

Akbar Nugroho

2211110015

SD03A

✓ Import Library

```
import pandas as pd
import numpy as np
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
import re
```

✓ Text-Preprocessing

```
def preprocess(text):
    text = text.lower()
    text = re.sub(r'\W+', ' ', text)
    text = re.sub(r'\d+', ' ', text)
    text = re.sub(r'\s+[a-z]\s+', ' ', text)
    text = re.sub(r'^\s+', '', text)
    return text

# Membaca dataset dari file CSV
df = pd.read_csv('dataset_mentalhealth.csv')

# Preprocessing pada kolom 'Questions' dan 'Jawaban'
df['Questions'] = df['Questions'].apply(preprocess)
df['Jawaban'] = df['Jawaban'].apply(preprocess)
```

✓ Normalisasi & Vektorisasi

```
# Proses vektorisasi
tfidf = TfidfVectorizer()
tfidf_matrix = tfidf.fit_transform(df['Questions'])
```

✓ Modelling

```
def answer_question(user question):
```

```

preprocessed_question = preprocess(user_question)
user_question_vec = tfidf.transform([preprocessed_question])
similarity = cosine_similarity(user_question_vec, tfidf_matrix).flatten()
index = similarity.argsort()[-1]
return df['Jawaban'].iloc[index]

```

✓ Evaluasi Model

```

eval_data = [
    ['apa yang dimaksud dengan penyakit mental?', 'penyakit mental adalah kondisi kese
    ['bagaimana saya tahu kalau saya tidak sehat?', 'jika keyakinan, pikiran, perasaar
    ['saya sudah dewasa dan saya pikir saya mungkin menderita adhd. Apa yang bisa saya
]

eval_df = pd.DataFrame(eval_data, columns=['Questions', 'Jawaban'])
eval_df

```

	Questions	Jawaban
0	apa yang dimaksud dengan penyakit mental?	penyakit mental adalah kondisi kesehatan yang ...
1	bagaimana saya tahu kalau saya tidak sehat?	jika keyakinan, pikiran, perasaan, atau perila...

```

predicted_answers = []
for question in eval_df['Questions']:
    ans = answer_question(question)
    predicted_answers.append(ans)

eval_df['PredictedAnswers'] = predicted_answers

```

```

from sklearn.metrics import accuracy_score

accuracy = accuracy_score(eval_df['Jawaban'], eval_df['PredictedAnswers'])
print('Akurasi model:', accuracy)

Akurasi model: 0.0

```

```

test_questions = ['apa yang disebut penyakit mental?']

for question in test_questions:
    print('Pertanyaan:', question)
    print('Jawaban Model:', answer_question(question))
    print()

Pertanyaan: apa yang disebut penyakit mental?
Jawaban Model: penyakit mental adalah kondisi kesehatan yang mengganggu pikiran en

```

✓ Uji Coba 10 Pertanyaan

Coba Pertanyaan

```
new_test_questions = [
    "Apa yang dimaksud dengan gangguan kecemasan?",
    "Bagaimana cara mengenali tanda-tanda depresi?",
    "Apa manfaat olahraga untuk kesehatan mental?",
    "Bagaimana cara mengatasi stres sehari-hari?",
    "Apakah terapi psikologis efektif untuk masalah mental?",
    "Apa yang dapat dilakukan untuk meningkatkan kesehatan mental di tempat kerja?",
    "Bagaimana cara mendukung teman yang sedang mengalami krisis mental?",
    "Apa peran nutrisi dalam mendukung kesehatan mental?",
    "Apakah tidur yang cukup berpengaruh pada kesehatan mental?",
    "Bagaimana cara menjaga keseimbangan kesehatan mental?"
]

new_predicted_answers = []

for question in new_test_questions:
    ans = answer_question(question)
    new_predicted_answers.append(ans)

# tampilkan pertanyaan dan jawaban yang diprediksi
for i in range(len(new_test_questions)):
    print('Pertanyaan : ', new_test_questions[i])
    print('Jawaban Model : ', new_predicted_answers[i])
    print()
```

Pertanyaan : Apa yang dimaksud dengan gangguan kecemasan?
Jawaban Model : penyakit mental adalah kondisi kesehatan yang mengganggu pikiran &

Pertanyaan : Bagaimana cara mengenali tanda-tanda depresi?
Jawaban Model : gejala gangguan kesehatan mental bervariasi tergantung pada jenis

Pertanyaan : Apa manfaat olahraga untuk kesehatan mental?
Jawaban Model : kita semua memiliki kesehatan mental yang terdiri dari keyakinan p

Pertanyaan : Bagaimana cara mengatasi stres sehari-hari?
Jawaban Model : sementara banyak orang memikirkan kesedihan dalam hal kehilangan s

Pertanyaan : Apakah terapi psikologis efektif untuk masalah mental?
Jawaban Model : seringkali lebih realistis dan bermanfaat untuk mengetahui apa yar

Pertanyaan : Apa yang dapat dilakukan untuk meningkatkan kesehatan mental di temp
Jawaban Model : masalah kesehatan mental adalah masalah serius pada usia berapa pu

Pertanyaan : Bagaimana cara mendukung teman yang sedang mengalami krisis mental?
Jawaban Model : ini mungkin tergantung pada hubungan anda dengan mereka dengan len

Pertanyaan : Apa peran nutrisi dalam mendukung kesehatan mental?
Jawaban Model : kita semua memiliki kesehatan mental yang terdiri dari keyakinan p

Pertanyaan : Apakah tidur yang cukup berpengaruh pada kesehatan mental?
Jawaban Model : seringkali lebih realistis dan bermanfaat untuk mengetahui apa yar

Pertanyaan : Bagaimana cara menjaga keseimbangan kesehatan mental?

Jawaban Model : merawat kesehatan fisik anda juga baik untuk kesehatan mental anda

```
import csv

# Pertanyaan dan jawaban prediksi
questions = new_test_questions
predicted_answers = new_predicted_answers

# Buat file CSV baru
with open('hasil_predictions.csv', 'w', newline='') as file:
    writer = csv.writer(file)

    # Tulis header kolom
    writer.writerow(['Questions', 'Jawaban'])

    # Tulis masing-masing pertanyaan dan jawaban prediksi ke baris csv
    for i in range(len(questions)):
        row = [questions[i], predicted_answers[i]]
        writer.writerow(row)

print("Hasil prediksi baru berhasil diekspor ke csv")
```

Hasil prediksi baru berhasil diekspor ke csv

✓ Kesimpulan

Dari hasil di atas, model dapat berjalan dan juga memberikan jawaban pada semua pertanyaan yang diberikan, tetapi dari hasil di atas hanya terdapat 5 jawaban yang relevan dengan pertanyaannya, 5 jawaban yang lainnya kurang relevan dengan pertanyaan yang diberikan.

✓ LINK GIT HUB

https://github.com/Akbar-Nugroho/praktikum-nlp/tree/main/responsi_2

