

Indicazioni per i progetti dei corsi di Statistical Data Analysis e Biomedical Statistics

- Ciascun progetto dovrà essere composto dalle 2 parti seguenti.
 - Applicazione ad un dataset reale delle tecniche di regressione studiate durante il corso.
 - Applicazione ad un dataset reale (che, a seconda delle applicazioni, può essere o meno lo stesso utilizzato al punto 1) delle tecniche di classificazione studiate durante il corso.
- In ciascuna delle due parti, gli studenti dovranno applicare le conoscenze acquisite durante il corso per l'ottimizzazione delle diverse tecniche di analisi dei dati. Ad esempio, per la preparazione dei dati (riduzione della dimensionalità, scelta delle feature, ecc.) o per la scelta dei parametri degli algoritmi (parametri di tuning di un algoritmo LASSO o di un algoritmo a discesa del gradiente, numero di "vicini" in un algoritmo k-NN, ecc.).
- In ciascuna delle due parti, gli studenti dovranno esplorare e confrontare le diverse tecniche studiate, motivando poi la scelta di una particolare tecnica sulla scorta di considerazioni teoriche e di evidenze sperimentali (ad es., confronto tra regressione lineare e k-NN, perché scelgo l'una o l'altra tecnica, ecc.).
- In particolare, è possibile (nonché consigliato) l'utilizzo di dataset sintetici a supporto dell'analisi (e **non** in sostituzione dei dataset reali).
- È suggerito l'utilizzo di grafici diagnostici per la valutazione della bontà di adattamento ai dati reali.
- I dataset reali devono corrispondere ad applicazioni verticalizzate sui rispettivi curricula/CdL. Per esempio, un dataset relativo ad un'applicazione attinente alla biostatistica è adeguato per gli studenti della laurea in Digital Health and Bioinformatic Engineering.
- Sono possibili gruppi con studenti appartenenti a diversi curricula/CdL, purché le applicazioni scelte siano calzanti per tutti i curricula/CdL coinvolti nel singolo gruppo.
- La scelta dei dataset e della particolare applicazione è lasciata agli studenti.
- La consegna del progetto (PDF e codice) deve avvenire, via e-mail agli indirizzi fpostiglione@unisa.it e vmatta@unisa.it, entro 5 giorni prima della data della particolare seduta alla quale si vuole partecipare. Qualora gli studenti lo ritenessero utile, per la presentazione orale del progetto in sede di esame è possibile predisporre una presentazione PowerPoint.