

## ANALISI DEI DATI

### MARCO ROMITO

Anno accademico	2019/20
CdS	MATEMATICA
Codice	699AA
CFU	6

Moduli	Settore	Tipo	Ore	Docente/i
ANALISI DEI DATI	MAT/06	LEZIONI	42	MARCO ROMITO

#### Obiettivi di apprendimento

##### *Conoscenze*

Al termine del corso lo studente avrà acquisito conoscenze di metodi di "statistical learning", sia nella parte di previsione che di inferenza.

##### *Modalità di verifica delle conoscenze*

Lo studente sarà valutato riguardo la sua abilità di valutare il modello statistico più opportuno per affrontare un problema di "statistical learning", e di tradurlo in una soluzione algoritmica.

##### *Capacità*

Al termine del corso lo studente

- sarà in grado di formulare il modello statistico più opportuno per l'analisi quantitativa di dati,
- saprà implementare l'analisi formulata per mezzo di un software statistico,
- sarà in grado di trarre conclusioni e formulare previsioni sul problema esaminato.

##### *Modalità di verifica delle capacità*

Il soggetto delle prove d'esame sarà l'analisi e implementazione di modelli statistici.

##### *Comportamenti*

Il corso permetterà di gestire l'analisi quantitativa di tabelle di dati mediante metodi statistici.

##### *Modalità di verifica dei comportamenti*

Durante le sessioni di esame lo studente sarà verificato sulle fasi di analisi statistica di una tabella di dati, dal riconoscimento del modello più efficace alla sua implementazione e previsione.

##### *Prerequisiti (conoscenze iniziali)*

Ci si aspetta che lo studente conosca i concetti e le idee di base della probabilità e della statistica, e di alcuni argomenti di base riguardanti l'analisi dei dati (regressione multivariata, analisi delle componenti principali, metodi autoregressivi per serie storiche). Ci si aspetta inoltre una conoscenza di base di R o python.

##### *Indicazioni metodologiche*

Il corso prevede lezioni frontali per la parte teorica. La parte implementativa è svolta parzialmente durante il corso, parzialmente come lavoro autonomo dello studente, attraverso lo svolgimento di progetti focalizzati via via su differenti argomenti scelti tra i contenuti del corso da svolgere in piccoli gruppi.

##### *Programma (contenuti dell'insegnamento)*

Introduzione al "statistical learning". Esempi di alcuni semplici esempi (regressione lineare, varianti non lineari, k-nearest-neighbour). Metodi di valutazione dei modelli (cross-validation, bootstrap, criteri di informazione). Problemi di classificazione mediante regressione logistica, analisi discriminante, support vector machines. Cenni su reti neurali. Metodi basati su alberi e foreste. Alcuni problemi di "unsupervised learning".

**Bibliografia e materiale didattico**

J. Gareth, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning

**Indicazioni per non frequentanti**

La frequenza del corso è fortemente consigliata.

**Modalità d'esame**

La prova d'esame consiste nel realizzare un progetto personale di analisi dei dati e di dare dettaglio della sua realizzazione in una relazione.

**Pagina web del corso**

<http://people.dm.unipi.it/romito/Teaching/2020/data>

*Ultimo aggiornamento 05/09/2019 13:42*