



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Bayesian Statistical Modeling

2425-2-F8204B042-F8204B042M

Obiettivi formativi

Il corso intende esplorare i principi fondamentali e le metodologie proprie dei modelli statistici adottando un approccio bayesiano, fornendo gli strumenti per affrontare problemi classici della modellazione statistica, quali regressione, classificazione e clustering, producendo un'accurata quantificazione dell'incertezza anche in presenza di dati complessi. I metodi considerati permetteranno di affrontare un'ampia gamma di problemi applicati mediante la specificazione e la stima di modelli bayesiani progrediti. Alla fine del corso lo studente avrà acquisito: 1) la conoscenza necessaria per specificare un modello bayesiano sulla base dello specifico problema in studio; 2) la capacità di stimare tale modello mediante l'uso di software open-source, come R e Stan; 3) valutare con adeguati strumenti teorici, grafici e quantitativi, la qualità del modello ottenuto. Il corso, introducendo modelli progrediti per l'analisi statistica bayesiana, contribuisce al raggiungimento degli obiettivi formativi nell'area di apprendimento del CdS: "Statistica".

Contenuti sintetici

Revisione dei metodi bayesiani e strategie computazionali; modelli e metodi per regressione, classificazione e clustering in ambito bayesiano; validazione e selezione dei modelli; distribuzioni a priori per regolarizzazione dei modelli; modelli per dati ad elevata complessità.

Programma esteso

Revisione dei principali concetti e strumenti di inferenza bayesiana e dei principali approcci computazionali MCMC. Modelli di regressione lineare in ambito bayesiano. Dal modello di regressione lineare ai modelli di regressione lineari generalizzati con approccio bayesiano. Valutazione della bontà delle stime e selezione dei modelli in ambito bayesiano. Modelli gerarchici e modelli ad effetti casuali. Clustering in ambito bayesiano mediante l'uso di modelli

mistura. Distribuzioni a priori per problemi di regressione sparsi e dati ad elevata dimensionalità. Modelli per dati ad elevata complessità: modelli per serie storiche e modelli per dati spaziali in ambito bayesiano.

Prerequisiti

Si richiede la conoscenza dei corsi base di Statistica ad un livello di laurea triennale, e la conoscenza della statistica bayesiana di base (insegnamento di "Statistica Bayesiana M" di Clamses). Il corso non è indicato per studenti undergraduate in mobilità internazionale (programma Erasmus) i quali sono invitati a contattare il docente prima dell'inizio del corso.

Metodi didattici

Il corso di 42 sarà interamente svolto in presenza con modalità erogativa. Applicazioni, esempi ed implementazioni verranno discussi durante specifiche lezioni in aula.

Modalità di verifica dell'apprendimento

L'esame finale consiste di due parti, un progetto da svolgere eventualmente in gruppo ed un esame scritto. Il progetto verte sulla valutazione delle competenze di analisi ed utilizzo dei modelli acquisite dallo studente, dove sarà richiesto di produrre un elaborato riguardante l'analisi di un problema applicato tramite metodi bayesiani. L'esame scritto mira a verificare le capacità metodologiche acquisite dallo studente, e consiste in 3 domande aperte. Durante la prova scritta non è ammesso l'uso di testi o altro materiale. Il voto finale sarà composto da una media pesata dei voti risultanti dal progetto, 30% del voto finale, e dall'esame scritto, 70% del voto finale.

Testi di riferimento

Testi principali:

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). Bayesian Data Analysis, Third Edition. CRC Press.

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Ulteriori testi di approfondimento:

Hoff, P. D. (2009). A First Course in Bayesian Statistical Methods. Springer.

Neal, P., Dellaportas, P., Polson, N. G., & Stephens, D. A. (2013). Bayesian theory and applications. Oxford University Press.

Congdon P. (2007). Bayesian Statistical Modelling, 2nd Edition. Wiley.

Robert, C.P., Casella, G. (2004). Monte Carlo Statistical Methods, 2nd Edition. Springer.

Ulteriore materiale:

Dispense, codici R, dataset ed ulteriori esempi saranno a disposizione dello studente tramite la piattaforma eLearning del corso.

Periodo di erogazione dell'insegnamento

Primo periodo del secondo semestre.

Lingua di insegnamento

Inglese

Sustainable Development Goals

ISTRUZIONE DI QUALITÀ
