

HENRY



Evaluación de modelos II

Data Science





Agenda



- Validación Cruzada
- K-fold Cross Validation
- Validación Cruzada Aleatoria
- Curvas de Validación
- Optimización de Hiperparámetros
- Random Search
- Descenso de Gradiente



OBJETIVOS DE LA CLASE

Al finalizar esta lecture estarás en la capacidad de...

- Comprender técnicas de Optimización de Entrenamiento de Modelos



Al **finalizar** cada uno de los temas,
tendremos un **espacio de consultas**.



Hay un **mentor** asignado para
responder el **Q&A**.

¡Pregunta, pregunta, pregunta! :D



Validación cruzada





¿Cómo?

¿Cómo podemos evaluar si el modelo está aprendiendo o no de nuestros datos?

Una forma práctica de hacerlo es observar su desempeño frente a nuevas instancias.



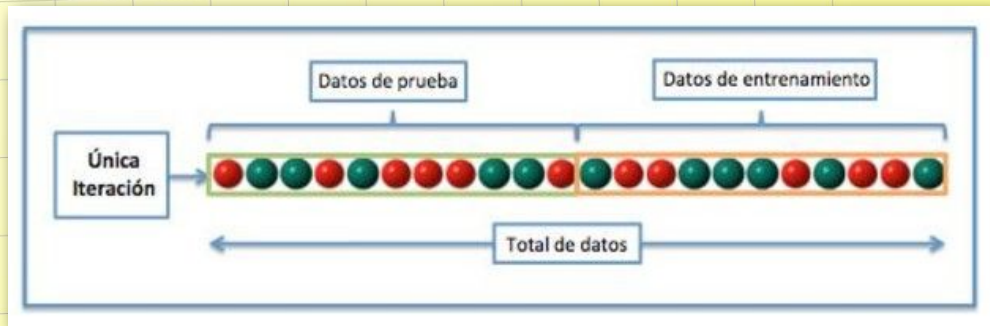


¿Cómo?

En nuestro flujo de trabajo, tendremos que emular una situación donde el modelo es entrenado con ciertos datos y luego es evaluado con datos nuevos.

Train Test Split:

- Separo los datos en dos conjuntos, Train y Test.
- Entreno con los datos de Train
- Evalúo el desempeño con los datos de Test.





Beneficios

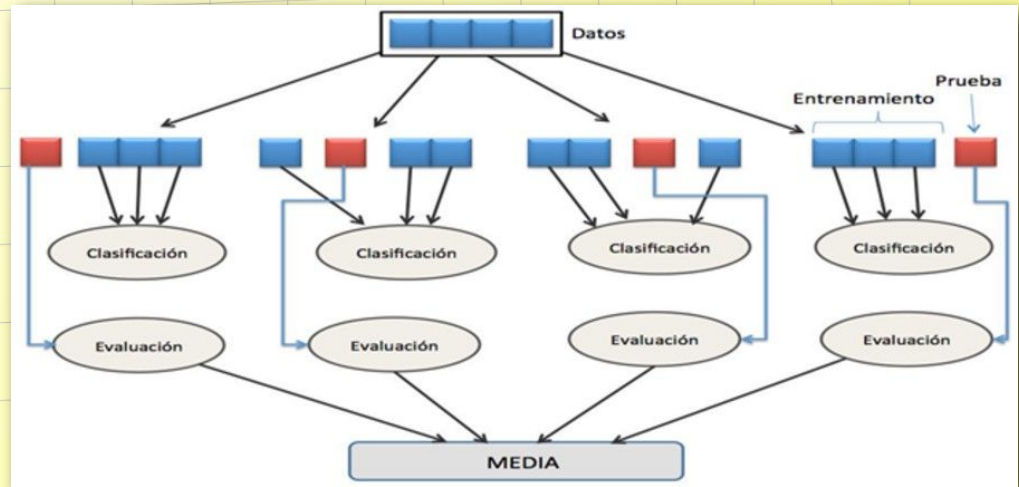
Evaluar el desempeño del sobreajuste de Test tiene varios usos:

- Obtenemos una evaluación realista del desempeño de nuestros modelos.
- Nos permite seleccionar el modelo que mejor desempeña sobre nuestros datos.

Objetivo

El objetivo de la validación cruzada es obtener una evaluación de performance de nuestro modelo que sea independiente de la partición en entrenamiento y prueba de los datos.

Haciendo muchas particiones esperamos que la medida de performance sea independiente de la partición de los datos.





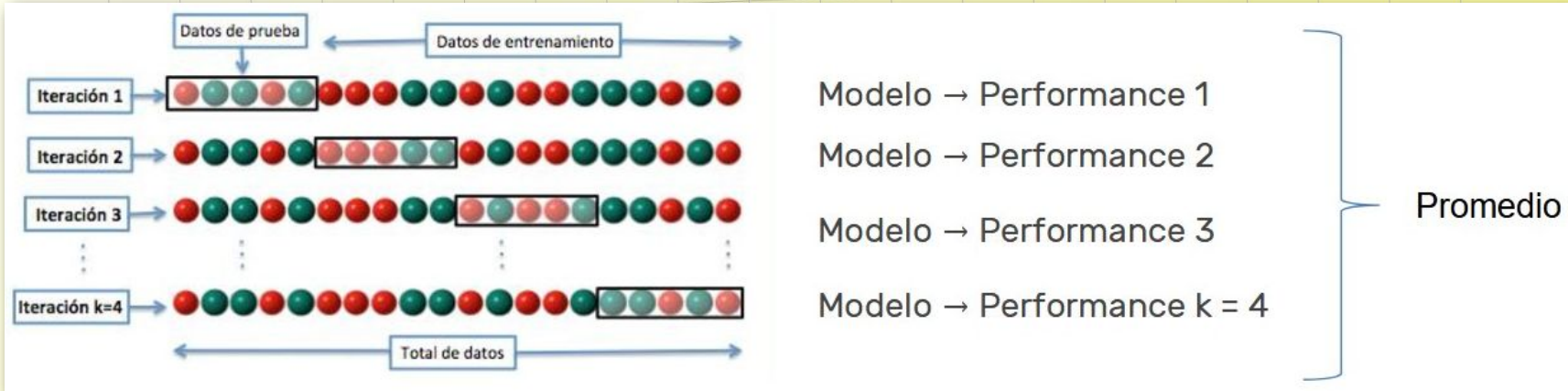
K-fold Cross validation





¿Cómo?

Es importante notar, que **cada dato aparece una sola vez** en los datos de prueba y $k-1$ en los datos de entrenamiento.





Conclusiones

- La validación cruzada es un **procedimiento de remuestreo** que se utiliza para evaluar modelos de aprendizaje automático en una muestra de datos limitada.
- El hiperparámetro más importante es **k** que se refiere al **número de grupos** en que se dividirá una muestra de datos dada.
- La validación cruzada está íntimamente relacionada con la **optimización de hiperparámetros**.



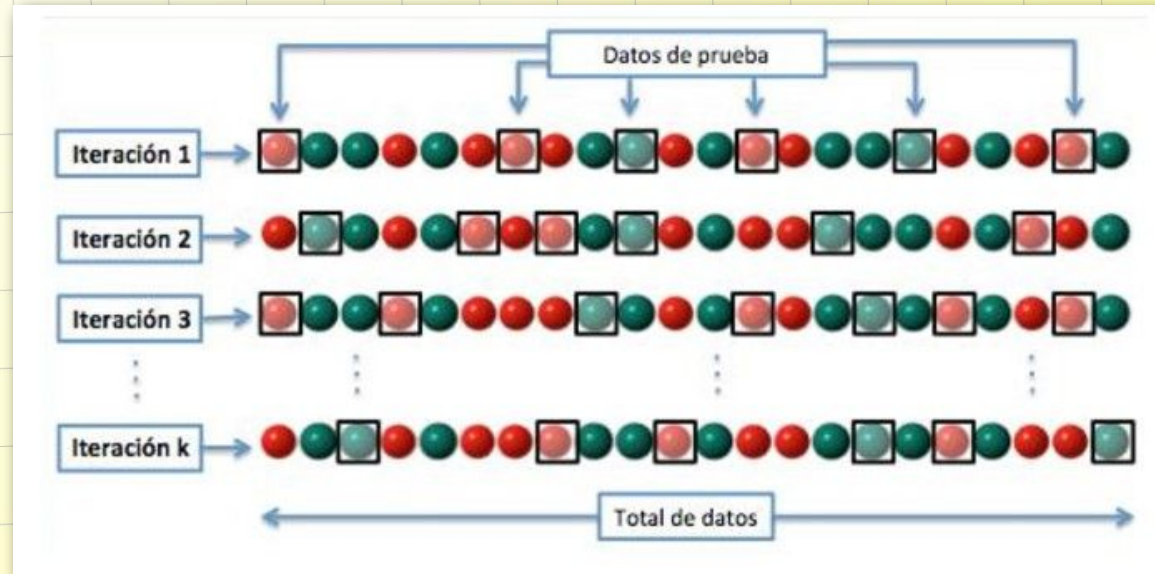
validación cruzada aleatoria





Evaluación de modelos

En este caso, cada dato puede aparecer más de una vez en el conjunto de prueba.





curvas de validación

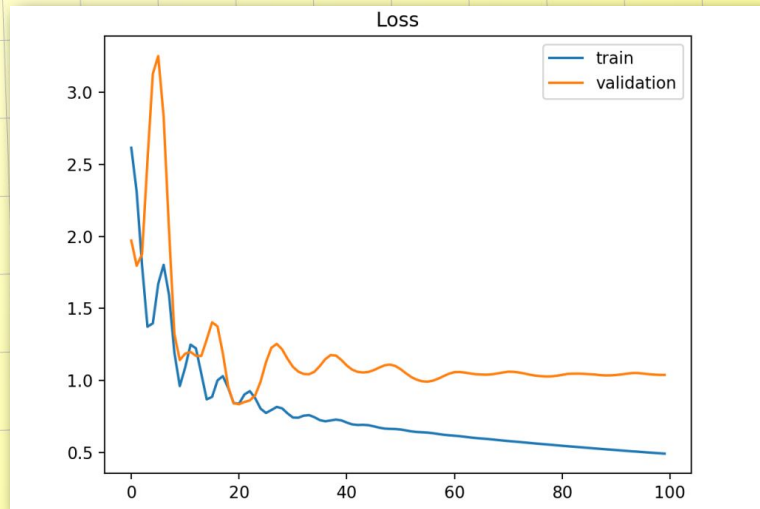




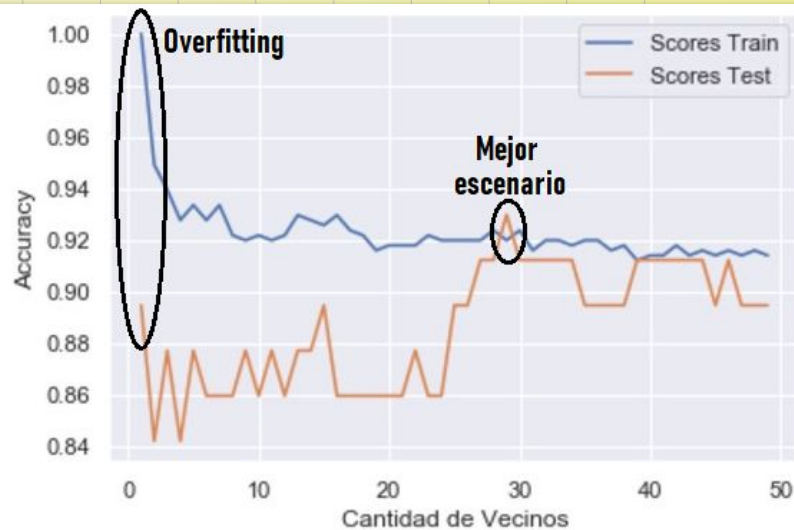
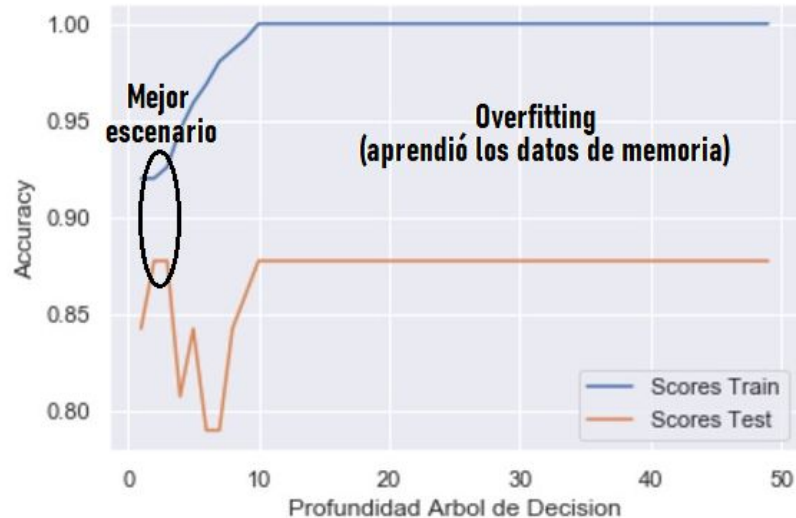
¿Para qué?

En general, el desempeño de un modelo depende de muchos hiperparámetros. Pero a veces hay uno que es el más importante, el que predomina sobre el resto.

Para elegir el valor de ese hiperparámetro - y también caracterizar mejor el desempeño de nuestro modelo -, es útil obtener las curvas de validación.



Ejemplos





Optimización de hiperparámetros





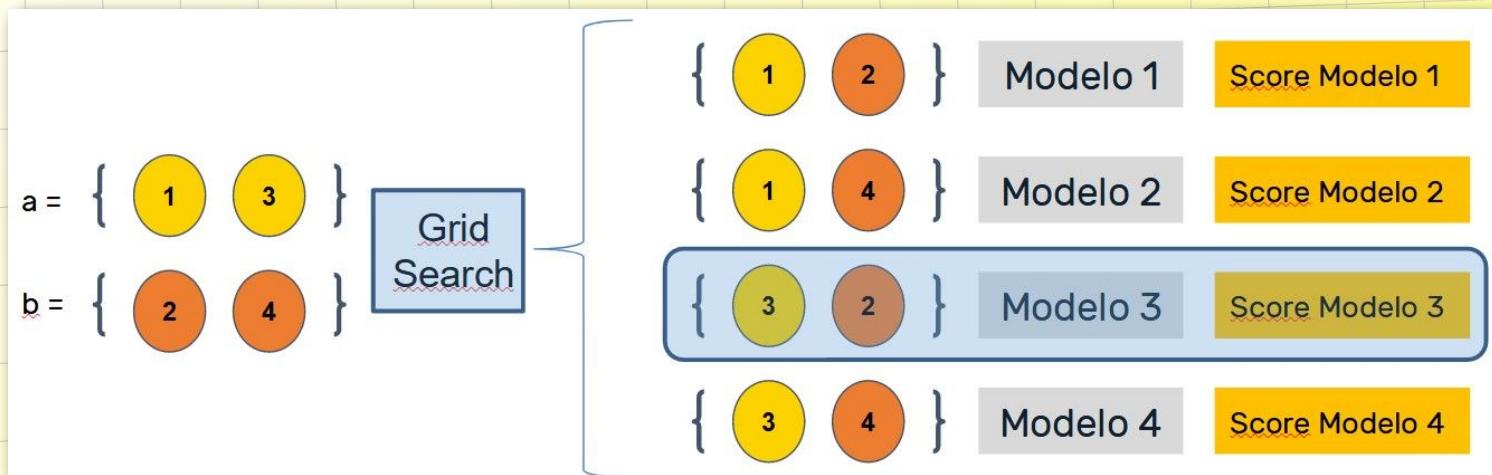
¿Cómo?

- ¿Cómo elegimos los mejores hiperparámetros para nuestro problema?
- ¿Qué es mejor, exactitud, precisión o exhaustividad? ¿Área bajo la curva ROC?
- Primero, se debe definir una métrica a optimizar. Una vez que se sabe cuál métrica optimizar, hay que probar los distintos valores de hiperparámetros.
- Se debe hacer una búsqueda exhaustiva. Es decir probando con todos los valores de los hiperparámetros que podamos y eligiendo la mejor combinación. Este método se llama Grid Search ("búsqueda de cuadrícula").



Ejemplo Grid Search

Si tenemos dos hiperparámetros, a y b , que pueden tomar valores $a = \{1,3\}$ y $b = \{2,4\}$





Ejemplo Grid Search

Pasos:

1. **Elegir** los valores que puede tomar cada hiperparámetro.
2. Armar las combinaciones “**todos con todos**” → Armar la grilla.
3. **Recorrer** la grilla entrenando el modelo para cada combinación y **evaluarlo**.
4. Elegir los hiperparámetros que **definen el mejor** modelo.



Random Search

Si por ejemplo, se tienen cinco hiperparámetros y cinco valores para probar por hiperparámetro, el tamaño de la grilla comienza a crecer.

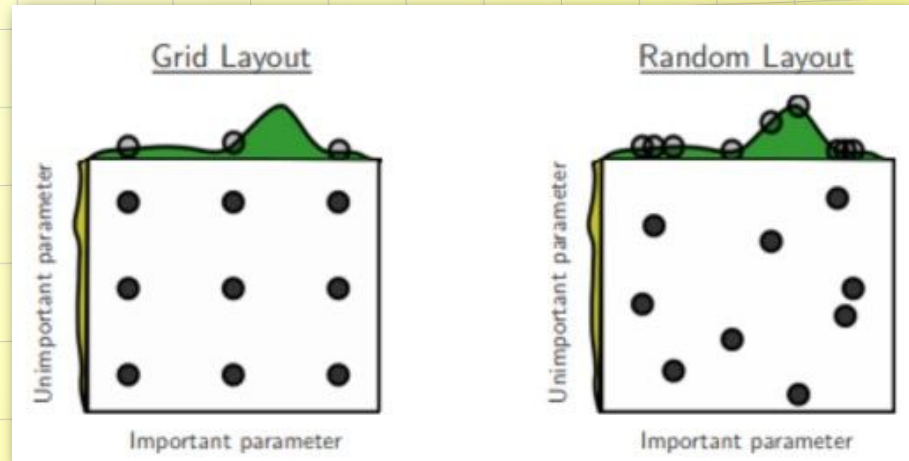
Además, para cada modelo se debe hacer la Validación Cruzada. Este proceso puede ser computacionalmente muy demandante.



Random Search

Random Search explora opciones y combinaciones al azar, de manera menos “ordenada”.

En muchas circunstancias, esto es más eficiente, tanto desde el punto de vista de performance del modelo como de desempeño computacional.





Conclusiones

- Es necesario **definir una métrica** a optimizar (exactitud, precisión, RMSE, ROC AUC, etc.).
- Un modelo (regresor o clasificador).
- Una grilla de hiperparámetros. Depende del tipo de modelo utilizado.
- Un método para buscar o muestrear los candidatos:
 - **Grid Search**: Plantea opciones y explora todas las combinaciones.
 - **Random Search**: explora opciones y combinaciones al azar.
- Crear un modelo lo antes posible, en cualquier caso, un modelo fallido muchas veces da tanta información sobre el proceso real como uno válido

¿PREGUNTAS?



¿Alguien dijo Homework?



~~HENRY~~



Próxima lecture
**Series de
tiempo**





¡Feedback!

Click on me



Dispones de un **formulario** en:



Homeworks



Guías de clase



Slack

HENRY

