusi menghadapi kompleksitas trans
```

14

Summary

In summary, this book has no content whatsoever.

References