



D17：特徵工程簡介



PDF 下載

全螢幕

Sample Code & 作業內容

請參考範例程式碼 `Day_017_Introduction_of_Feature_Engineering.ipynb`: 房價預測
請參閱作業範例`Day_017_HW.ipynb`: 鐵達尼生存預測

作業1: A~E五個程式區塊中, 哪一塊是特徵工程?
作業2: 對照程式區塊 B 與 C 的結果, 請問那些欄位屬於「類別型欄位」? (回答欄位英文名稱即可)
作業3: 續上題, 請問哪個欄位是「目標值」?

作業請提交`Day_017_HW.ipynb`

提醒: 作業與範例程式碼會使用到的資料, 請於檢視範例內data下載 **1.house_test** **2.house_train** 3.titanic_test
4.titanic_train

檢視範例

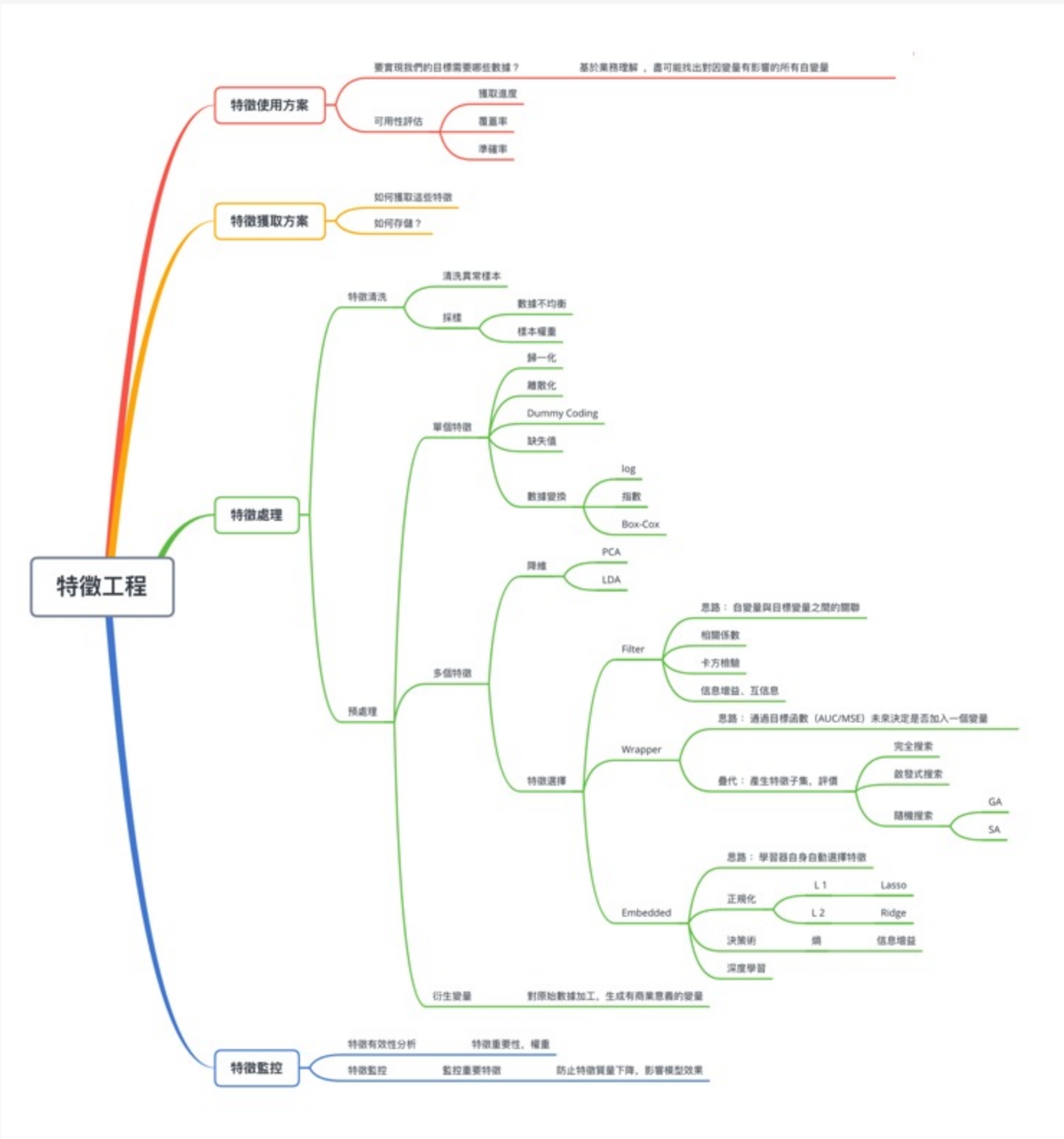
參考資料

課後補充:

知乎-特徵工程到底是什麼

網頁連結

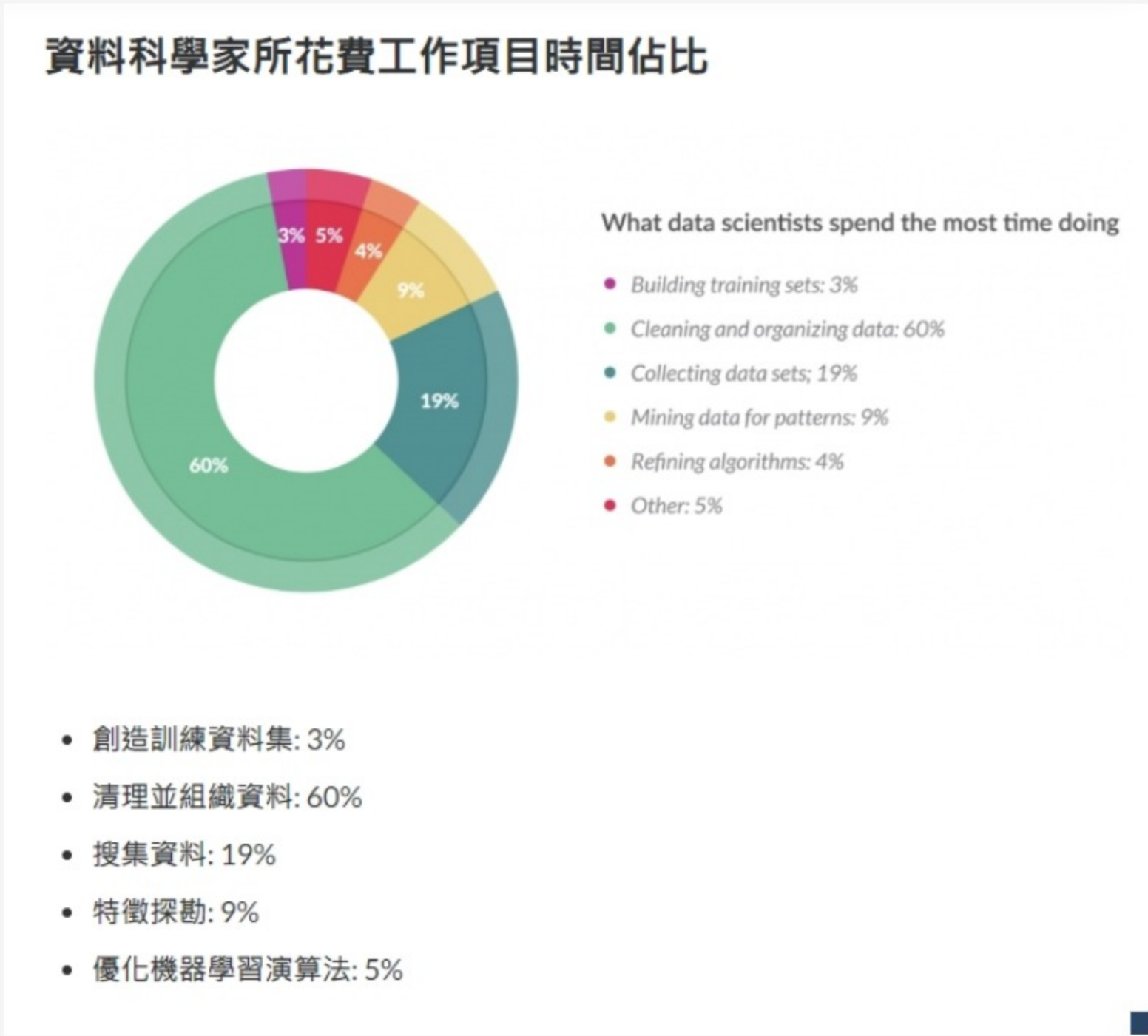
本文重點為右圖, 主要是希望同學大致知道特徵工程大致包含哪些部分, 若對細節有興趣, 還可以從這篇中了解一些概念其中一部分的內容, 會在後面的課程中說明並練習, 詳情請參閱百日馬拉松課綱。



痞客邦-2019鐵人賽：為什麼特徵工程很重要

網頁連結

本文主要在描述現實中資料科學工作的時間比重(下圖), 其中大部分的時間在於資料清理, 少部分為特徵探勘, 雖然這兩部份都是特徵工程, 但在學習階段與實務階段, 比重卻有著天壤之別。



提交作業

請將你的作業上傳至 Github，並貼上該網址，完成作業提交

<https://github.com/>

確定提交

如何提交

到 Cupoy 問答社區提問，讓教練群回答你的疑難雜症

向專家提問

如何提問