Université Paris Saclay

M2 Mathématiques et IA

Syllabus des UEs

université PARIS-SACLAY

FACULTÉ DES SCIENCES D'ORSAY

Mathématiques Orsay

CentraleSupélec

Mise à jour : 04/09/2024

Table des matières

Période 1

Méthodes supervisées avancées et data challenge (Orsay)

Responsable: Olivier Coudray

Langues d'enseignement : FRANCAIS

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.

Méthodes non supervisées avancées (Orsay)

Responsable: Christine Keribin

Langues d'enseignement : FRANCAIS

Objectif : 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.

Contenu :

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

Statistique en grande dimension (StatML)

Responsable : Christophe Giraud

Langues d'enseignement : FRANCAIS

Objectifs principaux :

- 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.

La principale difficulté du statisticien face aux données du XXIe siècle est de vaincre le fléau de la grande dimension. Ce fléau oppose aux statisticiens deux difficultés : d'une part il rend les méthodes statistiques classiques totalement inopérantes par manque de précision, d'autre part il oblige à développer des approches gardant sous contrôle la complexité algorithmique des procédures d'estimation.

Ce module est la première partie du cours du même nom du M2 StatML. Nous commençons par comprendre d'où vient ce fléau et comment le vaincre dans un contexte général. Ensuite, nous verrons comment rendre opérationnels ces concepts, avec une attention sur les frontières du possible.

website : https://www.imo.universite-paris-saclay.fr/ giraud/Orsay/HDPS.html

Plateformes et Langages de Programmation (CentraleSupelec, 3MD1060)

Responsable : Gianluca Quercini

Langues d'enseignement : FRANCAIS

Prérequis : Programmation Python

Contenu :

Exploiter tout type de données, structurées ou pas, y compris massives.

- Comprendre les concepts à la base du Big Data.
- Utiliser des paradigmes de calcul distribué : MapReduce et Spark.
- Concevoir des algorithmes de calcul distribué sur les données.
- Modéliser des données dans des bases de données NoSQL
- Ecosystems Hadoop, MongoDB, Neo4j

Optimisation (CentraleSupelec, 3MD1020)

Responsables : Vincent Lescarret

Langues d'enseignement : FRANÇAIS

Prérequis : Calcul Différentiel

Contenu :
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 et de méthodes numériques 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 et de méthodes numériques.

- Problèmes d'optimisation, Existence et unicité
- Optimisation sous contraintes, Multiplicateurs de Lagrange
- Lagrangien et dualité, Théorème KKT
- Algorithme d'uzawa
- Optimisation discrète en programmation linéaire, algorithme du simplexe

Probabilistic Generative Models (Info TC4)

Responsable: Caio Corro

page web: http://teaching.caio-corro.fr/2020-2021/TC4/

Pré-requis: Basics of linear algebra and probabilities

Contenu:

Recently, generative models have (again) become a hot topic in machine learning thanks to recent advances in deep learning. One of the benefits of these models is their ability to generate new data, see for example face or word generation. Moreover, they can be used for semi-supervised learning, feature extraction via latent variables, …

In this course, we will first review the theoretical background required to understand modern generative models:

- Probabilities,
- Latent variables models,
- Expectation Maximization algorithm,
- Change of variable theorem,
- Neural parameterization of probability distributions.

Based on this, we will study modern generative models based on neural networks:

- Variational Auto-Encoders (VAEs),
- Generative Adversarial Networks (GANs),
- Flow models,
- Energy networks.

Signal processing (info TC5)

Responsable: Matthieu Kowalski

page web: http://hebergement.universite-paris-saclay.fr/mkowalski/teaching.html

Pré-requis: linear algebra, calculus, probability. MatLab or Python programming

Contenu:

This course is a “flipped classroom”: a lot of materials is available through my webpage to study the theorical part of signal processing. Every technics are illustrated by Lab work on real applications. The class are devoted to work on the applications under MatLab or Python:

- Spectral analysis (Application: Guitar tuner and/or zoom on images)
- Filtering (Application: Delay effect and/or image filtering)
- Random signal (Application: Noise spectrum estimation)
- Time-Frequency (Application: audio denoising)
- Wavelets (Application: Image denoising)
- Final project on inverse problem

Période 2

Introduction to reinforcement learning.

Responsable : Joon Kwon

Reinforcement learning (RL) is an area of artificial intelligence and machine learning, where the goal is to learn through experimentation how to best interact with an environment. It is particularly attractive for tasks that require planning and adaptation. Applications are numerous: robotics, manufacturing, self-driving, game playing, consumer modeling, healthcare, etc.

The formal framework is a Markov Decision Process (MDP), which is a sequential decision problem. At each step, given the current state of the environment, an agent chooses an action, and depending on the latter, receives a reward and the state is updated. In a RL setting, the reward and state transition rules are unknown and must be estimated through experiment. The goal is to learn a near-optimal policy i.e. a decision rule that maximizes the cumulative reward.

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.

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

Content:

- 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

Bibliography:

- 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.

Sequential learning (StatML)

Responsable : Etienne Boursier

Langues d'enseignement : FRANCAIS , document en anglais

Objectifs: 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.

Langues d'enseignement : FRANCAIS, documents en anglais

Contenu: The course will be organised in 8x3h lessons. In particular, we will explore the following themes:

- 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

The evaluation will be based on a homework and a final written exam (both first and second sessions).

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.

Optimization for Computer Vision (CentraleSupelec)

Optimization is the workhorse of modern computer vision and responsible for much of its success, with first-order optimization offering a scalable way to solve many of the practical problems that arise in the field. This course explores the theoretical foundations and practical aspects of applying first-order optimization techniques to computer vision problems. The course covers some foundational concepts in convex analysis and optimization, wavelets, sparse feature learning, image processing, deep learning, and epipolar geometry. There will be practical sessions for implementing these techniques to solve real-world problems from image denoising, deblurring, inpainting, super resolution, and stereo matching.

Introduction au deep learning (CentraleSupelec)

Intervenants **:** Vincent Lepetit (ENCP) et Maria Vakalopoulou (CS))

Langues d'enseignement : FRANÇAIS/ANGLAIS

Prérequis : Il n'y a pas de prérequis officiel pour ce cours. Cependant, les étudiants doivent avoir une connaissance de base de Python.

Contenu :

Avec l'augmentation de la puissance de calcul et des quantités de données disponibles, mais aussi avec le développement de nouveaux algorithmes d'apprentissage et de nouvelles approches globales, de nombreuses percées ont eu lieu ces dernières années dans le domaine de l'apprentissage profond pour la reconnaissance d'objets et du langage parlé, la génération de textes et la robotique. Ce cours

couvrira les aspects fondamentaux et les développements récents de l'apprentissage profond dans différents domaines : Vision par ordinateur, traitement du langage naturel, et apprentissage profond par renforcement.

- 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.

Theoretical principles of deep learning (CentraleSupelec)

Responsable : Hédi Hadiji

Les algorithmes d'apprentissage automatique impliquant des réseaux neuronaux profonds ont accumulé les succès empiriques à un rythme spectaculaire au cours des dernières années. Nombre de ces réussites ne peuvent être expliquées par l'intuition issue de la théorie de l'apprentissage standard. En outre, à mesure que la popularité de l'apprentissage profond augmente, le fossé entre la théorie et la pratique ne cesse de se creuser. Construire les bases d'une théorie satisfaisante de l'apprentissage profond, avec l'objectif de guider les praticiens, est un défi majeur de la recherche moderne.

Dans ce cours, nous discuterons des progrès théoriques récents réalisés pour décrire les performances empiriques des méthodes de deep learning. Nous nous concentrerons principalement sur l'étude de la capacité de *généralisation* étonnamment bonne des réseaux profonds. Considérons une tâche de classification, dans laquelle, étant donné un ensemble de features et de labels d'entraînement, nous souhaitons prédire l'étiquette inconnue d'une nouvelle caractéristique de test. Une connaissance superficielle de théorie classique de l'apprentissage laisserait penser que des modèles très complexes doivent overfitter sur les données d'apprentissage, mais la pratique a prouvé à maintes reprises que les réseaux neuronaux donnent de bons résultats malgré un surparamétrage massif. Nous décrirons quelques idées qui ont été proposées pour expliquer ce phénomène ; les sujets **qui pourront être** abord**és** sont les suivants : théories de la généralisation (capacité, marge, stabilité, compression, ...), régularisation implicite par SGD et paysage d'optimisation, bornes PAC-bayes, approximation théorique des grands réseaux (NTK).

Modélisation en grande dimension (CentraleSupelec)

Responsable : Sarah Lemler

Langues d’enseignement : FRANCAIS

Dans ce cours, nous nous intéresserons à la question de la grande dimension lorsque le nombre de covariables (ou variables explicatives) est supérieur au nombre d’observations.

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.

Automatic Speech Recognition and Natural Language Processing (Info OPT 5)

Responsable : Kim Gerdes and Marc Evrard

page web: https://upsay-master-cs-ai.github.io/courses/m2/asr-nlp/

Contenu:

- Seminar 1:
  - General Introduction (Marc Evrard)
  - Voice Recognition (Claude Barras)
- Seminar 2: Speech processing + ASR Practical (Marc Evrard)
- Seminar 3: Automatic Emotion Recognition (Laurence Devillers)
- Seminar 4: From Linguistics to Natural Language Processing (Iona Vasilescu)
- Seminar 5: Treebanks and Oral Syntax (Kim Gerdes)
- Seminar 6: Speech Interaction (Samir Bennacef)
- Seminar 7: Conclusion, Preparation of Paper Reading Assignment (Kim Gerdes)

Deep Learning for NLP (Info Opt11)

Responsable: Caio Corro

page web: http://teaching.caio-corro.fr/2020-2021/OPT11/

Contenu:
Recently, neural networks trained end-to-end have obtained impressive results in many problems related to natural languages (e.g. machine translation). These deep learning techniques do not rely on manual feature extraction or rule-based systems. However, behind the scenes a large part of this success is due to the development of neural architectures that are able to handle structured inputs and outputs.

In this course, we will study how build neural networks for problems related to natural languages. Specifically, we will learn how to:

- build dynamic computation graphs for sequential inputs,
- predict structures (e.g. text generation, semantic parsing),
- introduce inductive bias.

Moreover, we will develop a critical analysis of state-of-the-art NLP models:

- do they actually learn what we expect?
- how does deep learning for NLP handle bias in training data?

Période 3

Guidelines in Machine learning

Responsables : Claire Boyer & Olivier Coudray

Langues d'enseignement : Français, Ressources pédagogiques en anglais

This course explores key themes in the context of real-world data processing and machine learning. It delves into various critical aspects, including data preprocessing, handling missing data, uncertainty

quantification using conformal prediction tools, addressing imbalanced data, and managing distribution shifts between training and test samples.

The following themes will be covered :

- data preprocessing: data encoding, feature scaling, normalization
- imbalanced data
- outliers detection
- handling missing data for imputation/estimation/prediction
- uncertainty quantification: conformal prediction
- distribution shift

Introduction à la théorie statistique de l'apprentissage

Responsable : Gilles Blanchard, Guillermo Durand

Langues d'enseignement : français, avec support en anglais.

Ce cours sera composé de deux parties correspondant aux deux intervenants : théorie statistique de l'apprentissage, et tests multiples.

Le but de la première partie est de donner une introduction aux techniques et outils mathématiques permettant de garantir d'un point de vue théorique que des algorithmes d'apprentissage machine vont bien se comporter d'un point de vue asymptotique, lorsque le nombre de données devient très grand, c'est-à-dire que la fonction de prédiction apprise va converger en un certain sens vers une fonction de prédiction optimale.

Les thèmes suivants seront abordés :

- outils de contrôle de l'erreur de généralisation : complexité de Rademacher, entropie métrique
- dimension de Vapnik-Chervonenkis : cas de la classification, cas de la régression
- bornes inférieures sur les vitesses d'apprentissage sur une classe de fonctions donnée
- certains thèmes spécialisés selon le temps disponible (vitesses de convergence « rapides », agrégation de prédicteurs, boosting, bornes PAC-Bayes…)

La partie tests multiples aura pour but d'introduire des outils de nature mathématique, statistique, et algorithmique, pour traiter du problème du contrôle de l'erreur quand un grand nombre de décisions doivent être prises (tests statistiques). Cette question de nature statistique apparaît fréquemment en combinaison avec l'utilisation de méthodes d'apprentissage statistique sur des données massives. On introduira d'abord le problème de l'explosion du nombre de faux positifs lorsque de nombreux tests statistiques sont réalisés sur un même jeu de données. On abordera ensuite les thèmes suivants :

- Critères d'erreur et procédures classiques
- Adaptivité
- Problème hétérogène discret
- Recherche exploratoire et bornes de confiance post hoc

Online learning, optimization and games (StatML)

Responsable : Joon Kwon

Langues d'enseignement : Français

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.

Recommended prerequisites are probability theory and notions of convex analysis.

Contents:

- 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

Bibliography:

- "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

Statistical theory of algorithmic fairness (StatML)

Responsable : Evgenii Chzen

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.

Modèles graphiques : inférence discrète et apprentissage (CentraleSupelec 3MD3040)

Responsable : Karteek Alahari

Langues d'enseignement : FRANCAIS

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

Contenu

Les modèles graphiques (ou modèles graphiques probabilistes) fournissent un paradigme puissant pour exploiter conjointement la théorie des probabilités et la théorie des graphes pour résoudre des problèmes complexes du monde réel. Ils constituent un élément indispensable dans plusieurs domaines de recherche, tels que les statistiques, l'apprentissage automatique, la vision par ordinateur, où un graphe exprime la dépendance conditionnelle (probabiliste) entre des variables aléatoires.

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.

Statistique bayésienne et applications (CentraleSupelec 3MD3210)

Responsable : Julien Bect

Langues d'enseignement : FRANCAIS

Mots clés :

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. Les grandeurs rentrant en compte (normalisation, loi a posteriori, etc…) nécessitent souvent des techniques d'échantillonnage par chaînes de Markov (Monte Carlo Markov Chains). Le cours comporte trois parties :

- Fondamentaux en 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.
- Quelques sujets "d'ouverture" (principalement du bayésien non-paramétrique).

Modèles géométriques (CentraleSupelec)

Intervenants: Frederic Cazals (Inria) et Mathieu Carriere (Inria)

Langues d'enseignement : FRANÇAIS

Prérequis : Il n'y a pas de prérequis officiel pour ce cours. Cependant, il est attendu des étudiants qu'ils aient une bonne connaissance :

- des bases en algorithmique (notions de complexité).
- des bases en algèbre linéaire, géométrie, théorie des probabilités.

La maîtrise d'un langage de programmation (C/C++, python, R) est également attendue.

Contenu :

L'analyse des donn.es est le processus de nettoyage, de transformation, de modélisation ou de comparaison des données, afin de déduire des informations utiles et de mieux comprendre des phénomènes complexes. D'un point de vue géométrique, lorsqu'une instance (un phénomène physique, un individu, etc.) est donnée comme une collection de taille fixe d'observations à valeurs réelles, elle est naturellement identifiée à un point géométrique ayant ces observations comme coordonnées. Toute collection de telles instances est alors considérée comme un nuage de points échantillonné dans un espace métrique ou normé.

Ce cours passe en revue les constructions fondamentales liées à la manipulation de tels nuages de points, en mélangeant les idées de la géométrie et de la topologie computationnelle, 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 donn.es d'exemple seront fournis pour illustrer les constructions.

- Les plus proches voisins dans les espaces euclidiens et métriques : structures de données et algorithmes de recherche
- Plus proches voisins dans les espaces euclidiens et métriques : analyse Algorithmes de réduction de la 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 forme : stabilité et aspects statistiques

Analyse de données multivariées avancée (CentraleSupelec 3MD3230)

Responsable : Arthur Tenenhaus

Les méthodes statistiques standards permettent l'analyse d'un tableau individus × 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 spatio-temporelle) 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.

Plan détaillé du cours (contenu)

- 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)
- Les méthodes présentées en cours seront mises en oeuvre devant machine au travers des logiciels R ou Python