



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Statistica Multivariata

2122-1-F8204B002-F8204B003M

Obiettivi formativi

Il corso si pone come obiettivo quello di introdurre tecniche di analisi statistica multivariata con finalità esplorativa, previsiva e inferenziale.

Alla fine del corso lo studente sarà in grado di applicare le suddette tecniche a insiemi di dati raccolti in molteplici contesti, quali quello economico, sociale e delle scienze naturali, e che possono essere caratterizzati da elevata dimensionalità e dipendenza temporale. Inoltre, lo studente avrà esperito situazioni tipiche del mondo del lavoro: dal risolvere un problema in tempo limitato alla presentazione, scritta ed orale, del proprio operato.

Contenuti sintetici

Il corso è composto da un'unica parte, ma al suo interno è possibile individuare tre parti in base alle finalità degli strumenti statistici trattati.

La prima parte del corso è dedicata allo screening iniziale dei dati, compresi alcuni richiami di visualizzazione dei dati. Nella seconda parte si introducono tecniche statistiche supervisionate e si contestualizza il loro impiego a fini previsivi fornendo, inoltre, i concetti e le procedure necessarie per valutare l'errore di generalizzazione di tali strumenti.

Programma esteso

Cenni di visualizzazione di dati multivariati

Metodi operativi per la valutazione dell'errore di generalizzazione: insieme di stima e verifica, K-fold cross validation

Classificatore di Bayes, analisi discriminante lineare e quadratica

Alberi di classificazione e regressione

Metodi di ensemble

Prerequisiti

Analisi esplorativa, modelli statistici, calcolo delle probabilità, inferenza statistica, programmazione.

Metodi didattici

L'intera attività formativa verrà svolta attraverso lezioni frontali e di laboratorio.

Modalità di verifica dell'apprendimento

L'esame si compone di una prova scritta

La prova scritta si compone di esercizi numerici, di domande aperte e a risposta chiusa. Gli esercizi numerici hanno lo scopo di verificare la capacità dello studente di applicare le tecniche statistiche introdotte durante il corso. Le domande aperte e a risposta chiusa hanno invece l'obiettivo di misurare il livello di conoscenza della teoria alla base delle suddette tecniche.

Testi di riferimento

James, G, Witten, D., Hastie, T., Tibshirani, R., An Introduction to Statistical Learning with applications in R, Springer, 2013

Everitt, B., Hothorn, T. An Introduction to Applied Multivariate Analysis with R, Springer, 2011

Hastie, T., Tibshirani, R., Friedman, J. ,The Elements of Statistical Learning, Springer, 2001

Izenman, A.J., Modern Multivariate Statistical Techniques, Springer, 2008

Azzalini, A. Scarpa, B. Analisi dei dati e data mining, Springer, 2009

Dobson, A., Barnett, A., An Introduction to Generalized Linear Models, CRC Press, 2008

Periodo di erogazione dell'insegnamento

Secondo semestre. Primo ciclo.

Lingua di insegnamento

Italiano
