Name : Kavya Choubey

Enrollment : 22SOEIT11064

Division :6ITE

Tutorial : 5

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score


# Load the dataset
df = pd.read_csv('/content/drive/MyDrive/SVM/std_marks_data.csv')

# Display the first few rows of the dataset
print(df.head())
```

```
   hours  age  internet  marks
0   6.84   15         0  78.64
1   6.56   20         1  88.80
2    NaN   21         1  88.90
3   8.67   22         1  98.99
4   7.55   17         1  92.34
```

```
# Check for missing values
print(df.isnull().sum())
```

```
hours       12
age          0
internet     0
marks        0
dtype: int64
```

```
# Fill missing values with the mean of the respective column
df.fillna(df.mean(), inplace=True)

# Verify if missing values are filled
print(df.isnull().sum())
```

```
hours       0
age         0
internet    0
marks       0
dtype: int64
```

```
# Input features (X) and target variable (y)
X = df[['hours', 'age', 'internet']]  # Input features
y = df['marks']  # Target variable


# Split the data into training and testing sets (80% train, 20% test)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)


# Initialize the Linear Regression model
model = LinearRegression()

# Train the model using the training data
model.fit(X_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
# Predict on the test data
y_pred = model.predict(X_test)

# Evaluate the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

```
print(f"Mean Squared Error: {mse}")
print(f"R-squared: {r2}")
```

```
Mean Squared Error: 296.9278739472545
R-squared: 0.04539029325368027
```

```
# New input data (hours, age, internet)
new_data = [[5, 20, 1]]  # Example: 5 hours, age 20, internet available

# Predict marks for the new input
predicted_marks = model.predict(new_data)
print(f"Predicted Marks: {predicted_marks[0]}")
```

```
Predicted Marks: 74.47098215271565
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LinearRe
  warnings.warn(
```