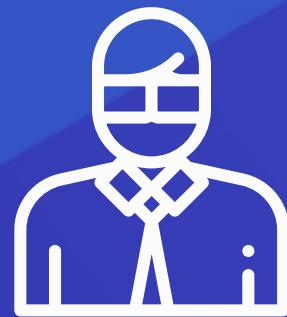




Day 37

機器學習

Regression 模型



楊証琨

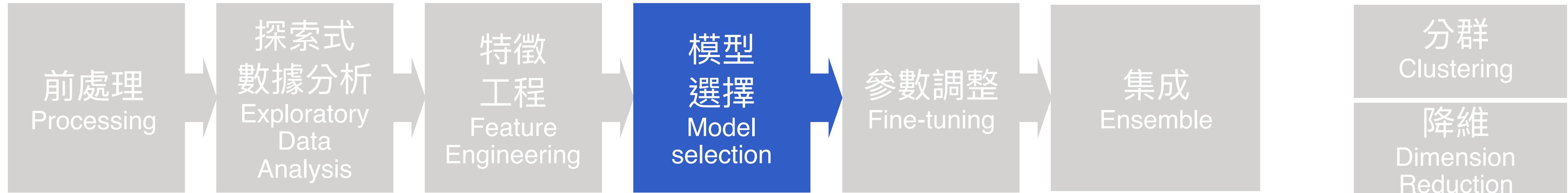
出題教練



知識地圖 機器學習- 模型選擇 - 線性模型 - 線性回歸/邏輯斯回歸

機器學習基礎模型建立

監督式學習 Supervised Learning



模型選擇 Model selection

概論

驗證基礎

預測類型

評估指標

基礎模型 Basic Model

線性回歸 Linear Regression

邏輯斯回歸 Logistic Regression

套索算法 LASSO

嶺回歸 Ridge Regression

樹狀模型 Tree based Model

決策樹 Decision Tree

隨機森林 Random Forest

梯度提升機 Gradient Boosting Machine

本日知識點目標

- 了解線性回歸與羅吉斯回歸的基本定義
- 線性回歸與羅吉斯回歸的差異
- 回歸模型使用上的限制

線性回歸模型 Linear Regression

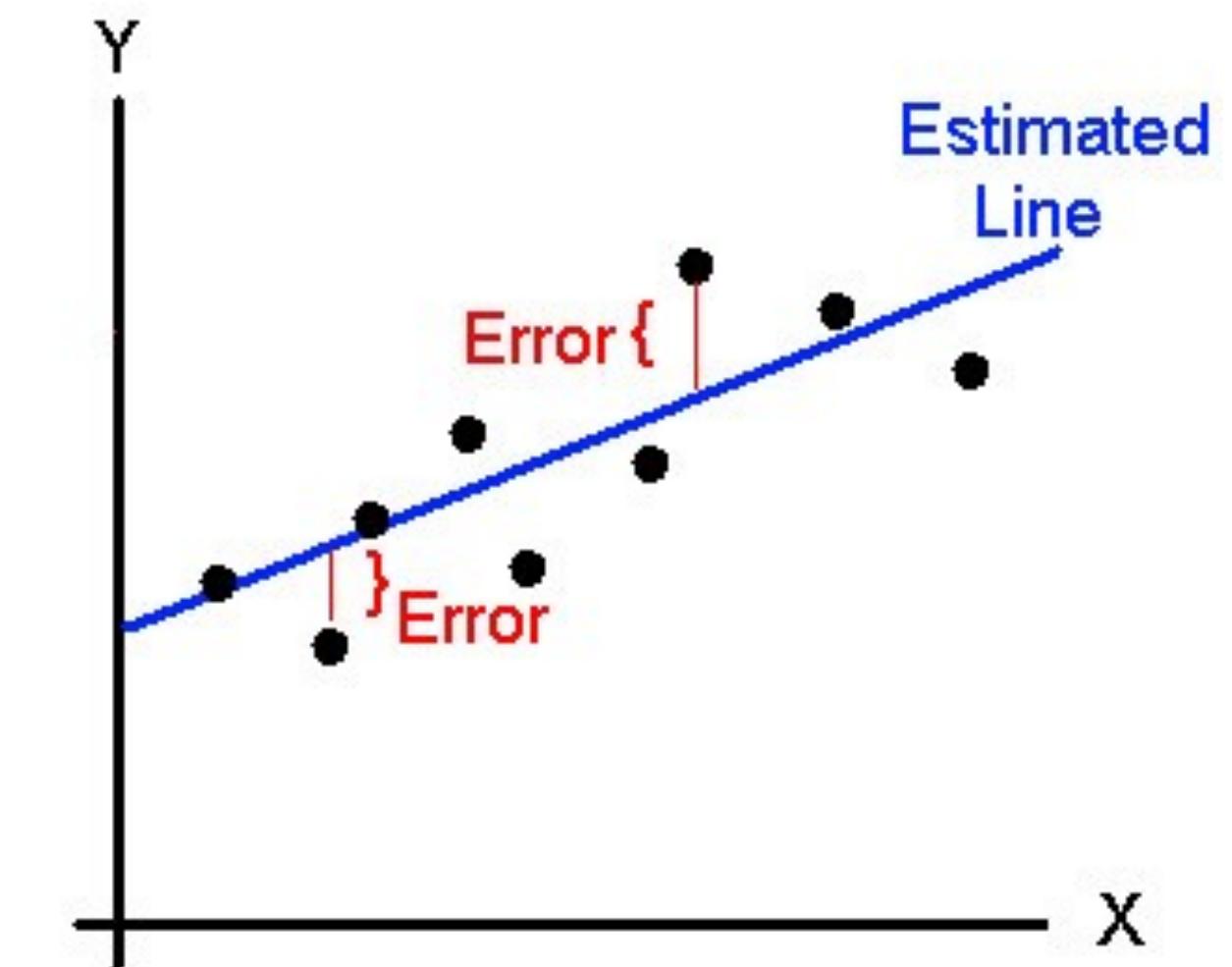
- 簡單常見的線性模型，可使用於回歸問題。

訓練速度非常快，但須注意資料共線性、資料標準化等限制

- 通常可作為 baseline 模型作為參考點

$$\hat{Y}_i = b_0 + b_1 X_i$$

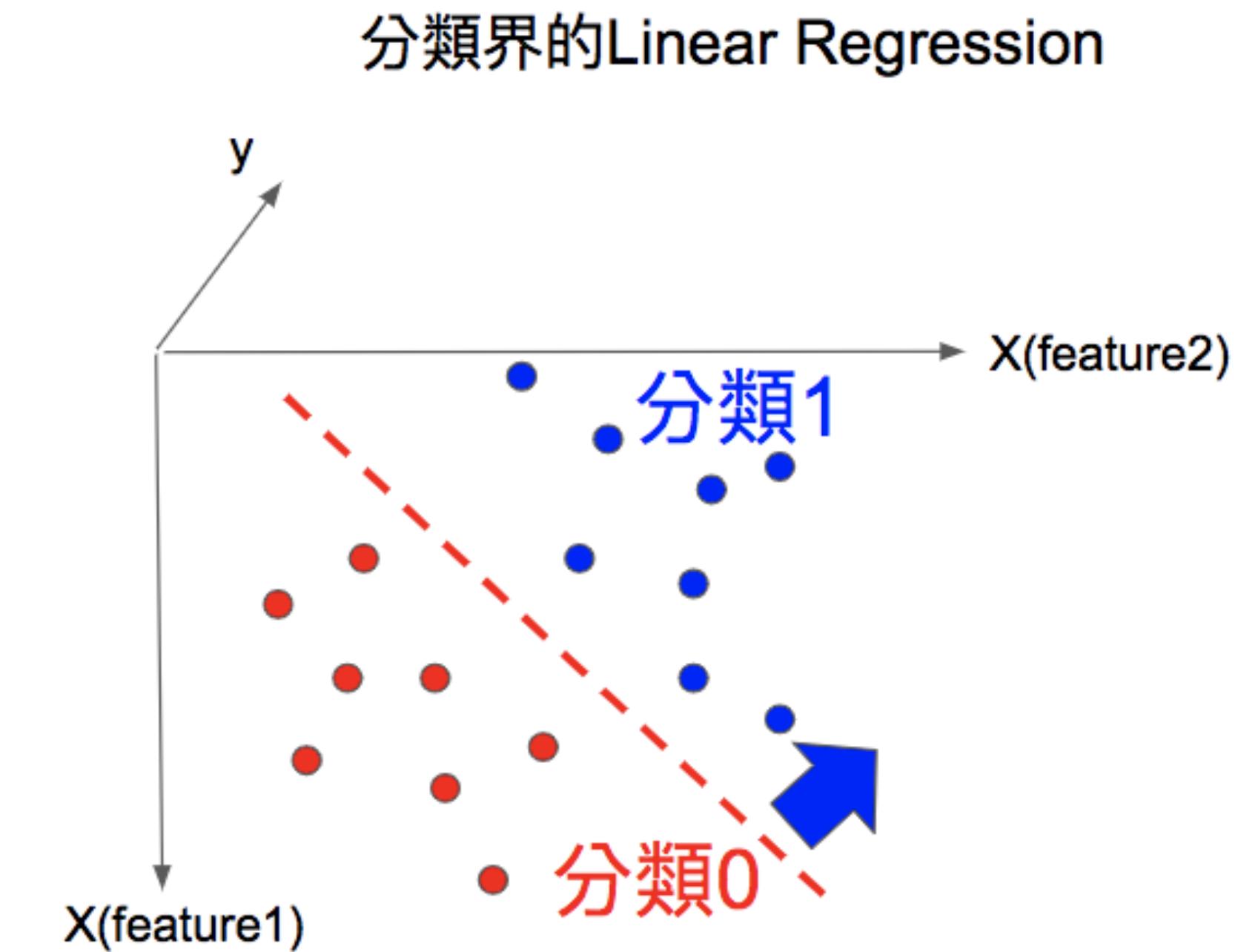
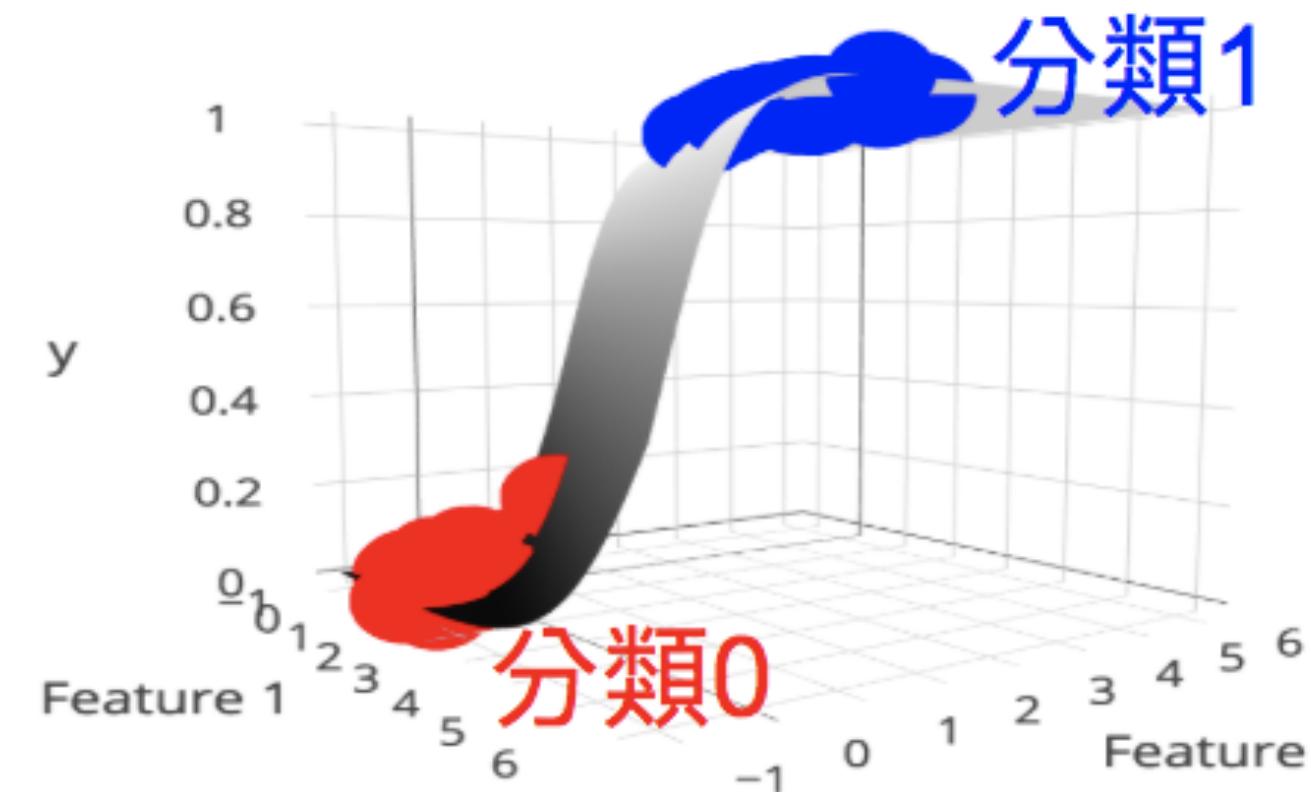
Estimated (or predicted) Y value for observation i
 Estimate of the regression intercept
 Estimate of the regression slope
 Value of X for observation i



羅吉斯回歸 Logistics Regression

- 雖然有回歸兩個字，但 Logistics 是分類模型
- 將線性回歸的結果，加上 Sigmoid 函數，將預測值限制在 0 ~ 1 之間，即為預測機率值。

$$P(y|x) = \frac{1}{1 + e^{-yw^T x}}$$



常見問題



Q：這些模型的數學式子都很多，一定要完全看懂才繼續往下嗎？不會推導可以嗎？

A：回歸模型是機器學習模型中的基礎，雖然實務上應用的機會不多（因為模型過於簡單），但是之後更複雜的模型都是基於回歸模型做加強，所以對基本原理有一定的了解會比較好。畢竟 Python 使用線性回歸只要一行程式碼，但是不了解原理，就會陷入當遇到錯誤不知如何修正的情況。另外有些數據公司的資料科學家職缺，在面試時很喜歡針對數學原理來問，或是請你現場推導，建議盡量熟悉模型的原理。



延伸 閱讀

- 超人氣 Stanford 教授 Andrew Ng 教你 Linear regression (強烈推薦觀看)
- Logistic regression 數學原理



解題時間

It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

