

# Pràctica de Ciència de Dades: Introducció a l'Anàlisi del Titanic amb Xarxes Neuronals

L'objectiu d'aquesta pràctica és introduir-se en l'anàlisi de dades i el modelatge predictiu mitjançant l'exploració de les dades dels passatgers del Titanic. Utilitzarem mètriques bàsiques per avaluar models i entrenarem una xarxa neuronal simple per predir la supervivència dels passatgers.

## Descripció de la Tasca:

S'ha de realitzar un anàlisi exploratòri bàsic de les dades, prepararà les dades per a l'entrenament d'un model, i desenvolupar una xarxa neuronal simple per a la tasca de classificació.

## Tasques:

1. Carrega el dataset i realitza una anàlisi exploratòria bàsica. Examina les primeres files, la distribució de les variables clau (com Survived, Sex, Age, etc.), i la relació entre elles.
2. Selecciona un conjunt de característiques simples per a l'entrenament del model, com Pclass, Sex, Age, i Fare. Converteix les variables categòriques en numèriques utilitzant codificació binària o one-hot encoding.
3. Divideix les dades en un conjunt d'entrenament i un de prova.
4. Utilitza una biblioteca com TensorFlow o Keras per dissenyar una xarxa neuronal amb una o dues capes ocultes com a màxim. Mantenim l'arquitectura simple: un

nombre modest de neurones i activació ReLU poden servir amb una funció de sortida sigmoid.

5. Entrena la xarxa amb el conjunt d'entrenament i utilitza el conjunt de prova per avaluar-la. Enfoca't en mètriques com la precisió (accuracy), la pèrdua (loss), i si vols, la precisió (precision).
6. Presenta els resultats de l'entrenament, incloent les mètriques seleccionades. Discuteix breument sobre el rendiment del model i possibles raons darrere dels resultats obtinguts.
7. Reflexiona sobre com les característiques seleccionades podrien afectar el rendiment del model i què podries fer per millorar els resultats en futures pràctiques.

#### **Lliurament:**

El treball es lliurarà en forma de notebook Jupyter, que inclourà totes les fases de l'anàlisi: des de la càrrega de dades fins a l'entrenament del model, incloent codi, gràfiques pertinents, i comentaris que expliquin el raonament darrere de cada decisió presa i els resultats obtinguts.