

支援向量機介紹

(Support Vector Machine, SVM)

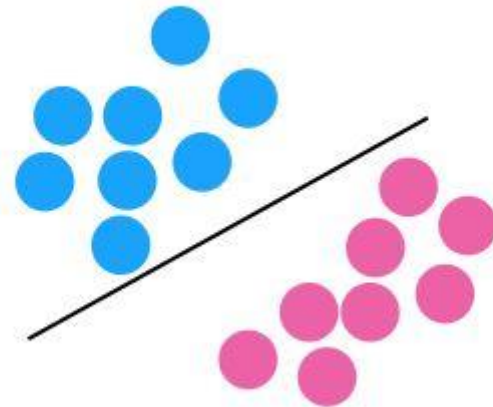
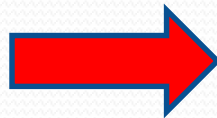
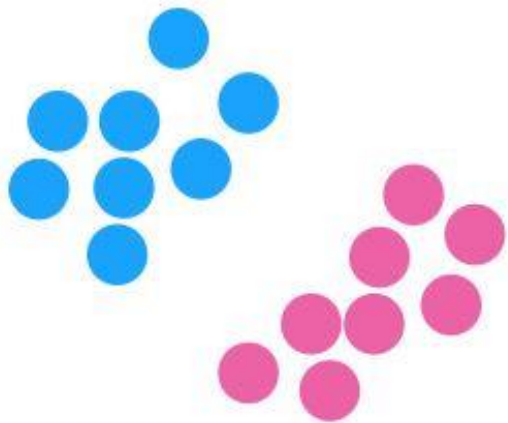
國立屏東科技大學資訊管理系

吳庭育 教授

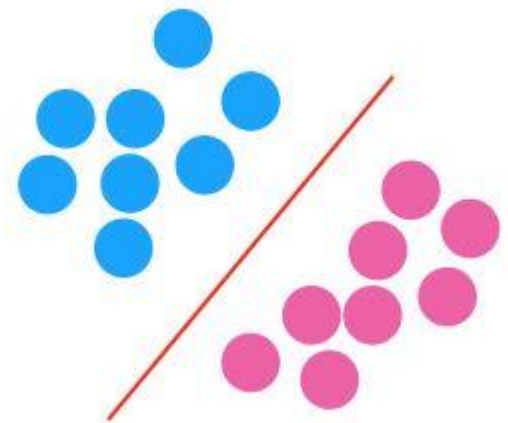
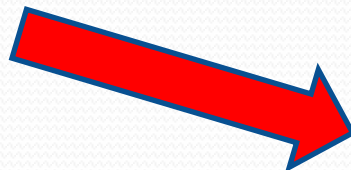
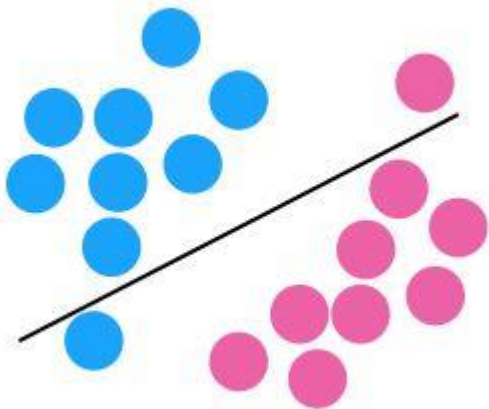
tyw@mail.npust.edu.tw

▼ 機器學習領域常見的類型

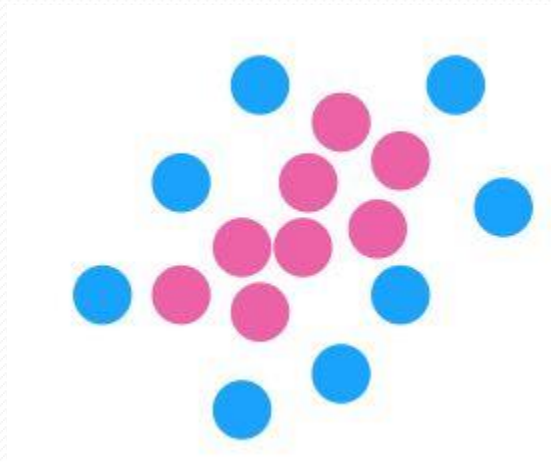
類型	線性迴歸分析	分類	分群
目的	趨勢預測分析	類別或等級的識別	類別或等級的區別
功能	連續值，預測數值為多少	非連續，負責識別出哪一種	非連續，負責區別出有幾群
學習方式	提供解答（標籤）	提供解答（標籤）	沒有解答（標籤）
應用	<ul style="list-style-type: none"> 下次考試成績會得幾分 最高氣溫預測冰品銷售量 最高氣溫與尖峰用電量 下一季的銷售額有多少 投入廣告費與銷售額 	<ul style="list-style-type: none"> 哪一個品種：依花的長寬分類 鐵達尼號船難者是否生還：依性別及艙等分類 判別是不是垃圾電子郵件 貓狗的識別 人臉、車牌等識別 	<ul style="list-style-type: none"> 班上同學分為跑得快跟跑得慢：依百米賽跑的秒數及身體的體脂肪率 哪些植物屬於相同的品種：依花的長寬分群 哪些動物屬於相同的品種：依體重及身長分群 哪些觀眾喜歡同一種類型的音樂或電影
常見的演算法	線性迴歸 複迴歸分析	KNN 決策樹 隨機森林 支援向量機	K-means

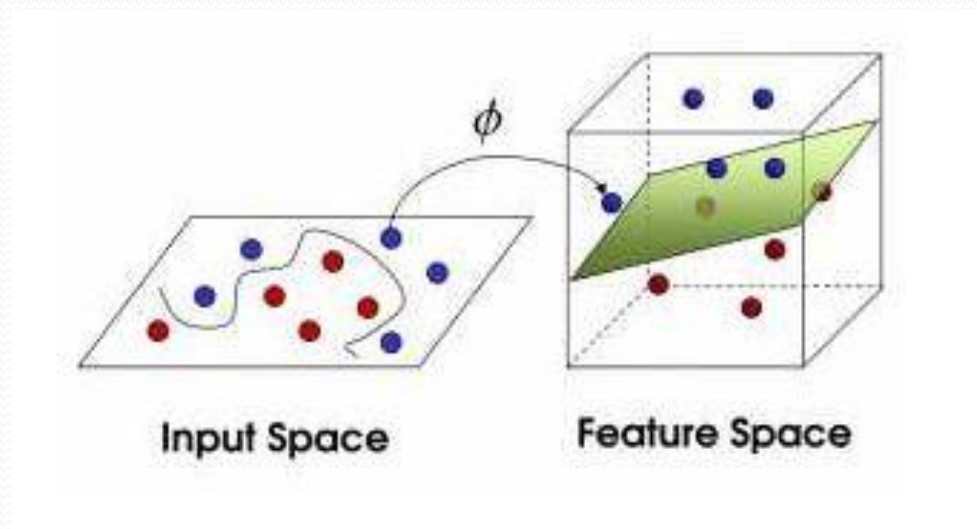


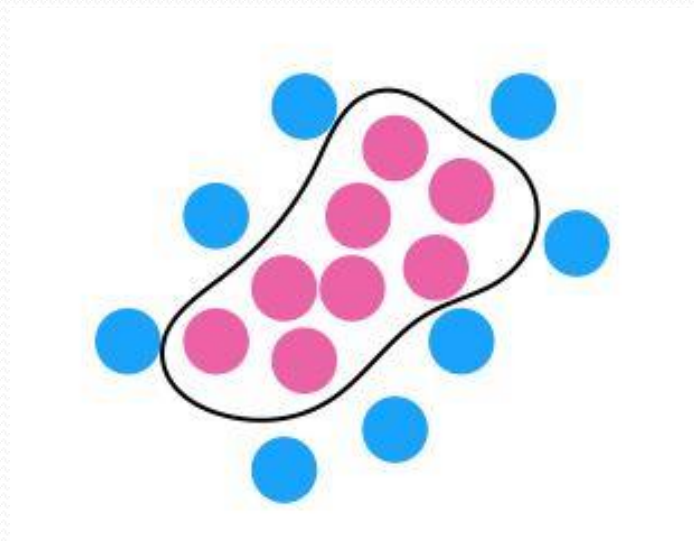
Logistic Regression



SVM



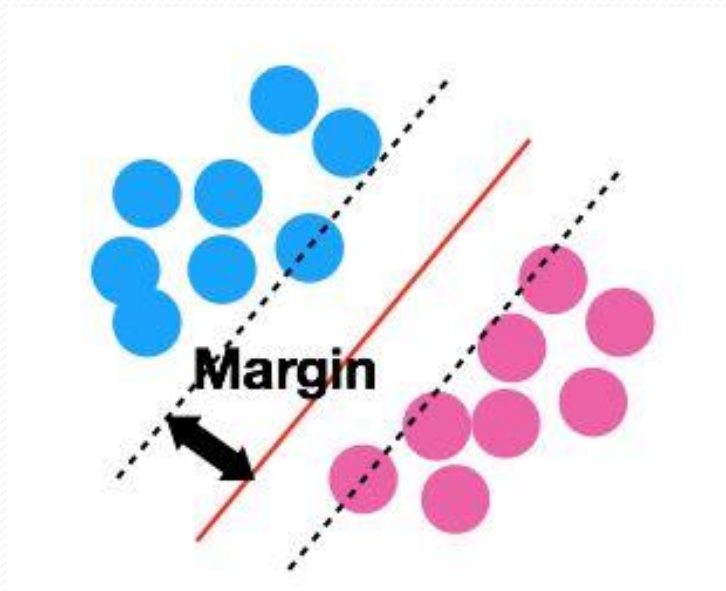






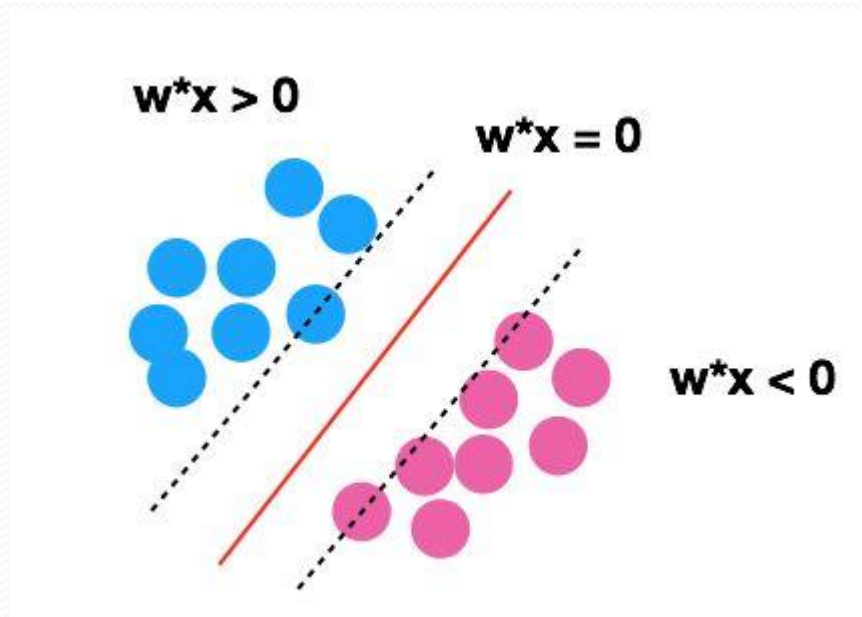
<https://www.youtube.com/watch?v=31iCbRZPrZA>

- SVM是怎麼得到那條很好的線呢？以直線來說，首先紅色的線會創造兩條黑色平行於紅色線的虛線，並讓黑線平移碰到最近的一個點，紅線到黑線的距離稱為Margin，而SVM

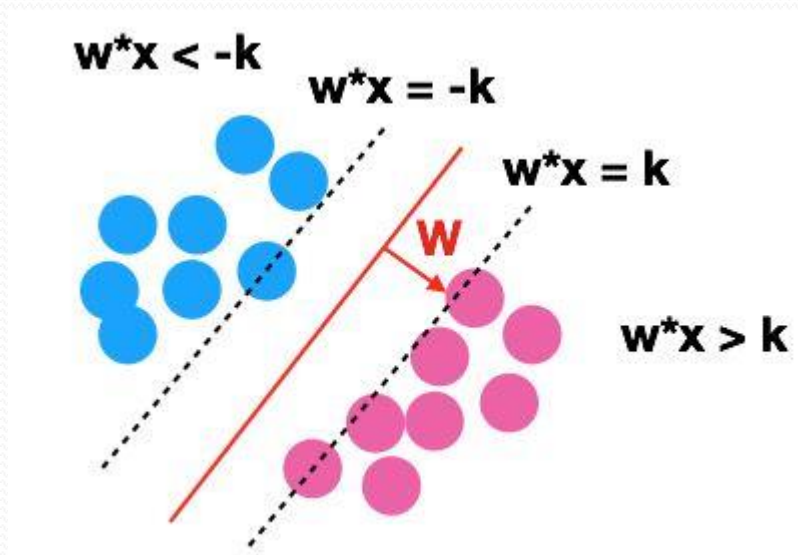


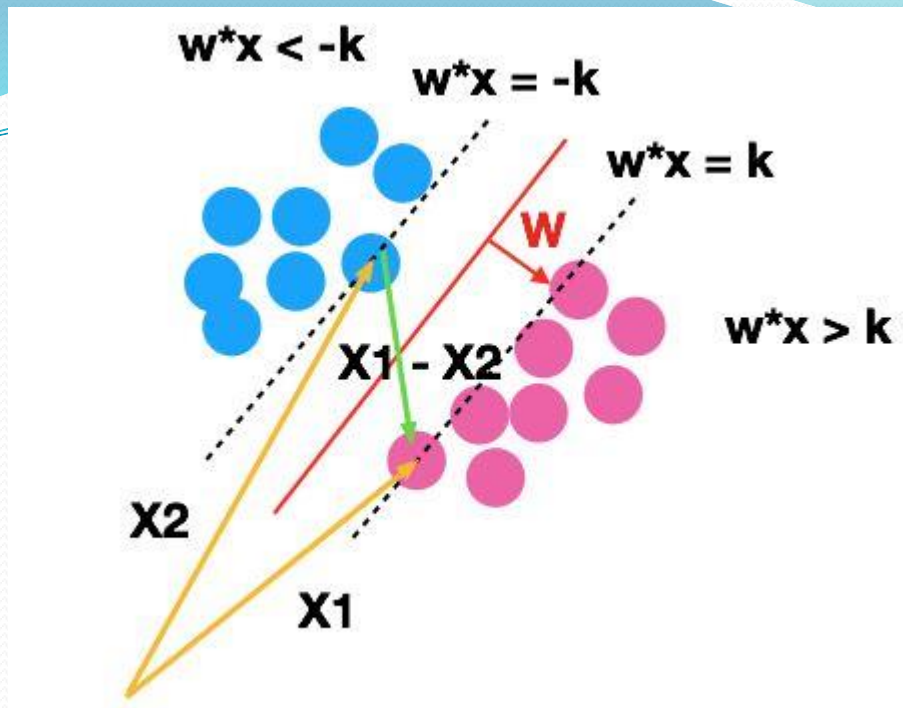
那要怎麼讓Margin最大呢？

- 假設紅線是 $w^*x = 0$ 在紅線上方的區域就是 $w^*x > 0$ 紅線下方的區域就是 $w^*x < 0$ 。



- 在左邊虛線上方的區域是 $w^*x < -k$ 在右邊虛線下方的區域是 $w^*x > k$ ，虛線中間不會有資料點。





$$\text{Margin} = \frac{W \cdot \overrightarrow{X1 - X2}}{2\|w\|}$$

$$= \frac{2k}{2\|w\|}$$

Condition = $y^*(w^*x) \geq k$

- 虛線上的點X1,X2 就是所謂的支援向量(Support Vector)
- 可以將X1向量-X2向量得到的向量投影到W 來計算Margin
- 接下來就是在 $Y^*(W^*X) \geq k$ 的條件下(虛線中間沒有點)，來最大化Margin。



END !!