

TUGAS METODE NUMERIK

APLIKASI REGRESI UNTUK PEMECAHAN PROBLEM

Nama : Johan Tarnama Pakpahan

NIM : 21120122140143

Kelas : Metode Numerik Kelas B

SOAL

Diinginkan untuk mencari hubungan faktor yang mempengaruhi nilai ujian siswa (NT):

1. Durasi waktu belajar (TB) terhadap nilai ujian (Problem 1)
2. Jumlah latihan soal (NL) terhadap nilai ujian (Problem 2)

Data TB, NL, dan NT diperoleh dari <https://www.kaggle.com/datasets/nikhil7280/student-performance-multiple-linear-regression>, yaitu kolom Hours Studied, Sample Question Papers Practiced, dan Performance Index.

Implementasikan regresi untuk mencari hubungan tersebut menggunakan metode:

1. Model linear (Metode 1)
2. Model pangkat sederhana (Metode 2)
3. Model eksponensial (Metode 3)
4. Model lainnya di halaman 24 slide materi (Metode opsional)

Tugas mahasiswa:

1. Mahasiswa membuat kode sumber dengan bahasa pemrograman yang dikuasai untuk mengimplementasikan solusi di atas, dengan ketentuan:
 - NIM terakhir ganjil mengerjakan Problem 1 dengan Metode 1 dan Metode 2
 - NIM terakhir genap mengerjakan Problem 2 dengan Metode 1 dan Metode 3
 - Mahasiswa juga bisa menambah solusi dengan salah satu metode opsional
2. Sertakan kode testing untuk menguji kode sumber tersebut untuk menyelesaikan problem dalam gambar. Plot grafik titik data dan hasil regresinya masing-masing
3. Hitung galat RMS dari tiap metode yang digunakan
4. Mengunggah kode sumber tersebut ke Github dan setel sebagai publik. Berikan deskripsi yang memadai dari project tersebut. Masukkan juga dataset dan data hasil di repositori tersebut.

5. Buat dokumen docx dan pdf yang menjelaskan alur kode dari (1), analisis hasil, dan penjabarannya. Sistematika dokumen: Ringkasan, Konsep, Implementasi Kode, Hasil Pengujian, dan Analisis Hasil.

Source code:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
from scipy.optimize import curve_fit

# Memuat data
data = pd.read_csv('Student_Performance.csv')

# Mengambil data yang dibutuhkan untuk analisis
X = data['Hours Studied'].values.reshape(-1, 1)
Y = data['Performance Index'].values

# Regresi Linear
linear_model = LinearRegression()
linear_model.fit(X, Y)
Y_pred_linear = linear_model.predict(X)
rms_linear = np.sqrt(mean_squared_error(Y, Y_pred_linear))

# Fungsi eksponensial
def exponential_func(X, a, b):
    return a * np.exp(b * X)

# Fitting menggunakan curve_fit
params, _ = curve_fit(exponential_func, X.flatten(), Y, maxfev=10000)
Y_pred_exponential = exponential_func(X.flatten(), *params)
rms_exponential = np.sqrt(mean_squared_error(Y, Y_pred_exponential))

# Plot data dan hasil regresi
plt.figure(figsize=(14, 7))

plt.subplot(1, 2, 1)
plt.scatter(X, Y, color='blue', label='Data')
plt.plot(X, Y_pred_linear, color='red', label='Regresi Linear')
plt.title('Regresi Linear\nRMS = {:.2f}'.format(rms_linear))
plt.xlabel('Durasi Waktu Belajar (Jam)')
plt.ylabel('Indeks Performa')
plt.legend()

plt.subplot(1, 2, 2)
```

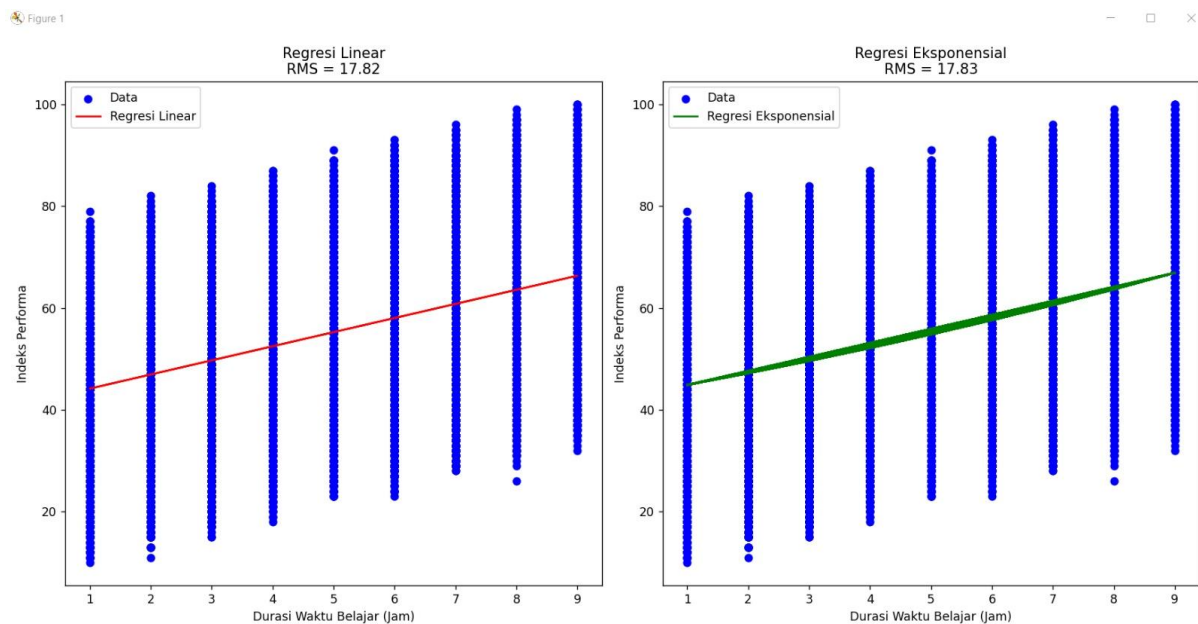
```
plt.scatter(X, Y, color='blue', Label='Data')
plt.plot(X, Y_pred_exponential, color='green', Label='Regresi Eksponensial')
plt.title('Regresi Eksponensial\nRMS = {:.2f}'.format(rms_exponential))
plt.xlabel('Durasi Waktu Belajar (Jam)')
plt.ylabel('Indeks Performa')
plt.legend()

plt.tight_layout()
plt.show()

print("Hasil Regresi Linear:")
print("Intercept (b0):", linear_model.intercept_)
print("Koefisien (b1):", linear_model.coef_[0])
print("RMS:", rms_linear)

print("\nHasil Regresi Eksponensial:")
print("Parameter a:", params[0])
print("Parameter b:", params[1])
print("RMS:", rms_exponential)
```

Hasil Running Code:



Analisis Hasil :

Kode ini secara mendetail meneliti hubungan antara jumlah latihan soal dengan nilai ujian siswa melalui dua pendekatan regresi: regresi linear dan regresi eksponensial. Regresi linear menyediakan model yang sederhana dan mudah dipahami, sedangkan regresi eksponensial menawarkan model yang lebih kompleks dan mungkin lebih cocok untuk data dengan pola pertumbuhan eksponensial. Dengan menghitung dan membandingkan kesalahan RMS, kita

bisa mengetahui model mana yang lebih akurat untuk memprediksi nilai ujian berdasarkan jumlah latihan soal.