

Enseignants	Hafner Christian ;
Langue d'enseignement	Anglais
Lieu du cours	Louvain-la-Neuve
Préalables	Concepts et outils équivalents à ceux enseignés dans les UEs LSTAT2020 Logiciels et programmation statistique de base LSTAT2120 Linear models LSTAT2110 Analyse des données
Thèmes abordés	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
Acquis d'apprentissage	
Modes d'évaluation des acquis des étudiants	Project using a real data set (40%), and an oral exam (60%)
Méthodes d'enseignement	The lectures provide the theoretical material, give many practical examples, and show how to implement the methods in common programming packages.
Contenu	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
Ressources en ligne	Transparents, codes R, données
Bibliographie	- 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.

Faculté ou entité en charge:	LSBA
------------------------------	------

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	3		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	3		