

5.0 crédits	30.0 h + 30.0 h	2q
-------------	-----------------	----

Enseignants:	Saerens Marco ;
Langue d'enseignement:	Anglais
Lieu du cours	Louvain-la-Neuve
Préalables :	<p>Ce cours présuppose que l'étudiant a suivi</p> <ul style="list-style-type: none"> -- un cours de calcul des probabilités, -- un cours de statistique mathématique, -- un cours d'analyse statistique multivariée et -- un cours de calcul matriciel.
Thèmes abordés :	<p>Le cours est articulé autour de quatre thèmes,</p> <ul style="list-style-type: none"> -- Compléments de fouille de données -- Prise de décision, -- Recherche d'information, -- Analyse de liens et l'exploration du web / graphique.
Acquis d'apprentissage	<ul style="list-style-type: none"> -- Comprendre et appliquer des méthodes de fouille de données (data mining), tant qualitatives que quantitatives, dans le cadre de la prise de décision. -- Avoir un jugement critique sur les méthodes de fouilles de données en fonction de leur champ d'application. -- Maîtriser des méthodes d'extraction et recherche d'informations (information retrieval) dans de très larges collections de données, éventuellement enrichies d'une structure de liens (WEB, réseaux sociaux...). -- Comprendre l'application de ces méthodes dans les moteurs de recherche et les systèmes de recommandation automatisée. -- Mettre en oeuvre les algorithmes de fouille de données et d'extraction d'informations dans des logiciels de data mining ou de traitement statistique tels que S-Plus, R, SAS, Weka ou Matlab. <p><i>La contribution de cette UE au développement et à la maîtrise des compétences et acquis du (des) programme(s) est accessible à la fin de cette fiche, dans la partie « Programmes/formations proposant cette unité d'enseignement (UE) ».</i></p>
Contenu :	<p>Compléments de fouille de données</p> <ul style="list-style-type: none"> -- Analyse des corrélations canoniques -- Analyse des correspondances -- Régression partielle des moindres carrés -- Modèles log-linéaires -- Règles de l'Association <p>Prise de décision</p> <ul style="list-style-type: none"> -- Processus de décision de Markov et l'apprentissage par renforcement -- Exploration / exploitation et les problèmes hostiles -- Théorie de l'utilité -- Modélisation des préférences multi-critères - la méthode Prométhée -- Raisonnement probabiliste avec des réseaux bayésiens -- Théorie des possibilités -- Théorie des jeux à deux joueurs -- Décisions collectives <p>Recherche d'information</p> <ul style="list-style-type: none"> -- Modèle de base d'espace vectoriel -- Modèle probabiliste -- Pages Web Ranking: PageRank, HITS, etc. -- Modèles collaboratifs par recommandations (systèmes de recommandation). <p>Analyse de liens et l'exploration du web / graphique</p> <ul style="list-style-type: none"> -- Détection de la communauté réseau -- Mesures de similarité entre les n'uds -- Partitionnement de graphe spectral et cartographie -- Modèles de réputation

<p>Bibliographie :</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. -- Bouroche & 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, Wuillemain, Leray, Pourret & mp; 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 (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 et année d'étude: :</p>	<p>> Master [120] : ingénieur civil en informatique > Master [120] en sciences informatiques > Master [120] en ingénieur de gestion > Master [120] en ingénieur de gestion > Certificat universitaire en statistique > Master [120] en statistiques, orientation générale</p>
<p>Faculté ou entité en charge:</p>	<p>INFO</p>