

Programme du M1 Math&AS (2023-2024) :

Les UEs dotées d'une astérisque sont mutualisées avec les étudiant.e.s du master ISADS.

Les UEs dotées de deux astérisques sont mutualisées avec les étudiant.e.s du master AMS. L'UE dotée de trois astérisques est mutualisée avec les étudiant.e.s du master 1 Informatique.

Optimisation numérique ()** - 48h – 6 ECTS [Laurent Dumas]

De très nombreux problèmes en industrie, en physique et en économie consistent en la minimisation (ou la maximisation) d'une fonction objective. Ce cours vise à présenter un grand nombre de méthodes numériques qui ont été développées pour résoudre de tels problèmes. Ces méthodes peuvent être locales ou globales, déterministes ou stochastiques. De nombreux exemples seront implémentés sur machine afin d'illustrer l'emploi en pratique de ces méthodes.

Probabilités ()** - 30h – 4 ECTS [Catherine Donati-Martin]

Le module est consacré principalement à l'étude des chaînes de Markov à espace d'états discret, avec des applications aux marches aléatoires et à des processus à valeurs dans un espace d'états discret. Dans le cadre de cette étude, nous approfondirons les notions d'espérance conditionnelle et de loi conditionnelle. Le cours se terminera par des théorèmes limites incluant des rappels sur les différents modes de convergence possibles dans le domaine des probabilités.

Algorithmique randomisée et recherche opérationnelle (*)** - 45h – 5 ECTS [Thierry Mautor et Ider Tseveendorj]

Principalement, ce cours concerne l'étude des méthodes de résolution des problèmes d'optimisation difficiles pour lesquels il n'existe pas de méthode de résolution polynomiale connue. Dans ce cadre, ce cours comprend:

** Programmation linéaire, méthode du simplexe, dualité.*

** Heuristiques et meta-heuristiques (Recherche Tabou, Recuit Simulé, Algorithme Génétique). * Méthodes de résolution exacte: Branch and Bound et Branch and Cut.*

Par ailleurs, ce cours aborde également l'algorithmique randomisée à travers l'étude des méthodes de Monte-Carlo et de Las-Vegas.

Traitement des données et machine learning (* mutualisation partielle, avec l'M1 ISADS) - 72h – 8 ECTS [Ester Mariucci et Charles Tillier]

Le but de ce cours est de prédire quels types d'algorithmes de machine learning existants seront les plus appropriés pour quels types de tâches, sur la base de propriétés formelles et de résultats expérimentaux. Et également d'évaluer et analyser les algorithmes d'apprentissage existants. Dans un premier temps on s'intéressera à l'analyse des données, à comment extraire ou synthétiser l'information contenue dans un jeu de données. Nous nous intéresserons à plusieurs classes de méthodes et tout particulièrement aux méthodes d'analyse factorielle (ACP, AFC et ACM). Ensuite, on se focalisera sur les méthodes d'apprentissage supervisé et on abordera plusieurs algorithmes : K-plus proches voisins, régression logistique, SVM, et arbres de décision. Une partie de cette UE est propre aux étudiants du M1 Math&AS (non mutualisée avec l'M1 ISADS) et se consacre à l'étude théorique de certaines séries chronologiques.

Inférence statistique et théorie de l'apprentissage – 72h – 8 ECTS [Ester Mariucci et Emmanuel Rio]

Le but de ce cours est de fournir les outils pour comprendre des travaux avancés en statistique et en apprentissage automatique. Une première partie sera consacrée aux tests statistiques et intervalles de confiance, à l'étude des propriétés des estimateurs classiques, à la statistique asymptotique et à la théorie de la décision. Ensuite on traitera la théorie de l'apprentissage statistique : formulation d'un problème d'apprentissage, minimisation du risque empirique, inégalités de concentration et théorie des processus empiriques.

Modèle linéaire (*) - 54h – 6 ECTS [Julien Worms]

Modèle linéaire général et cas particulier de la régression simple ou multiple, traitement par la méthode des moindres carrés. Propriétés des estimateurs avec ou sans normalité. Décomposition de la variance. Ratios de Student et tests basés sur les matrices de contraste. Intervalles de

prédiction. Test de Fisher général pour modèles emboîtés. Approche par vraisemblance et lien entre le LRT et le test de Fisher. Diagnostics (résidus studentisés, QQ-plots, autres indicateurs usuels). Introduction à la pénalisation de la dimension. Une attention particulière sera portée à la gestion des régresseurs catégoriels (interprétation des coefficients et tests associés).

Logiciels statistiques et Python (*) - 72h – 8 ECTS [Agnès Grimaud, Charles Tillier et Julien Worms]

Dans cette UE nous considérerons trois langages R, SAS et Python.

Partie R : programmation R, manipulation de différents objets, boucles, fonctions. Applications en statistique.

Partie SAS : concepts généraux de la programmation SAS (étapes, librefs, formats, options de noms de tables), fonctionnement fin de l'étape data, jointures, manipulation de données textuelles, arrays, macro-variables et macro-langage.

Partie Python : introduction générale sur Python et la programmation object; utilisation et manipulation des bibliothèques de bases (NumPy, Math, Scipy); initiation à l'apprentissage statistique avec Scikit-Learn.

Anglais – 27h – 3 ECTS

Dans un contexte à caractère professionnel, les cours en anglais Master visent à aider les étudiants à faire face aux exigences du monde du travail. Contenu : Job Interview, Debating , CV - Cover letter, Essay writing, Listening Comprehension, TOEIC training.

Stage en entreprise – 12 ECTS

Programme du M2 Math&AS (2023-2024) :

Le Master 2 Math&AS s'adresse aux étudiants titulaires d'un Master 1 Math&AS ou d'un Master 1 de Mathématiques avec des bases solides en statistique théorique, probabilités, langage python et sachant utiliser le logiciel R.

Les UEs dotées d'une astérisque sont partiellement mutualisées avec les étudiant.e.s du master ISADS. Des heures d'approfondissement théorique sont ajoutées pour les étudiant.e.s de la formation Math&AS.

Méthodes de régression (*) - 42h – 5 ECTS [Charles Tillier et intervenant extérieur]

Rappel sur le modèle linéaire; principe de pénalisation; régression LASSO, Ridge et Elasticnet; régression PLS; introduction aux régressions non-linéaires.

Apprentissage non supervisé (*) - 42h – 5 ECTS [Charles Tillier et intervenant extérieur]

Introduction à l'intelligence artificielle et au concept d'apprentissage; algorithme K-means et clustering; classification hiérarchique; réduction de la dimensionnalité ; introduction à la détection de nouveautés/ anomalies, one-class-classification.

Statistique non paramétrique (*) - 51h – 6 ECTS [Ester Mariucci, Emmanuel Rio et intervenant extérieur]

Le module est une introduction à la statistique non paramétrique. Les sujets traités seront l'estimation de la fonction de répartition, la construction des estimateurs à noyau et des estimateurs par projection de la fonction de densité, estimateurs adaptatifs, modèles de régression,

régions de confiance non-paramétriques, tests non paramétriques. Les objectifs majeurs de cette UE sont la compréhension des hypothèses sous-jacentes aux méthodes non paramétriques et l'acquisition d'une part de la capacité à appliquer les méthodes non paramétriques aux problèmes modernes d'analyse de données et d'autre part d'une expérience pratique dans la mise en œuvre des méthodes.

Techniques de prévision (*) - 36h – 4 ECTS [Agnès Grimaud]

Le but est d'effectuer des prévisions à partir d'une série chronologique, les différentes étapes sont : modélisation par un processus, estimation du modèle, validation et prévisions. Programme : Décomposition d'une série chronologique : tendance, saisonnalité et bruit. Estimation des différentes composantes. Étude du bruit : processus stationnaires, ARMA. Savoir les identifier. Processus ARIMA et SARIMA. Mise en pratique sur des exemples avec le logiciel R.

Statistique computationnelle – 30h – 4 ECTS [Agnès Grimaud]

Méthodes de Monte-Carlo, algorithme de Metropolis-Hastings, échantillonneur de Gibbs.

Algorithme EM.

Modèle de mélange.

Mise en application avec le logiciel R sur des exemples.

Statistique bayésienne (*) - 30h – 4 ECTS [Ester Mariucci et Mariane Pelletier]

Dans cette UE on commencera par étudier le modèle paramétrique bayésien : Calcul d'une loi a posteriori; Prédicteur et intervalles de prédiction; Estimateur de Bayes; Intervalles de confiance et intervalles de crédibilité. Enfin, nous nous intéresserons également à des modèles non paramétriques et on discutera le choix de la loi à priori dans différentes situations.

Ouverture sur la recherche – 30h – 4ECTS [Alexis Devulder et Julien Worms]

Cette UE se constitue de 2 mini-cours théoriques en statistique mathématique ou appliquée dont les contenus peuvent changer d'une année à l'autre.

Deep learning – 30h – 4 ECTS [Ester Mariucci et intervenant extérieur]

Dans ce cours nous explorerons la théorie et la pratique du deep learning, en donnant les bases mathématiques de ce sujet, mais toujours en assurant de savoir implémenter les algorithmes mentionnés. La première partie du cours se concentre sur les concepts de base des réseaux de neurones artificiels : couches linéaires suivies par des fonctions d'activation, comment elles généralisent les algorithmes linéaires (par exemple la régression logistique), comment on peut les entraîner grâce à l'algorithme du gradient. Ensuite, nous allons présenter les différentes couches qui sont le plus souvent utilisées dans des contextes spécifiques pour obtenir la plus grande précision dans la pratique : couches de convolution, les différents réseaux récurrents, les couches de normalisation et, si le temps les permet, les concepts d'attention et l'architecture « transformer ».

Séminaire invité entreprise – 18h – 2 ECTS [Ester Mariucci et intervenant extérieur]

Projet data challenge – 4 ECTS [Ester Mariucci et intervenant extérieur]

Dans ce cours de nature pratique, on verra comment résoudre un problème basé sur des données grâce au deep learning. On discutera d'abord comment obtenir un dataset et s'assurer de sa qualité pour pouvoir commencer un traitement basé sur le deep learning, et comment s'assurer une plateforme où entraîner des réseaux de neurones. Plusieurs challenges seront proposés aux étudiants, qui pourront choisir quel algorithme utiliser pour les résoudre.

Stage en entreprise – 18 ECTS