LLSMF2013
2013-2014

Quantitative data analysis(in English)

5.0 credits	30.0 h	2q

Teacher(s) :	Saerens Marco ;
Language :	Anglais
Place of the course	Louvain-la-Neuve
Main themes :	Presentation of quantitative data analysis methods, in particular scoring methodology and classification; Presentation of some decision making models; Reading texts containing data analysis methods; Exercises in appropriation by a group work, in analysing methods of qualitative and quantitative materials collected personally or placed at the disposal; Initiation to professional data analysis software such as Atlas-TI, SAS/JMP and R.
Aims :	To be able to make decisions on the basis of quantitative information, and to assess accurately the performances of the mobilized models. *The contribution of this Teaching Unit to the development and command of the skills and learning outcomes of the programme(s) can be accessed at the end of this sheet, in the section entitled "Programmes/courses offering this Teaching Unit".*
Content :	Content The study of data analysis and decision-making methods, with a focus on the interpretation of the results; in particular, classification, scoring methodology: clustering, factorial and projection methods, decision trees, logistic regression, A discussion on which method to use in function of the problem at hand and the available data. Methods A combination of lectures, practical exercises and a project dealing with real data. Content A review of the main subspace projection and feature extraction of data analysis/modeling, and their interpretation: - Categorical data: subspace projection and latent variable techniques techniques, log-linear models, etc. - Numerical data: subspace projection and latent variable techniques, clustering techniques, discriminant analysis, etc. Supervised classification: naïve Bayes, artificial neural networks, decision trees, combining classifiers, etc. Unsupervised classification (clustering) methods. Decision-making from data: a short introduction to Bayes decision theory, Bayesian networks, Markov decision processes, reinforcement learning, multicriteria decision analysis. Application to "information retrieval" and to "web mining" (PageRank, Hits, collaborative recommendation, etc). A discussion of which method to use in function of the data and the problem at hand. Projects (for instance scoring) based on real data, with SAS/JMP, S-Plus or R. Methods In-class activities 0 Lectures 0 Project based learning At home activities 0 Readings to prepare the lecture 0 Paper work

Other infos :	Prerequisites (ideally in terms of competencies): A course in multivariate statistical analysis, on probability theory, on mathematical statistics, on matrix algebra and on multivariate analysis. Evaluation : Writing of two papers. Support : Book chapters provided to the students References : Provided during the class - Duda, Hart & Stork (2001), "Pattern classification, 2nd ed". John Wiley & Sons. - Bardos (2001), "Analyse discriminante. Application au risque et scoring financier. Dunod. - Lebart, Morineau & Piron (1995), "Statistique exploratoire multidimensionnelle". Dunod. - Webb (2002), "Statistical pattern recognition, 2nd ed". John Wiley and Sons. - Theodoridis & Koutroumbas (2003), "Pattern recognition". Academic Press. - Alpaydin (2004), "Introduction to machine learning". MIT Press. - Han & Kamber (2000), "Data mining: Concepts and techniques". Morgan Kaufmann. - etc.
Cycle and year of study :	> Master [120] in Business engineering > Master [120] in Population and Development Studies > Bachelor in Computer Science
Faculty or entity in charge:	CLSM