



# CIENCIA DE • DATOS •

# CIENCIA DE • DATOS •

## **Evaluación de modelos: métricas de rendimiento, validación cruzada, matriz de confusión.**

La evaluación de modelos es una etapa esencial en el proceso de construcción de modelos en Ciencia de Datos. Consiste en medir y comprender el rendimiento de los modelos en función de métricas específicas, validar su capacidad de generalización a nuevos datos y comprender la calidad de las predicciones realizadas. En esta introducción, exploraremos la importancia de la evaluación de modelos, así como algunas métricas de rendimiento comunes, técnicas de validación cruzada y el uso de la matriz de confusión en esta fase.





## Importancia de la evaluación de modelos:

La evaluación de modelos es crucial para determinar qué tan bien un modelo puede generalizar y realizar predicciones precisas en nuevos datos. Permite medir el rendimiento del modelo en función de métricas específicas, como la precisión, el recall, el F1-score y la exactitud. Además, ayuda a identificar problemas como el sobreajuste (overfitting) o subajuste (underfitting) y proporciona información para mejorar el modelo.



# CIENCIA DE DATOS

## Métricas de rendimiento:

Existen diversas métricas de rendimiento utilizadas para evaluar modelos en diferentes escenarios. Algunas de las métricas más comunes incluyen:

**Precisión (Precision):** Mide la proporción de instancias positivas correctamente clasificadas sobre todas las instancias clasificadas como positivas. Es útil cuando el objetivo es minimizar los falsos positivos.

**Recall (Recall o Sensibilidad):** Mide la proporción de instancias positivas correctamente clasificadas sobre todas las instancias que son realmente positivas. Es útil cuando el objetivo es minimizar los falsos negativos.

**F1-score:** Es una medida que combina la precisión y el recall en un solo valor, proporcionando una medida general del rendimiento del modelo.

**Exactitud (Accuracy):** Mide la proporción de instancias correctamente clasificadas sobre el total de instancias. Es útil cuando las clases están equilibradas en el conjunto de datos.

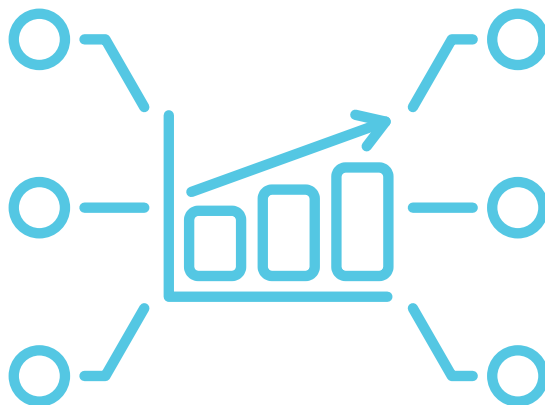


# CIENCIA DE • DATOS •

## Validación cruzada:

La validación cruzada es una técnica utilizada para evaluar el rendimiento de los modelos y su capacidad de generalización a nuevos datos. Consiste en dividir el conjunto de datos en múltiples subconjuntos, generalmente llamados "folds". Luego, se entrena y evalúa el modelo en diferentes combinaciones de subconjuntos, asegurando que cada subconjunto se utilice tanto para entrenamiento como para evaluación en diferentes iteraciones.

Esta técnica proporciona una estimación más confiable del rendimiento del modelo al reducir la dependencia de una sola partición de los datos. La validación cruzada también ayuda a detectar problemas como el sobreajuste y subajuste.





## Matriz de confusión:

La matriz de confusión es una herramienta visual que muestra la cantidad de instancias clasificadas correctamente e incorrectamente por un modelo de clasificación. La matriz se organiza en filas y columnas, representando las clases reales y las clases predichas, respectivamente.

La matriz de confusión permite analizar el rendimiento del modelo en términos de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos. Basándose en estos valores, se pueden calcular diversas métricas, como la precisión, el recall y el F1-score.

### Ejemplo de uso de métricas y matriz de confusión:

Supongamos que hemos construido un modelo de clasificación para predecir si un correo electrónico es spam o no spam. Después de entrenar y evaluar el modelo, obtenemos la siguiente matriz de confusión:

|                 | Predicted Not Spam | Predicted Spam |
|-----------------|--------------------|----------------|
| Actual Not Spam | 900                | 20             |
| Actual Spam     | 25                 | 855            |

Basándonos en esta matriz de confusión, podemos calcular varias métricas de rendimiento. Por ejemplo, la precisión se calcula como  $900 / (900 + 20) = 0.978$ , el recall se calcula como  $900 / (900 + 25) = 0.973$  y el F1-score se calcula como  $2 * ((0.978 * 0.973) / (0.978 + 0.973)) = 0.975$ .



# CIENCIA DE DATOS

La evaluación de modelos es una etapa fundamental en el proceso de construcción de modelos en Ciencia de Datos. Permite medir el rendimiento y la calidad de los modelos, identificar problemas y comprender su capacidad de generalización. Las métricas de rendimiento, como la precisión, el recall y el F1-score, proporcionan información cuantitativa sobre el rendimiento del modelo, mientras que la validación cruzada asegura una evaluación confiable y robusta.

La matriz de confusión es una herramienta visual útil para analizar el rendimiento del modelo y comprender el equilibrio entre los diferentes tipos de errores. Al utilizar estas técnicas y herramientas de evaluación, se puede tomar decisiones informadas y mejorar continuamente los modelos en el campo de la Ciencia de Datos.

