



Jurusan Teknologi Informasi Politeknik Negeri Malang
Jobsheet-3: Regresi Linear
Mata Kuliah Pembelajaran Mesin

Ketentuan;

Simpanlah file lembar jawaban ini dengan format; **Kelas_Nama Lengkap**

Contoh; **TI 3C_Rifki Fakhruhin**

Upload tugas jobsheet ini dengan batas maksimum sesuai jadwal perkuliahan masing masing kelas

Upload file tugas jobsheet di website **lms.polinema**

Nama	:	Rifki Fakhruhin
Nim	:	2241720218
Kelas	:	3C

Tulislah Jawaban Pada Kolom Yang tersedia di bawah ini;

LEMBAR JAWABAN JOBSHEET-3 Tugas

Soal No	Jawaban
1	<pre># import package import numpy as np import pandas as pd import matplotlib.pyplot as plt import seaborn as sns from sklearn.linear_model import LinearRegression from sklearn.model_selection import train_test_split from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error from sklearn.preprocessing import StandardScaler</pre>
2	<pre># baca data dari file CSV from google.colab import files uploaded = files.upload() # upload dataset df = 'insurance.csv' data = pd.read_csv(df)</pre>
3	Tugas 1



Jurusan Teknologi Informasi Politeknik Negeri Malang
Jobsheet-3: Regresi Linear
Mata Kuliah Pembelajaran Mesin

	<pre># 1. Identifikasi variabel-variabel yang akan digunakan sebagai variabel bebas (fitur) dan variabel target (biaya medis personal) print("1. Identifikasi variabel-variabel:") # Read data from CSV df = 'insurance.csv' data = pd.read_csv(df) # Encode 'smoker' variable data['smoker'] = data['smoker'].map({'yes': 1, 'no': 0}) # Select independent variables (features) and the target variable (medical costs) X = data[['age', 'bmi', 'children', 'smoker']] y = data['charges'] print(f"Variabel bebas (fitur): {X.columns.tolist()}") print("Variabel target: 'charges'")</pre> <p>1. Identifikasi variabel-variabel: Variabel bebas (fitur): ['age', 'bmi', 'children', 'smoker'] Variabel target: 'charges'</p>
4	<p>Tugas 2</p> <pre># 2. Bagi dataset menjadi data latih (train) dan data uji (test) dengan proporsi yang sesuai print("\n2. Membagi dataset menjadi data latih dan data uji:") X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42) print(f>Data latih: {len(X_train)} baris, Data uji: {len(X_test)} baris")</pre> <p>2. Membagi dataset menjadi data latih dan data uji: Data latih: 1070 baris, Data uji: 268 baris</p>
5	<p>Tugas 3</p> <pre># 3. Lakukan feature scaling jika diperlukan print("\n3. Melakukan feature scaling (standarisasi fitur):") scaler = StandardScaler() X_train_scaled = scaler.fit_transform(X_train) X_test_scaled = scaler.transform(X_test) print("Fitur telah distandarisasi menggunakan StandardScaler.")</pre> <p>3. Melakukan feature scaling (standarisasi fitur): Fitur telah distandarisasi menggunakan StandardScaler.</p>
6	<p>Tugas 4</p> <pre># 4. Buat model multiple linear regression menggunakan Scikit-Learn print("\n4. Membuat model multiple linear regression:") model = LinearRegression()</pre> <p>4. Membuat model multiple linear regression:</p>
7	<p>Tugas 5</p>



Jurusan Teknologi Informasi Politeknik Negeri Malang
Jobsheet-3: Regresi Linear
Mata Kuliah Pembelajaran Mesin

```
# 5. Latih model pada data latih dan lakukan prediksi pada data uji
print("\n5. Melatih model dan melakukan prediksi pada data uji:")
model.fit(X_train_scaled, y_train)
y_pred = model.predict(X_test_scaled)
```

5. Melatih model dan melakukan prediksi pada data uji:

8 **Tugas 6**

```
# 6. Evaluasi model dengan menghitung metrik seperti R-squared, MSE, dan MAE
print("\n6. Evaluasi model:")
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
print(f"Mean Squared Error (MSE): {mse}")
print(f"R-squared: {r2}")
print(f"Mean Absolute Error (MAE): {mae}")

# Visualisasi hasil
plt.figure(figsize=(12, 6))

for i, feature in enumerate(['age', 'bmi', 'children', 'smoker']):
    plt.subplot(2, 2, i + 1)
    plt.scatter(X_test[feature], y_test, color='blue', label='Actual')
    plt.scatter(X_test[feature], y_pred, color='red', label='Predicted')
    plt.xlabel(feature.capitalize())
    plt.ylabel('charges')
    plt.legend()

plt.tight_layout()
plt.show()
```



Jurusan Teknologi Informasi Politeknik Negeri Malang
Jobsheet-3: Regresi Linear
Mata Kuliah Pembelajaran Mesin

6. Evaluasi model:

Mean Squared Error (MSE): 33981653.95019775

R-squared: 0.7811147722517887

Mean Absolute Error (MAE): 4213.798594527246

