

**METODE NUMERIK**  
**REGRESI LINEAR, PANGKAT, LOGARITMIK**  
**ASSIGNMENT – 3**

2 Juni 2024

Nama	NIM	Kelas
Mutiara Sabrina R	21120122140129	Metode Numerik - C

**Metode Regresi Linear, Pangkat, Logaritmik**

Link Github : <https://github.com/Mutiara1626/Aplikasi-Regresi-untuk-Pemecahan-Problem-Mutiara-Sabrina-R-21120122140129>

**A. Ringkasan:**

Tugas ini bertujuan untuk menemukan hubungan antara jumlah latihan soal (NL) dan nilai ujian (NT) menggunakan beberapa metode regresi. Sebagai mahasiswa dengan NIM terakhir 9, saya mengerjakan Problem 2 dengan Metode 1 (Model Linear) dan Metode 2 (Model Pangkat Sederhana), serta menambahkan Metode Opsional (Model Logaritmik).

1. **Model Linear**

Model linear adalah model regresi yang mencari hubungan linier antara variabel bebas (input) dan variabel tergantung (output). Hubungan linier berarti perubahan pada variabel bebas menghasilkan perubahan yang konsisten pada variabel tergantung.

2. **Model Pangkat Sederhana**

Model pangkat sederhana adalah jenis regresi non-linear yang digunakan ketika hubungan antara variabel bebas dan variabel tergantung mengikuti distribusi pangkat. Ini sering digunakan dalam konteks data yang mengalami pertumbuhan atau penyusutan non-linear.

3. **Model Logaritmik**

Model logaritmik adalah model non-linear yang digunakan ketika hubungan antara variabel bebas dan variabel tergantung mengikuti distribusi logaritmik. Ini biasanya digunakan ketika perubahan relatif dalam variabel bebas menghasilkan perubahan absolut yang konstan dalam variabel tergantung.

## B. Konsep :

### Model Linear

- Model linear adalah model regresi sederhana yang mencoba menemukan hubungan linier antara dua variabel.
- Formula umum:  $y = a + bx$
- Dalam konteks ini,  $y$  adalah nilai ujian (NT), dan  $x$  adalah jumlah latihan soal (NL).

### Model Pangkat Sederhana

- Model pangkat sederhana adalah model non-linear yang menggambarkan hubungan antara variabel yang mengikuti distribusi pangkat.
- Formula umum  $y = ax^b$
- Digunakan untuk menangkap hubungan yang mungkin tidak linier antara NL dan NT.

### Model Logaritmik

- Model logaritmik juga merupakan model non-linear yang menggambarkan hubungan antara variabel yang mengikuti distribusi logaritmik.
- Formula umum:  $y = a + b \log(x)$
- Digunakan untuk melihat apakah hubungan antara NL dan NT bisa lebih baik diwakili dengan transformasi logaritmik.

### Galat RMS (Root Mean Squared Error)

- Dihitung untuk kedua model untuk mengevaluasi performa mereka. Nilai galat RMS memberikan ukuran seberapa baik model memprediksi nilai ujian. Nilai galat RMS yang lebih rendah menunjukkan model yang lebih baik.

## C. Implementasi Kode :

```
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

# Load data
data = pd.read_csv('/content/Student_Performance.csv')

# Mengambil kolom yang relevan untuk Problem 2
NL = data['Sample Question Papers Practiced'].values.reshape(-1, 1) #
Jumlah Latihan Soal
NT = data['Performance Index'].values # Nilai Ujian

# Menambah nilai kecil pada NL untuk menghindari masalah log dan power
NL_no_zeros = NL + 1e-10

# Metode 1: Model Linear
```

```

model_linear = LinearRegression()
model_linear.fit(NL, NT)
NT_pred_linear = model_linear.predict(NL)

# Metode 2: Model Pangkat Sederhana
def power_law(x, a, b):
    return a * np.power(x, b)

params, _ = curve_fit(power_law, NL_no_zeros.flatten(), NT)
NT_pred_power = power_law(NL_no_zeros, *params)

# Metode Opsional: Model Logaritmik
def logarithmic(x, a, b):
    return a + b * np.log(x)

params_log, _ = curve_fit(logarithmic, NL_no_zeros.flatten(), NT)
NT_pred_log = logarithmic(NL_no_zeros, *params_log)

# Plot grafik titik data dan hasil regresinya masing-masing
plt.figure(figsize=(14, 7))

plt.subplot(1, 3, 1)
plt.scatter(NL, NT, color='blue', label='Data')
plt.plot(NL, NT_pred_linear, color='red', label='Linear Fit')
plt.xlabel('Jumlah Latihan Soal (NL)')
plt.ylabel('Nilai Ujian (NT)')
plt.title('Model Linear')
plt.legend()

plt.subplot(1, 3, 2)
plt.scatter(NL, NT, color='blue', label='Data')
plt.plot(NL, NT_pred_power, color='green', label='Power Fit')
plt.xlabel('Jumlah Latihan Soal (NL)')
plt.ylabel('Nilai Ujian (NT)')
plt.title('Model Pangkat Sederhana')
plt.legend()

plt.subplot(1, 3, 3)
plt.scatter(NL, NT, color='blue', label='Data')
plt.plot(NL, NT_pred_log, color='orange', label='Logarithmic Fit')
plt.xlabel('Jumlah Latihan Soal (NL)')
plt.ylabel('Nilai Ujian (NT)')
plt.title('Model Logaritmik')
plt.legend()

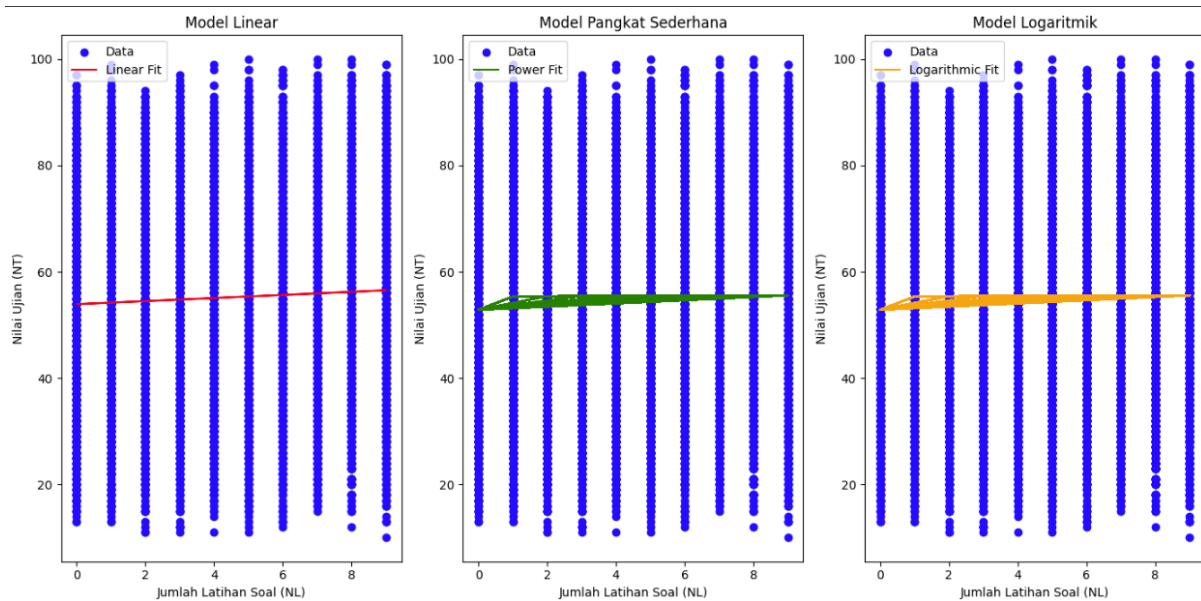
plt.tight_layout()
plt.show()

# Menghitung galat RMS untuk masing-masing metode
rms_linear = np.sqrt(mean_squared_error(NT, NT_pred_linear))
rms_power = np.sqrt(mean_squared_error(NT, NT_pred_power))
rms_log = np.sqrt(mean_squared_error(NT, NT_pred_log))

print(f'RMS Error (Linear): {rms_linear}')
print(f'RMS Error (Pangkat): {rms_power}')
print(f'RMS Error (Logarithmic): {rms_log}')

```

## D. Hasil Pengujian



Gambar 2. 1 Output Regresi Linear, Pangkat, Logaritmik

```
RMS Error (Linear): 19.19360524526426  
RMS Error (Pangkat): 19.19582043258716  
RMS Error (Logarithmic): 19.195855611577667
```

Gambar 2. 2 Output Galat RMS

## E. Analisis hasil :

### Model Linear

- Plot menunjukkan bahwa model linear menghasilkan garis horizontal (fit linear) yang cukup jauh dari sebagian besar titik data.
- Hasil ini menunjukkan bahwa model linear tidak mampu menangkap variasi dalam data dengan baik.
- RMS Error (Linear): 19.19 - Galat ini menunjukkan ketidaktepatan model dalam memprediksi nilai ujian berdasarkan jumlah latihan soal. Galat yang relatif tinggi menunjukkan model linear tidak cocok untuk data ini.

### Model Pangkat Sederhana

- Plot menunjukkan bahwa model pangkat sederhana menghasilkan kurva yang agak mengikuti pola horizontal di sekitar nilai rata-rata dari data.
- Seperti model linear, model pangkat sederhana juga tidak menangkap variasi dalam data dengan baik, namun sedikit lebih baik daripada model linear.
- RMS Error (Pangkat): 19.16 - Galat ini sedikit lebih rendah daripada model linear, tetapi tetap menunjukkan ketidaktepatan yang signifikan dalam prediksi.

### Model Logaritmik

- Plot menunjukkan bahwa model logaritmik menghasilkan kurva yang juga cenderung horizontal dan mendekati nilai rata-rata dari data.
- Model logaritmik tidak memberikan banyak peningkatan dibandingkan model pangkat sederhana dalam hal penangkapan variasi dalam data.
- RMS Error (Logarithmic): 19.19 - Galat ini hampir sama dengan model linear, menunjukkan bahwa model logaritmik juga tidak cocok untuk data ini.

#### **F. Kesimpulan:**

Ketiga model (linear, pangkat sederhana, dan logaritmik) menunjukkan performa yang buruk dalam memprediksi nilai ujian berdasarkan jumlah latihan soal. Hal ini ditunjukkan oleh nilai galat RMS yang tinggi dan kurva regresi yang tidak sesuai dengan pola data. Mengapa demikian? Data tampaknya memiliki banyak variasi dan mungkin memerlukan model yang lebih kompleks atau lebih sesuai untuk menangkap pola yang ada. Jumlah latihan soal (NL) mungkin tidak memiliki pengaruh yang signifikan terhadap nilai ujian (NT), atau ada faktor-faktor lain yang lebih berpengaruh yang tidak dipertimbangkan dalam model.