UCL
Université catholique de Louvain

LLSMF2013
2016-2017

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 :	Having regard to the LO of the programme X, this activity contributes to the development and acquisition of the following LO: -- 2. Knowledge and reasoning 2.1. Master the core knowledge of each area of management. 2.2. Master highly specific knowledge ' 2.4. Activate and apply the acquired knowledge ' 3. A scientific and systematif approach 3.1. Conduct a clear, structured, analytical reasoning ' 3.2. Collect, select and analyze relevant information ' 3.3.Consider problems using a systemic and holistic approach ' 3.4. Perceptively synthesize 'demonstrating a certain conceptual distance ' 3.5.Produce, through analysis and diagnosis, implementable solutions' 6. Teamwork and leadership 6.1. Work in a team... 7. Project management 7.1.Analyse a project within its environment and define the expected outcomes' 8. Communication and interpersonal skills 8.1. Express a clear and structured message' 8.2. Interact and discuss effectively ' 9. Personal and professional development 9.1. Independent self-starter ' 9.4. Quick study, lifelong learner ' *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

Bibliography :	: No TEXTBOOK. and available on line . BOOK : Alpaydin (2004), 'Introduction to machine learning'. MIT Press. Bardos (2001), Analyse discriminante. Application au risque et scoring financier. Dunod. Bishop (1995), Neural networks for pattern recognition . Clarendon Press. Bishop (2006), 'Pattern recognition and machine learning'. Springer-Verlag. Bouroche & Saporta (1983), L analyse des données . Que Sais-je. Cornuéjols & Miclet (2002), Apprentissage artificiel. Concepts et algorithmes . Eyrolles. Duda, Hart & Stork (2001), Pattern classification, 2nd ed . John Wiley & Sons. Dunham (2003), Data mining. Introductory and advanced topics . Prentice-Hall. Greenacre (1984), Theory and applications of correspondence analysis . Academic Press. Han & Kamber (2006), Data mining: Concepts and techniques, 2nd ed . Morgan Kaufmann. Hand (1981), Discrimination and classification . John Wiley & Sons. Hardle & Simar (2003), Applied multivariate statistical analysis . Springer-Verlag. Disponible à http://www.quantlet.com/mdstat/scripts/mva/htmlbook/mvahtml.html Hastie, Tibshirani & Friedman (2001), The elements of statistical learning . Springer-Verlag. Johnson & Wichern (2002), Applied multivariate statistical analysis, 5th ed . Prentice-Hall. Lebart, Morineau & Piron (2006), Statistique exploratoire multidimensionnelle, 4e ed . Dunod. Mitchell (1997), Machine learning . McGraw-Hill. Naim, Wuillemin, Leray, Pourret & Becker (2004), 'Réseaux bayesiens'. Editions Eyrolles. Nilsson (1998), 'Artificial intelligence: A new synthesis'. Morgan Kaufmann. Ripley (1996), Pattern recognition and neural networks . Cambridge University Press. Rosner (1995), Fundamentals of biostatistics, 4th ed .Wadsworth Publishing Company. Saporta (2006), Probabilités, analyse des données et statistique, 2nd ed . Editions Technip. Tan, Steinbach & Kumar (2005), 'Introduction to data mining'. Addison Wesley. Theodoridis & Koutroumbas (2006), Pattern recognition, 3nd ed . Academic Press. Therrien (1989), Decision, estimation and classification . Wiley & Sons. Venables & Ripley (2002), Modern applied statistics with S. Springer-Verlag. Vincke (1989), L aide multicritere a la decision . Editions Ellipses. Wasserman (2004), 'All of statistics'. Springer. Webb (2002), Statistical pattern recognition, 2nd ed . John Wiley and Sons. not compulsory and available on line Supports available on line are on ICAMPUS.
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. 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.
Faculty or entity in charge:	CLSM

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Credits	Prerequis	Acquis d'apprentissage
Additionnal module in Computer Sciences	LSINF110P	5	-	🔍
Master [120] in Population and Development Studies	SPED2M	5	-	🔍
Master [120] in Business Engineering	INGE2M	5	-	🔍
Master [120] in Management	GEST2M	5	-	🔍
Master [120] in Management	GESM2M	5	-	🔍