

- 1) Descrivere la Pipeline di classificazione
- 2) Domande sulla rappresentazione di feature categoriche. p.e. Come rappresentare l'età o il porto di partenza delle persone nel dataset TITANIC
- 3) Misure di correlazione
- 4) Underfitting e Overfitting. Come modificare i parametri esterni di un classificatore.
- 5) Model selection, Holdout, CrossValidation, LeaveOneOut CV
- 6) Valutazione per problemi di Classificazione (acc,precision,recall, F1)
- 7) Descrivere il problema del clustering (flat o gerarchico)
- 8) Criteri esterni e interni per valutare un sistema di clustering
- 9) Tipi di Linkage per lo HAC e loro peculiarità
- 10) Sistemi di raccomandazione, tipologie, approcci, misure di valutazione