

Лабораторна робота №3 ІАД

Виконав студент групи МІТ-31
Тимохін Роман Миколайович

Тема: Регресійний аналіз

Мета:

Ознайомитися з методами регресійного аналізу, навчитися працювати з реальними і штучно створеними наборами даних для класифікації, застосовувати моделі для класифікації даних та аналізувати результати.

Завдання:

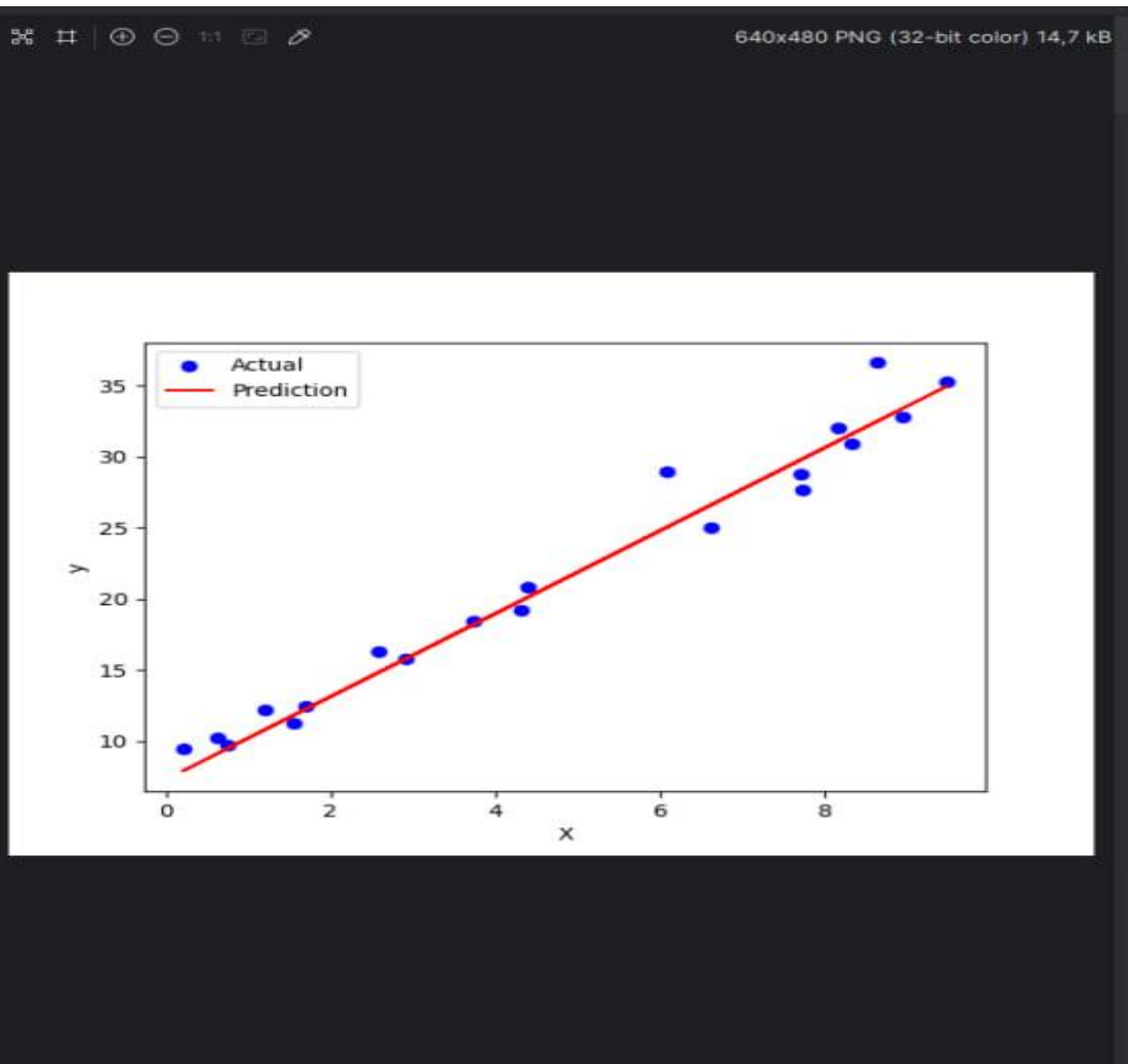
1. Створити набір даних за допомогою [DatasetGenerator](#).
2. Завантажити набір даних для регресії з сайту [Kaggle](#). Кожен студент має обрати *унікальний набір даних*.
3. (Опціонально для отримання відмінної оцінки) Згенерувати набір даних для регресії, використовуючи стандартні можливості Python (наприклад, з допомогою бібліотек Numpy або random).
4. Провести моделювання для отриманих наборів даних з використанням регресійного аналізу.
5. Викласти результати на GitHub у власний репозиторій в основну гілку в папці [Lab3](#).

Результати і код

```

1 import numpy as np
2 import pandas as pd
3 from sklearn.model_selection import train_test_split
4 from sklearn.linear_model import LinearRegression
5 from sklearn.metrics import mean_squared_error, r2_score
6 import matplotlib.pyplot as plt
7
8 # Генерація набору даних
9 np.random.seed(42)
10 X = np.random.rand(100, 1) * 10 # Генеруємо випадкові числа для X
11 y = 3 * X.squeeze() + 7 + np.random.randn(100) * 2 # Залежна змінна з шумом
12
13 # Створення DataFrame
14 generated_data = pd.DataFrame({'X': X.squeeze(), 'y': y})
15
16 # Завантаження даних з Kaggle (приклад)
17 # kaggle_data = pd.read_csv('path_to_kaggle_data.csv') # Завантажте набір даних
18
19 # Розподілення на тренувальну і тестову вибірки
20 X_train, X_test, y_train, y_test = train_test_split(*arrays: X, y, test_size=0.2, random_state=42)
21
22 # Створення та навчання моделі
23 model = LinearRegression()
24 model.fit(X_train, y_train)
25
26 # Прогнозування
27 y_pred = model.predict(X_test)
28
29 # Оцінка моделі
30 print(f"Mean Squared Error: {mean_squared_error(y_test, y_pred):.2f}")
31 print(f"R^2 Score: {r2_score(y_test, y_pred):.2f}")
32
33 # Візуалізація
34 plt.scatter(X_test, y_test, color="blue", label="Actual")
35 plt.plot(*args: X_test, y_pred, color="red", label="Prediction")
36 plt.xlabel("X")
37 plt.ylabel("y")
38 plt.legend()
39 plt.show()

```



```
C:\Users\roman\PycharmProjects\Pyth  
Mean Squared Error: 2.61  
R^2 Score: 0.97  
  
Process finished with exit code 0
```

Висновок:

Ця робота сприяє розвитку аналітичного мислення, вмінню обробляти та аналізувати дані, а також інтегрувати сучасні бібліотеки Python для вирішення регресійних задач.