Supervised learning 🡪È un analisi caratterizzata da attributi🡪 esplicative e classi di appartenenza 🡪 target

Classificazione lineare semplice 🡪 linearmente separabili

Target 2 valori 🡪 classificazione binaria

**Hold out** : ripartizione in due insiemi, training e test di norma 2/3 delle osservazioni. L’accuratezza dipende dalla scelta del test et, è possibile sovra o sotto stimare l’accuratezza

**Cross Validation**: si randomizzano le osservazioni, si dividono le osservazioni in k gruppi (es 4-fold cross validation k=4) e quindi otterrò (che poi chiamerò G1,G2,G3,G4…)

Addestramento G1 + G2 + G3. Calcolo Accuratezza del classificatore su G4. E ottengo il valore A4

Addestramento G1 + G2 + G4. Calcolo Accuratezza del classificatore su G3. E ottengo il valore A3

Addestramento G1 + G3 + G4. Calcolo Accuratezza del classificatore su G2. E ottengo il valore A2

Addestramento G2 + G3 + G4. Calcolo Accuratezza del classificatore su G1. E ottengo il valore A1

Accuratezza finale : (A1 + A2 + A3 + A4)/4

Matrici di confusione