

Métodos estadísticos en Data Science

Regresión lineal

> La cuna del machine learning

- Algoritmo básico
- Es un conocimiento indispensable
- Conceptos como la beta de las variables, p-valor o R Cuadrado

Regresión logística

> Para resolver problemas de clasificación

- Modificar la regresión lineal
- El 80% de los proyectos son problemas de clasificación

Series temporales

> Estudio de variables temporales

- Modificación de las regresiones lineales
- Conceptos como estacionalidad y tendencias a través de ARIMA

Árbol de decisión

> De cara a presentar resultados

Algoritmos muy usados:

- Random Forest
- XGBoost

Red Neuronal

> Algoritmos potentes y oscuros

- Acierta, pero no el porqué
- Las técnicas de Deep Learning se basan en redes neuronales

K-means

> Para agrupar y clasificar por características

- Conceptos como normalización o distancias

k-NN

> También llamado k-vecinos

- Se usa para los recomendadores

PCA

> Reducción de dimensiones

- Muy usada en las encuestas

Cadenas de markov

> Para pocos datos

- Estadística bayesiana
- Conceptos como la simulación
- Pistas al modelo para que aprenda

Simplex

> Búsqueda de soluciones óptimas

- Último paso de los modelos
- De predicción a descripción: Para maximizar las ventas



Firmado por: **Iván Robles**

Senior Data Scientist en Orange España
Profesor en el Master en Big Data Management de ICEMD.

¿Quieres especializarte en Big Data?
Master en **Big Data Management**

