

UNIDAD

V2

**DIPLOMATURA EN PYTHON APLICADO A LA
CIENCIA DE DATOS**

Métricas Modelos de Clasificación

Unidad V2: Métricas Modelos de Clasificación

Matriz de Confusión

Una matriz de confusión, también conocida como matriz de errores, es una tabla resumida que se utiliza para evaluar el desempeño de un modelo de clasificación.

El número de predicciones correctas e incorrectas se resume con valores de recuento y se desglosa por clase.

A continuación se muestra una imagen de la estructura de una matriz de confusión de 2x2. Para dar un ejemplo, digamos que hubo diez casos en los que un modelo de clasificación predijo "Sí" en el que el valor real fue "Sí".

Luego, el número diez iría en la esquina superior izquierda en el cuadrante Verdadero Positivo. Esto nos lleva a algunos términos clave:

		Valores Actuales	
		Sí	No
Valores Predichos	Sí	Verdadero Positivo	Falso Positivo
	No	Falso Negativo	Verdadero Negativo

- Positivo (P): la observación es positiva
- Negativo (N): la observación no es positiva
- Verdadero Positivo (TP): Resultado en el que el modelo predice correctamente la clase positiva.
- Verdadero Negativo (TN): Resultado en el que el modelo predice correctamente la clase negativa.
- Falso positivo (FP): también llamado error de tipo 1, un resultado en el que el modelo predice incorrectamente la clase positiva cuando en realidad es negativa.
- Falso negativo (FN): también llamado error de tipo 2, un resultado en el que el modelo predice incorrectamente la clase negativa cuando en realidad es positivo.

Precisión

Con la métrica de precisión podemos medir la **calidad** del modelo de machine learning en tareas de clasificación.

Responde a la pregunta ¿qué porcentaje de lo identificado como positivo es realmente correcto?

Accuracy (Exactitud)

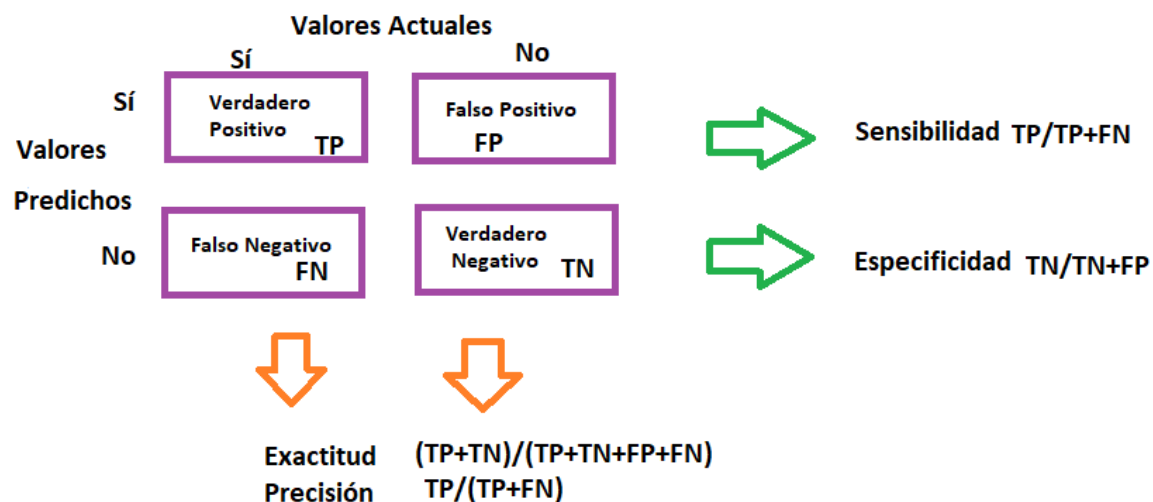
La exactitud (accuracy) mide el porcentaje de casos que el modelo ha acertado. Evitar el abuso de solo aplicar esta métrica. El problema con la exactitud es que nos puede llevar al engaño, es decir, puede hacer que un modelo malo parezca que es mucho mejor de lo que es.

Recall (Exhaustividad- Sensibilidad)

La métrica de exhaustividad nos va a informar sobre la **cantidad** que el modelo de machine learning es capaz de identificar. Responde a la pregunta ¿Qué porcentaje de los valores positivos fueron bien identificados?

Especificidad:

También conocida como la Tasa de Verdaderos Negativos, ("true negative rate") o TN. Se trata de los casos negativos que el algoritmo ha clasificado correctamente. Expresa cuán bien puede el modelo detectar esa clase.-



F1

El valor F1 se utiliza para combinar las medidas de precisión y sensibilidad en un sólo valor. Esto es práctico porque hace más fácil el poder comparar el rendimiento combinado de la precisión y la sensibilidad entre varias soluciones. F1 se calcula haciendo la media armónica entre la precisión y la sensibilidad:

$$F1 = 2 * (\text{precisión} * \text{sensibilidad}) / (\text{precisión} + \text{sensibilidad})$$

¿Cómo usamos en Python?

Confusion Matrix

```
from sklearn.metrics import confusion_matrix  
confusion_matrix(y_true, y_pred)
```

Accuracy

```
from sklearn.metrics import accuracy_score  
accuracy_score(y_true, y_pred)
```

Recall

```
from sklearn.metrics import recall_score  
recall_score(y_true, y_pred, average=None)
```

Precision

```
from sklearn.metrics import precision_score  
precision_score(y_true, y_pred, average=None)
```

Método 1: sklearn

```
from sklearn.metrics import f1_score  
f1_score(y_true, y_pred, average=None)
```

Método 2: Cálculo Manual

$F1 = 2 * (\text{precision} * \text{sensibilidad}) / (\text{precision} + \text{sensibilidad})$

Method 3: Classification report [BONUS]

```
from sklearn.metrics import classification_report  
print(classification_report(y_true, y_pred, target_names=target_names))
```

Usando Pandas

Version de pandas V23 o superior

```
from pandas_ml import ConfusionMatrix
```

```
Confusion_Matrix = ConfusionMatrix(df['y_Actual'], df['y_Predicted'])  
Confusion_Matrix.print_stats()
```

```
population: 12
P: 5
N: 7
PositiveTest: 6
NegativeTest: 6
TP: 4
TN: 5
FP: 2
FN: 1
TPR: 0.8
TNR: 0.7142857142857143
PPV: 0.6666666666666666
NPV: 0.8333333333333334
FPR: 0.2857142857142857
FDR: 0.3333333333333333
FNR: 0.2
ACC: 0.75
F1_score: 0.7272727272727273
MCC: 0.50709255283711
informedness: 0.5142857142857142
markedness: 0.5
prevalence: 0.4166666666666667
LRP: 2.8000000000000003
LRN: 0.28
DOR: 10.0
FOR: 0.16666666666666666
```