

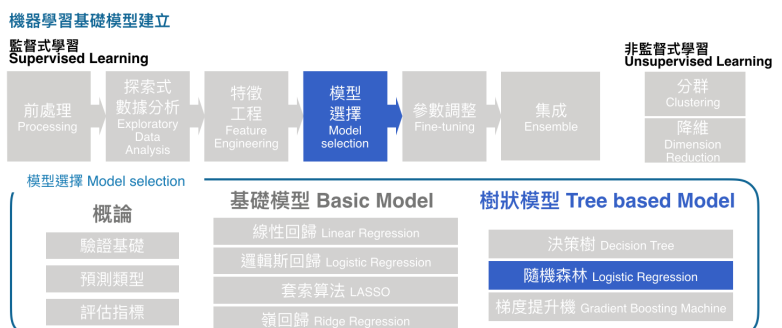
D43 : tree based model - 隨機森林 (Random Forest) 介紹

[簡報閱讀](#)[範例與作業](#)[問題討論](#)[隨機森林](#)

隨機森林

[知識地圖](#)[本日知識點目標](#)[決策樹的缺點](#)[集成模型 - 隨機森林 \(Random Forest\)](#)

知識地圖

[隨機森林 \(Random Forest\), 隨機在哪？](#)[常見問題](#)[解題時間](#)

本日知識點目標

本日知識點目標

- 了解隨機森林的基本原理與架構
- 決策樹與隨機森林的差異
- 隨機森林如何彌補了決策樹的缺點

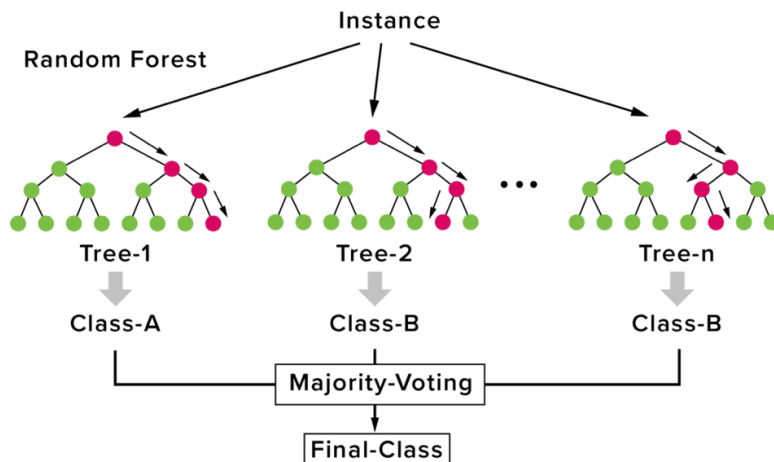
決策樹的缺點

- 若不對決策樹進行限制 (樹深度、葉子上至少要有多少樣本等)，決策樹非常容易 Over-fitting
- 為了解決決策樹的缺點，後續發展出了隨機森林的概念，以決策樹為基底延伸出的模型



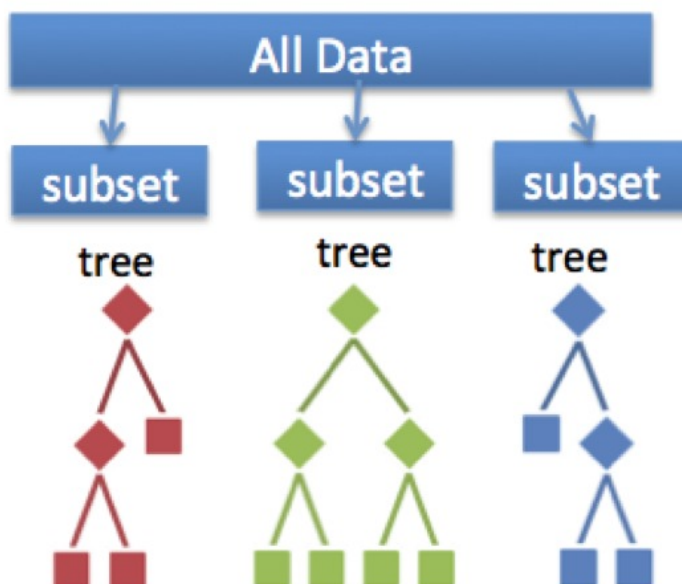
集成模型 - 隨機森林 (Random Forest)

Random Forest Simplified



隨機森林 (Random Forest), 隨機在哪？

- 決策樹生成時，是考慮所有資料與特徵來做切分的
- 而隨機森林的每一棵樹在生成過程中，都是隨機使用一部份的訓練資料與特徵 代表每棵樹都是用隨機的資料訓練而成的



常見問題



Q：隨機森林的模型準確率會比決策樹來的差嗎？

A：若隨機森林中樹的數量太少，造成嚴重的 Overfit，是有可能會比較差。但如果都是用預設的參數，實務上不太會有隨機森林比決策樹差的情形，要特別注意程式碼是否有誤

解題時間



Sample Code & 作業
開始解題



