COURSE DETAIL

BIG DATA: NEW TOOLS FOR ECONOMETRICS
Country Italy
Host Institution University of Bologna
Program(s) University of Bologna
UCEAP Course Level Upper Division
UCEAP Subject Area(s) Economics
UCEAP Course Number 172
UCEAP Course Suffix
UCEAP Official Title BIG DATA: NEW TOOLS FOR ECONOMETRICS
UCEAP Transcript Title BIG DATA:ECONMETRCS
UCEAP Quarter Units 6.00
UCEAP Semester Units 4.00

Course Description

This is a graduate level course that is part of the Laurea Magistrale program. The course is intended for advanced level students only. Enrollment is by consent of the instructor. The course focuses on the main tools used by economists and statisticians in machine learning and statistical learning to analyze large/huge data sets coming from several domains. The course highlights how to apply key aspects of machine and statistical learning, such as out-of-sample cross-validation, regularization, and scalability. Special attention is placed on the concepts of supervised and unsupervised learning, classification, regression, and clustering analysis as well as the detection of association rules. The course also focuses on the main learning tools such as lasso and ridge regression, regression trees, boosting, bagging and random forests, principal components, mixture models and the k-means algorithm. The course places emphasis on the application of the techniques discussed using dedicated open-source software packages on training datasets. Course topics: introduction and overview of statistical learning; linear regression as a prediction tool; binary and multinomial classification: logistic regression, linear discriminant analysis and k-nearest neighbors; resampling methods: cross-validation and the bootstrap; linear model selection and regularization: ridge regression, the lasso, and principal components; moving beyond linearity: regression splines, smoothing splines and general additive models; tree-based methods: CART, bagging, boosting, and random forests; support vector machines and neural networks; unsupervised learning: hierarchical and k-means clustering. The relevant theory will be applied to each topic and subsequently the analysis will move to its empirical application in the R language. Special emphasis is placed on the economic interpretation of the results. The course focuses in several empirical analyses and replicates the results of a few case studies using the statistical software R and several of its packages.

Language(s) of Instruction
English

Host Institution Course Number
81618

Host Institution Course Title BIG DATA: NEW TOOLS FOR ECONOMETRICS (LM)
Host Institution Course Details
Host Institution Campus ECONOMIA E MANAGEMENT
Host Institution Faculty
Host Institution Degree
Host Institution Department Economics
Course Last Reviewed

Print