

# Orange 《Linear Regression》模型設定說明

——以「學生成績預測」模型為例

參考：ChatGPT

編撰：屏東大學 周國華老師 2025/11/18

這個視窗負責設定 **線性迴歸 (Linear Regression) 模型** 的細節，包括截距、正規化方式、L1/L2 混合比例等。

---

## ① 模型名稱 (Name)

您將模型命名為：

👉 「學生成績預測」

這只是方便辨識的名稱，

在 Test & Score 或 Predictions 中會直接顯示這個名字。

---

## ② Parameters (參數)

✓ Fit intercept (是否包含截距)

- 打勾 = 模型會學習一個截距 (intercept)
- 不打勾 = 強迫模型的截距為 0

✦ 一般都建議保持打勾

因為現實世界中輸入特徵全為 0 時，目標值通常不會剛好是 0。

➔ 本模型有包含截距 → 合理設定

---

## ③ Regularization (正規化)

正規化用來避免過度擬合 (overfitting) ，  
並處理多重共線性 (multicollinearity) 。

您的設定選擇了：

## ★ Elastic Net Regression (彈性網迴歸)

Elastic Net = L1 (Lasso) + L2 (Ridge) 的組合，是一種更強健的正規化方式。

---

### 🔍 為什麼要使用正規化？

因為學生成績預測模型可能包含：

- 過多特徵 (study\_hours、sleep\_hours、prior\_grade...)
- 特徵之間可能高度相關 (如 study\_hours vs homework\_avg)

正規化能避免模型係數過大、過度依賴某些欄位，提高泛化能力。

---

## ④ Regularization Strength (正規化強度)

這個滑桿控制：

要多強的限制模型係數大小？

- 越往右 → 正規化越強，模型越保守
- 越往左 → 正規化越弱，模型更自由但較容易 overfitting

一般建議使用預設值即可，除非驗證結果顯示過度擬合。

---

## ⑤ Alpha (正則化係數)

Alpha 表示 正規化的整體強度

數值越大 → 正規化越強烈

數值越小 → 正規化較弱

在您圖中：

$\text{Alpha} = 0.0001$ （非常弱的正規化）

這通常意味著資料品質不錯、相關性適中，因此只需輕微正規化即可。

---

## Elastic Net Mixing（L1 / L2 比例）

Elastic Net =

- L1（Lasso）→ 會讓不重要特徵係數變成 0（特徵選擇）
- L2（Ridge）→ 讓模型更穩定，避免係數震盪

您的設定：

L1 : L2 = 0.50 : 0.50

這表示：

- 一半使用 Lasso（可自動做特徵選擇）
- 一半使用 Ridge（提高模型穩定性）

- 👉 這是非常實務且安全的設定
  - 👉 適合多特徵、多量表、多欄位的學生學習資料
- 

## 模型總結（教學版）

您的設定建立了一個：

- ★ 「包含截距」
- ★ 「同時使用 L1 與 L2 正規化」
- ★ 「正規化強度低、適度調整」

★ 「適用於多變量學生成績預測」

★ 「擁有較佳泛化能力」

的 Elastic Net 線性迴歸預測模型。

---

## 教學建議（可加入課程內容）

Elastic Net 為什麼適合「學生成績預測」？

- ✓ 特徵多、尺度不同（睡眠、作業、滿意度）
- ✓ 可能有共線性（study\_hours vs homework\_avg）
- ✓ 又希望模型能自動挑出重要特徵（L1）
- ✓ 同時保持穩定性（L2）

因此這是一個高品質、適合初學者的機器學習設定示範。

---