eupay

Day 62 非監督式機器學習

t-sne 觀察:

分群與流形燙原





周俊川/陳明佑

知識地圖非監督學習



非監督學習

監督式學習 Supervised Learning

前處理 Processing 探索式 數據分析 Exploratory Data Analysis

特徵 工程 Feature Engineering 模型 選擇 Model selection

參數調整 Fine-tuning

集成 Ensemble 非監督式學習 Unsupervised Learning

> 分群 Clustering

降維 Dimension Reduction

非監督學習

Unsupervised learning

非監督簡介

分群 Clustering K-平均算法 K-Mean

階層分群法 Hierarchical Clustering

降維 Dimension Deduction 主成分分析PCA(Principal components analysis)

T分佈隨機近鄰嵌入 t-SNE



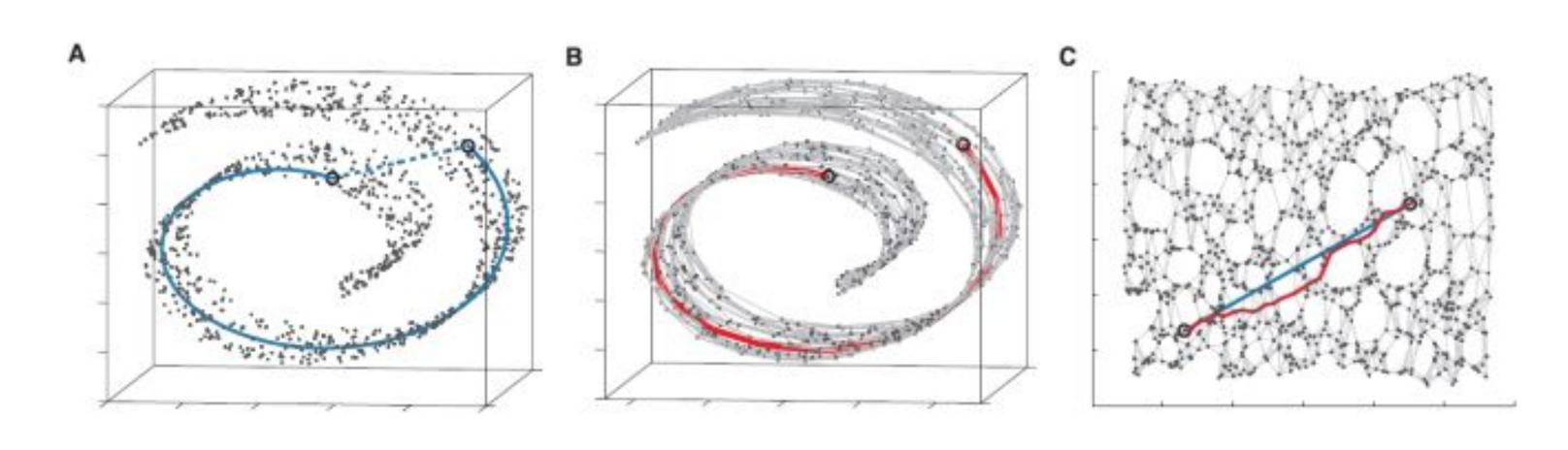
本目知識點目標

- 非監督問題類類型中,流形還原意義是什麼?
- ◎ 除了 t-sne 外,還有哪些可以用在流形還原上的方法

什麼是流形還原



- 流形還原的意義
 - 流形還原就是將高維度上相近的點,對應到低維度上相近的點,沒有資料點的地方不列入考量範圍
 - 簡單的說,如果資料結構像瑞士捲一樣,那麼流形還原就是把它攤開鋪平(流形還原資料集的其中一種,就是叫做瑞士捲-Swiss Roll)



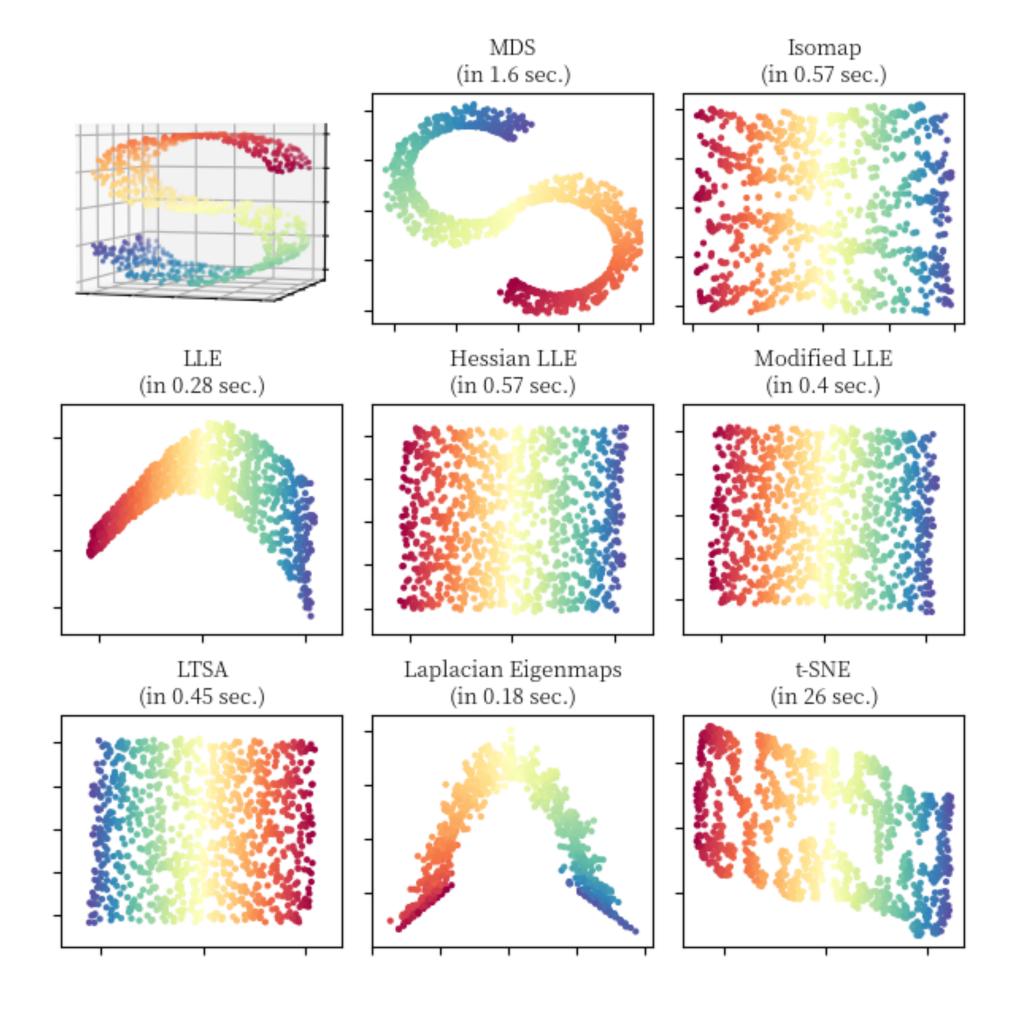
圖片來源:知乎

流形還原-其他模型(1/2)



• 流形還原的其他模型

其實流形還原的模型相當多種,只是應用上較少,常見的除了 t-sne 之外,還有 Isomap / LLE / MDS 等等,因為實用度不高,之後的課程中我們也不會教,因此只在此展示幾種流形還原的結果圖

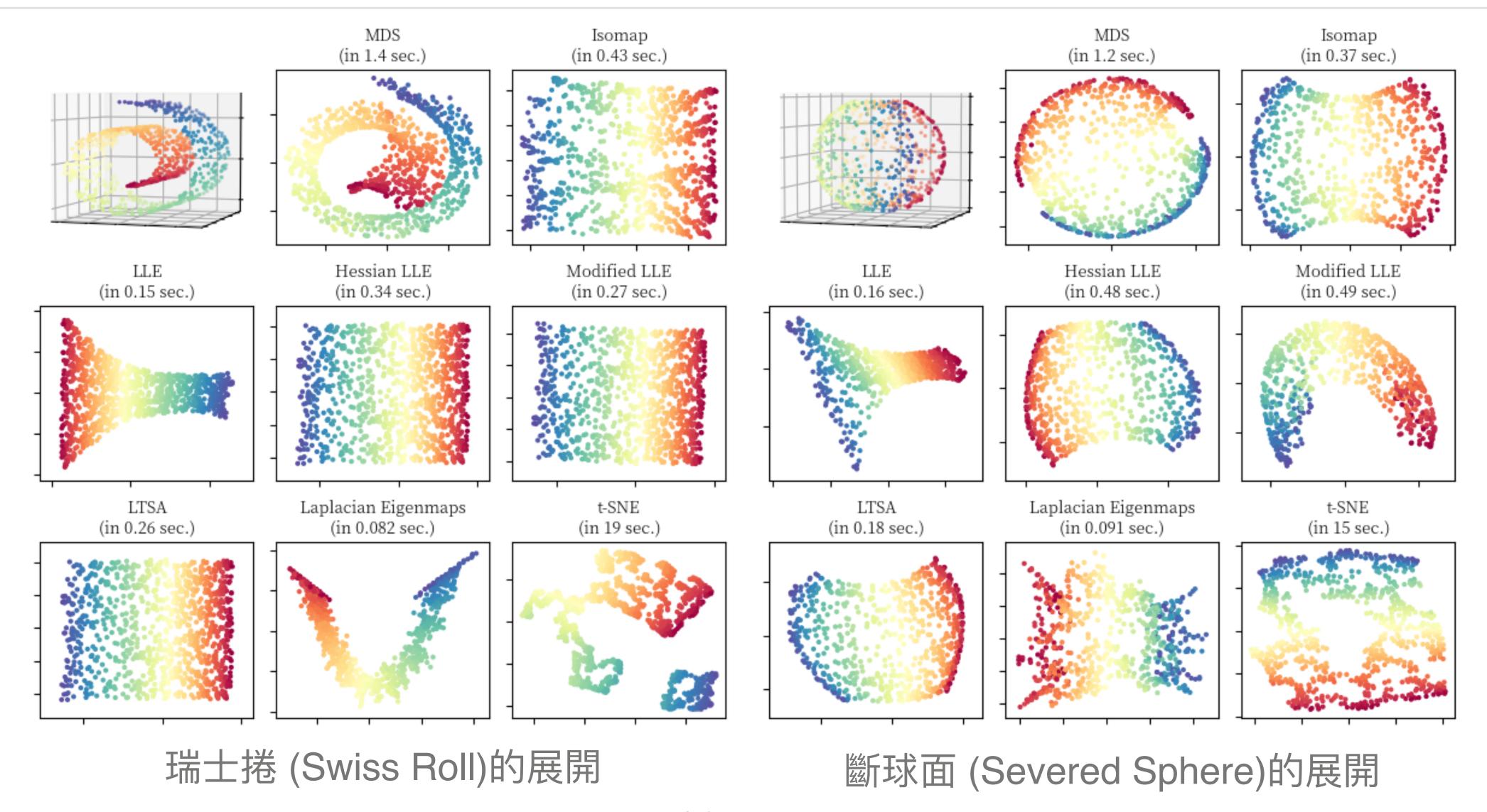


S型 (S Curve)的展開

圖片來源:leovan

流形還原-其他模型(2/2)

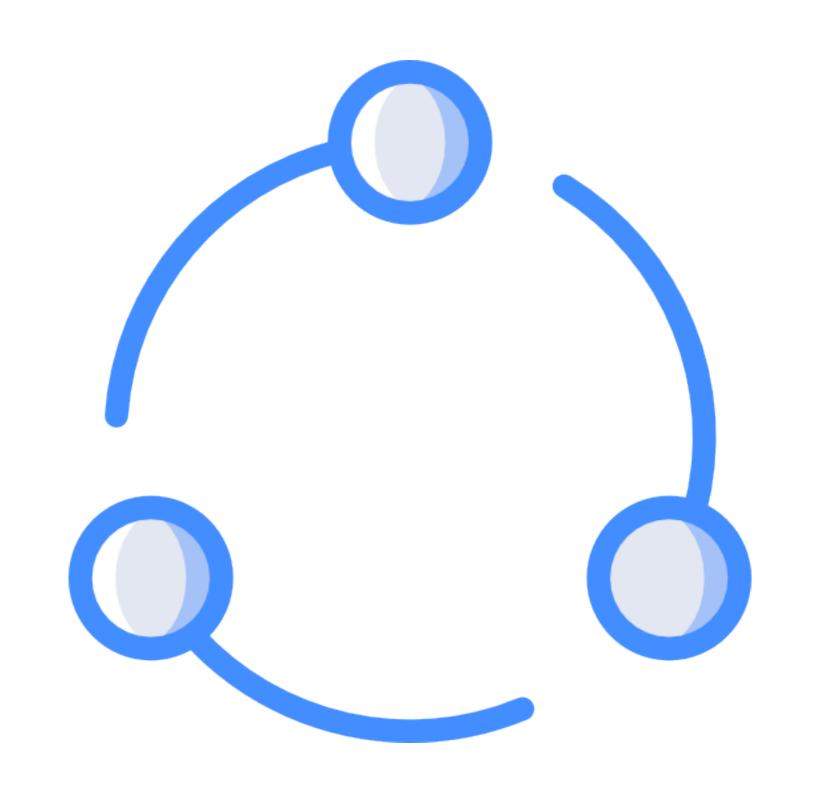




圖片來源:<u>leovan</u>

重要知識點複習





流形還原就是在高維度到低維度的對應中,盡量 保持資料點之間的遠近關係,沒有資料點的地 方,就不列入考量範圍

● 除了 t-sne 外,較常見的流形還原還有 Isomap 與 LLE (Locally Linear Embedding) 等工具

角程題時間 Coding Time

請跳出PDF至官網Sample Code&作業 開始解題

