Untitled

林茂廷

6/10/2019

## 分類問題

對學生i，給定預測二一特徵，要如何產生預測結果，不同分類工具可從「模型」、「Training方式」、及「預測方式」來區分說明。

## 模型

### Logistic regression

Model：

Training：

預測法：

### 決策樹

Model：一連串的樹狀決策結構，從根結點出發開始一連串由單一特徵變數所形成的「是/否」分枝，一直下去直到判斷出所屬分類。

example:

“過去是否有二一”

* 是：預測會二一
* 不：預測不會二一

Training:

每一筆資料依決策樹分類後，會落在其中一個葉結點，因此每個葉結果會搜集到一群資料。完美的分類必需是每個葉結點資料群都是同類的，即都被二一或都不被二一，因此，對於一棵決策樹可計算它每個葉節點群的異類混雜程度，即不純度的概念，接著再進一步去考慮要不要對某個葉節點改成特徵變數子結點，進一步分類，以降低整體的不純度，直到不純度夠低為止。

## 支持向量機

Model：

預測法

Training： 若training data可被特徵變數超平面完美區隔成2類，則為線性可分，其訓練過程在於：

若training data不可被特徵變數超平面完美區隔成2類，則為線性不可分，其訓練過程在於：

## 人工神經網路

Model：

其中

函數a(z)一般稱為激活函數，用來控制z是否輸出，常用的函數型態為; 此外, 我們選定為一logistic函數:

Training：

預測法：