



# **EVALUACIÓN MODELOS CLASIFICACIÓN**



## Métricas de clasificación

Exactitud (Accuracy)

Precisión (Precision)

Exhaustividad (Recall)

F-score

Kappa

Matriz de confusión

Curvas ROC

AUC



## Exactitud (Accuracy)

If  $\hat{y}_i$  is the predicted value of the  $i$ -th sample and  $y_i$  is the corresponding true value, then the fraction of correct predictions over  $n_{\text{samples}}$  is defined as

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

¿Qué proporción de muestras fueron clasificadas correctamente?

```
from sklearn.metrics import accuracy_score  
accuracy=accuracy_score(y_true=iris_class_test, y_pred=predicted)  
print('Accuracy: %f' % accuracy)
```



## Precisión (Precision)

Precision is the ability of the classifier not to label as positive a sample that is negative

$$\textit{Precisión} = \frac{VP}{VP + FP}$$

¿Qué proporción de identificaciones positivas fue correcta?

```
from sklearn.metrics import precision_score  
precision=precision_score(y_true=iris_class_test, y_pred=predicted, average='micro')  
print('Precision:', precision)
```



## Exhaustividad (Recall)

Recall is the ability of the classifier to find all the positive samples

$$\text{Exhaustividad} = \frac{VP}{VP + FN}$$

¿Qué proporción de positivos reales se identificó correctamente?

```
from sklearn.metrics import recall_score  
recall = recall_score(y_true=iris_class_test, y_pred=predicted, average = 'micro')  
print('Recall: %f' % recall)
```



## F-score

F-Score (también llamado Valor-F o Medida-F en español) (Lewis y Gale, 1994) combina las medidas de precisión y exhaustividad para devolver una medida de calidad más general del modelo. Se calcula como la media armónica de las métricas mencionadas

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}} \quad F_1 = \left( \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} \right) = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

```
from sklearn.metrics import f1_score  
f1 = f1_score(y_true=iris_class_test, y_pred=predicted, average= 'micro')  
print('F1 score: %f' % f1)
```



## Kappa

El índice de Kappa, un instrumento diseñado por Cohen que ajusta el efecto del azar en la proporción de la concordancia observada

$$K = \frac{P_o - P_e}{1 - P_e}$$

Donde  $P_o$  es la proporción de concordancia observada,  $P_e$  es la proporción de concordancia esperada por azar y  $1 - P_e$ , representa el acuerdo o concordancia máxima posible no debida al azar

```
from sklearn.metrics import cohen_kappa_score  
kappa=cohen_kappa_score (y1= iris_class_test, y2= predicted)  
print('Cohens kappa: %f' % kappa)
```



## Matriz de confusión

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Ejemplo:

n=165	Predicted: NO	Predicted: YES
Actual: NO	50	10
Actual: YES	5	100

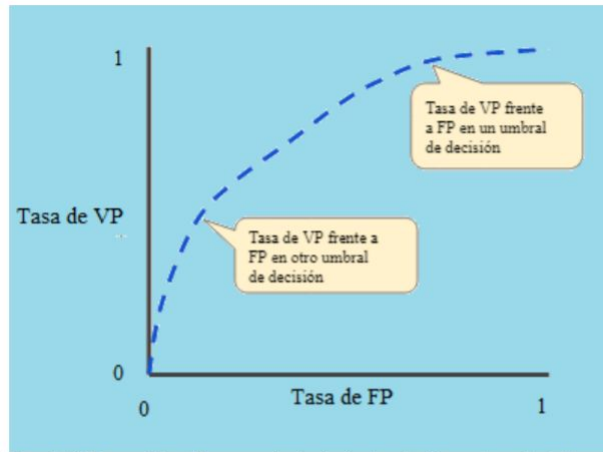
```
from sklearn.metrics import confusion_matrix
y_true = [2, 0, 2, 2, 0, 1]
y_pred = [0, 0, 2, 2, 0, 2]
confusion_matrix(y_true, y_pred)
array([[2, 0, 0],
       [0, 0, 1],
       [1, 0, 2]])
```





## Curvas ROC - AUC

$$TPR = \frac{VP}{VP + FN}$$



$$FPR = \frac{FP}{FP + VN}$$

