

Day 12

KNN (迴歸器)

[全民瘋AI系列]



第12屆 iT邦幫忙 鐵人賽

Day 12 學習目標

01

KNN 迴歸

學習 KNN 方法如何處理連續性輸出

02

實作 KNN 迴歸器

查看KNN方法在簡單線性迴歸和非線性迴歸表現

Part 1

KNN (迴歸器) 觀念講解

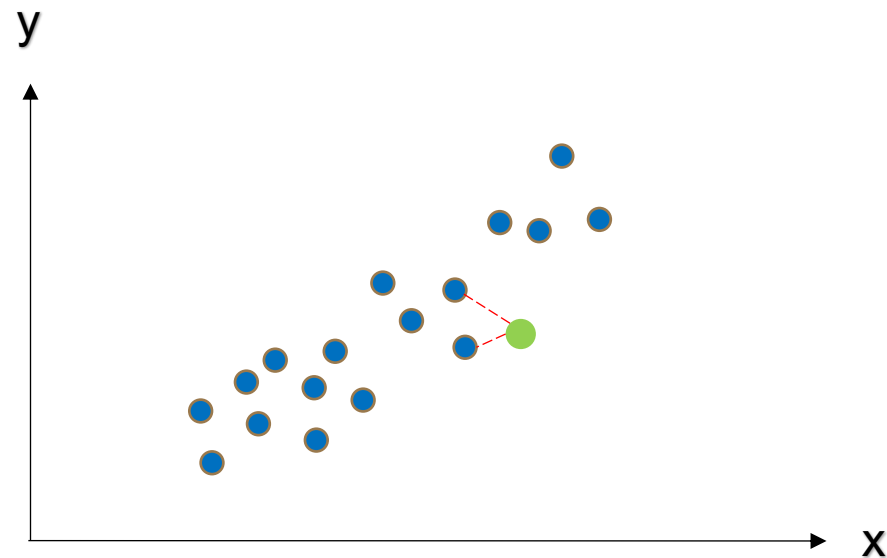


第12屆 iT邦幫忙 鐵人賽

/// KNN (Regression)

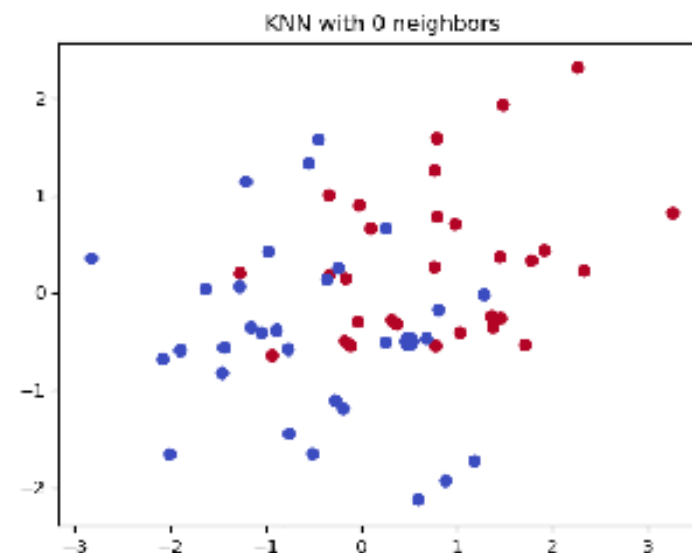
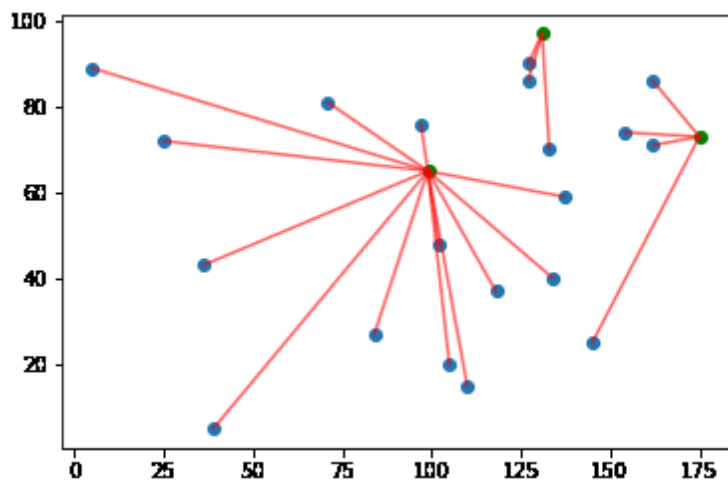
在sklearn中KNN提供迴歸(Regression)演算法

- 輸出的結果是一個連續性數值
- 該值是其k個最近鄰居的值的平均值。



//// KNN 與 k-means 勿混淆

- k-近鄰演算法的缺點是對資料的局部結構非常敏感
- KNN 與 K means 無任何關聯



迴歸模型評估

迴歸問題 (Regression)

- mean squared error (MSE)
- mean absolute error (MAE)
- R2 score

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=0}^{n_{\text{samples}}-1} (y_i - \hat{y}_i)^2}{\sum_{i=0}^{n_{\text{samples}}-1} (y_i - \bar{y})^2}$$



Part 2

KNN (迴歸器) 程式實作

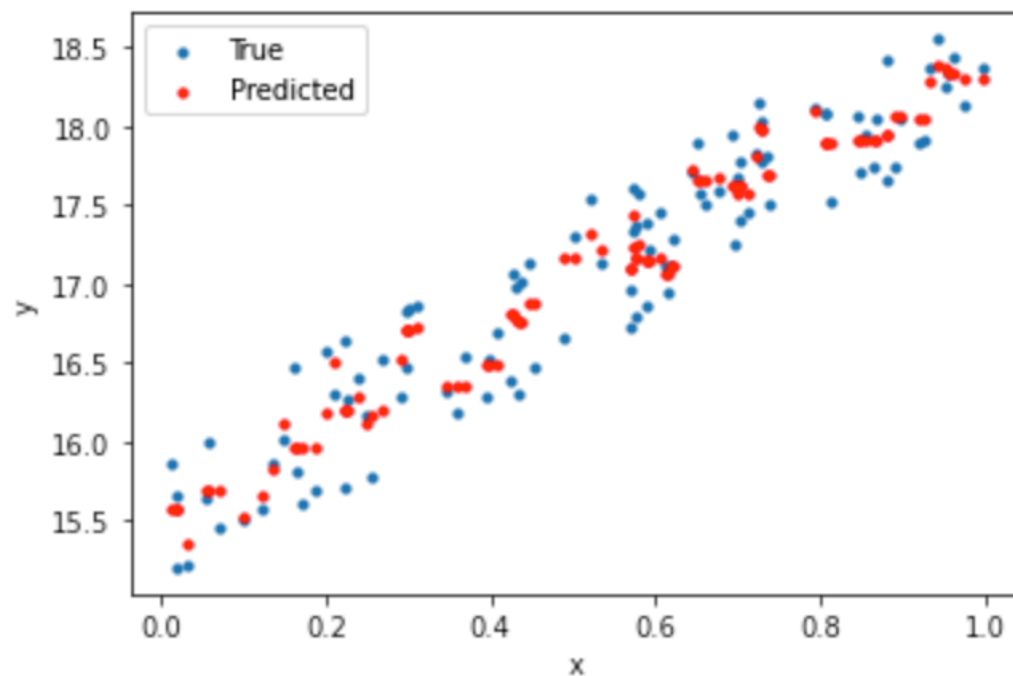


第12屆 iT邦幫忙 鐵人賽

//// 簡單線性迴歸

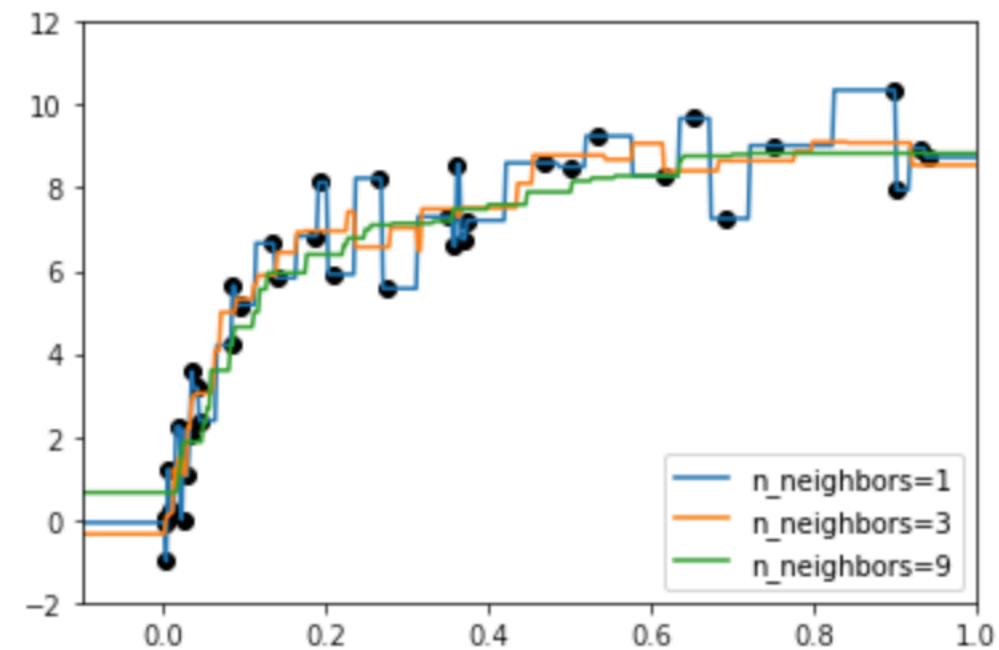
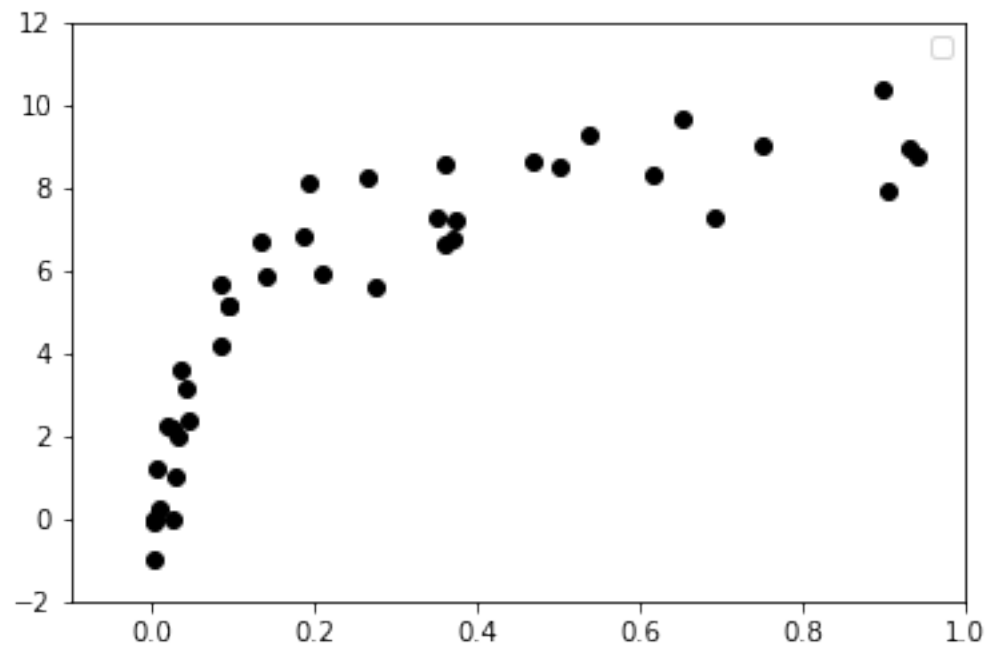
- 目標函式: $y=3x+15$
- 隨機添加 noise 讓資料分散
- x 值域介於 0~1

```
from sklearn.neighbors import KNeighborsRegressor  
  
# 建立KNN模型  
knnModel = KNeighborsRegressor(n_neighbors=3)  
# 使用訓練資料訓練模型  
knnModel.fit(x,y)  
# 使用訓練資料預測  
predicted= knnModel.predict(x)
```



非線性迴歸

在非線性的資料中可以發現 KNN 所預測出來的結果, k 值越大曲線越平滑



Thanks

PRESENTED BY 10程式中



第12屆 iT邦幫忙 鐵人賽