

# Day 22

## Stacking

[全民瘋AI系列]



第12屆 iT邦幫忙 鐵人賽

# Day 22 學習目標

---

01

## 了解 Stacking 方法

堆疊法的學習機制為何？

02

## 利用 Stacking 實作分類器

透過 Stacking Classifier 訓練一個基於預測分類器

# Part 1

## Stacking

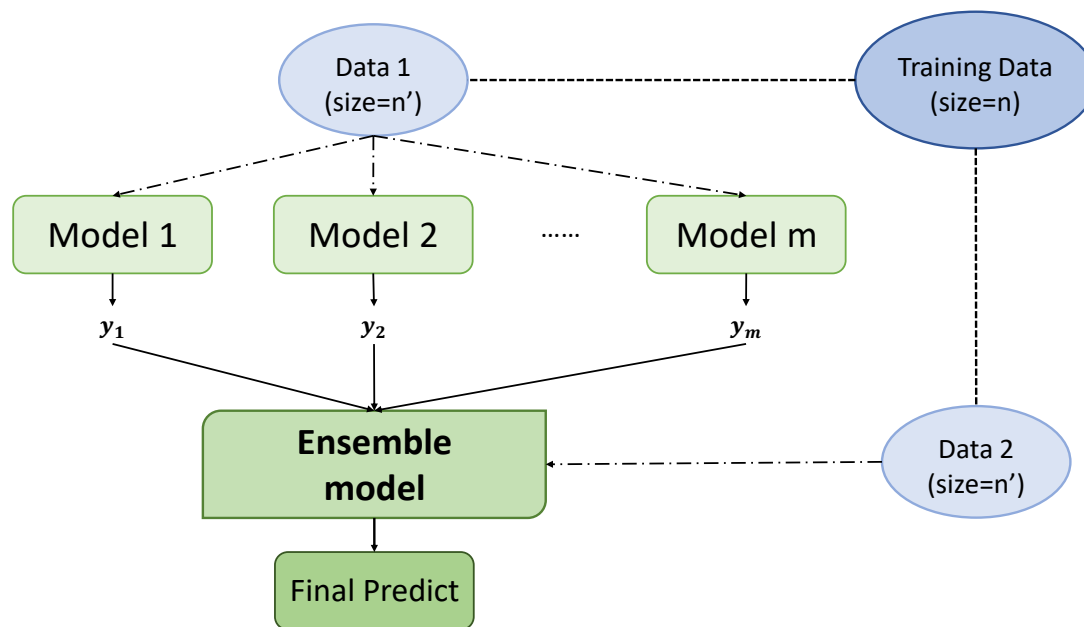
### 觀念講解



第12屆 iT邦幫忙 鐵人賽

# Stacking 堆疊法

- 產生出  $m$  個不同演算法的 base learners 模型
- base learner 模型彼此間並無互相關連
- 把  $m$  個 base learner 的輸出當成新的模型的輸入，另外再訓練一個模型



# Ensemble learning 小結

Ensemble learning 有三種類型第一種為 Bagging，第二種為 Boosting，第三種為 Stacking。

- **Bagging 採用重新採樣技巧**
  - Random forest
- **Boosting 一次產生一個新的模型，並補足前幾棵樹表現不好的部分**
  - AdaBoost
  - Gradient Boosting
- **Stacking 結合許多弱學習器**



# Part 2

## Stacking (分類器) 程式實作

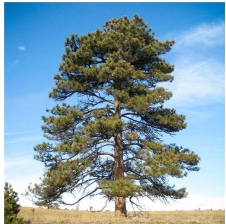


第12屆 iT邦幫忙 鐵人賽

# 森種類預測

總共約有一萬四千多筆森林的資料來做分類問題(共7類)。

Spruce/Fir	Lodgepole Pine	Ponderosa Pine	Cottonwood/ Willow	Aspen	Douglas-fir	Krummholz
雲杉木	海灘松	西黃松	三角葉楊	白楊樹	花旗松	高山矮曲林
1	2	3	4	5	6	7



# //// 特徵一覽

## 輸入

Elevation (海拔)	Aspect (方位角)	Slope (坡度)	Horizontal_Distance_To_Hydrology (水源水平距離)	Hillshade_Noon (夏至中午丘索引)
Vertical_Distance_To_Hydrology (水源垂直距離)		Horizontal_Distance_To_Roadways (最近道路水平距離)		Hillshade_9am (夏至上午9點山丘索引)
Wilderness_Area (指定荒野地區種類)*4	Soil_Type (指定土壤種類)*40	Horizontal_Distance_To_Fire_Points (水平火源距離)		Hillshade_3pm (夏至下午3點山丘索引)

## 輸出

Cover\_Type  
(樹木種類)



共有54個輸入特徵，1個輸出





# Thanks

PRESENTED BY 10程式中



第12屆 iT邦幫忙 鐵人賽