

Partager



Email destinataire :

Vous devez renseigner un email valide

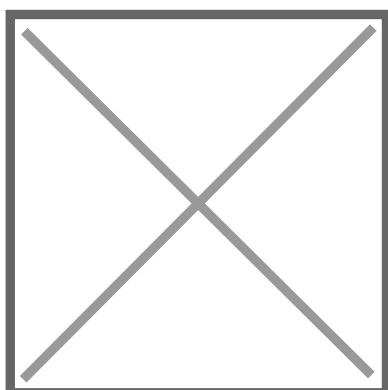
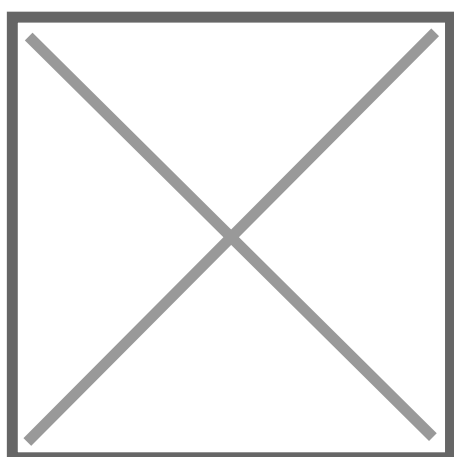
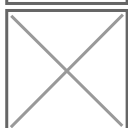
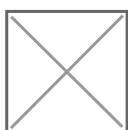
Envoyer

Une erreur est survenue. Merci de réessayer ultérieurement

Le mail de partage a bien été envoyé.

Ajouter au comparateur

M2 Mathématiques et Intelligence Artificielle



Master

Mention Mathématiques et applications

Régime(s) d'inscription

Formation initiale Formation continue

Langue(s) d'enseignement

Anglais Français

Le master Mathématiques et Intelligence Artificielle vise à fournir une compréhension approfondie des fondements théoriques et pratiques des algorithmes d'apprentissage statistique, et à apporter les compétences nécessaires à leur analyse critique et à leur mise en oeuvre.

Le M2 Mathématiques et Intelligence artificielle est porté conjointement par le département de Mathématiques d'Orsay et Centrale-Supelec, avec le soutien du département d'informatique et du programme SaclAI-school.

Il est structuré en trois périodes de cours (septembre à début novembre, début novembre à mi-janvier, mi-janvier à fin mars). La fin de l'année (période allant de début avril à fin septembre) est dédiée au stage.

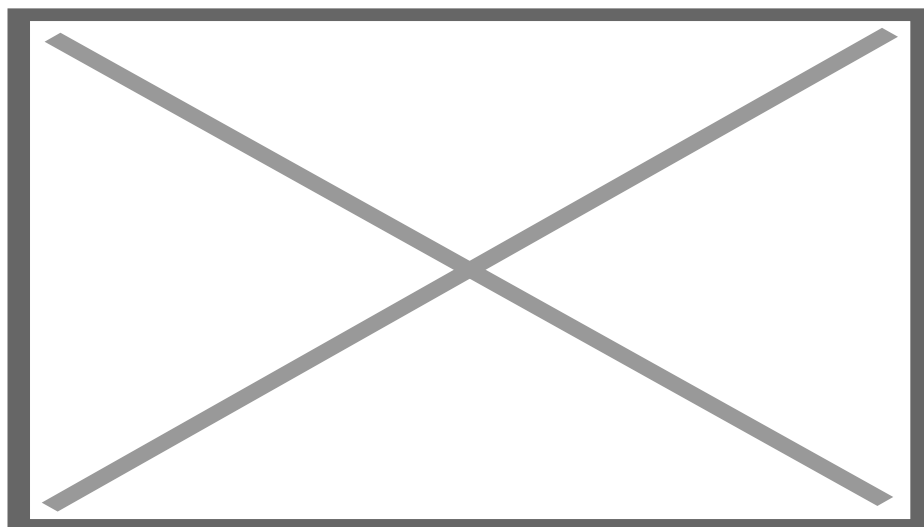
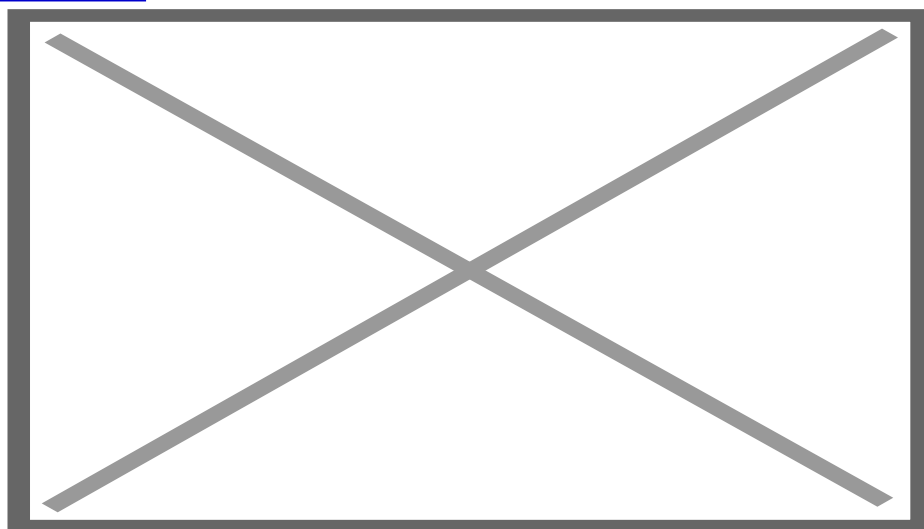
Son programme propose des modules obligatoires et des modules optionnels. En fonction des modules, les cours ont lieu au laboratoire de mathématiques d'Orsay, à CentraleSupélec ou laboratoire d'informatique de Paris-Saclay.

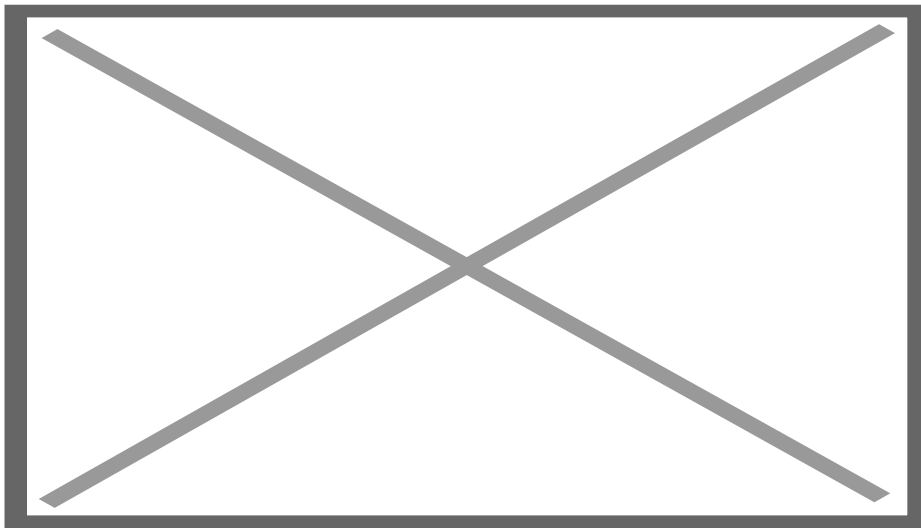
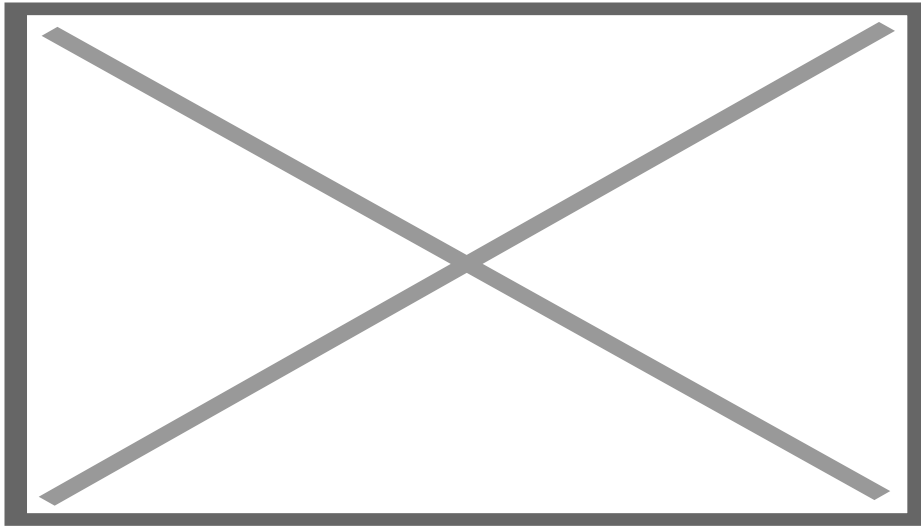
Les modules optionnels sont conçus pour permettre à l'étudiant.e de choisir son curseur entre un programme très mathématique et un programme plus appliqué.

Pour plus d'informations, vous pouvez consulter le site web de cette formation M2 Mathématiques et Intelligence artificielle.

[Télécharger la fiche détaillée](#)

[Je candidate](#)





■ ■

Informations

- [Présentation](#)
- [Admission](#)
- [Programme](#)

Présentation

Compétences

Maîtriser et mettre en oeuvre des outils et méthodes mathématiques de haut niveau, en particulier en lien avec le machine learning et l'IA

Objectifs pédagogiques de la formation

Contexte. Les mathématiques jouent un rôle important en intelligence artificielle (IA), et notamment en apprentissage. Les sciences des données, qui allient modélisation mathématique, statistique, informatique, visualisation et applications, ont pour objectif de passer du stockage et de la diffusion de l'information à la création de connaissances. Ce passage des données aux connaissances requiert une approche interdisciplinaire qui s'appuie fortement sur le traitement statistique de l'information (statistiques mathématiques, statistiques numériques, apprentissage statistique ou machine learning).

La grande dimension pousse à l'utilisation de nouveaux outils issus de différentes branches des mathématiques (analyse fonctionnelle, analyse numérique, optimisation convexe et non convexe) dont il s'agit d'acquérir la compréhension.

Objectifs. Ce parcours permet de maîtriser les enjeux et techniques mathématiques qui fondent l'apprentissage automatique, tout en donnant de solides compétences informatiques pour le développement de projets en apprentissage, sciences des données et IA. Le M2 Mathématiques de l'Intelligence Artificielle à la carte a une double visée :

- acquérir, comprendre et maîtriser un grand nombre d'outils et de méthodes mathématiques et d'algorithmes d'apprentissage en proposant un spectre étendu de cours dans l'ensemble des domaines des mathématiques et de leurs interfaces (notamment l'informatique) avec l'intelligence artificielle.
- permettre aux étudiants de découvrir de l'intérieur l'activité de recherche en mathématiques aux interfaces avec l'intelligence artificielle en se confrontant à des problématiques ouvertes dès le début des cours, d'une activité de recherche sous forme de stages, mais aussi de cours proposant des mini-projets.

Débouchés

Poursuite d'études

**Chercheur/chercheuse en R&D ou expert·e en modélisation et analyse de données dans des entreprises ou laboratoires de pointe.
Data Scientist, Data Analyst, Ingénieur·e en Machine Learning dans des secteurs innovants (tech, finance, santé, énergie, etc.) ;
domaines de l'apprentissage statistique, de l'intelligence artificielle ou de l'analyse de données avancée
Ingénierie études, recherche et développement
Poursuite en thèse en apprentissage statistique, mathématiques appliquées ou intelligence artificielle ;**

Admission

Capacité d'accueil

Places

30

Public visé et prérequis

Le public visé est celui des étudiant.es ayant suivi un cursus de M1 de mathématiques avec une composante statistique, de sciences de données et des compétences en informatique ou un cursus de M1 d'informatique ou d'intelligence artificielle avec une solide composante mathématique.

Le cursus est en particulier pensé en continuité du M1 Mathématiques et IA.

Il est aménagé pour les élèves de CentraleSupélec ayant suivi le cursus SDI, mais peut accueillir, en césure par exemple, des élèves d'autres écoles d'ingénieur ayant les prérequis mathématiques.

Période(s) de candidature

Plateforme Inception

Du 01/04/2026 au 28/06/2026

Pièces justificatives

Obligatoires

Lettre de motivation.

Tous les relevés de notes des années/semestres validés depuis le BAC à la date de la candidature.

Curriculum Vitae.

Facultatives

Lettre de recommandation ou évaluation de stage.

Une lettre de recommandation n'est pas obligatoire mais très fortement conseillée.

Attestation de français (obligatoire pour les non francophones).

Dossier VAPP (obligatoire pour toutes les personnes demandant une validation des acquis pour accéder à la formation) <https://www.universite-paris-saclay.fr/formation/formation-continue/validation-des-acquis-de-lexperience>.

D'éventuelles lettres de recommandation.

Document justificatif des candidats exilés ayant un statut de réfugié, protection subsidiaire ou protection temporaire en France ou à l'étranger (facultatif mais recommandé, un seul document à fournir) :

- Carte de séjour mention réfugié du pays du premier asile
- OU récépissé mention réfugié du pays du premier asile
- OU document du Haut Commissariat des Nations unies pour les réfugiés reconnaissant le statut de réfugié
- OU récépissé mention réfugié délivré en France
- OU carte de séjour avec mention réfugié délivré en France
- OU document faisant état du statut de bénéficiaire de la protection subsidiaire en France ou à l'étranger.

Programme

Modules Fondamentaux (S1)

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Méthodes supervisées et data challenge	5	Semestre 1	24								



Méthodes supervisées et data challenge

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

Dans un premier temps, les acquis de M1 en apprentissage supervisé seront consolidés sur plusieurs jeux de données de taille conséquente donnant lieu à data challenge entre les étudiants. Les stratégies seront comparées. Ces expériences pratiques seront sous-tendues par des mises en perspective théoriques sur les bornes de risque des méthodes ; la palette des méthodes sera étendue à des méthodes non encore étudiées, comme par exemple les SVM et méthodes à noyau. Dans un second temps seront étudiés les fondements théoriques et pratiques de l'apprentissage semi-supervisé.

Nature de l'évaluation

Evaluation Continue non Intégrale

[Méthodes non
supervisées
avancées](#)

5

Semestre 1 24

Méthodes non supervisées avancées

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

Méthodes de recherche de représentation parcimonieuse de données • minimisant l'erreur de reconstruction : ACP non linéaire, NMF, ACP sparse, auto-encoder, ... • minimisant la perte d'une relation : MDS, isomap, tSNE , ... Méthodes de clustering • clustering par modèle de mélange, algorithme EM • clustering spectral • méthodes de clustering de graphes, détection de communautés • co-clustering, systèmes de recommandation

Objectifs d'apprentissage

Les méthodes non supervisées travaillent sur des données non étiquetées. Elles diffèrent suivant l'objectif poursuivi : (1) réduire leur dimension de façon optimale (soit à des fins de visualisation, soit pour l'utilisation dans des étapes supervisées ultérieures) (2) réduire leur nombre en les regroupant en classes d'observations homogènes. Des méthodes de base (ACP pour le premier objectif, Kmeans et classification hiérarchique) ont été vues en première année. Il s'agit d'élargir le spectre de ces méthodes, en comprenant leurs fondements mathématiques, les éventuels liens entre elles, leurs domaines de validité et leurs écueils potentiels afin de les utiliser à bon escient sur des jeux de données réels.

Nature de l'évaluation

Evaluation Continue non Intégrale

[Introduction à](#)

[l'apprentissage](#)

5

Semestre 1

24

[par renforcement](#)

Introduction à l'apprentissage par renforcement

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

Recommended prerequisites are a strong understanding of probability theory, linear algebra; familiarity with optimization, machine learning; and some experience with deep learning.

Programme / plan / contenus

- Markov decision processes
- Dynamic programming
- Monte Carlo methods
- Temporal-Difference Learning
- Value function approximation
- Planning and models
- Policy Gradient
- Actor-critic methods
- Exploration and exploitation
- Multi-step & Off Policy
- Deep RL

Objectifs d'apprentissage

The goal of the course is to understand the theoretical foundations of RL: MDP, value functions, and policies; implement and apply common RL algorithms and apply RL methods to real-world problems.

Bibliographie

- Sutton & Barto (2018) Reinforcement learning: An introduction, MIT press.
- François-Lavet, Henderson & Islam et al. (2018) An introduction to deep reinforcement learning, Foundations and Trends in Machine Learning.

Nature de l'évaluation

Evaluation Continue non Intégrale

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
[M2 Inf-AI] Deep Learning for NLP	2.5	Semestre 1	10.5	10.5							



[M2 Inf-AI] Deep Learning for NLP

ECTS : 2.5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 10.5

Travaux dirigés : 10.5

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

keywords: Machine learning, statistics, probability theory, python

Programme / plan / contenus

1. Introduction, language models
2. Convolutional neural nets
3. Recurrent neural networks
4. Attention mechanisms
5. Word representations, Transformers
6. Generative models

Nature de l'évaluation

Evaluation Continue non Intégrale

Modélisation en grande dimension	5	Semestre 1	24
--	---	------------	----

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
■											

Modélisation en grande dimension

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

Nous montrerons les limites des procédures usuelles dans ce contexte. Nous présenterons des méthodes de sélection de variables dont nous mettrons en évidence les avantages et inconvénients à la fois d'un point de vue théorique et pratique. Nous présenterons des méthodes de régularisation adaptées à différentes problématiques. Enfin, nous introduirons des méthodes de screening permettant de gérer le cas de l'ultra grande dimension lorsque les méthodes de régularisation ne sont pas suffisantes. Des TP en R permettront de mettre en pratique les différentes notions vues en cours.

Objectifs d'apprentissage

comprendre les problèmes de la grande dimension (grand nombre de variables)

connaître les méthodes adaptées et savoir les utiliser à bon escient sur des jeux de données de grande dimension

Nature de l'évaluation

Evaluation Continue non Intégrale

[\[AI\] TC4:](#)

[Probabilistic](#)

[Generative](#)

[Models](#)

2.5

Semestre 1

16.5

4.5

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
■											

[AI] TC4: Probabilistic Generative Models

ECTS : 2.5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 16.5

Travaux dirigés : 4.5

Langue d'enseignement

Français

Enseignement à distance

non

Nature de l'évaluation

Evaluation Continue non Intégrale

[\[M2 Inf-AI\]](#)

[Speech and](#)

[Language](#)

[Processing](#)

2.5

Semestre 2 21

[M2 Inf-AI] Speech and Language Processing

ECTS : 2.5

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 21

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

good knowledge of python notebooks, datavisualization, pandas, matplotlib, ... as shown in the T1 class Hands-On Machine Learning
some knowledge from M1 Hands-On NLP (tokens, embeddings, data structure of raw and annotated texts, notebooks)

Objectifs d'apprentissage

various talks on the following subjects:

1. From linguistics to NLP
2. Treebanks and oral syntax
3. NLP, semantic, mutlimots
4. Speech-Audio processing + ASR + TP
5. Emotion detection
6. Speech interaction

Nature de l'évaluation

Evaluation Continue non Intégrale

[Sequential Learning](#)

5

Semestre 1 24

Sequential Learning

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

Probability Theory Measure Theory

Programme / plan / contenus

- Online learning with experts
- Azuma-Hoeffding inequality
- Stochastic multi-armed bandits : regret bounds of main algorithms
- Kullback Leibler divergence
- Lower bounds
- Linear bandits
- Contextual and continuous bandits
- Best arm identification

Objectifs d'apprentissage

Sequential learning deals with data observed on the fly. Algorithms treat these data in a sequential (or online) manner, hoping to maximise their cumulated reward. It has been popularised in the last decade due to its numerous applications, which include online recommendation, repeated auctions or spam detection. The objective of this course is to introduce the main concepts of stochastic online learning (regret, adaptivity...) from a rigorous mathematical point of view. In particular, the course will focus on stochastic multi-armed bandits, aiming at understanding the main proofs techniques and algorithms. This course will also introduce the technical tools used in the regret bounds proof, including concentration inequalities and information theory.

Bibliographie

"Prediction, learning, and games" Nicolò Cesa-Bianchi and Gábor Lugosi. Cambridge University Press, 2006

"Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems" Sebastien Bubeck and Nicolò Cesa-Bianchi. In Foundations and Trends in Machine Learning, 2012 "Bandit algorithms" Tor Lattimore and Csaba Szepesvári. Cambridge University Press, 2020.

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Introduction à l'apprentissage profond	5	Semestre 1	24								
■											

Introduction à l'apprentissage profond

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

Connaissances de base en Python.

Programme / plan / contenus

- histoire de l'apprentissage profond et sa relation avec les sciences cognitives ;
- réseaux feedforward, régularisation et optimisation ;
- apprentissage par représentation et réseaux siamois ;
- GAN et apprentissage par transfert ;
- réseaux récurrents et LSTM pour le traitement du langage naturel ;
- détection d'objets ;
- méthodes auto-supervisées ;
- apprentissage par renforcement profond.

Objectifs d'apprentissage

Le cours vise à initier les étudiants au Deep Learning par le développement de projets pratiques. Nous espérons qu'à la fin du cours, les étudiants seront en mesure de mettre en œuvre et de développer de nouvelles méthodes dans différents domaines.

Bibliographie

Apprentissage profond d'Ian Goodfellow, Yoshua Bengio, et Aaron Courville, MIT Press, 2016.

Nature de l'évaluation

Evaluation Continue non Intégrale

[Optimisation \(M2 Maths et IA\)](#) 5 Semestre 1 24

Optimisation (M2 Maths et IA)

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

- Problèmes d'optimisation, Existence et unicité, Caractérisation des solutions
- Théorème de Fritz John
- Méthodes numériques pour la recherche de minima locaux sans contraintes
- Optimisation sous contraintes : projection, pénalisation, dualité
- Contrôle optimal
- Méthodes heuristiques
- Résolution exacte de problèmes d'optimisation discrète : branch and bound
- Principaux problèmes d'optimisation sur les graphes

Objectifs d'apprentissage

L'optimisation est le domaine étudiant la minimisation ou la maximisation d'un critère à valeurs réelles. Pour l'optimisation continue, le critère est défini sur un ensemble fermé, d'intérieur non vide. Pour l'optimisation discrète, le critère est défini sur un ensemble fini ou dénombrable.

L'objectif de ce cours est tout d'abord de présenter le cadre formel des problèmes d'optimisation et d'étudier les questions d'existence et d'unicité, de caractérisation des solutions.

Organisation générale et modalités pédagogiques

Ce cours d'adresse principalement aux élèves de CentraleSupélec qui ne l'auraient pas déjà suivi

Bibliographie

Culioli, J. (1994). Introduction à l'Optimisation. Paris : Ellipses.

Evans, L.C. (1987). An Introduction to Mathematical Optimal Control Theory. Berkeley Lecture Notes.

Hiriart-Urruty, J. and Lemaréchal, C. (2001).

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
<u>Plateformes et langages de programmation en données massives</u>	5	Semestre 1	24								

Plateformes et langages de programmation en données massives

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

Programmation Python

Programme / plan / contenus

Chapitre 1. Programmation MapReduce.

- Introduction à MapReduce: motivations et exemples.
- Mise en œuvre de MapReduce.

Chapitre 2. Hadoop.

- Introduction à Hadoop et ses fonctionnalités.
- Hadoop Distributed File System (HDFS)
- Implémentation de MapReduce en Hadoop.
- L'écosystème Hadoop (Hive, Impala, Pig, Kafka, HBase....)

Chapitre 3. Introduction à Apache Spark.

- Introduction à Spark et ses fonctionnalités (RDD, transformations, actions, dataframes, datasets).
- Architecture de Spark et modèle d'exécution.
- Mise en œuvre d'Apache Spark.

Chapitre 4. Composantes d'Apache Spark.

- SparkSQL, Spark Streaming, Machine learning and graph analysis.

Chapitre 5. NoSQL databases.

- Introduction à NoSQL (bases de données distribuées, familles de bases de données NoSQL).
- Bases de données orientées document : MongoDB.
- Bases de données graphe : Neo4j

Objectifs d'apprentissage

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Principes théoriques de l'apprentissage profond	5	Semestre 1	24								
■											

Principes théoriques de l'apprentissage profond

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

- Rappels d'apprentissage statistique et optimisation
- Bornes PAC pour les réseaux de neurones
- Bornes PAC-Bayes et bornes de compression
- Limites du framework PAC et `biais inductifs' liés à l'optimisation
- NTK et dynamique d'apprentissage

Objectifs d'apprentissage

- une compréhension approfondie des aspects spécifiques à l'apprentissage profond qui le différencient d'autres techniques de machine learning.
- une connaissance de différents critères d'évaluation théoriques pour les méthodes de machine learning
- une connaissance des principales hypothèses avancées par la communauté scientifique pour expliquer des succès du deep learning, ainsi que des limites de ces hypothèses.

Nature de l'évaluation

Evaluation Continue non Intégrale

Optimisation pour la vision par ordinateur	5	Semestre 1	24
--	---	------------	----

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
■											

Optimisation pour la vision par ordinateur

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

L'optimisation est un pilier de la vision par ordinateur moderne et est à l'origine d'une grande partie de son succès. L'optimisation du premier ordre offre une méthode évolutive pour résoudre de nombreux problèmes pratiques rencontrés dans ce domaine. Ce cours explore les fondements théoriques et les aspects pratiques de l'application des techniques d'optimisation du premier ordre aux problèmes de vision par ordinateur. Il aborde des concepts fondamentaux tels que l'analyse et l'optimisation convexes, les ondelettes, l'apprentissage de caractéristiques parcimonieuses, le traitement d'images, l'apprentissage profond et la géométrie épipolaire. Des séances pratiques permettront de mettre en œuvre ces techniques pour résoudre des problèmes concrets, notamment le débruitage, le défloutage, la restauration d'images, la super-résolution et la mise en correspondance stéréoscopique.

Objectifs d'apprentissage

Résoudre des problèmes d'optimisation convexe

mettre en oeuvre sur des problématiques de traitement d'image

Nature de l'évaluation

Evaluation Continue non Intégrale

[Module
d'ouverture](#)

5

Semestre 1

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
■											

Module d'ouverture

ECTS : 5

Semestre calendaire : Semestre 1

Langue d'enseignement

Anglais

Enseignement à distance

non

Organisation générale et modalités pédagogiques

UE d'un autre master.

Nature de l'évaluation

Evaluation Continue non Intégrale

[Probabilités et](#)

[statistiques en](#)

[grande dimension](#)

[\(M2 Maths et IA\)](#)

5

Semestre 1 24

Probabilités et statistiques en grande dimension (M2 Maths et IA)

ECTS : 5

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Anglais

Enseignement à distance

non

Programme / plan / contenus

La principale difficulté du statisticien face aux données du XXIème siècle est de vaincre le fléau de la grande dimension. Dans la première partie du cours, nous commencerons par comprendre d'où vient ce fléau et quels concepts permettent de le vaincre. Ensuite, nous verrons comment rendre opérationnels ces concepts, avec une attention sur les frontières du possible. Pour l'essentiel, nous resterons dans un cadre gaussien afin que les aspects techniques ne viennent pas masquer les principales idées. La seconde partie du cours, sera principalement consacrée aux outils probabilistes fondamentaux indispensables pour analyser des problèmes en grande dimension. Du temps sera consacré à démontrer des inégalités de concentration sur de grandes matrices aléatoires, à obtenir des résultats de chaînage pour montrer des inégalités maximales et faire des liens avec la géométrie des espaces de Banach. L'accent sera mis sur les techniques mathématiques et sur la généricité des approches déployées.

Objectifs d'apprentissage

L'objectif principal de ce cours est de :

- comprendre les phénomènes liés à la grande dimension;
- fournir des bases conceptuelles et méthodologiques solides;
- acquérir des techniques mathématiques fondamentales en vue d'une thèse.

Bibliographie

Christophe Giraud, Introduction to High-Dimensional Statistics (2nd Edition), CRC Press (2021).

Nature de l'évaluation

Evaluation Continue non Intégrale

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
[M2 Inf-AI] Signal Processing	2.5	Semestre 2									
■											

[M2 Inf-AI] Signal Processing

ECTS : 2.5

Semestre calendaire : Semestre 2

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

keywords:

Hilbert analysis (norms, inner product)

Linear algebra

Probabilty (random variables)

Objectifs d'apprentissage

<http://hebergement.universite-paris-saclay.fr/mkowalski/SigPro/>

[https://mycore.core-](https://mycore.core-cloud.net/index.php/s/Df9IMQK5Zun0PAU)

[cloud.net/index.php/s/Df9IMQK5Zun0PAU](https://mycore.core-cloud.net/index.php/s/Df9IMQK5Zun0PAU)

1. Introduction & Fourier analysis

2. Filtering

3. Random signals

4. Time-Frequency analysis

5. Time-Scale analysis

6. Introduction to linear inverse problems

Nature de l'évaluation

Détail renseigné au niveau Élément Constitutif

Parcours à la carte (S2)

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Apprentissage séquentiel, optimisation et jeux	4	Semestre 2	20								

Apprentissage séquentiel, optimisation et jeux

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 20

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

probability theory and notions of convex analysis

Programme / plan / contenus

- Regret minimizing algorithms
- Online linear/convex optimization
- Adaptive regret bounds
- Blackwell's approachability
- Adversarial bandits & partial monitoring
- Application to optimization, variational inequalities, fixed-point iterations
- Application to learning in games

Objectifs d'apprentissage

Online learning is an important topic at the intersection of machine learning, optimization and sequential decision.

This course focuses on a general class of sequential decision problems in adversarial environments. We first introduce the mathematical tools to design and analyze a wide range of regret minimizing algorithms in online linear/convex optimization. These results are then extended to a wider range of problems through Blackwell's approachability. We then explore deep connections with various optimization problems: (stochastic) optimization, minimax problems, variational inequalities, learning in games, fixed-point iterations.

Bibliographie

- "Prediction, learning, and games", Nicolò Cesa-Bianchi and Gábor Lugosi. Cambridge University Press, 2006
- "Approachability, regret and calibration: Implications and equivalences" ,Vianney Perchet, Journal of Dynamics & Games (2014)
- "A modern introduction to online learning", Francesco Orabona, lecture notes, 2023

Nature de l'évaluation

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Guidelines in Statistical Learning	4	Semestre 2	32								
■											

Guidelines in Statistical Learning

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 32

Langue d'enseignement

Anglais

Enseignement à distance

non

Programme / plan / contenus

- valeurs manquantes
- prédiction conformelle
- robustesse

Nature de l'évaluation

Evaluation Continue non Intégrale

[Mathématiques
des modèles
génératifs](#)

4 Semestre 2 24

Mathématiques des modèles génératifs

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

Le cours est structuré en quatre grands blocs, alternant théorie (CM) et pratique (TP).

1. Panorama et modèles à vraisemblance explicite :

- Histoire, taxonomie et relations entre les familles de modèles.
- Distinction génératif/discriminant.
- Modèles autorégressifs.
- Auto-Encodeurs (AE) et Auto-Encodeurs Variationnels (VAE).

2. L'approche adversariale :

- Théorie des Réseaux Antagonistes Génératifs (GAN).
- Architectures (DCGAN) et défis pratiques.

3. Modèles de l'état de l'art basés sur les flux et l'énergie :

- Modèles à flux normalisant (Normalizing Flows).
- Modèles basés sur l'énergie (EBM) et dynamique de Langevin.
- Modèles de diffusion et Flow Matching.

4. Synthèse, applications et perspectives :

- Modèles de fondation.
- Évaluation des modèles et mise en application.
- Perspectives de recherche actuelles du domaine.

Objectifs d'apprentissage

- Classifier les grandes familles de modèles génératifs (autorégressifs, variationnels, adversariaux, basés sur les flux, basés sur l'énergie, etc.) et comprendre leurs liens.
- Comprendre les fondements théoriques des architectures clés : VAE, GAN, Modèles de Diffusion, Normalizing Flows, Flow Matching et les modèles de fondation.
- Maîtriser les concepts avancés inspirés de la dynamique de Langevin.
- Implémenter des versions fonctionnelles des modèles (VAE, GAN, Diffusion)

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
<u>Analyse topologique des données</u>	4	Semestre 2	20								

Analyse topologique des données

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 20

Langue d'enseignement

Anglais

Enseignement à distance

non

Prérequis

Les notions nécessaires de topologie et de géométrie seront introduite ou rappelée au fil du cours.

Programme / plan / contenus

Plan du cours :

1. Introduction et rappels
2. Fonctions distance et reconstruction d'ensembles compacts.
3. Application à l'estimation de l'homologie de de sous-variétés à partir d'échantillon i.i.d. Aspects algorithmiques et statistiques.
4. Homologie persistente : définition, stabilité et applications (clustering, signatures topologiques,...).
5. Propriétés statistiques de l'homologie persistente.
6. Inférence géométrique pour les mesures de probabilité.

Objectifs d'apprentissage

L'estimation des propriétés de nature topologique et géométrique de données souvent représentées par des ensembles de points dans des espaces euclidiens de grande dimension (ou plus généralement des variétés riemanniennes ou des espaces métriques) connaît un développement important depuis quelques années.

L'analyse topologique des données est motivée par l'observation que généralement les données, qui ne sont pas distribuées uniformément dans l'espace ambiant mais se concentrent autour de structures géométriques de plus petite dimension qu'il est important de comprendre.

L'objectif de ce cours est de donner une introduction à ce sujet en pleine expansion.

Nature de l'évaluation

Evaluation Continue non Intégrale

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Analyse de données multivariées avancée	4	Semestre 2	24								

Analyse de données multivariées avancée

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

Connaissances de bases en statistique, optimisation, algèbre linéaire.

Programme / plan / contenus

• Données tensorielle (PARAFAC, Tucker, Coupled Matrix Tensor Factorization, Régression Tensorielle, etc...) • Données multi-tableau (Analyse canonique généralisée régularisée et cas particuliers) • Méthodes à noyaux (e.g. kernel PCA, kernel GCCA). • Modèles à équations structurels (SEM)

Objectifs d'apprentissage

Les méthodes statistiques standards permettent l'analyse d'un tableau individus $\times$ variables. Cette structure de données est souvent trop limitée pour permettre de représenter des données plus complexes. A titre d'exemples, citons (i) les données de nature tensorielle où les mêmes variables sont observées selon plusieurs modalités (e.g. données spatiotemporelle) ou (ii) les données multi-tableau ou les individus sont caractérisés par des variables de nature hétérogène (e.g. données d'imagerie-génétique). Ce cours dresse un panorama des méthodes permettant l'analyse de ce type de données complexes.

Bibliographie

Adachi, K. (2016). Matrix-based introduction to multivariate data analysis. Springer

Nature de l'évaluation

Evaluation Continue non Intégrale

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Statistique bayésienne et applications	4	Semestre 1	24								

Statistique bayésienne et applications

ECTS : 4

Semestre calendaire : Semestre 1

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Programme / plan / contenus

• Statistique Bayésienne -- Modèle paramétrique. Loi a priori. Loi a posteriori. Performance d'un estimateur en moyenne (fonction de perte, risque). A priori diffus. A priori conjugués. • Echantillonnage par chaînes de Markov -- Intégration en grande dimension. Principes élémentaires de simulations d'une variable aléatoire scalaire. Principe d'acceptation-rejet. Principes et variantes de l'algorithme de Metropolis-Hastings. Introduction aux chaînes de Markov et éléments d'analyse des algorithmes MCMC. • Problèmes inverses -- Définition et exemples de problèmes inverses, problèmes mal posés. Régularisation et méthodes classiques d'inversion. Approche bayésienne de la résolution des problèmes inverses. Inversion dans le cas gaussien et filtre de Kalman. Approches MCMC et bayésiennes variationnelles. Applications.

Objectifs d'apprentissage

Ce cours comprend trois parties. La première est consacrée aux fondements de la statistique bayésienne. L'inférence bayésienne repose sur le choix d'un modèle bayésien, lui-même constitué d'un modèle des observations et d'une loi de probabilité a priori pour les paramètres du modèle. Étant donnée des observations, on calcule ce que l'on appelle une loi a posteriori des paramètres du modèle bayésien (à l'aide de la règle de Bayes). La connaissance de cette loi a posteriori permet ensuite d'accéder à la loi a posteriori des quantités d'intérêt. La deuxième partie du cours est consacrée aux techniques d'échantillonnage par chaînes de Markov (Monte Carlo Markov Chains). La troisième partie est consacrée aux applications de l'inférence bayésienne, en particulier aux problèmes inverses. Par problème inverse, on entend la détermination d'une grandeur (scalaire ou vectorielle) à partir d'observations indirectes (lorsque l'observation directe d'une grandeur physique n'est pas possible).

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Statistical Theory of Algorithmic Fairness	4	Semestre 2	24								
■											

Statistical Theory of Algorithmic Fairness

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Anglais

Enseignement à distance

non

Objectifs d'apprentissage

With the ubiquitous deployment of machine learning algorithms in nearly every area of our lives, the problem of unethical or discriminatory algorithm-based decisions becomes more and more prevalent. To partially address these concerns, a new sub-field of machine learning has emerged. The goal of the course is to introduce the audience to recent developments of fairness aware algorithms. The emphasise will be made on those methods which are supported by statistical guarantees and that can be implemented in practice. We will study classification and regression problems under the so called demographic parity constraint—a popular way to define fairness of an algorithm. Several research directions will be proposed thought out the course.

Nature de l'évaluation

Evaluation Continue non Intégrale

[Modèles](#)

[graphiques :](#)

[inférence discrète](#)

[et apprentissage](#)

4

Semestre 2 24

Modèles graphiques : inférence discrète et apprentissage

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

- Solide compréhension des modèles mathématiques
- Algèbre linéaire
- Transformations intégrales
- Équations différentielles
- Idéalement, un cours de base en optimisation discrète.

Programme / plan / contenus

- Modèles graphiques
- Propagation de croyance
- Introduction aux coupes de graphes
- Coupes de graphes + Primal-dual, Partie I
- Coupes de graphes + Primal-dual, Partie II
- Systèmes de recommandation
- Causalité
- Apprentissage
- Réseaux bayésiens

Objectifs d'apprentissage

Ce cours se concentrera sur les modèles discrets, c'est-à-dire les cas où les variables aléatoires des modèles graphiques sont discrètes. Après une introduction aux bases des modèles graphiques, le cours se concentrera sur les problèmes de représentation, d'inférence et d'apprentissage des modèles graphiques. Nous couvrirons les algorithmes classiques ainsi que l'état de l'art utilisés pour ces problèmes. Plusieurs applications en apprentissage automatique et en vision par ordinateur seront étudiées dans le cadre du cours.

Nature de l'évaluation

Evaluation Continue non Intégrale

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
Méthodes géométriques dans l'analyse des données	4	Semestre 2	24								

Méthodes géométriques dans l'analyse des données

ECTS : 4

Semestre calendaire : Semestre 2

Détail du volume horaire

Cours magistraux : 24

Langue d'enseignement

Français

Enseignement à distance

non

Prérequis

Il est attendu des étudiants qu'ils aient une bonne connaissance de :

- Des bases en algorithmique (notions de complexité).
- Bases en algèbre linéaire, géométrie, théorie des probabilités.
- La maîtrise d'un langage de programmation (C/C++, python, R) est également attendue.

Programme / plan / contenus

- Plus proches voisins dans les espaces euclidiens et métriques : structures de données et algorithmes de recherche
- Proches voisins dans les espaces euclidiens et métriques : analyse
- Algorithmes de réduction de dimensionnalité
- Couvertures et nerfs : inférence géométrique et l'algorithme Mapper
- Algorithmes de classification et introduction à l'homologie persistante
- Tests d'hypothèses statistiques et tests à deux échantillons (TST)
- Comparaison de distributions à haute dimension, comparaison de clustering
- Signatures de formes : stabilité et aspects statistiques

Objectifs d'apprentissage

Ce cours passe en revue les constructions fondamentales liées à la manipulation de nuages de points échantillonnés, en mélangeant les idées de la géométrie et de la topologie computationnelles, des statistiques et de l'apprentissage automatique. L'accent est mis sur les méthodes qui non seulement présentent des garanties théoriques, mais qui fonctionnent également bien dans la pratique. En particulier, des références logicielles et des jeux de

[illegible]

Module d'ouverture

Semestre calendaire : Semestre 2

Langue d'enseignement

Anglais

Enseignement à distance

non

UE d'un autre master.

Evaluation Continue non Intégrale

Séminaire des étudiants

Semestre 2

Séminaire des étudiants

Semestre calendaire : Semestre 2

Langue d'enseignement

Français

Enseignement à distance

non

Nature de l'évaluation

Evaluation Continue Intégrale

[illegible]

Matières	ECTS	Semestre	Cours	TD	TP	Cours- TD	Cours- TP	TD- TP	A distance	Projet	Tutorat
■											

Stage ou mémoire de recherche

ECTS : 16

Semestre calendaire : Semestre 2

Langue d'enseignement

Français

Enseignement à distance

non

Nature de l'évaluation

Evaluation Continue non Intégrale

Lieu(x) d'enseignement

ORSAY

GIF SUR YVETTE

PALaiseau

Contact(s)

- **Gilles BLANCHARD**

PR

Liens

[Blog Destination études supérieures : quel parcours pour être data scientist?](#)

[Présentation vidéo du master](#)

<http://www.universite-paris-saclay.fr/formation/master/mathematiques-et-applications/m2-mathematiques-et-intelligence-artificielle>