SORBONNE UNIVERSITÉ

Département des Masters de Sciences de l'Ingénieur
Mention Automatique, Robotique

MASTERS Sdi Sciences de l'ingénieur

Intitulé de l'Unité d'Enseignement (UE) : Deep learning avancé et apprentissage multimodal

Responsable de l'UE : Nicolas Obin (nicolas.obin@sorbonne-universite.fr)
Code UE : UM5RBI30
Volumes horaires / étu : 16 h Cours, 24 h TP
Nombre de crédits de l'UE/EC : 6 ECTS
Parcours-type : ☐ ROBOT ☒ SI ☐ MeDH
Semestre où l'enseignement est proposé : ☐ S1 ☐ S2 ☒ S3 ☐ S4
Langue d'enseignement : ☒ Français ☐ Anglais
Public concerné : ☒ Département SDI ☒ Autre (préciser) : Poltech Sorbonne
Localisation : ☒ Campus PMC ☐ Autre (préciser) :

Thématique

Cette UE propose un approfondissement en apprentissage profond (réseaux de neurone et apprentissage par renforcement), avec un focus sur l'apprentissage à partir de données issues de plusieurs modalités (par exemple : texte, son, image). Les enseignements présentent en particulier un approfondissement sur les architectures neuronales actuelles (transformeurs, grands modèles de langage, modèles de fondation, modèles génératifs) et des ouvertures sur des thématiques émergentes ou complémentaires pour les systèmes intelligents. Par ailleurs et aussi : en amont sur l'empreinte écologique utilisées ou en aval pour l'interprétation et l'explicabilité des sorties d'un réseau et sur des applications concrètes d'IA embarquée ou pour des systèmes robotiques.
Mots-clés : Apprentissage profond. Modélisation générative. Apprentissage multimodal. Interprétabilité et explicabilité. Optiimisation matérielle. Coût énergétique et écologique.

Prérequis

Afin de suivre cette UE, doivent être acquis précédemment :

- Mathématiques : Algèbre linéaire, calcul vectoriel et matriciel, optimisation numérique de fonctions continues à plusieurs variables, calcul différentiel, probabilités
- Informatique : Installer et utiliser un environnement de programmation (langage, ide), utiliser un environnement virtuel et un logiciel de gestion de version de code, utiliser les environnements de programmation dédié au deep learning (TensorFlow, PyTorch)
- Apprentissage machine : Maitriser les concepts et méthodes fondementaux de l'apprentissage machine et de l'apprentissage profond. Maitriser la préparation des données pour l'apprentissage. Maitriser les méthodologies d'évaluation de modèles.

Acquis d'apprentissage visés (AAV)

À l'issue de cette UE, les étudiantes ou étudiants seront capables de :

1. C1.2. Analyser un problème : Choisir un algorithme adapté à la tâche, aux données, et aux ressources
2. C1.4 Analyser un problème : Intégrer les contraintes de matériel, de consommation d'énergie et d'empreinte écologique, lors de la conception d'une solution à un problème d'apprentissage
3. C2.1 Conception : Fondements algorithmiques
4. C2.3 Conception : Spécifier les méta-paramètres pour optimiser l'apprentissage des paramètres du modèle
5. C2.4 Conception : Choisir le processus d'apprentissage
6. C3.3 Mise en oeuvre : Préparer les données pour l'apprentissage
7. C3.4 Mise en oeuvre : Réaliser l'apprentissage défini dans la phase de conception pour optimiser les paramètres et les méta-paramètres de l'algorithme
8. C3.5 Mise en oeuvre : Reproduire des apprentissages et des résultats expérimentaux

9. C4.3 Evaluer : Diagnostiquer les causes d'un comportement sous-optimal du réseau afin de proposer des solutions y remédiant
10. C4.4 Evaluer : Interpréter et expliquer les décisions d'un modèle : carte d'attention, vérification des biais

Séquencement et modalités de l'enseignement (indicatif)

Semaine	C/TD/TP	Contenu	Préparation	AAV associé(s)
S1	C1 (2h)	C1 - Transformeurs		AAV 1, 3, 4, 5
S2	C2 (2h) - TP1 (4h)	C2 - Grand modèles de langage. Apprentissage auto-supervisé. Masquage. Multimodalité TP1 - Transformeurs		AAV 1, 3, 4, 5, 6, 7
S3	C3 (2h) - TP2 (4h) P1 (4h)	C3 - Modèles de fondation. Texte, son, image. Apprentissage de représentations multimodales : CLIP, CLAP TP2 - Grands modèles de langage P1 - Séance de projet 1		AAV 1, 3, 4, 5, 6, 7
S4	C4 (2h) - TP3 (4h)	C4 - Models génératifs : GAN, diffusion, modèles de langage TP3 - Modèles de fondation		AAV 1, 3, 4, 5, 6, 7
S5	TP4 (4h)	TP4 - Modèles génératifs		AAV 1, 3, 4, 5, 6, 7, 10
S6	C6 / TP6 (2h+2h) Quiz 1 P2 (4h)	C6 / TP6- Empreinte environnementale de l'apprentissage profond P2 - Séance de projet 2		AAV 2
S7	C7 / TP7 (2h+2h) Quiz 2	C7 / TP7 - Compression et accélération		AAV 2
S8	C8 / TP8 (2h+2h) Quiz 3	C8 /TP 8 - Interprétabilité et explicabilité		AAV 2
S9	C9 / TP9 (2h+2h) Quiz 4 P3 (4h)	C9 / TP9 - Modèles de fondation pour la robotique P3 - Séance de projet 3		AAV 1, 3, 4, 5
S10				
S11	P4 (4h)	P3 - Séance de projet 4		
S14	Soutenance projet			
S15	ER (2h)			

Déroulement de l'UE

L'UE se déroule en 3 séquences :

- Séquence 1 : 4 cours d'approfondissement sur les fondamentaux des architectures actuelles d'apprentissage profond (C1-5, TP1-5, examen de TP)
- Séquence 2 : 4 cours d'ouverture sur des thématiques émergentes ou complémentaires associés aux systèmes intelligents (C6-9, ER)
- Séquence 3 : un projet dans lequel les étudiants doivent reproduire les résultats d'un article de recherche (P1-4, projet)

Les deux premières de succèdent. La séquence de projet commence dès octobre.

Modalités d'évaluation des acquis d'apprentissage (1ère session)

Semaine	Indiv/collectif	Présentiel/distanciel	Type d'examen	AAV évalué(s)	Barème %
S7-10	Individuel	Présentiel	Quiz	AAV 6-10	20 %
S14	Collectif	Présentiel	Autre	AAV 1-10	40 %
S15	Individuel	Présentiel	Écrit	AAV 1-5	40 %

2nde session

Session	Indiv/collectif	Présentiel/distanciel	Type d'examen	AAV évalué(s)	Barème %

Références bibliographiques

— I. Goodfellow, Y. Bengio, A. Courvielle, Deep Learning, MIT Press, 2016
— Richard S. Sutton et Andrew G. Barto, Reinforcement Learning : An Introduction, second edition , 2018
— François Chollet, Deep Learning with Python. Mannan, 2017-2021
— Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2006

Date de génération de la fiche d'UE : 6 octobre 2026