UCLouvain

ldats2470

2026

Statistical Machine Learning and High Dimensional Data Analysis

3.00 credits	15.0 h	Q2

Teacher(s)	Hafner Christian ;
Language :	English
Place of the course	Louvain-la-Neuve
Prerequisites	Concepts and tools equivalent to those taught in teaching units LSTAT2020 Logiciels et programmation statistique de base LSTAT2120 Linear models LSTAT2110 Analyse des données
Main themes	1. Partitioning methods for clustering 2. Statistical approaches for dimension reduction and feature extraction 3. Regularization methods in high dimensions, including linear and nonlinear shrinkage 4. Applications
Learning outcomes	
Evaluation methods	Project using a real data set (40%) , and an oral exam (60%)
Teaching methods	The lectures provide the theoretical material, give many practical examples, and show how to implement the methods in common programming packages.
Content	1. Partitioning methods for clustering · k-means and variants · Nonlinear k-means with kernels · Support Vector Machines and other multiple kernel learning machines · Spectral clustering 2. Statistical approaches for dimension reduction and feature extraction · Factor models and probabilistic PCA · Kernels for non-linear PCA · Kernels for non-linear ICA 3. Regularization methods in high dimensions, including linear and nonlinear shrinkage 4. Applications
Inline resources	Slides, R codes and data
Bibliography	- Everitt, B. and Hothorn, T. (2011). An Introduction to Applied Multivariate Analysis with R, Springer Verlag. - Härdle, W. and Simar, L. (2015). Applied Multivariate Statistical Analysis, Springer Verlag. - Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning, Springer Verlag. - Izenman, A.J. (2008) Modern multivariate statistical techniques, Springer - James, Witten, Hastie, Tibshirani (2013) An Introduction to statistical learning with applications in R, Springer - Koch, I. (2014) Analysis of multivariate and high-dimensional data, Cambridge - Ledolter, J. (2013), Data Mining and Business Analytics with R, Wiley - Zaki, M.J. and Meira, W. (2020) Data Mining and Machine Learning, fundamental concepts and algorithms, 2nd ed., Cambridge.
Faculty or entity in charge	LSBA

Programmes containing this learning unit (UE)

Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	3		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	3		