

VALIDACIÓN CRUZADA



OCCAM'S PROFESSOR

"WHEN FACED WITH TWO POSSIBLE WAYS OF DOING SOMETHING, THE MORE COMPLICATED ONE IS THE ONE YOUR PROFESSOR WILL MOST LIKELY ASK YOU TO DO."



Edición 2025

Angela Gorgoglione

Instituto de Mecánica de los Fluidos e Ingeniería Ambiental (IMFIA)
Facultad de Ingeniería, Universidad de la República, Uruguay

agorgoglione@fing.edu.uy

BIBLIOGRAFÍA

Handbook of HydroInformatics, Volume I: Classic Soft-Computing Techniques, Chapter 5. Elsevier, ISBN: 978-0-12-821285-1

OBJETIVOS

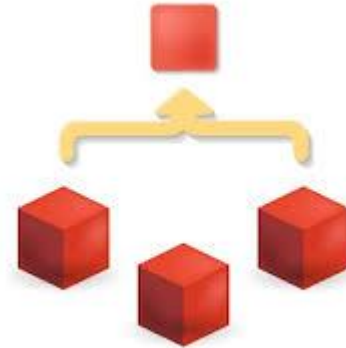
Comprender la importancia de la **validación** en modelos de **aprendizaje automático**, conocer los distintos tipos de validación y aplicarlos según el problema.

Al final, los estudiantes podrán elegir e interpretar la técnica adecuada para evaluar la **generalización de un modelo**.

IMPORTANCIA DE LA VALIDACIÓN

¿Cómo sabes eso?

¿Estás seguro?



En la ciencia y la ingeniería, la **validación** es esencial cuando se adquieren datos experimentales o se proponen modelos **para generalizar la aplicabilidad** de estos datos. En modelos predictivos, la validación garantiza que las predicciones sean fiables antes de su aplicación en diseños, inversiones o pronósticos.

En **modelos basados en aprendizaje automático**, la validación es clave para evaluar su capacidad de generalización. Un modelo válido debe funcionar adecuadamente con datos independientes, es decir, aquellos que no se utilizaron en su entrenamiento.

IMPORTANCIA DE LA VALIDACIÓN

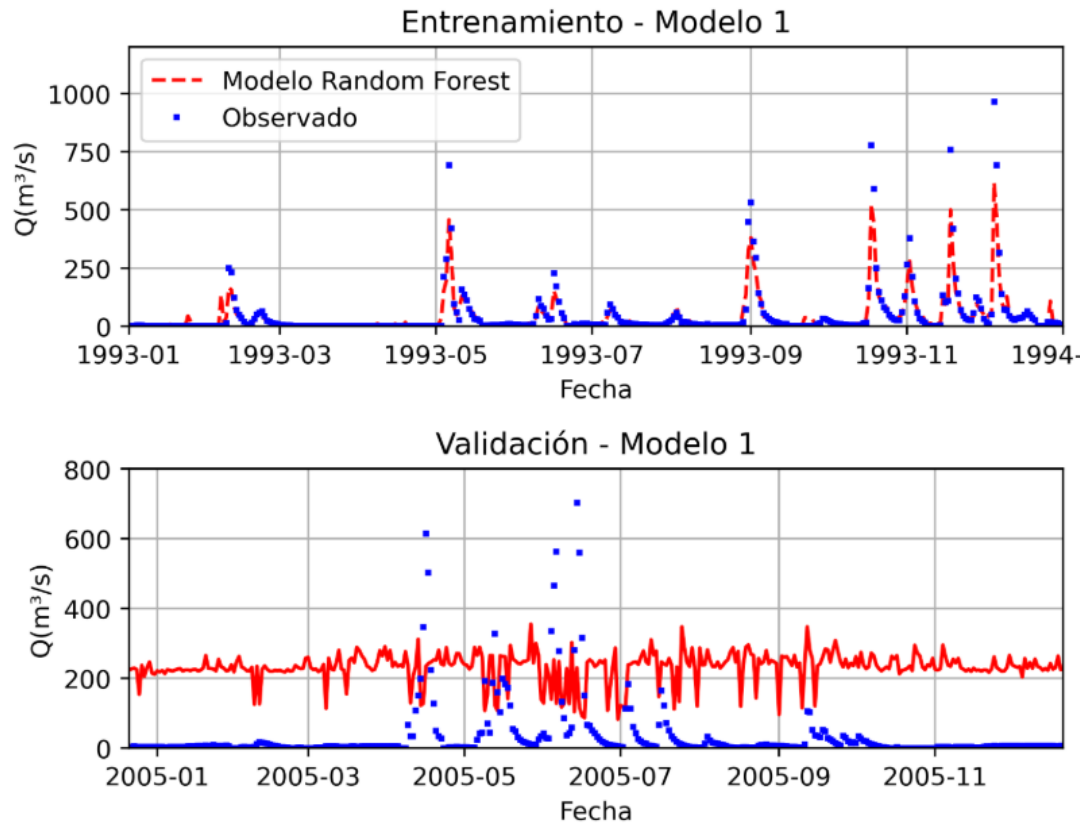
Validación del Proceso de Entrenamiento

En aprendizaje automático, la fase de **entrenamiento** utiliza la mayor parte de los datos disponibles para encontrar los parámetros del modelo que minimicen el error.

Sin embargo, un modelo que se ajusta demasiado a los datos de entrenamiento puede fallar con datos nuevos. Por ello, es necesario establecer criterios de detención para evitar el **sobreajuste**.

IMPORTANCIA DE LA VALIDACIÓN

Validación del Proceso de Entrenamiento



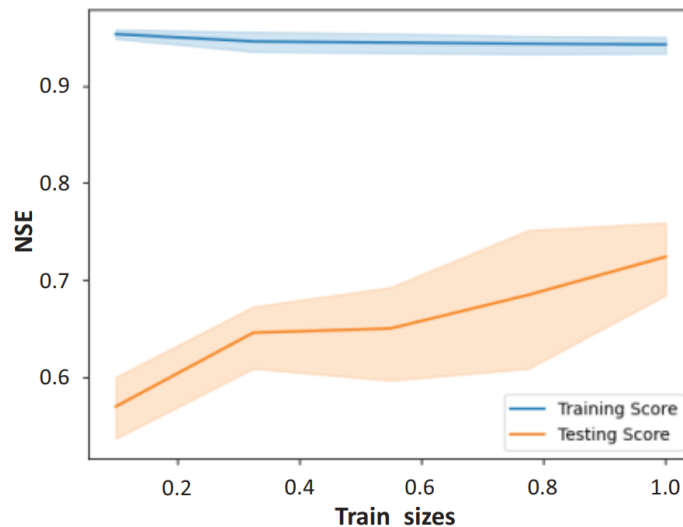
Ejemplo de sobreajuste

Parámetro	Entrenamiento	Validación
NS	0.89	-5.42
R ²	0.89	-5.42
RMSE	33.6	229.1
KGE	0.5	-0.62

IMPORTANCIA DE LA VALIDACIÓN

Validación del Proceso de Entrenamiento

Una estrategia común es **evaluar el error en un conjunto de prueba** durante el entrenamiento y detener el proceso cuando este error deja de mejorar.



**Learning
curve**

La **validación cruzada** es una técnica más robusta que permite evaluar el modelo utilizando diferentes particiones de los datos disponibles.

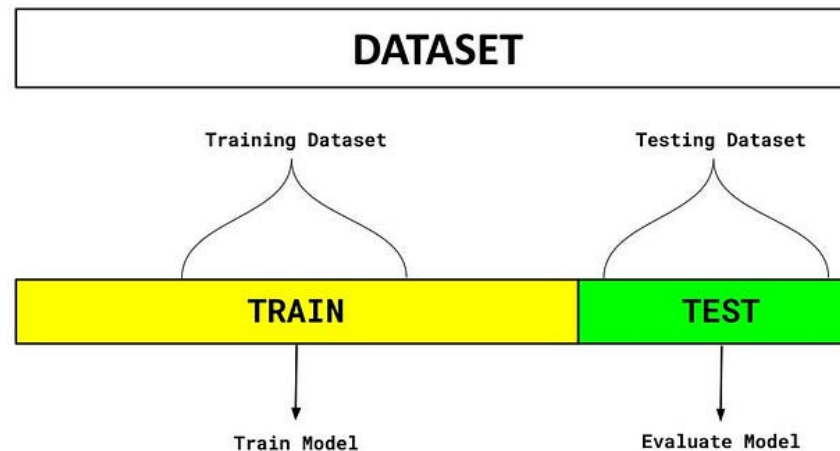
VALIDACIÓN CRUZADA

La validación cruzada consiste en dividir el conjunto de datos en conjuntos de entrenamiento y prueba de diversas maneras. Puede ser:

1. Validación Cruzada Exhaustiva y No Exhaustiva:

❖ **No Exhaustiva:** Utiliza solo algunas combinaciones, reduciendo el costo computacional.

📌 *Ejemplo:* Separamos un conjunto de datos de pacientes en 70% para entrenar un modelo de diagnóstico médico y 30% para probarlo. Si el modelo tiene buen desempeño en los datos de prueba, podemos confiar en él.



VALIDACIÓN CRUZADA

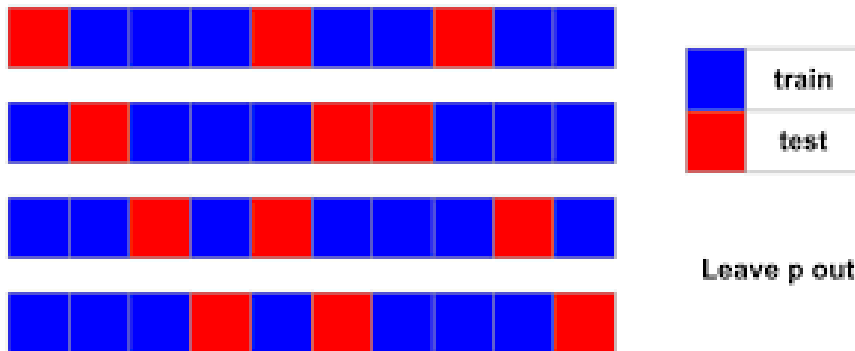
La validación cruzada consiste en dividir el conjunto de datos en conjuntos de entrenamiento y prueba de diversas maneras. Puede ser:

1. Validación Cruzada Exhaustiva y No Exhaustiva:

❖ **Exhaustiva:** Considera todas las combinaciones posibles de división de los datos.

1.1 *Leave-p-out Cross-Validation (LpO-CV)*

- Se retiran "p" muestras del conjunto total de datos y se entrena con las restantes.
- Puede volverse computacionalmente costoso si "p" es grande.



📌 *Ejemplo:* En un conjunto con 50 mediciones de contaminación, si usamos $p=2$, entrenamos el modelo con 48 valores y lo probamos con los 2 restantes, repitiendo este proceso hasta haber probado todas las combinaciones posibles.

VALIDACIÓN CRUZADA

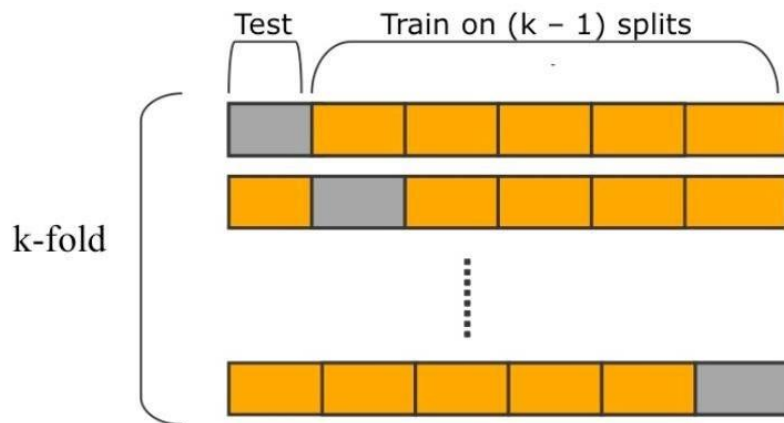
La validación cruzada consiste en dividir el conjunto de datos en conjuntos de entrenamiento y prueba de diversas maneras. Puede ser:

1. Validación Cruzada Exhaustiva y No Exhaustiva:

❖ **Exhaustiva:** Considera todas las combinaciones posibles de división de los datos.

1.2 *Leave-One-Out Cross-Validation (LOO-CV)*

- Variante de LpO-CV donde " $p=1$ ".
- Cada muestra actúa como conjunto de prueba una vez.
- Ofrece alta fidelidad pero puede ser computacionalmente demandante.



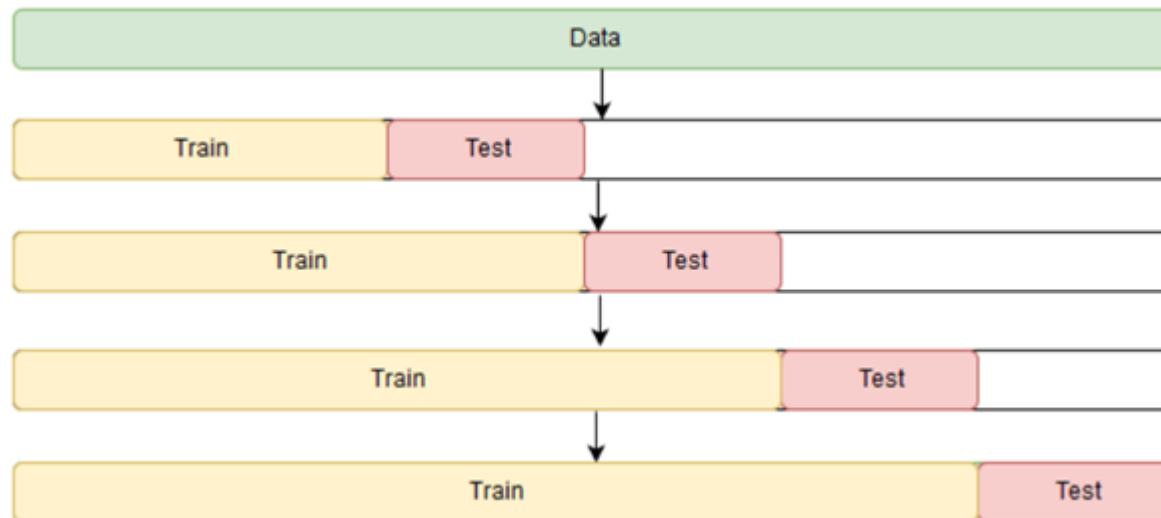
📌 *Ejemplo:* Para predecir el rendimiento académico de un estudiante, entrenamos con todos los registros excepto uno y usamos ese único dato restante para evaluar el modelo. Repetimos esto para cada estudiante.

VALIDACIÓN CRUZADA

3. Validación Cruzada para Series Temporales:

- ❖ Se respeta el orden temporal de los datos.
- ❖ Se entrena con datos pasados y se prueba con datos futuros.
- ❖ Conocida como "validación rodante" o "método de cadena progresiva".

📌 *Ejemplo:* Para predecir la demanda eléctrica, entrenamos con datos de enero a marzo y validamos con abril. Luego, usamos enero a abril y validamos con mayo, asegurándonos de no usar información del futuro

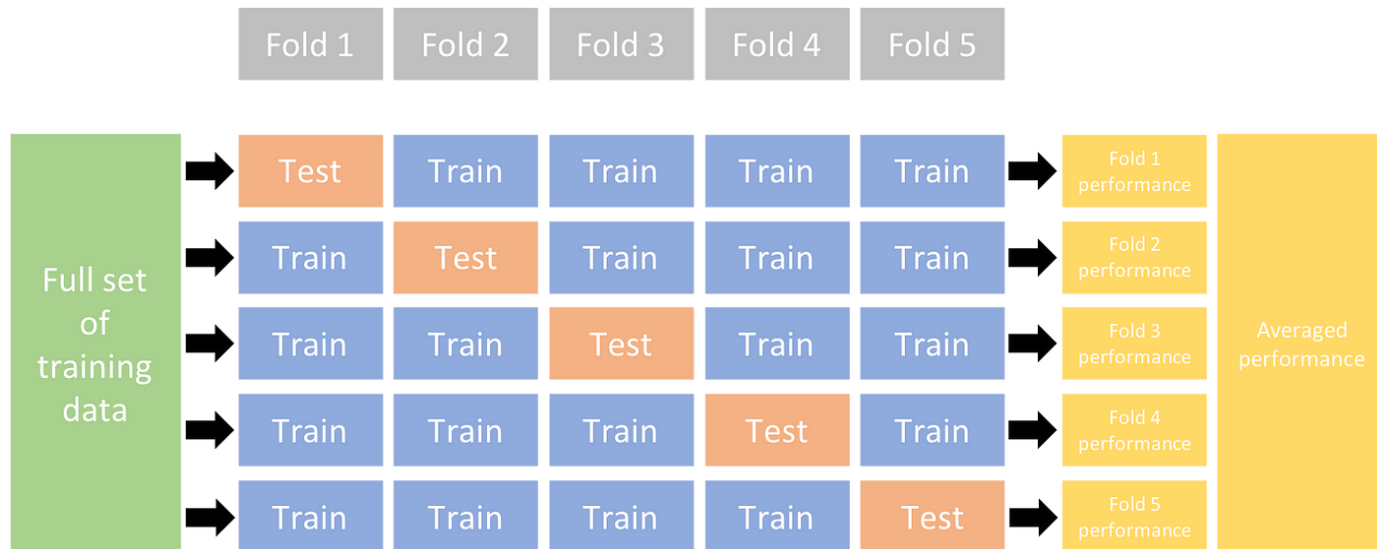


VALIDACIÓN CRUZADA

4. Validación Cruzada *k-Fold*:

- ❖ Se divide el conjunto de datos en "*k*" subconjuntos o "*folds*".
- ❖ Se entrena el modelo con "*k-1*" subconjuntos y se prueba con el restante.
- ❖ Se repite "*k*" veces para que cada subconjunto actúe como conjunto de prueba.
- ❖ Se promedian los resultados para obtener el desempeño final.

📌 *Ejemplo*: Si tenemos 1000 muestras para predecir la calidad del aire, las dividimos en 5 partes (*folds*). Entrenamos con 4 y validamos con la restante, repitiendo el proceso hasta que cada *fold* haya sido test una vez.

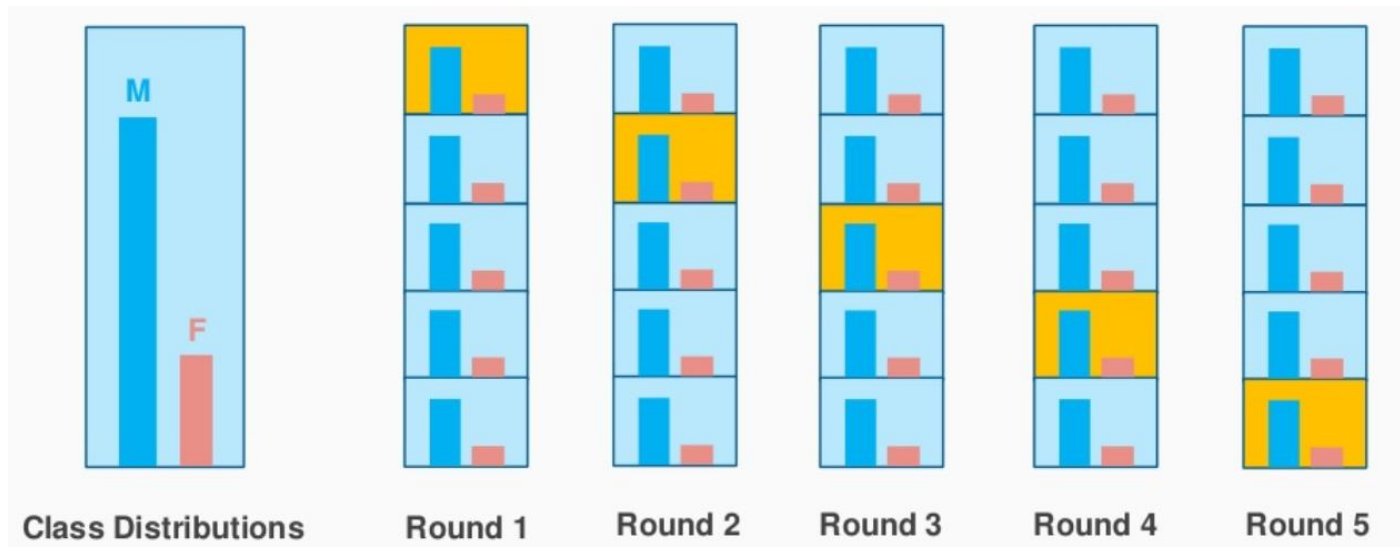


VALIDACIÓN CRUZADA

5. Validación Cruzada *k-Fold* Estratificada:

- ❖ Se utiliza para conjuntos de datos desbalanceados.
- ❖ Se asegura que cada *fold* tenga una distribución similar a la del conjunto original.
- ❖ No es adecuada para datos de series temporales.

📌 *Ejemplo:* Para un modelo de detección de cáncer donde solo el 5% de los datos son positivos, aseguramos que cada *fold* tenga aproximadamente un 5% de casos positivos y 95% negativos, para evitar sesgos.



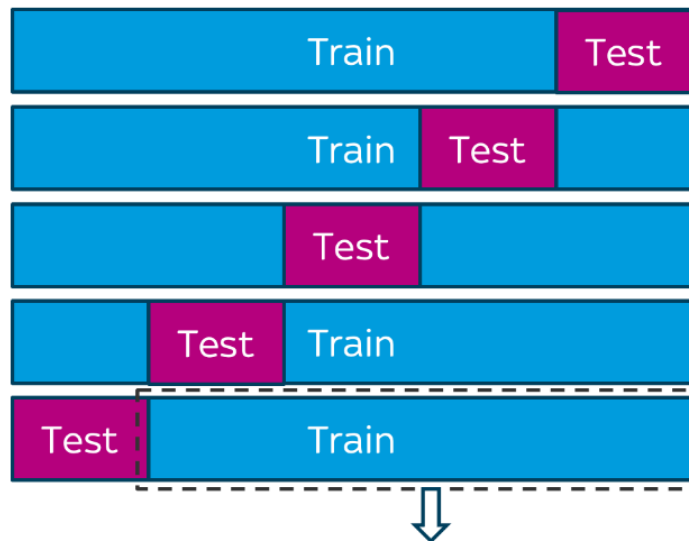
VALIDACIÓN CRUZADA

6. Validación Cruzada *k-Fold* Anidada:

- ❖ Se emplea cuando se necesita ajustar hiperparámetros y evaluar el modelo simultáneamente.
- ❖ Se usa un esquema de doble validación cruzada (una para selección de hiperparámetros y otra para evaluación del modelo).

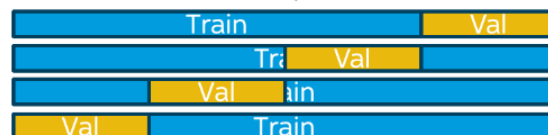
Outer loop:

Evaluate model on each `Test` split



Inner loop:

Tune hyperparameters



📌 *Ejemplo:* Para un modelo de clasificación de especies de plantas con múltiples hiperparámetros, usamos un *k-fold* externo para evaluar el modelo final y un *k-fold* interno dentro de cada *fold* externo para ajustar los hiperparámetros.