

5.0 credits

30.0 h + 30.0 h

2q

Teacher(s) :	Saerens Marco ;
Language :	Anglais
Place of the course	Louvain-la-Neuve
Inline resources:	 > http://icampus.uclouvain.be/claroline/course/index.php?cid=sinf2275
Prerequisites :	-- A first course on probability theory -- A first course on mathematical statistics -- An undergraduate course on matrix algebra -- An undergraduate course on multivariate analysis
Main themes :	The course is structured around four themes -- Complements of data mining, -- Decision making, -- Information retrieval, -- Link analysis and web/graph mining .
Aims :	-- to understand quantitative and qualitative data mining methods and to apply them to decision making -- to develop a critical view of data mining techniques in specific application domains -- to master information retrieval techniques from very large data collection, possibly enriched with link structures (WEB, social networks, ...) -- to apply information retrieval techniques in the context of search engines and automated recommendation systems -- to implement data mining and information retrieval algorithms within standard software environments such as S-Plus, R, SAS, Weka or Matlab <i>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".</i>
Evaluation methods :	-- Two projects / case studies each counting for 3 points on 20. -- An oral exam in session will count for 14 points on 20
Teaching methods :	-- 30 hours of magistral courses. -- Two projects / case studies on solving two problems.
Content :	Complements of data mining -- Canonical correlation analysis -- Correspondence analysis -- Partial least squares regression -- Log-linear models -- Association rules Decision making -- Markov decision processes and reinforcement learning -- Exploration/exploitation and bandit problems --

Utility theory
--
Multi-criteria preference modeling - the Promethee method
--
Probabilistic reasoning with bayesian networks
--
Possibility theory
--
Two-players game theory
--
Collective decisions
Information retrieval
--
The basic vector-space model
--
The probabilistic model
--
Ranking web pages :PageRank, HITS, etc.
--
Collaborative recommendation models (recommender systems) .
Link analysis and web/graph mining
--
Network community detection
--
Similarity measures between nodes
--
Spectral graph partitioning and mapping
--
Reputation models

<p>Bibliography :</p>	<p>-- 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. -- Bourouche & mp; Saporta (1983), "L'analyse des données". Que Sais-je. -- Cornuéjols & mp; Miclet (2002), "Apprentissage artificiel. Concepts et algorithmes". Eyrolles. -- Duda, Hart & mp; Stork (2001), "Pattern classification, 2nd ed". John Wiley & mp; Sons. -- Dunham (2003), "Data mining. Introductory and advanced topics". Prentice-Hall. -- Greenacre (1984), "Theory and applications of correspondence analysis". Academic Press. -- Han & mp; Kamber (2005), "Data mining: Concepts and techniques, 2nd ed.". Morgan Kaufmann. -- Hand (1981), "Discrimination and classification". John Wiley & mp; Sons. -- Hardle & mp; Simar (2003), "Applied multivariate statistical analysis". Springer-Verlag. Disponible à http://www.quantlet.com/mdstat/scripts/mva/htmlbook/mvahtml.html -- Hastie, Tibshirani & mp; Friedman (2001), "The elements of statistical learning". Springer-Verlag. -- Johnson & mp; Wichern (2002), "Applied multivariate statistical analysis, 5th ed". Prentice-Hall. -- Lebart, Morineau & mp; Piron (1995), "Statistique exploratoire multidimensionnelle". Dunod. -- Mitchell (1997), "Machine learning". McGraw-Hill. -- Naim, Wuillemin, Leray, Pourret & mp; Becker (2004), "Réseaux bayésiens". 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 (1990), "Probabilités, analyse des données et statistique". Editions Technip. -- Tan, Steinbach & mp; Kumer (2005), "Introduction to data mining". Pearson. -- Theodoridis & mp; Koutroumbas (2003), "Pattern recognition, 3th ed". Academic Press. -- Therrien (1989), "Decision, estimation and classification". Wiley & mp; Sons. -- Venables & mp; Ripley (2002), "Modern applied statistics with S. Springer-Verlag. -- Webb (2002), "Statistical pattern recognition, 2nd ed". John Wiley and Sons.</p>
<p>Cycle and year of study :</p>	<p>> Certificat universitaire en statistique > Master [120] in Business engineering > Master [120] in Business Engineering > Master [120] in Environmental Bioengineering > Master [120] in Computer Science and Engineering > Master [120] in Computer Science > Master [120] in Chemistry and Bio-industries > Master [120] in Statistics: General > Master [120] in Agricultural Bioengineering > Master [120] in Forests and Natural Areas Engineering</p>
<p>Faculty or entity in charge:</p>	<p>INFO</p>