

Taller validación cruzadas y métricas

Este taller es una continuación del anterior. En los grupos con las mismas personas que hiciste el taller anterior (en el cuál analizamos diferentes Dataset en Kaggle) desarrolla los siguientes ejercicios.

Punto 1.

Si el modelo que entrenaron fue de clasificación, responde la siguiente pregunta ¿Los datos en el conjunto de clases están balanceados? Has una gráfica que muestre la distribución de los datos en las clases de tu problema.

Punto 2.

¿Cuál métrica es la mejor para estudiar el rendimiento de modelo? ¿Por qué? Calcula dicha métrica y en base a este valor da una valoración del rendimiento del modelo.

Punto 3.

¿Sospechas que tu modelo se está sobre-ajustando? ¿Por qué?

Punto 4.

Calcula todos las pruebas de validación cruzada vistas en clase en tu modelo, usando la métrica escogida en punto 2 ¿Qué interpretación puedes hacer de las pruebas de validación cruzada realizadas? ¿Cuál sería la más adecuada para tu dataset?