



Aprendizaje Supervisado - Clasificación

Tutoría

¿Qué es el aprendizaje supervisado?



Autoaprendizaje

Recursos asincrónicos

- ¿Revisaste los recursos de la semana 4 (Guía y desafío)?
- ¿Tienes dudas sobre alguno de ellos?



Ideas fuerza



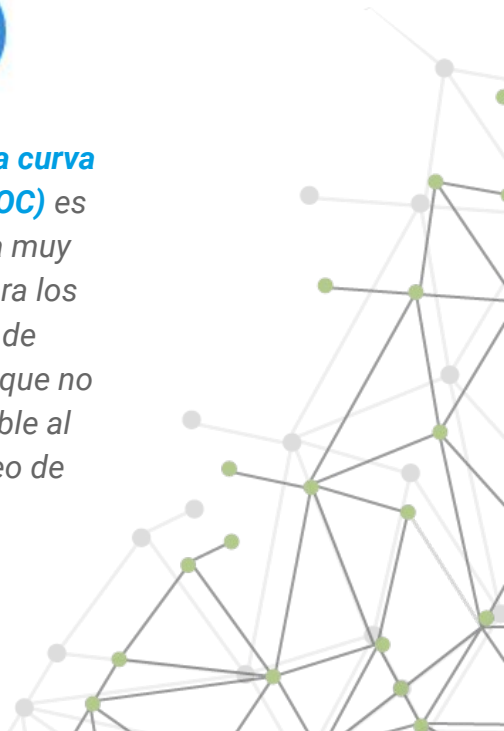
Los conceptos de **Overfitting** y **Underfitting** son muy importantes para reconocer la generalización de los modelos.



Aprender algoritmos muy importantes para la clasificación como son **regresión logística** y **support vector machine**.



El **área bajo la curva ROC (AUC ROC)** es una métrica muy relevante para los modelos de clasificación que no es susceptible al desbalanceo de clases.

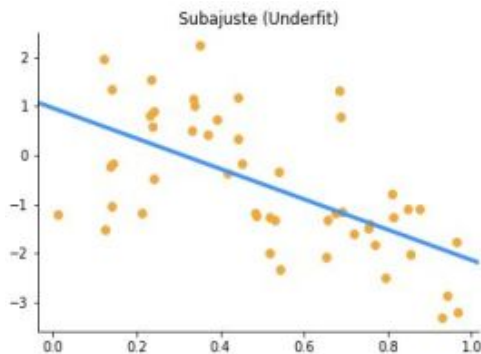


/* Overfitting y Underfitting */

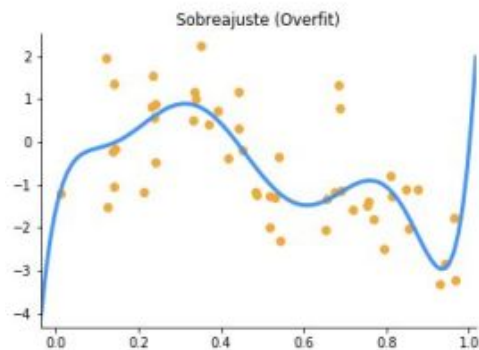
Sobreajuste y Subajuste

Comparación

1. Alto Sesgo (forma inflexible)
2. Menor desempeño en clasificación
3. Error generalizable



1. Alta Varianza (forma acoplada)
2. Mayor desempeño en clasificación
3. Error poco generalizable



/* Regresión Logística */

Regresión Logística

¿En qué consiste?

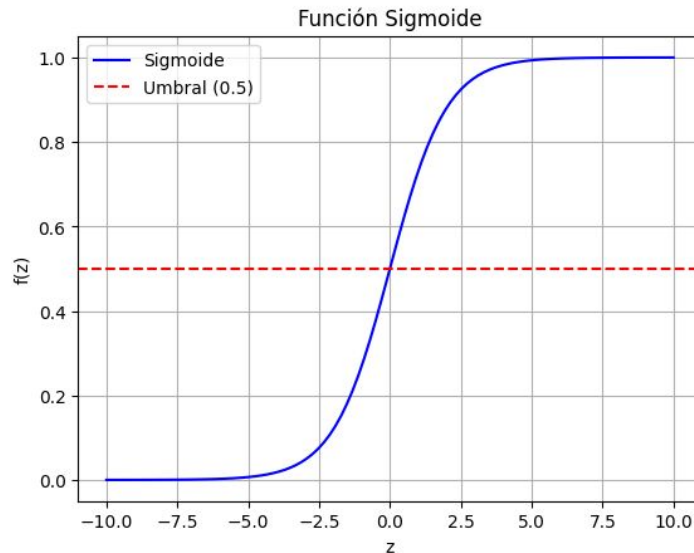
Modelo de clasificación que utiliza la **función logística** (función sigmoide) para predecir la probabilidad de que una observación pertenezca a una clase.

$$f(z) = \frac{1}{1+e^{-z}}$$

$$z = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$$

Probabilidad de pertenecer a la clase 1

$$P(y = 1|X) = f(z)$$



Regresión Logística

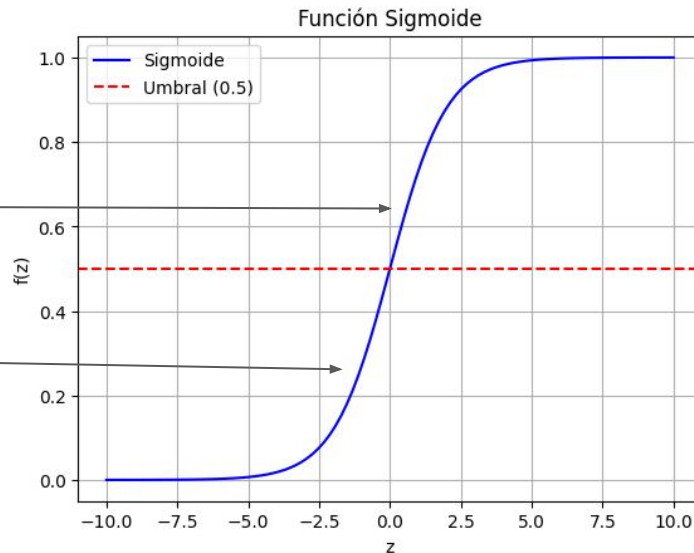
¿En qué consiste?

Modelo de clasificación que utiliza la **función logística** (función sigmoide) para predecir la probabilidad de que una observación pertenezca a una clase.

Probabilidad de pertenecer a la clase 1

$$P(y = 1|X) = f(z)$$

$$P(0) = 1 - P(y = 1|X)$$



Ecuación de Máxima Verosimilitud

¿Cómo se aplica?

Función de probabilidad conjunta

La probabilidad conjunta de los datos observados es la multiplicación de las probabilidades individuales de los datos.

$$P(Y_i = 1|X) = p_i$$

$$L(\beta) = \prod P(Y_i = 1|X)$$

$$\log(L(\beta)) = \sum \log(P(Y_i|X_i; \beta))$$

¡Veamos con Python!



Regresión Logística con Python

Aplicando nuestro código

Vamos a ver de manera práctica la implementación de la regresión logística con Python. Para ello, puedes abrir tu propio archivo y seguir los pasos que te mostrará tu profesor en la presentación. En esta abordaremos:

1. Regresión logística
2. Support vector machine
3. Hiperparámetros
4. Métricas - ROC AUC



Desafío

“Clasificación de calidad de ramen”

- ¿Leíste el desafío de esta semana? ¿Comprendes bien lo que se solicita en cada caso?
- ¿Hay contenidos que necesitas repasar antes de comenzar este desafío?
- ¿Necesitas algún ejemplo o indicación para alguna pregunta o requerimiento específico?





Próxima sesión...

*Veremos modelos de aprendizaje supervisado,
específicamente estudiaremos modelos de regresión en
conjunto con las métricas correspondientes.*

{desafío}
latam_

*Academia de
talentos digitales*

