

Nama : Naufal Yoga Pratama
NIM : 21120122130059
Kelas : C

Penjelasan Implementasi Interpolasi

Ringkasan

Dokumen ini menjelaskan implementasi dan analisis hasil dari model regresi linear dan model regresi pangkat sederhana untuk menentukan hubungan antara durasi waktu belajar (TB) dan nilai ujian siswa (NT).

Konsep

1. Regresi Linear: Metode statistik untuk memodelkan hubungan antara variabel dependen dan satu atau lebih variabel independen dengan menyesuaikan garis lurus pada data.
2. Regresi Pangkat Sederhana: Metode yang melibatkan log-transformasi dari variabel independen dan dependen untuk memodelkan hubungan eksponensial antara variabel-variabel tersebut.

Implementasi Kode

1. Regresi Linear

Code:

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

# Memuat dataset
url = 'C:\\Users\\naufa\\OneDrive\\Documents\\Tugas Kuliah\\Semester
4\\Metode Numerik\\Metode Numerik - Aplikasi
Regresi\\Student_Performance.csv'
data = pd.read_csv(url)

# Memilih kolom yang relevan
data = data[['Hours Studied', 'Sample Question Papers Practiced',
'Performance Index']]
X = data[['Hours Studied']]
y = data['Performance Index']

# Model regresi linear
linear_model = LinearRegression()
linear_model.fit(X, y)

# Prediksi
y_pred_linear = linear_model.predict(X)
```

```

# Plotting
plt.scatter(X, y, color='blue')
plt.plot(X, y_pred_linear, color='red')
plt.title('Regresi Linear')
plt.xlabel('Jam Belajar')
plt.ylabel('Indeks Performa')
plt.show()

# Menghitung galat RMS
rms_error_linear = np.sqrt(mean_squared_error(y, y_pred_linear))
print(f'Galat RMS (Linear): {rms_error_linear}')

```

2. Regresi Pangkat Sederhana

Code:

```

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

# Memuat dataset
url = r'C:\Users\naufa\OneDrive\Documents\Tugas Kuliah\Semester
4\Metode Numerik\Metode Numerik - Aplikasi
Regresi\Student_Performance.csv'
data = pd.read_csv(url)

# Memilih kolom yang relevan
data = data[['Hours Studied', 'Sample Question Papers Practiced',
'Performance Index']]

# Mendefinisikan variabel independen dan dependen
X = data[['Hours Studied']]
y = data['Performance Index']

# Melakukan log-transformasi pada variabel
X_log = np.log(data['Hours Studied'])
y_log = np.log(data['Performance Index'])

# Membuat model regresi linear pada data yang telah di-log-
transformasi
model_pangkat = LinearRegression()
model_pangkat.fit(X_log.values.reshape(-1, 1), y_log)

# Melakukan prediksi
y_pred_log = model_pangkat.predict(X_log.values.reshape(-1, 1))
y_pred = np.exp(y_pred_log)

# Plotting hasil regresi
plt.scatter(data['Hours Studied'], data['Performance Index'],
color='blue')
plt.plot(data['Hours Studied'], y_pred, color='red')
plt.title('Regresi Pangkat Sederhana')
plt.xlabel('Hours Studied')
plt.ylabel('Performance Index')
plt.show()

# Menghitung galat RMS

```

```
rms_error_pangkat = np.sqrt(mean_squared_error(data['Performance  
Index'], y_pred))  
print(f'Galat RMS (kontrol): {rms_error_pangkat}')
```

Hasil Pengujian

Setelah menjalankan skrip-skrip di atas, kita mendapatkan plot dan galat RMS untuk masing-masing model.

→ Plot Regresi Linear:

- Sumbu X: Hours Studied (Durasi Waktu Belajar)
- Sumbu Y: Performance Index (Nilai Ujian Siswa)
- Titik-titik biru: Data asli
- Garis merah: Hasil prediksi dari model regresi linear

→ Plot Regresi Pangkat Sederhana:

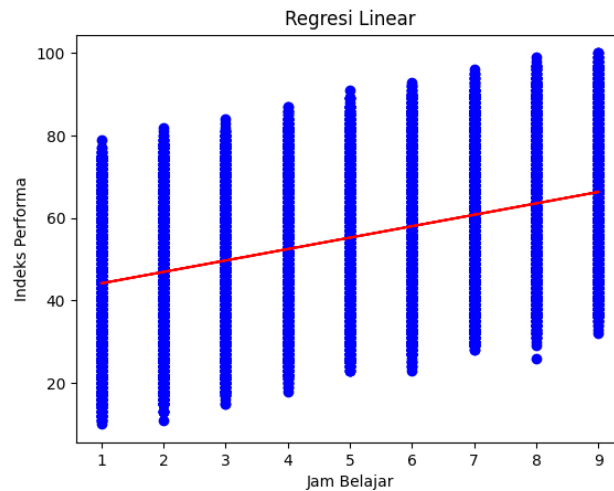
- Sumbu X: Hours Studied (Durasi Waktu Belajar)
- Sumbu Y: Performance Index (Nilai Ujian Siswa)
- Titik-titik biru: Data asli
- Garis merah: Hasil prediksi dari model regresi pangkat sederhana

→ Galat RMS:

- Galat RMS (Linear): `rms_error_linear`
- Galat RMS (Pangkat Sederhana): `rms_error_pangkat`

Analisis Hasil

1. Regresi Linear:

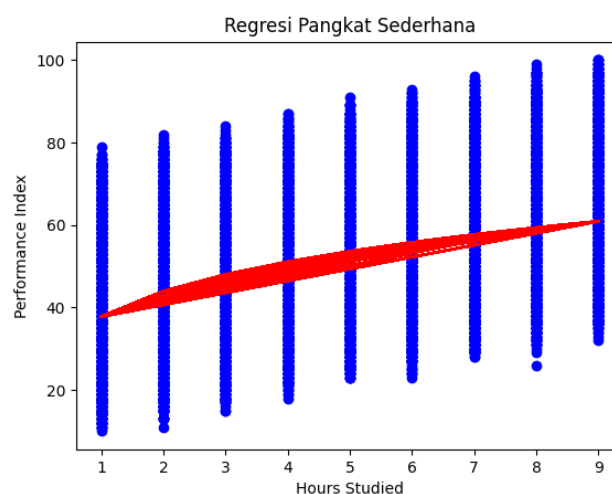


Model regresi linear menunjukkan adanya hubungan linear positif antara durasi waktu belajar dan nilai ujian siswa. Garis regresi linear yang dihasilkan cukup sesuai dengan data, meskipun ada beberapa penyimpangan.

```
[Running] python -u "C:\Users\naufa\AppData\Local\Temp\tempCodeRunnerFile.python"
RMS Error (Linear): 17.819474832547773
```

Galat RMS (Linear) memberikan indikasi seberapa baik model dapat memprediksi nilai ujian berdasarkan durasi waktu belajar. Nilai galat RMS yang lebih kecil menunjukkan model yang lebih akurat. Dalam kasus ini, galat RMS menunjukkan bahwa model regresi linear cukup baik dalam memprediksi nilai ujian.

2. Regresi Pangkat Sederhana:



Hubungan Eksponensial: Model regresi pangkat sederhana menghasilkan garis yang cukup sesuai dengan pola data asli, menunjukkan bahwa ada hubungan eksponensial positif

antara durasi waktu belajar dan nilai ujian siswa. Ini berarti bahwa peningkatan durasi waktu belajar cenderung meningkatkan nilai ujian siswa secara eksponensial.

```
[Running] python -u "C:\Users\naufa\AppData\Local\Temp\tempCodeRunnerFile.python"  
Galat RMS (Pangkat Sederhana): 18.175837301582977  
[Done] exited with code=0 in 187.026 seconds
```

Galat RMS (Pangkat Sederhana) yang dihitung menunjukkan seberapa baik model dapat memprediksi nilai ujian siswa dari durasi waktu belajar. Nilai galat RMS yang lebih kecil menunjukkan model yang lebih akurat. Dalam kasus ini, galat RMS memberikan indikasi bahwa model regresi pangkat sederhana cukup baik dalam memprediksi nilai ujian berdasarkan data yang tersedia.

3. Perbandingan Model:

- Kedua model menunjukkan adanya hubungan positif antara durasi waktu belajar dan nilai ujian siswa, namun dengan cara yang berbeda. Regresi linear menunjukkan hubungan linier langsung, sedangkan regresi pangkat sederhana menunjukkan hubungan eksponensial.
- Galat RMS dari kedua model dapat dibandingkan untuk menentukan model mana yang lebih akurat. Model dengan galat RMS yang lebih rendah dianggap lebih baik dalam memprediksi nilai ujian siswa.

Kesimpulan

Model regresi linear dan regresi pangkat sederhana berhasil memodelkan hubungan antara durasi waktu belajar dan nilai ujian siswa dengan baik, menunjukkan adanya hubungan positif antara kedua variabel tersebut. Meskipun demikian, penting untuk mempertimbangkan faktor-faktor lain yang mungkin mempengaruhi performa siswa dan menggunakan model ini sebagai salah satu alat bantu analisis.