



# Business Intelligence per Big Data

*Progetti di analisi di dati*



Data Base and Data Mining Group of Politecnico di Torino



AA 2022-2023 - *Politecnico di Torino*



# Progetti di analisi di dati

- Attività da svolgere
  - Scegliere uno *use case* tra i tre proposti
  - Caratterizzare il dataset
  - Effettuare diverse sessioni di analisi sul dataset utilizzando il tool RapidMiner e/o altri tools noti al team
  - Analizzare i risultati e sintetizzarli in grafici
  - Discutere come sfruttare la conoscenza estratta in un'applicazione di business
  - Scrivere il report scientifico sintetico sulle attività svolte
    - Utilizzare il template overleaf [link](#)
    - Rispettare la struttura data
    - Rispettare il numero di pagine massimo (4 pagine massimo)



# Formazione gruppi

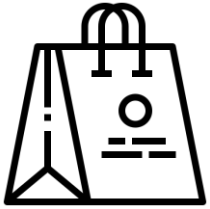
- Formazione dei gruppi
  - Compilare il google form al link <https://forms.gle/QMSBGAjEQ45JLxth9> entro 2 Maggio 2023
    - Tutti i campi sono obbligatori
    - 1 compilazione per gruppo
  - Scelta dello use case
    - Prima scelta
    - Seconda scelta
  - Informazioni dei componenti del gruppo
    - Matricola, Nome, Cognome, Genere, Corso di Laurea
  - Le assegnazioni degli use case ai gruppi sarà resa disponibile entro il 3 Maggio 2023



# Use Case: Clustering



Anagrafica



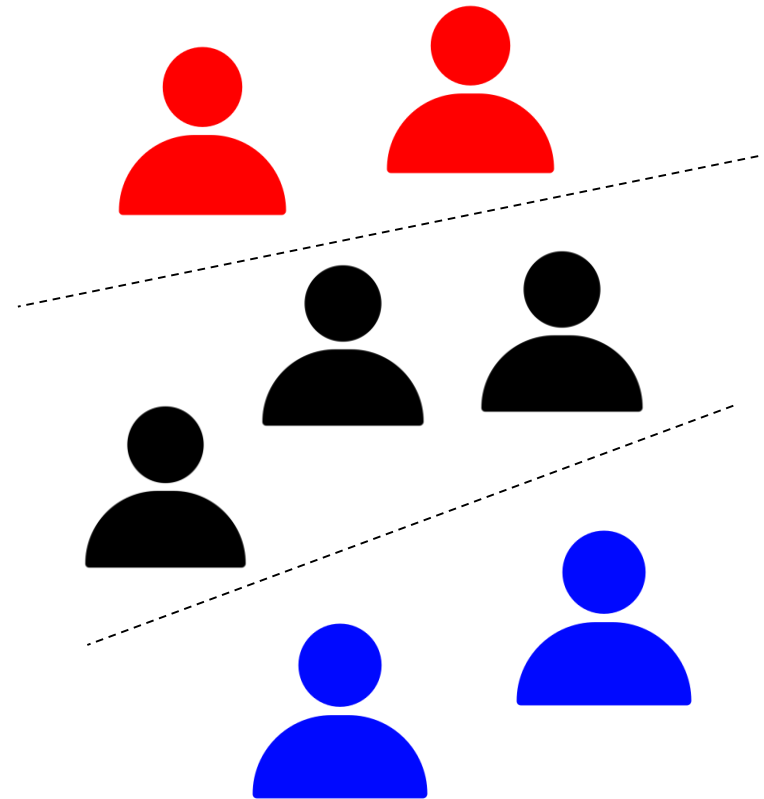
Prodotti



Sconti



Luogo di acquisto





# Use Case: Classificazione Binaria



Temperatura



Coppia



Velocità rotazionale





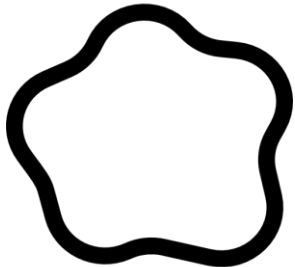
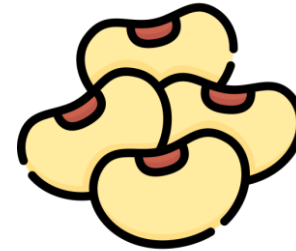
# Use Case: Classificazione Multiclasse



Area



Lunghezza Assi



Fattore di forma





# Consulenze per le attività progettuali

## ■ Consulenza

- Martedì mattina slot orario 10:00-11:30
- Dal 9/5 e fino al 6/6
- Aula da definire



# Materiale da consegnare

- Il gruppo deve consegnare
  - Report scientifico sintetico
    - pdf e progetto overleaf
  - Processo di rapid miner e file memorizzati nel repository e/o codice
  - File excel/csv con i completi dati degli esperimenti
- Il gruppo deve compilare
  - 1-2 questionari online (obbligatori)





# Date di consegna

- Entro 7 giorni prima della data della prova scritta
- Il gruppo deve consegnare un'unica cartella zip con il materiale indicato in slide 8 effettuando l'upload al link Dropbox (<https://www.dropbox.com/request/0mTzCZWD5XfuByC29ehi>).



# Valutazione

- Ogni studente del gruppo sarà valutato con un punteggio in trentesimi
  - Con valutazione pari 30 e lode, viene considerato 32 (per calcolare il voto finale)
- Il voto della tesina sarà mediato con il seguente punteggio
  - Voto conseguito all'esame scritto incrementato di
    - 1/30 in caso di consegna dell'homework su Google Data studio
    - 1/30 in caso di consegna dell'homework su MapReduce/MongoDB
  - La lode viene riconosciuta se il voto non arrotondato è  $\geq 31$