

[100 道題目](#) > [D1 資料介紹與評估資料\(串論+程式碼\)](#) > [範例下載](#) > [檔案詳情](#)

檔案詳情

[下載檔案](#)

# 統計指標實作範例

## 常見於迴歸問題的評估指標

平均絕對誤差 - Mean Absolute Error (MAE)

平均平方誤差(均方差) - Mean Squared Error (MSE)

## 常見於分類問題的指標

Binary Cross Entropy (CE)

後面的課程還會有更詳細的說明

## [教學目標]

開始的第一堂課：我們先一些機器學習的基礎開始，需要用到一些 Python 語法

如果不熟 Python，但是至少熟悉過一門語言，可以從這些範例開始熟悉

所謂評價函數 (Metric)，就是機器學習的計分方式，範例會展示平均絕對誤差 (MAE) 的寫法

我們來了解意義並寫作一個函數吧!!



複習 / 熟悉 Python 載入套件的語法, 了解什麼是代稱 (In[1], Out[1])

了解 Python 如何使用 Numpy 套件, 計算我們所需要的數值與繪圖 (In[2], Out[2], In[3], Out[3])

如何寫作平均絕對誤差 (MAE) 函數 (In[4], Out[4])

```
In [1]: # import 是 Python 載入套件的基本語法 (類似 C 語言的 i
# import AAAAA as BB, 其中 BB 是代稱, 表示除了載入 AAAA
# 常用套件往往有其對應代稱, numpy的代稱是np, pandas的代
# numpy 常用於數值/陣列運算, pandas 擅長資料格式的調整,
import numpy as np
import matplotlib.pyplot as plt
```

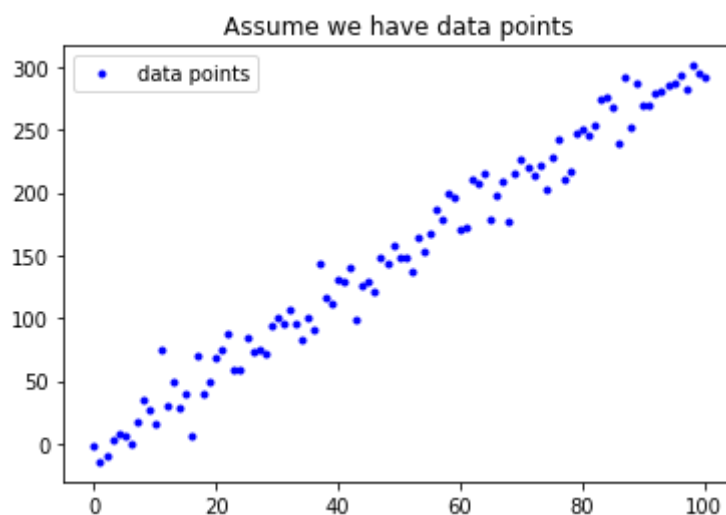


```
# w, b 是數值
w = 3
b = 0.5

# np.linspace 是 numpy.linspace 的意思
# np.linspace(0, 100, 101)是指 0~100 劃分成 101 個刻度
# 這時候, x_lin 因為要記錄不只一個數, 因為 np.linspace(
x_lin = np.linspace(0, 100, 101)

# np.random.randn() 就是 numpy.random.randn(), 會隨機值
# np.random.randn(101) 表示取樣了101次, 型態是 Array,
# 所以最後的結果 y, 也是一個長度 101 的 Array
y = (x_lin + np.random.randn(101) * 5) * w + b

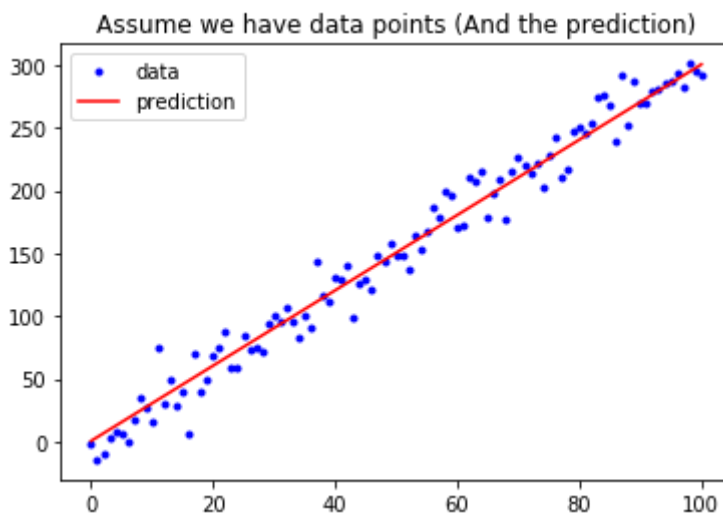
# 這邊就是將 x_lin 以及剛剛算完的 y, 當作座標值, 將101個
# b. : b 就是 blue, 點(.) 就是最小單位的形狀, 詳細可以查
plt.plot(x_lin, y, 'b.', label = 'data points')
plt.title("Assume we have data points")
plt.legend(loc = 2)
plt.show()
```



<Figure size 432x288 with 1 Axes>



```
plt.plot(x_lin, y, 'b.', label = 'data')  
# 上面的 'b.' 是藍色點狀，下面的 'r-' 是紅色線狀，label  
plt.plot(x_lin, y_hat, 'r-', label = 'prediction')  
plt.title("Assume we have data points (And the predic  
plt.legend(loc = 2)  
plt.show()
```



<Figure size 432x288 with 1 Axes>



```
# 以本例來說，mean_absolute_error 這個函數的定義範圍到
# 函數中，sum(), abs(), len() 都是 Python 原有的方法，
def mean_absolute_error(y, yp):
    """
    計算 MAE
    Args:
        - y: 實際值
        - yp: 預測值
    Return:
        - mae: MAE
    """
    # MAE : 將兩個陣列相減後，取絕對值(abs)，再將整個陣列
    mae = MAE = sum(abs(y - yp)) / len(y)
    return mae

# 呼叫上述函式，傳回 y(藍點高度)與 y_hat(紅線高度) 的 MAE
MAE = mean_absolute_error(y, y_hat)
print("The Mean absolute error is %.3f" % (MAE))
```

```
The Mean absolute error is 11.529
```

## 練習時間

請寫一個函式用來計算 Mean Square Error

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Hint: [如何取平方](#)

In [ ]:



