** **

**KÜTAHYA DUMPLUPİNAR ÜNİVERSİTESİ**

**BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ**

YUKSEK DUZEY PROGRAMLAMA DERSİ

**PREDICT FUTURE SALES**

FAİSAL OSMAN MOHAMUD

202013172122

**KÜTÜPHANELER**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_squared\_error, r2\_score

***Step 1: Generate Synthetic Sales Data***

np.random.seed(42)

days = np.array(range(1, 101))

sales = 100 + days \* 3 + np.random.normal(0, 10, size=100)

data = pd.DataFrame({"Day": days, "Sales": sales})

***Step 2: Prepare Data for Model***

X = data["Day"].values.reshape(-1, 1)

y = data["Sales"].values

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

***Step 3: Train the Linear Regression Model***

model = LinearRegression()

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

***Step 4: Make Predictions***

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

future\_days = np.array(range(101, 131)).reshape(-1, 1)

future\_sales = model.predict(future\_days)

***Step 5: Evaluate Model***

mse = mean\_squared\_error(y\_test, y\_pred)

r2 = r2\_score(y\_test, y\_pred)

*# Display Metrics*

print(f"Mean Squared Error: {mse:.2f}")

print(f"R-squared: {r2:.2f}")

***Step 6: Visualize Results***

plt.figure(figsize=(10, 6))

*# Historical data*

plt.scatter(data["Day"], data["Sales"], label="Historical Sales", color="blue")

*# Predictions on test data*

plt.scatter(X\_test, y\_pred, label="Predicted Test Sales", color="green")

*# Future predictions*

plt.plot(range(101, 131), future\_sales, label="Future Sales Prediction", color="red", linestyle="--")

plt.title("Sales Prediction")

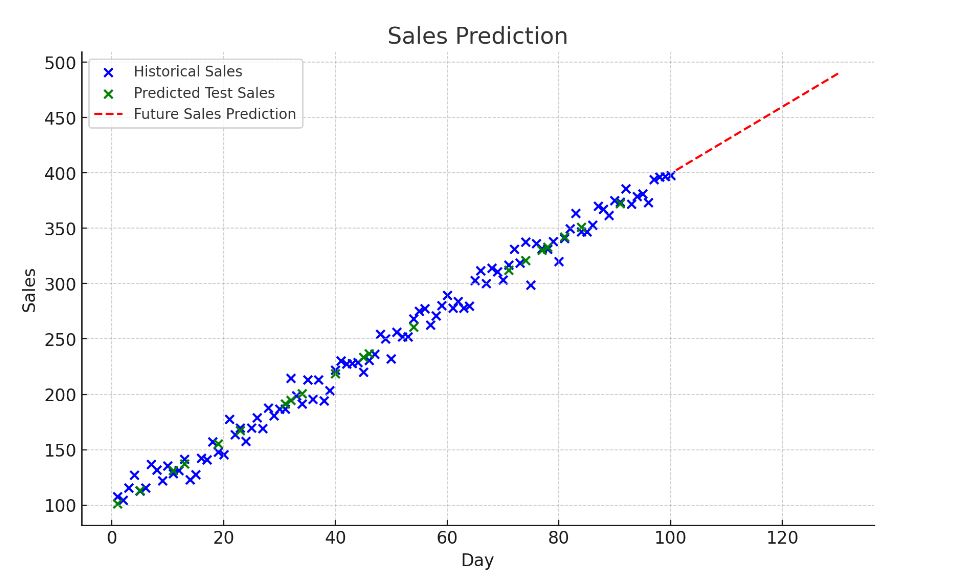
plt.xlabel("Day")

plt.ylabel("Sales")

plt.legend()

plt.grid(True)

plt.show()

****

Kod başarıyla çalıştırıldı ve işte sonuçlar:

* **Ortalama Kare Hatası (MSE):** 62.20
* **R-kare (R²):** 0.99

Görselleştirme şunları içerir:

1. Tarihsel satış verileri için mavi noktalar.
2. Test veri seti tahminleri için yeşil noktalar.
3. Önümüzdeki 30 gün için gelecekteki satış tahminlerini gösteren kırmızı kesik çizgi.

Yüksek R² değeri, modelin verileri çok iyi uyum sağladığını gösterir.