



巨匠線上真人

Python 資料科學應用開發

第十四堂：模型評估與效能調校 (Model Evaluation and Performance Tuning)

同學，歡迎你參加本課程

- ☑ 請關閉你的FB、Line等溝通工具，以免影響你上課。
- ☑ 考量頻寬、雜音，請預設關閉攝影機、麥克風，若有需要再打開。
- ☑ 隨時準備好，老師會呼叫你的名字進行互動，鼓勵用麥克風提問。
- ☑ 如果有緊急事情，你必需離開線上教室，請用聊天室私訊給老師，以免老師癡癡呼喚你的名字。
- ☑ 軟體安裝請在上課前安裝完成，未完成的同學，請盡快進行安裝。

課程檔案下載

The screenshot displays the homepage of the Juei Computer Online Live website. The top navigation bar includes links for '開課查詢', '免費體驗專區', '課程總覽', '專業師資', '學員專區', '講師專區', and '最新消息'. Social media icons for Line, Facebook, and YouTube are on the right. A user is logged in as '您好!' with a '登出' button. The main banner features the text '程式語言好難學? 那是因為你還沒學過Python! (線上老師 LIVE 直播教學 · 搶先看)'. A dropdown menu is open from the '學員專區' link, listing various resources, with '課程檔案下載' highlighted. An orange callout bubble points to this menu item with the text '課程檔案下載'. The background of the banner has a technical, circuit-like aesthetic.

巨匠電腦線上真人 開課查詢 免費體驗專區 課程總覽 專業師資 學員專區 講師專區 最新消息

您好! 登出

點數卡產品兌換
APCS檢測專區
公告專區
我的課表
IT真人課程劃位
電腦分校課程劃位
外語真人課程劃位
美語分校課程劃位
取消劃位
課程檔案下載
上課權益查詢
教學平台測試
學習諮詢
常見問題
個資維護
忘記密碼
登出

程式語言好難學?
那是因為
你還沒學過Python!
(線上老師 LIVE 直播教學 · 搶先看)

巨匠電腦真人課程

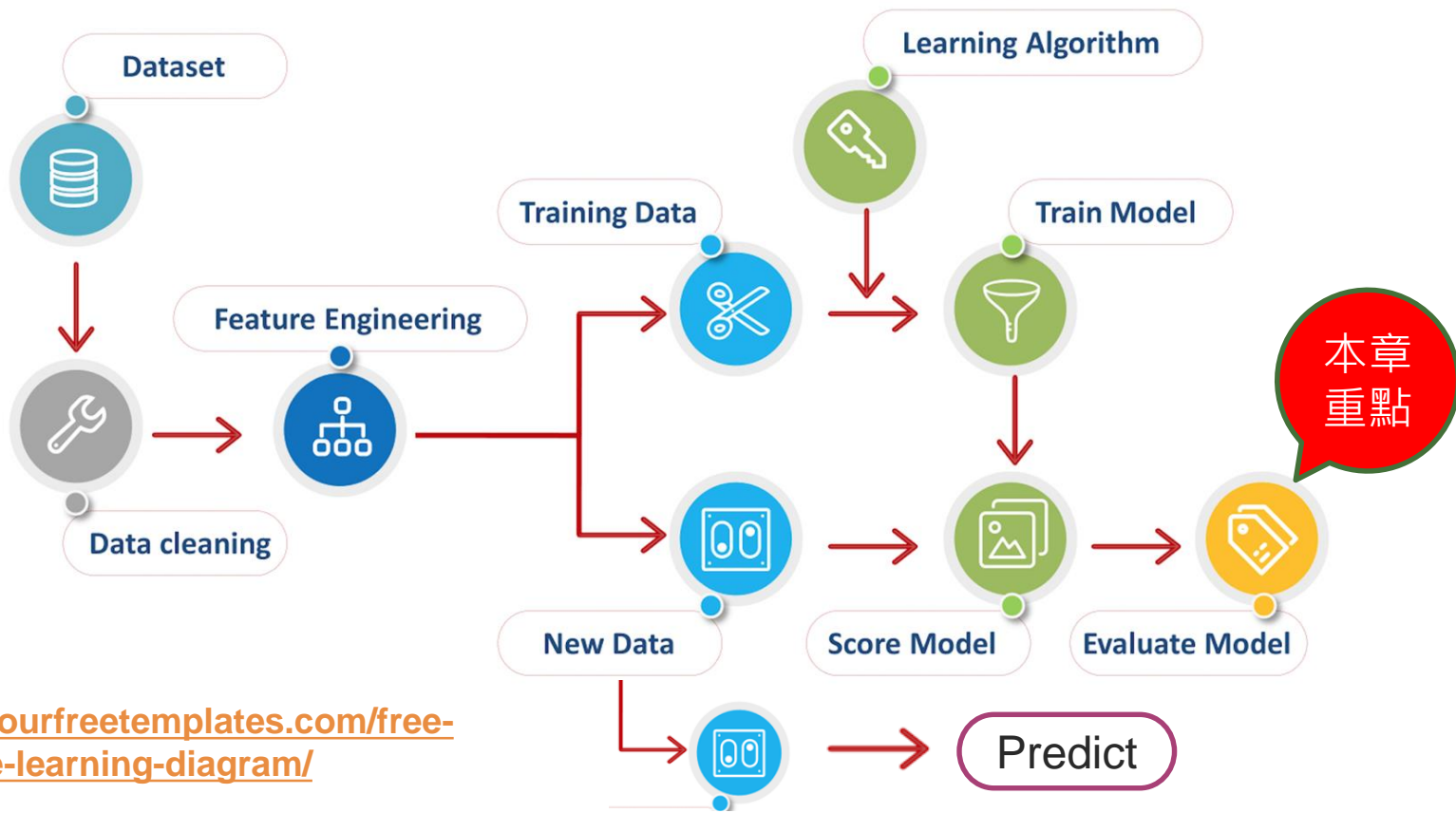
ZOOM 學員操作說明

The screenshot shows the Zoom interface with several callouts:

- 5 查看選項/共同註記/筆 (連連看)**: A dropdown menu is open, showing options: 原始大小, 請求遠端控制, 共同註記, and 退出至螢幕. An arrow points from the '共同註記' (Annotate) icon in the top toolbar to this menu.
- 2 共享螢幕 (指導演練; 點評作品)**: A callout box titled '與會者 (15)' (Participants 15) is shown. It lists participants: 張齡月 (我) and 婷婷. A '舉手' (Raise Hand) button is highlighted with a box and an arrow pointing to it.
- 1 聊天**: A callout points to the '聊天' (Chat) button in the bottom toolbar.
- 3 與會者/舉手**: A callout points to the '與會者' (Participants) button in the bottom toolbar.
- 4 解除靜音**: A callout points to the '解除靜音' (Unmute) button in the bottom toolbar.

The bottom toolbar also includes buttons for 邀請 (Invite), 與會者 (Participants), 共享螢幕 (Share Screen), 聊天 (Chat), and 錄影 (Record).

機器學習流程



<https://yourfreetemplates.com/free-machine-learning-diagram/>

課程內容

效能調校方法

- 簡化效能調校流程
- 各種效能調校的方法

效能衡量指標

- 混淆矩陣
- 準確率、精準率、召回率
- ROC、AUC

課程內容

效能調校方法

- 簡化效能調校流程
- 各種效能調校的方法

效能衡量指標

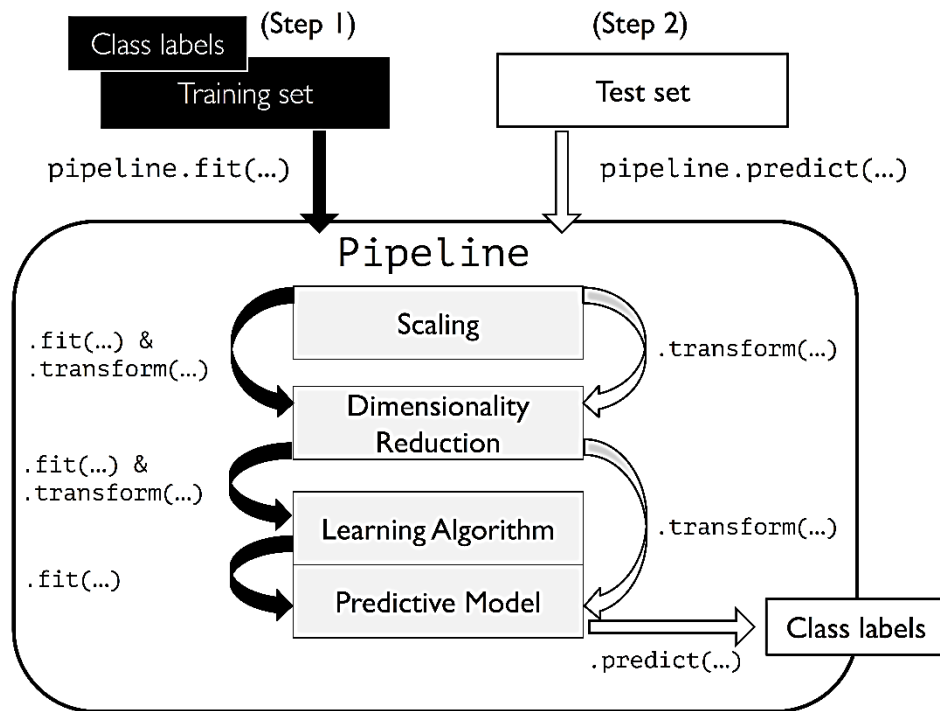
- 混淆矩陣
- 準確率、精準率、召回率
- ROC、AUC

簡化效能調校流程

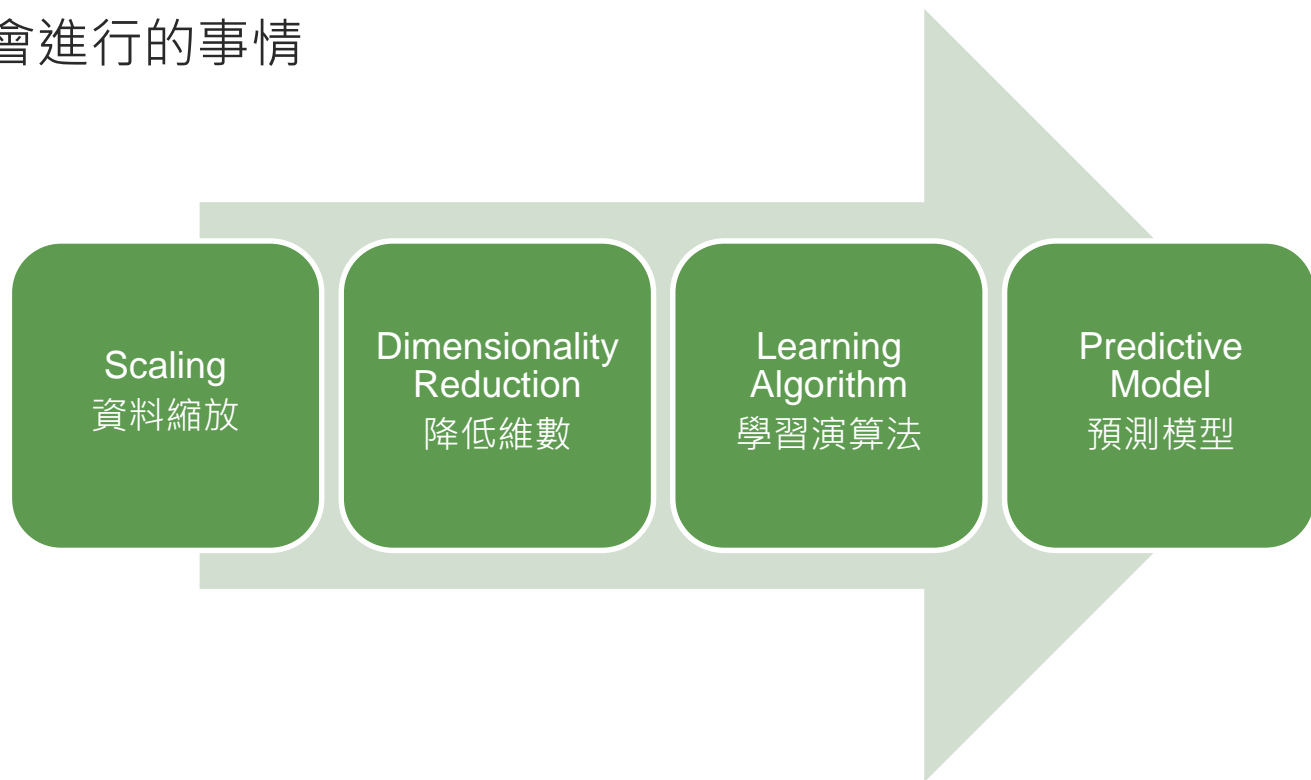
- ◆ 管線 (Pipeline) : 以管線定義工作流程，簡化程式，方便調校『超參數』 (Hyperparameter)，找出最佳效能
- ◆ 例如：下列指令提供三項工作串成一個管線，依序執行

```
pipe_lr = make_pipeline(StandardScaler(),  
                          PCA(n_components=2),  
                          LogisticRegression(random_state=1))
```


圖解



◆ 管道內會進行的事情



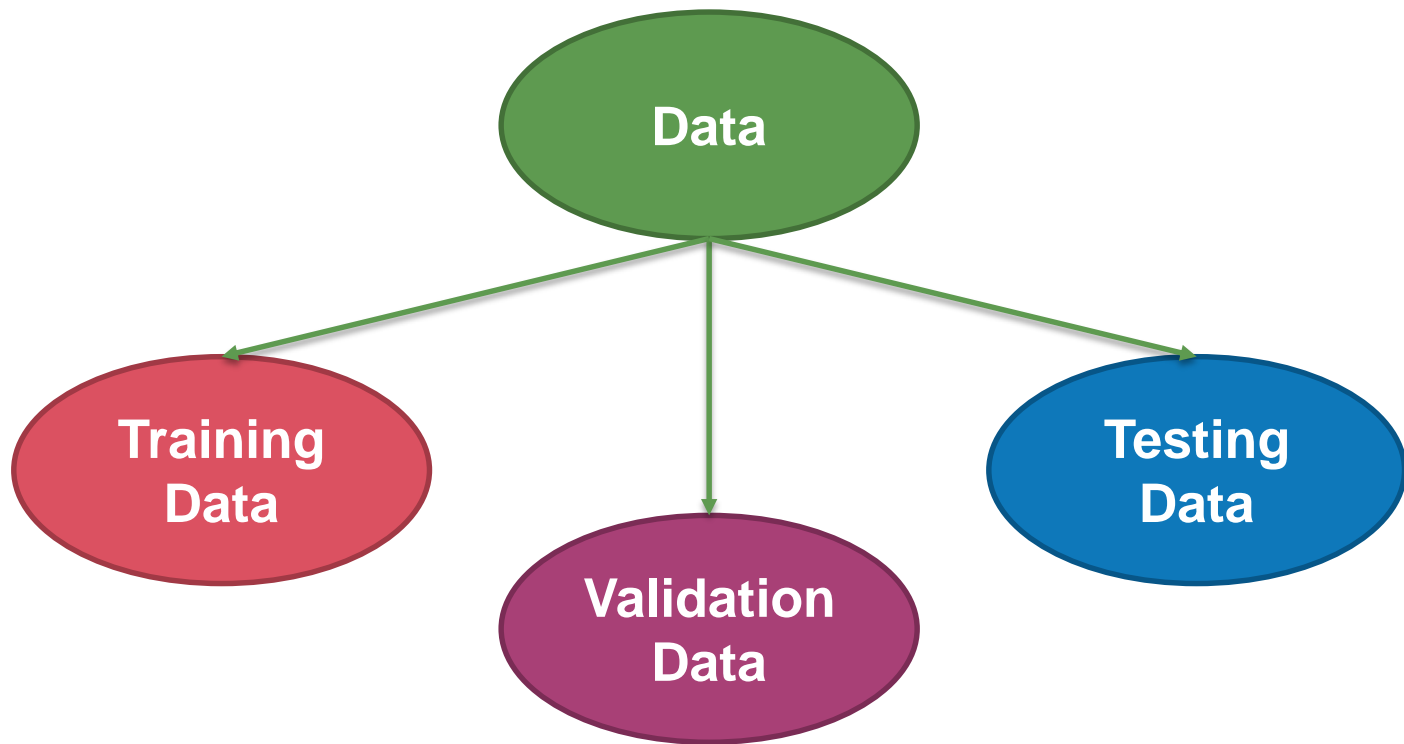
實作

- ◆ 程式碼：ch06.ipynb 第 4 ~ 9 格
 - ◇ 以 Breast Cancer Wisconsin dataset 為例
- ◆ 請問準確率 = ?
- ◆ 若 PCA 的主成分數目提高，準確率是否會隨之提高？

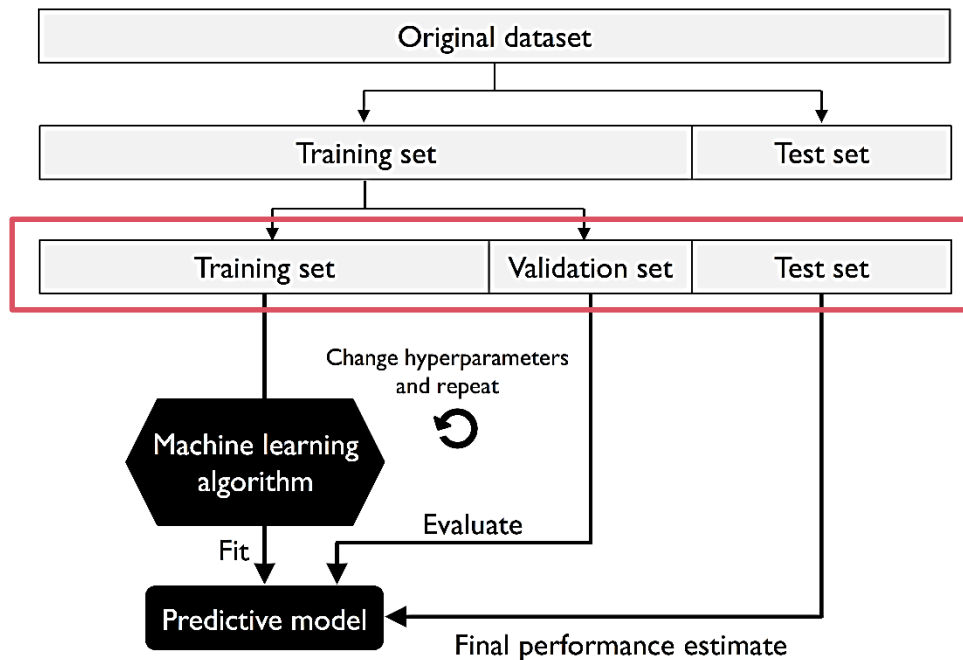
K 折交叉驗證法 (K-fold Cross-Validation)

- ◆ 正規化 (Regularization) 可以解決過度擬合 (Overfitting) 的問題。
反之，若模型太過簡單，則會產生低度擬合 (Underfitting)，偏差過大。
- ◆ 為折衷偏差 (Bias) 與變異 (Variance)，可以採用『交叉驗證法』 (Cross Validation)，分兩種：
 - ◆ 保留交叉驗證法 (Holdout Cross Validation)
 - ◆ K 折交叉驗證法 (K fold Cross Validation)

保留交叉驗證法 (Holdout Cross Validation)



用法

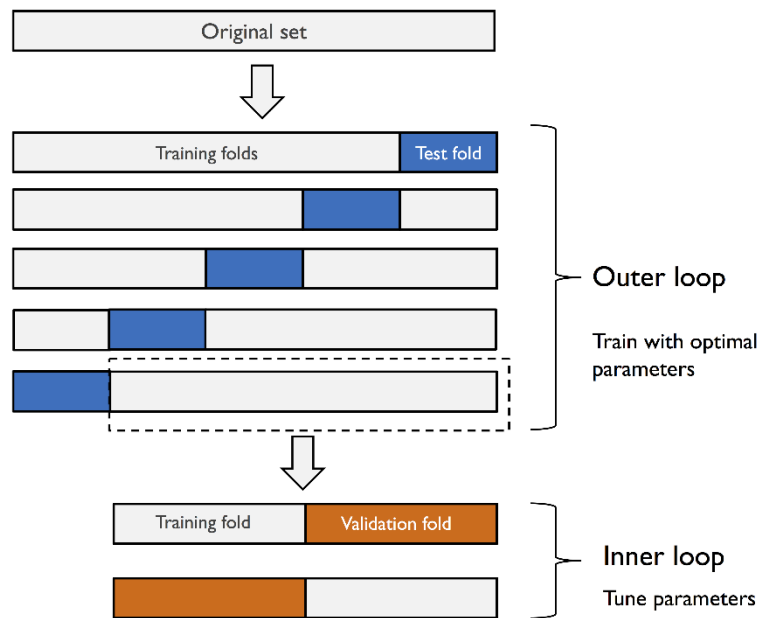


缺點

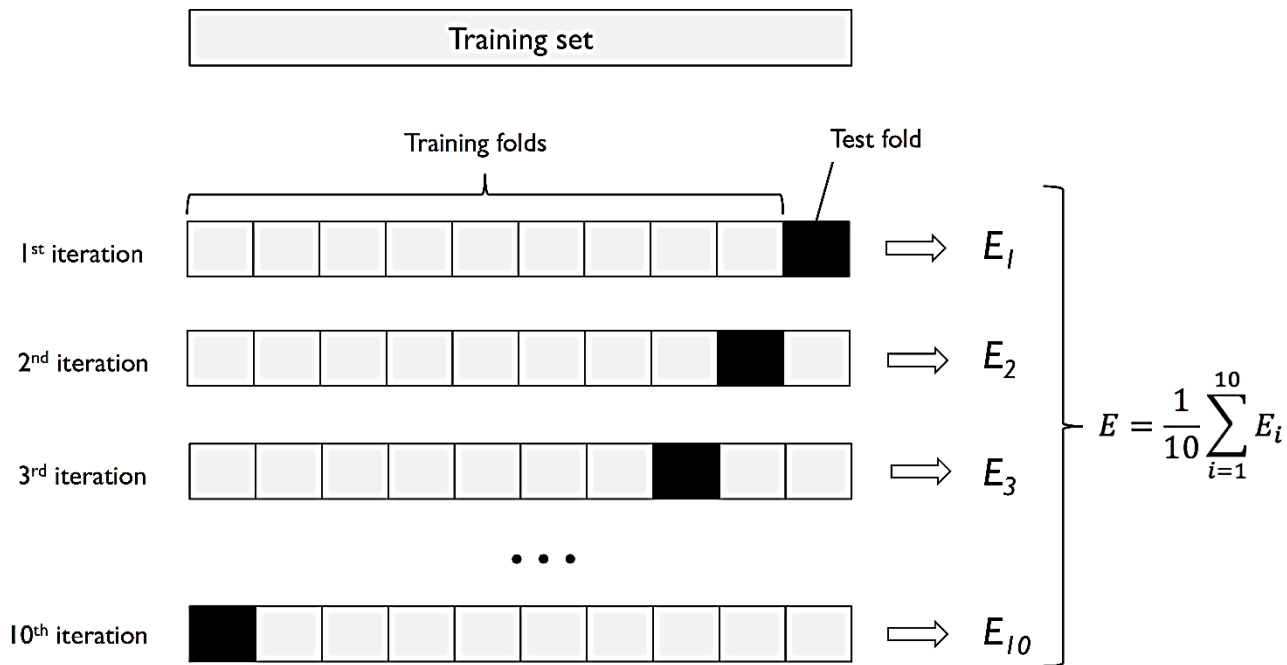
- ◆ 資料切割不當，會造成效能評估不穩定
- ◆ K 折交叉驗證法 (K fold Cross Validation) 可修正此一問題

K 折交叉驗證法 (K fold Cross Validation)

- ◆ 隨機將資料切割成 K 份
 - ◆ K-1 份當作訓練資料
 - ◆ 1 份當作測試資料
- ◆ 重複 K 次，放入模型中，會得到 K 個模型及效能評估值
- ◆ 效能評估值取平均值，即為最終模型的效能



模型效能



K 值的選擇

然而，如果我們原始的「訓練數據集」相對來說是很小的時候，那麼增加「折數」就會很有用了。如果我們增加 k 的值，則在每次的迭代中，就可以使用更多的「訓練數據」。那麼當我們計算這些模型（對個別折數據）效能評估的平均值，並以此平均值當作該模型的一般「效能估計」值的時候，這個估計值會有較低的「偏誤」。然而， k 值越大時，也會同時增加「交叉驗證演算法」的執行時間，並且會產生較大的「變異數」，因為不同迭代的訓練折中，重疊的數據比較多；折與折之間，彼此會非常類似。另一方面，如果我們有個很大的「數據集」，我們可以選擇一個相對較小 k 值，例如， $k=5$ ，並且仍然可以獲得該「模型平均效能」的正確估計，同時減少對每折中，重複「適合」與「模型評估」的計算成本。

實作

- ◆ 程式碼：ch06.ipynb 第 13 ~ 14 格
 - ◇ scikit-learn 提供 `cross_val_score` 函數，進行 K 折交叉驗證法
- ◆ Lab (275x)
 - ◇ Module5-275 / FeatureSelection.ipynb
- ◆ Lab 流程
 - ◇ Load the dataset
 - ◇ Eliminate low variance features
 - ◇ Select k best features
 - ◇ Apply nested cross validation to create model
 - ◇ Test the model

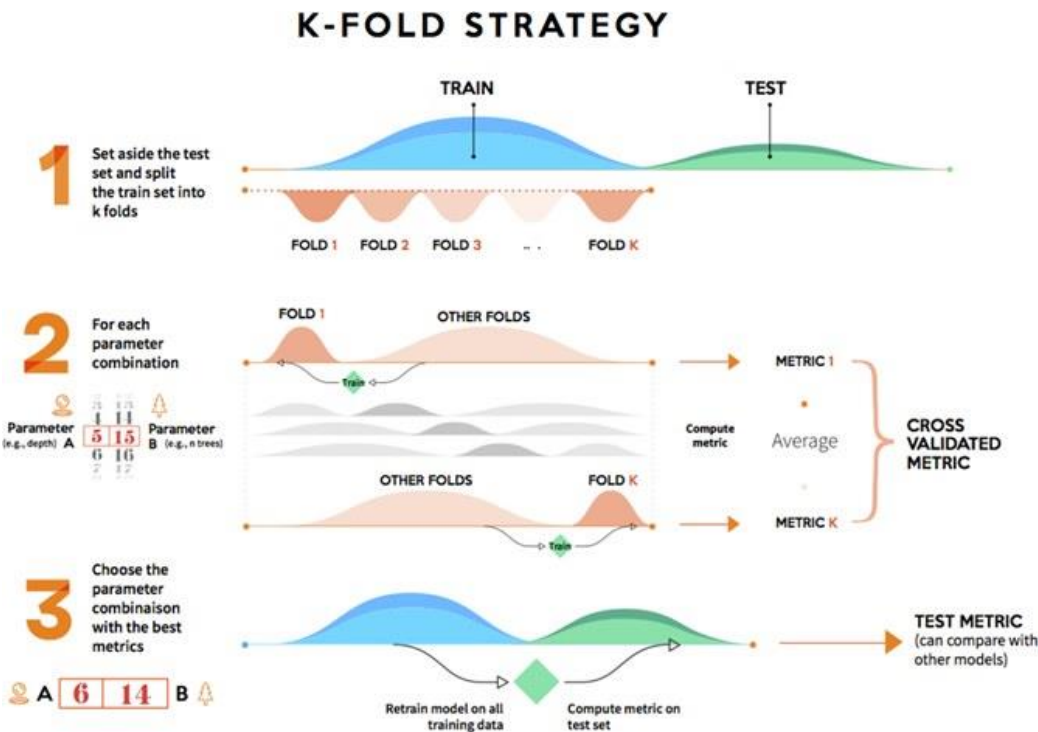
K 折交叉驗證法 (K fold Cross Validation)

1. Set aside the test set and split the train set into k folds
2. For each parameter combination
3. Choose the parameter combination with the best metrics

1. 保留測試區塊並將資料分成k個折疊
2. 執行k個區域的資料訓練，進行參數組合
3. 選擇具有最佳指標的參數組合後進行測試

- 優點： 使用整個訓練資料集而降低擬合。
- 缺點： 更高的計算成本，模型需要在驗證步驟中訓練K次（在測試步驟再加一次）。

學習曲線與驗證曲線的交叉驗證法預設為 K 折交叉驗證法

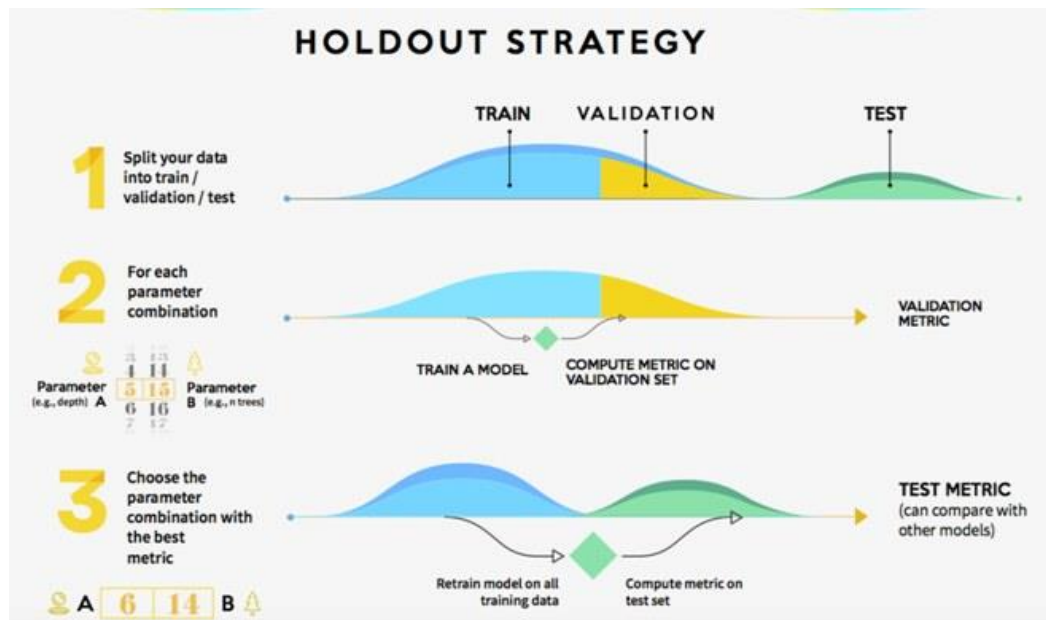


保留交叉驗證法 (Holdout Cross Validation)

- 1.Split your data into train/validation /test
- 2.For each parameter combination
- 3.Choose the parameter combination with the best metric

- 1.將您的資料拆解為訓練 / 驗證 / 測試
- 2.對於每個參數進行組合。訓練每一個模型，然後送到驗證資料集進行驗證與計算準確度。
- 3.選擇具有最佳效果的參數組合，計算所有訓練資料內的各模型的準確度，挑選一個出來進行後續的測試。

- 優點：完全獨立的資料；只需運行一次，因此計算成本較低。
- 缺點：適合較小規模的資料，較大規模資料可能出現較大差異的評估。



不均衡資料

- ◆ 不均衡的資料預測起來很簡單，以較多的那一面進行即可。
- ◆ 所以機器學習可能到最後就以預測較多的那一部分為主。
- ◆ 調整方式
 - ◇ 獲取更多資料
 - 有時候只是因為前半段時期的資料大都是某一種趨勢，也許到了後半段時間的資料走向就不同。
 - ◇ 更換判斷成果的方式
 - 通常我們會使用**準確率** accuracy 或者誤差 cost 來判斷機器學習的成果，可是當這些判斷成果方法在不均衡資料前，高的準確率與低的誤差變得不重要，建議改用別的方式。
 - 使用混淆矩陣 Confusion Matrix、精確率 Precision 與召回率 Recall，藉由計算 F1 分數方式區分。

不均衡資料

◆ 調整方式

◆ 重組資料

- 最簡單但粗暴的方法
- 第一個是複製或著合成少數部分的樣本，使樣本數與多數部分差不多。
- 第二個就是砍掉一些多數部分資料。

◆ 使用其他機器學習方法

- 如果使用神經網路等方法，無法處理不均衡資料，但如果是決策樹等就不會受到影響。

◆ 修改演算法參數

課程內容

效能調校方法

- 簡化效能調校流程
- 各種效能調校的方法

效能衡量指標

- 混淆矩陣
- 準確率、精準率、召回率
- ROC、AUC

績效衡量指標 (Performance Metrics)

		True condition			
Total population		Condition positive	Condition negative	Prevalence $= \frac{\Sigma \text{Condition positive}}{\Sigma \text{Total population}}$	Accuracy (ACC) = $\frac{\Sigma \text{True positive} + \Sigma \text{True negative}}{\Sigma \text{Total population}}$
Predicted condition	Predicted condition positive	True positive , Power	False positive , Type I error	Positive predictive value (PPV), Precision = $\frac{\Sigma \text{True positive}}{\Sigma \text{Predicted condition positive}}$	False discovery rate (FDR) = $\frac{\Sigma \text{False positive}}{\Sigma \text{Predicted condition positive}}$
	Predicted condition negative	False negative , Type II error	True negative	False omission rate (FOR) = $\frac{\Sigma \text{False negative}}{\Sigma \text{Predicted condition negative}}$	Negative predictive value (NPV) = $\frac{\Sigma \text{True negative}}{\Sigma \text{Predicted condition negative}}$
		True positive rate (TPR), Recall, Sensitivity, probability of detection $= \frac{\Sigma \text{True positive}}{\Sigma \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm $= \frac{\Sigma \text{False positive}}{\Sigma \text{Condition negative}}$	Positive likelihood ratio (LR+) $= \frac{\text{TPR}}{\text{FPR}}$	Diagnostic odds ratio (DOR) = $\frac{\text{LR+}}{\text{LR-}}$ F₁ score = $\frac{2}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}}$
		False negative rate (FNR), Miss rate $= \frac{\Sigma \text{False negative}}{\Sigma \text{Condition positive}}$	True negative rate (TNR), Specificity (SPC) $= \frac{\Sigma \text{True negative}}{\Sigma \text{Condition negative}}$	Negative likelihood ratio (LR-) $= \frac{\text{FNR}}{\text{TNR}}$	

https://en.wikipedia.org/wiki/Precision_and_recall

績效衡量指標 (Performance Metrics)

- ◆ 混淆矩陣
- ◆ 準確率 (Precision)
- ◆ 覆蓋率 (Recall , 也叫作召回率)
- ◆ 皮氏 F-分數 (Piotroski F-Score ; FSC)
- ◆ ROC 曲線、AUC 曲線

常用績效衡量指標

- ◆ 準確率 (Accuracy)
 - ◆ 精確率 (Precision)
 - ◆ 召回率 (Recall)
 - ◆ F1 score = harmonic mean (調和平均數) of precision and sensitivity = $2 / ((1/ \text{Precision}) + (1/ \text{Recall}))$, 兼顧精確率與召回率
 - ◆ 為什麼要有其他比率 ? (unbalanced samples)
 - ◇ 100個樣本，其中95個是真， 5個是假
 - 全部猜真，準確率：95%
- ➔ 事實：【假】的準確率=0%
- ◇ 美國一年有800萬個航空旅客，經統計2000 ~ 2017有19個恐怖份子搭機
 - 沒抓到半個 ➔ 準確率=99.9999999%

Sample
classification_report.py

實作

◆ 程式碼：classification_report.ipynb

◆ scikit-learn 提供 classification_report 函數

混淆矩陣 (Confusion Matrix)

		真實狀況	
		事實為真	事實為假
預測狀況	預測為真	tp	fp (Type I error)
	預測為假	fn (Type II error)	tn

◆ 以預測的角度看

- ◆ tp：預測為真，預測正確
- ◆ tn：預測為假，預測正確
- ◆ fp：預測為真，預測錯誤
- ◆ fn：預測為假，預測錯誤

- 準確率 (Accuracy) = $(tp+tn)/(tp+fp+fn+tn)$
- 精確率 (Precision) = $tp/(tp+fp)$
判斷致癌的病人正確的比率，有幾個被誤診？
- 召回率 (Recall) = $tp/(tp+fn)$
所有恐怖分子被識別的比率，有幾個沒被找到？

各種績效衡量指標

◆ Confusion matrix -Wiki

		True condition			
		Total population	Condition positive	Condition negative	
Predicted condition	Predicted condition positive	True positive, Power	False positive, Type I error	Positive predictive value (PPV), Precision = $\frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}}$	Accuracy (ACC) = $\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$
	Predicted condition negative	False negative, Type II error	True negative	False omission rate (FOR) = $\frac{\sum \text{False negative}}{\sum \text{Predicted condition negative}}$	False discovery rate (FDR) = $\frac{\sum \text{False positive}}{\sum \text{Predicted condition positive}}$
		True positive rate (TPR), Recall, Sensitivity, probability of detection = $\frac{\sum \text{True positive}}{\sum \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm = $\frac{\sum \text{False positive}}{\sum \text{Condition negative}}$	Positive likelihood ratio (LR+) = $\frac{\text{TPR}}{\text{FPR}}$	Negative predictive value (NPV) = $\frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}}$
		False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$	True negative rate (TNR), Specificity (SPC) = $\frac{\sum \text{True negative}}{\sum \text{Condition negative}}$	Negative likelihood ratio (LR-) = $\frac{\text{FNR}}{\text{TNR}}$	Diagnostic odds ratio (DOR) = $\frac{\text{LR+}}{\text{LR-}}$
				F ₁ score = $\frac{2}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}}$	

sensitivity, recall, hit rate, or true positive rate (TPR)

$$\text{TPR} = \frac{\text{TP}}{P} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

precision or positive predictive value (PPV)

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$F1 = 2 \frac{\text{PRE} \times \text{REC}}{\text{PRE} + \text{REC}}$$

實作

◆ 程式碼：confusion matrix.ipynb

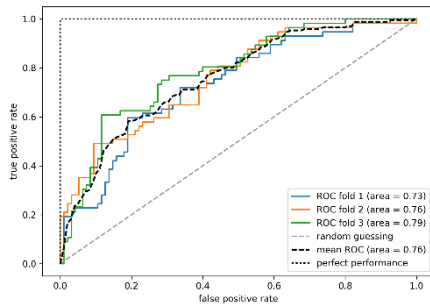
◆ 請計算

- 準確率 (Accuracy)
- 精確率 (Precision)
- 召回率 (Recall)
- F1 score

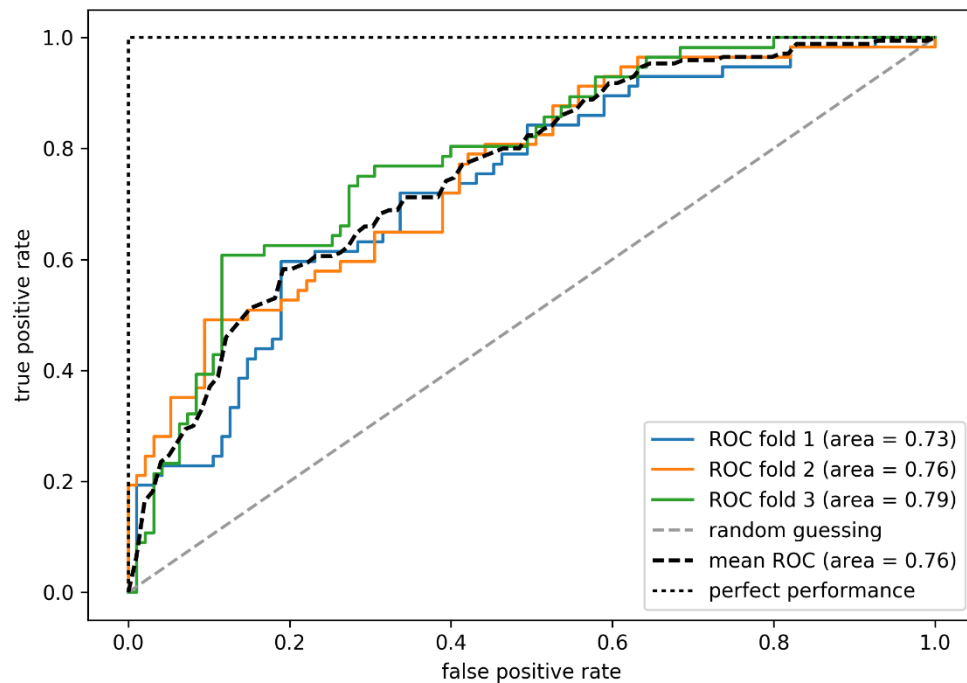
◆ 並畫出 confusion matrix

ROC / AUC

- ◆ 接收操作特徵圖 (Receiver operator characteristic ; ROC) : 在各種決策門檻 (decision threshold) 下 , 比較『真陽率』 (True Positive Rate ; TPR) 與『假陽率』 (False Positive Rate ; FPR) 間的變化
- ◆ 曲線下面積 (Area Under the Curve ; AUC) : AUC面積大小可視為模型的效能表現 , 常用於比較多個模型的效能表現
- ◆ 對角線為隨機猜測的效能
低於對角線表示模型效能比隨機猜測還差



ROC / AUC 圖解



實作

- ◆ 程式碼：ch06.ipynb 第 32 格
 - ◆ 使用 K Fold 驗證資料集各種子集合的效能表現

實作

- ◆ Lab (275x)
 - ◆ Module4-275 / Classification.ipynb
 - ◆ Module5-275 / CrossValidation.ipynb
 - ◆ Module6-275 / SupportVectorMachines.ipynb
 - ◆ Module6-275 / NeuralNetworks.ipynb

其他效能指標

◆ precision-recall curves

◆ 請參見 『[sklearn.metrics.precision_recall_curve](#)』

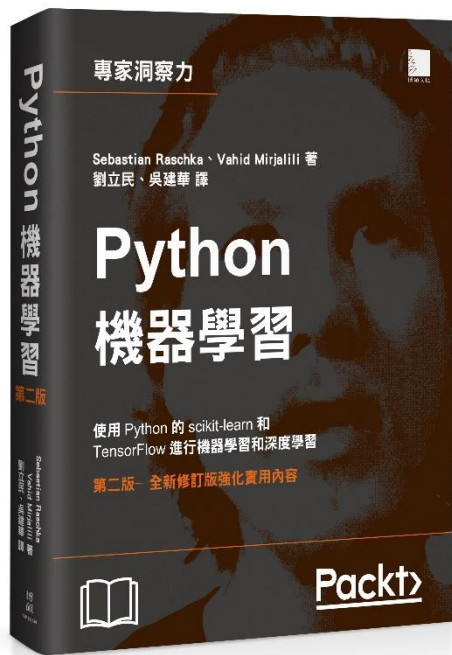
作品：鐵達尼資料集實作

- 資料集：鐵達尼 (Titanic)
- `import seaborn as sns`
- `titanic = sns.load_dataset("titanic")`
- 以 K 折交叉驗證法實作，比較採用前後之模型準確率比較

展示



參考用書



- ◆ 書名：Python機器學習（第二版）

<http://www.drmaster.com.tw/bookinfo.asp?BookID=MP11804>

- ◆ 作者：Sebastian Raschka, Vahid Mirjalili ISBN
- ◆ 譯者：劉立民、吳建華
- ◆ 出版社：博碩

問卷

<http://www.pcschoolonline.com.tw>

開課查詢

免費體驗專區

課程總覽

專業師

1

學員專區

講師專區



➤ 課程檔案下載：

學員的「上課教材」，下載檔案為壓縮檔 ([解壓縮操作步驟](#))。
如無法觀看上課教材，請安裝 [PDF閱讀軟體](#)。

公告專區

我的課表

課程劃位

取消劃位

2

課程檔案下載

自107年1月1日起，課程錄影檔由180天改為365天(含)內無限次觀看 (上課隔日18:00起)。

問
卷

上課日期	課程名稱	課程節次	教材下載		
2017/12/27 2000 ~ 2200	線上真人-ZBrush 3D動畫造型設計	18	上課教材	錄影檔	課堂問卷
2017/12/20 2000 ~ 2200	線上真人-ZBrush 3D動畫造型設計	17	上課教材	錄影檔	
2017/12/18 2000 ~ 2200	線上真人-ZBrush 3D動畫造型設計	16	上課教材	錄影檔	



巨匠線上真人

www.pcschoolonline.com.tw