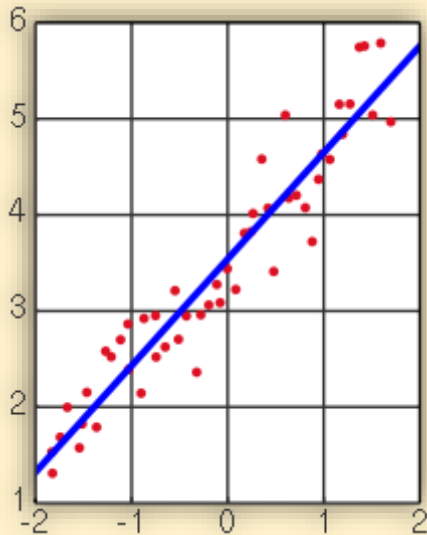


## 第五週 延伸資料



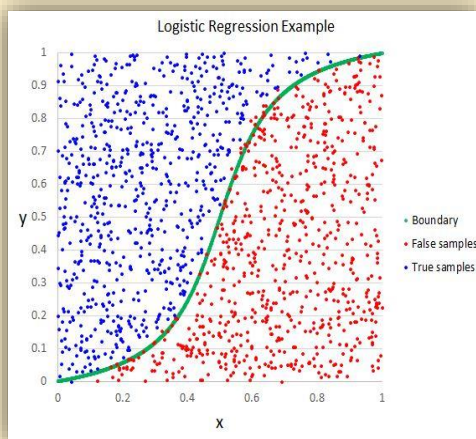
### 統計學



#### 線性迴歸

在統計學中，線性迴歸（英語：linear regression）是利用稱為線性迴歸方程式的最小平方函數對一個或多個自變數和應變數之間關係進行建模的一種迴歸分析。這種函數是一個或多個稱為迴歸係數的模型參數的線性組合。

更多介紹：[線性迴歸 wiki](#)



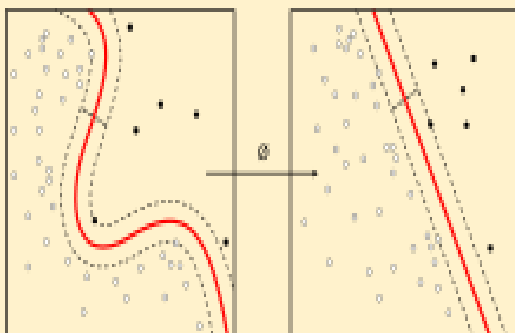
#### 羅吉斯迴歸

Logistic 迴歸是機器學習從統計領域借用的另一種技術。它是二進制分類問題（具有兩個類值的問題）的首選方法。Logistic 迴歸就像線性迴歸一樣，其目標是找到加權每個輸入變量的係數的值。

更多介紹：[羅吉斯迴歸 wiki](#)



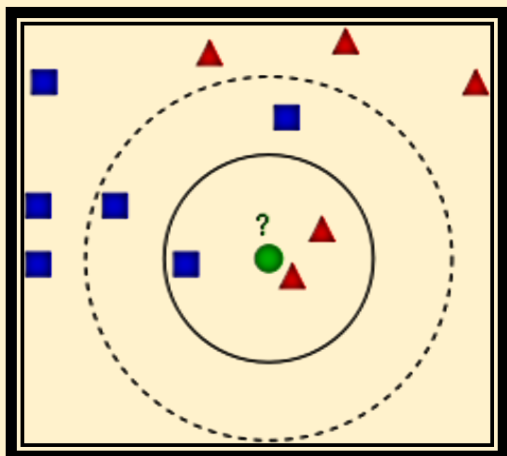
### 機器學習



#### 支援向量機

在機器學習中，支援向量機（英語：support vector machine，常簡稱為 SVM，又名支援向量網路[1]）是在分類與迴歸分析中分析資料的監督式學習模型與相關的學習演算法。

更多介紹：[支援向量機 wiki](#)



## K-近鄰演算法

最近鄰居法（KNN 演算法，又譯 K-近鄰演算法）是一種用於分類和回歸的無母數統計方法。

K-NN 是一種基於實例的學習，或者是局部近似和將所有計算推遲到分類之後的惰性學習。

K-NN 是所有的機器學習演算法中最簡單的之一。

更多介紹：[K-近鄰演算法 wiki](#)

## 單純貝氏分類器

在機器學習中，單純貝氏分類器是一系列以假設特徵之間強（樸素）獨立下運用貝葉斯定理為基礎的簡單機率分類器。在統計學和電腦科學文獻中，單純貝氏模型有各種名稱，包括簡單貝葉斯和獨立貝葉斯。

更多介紹：[單純貝氏分類器 wiki](#)



## scikit-learn

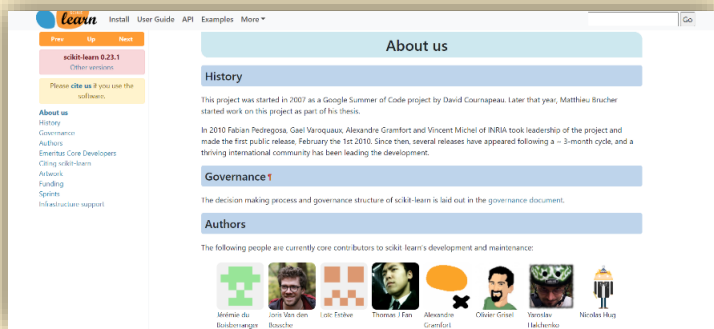


**Scikit-learn**（以前稱為 `scikits.learn`，也稱為 `sklearn`）是針對 Python 編程語言的免費軟件 機器學習 庫。它具有各種分類，回歸和聚類算法，包括支持向量機，隨機森林，梯度提升，k 均值和 DBSCAN，並且旨在與 Python 數值科學圖書館 NumPy 和科學。

更多介紹：[scikit-learn wiki](#)



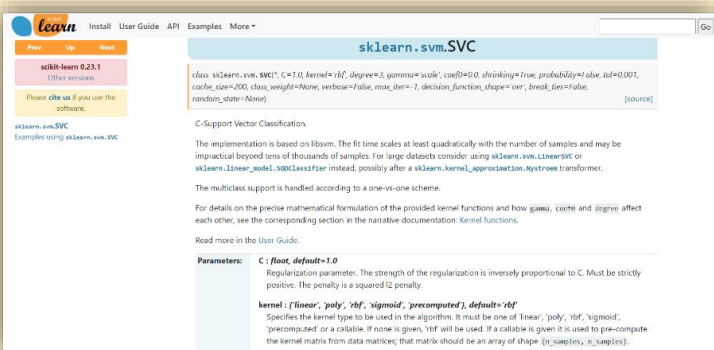
[Scikit-learn 官方網站](https://scikit-learn.org/)



[Scikit-learn](https://scikit-learn.org/)

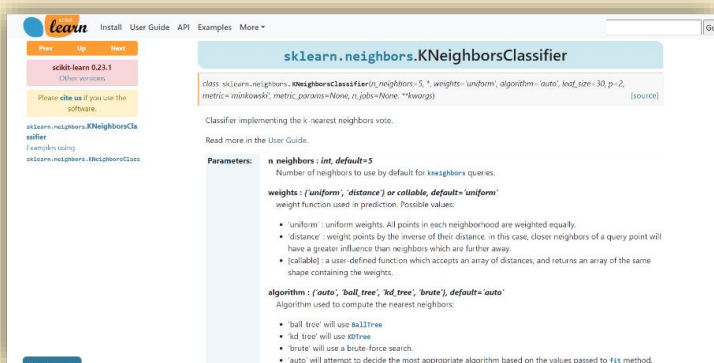


[Scikit-learn SVC](https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html)



[Scikit-learn](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html)

[KNeighborsClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html)



## 隨機森林分類器

隨機森林是一種元估計量，它適合數據集各個子樣本上的許多決策樹分類器，並使用平均數來提高預測準確性和控制過度擬合。子樣本大小由 `max_samples` 參數 if `bootstrap=True`（默認）控制，否則整個數據集將用於構建每棵樹。

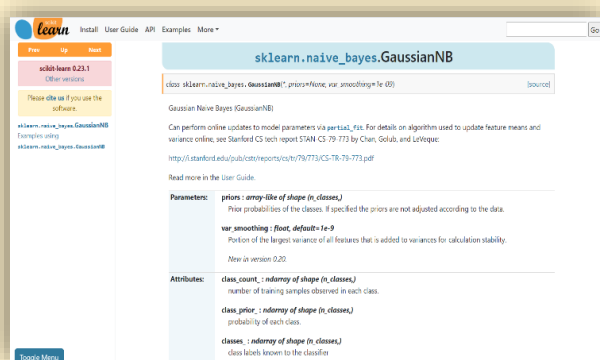
更多介紹：[Scikit-learn RandomForestClassifier](#)



## 高斯樸素貝葉斯

可以通過進行在線更新以更新模型參數 `partial_fit`。

更多介紹：[Scikit-learn GaussianNB](#)



## 樸素貝葉斯分類器用於多項模型

多項式樸素貝葉斯分類器適用於具有離散特徵的分類（例如，用於文本分類的字數統計）。多項式分佈通常需要整數特徵計數。但是，實際上，小數計數（例如 `tf-idf`）也可能起作用。

更多介紹：[Scikit-learn MultinomialNB](#)

