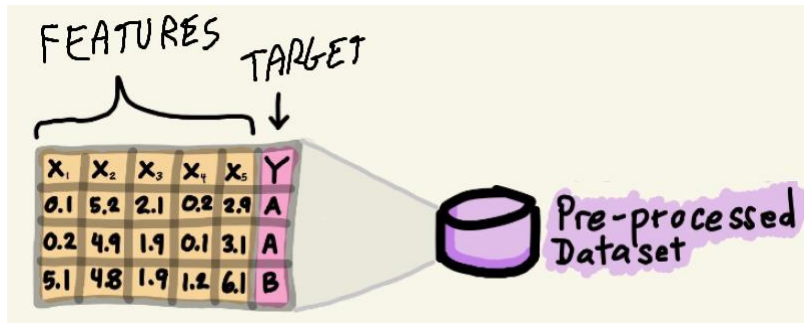


## FLOW CLASSIFICAZIONE ML

- 1) Data cleaning + fill/drop missing data + remove not useful features
- 2) Verify data dtypes
- 3) Trasformare la colonne categoriche in numeriche(compresa target)



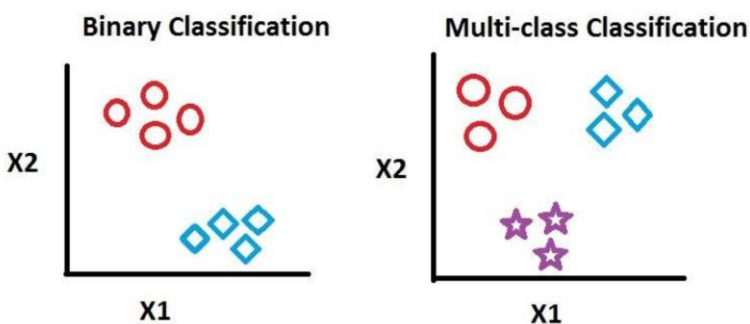
- 4) Selezione matrice features e vettore target



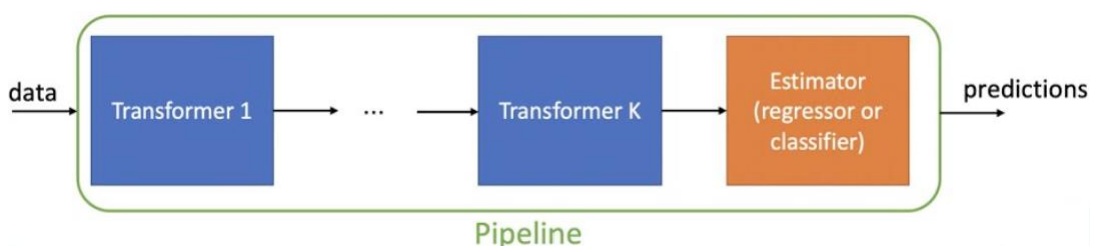
- 5) Normalizzo i dati (opzionale)
- 6) Divido il mio dataset in train e test
- 7) Dichiaro il classificatore che allenerò
- 8) Fitto il modello sulla parte di training
- 9) Verifico l'accuracy del modello



- 10) Stampo la confusion matrix del modello
- 11) Stampo il classification report per f1 score



- 12) Se il problema è **BINARIO** stampo la ROC curve
- 13) Se il problema è **BINARIO** stampo la AUC
- 14) Se ho più modelli seleziono quello con la AUC più grande.



- 15) Creare la pipeline finale di inference
- 16) Generare un array compatibile con il modello che ci restituisca la corretta prediction.
- 17) Salvare la pipeline come file \*.pkl