輔大智慧金融實驗室

Class 4
找出最佳的模型參數

主講人：黃弘開
Email：kai830227@gmail.com

本週課程重點

1. 模型的訓練、驗證、測試
2. 交叉驗證
3. Grid Search

模型的訓練、驗證、測試

1. 訓練集、驗證集、測試集的用途
2. 交叉驗證

在前幾週的課程中

我們要將預測結果上傳到 Kaggle
才能知道模型的表現績效…

這樣做不只是麻煩而已，還會遇到一些問題

- 得到的評估分數是可靠的嗎？
- 有可能過度配適測試集（猜出正確答案）

Dataset

使用訓練及訓練模型

使用測試集確認
模型的效果

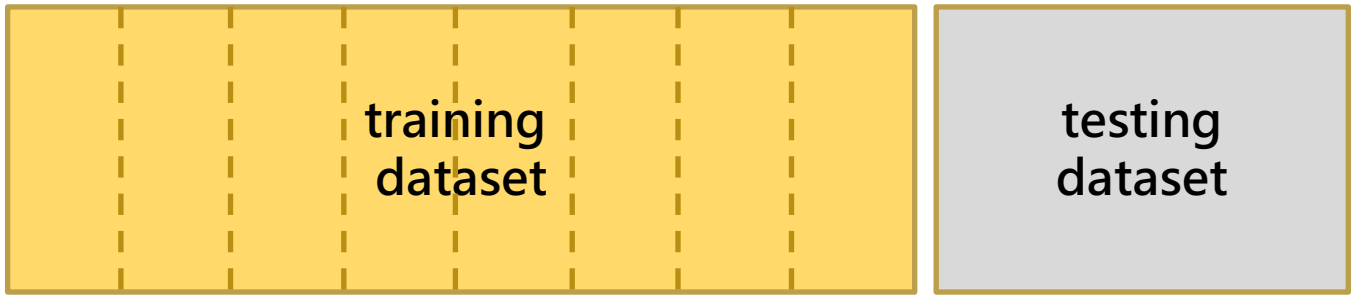
Dataset

training dataset	validation dataset	testing dataset
使用訓練及訓練模型	透過驗證集的驗證結果，調整模型超參數	使用測試集再次確認模型的效果

驗證集

使用驗證集的驗證結果，作為調整模型超參數的依據，這樣就可以避免模型過度配適訓練集

K-fold Cross Validation



1. 把訓練集隨機分為 K 等份
2. 進行 K 次訓練，其中一份為驗證集，其餘為訓練集
3. 計算 K 次驗證結果的平均作為模型分數

交叉驗證

使用交叉驗證來評估模型的訓練效果有兩個優點：

1. 減少資料分割時所帶來的變異
2. 所有訓練集中的資料都會參與模型訓練

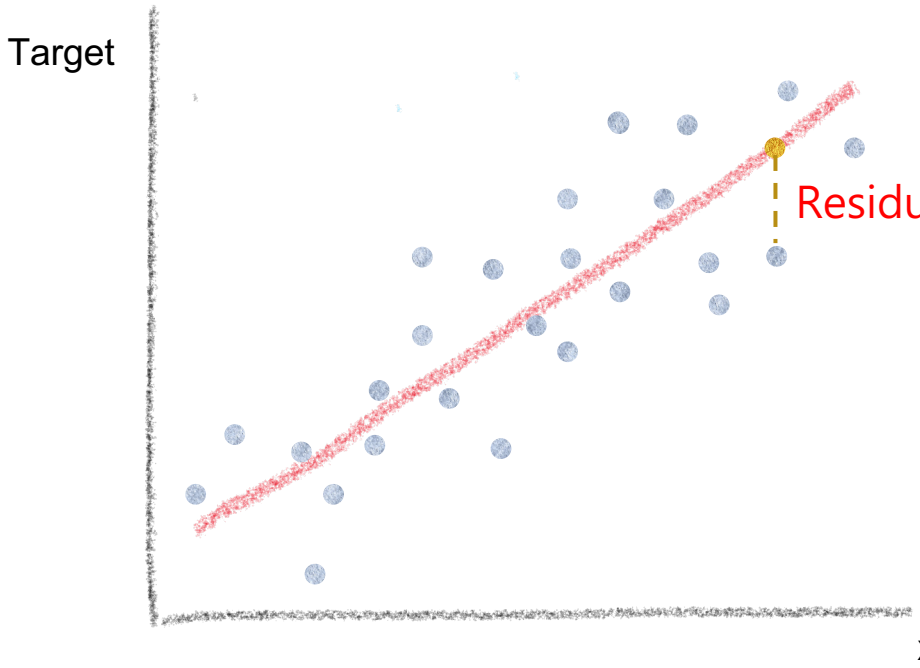
模型評估指標

1. MSE
2. Accuracy
3. Confusion Matrix
4. F1-Score

這樣做不只是麻煩而已，還會遇到一些問題

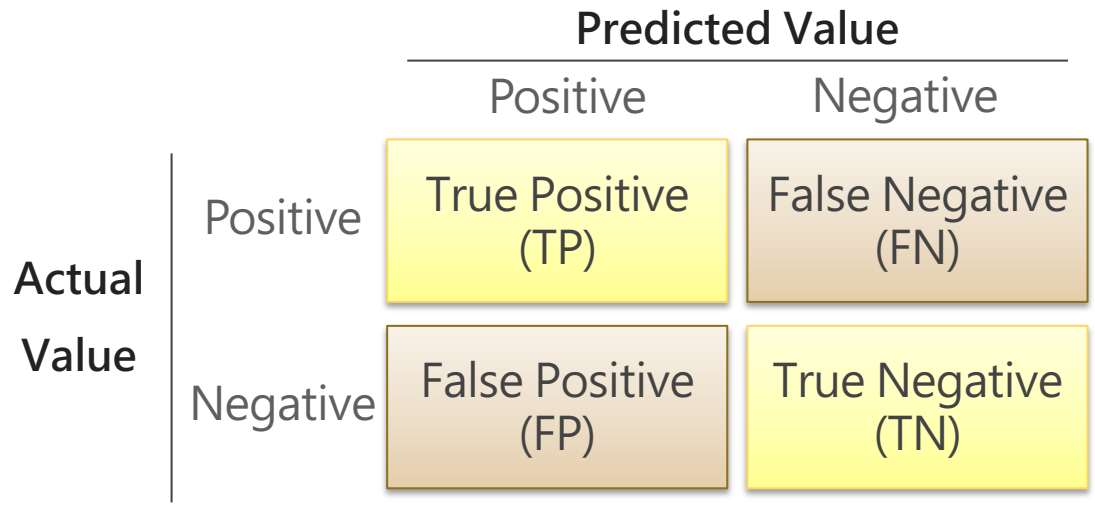
- 得到的評估分數是可靠的嗎？
- 有可能過度配適測試集（猜出正確答案）

Mean Squared Error 均方差

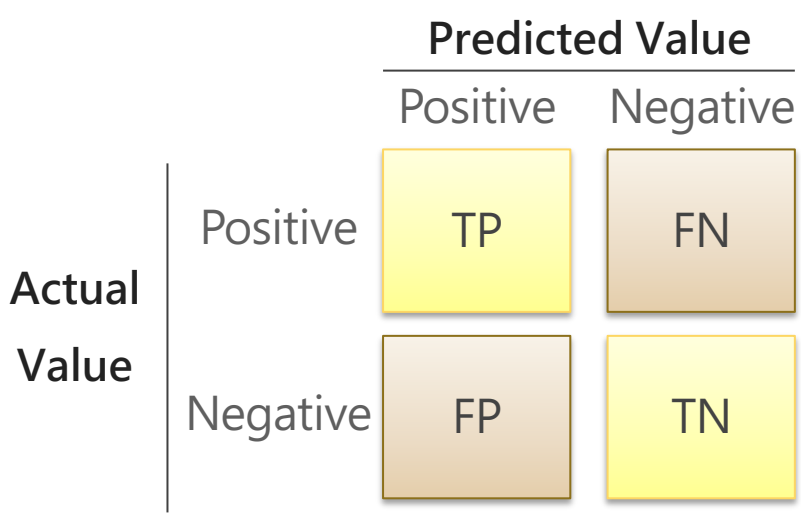


$$\mathrm{MSE}(y, \hat{y}) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2$$

Confusion Matrix 混淆矩陣



Confusion Matrix 混淆矩陣



Accuracy 正確率

= (TP+TN)/(TP+FN+FP+TN)

Accuracy是最常用的指標，但是在樣本不平衡的情況下，有可能失效。

Precision 精確率

預測為正的樣本中有多少是實際為正

= TP / (TP+FP)

Recall 召回率

實際為正的樣本有多少被預測為正

= TP / (TP+FN)

F1-Score

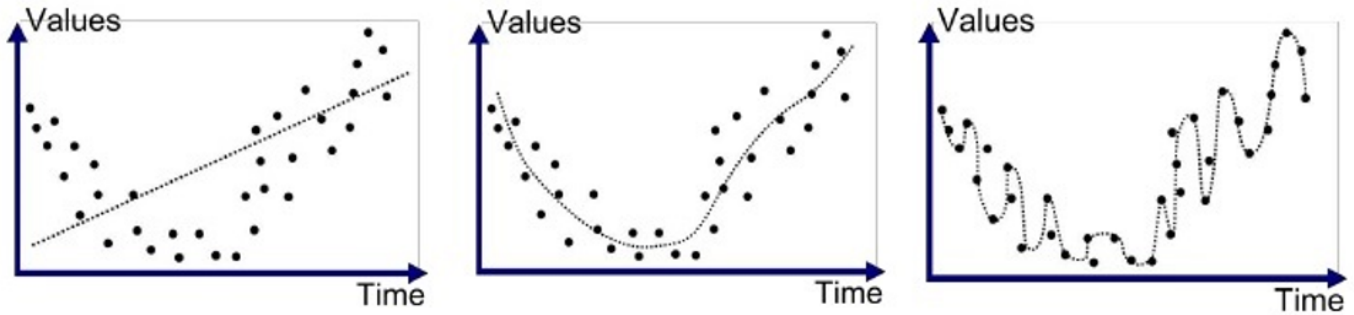
= 2PR / (P+R)

模型擬合

1. 欠擬合 Underfitting
2. 過擬合 Overfitting

模型訓練完了，怎麼樣的模型才夠好呢？

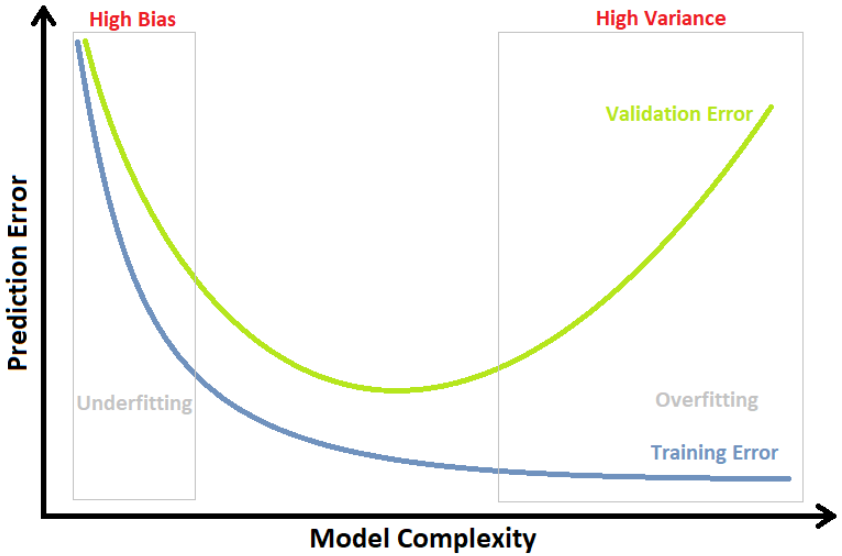
可能會出現三種狀況



Underfitted　　Good Fit/Robust　　Overfitted

模型訓練完了，怎麼樣的模型才夠好呢？

可以觀察模型在訓練集的表現與模型在測試集的表現差距多少？



Underfitting & Overfitting

Underfitting 欠擬合

模型在訓練集與測試集的表現都很差。

原因：模型太簡單

例如：決策樹的深度太淺

Overfitting 過擬合

模型在訓練集表現良好，但是在測試集表現很差。

原因：模型太複雜

例如：輸入特徵太多

模型超參數調整

1. Grid Search

Grid Search

1. 給定模型的超參數數列
2. 遍歷所有超參數組合，找出最佳組合

Thanks!

Does anyone have any questions?