

# Introduzione al Corso

*Vers. 1.0.0*

Gianluca Mastrantonio

[gianluca.mastrantonio@polito.it](mailto:gianluca.mastrantonio@polito.it)

- comprendere la rappresentazione gerarchici dei modelli probabilistici
- comprendere il concetto di variabile latente e la loro utilità nei modelli
- essere in grado di formalizzare un modello gerarchico, sia in ottica Bayesiana che frequentista
- essere in grado di implementare un algoritmo per stimare un modello gerarchico
- essere in grado di formalizzare e applicare modelli a dati reali complessi
- essere in grado di simulare dati da modello utilizzando metodi Monte Carlo
- sviluppare la capacità di formalizzare ipotesi circa le relazioni tra variabili osservate e latenti
- sviluppare la capacità di valutare criticamente la bontà e aderenza del modello ai dati

## **Modulo 1 - Simulazione**

- Stimatori Monte Carlo
- Markov Chains (MC)
- Directed Acyclic Graphs (DAG)
- Algoritmi Markov chain Monte Carlo (MCMC)
- Expectation Maximization
- Bootstrap

## **Modulo 2 - Dipendenza Strutturata**

- Processi Gaussiani (GP)
- Geostatistica
- Accenni di Serie Storiche
- Teorema di Hammersley–Clifford e Lemma di Brooks
- Automodelli e Modelli su Grafo

## **Modulo 3 - Clustering**

- Modelli mistura base
- Modelli mistura per serie temporali (Hidden Markov Model)
- Algoritmo di Viterbi
- Change-point model

## **Modulo 1 - Simulazione**

- Introducing Monte Carlo Methods with R. By Christian Robert, George Casella
- A Student's Guide to Bayesian Statistics 1st Edition, by Ben Lambert
- The EM Algorithm and Extensions 2nd Edition. By Geoffrey J. McLachlan, Thriyambakam Krishnan

## **Modulo 2 - Dipendenza Strutturata**

- Hierarchical Modeling and Analysis for Spatial Data, 2nd Edition. By Sudipto Banerjee, Bradley P. Carlin, Alan E. Gelfand

## **Modulo 3 - Clustering**

- Hidden Markov Models for Time Series. An Introduction Using R, Second Edition. By Walter Zucchini, Iain L. MacDonald, Roland Langrock

Vedremo

- R
- STAN
- NIMBLE
- JAGS
- ...

Tesine e assignment possono essere fatti in qualsiasi linguaggio di programmazione

**Elaborato Scritto in gruppo** con Dati Simulati o Reali. Al massimo 10 persone per gruppo - orale in gruppi più piccoli. Deve contenere

- Descrizione dei Dati (o come sono stati simulati) - Un solo file per gruppo
- Stima dei modelli - Ogni studente deve scrivere il suo codice

## **Orale**

- aspetti teorici
- script mostrati durante le lezioni
- scrittura codice

**Alternative:** Consegna di un breve report alla fine di ogni modulo

- **Elaborato:** non bisogna scrivere il proprio codice dell'elaborato
- **Orale:** non vengono chiesti script e scrittura codice
- In entrambi i casi aspettatevi qualche domanda utile a capire se il codice è stato scritto da voi.

