CRISP-DM 流程概述

CRISP-DM 流程分為六個步驟:

- 1. 商業理解(Business Understanding):理解問題目標,這裡我們的目標是建立一個簡單的線性回歸模型,並讓使用者通過 Web 應用來調整模型參數,查看回歸結果的變化。
- 2. 數據理解 (Data Understanding):生成模擬數據,基於公式 y = ax + b + noise , 並可調整 斜率 a 、截距 b 以及資料點數和雜訊。
- 3. 數據準備(Data Preparation): 我們將生成隨機數據,並整理為適合建模的結構。
- 4. 建模 (Modeling):使用 scikit-learn 中的線性回歸模型來進行擬合。
- 5. 評估 (Evaluation):透過視覺化工具 (如 matplotlib),展示模型的回歸線及資料點,並輸出模型的評估結果 (斜率 a 、截距 b)。
- 6. 部署(Deployment):透過 Streamlit 或 Flask 將應用部署為 Web 應用,讓使用者可以調整參數並即時杳看結果。

步驟 1:安裝環境與工具

- 安装 Python:如果尚未安裝,請前往 Python 官網 下載並安裝,確保在安裝過程中勾選「Add Python to PATH」。
- 2. 安裝 Visual Studio Code: Visual Studio Code 下載 並安裝。
- 3. 安装必要的 Python 套件:打開 VS Code, 並打開終端(ctrl +), 執行以下命令安裝必要的庫:

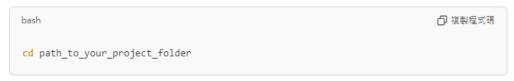
bash
pip install streamlit scikit-learn pandas matplotlib numpy

步驟 2:編寫線性回歸模型應用

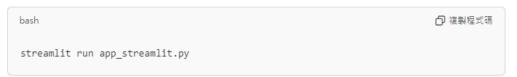
- 1. 創建專案資料夾:在桌面或其他位置創建一個資料夾,例如 linear_regression_project 。
- 2. 打開專案資料夾:在 VS Code 中打開該資料夾,然後創建一個新的 Python 檔案 app_streamlit.py。
- 3. 編寫程式碼:以下是簡單的線性回歸應用程式碼,請將其貼到 app_streamlit.py 中。

步驟 3: 運行 Streamlit 應用

1. 打開 VS Code 的終端 (Ctrl +) ,進入專案資料夾:



2. 執行 Streamlit 應用:



3. 應用程式會自動打開瀏覽器,並在 http://localhost:8501 上顯示。



13	2.1234	12.0371
14	1.8182	10,5431
15	1.834	13.0584
16	3.0424	15.8811
17	5.2476	23.3928
18	4.3195	21.672
19	2.9123	14.1688

模型斜率 (a): 3.8283465266567656

模型截距 (b): 4.053605003000303

