

# D54 : clustering 1 非監督式機器學習簡介



簡報閱讀



範例與作業

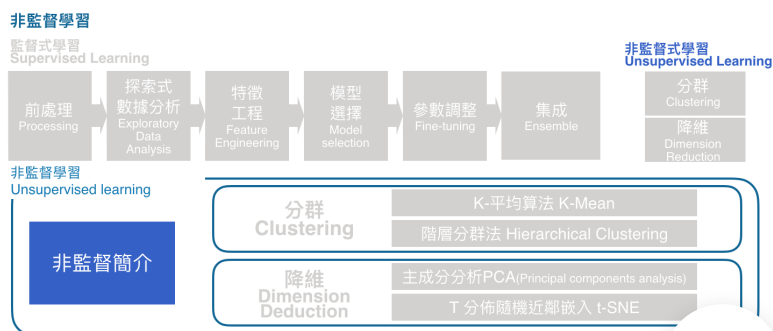


問題討論

## 非監督式機器學習簡介



## 知識地圖



## 本日知識點目標

## 本日知識點目標

- 瞭解非監督式學習 (unsupervised learning) 相關技術概要。
- 瞭解非監督式學習的應用場景。

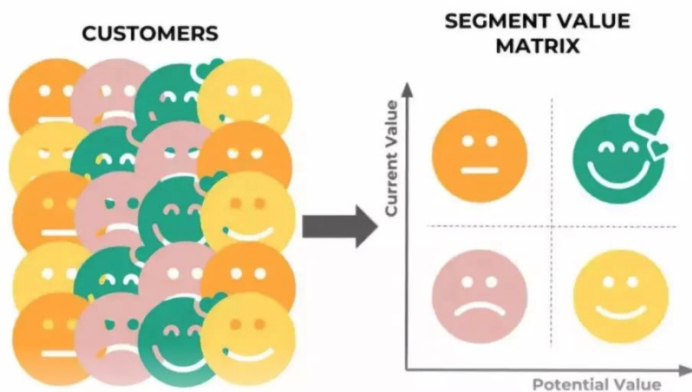
## 什麼是非監督式學習？

非監督學習允許我們在對結果無法預知時接近問題。非監督學習演算法只基於輸入資料找出模式。當我們無法確定尋找內容，或無標記 (y) 資料時，通常會用這個演算法，幫助我們了解資料模式。

## 應用案例

### (一)：客戶分群

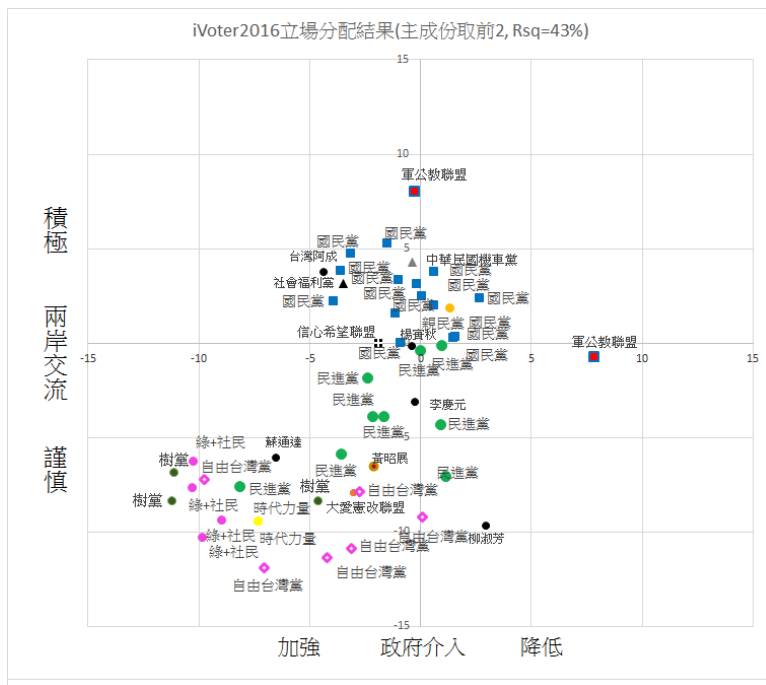
在資料沒有任何標記，或是問題還沒定義清楚前，可用分群的方式幫助理清資料特性。



### (二)：特徵抽象化

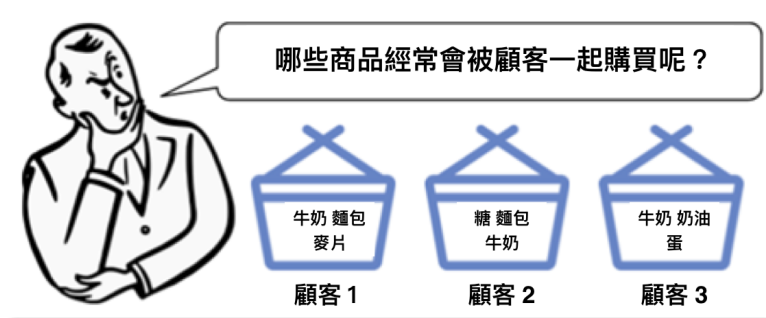


訊，組合成新的特徵。



### (三)：購物籃分析

資料探勘的經典案例，適用於線下或線上零售的商品組合推薦。



(四)：非結構化資料分析

非結構化資料如文字、影像等，可以藉由一些非監督式學習的技術，幫助呈現及描述資料。

- 主題模型 (topic model)



## 非監督學習算法概要

- 聚類分析：尋找資料的隱藏模式
- 降低維度：特徵數太大且特徵間相關性高，以此方式縮減特徵維度
- 其他：關聯法則 (購物籃分析)、異常值偵測、探索性資料分析等

## 重要知識點複習

- 在不清楚資料特性、問題定義、沒有標記的情況下，非監督式學習技術可以幫助我們理清資料脈絡
- 特徵數太龐大的情況下，非監督式學習可以幫助概念抽象化，用更簡潔的特徵描述資料
- 非監督式學習以聚類算法及降低維度算法為主，本課程也以這兩門技術進行探究

## 延伸閱讀



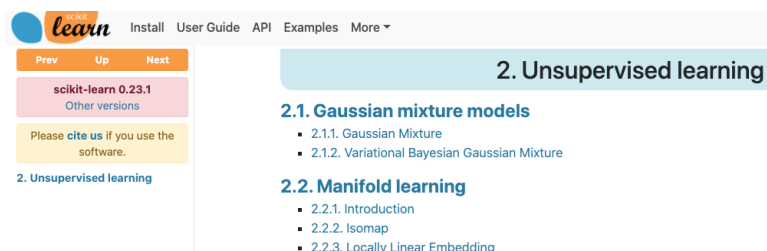
## 閱讀

### • Unsupervised learning : PCA (英文)

$$\begin{aligned}
 z_1 &= w^1 \cdot x \\
 \text{PCA} \quad \bar{z}_1 &= \sum z_1 = \sum w^1 \cdot x = w^1 \cdot \sum x = w^1 \cdot \bar{x} \\
 \text{Var}(z_1) &= \sum_{z_1} (z_1 - \bar{z}_1)^2 \\
 &= \sum_x (w^1 \cdot x - w^1 \cdot \bar{x})^2 \\
 &= \sum (w^1 \cdot (x - \bar{x}))^2
 \end{aligned}
 \quad
 \begin{aligned}
 (a \cdot b)^2 &= (a^T b)^2 = a^T b a^T b \\
 &= a^T b (a^T b)^T
 \end{aligned}$$



### • Scikit-learn unsupervised learning (英文)



## 解題時間



Sample Code & 作業  
開始解題



非監督式機器學習簡介 >

知識地圖 >

本日知識點目標 >

什麼是非監督式學習？ >

應用案例 >

非監督學習算法概要 >

重要知識點複習 >

延伸閱讀 >

解題時間 >

[下一步：閱讀範例與完成作業](#)

