

撰寫線性回歸預測程式 (2025-11-05 07:23Z)

User (2025-11-05 07:23Z)

接下來的對話我會請你幫我撰寫一隻程式，需要遵守下面幾點：

1. 使用CRISP-DM流程建立
2. 程式碼使用jupyter notebook格式，code cell之間寫入對應的markdown來說明程式碼
3. 建立readme.md，詳細記錄程式設計流程和修改紀錄

程式需求：

1. 使用線性回歸模型預測學生表現
2. student_habits_performance.csv，包含15個欄位：
 - 第一欄是學生ID，可以去除
 - 最後一欄是成績，也就是要預測的值
 - 剩餘的14欄是特徵
3. 程式內容需要包含執行"特徵選擇"以及"模型評估"

Agent (model default, mode Agent)

先檢查現有資料檔案，接著依 CRISP-DM 流程建立 Jupyter Notebook 與 README.md。

► Tool use: **list_directory** • Listed directory undefined • **0** results

► Tool use: **glob_file_search** • Searched codebase "*.csv" • **1** result in **1** directory

讀取資料檔案以了解結構：

► Tool use: **read_file** • Read file: student_habits_performance.csv

建立符合 CRISP-DM 流程的 Jupyter Notebook 與 README.md。

► Tool use: **todo_write** • Todo List

► Tool use: **read_file** • Read file: student_habits_performance.csv

依 CRISP-DM 流程建立 Jupyter Notebook：

[illegible]

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

▶ Tool use: `**code_edit**` • Edit file: student_performance_prediction.ipynb

建立 README.md 文件：

▶ Tool use: `**todo_write**` • Todo List

▶ Tool use: `**code_edit**` • Edit file: README.md

完成所有任務。檢查是否有需要修正的問題：

▶ Tool use: `**todo_write**` • Todo List

▶ Tool use: `**read_lints**` • Read lints for 1 file

已完成。已建立符合需求的程式專案，內容如下：

已完成的工作

1. Jupyter Notebook

(student_performance_prediction.ipynb)

依 CRISP-DM 流程建立，包含：

- 商業理解：專案目標與成功標準
- 資料理解：資料探索、缺失值檢查、分佈分析、相關係數分析
- 資料準備：
 - 資料清理（移除學生ID）
 - 特徵編碼（處理類別變數）
 - 特徵選擇（三種方法：相關係數、SelectKBest、RFE）
 - 資料標準化與分割
- 建模：線性回歸模型建立與訓練
- 評估：多種評估指標（ R^2 、RMSE、MAE）與視覺化
- 總結：模型表現總結與後續建議

每個程式碼單元格之間都有 Markdown 說明。

2. README.md 文件

包含：

- 專案簡介與資料說明
- CRISP-DM 流程說明
- 程式設計流程與執行步驟
- 修改紀錄
- 注意事項與未來改進方向

主要特色

1. 特徵選擇：使用三種方法並以投票機制綜合結果
2. 模型評估：包含 R^2 、RMSE、MAE，並提供視覺化
3. 程式碼說明：每個步驟都有對應的 Markdown 說明
4. 文件完整：README.md 記錄設計流程與修改紀錄

可直接開啟 `student_performance_prediction.ipynb` 執行所有程式碼單元格。程式會自動：

- 載入並探索資料
- 進行特徵選擇
- 訓練線性回歸模型
- 評估模型表現
- 產生視覺化結果

如需調整或補充，請告知。
