

# Sapore DiVino

Progetto realizzato per il corso di **Fondamenti di Intelligenza  
Artificiale**



UNIVERSITÀ DEGLI  
STUDI DI SALERNO

---

FABIO SESSA – MAT.0512114634

---

2023/2024



# INTRODUZIONE

---

L'universo del vino è un affascinante viaggio sensoriale che coinvolge tutti i nostri sensi, dalla vista al gusto, fino all'olfatto. Gli appassionati di vino apprezzano non solo la bevanda in sé, ma anche la complessità e la raffinatezza che si cela dietro ogni bottiglia.

# FORMULAZIONE DEL PROBLEMA



## Obiettivi

- Analisi dettagliata di un dataset di vini rossi.
- L'identificazione delle varie feature associate ai vini.
- L'implementazione di un modello di apprendimento in grado di classificare in modo affidabile la qualità di vini almeno sufficienti e insufficienti.

# IL DATASET

---

## LA SCELTA DEL DATASET

- Una selezione oculata del dataset contribuisce a garantire che il modello sia in grado di generalizzare bene su nuovi dati e affrontare le sfide del mondo reale.

---

## KAGGLE

- Una piattaforma online dedicata alle competizioni di data science, Machine Learning e analisi dei dati.

---

## RISULTATO

- La scelta del dataset è ricaduta su un elenco di vini rossi provenienti dal Portogallo:
  - Formato da ben 1599 righe e 12 colonne.

# DATA UNDERSTANDING



---

## COSA È

- Implica un esame più attento dei dati disponibili per il data mining.

---

## OBIETTIVO

- Questo passaggio è di fondamentale importanza per evitare problemi imprevisti durante la fase successiva di Data Preparation, che è in genere la parte più lunga di un progetto.

---

## STRUTTURA

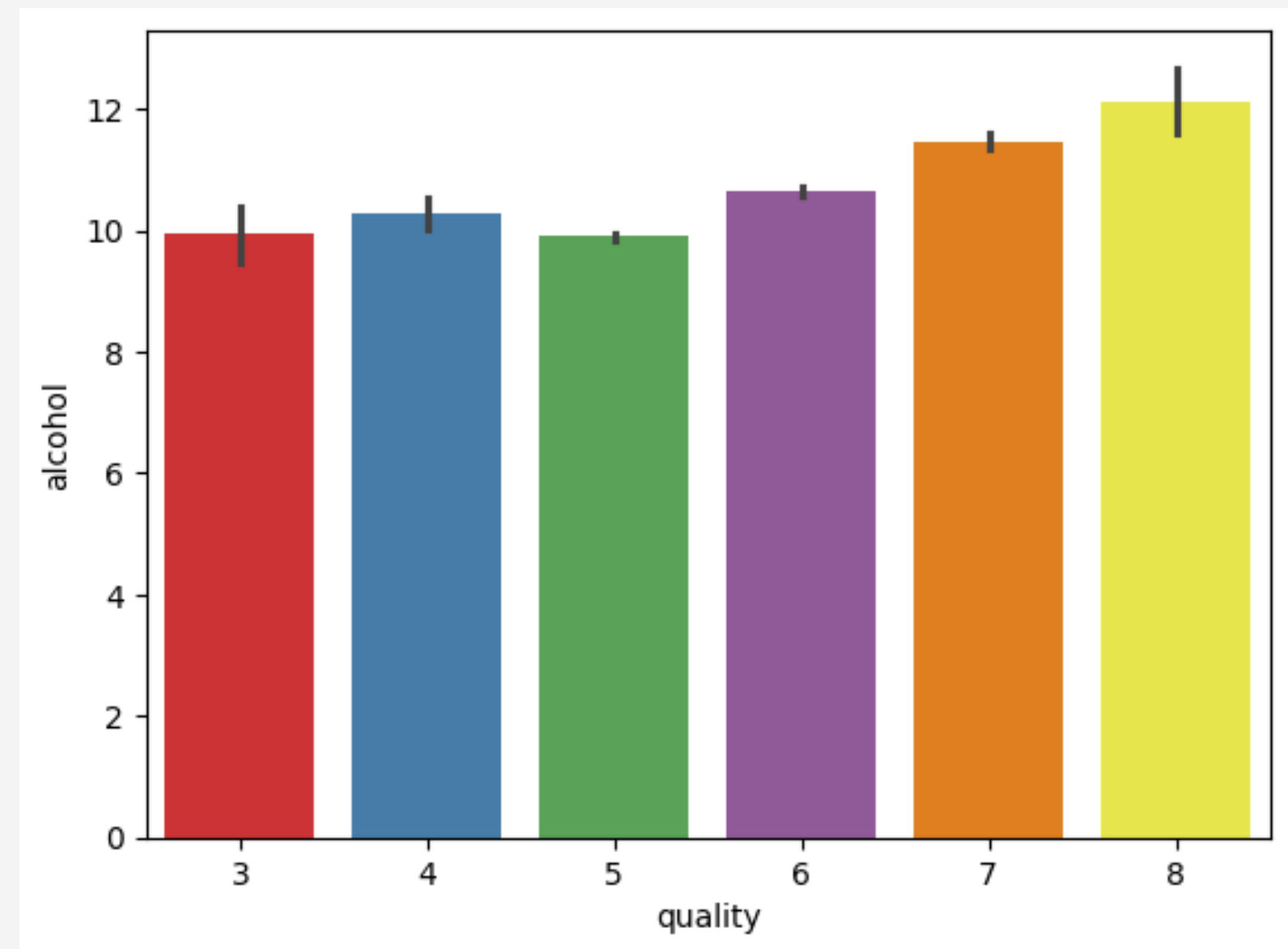
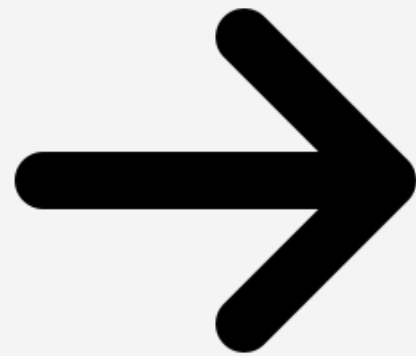
- Contiene tre differenti fasi:
  - Data Collection
  - Data Description
  - Data Quality

# DATA DESCRIPTION

## QUANTITÀ DI ALCOHOL

---

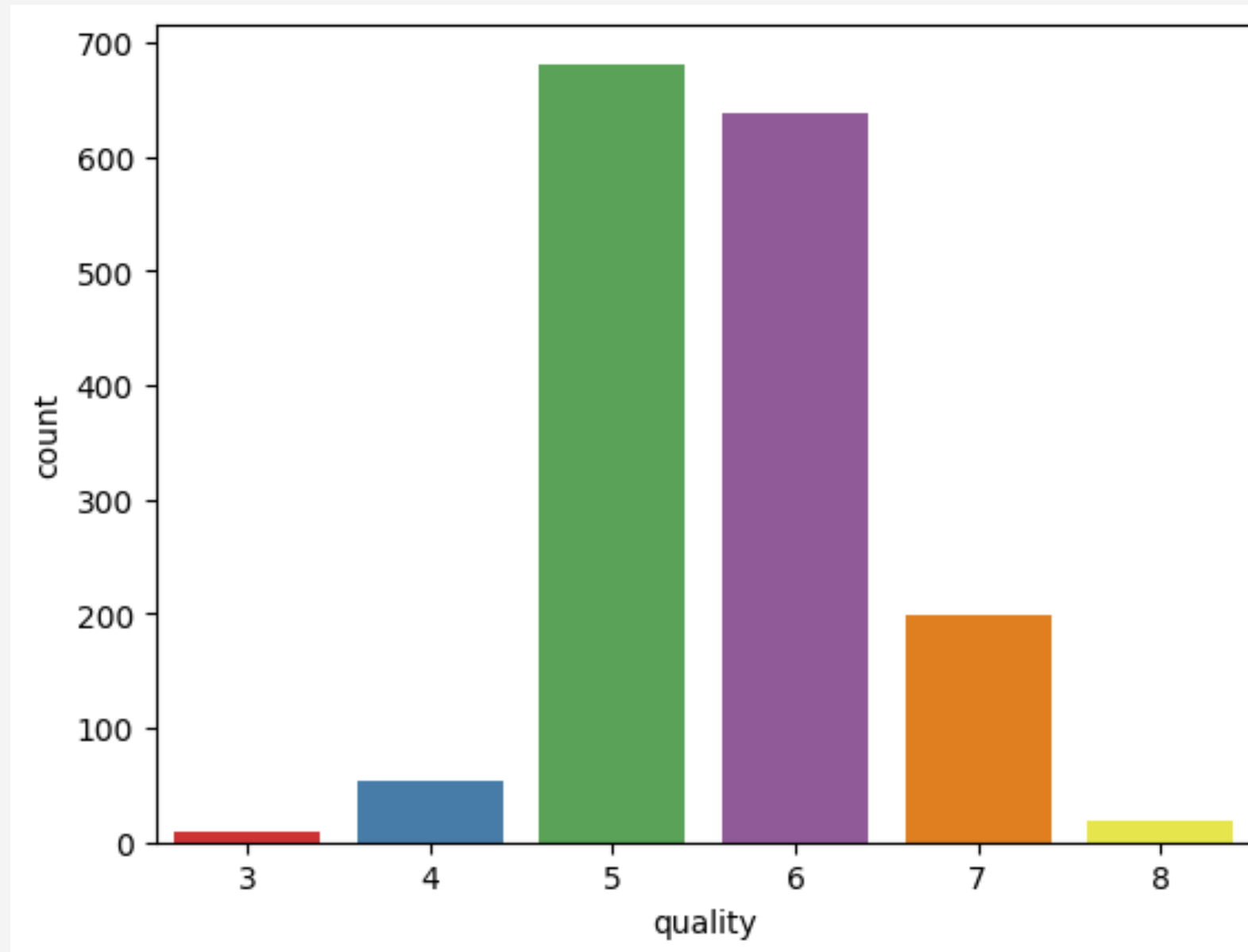
- Rappresentazione grafica della quantità di alcohol nei vini suddivisi per qualità.



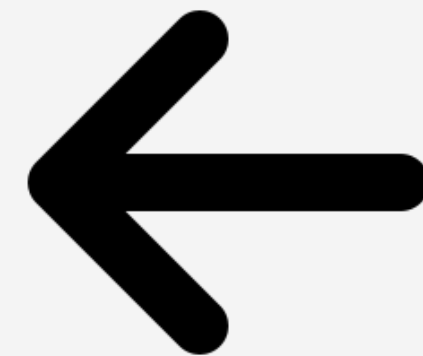
# DATA DESCRIPTION

## QUALITÀ DEI VINI

---



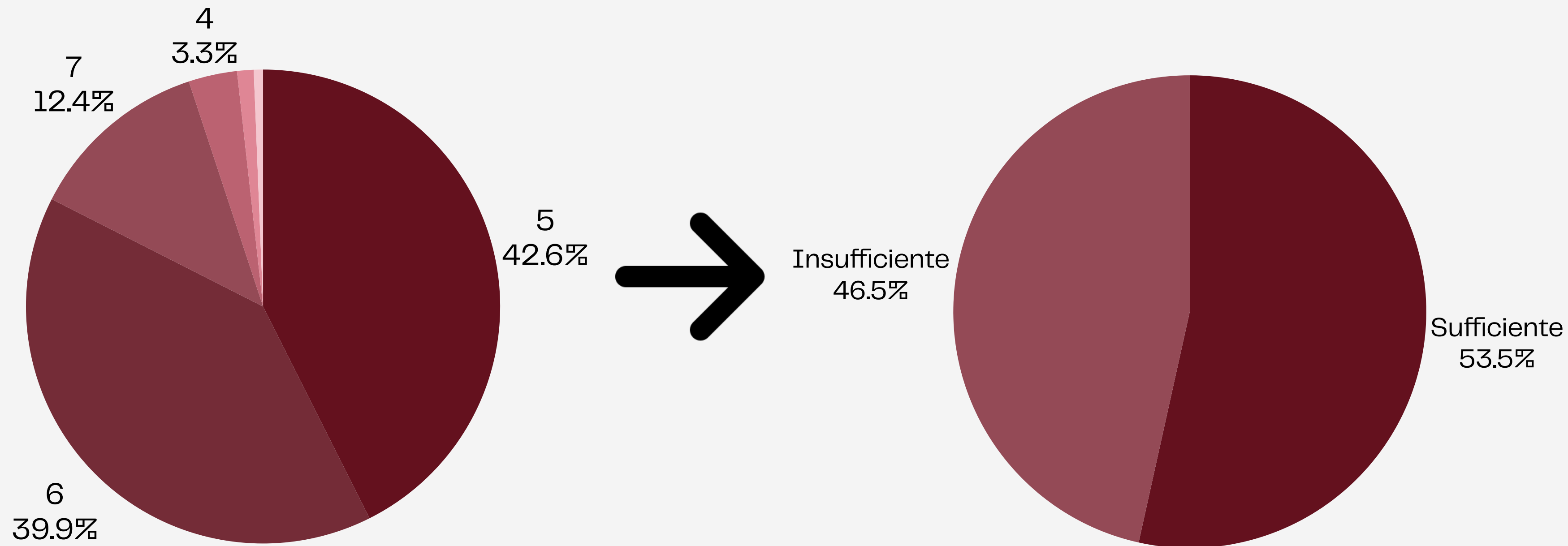
- Rappresentazione grafica riguardante la frequenza dei vini suddivisi per qualità.





# DATA QUALITY

CATEGORIZZAZIONE IN DUE DELLA QUALITÀ







# DATA PREPARATION

---

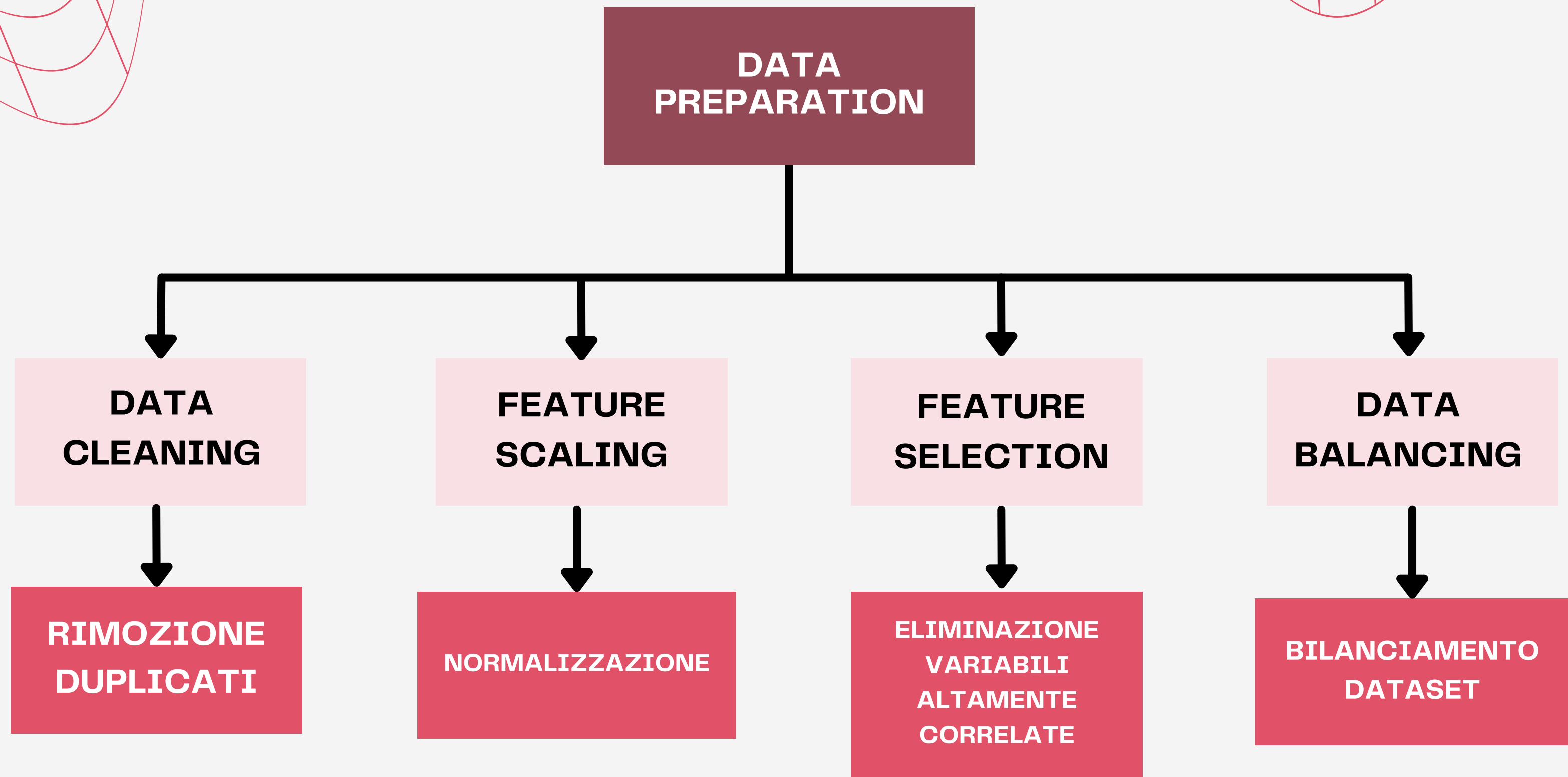
## COSA È

- È il processo di pulizia, trasformazione e organizzazione dei dati grezzi in un formato adatto all'analisi. Include operazioni come la rimozione dei dati mancanti, la normalizzazione e la gestione delle anomalie.

---

## OBIETTIVO

- Garantire che i dati siano accurati, completi e pronti per l'elaborazione da parte degli algoritmi di analisi. Ciò migliora l'affidabilità e l'efficacia delle future analisi e modelli predittivi.





# DATA MODELING

---

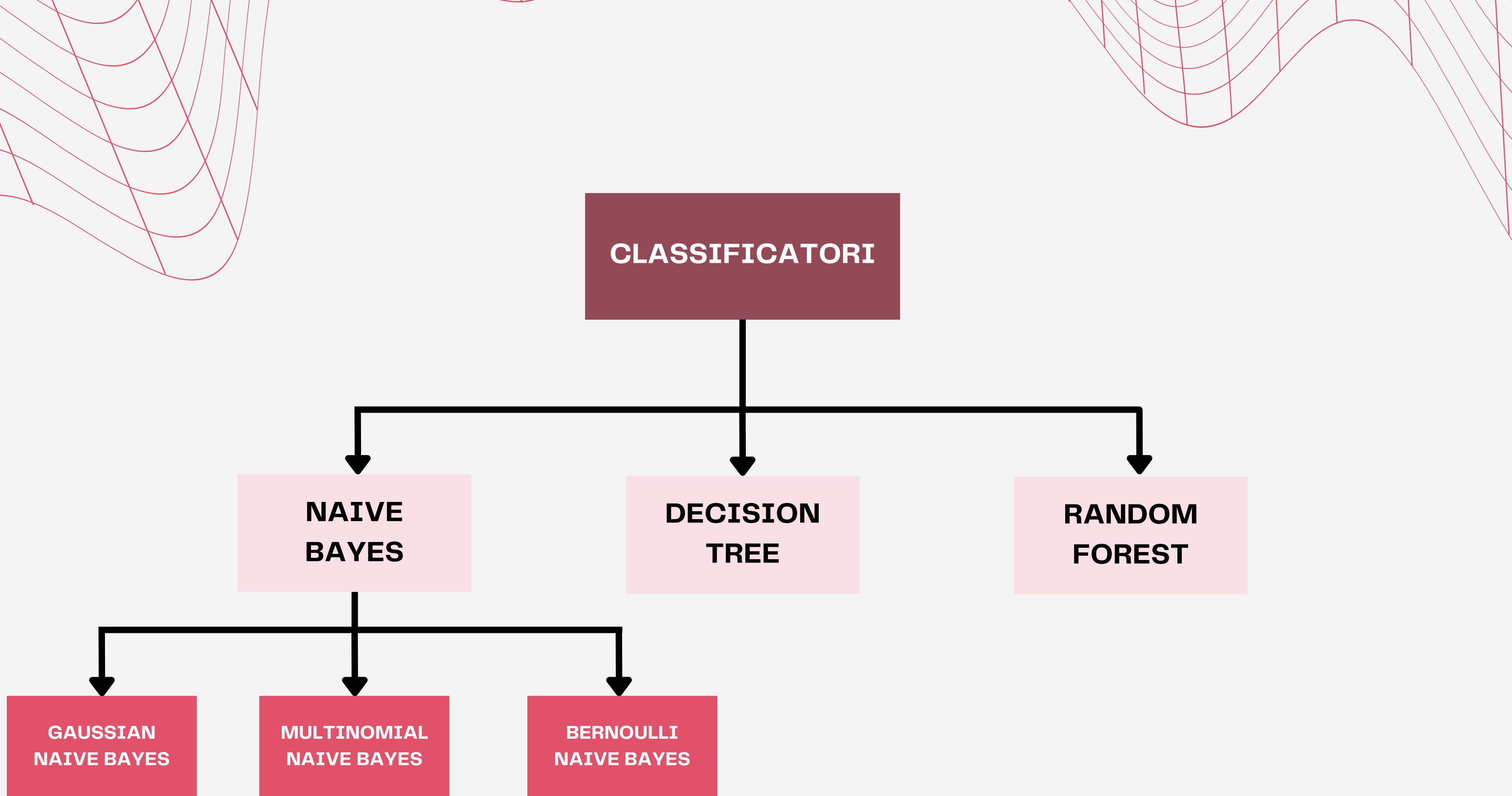
## COSA È

- Comprende diverse attività cruciali che si basano sulla rappresentazione dei dati.

---

## OBIETTIVO

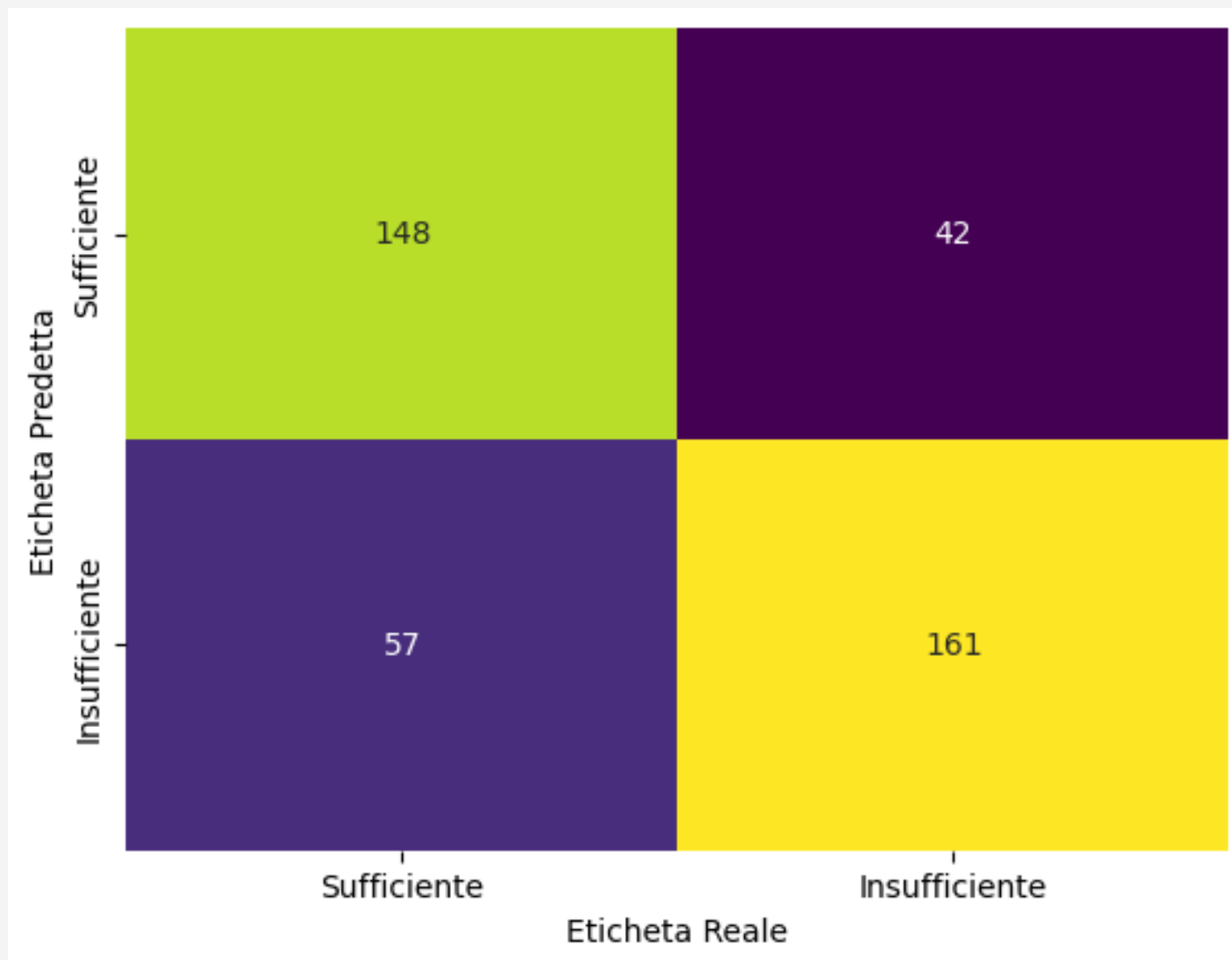
- Sviluppare modelli accurati e significativi che facilitino l'interpretazione, la manipolazione e l'analisi dei dati.



# CONFRONTO CLASSIFICATORI

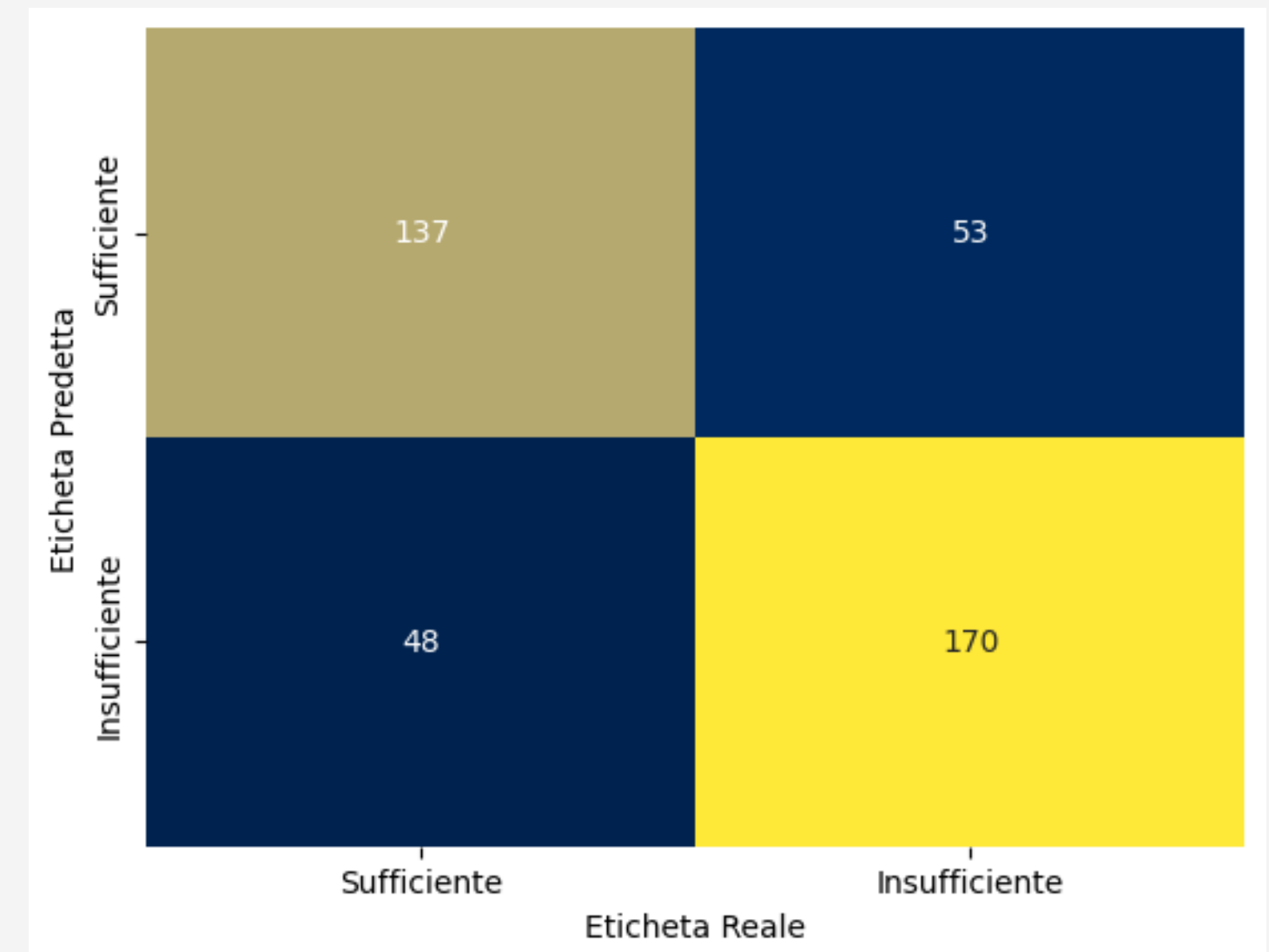
Matrice di confusione

**Gaussian Naive Bayes**



vs

**Random Forest**



# LE METRICHE

---

## COSA SONO

- Sono specifiche per il tipo di attività di Machine Learning eseguite da un modello.

---

## TIPI

- Utilizzeremo quattro metriche:
  - Precision
  - Recall
  - Accuracy
  - F-score

# CONFRONTO CLASSIFICATORI

Metriche di valutazione

## Gaussian Naive Bayes

PRECISION	RECALL	ACCURACY	F-SCORE
0.79	0.73	0.75	0.76

vs

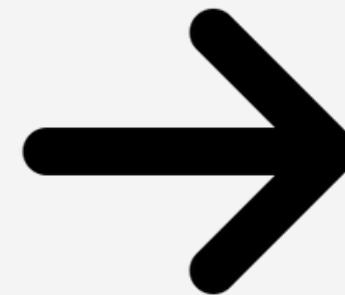
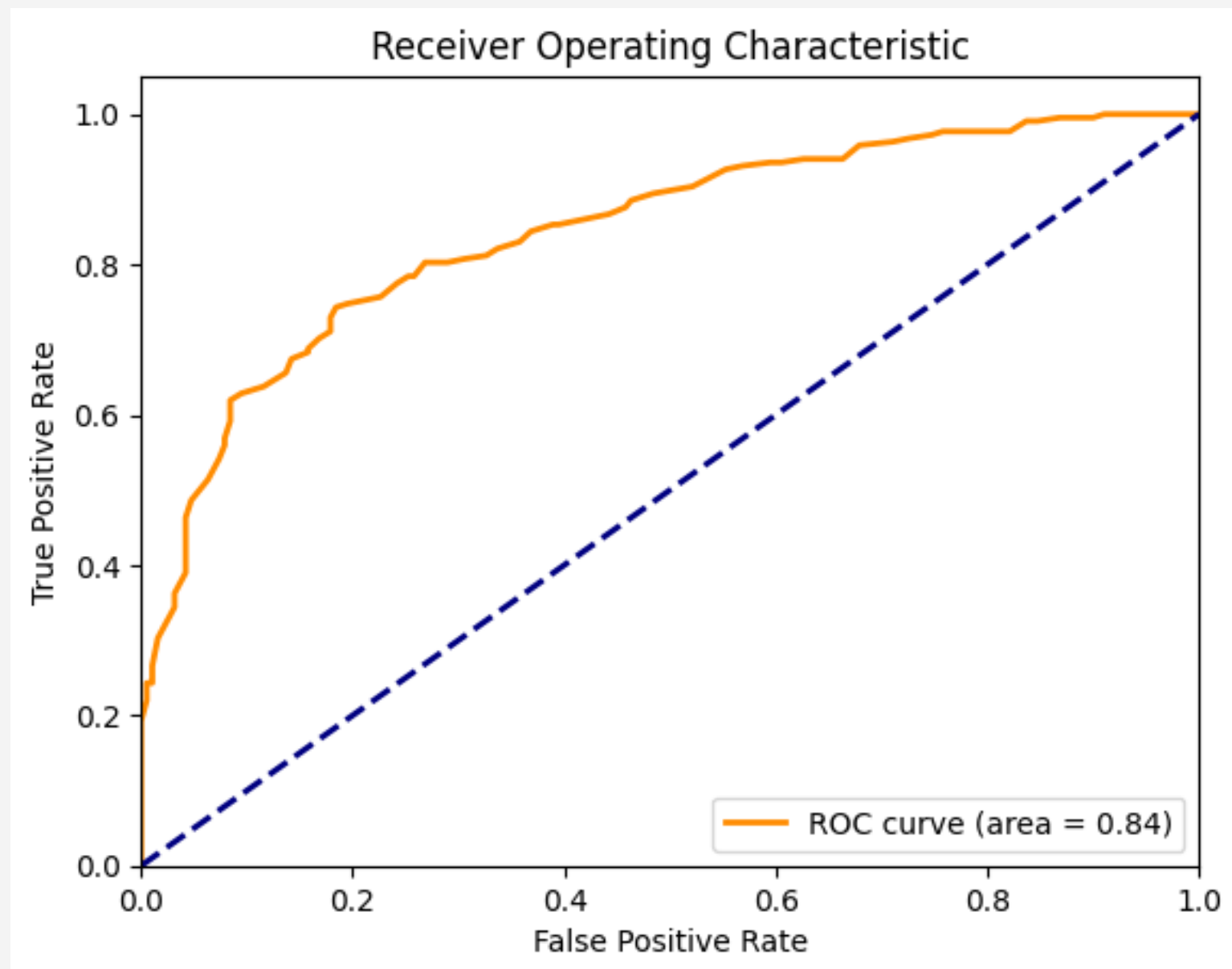
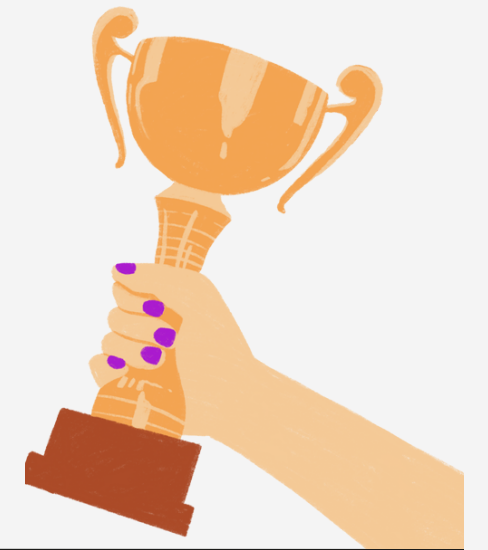
## Random Forest

PRECISION	RECALL	ACCURACY	F-SCORE
0.76	0.77	0.75	0.77



# EVALUATION

## Curva ROC – AUC



Il mio modello ha ottenuto un valore pari a 0.84, che è un ottimo risultato. Quindi posso considerare la costruzione del modello e in generale dell'approccio completa.

A warm, slightly blurred photograph of a group of people at a social gathering. Several hands are visible, each holding a glass of red wine. The glasses are being clinked together in a toast. In the background, a string of warm-toned, round light bulbs is visible, adding to the cozy atmosphere. The overall color palette is soft and warm, with a light beige or cream overlay.

**GRAZIE PER L'ATTENZIONE!**