

```
"""
Bir veri kümesi kullanarak (hazır veri kümesi de kullanabilirsiniz) basit ve çoklu lineer regresyon modelleri kurun. Aşağıdaki adımları izleyin:

Veri Kümesini Yükleğin:
sklearn.datasets.load_diabetes() veri kümesini kullanın.

Basit Lineer Regresyon:

Hedef değişken: target

Bağımsız değişken: Veri kümesinden bir sütun seçin (örneğin: BMI)

Modeli eğitin ve R2 skorunu yazdırın.

Çoklu Lineer Regresyon:

Tüm bağımsız değişkenleri kullanarak bir model kurun.

R2 skorunu yazdırın ve basit modelle karşılaştırın.

Hata Metrikleri:
Her iki model için aşağıdaki metrikleri hesaplayın:

MAE (Ortalama mutlak hata)

MSE (Ortalama kare hata)

Yorumlayın:

Hangi model daha başarılı? Neden?

R2 değerleri ne ifade ediyor?
"""

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

diabetes = load_diabetes()
X = pd.DataFrame(diabetes.data, columns=diabetes.feature_names)
y = diabetes.target

# bmi (vücut kitle indeksi)
X_bmi = X[['bmi']]
X_train_bmi, X_test_bmi, y_train_bmi, y_test_bmi = train_test_split(X_bmi, y, test_size=0.2, random_state=42)

simple_model = LinearRegression()
simple_model.fit(X_train_bmi, y_train_bmi)
y_pred_simple = simple_model.predict(X_test_bmi)

# R2 skoru
r2_simple = r2_score(y_test_bmi, y_pred_simple)

# 2. Çoklu Lineer Regresyon
X_train_multi, X_test_multi, y_train_multi, y_test_multi = train_test_split(X, y, test_size=0.2, random_state=42)

multi_model = LinearRegression()
multi_model.fit(X_train_multi, y_train_multi)
y_pred_multi = multi_model.predict(X_test_multi)

# R2 skoru
r2_multi = r2_score(y_test_multi, y_pred_multi)

# Hata Metrikleri
mae_simple = mean_absolute_error(y_test_bmi, y_pred_simple)
mse_simple = mean_squared_error(y_test_bmi, y_pred_simple)

mae_multi = mean_absolute_error(y_test_multi, y_pred_multi)
mse_multi = mean_squared_error(y_test_multi, y_pred_multi)

# Sonuçları Yazdıralım
print("***Basit Lineer Regresyon (BMI)**")
print(f"R^2: {r2_simple:.4f}")
print(f"MAE: {mae_simple:.2f}")
print(f"MSE: {mse_simple:.2f}")

print("\n***Çoklu Lineer Regresyon***")
print(f"R^2: {r2_multi:.4f}")
print(f"MAE: {mae_multi:.2f}")
print(f"MSE: {mse_multi:.2f}")

# Çoklu lineer regresyon modeli daha başarılı çünkü: R^2 skoru daha yüksek, MAE ve MSE değerleri daha düşük
# R^2 verilerin birbirinden ne kadar farklı olduğunu gösterir
```