Nama : Nanda Zayyan Nurirfan

NIM : 21120122130079

Metode Numerik B

Aplikasi Regresi Untuk Pemecahan Problem

1. Ringkasan

Proyek ini bertujuan untuk menemukan hubungan antara durasi waktu belajar (TB) dan nilai ujian siswa (NT) menggunakan regresi linear dan regresi eksponensial. Kedua model digunakan untuk melakukan prediksi nilai ujian berdasarkan durasi waktu belajar.

**Regresi Linear:** Regresi linear adalah pendekatan statistik yang digunakan untuk memodelkan hubungan linier antara variabel independen (X) dan variabel dependen (y). Model regresi linear dapat diwakili oleh persamaan: , di mana m adalah gradien garis regresi dan c adalah intersep.

**Regresi Eksponensial:** Regresi eksponensial digunakan ketika hubungan antara variabel independen dan dependen ditentukan oleh fungsi eksponensial. Model regresi eksponensial diwakili oleh persamaan: di mana a adalah parameter skala dan b adalah laju pertumbuhan.

1. Konsep:

Konsep dasar dari proyek ini adalah memahami hubungan antara dua variabel, yaitu durasi waktu belajar (TB) dan nilai ujian siswa (NT), dan mencari pola atau tren dalam data yang memungkinkan pembuatan model prediksi. Regresi linear digunakan untuk menemukan hubungan linier antara kedua variabel, sementara regresi eksponensial memodelkan hubungan eksponensial antara keduanya.

1. Implementasi Code
2. Regresi Linier

import pandas as pd

import numpy as np

from sklearn.linear\_model import LinearRegression

import matplotlib.pyplot as plt

from sklearn.metrics import mean\_squared\_error

# Load data

data = pd.read\_csv('Student\_Performance.csv')

# Persiapan data

X = data.iloc[:, 1].values.reshape(-1, 1) # Durasi waktu belajar

y = data.iloc[:, 6].values.reshape(-1, 1) # Nilai ujian siswa

# Membuat model regresi linear

model\_linear = LinearRegression()

model\_linear.fit(X, y)

# Prediksi nilai ujian menggunakan model linear

y\_pred\_linear = model\_linear.predict(X)

# Visualisasi data dan hasil regresi linear

plt.scatter(X, y, color='blue', label='Data')

plt.plot(X, y\_pred\_linear, color='red', label='Regresi Linear')

plt.title('Hubungan Durasi Waktu Belajar dan Nilai Ujian (Regresi Linear)')

plt.xlabel('Durasi Waktu Belajar (TB)')

plt.ylabel('Nilai Ujian (NT)')

plt.legend()

plt.show()

# Menghitung galat RMS regresi linear

rms\_linear = np.sqrt(mean\_squared\_error(y, y\_pred\_linear))

print("Galat RMS Regresi Linear:", rms\_linear)

1. Regresi Eksponensial

import pandas as pd

import numpy as np

from sklearn.linear\_model import LinearRegression

import matplotlib.pyplot as plt

from sklearn.metrics import mean\_squared\_error

*# Load data*

data = pd.read\_csv('Student\_Performance.csv')

*# Persiapan data*

X = data.iloc[:, 1].values.reshape(-1, 1)  *# Durasi waktu belajar*

y = data.iloc[:, 6].values.reshape(-1, 1)  *# Nilai ujian siswa*

*# Membuat model regresi eksponensial*

X\_exp = np.exp(X)

model\_exp = LinearRegression()

model\_exp.fit(X\_exp, y)

*# Prediksi nilai ujian menggunakan model eksponensial*

y\_pred\_exp = model\_exp.predict(X\_exp)

*# Visualisasi data dan hasil regresi eksponensial*

plt.scatter(X, y, color='blue', label='Data')

plt.plot(X, y\_pred\_exp, color='green', label='Regresi Eksponensial')

plt.title('Hubungan Durasi Waktu Belajar dan Nilai Ujian (Regresi Eksponensial)')

plt.xlabel('Durasi Waktu Belajar (TB)')

plt.ylabel('Nilai Ujian (NT)')

plt.legend()

plt.show()

*# Menghitung galat RMS regresi eksponensial*

rms\_exp = np.sqrt(mean\_squared\_error(y, y\_pred\_exp))

print("Galat RMS Regresi Eksponensial:", rms\_exp)

1. Hasil Pengujian dan Analisis

Dari hasil pengujian dengan dataset yang berbeda ukuran (N), dapat disimpulkan bahwa regresi eksponensial cenderung memberikan hasil yang lebih baik dibandingkan dengan regresi linear, terutama saat ukuran dataset meningkat. Hal ini terlihat dari penurunan galat RMS dan peningkatan dalam pemodelan pola yang sesuai dengan data aktual. Namun, perlu dicatat bahwa waktu eksekusi mungkin juga meningkat dengan ukuran dataset, terutama untuk regresi eksponensial yang mungkin memerlukan lebih banyak komputasi. Oleh karena itu, evaluasi performa model tidak hanya berdasarkan pada hasil regresi dan galat, tetapi juga memperhitungkan faktor waktu eksekusi, terutama dalam skenario di mana kinerja real-time atau pemrosesan cepat diperlukan.