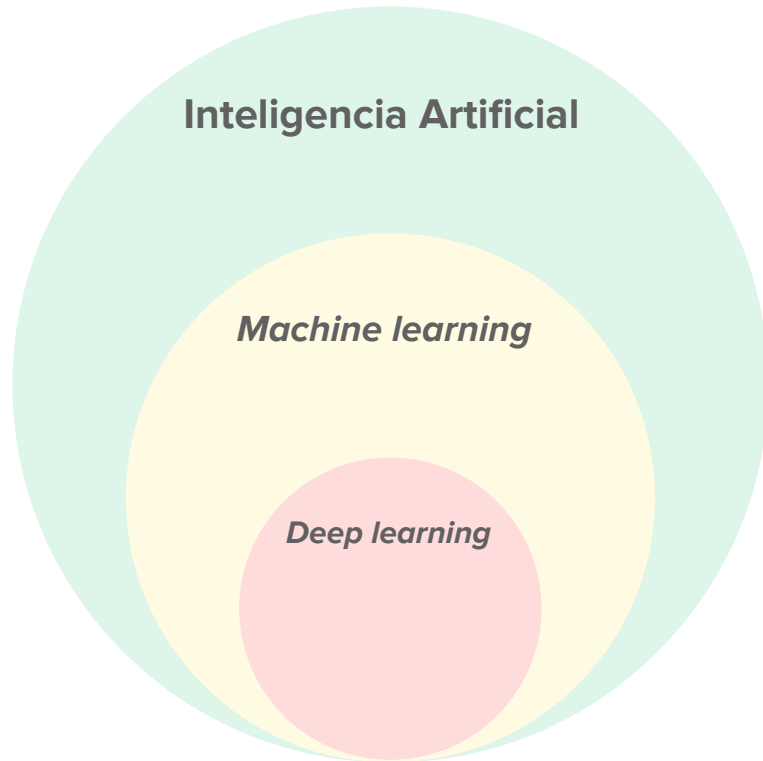


Clasificación binaria, multiclase y métricas. Curva ROC

Juan José Silva Torres

Qué es la modelización supervisada

IA, Machine Learning y Deep Learning



¿Qué es el *machine learning*?

- **Arthur Samuel.** Campo de estudio que da a las computadoras la capacidad de aprender sin ser programadas explícitamente.
- **Tom Mitchell.** El estudio de algoritmos de computación que mejoran automáticamente su rendimiento gracias a la experiencia. Se dice que un programa informático aprende sobre un conjunto de tareas, gracias a la experiencia y usando una medida de rendimiento, si su desempeño en estas tareas mejoran con la experiencia.




Algoritmos supervisados

Existe una **variable objetivo** o target que se desea estimar o predecir.

El modelo aprende a partir de un conjunto de datos **etiquetados**.

Es posible medir la **precisión** y **bondad del modelo**.

Precio de una vivienda

	Variables explicativas				Target
	Metros cuadrados	Nº de habitaciones	Terraza	...	Precio
	102	4	Si		375.000 €
	81	3	No		250.000 €
...		...			
	77	2	No		220.000 €

Regresión vs Clasificación

- La principal diferencia entre la regresión y la clasificación reside en la **naturaleza del target**.
- En el caso de que la variable sea de tipo **cuantitativa continua** se utilizan **algoritmos de regresión**.
- Cuando la variable es de **tipo cualitativo o cuantitativo discreto** entonces se habla de **algoritmos de clasificación**.

Clasificación binaria vs multiclase

- Dependiendo de la naturaleza del target la clasificación será binaria o multiclase.
- Si solo existen **dos posibles valores** para el target entonces se habla de clasificación **binaria**.
- En el caso de que el target pueda tomar **más de dos valores categóricos distintos** entonces se habla de **multiclase**.

Algunos ejemplos

Binaria

- Detección de enfermedades como cáncer o tumores.
- Detección de una transacción fraudulenta.
- Propensión de fuga
- Propensión a accidente
- ...

2 categorías

Multiclase

- Identificación de objetos.
- Clasificación del estadio de una enfermedad.
- Detección de lengua.
- ...

Más de 2 categorías

Conclusiones

En función de target a modelizar
se escogerá un tipo de algoritmo
u otro

¡Gracias!

Contacto:

juanjo.silva.t@gmail.com