

ASML: Laboratorio 8

Confronto fra tecniche di classificazione

N. Torelli

2024

Università di Trieste - Dipartimento di Scienze Economiche Aziendali Matematiche e Statistiche "Bruno de Finetti"

Scopo del laboratorio è quello di utilizzare gli alberi di regressione e altri classificatori e di confrontarne le performance

1. Si considerino i dati su kaggle in <https://www.kaggle.com/datasets/shrutimechlearn/churn-modelling>. Si tratta di prevedere la decisione dei clienti di abbandonare la banca.
2. Si suddivida casualmente il dataset in due parti: l'insieme di apprendimento (training) e quello di valutazione (Test set)
3. Si usino diversi modelli di classificazione (Alberi di classificazione, GLM, GAM, LDA, LQA, NaiveB, KKN) eventualmente usando diversi parametri di tuning.
4. Si ottengano opportuni indici per valutare la qualità della classificazione (accuratezza, recall, specificità, F1, AUC)
5. Si disegnino le curve ROC per i diversi classificatori