

5.00 crédits

30.0 h + 15.0 h

Q2

Enseignants	Cappart Quentin ;Piette Eric ;
Langue d'enseignement	Anglais > Facilités pour suivre le cours en français
Lieu du cours	Louvain-la-Neuve
Préalables	Intelligence artificielle, tel que couvert par LINFO1361
Thèmes abordés	• Fondements de l'apprentissage par renforcement (RL) • Bandits multi-bras et exploration/exploitation • Processus de décision de Markov (MDP) • Résolution par programmation dynamique • Méthodes de Monte Carlo • Méthodes de <i>Temporal Difference Learning (Q-learning)</i> • Apprentissage par renforcement profond • Approximations de valeur (DQN et variantes) • Méthodes de <i>policy gradient</i> (REINFORCE, AC, PPO) • Arbres de recherche de Monte-Carlo • <i>Large Reasoning Models</i> et <i>RL from Human Feedback</i> • Applications aux jeux et environnements simulés • Enjeux contemporains, limites et perspectives du RL
Acquis d'apprentissage	A la fin de cette unité d'enseignement, l'étudiant est capable de : Eu égard au référentiel AA du programme « Master en sciences informatiques », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants : • SINF1.M4 • SINF2.1-4 • SINF5.3-4 • SINF6.1, SINF6.4, SINF6.5 Les étudiants ayant suivi ce cours seront capables de : • Modéliser un problème en termes de processus de décision de Markov • Implémenter des algorithmes de RL classiques (Q-Learning, Monte Carlo, etc.) • Comprendre les enjeux de l'exploration et de l'approximation des fonctions de valeur • Implémenter des algorithmes de RL contemporains (DQN, REINFORCE, PPO, etc.) • Décrire le fonctionnement des <i>Large Reasoning Models</i> et leur usage dans le <i>RL from Human Feedback</i> • Appliquer le RL sur des environnements simulés (jeux, tâches de contrôle) • Lire, comprendre, et analyser des articles scientifiques dans le domaine du RL • Analyser les performances et les limites des approches implémentées
Contenu	• Introduction générale au RL (agent, environnement, états, actions, récompenses, politique, fonctions de valeurs, convergence). • Bandits multi-bras (Exploration/Exploitation, # -greedy , upper confidence bound, softmax, Thompson sampling, Regrets) • Processus de décision de Markov : formalisme et dynamique (Propriété de Markov, Politiques stochastiques vs déterministes, Fonctions action-valeur, équation de Bellman, Optimalité) • Résolution par programmation dynamique (policy evaluation , policy iteration , value iteration) • Méthodes de Monte Carlo (estimation de valeurs d'états et valeur d'actions, convergence) • <i>Temporal Difference Learning</i> (Bootstrap, TD(0), Variance, Apprentissage en ligne) • Algorithmes <i>Q-Learning</i> • Approximation de fonction et Deep Q-Networks (Gradient, Approximation non linéaire, DQN) • Arbres de recherche de Monte-Carlo et variantes profondes

	• Exploration avancée (REINFORCE, <i>Actor-Critic</i> , <i>Proximal Policy Optimization</i>) • Introduction aux <i>Large Reasoning Models</i> (LRMs) et <i>RL from Human Feedback</i> (RLHF) - <i>Language Modeling</i>, <i>Direction Preference Optimization</i> (DPO), <i>supervised Fine-Tuning</i> • Applications aux jeux et environnements simulés avec la librairie open-source Gymnasium • Études de cas (Atari, CartPole, LunarLander) et/ou projet pratique d'implémentation et analyse comparative des méthodes
Ressources en ligne	Disponibles sur la page Moodle du cours
Bibliographie	Some recommended reference books : • 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 (2005), "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 (1995), "Statistique exploratoire multidimensionnelle". Dunod. • Mitchell (1997), "Machine learning". McGraw-Hill. • Naim, Willemin, Leray, Pourret & 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 & Kumer (2005), "Introduction to data mining". Pearson. • Theodoridis & Koutroumbas (2003), "Pattern recognition, 3th ed". Academic Press. • Therrien (1989), "Decision, estimation and classification". Wiley & Sons. • Venables & Ripley (2002), "Modern applied statistics with S. Springer-Verlag. • Webb (2002), "Statistical pattern recognition, 2nd ed". John Wiley and Sons.
Autres infos	Préalables : • LBIR1304 ou LFSAB1105 : un cours de calcul des probabilités et de statistique mathématique, • LBIR1200 ou LFSAB1101 : un cours de calcul matriciel et d'algèbre linéaire, • LFSAB1402 : un bon cours de programmation Python, • Un cours d'analyse multivariée (mathématiques).
Faculté ou entité en charge:	INFO

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	5		
Master [120] : ingénieur civil en chimie et science des matériaux	KIMA2M	5		
Master [120] : ingénieur civil des constructions	GCE2M	5		
Master [120] : ingénieur civil biomédical	GBIO2M	5		
Master [120] : bioingénieur en gestion des forêts et des espaces naturels	BIRF2M	5		
Master [120] : bioingénieur en sciences et technologies de l'environnement	BIRE2M	5		
Master [120] : ingénieur civil mécanicien	MECA2M	5		
Master [120] : ingénieur civil électricien	ELEC2M	5		
Master [120] : ingénieur civil physicien	FYAP2M	5		
Master [120] : bioingénieur en chimie et bioindustries	BIRC2M	5		
Master [120] : ingénieur civil en informatique	INFO2M	5		
Master [120] en sciences informatiques	SINF2M	5		
Master [120] : ingénieur civil électromécanicien	ELME2M	5		
Master [120] : ingénieur civil en mathématiques appliquées	MAP2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	5		

Master [120] : bioingénieur en sciences agronomiques	BIRA2M	5		
Master [120] en science des données, orientation technologies de l'information	DAT12M	5		
Master [120] : ingénieur civil en génie de l'énergie	NRGY2M	5		