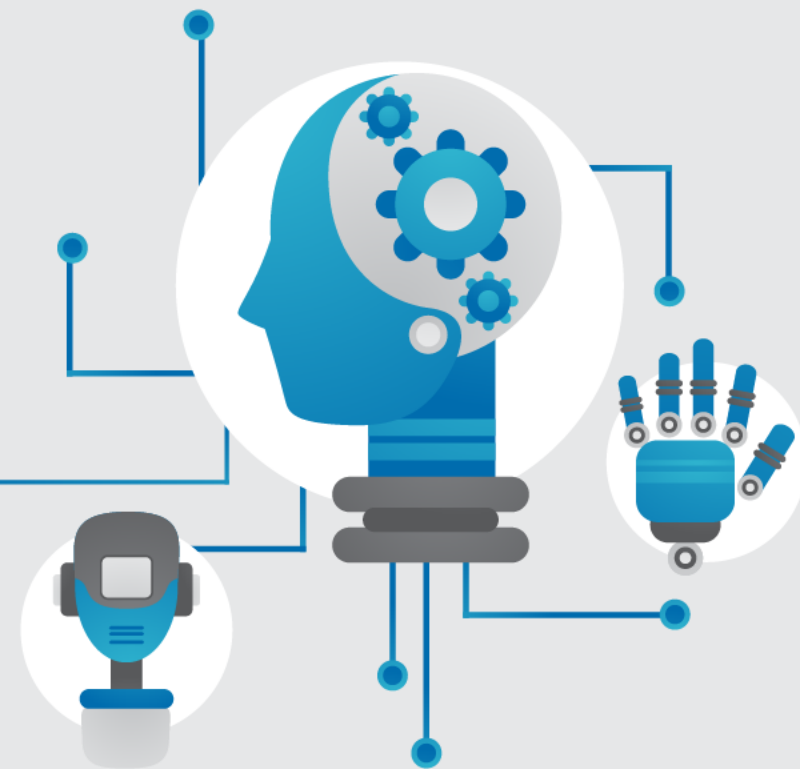


模型評估





混淆矩陣



› 混淆矩陣 (Confusion Matrix)

也稱為**錯誤矩陣**，是一種特殊的表格，
可將學習演算法的效率視覺呈現。

	實際 YES	實際 NO
預測 YES	True Positive (TP)	False Positive (FP) Type I Error
預測 NO	False Negative (FN) Type II Error	True Negative (TN)



Accuracy



› Accuracy (準確率)

$$\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total}$$

› 當TP案例數量很少時 (資料不平均) ，此指標會造成誤導。

	實際YES	實際NO
預測YES	True Positive (TP)	False Positive (FP) Type I Error
預測NO	False Negative (FN) Type II Error	True Negative (TN)



Precision



› Precision (精確率)

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

› 預測正向的案例，有多少是正確預測？

	實際YES	實際NO
預測YES	True Positive (TP)	False Positive (FP) Type I Error
預測NO	False Negative (FN) Type II Error	True Negative (TN)



Recall



› Recall (召回率) , 就是Sensitivity (靈敏度)

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

› 實際正向的案例，有多少被正確預測？

	實際YES	實際NO
預測 YES	True Positive (TP)	False Positive (FP) Type I Error
預測 NO	False Negative (FN) Type II Error	True Negative (TN)



Specificity



› Specificity (特異度)

$$\text{Specificity} = \text{TN} / (\text{FP} + \text{TN})$$

› 實際負向的案例，有多少被正確預測？

	實際YES	實際NO
預測YES	True Positive (TP)	False Positive (FP) Type I Error
預測NO	False Negative (FN) Type II Error	True Negative (TN)



F-measure



- › F1 score

$$\text{F1 score} = 2 / (1 / \text{precision} + 1 / \text{recall})$$

- › F-measure

$$F_{\beta} = (1 + \beta^2) \times \text{precision} \times \text{recall} / (\beta^2 \times \text{precision} + \text{recall})$$

- › Precision和Recall的統合指標，
F1 score將兩者視為同等重要



TPR & FPR



› TPR (True Positive Rate 真陽率) , 即Sensitivity

$$\text{TPR} = \text{TP} / (\text{TP} + \text{FN})$$

› FPR (False Positive Rate 假陽率)

$$\text{FPR} = \text{FP} / (\text{FP} + \text{TN})$$

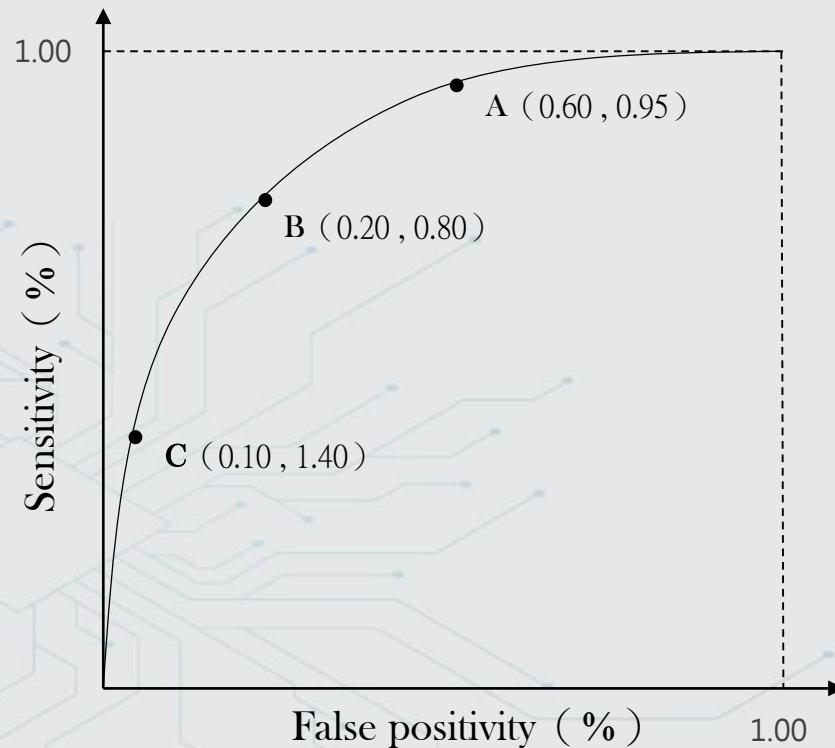
	實際YES	實際NO
預測YES	True Positive (TP)	False Positive (FP) Type I Error
預測NO	False Negative (FN) Type II Error	True Negative (TN)



ROC



- › ROC (受試者操作特徵曲線 Receiver Operating Characteristic curve)
ROC是由點 (TPR,FPR) 組成的曲線，
表示不同模型的真陽性率對假陽性率的函數關係。

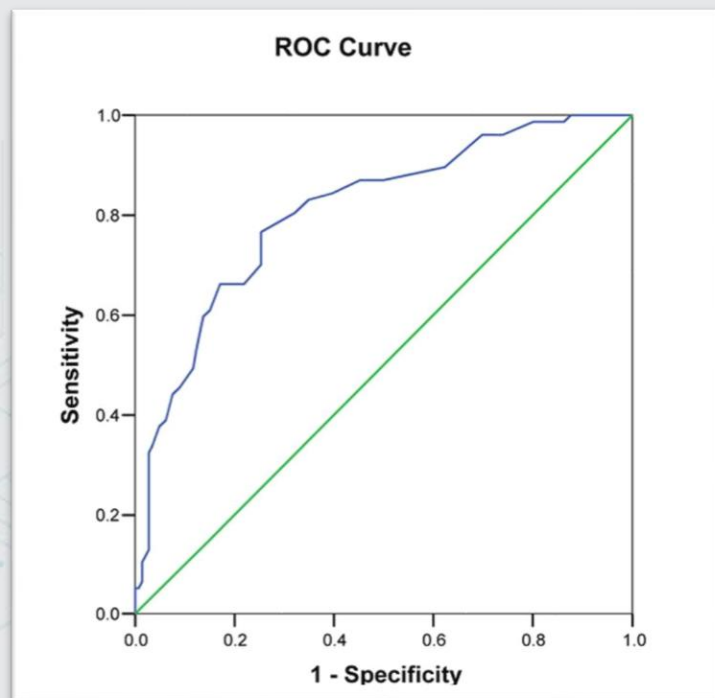




AUC



- › AUC (Area Under the Curve) = ROC的面積
- › AUC (面積) 愈大愈好
AUC值越大，當前的分類演算法愈有可能將正樣本排在負樣本前面，即能夠更好的分類。





Confusion Matrix實作



› 使用Pandas套件中的cross table

```
# 模型評估 Confution Matrix
prediction = model.predict_classes(x_test)
print(prediction)

import pandas as pd
print(pd.crosstab(y_test.reshape(-1) , prediction ,
                  rownames = ['label'], colnames=['predict']))
```



Confusion Matrix實作

機器學習實務



Predict label	0	1	2	3	4	5	6	7	8	9
0	973	0	1	2	0	0	1	1	2	0
1	0	1121	2	0	0	2	3	0	7	0
2	2	1	1012	2	1	0	2	9	3	0
3	1	0	7	977	0	15	0	6	4	0
4	2	1	2	0	953	0	5	5	2	12
5	3	0	0	2	1	873	6	3	3	1
6	6	3	0	0	3	12	934	0	0	0
7	1	5	8	2	0	0	0	1009	0	3
8	4	0	3	11	3	11	2	6	930	4
9	4	4	1	11	6	13	2	10	1	957