

Day 39

機器學習

LASSO, Ridge Regression



機器學習模型中的目標函數

- 機器學習模型的目標函數中有兩個非常重要的元素
 - 損失函數 (Loss function)
 - 正則化 (Regularization)
- 損失函數衡量預測值與實際值的差異，讓模型能往正確的方向學習
- 正則化則是避免模型變得過於複雜，造成過擬合 (Over-fitting)

回歸模型與正規化

- 前兩天學習到的回歸模型，我們只有提到損失函數會用 MSE 或 MAE
- 為了避免 Over-fitting，我們可以把正則化加入目標函數中，此時目標函數 = 損失函數 + 正則化
- 正則化可以懲罰模型的複雜度，當模型越複雜時其值就會越大

正則化函數

- 正則化函數是用來衡量模型的複雜度
- 該怎麼衡量？有 L1 與 L2 兩種函數
- L1： $\alpha \sum |weights|$
- L2： $\alpha \sum (weights)^2$
- 這兩種都是希望模型的參數數值不要太大，原因是參數的數值變小，噪音對最終輸出的結果影響越小，提升模型的泛化能力，但也讓模型的擬合能力下降

LASSO, Ridge Regression

- LASSO 為 Linear Regression 加上 L1
- Ridge 為 Linear Regression 加上 L2
- 其中有個超參數 α 可以調整正則化的強度
- 簡單來說，LASSO 與 Ridge 就是回歸模型加上不同的正則化函數

解題時間 It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

