

Année universitaire 2024/2025

MIAGE Informatique Décisionnelle - 2e année de Master

Crédits ECTS : 60

LES OBJECTIFS DE LA FORMATION

Le parcours MIAGE Informatique Décisionnelle vise à doter les étudiants de compétences scientifiques pour améliorer le fonctionnement des organisations par l'utilisation des outils de l'informatique décisionnelle et l'application d'une approche scientifique. Il s'agit donc de former des professionnels ayant une double compétence en informatique.

Les objectifs de la formation :

- Avoir des compétences approfondies en informatique et informatique décisionnelle, (bases de données, data warehouse, nouvelles technologies de l'information)
- Avoir des compétences scientifiques pour comprendre les organisations et leurs impératifs stratégiques et améliorer leur fonctionnement
- Etre en capacité de modéliser et résoudre des problèmes de décision à travers la maîtrise de différents outils décisionnels

PRÉ-REQUIS OBLIGATOIRES

- Titulaires d'un diplôme BAC+4 (240 crédits ECTS) ou équivalent à Dauphine, d'une université, d'une école de commerce, d'ingénieurs ou d'un autre établissement de l'enseignement supérieur dans les domaines suivants : informatique, mathématiques, mathématiques appliquées aux sciences sociales, gestion (avec une appétence pour l'informatique et les systèmes d'information)

POURSUITE D'ÉTUDES

Les débouchés de la spécialité sont principalement orientés vers :

- Les sociétés de service en informatique notamment spécialisées dans le décisionnel,
- Les sociétés de conseil et bureau d'études,
- Les départements fonctionnels et d'études des entreprises et administrations.

PROGRAMME DE LA FORMATION

- Semestre 3
 - UE fondamentales S3
 - Entrepôts de données
 - Sécurité dans les SI
 - Outils d'optimisation pour les sciences des données et de la décision
 - Data Mining/Machine learning
 - Systèmes, Langages et Paradigmes pour le Big Data
 - UE optionnelles S3
 - Transparence des algorithmes pour la décision
 - Gouvernance des Systèmes d'Information

- Capitalisation et management des connaissances
- Qualité des données
- Aide à la Décision : construction d'indicateur
- Développement numérique Éco-responsable
- Semestre 4
 - UE fondamentales S4
 - Management de projet informatique
 - Droit, Données et Ethique
 - Le décisionnel en entreprise, Cas pratiques
 - Projet en Business Intelligence
 - Visualisation de données
 - Anglais
 - UE optionnelles S4
 - Large scale machine learning in Spark
 - Management des Organisations
 - Marketing digital
 - Négociation
 - Méthodes agiles d'ingénierie logicielle
 - Bloc stage
 - Stage

DESCRIPTION DE CHAQUE ENSEIGNEMENT

Aide à la Décision : construction d'indicateur

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Modélisation (définir les composants de l'indicateur) : techniques de travail collectif pour la structuration du problème : carte cognitives, focus groupe, etc.

Choix des fonctions d'agrégation : présentation de différents types d'agrégation (additive, ordinale). Analyse de leurs propriétés mathématiques

Mise ensemble des différents types de données venant de différentes sources avec différents niveaux de qualités (avec une approche de "théorie de mesurage" et analyse de signifiante)

Etude de cas avec une présentation d'un cas d'application industriel avec un intervenant industriel

Compétence à acquérir :

Fournir les fondements mathématiques et les outils de gestion nécessaires pour la conception et la mise en place des indicateurs composites.

Anglais

ECTS : 3

Volume horaire : 39

Capitalisation et management des connaissances

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Le cours de capitalisation sur les connaissances de l'entreprise est une introduction au Knowledge Management (KM). Il prodigue les concepts académiques fondamentaux aussi bien que des études de cas réelles permettant d'apprécier les enjeux

du management des connaissances dans les entreprises.

- Interprétations divergentes dans une entreprise étendue : contextes, cultures, intentions
- Cadre académique de l'ingénierie des connaissances, outils de cartographies cognitives
- Expertise, pouvoir, partage et rétention de connaissances
- Devenir des métiers, design thinking

Compétence à acquérir :

Concepts et enjeux du management des connaissances dans les entreprises

Data Mining/Machine learning

ECTS : 4

Volume horaire : 36

Description du contenu de l'enseignement :

L'évaluation se fait par examen et par un challenge Kaggle (kaggle.com)

- Introduction
- Objectifs et panorama du datamining et du Machine learning
- Méthodes non supervisées :
Réduction de dimensionnalité
Clustering :
K-means, CAH
Approches probabilistes : EA
Approches spectrales
Application à une segmentation marketing
Application au Text Mining
Règles d'association
- Méthodes supervisées :
Régression logistique
Arbre de Décision
Méthodes à Noyaux
Approches neuronales
Application au scoring

Compétence à acquérir :

Il s'agit d'initier les étudiants à l'apprentissage automatique (machine learning) et à la pratique de la fouille (data mining) et l'extraction de connaissances à partir des grandes masses de données. Il sera illustré par des cas concrets des exemples réalisés en session avec Python

Droit, Données et Ethique

ECTS : 3

Volume horaire : 36

Description du contenu de l'enseignement :

Introduction :

- Qu'est-ce que l'éthique (distinction, articulation entre éthique, droit, morale).
- De quelles données parle-t-on ?
- Pourquoi une éthique des données aujourd'hui ? Les principaux enjeux : le développement de l'Intelligence artificielle et la question du respect de dignité humaine, dépersonnalisation, déresponsabilisation, encadrement nécessaire de la gouvernance algorithmique (Cambridge Analytica)

Partie 1 : Les différentes formes d'encadrement de l'éthique des données

- L'encadrement collectif
Des obligations actuelles fixées par les Etats et l'UE (RGDP)
Des tentatives de régulation à l'échelle internationale
Une forme de responsabilité sociétale des entreprises (développement de code de conduite, d'outils de privacy by design ou by default)

- Une responsabilité individuelle : Du cas de conscience à la revendication du statut de lanceur d'alerte

Partie 2 : L'émergence de principes éthiques des données

- Les initiatives en cours :

Les recommandations de la CNIL : Vers la concrétisation d'un principe de loyauté des algorithmes et d'un principe de précaution en matière d'Intelligence Artificielle ?

Le projet de lignes directrices en matière d'éthique dans le domaine de l'IA développée par la Commission européenne
Vers une charte constitutionnelle française du numérique ?

- Les questionnements éthiques spécifiques :

La revendication d'un droit des robots

Vers la création de biens communs numériques ?

- Les bonnes pratiques sectorielles : Ethique des données dans le domaine de la finance, de l'éducation, de la santé etc.

Compétence à acquérir :

La profonde transformation numérique de nos sociétés rend nécessaire comme l'a encore récemment rappelé la CNIL de « former à l'éthique tous les maillons de la chaîne algorithmique, concepteurs, professionnels, citoyens ». La création de ce nouveau cours Données et Ethique s'inscrit dans cette optique. Il s'agit de permettre aux informaticiens de prendre conscience des implications éthiques et sociales du traitement des données qu'ils opèrent. Il s'agit aussi les informer sur l'évolution du cadre dans lesquels ils vont pouvoir ou devoir mettre en œuvre le traitement de leurs données tant à titre personnel que professionnel. L'éthique est une notion souvent difficile à cerner. Elle repose sur un ensemble de règles morales susceptibles d'orienter le comportement des individus voire des entreprises. Ces règles qu'ils s'imposent à eux-mêmes peuvent soit préfigurer, soit en réalité se superposer ou se confondre avec des règles juridiques contraignantes. Ainsi, l'éthique des données se développe à titre collectif qu'à titre individuel notamment avec la figure des lanceurs d'alerte. En outre, l'éthique des données peut varier en fonction de la catégorie de données en cause (données personnelles, non personnelles, données dite d'intérêt général etc.). Ces prémisses sont importantes pour comprendre les enjeux économiques et sociétaux autour de l'émergence actuelle des principes éthiques des données.

Développement numérique Éco-responsable

ECTS : 2

Volume horaire : 18

Entrepôts de données

ECTS : 4

Volume horaire : 36

Description du contenu de l'enseignement :

Introduction et définition d'un entrepôt de données, Architecture fonctionnelle, Modélisation conceptuelle, Alimentation, Stockage, Gestion et exploitation de l'entrepôt. Utilisation de l'ETL Talend en TP (Introduction Talend / Construction des premiers jobs (tri & jointure); Découverte d'autres composants Talend (filtre, condition, agrégation, gestion des erreurs); Modélisation et interaction avec la base de données; Cas pratique complet)

Compétence à acquérir :

Acquérir les notions de base relatives à l'acquisition, l'intégration, la modélisation et au traitement de données multidimensionnelles.

Gouvernance des Systèmes d'Information

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Description :

Introduction - concepts

Entreprise et système d'information

· Définition du système d'informations

· Les composantes clés du SI

Gouvernance : définition et enjeux

- La gouvernance : pourquoi, à quoi cela sert-il ?
- Gouvernance au niveau de l'entreprise
- Gouvernance des SI : pour leur pilotage interne
- SI et entreprise : alignement du SI sur la stratégie et la politique de l'entreprise 360° de la gouvernance
- La gouvernance à mettre en œuvre dépend de comment l'organisation entend piloter son activité et se contrôler
- La gouvernance dépend de l'organisation voire de l'environnement socio-économique
- Notion de « Compliance » : empilage des législations, réglementations, bonnes pratiques et procédures qui conditionnent la gouvernance

Les référentiels

Les grands standards : Itil, Cobit, CMMi, ...

- Orientations plus particulières de certains référentiels : développement logiciel, mode projet, exploitation, etc.

Focus sur CoBit

- Applications types, intérêts et limites
- Limites par rapport à la gouvernance de l'entreprise sur son SI

Modèle d'organisation : la matrice des fonctions du département informatique.

Focus sur ITIL

- Applications types intérêts et limites
- Limites par rapport à la gouvernance de l'entreprise sur son SI

Exemple de mise en place d'une démarche ITIL

Organisation de la DSI

Gouvernance des systèmes d'information-gouvernance d'entreprise

- Comment s'articulent la gouvernance interne du SI et le niveau de gouvernance du SI requis pour le pilotage de l'entreprise ?
- Quel recours pertinent aux référentiels métiers SI, quelles autres leviers et indicateurs pour répondre à toutes les exigences de gouvernance au niveau de l'entreprise
- Et si la Direction de l'entreprise n'envisage même pas le SI comme objet/sujet de la gouvernance d'entreprise ?

Les indicateurs de la Direction des Systèmes d'informations

Ce qui se mesure s'améliore

- Des indicateurs, pourquoi faire ?
- Quelques indicateurs incontournables

Tendances Cloud - Digital Marketing et BYOD ...

- Quel impact sur la gouvernance et le contrôle de risque ?

Etude de cas - Elaboration d'un Plan Directeur Informatique

- Le plan directeur
- Exemple d'élaboration du Plan Directeur Informatique d'une PME Français

Compétence à acquérir :

Présenter les principes de la gouvernance des SI

Large scale machine learning in Spark

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Les algorithmes d'apprentissage automatique sont de plus en plus utilisés de nos jours, et il existe une demande croissante d'algorithmes d'apprentissage qui sont capable de passer à l'échelle et de traiter des données massives.

Plutôt qu'offrir une introduction exhaustive à l'apprentissage automatique, ce cours se concentre sur les aspects typiques qui doivent être traités dans la conception d'algorithmes distribués pour l'apprentissage, et qui peuvent être exécutés sur les clusters Hadoop, afin d'analyser des grands jeux de données, en tenant compte l'adaptabilité à la croissance du volume des données ainsi que la robustesse en cas de pannes.

Le focus sera sur des algorithmes de réduction de dimension, de clustering et de classification, en tenant compte les aspects

suivants. Le premier est la conception d'algorithmes en s'appuyant sur des paradigmes basés sur map-reduce, à cette fin Spark sera utilisé. Le second aspect est l'analyse expérimentale des algorithmes implémentés en Spark, afin de tester leur capacité de passer à l'échelle (scalabilité). Le troisième aspect concerne l'étude et l'application de techniques d'optimisation afin de pallier le manque éventuel de scalabilité.

Bien que le cours se focalise sur certains algorithmes d'apprentissage, les techniques étudiées sont transversales et trouvent application dans un large éventail d'algorithmes d'apprentissage automatique. Pour certains des algorithmes étudiés le cours présentera des techniques pour une implémentation à partir de zéro en Spark-core, tandis que pour d'autres algorithmes Spark ML sera utilisé, et des pipelines de bout en bout seront conçus. Dans les deux cas, les algorithmes seront analysés et optimisés sur des jeux de données réels, sur un cluster Hadoop local, ainsi que sur un cluster sur le cloud Amazon WS.

Compétence à acquérir :

Savoir concevoir des applications efficaces pour l'apprentissage machine sur les données massives.

Mode de contrôle des connaissances :

CC et Projet

Le décisionnel en entreprise, Cas pratiques

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Donner une visibilité concrète de l'utilisation des applications décisionnelles en entreprise : Qu'apporte un logiciel décisionnel à un PDG, à un Directeur Marketing, à un Directeur commercial ? Quelles formes prennent les outils de l'informatique décisionnelle en entreprise ? Comment transformer un entrepôt de données en tableau de bord fonctionnel permettant d'expliquer le passé et d'anticiper l'avenir ? Ce cours vous donnera les réponses à ces questions à travers la présentation de projets concrets mis en place récemment dans de grands groupes.

Compétence à acquérir :

Acquérir les concepts de l'analyse décisionnelle

Comprendre l'apport d'un outil décisionnel aux différents acteurs de l'entreprise

Démarche et présentation de projets décisionnels concrets

Initiation sur un outil décisionnel

Mise en place d'un rapport à partir d'un jeu de données consolidées

Réalisation d'un cas pratique

Management de projet informatique

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

- Les différents processus qui accompagnent les méthodes de conduite de projet informatique
- Rôle de la maîtrise d'ouvrage et de la maîtrise d'œuvre
- Les activités de support
- Exposer et développer les méthodologies de conduite de projet informatique
- Conception agile : Lean, Scrum et Kanban
- Rétro-conception outillée d'un système existant : impact mapping, story mapping et backlog prêt
- Rétrospective méthodologique
- Définition et domaines d'usage d'un ERP
- Pourquoi un ERP ? Les enjeux Business
- Fonctionnement d'un ERP,
- Analyse et Stratégie de l'Entreprise
- Critères de choix d'un ERP,
- Organisation du projet et facteurs clés du succès
- Développement de la solution, test de la solution, mise en production de la solution
- Évaluation de la solution.

Prérequis : UML

Compétence à acquérir :

Développer les compétences nécessaires en matière de lancement, planification, et gestion de projets de manière efficace ainsi qu'à la consolidation des notions de conception et de modélisation et de leur intégration dans une démarche agile.

Management des Organisations

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Ce cours vise à faire comprendre le contexte actuel de management dans les organisations et favoriser l'intégration des futurs diplômés dans toute structure en s'appuyant sur des outils concrets.

Le cours se concentre ainsi sur le savoir-être qui est indispensable pour réussir dans le monde du travail.

Compétence à acquérir :

Comprendre le contexte de management des organisations

Préparer les futurs diplômés à la prise de fonction d'un poste en management

Acquérir des outils concrets pour animer une équipe

Introduire les notions de gestion du changement

Travailler la prise de parole en public et les présentations

Mode de contrôle des connaissances :

Contrôle continu 40%

Projet 30%

Examen final 30%

Marketing digital

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

1. Comprendre la révolution numérique

Le marketing à l'ère numérique

L'économie numérique

Le e-consommateur

2. Définir la stratégie digitale

Marché, concurrence, compétences, organisation

Proposition de valeur, mix-marketing

Business model

3. Mettre en œuvre la stratégie digitale

Les acteurs de l'écosystème publicitaire

Créer du trafic vers son site, les leviers du e-marketing

Ergonomie, design, expérience client, parcours clients, cross-canal

Créer et gérer son site internet et mobile

Gérer les moyens de paiement

Fidéliser ses clients internautes

Gérer le cycle de vie client

4. Le marketing des réseaux sociaux

Panorama des médias sociaux

Présence de l'entreprise sur les réseaux sociaux

Animer des campagnes sur les réseaux sociaux

5. Le marketing mobile

L'équipement et les usages des consommateurs

Les stratégies publicitaires sur le mobile
Les applications stores

6. Les tendances, les opportunités du digital
Le brand content, le storytelling, l'UGC
L'économie collaborative et participative
Les objets connectés, le big data

Compétence à acquérir :

Ce cours a pour objectif de vous faire acquérir les bases du marketing de l'internet et de la communication sur le web. Les différentes techniques permettant de recruter et de fidéliser le e-consommateur seront étudiés.

Méthodes agiles d'ingénierie logicielle

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

La nouvelle économie met à rude épreuve les équipes des projets informatiques. L'instabilité notoire des spécifications et le raccourcissement des délais de livraison imposent une amélioration drastique de la réactivité et de la productivité du développement, sans tolérer pour autant une quelconque négligence en matière de stabilité et d'évolutivité des applications déployées. Les méthodologies dites "agiles" tentent de répondre à ces exigences, en adoptant un investissement minimaliste et prônant des résultats tangibles et fréquents, par opposition aux méthodologies classiques qui - tout en imposant des organisations coûteuses - n'arrivent plus à prouver leur efficacité dans ces conditions extrêmement dynamiques. Ce cours propose un panorama des méthodes et techniques efficaces applicables dès les projets de petite taille, y compris pour assurer la traçabilité complète entre les besoins et les composants techniques en continue évolution. Les étudiants pourront les mettre en œuvre avantageusement, preuve comprise, aussi bien pour leurs projets pédagogiques qu'industriels, une fois dans l'entreprise.

- Introduction. Evolution des modèles du cycle de vie des projets. Manifeste agile: valeurs et principes.
- Offre des méthodologies agiles : XP, Scrum, FDD, Kanban, Lean, etc.
- eXtremeProgramming : cycle de négociation, métaphore & storytelling, développement piloté par les tests, refactoring permanent, etc.
- Gestion des exigences et de configuration logicielles. CMMi vs agile.
- Zoom sur le testing automatique fonctionnel : Behaviour Driven Développement utilisant des solutions en Gherkin avec Cucumber ou Behave.
- Modélisation, patterns, stratégies et styles architecturaux agiles: Design Patterns, Analysis Patterns, Metapatterns, Hexagonal, CQRS, DDD.
- Techniques encourageant l'agilité : RAD, AOP, Assemblage de composants, MDA, métriques & audits, EventModeling, Impact et StoryMapping.

Illustrations en Java et Python, tout au long du cours.

Compétence à acquérir :

- Découvrir, comprendre et appliquer les principes et les pratiques agiles les plus efficaces.
- Collaborer en équipe pour atteindre rapidement un objectif fonctionnel ou technique.
- Renforcer les compétences de spécification, conception, planification, communication, développement et tests automatisés.
- Fournir des résultats tangibles et fréquents, avec le sourire

Mode de contrôle des connaissances :

Projet évolutif avec soutenance validant les connaissances individuelles acquises.

Bibliographie, lectures recommandées :

Agile Ownership in a nutshell (vidéo) — <https://youtu.be/5021LHjX9EE>

Scrum and XP from the Trenches — <https://leanagiletraining.com/wp-content/uploads/2020/03/Scrum-and-XP-from-the-Trenches-2nd-edition.pdf>

The Pragmatic Programmer — <https://pragprog.com/titles/tpp20/the-pragmatic-programmer-20th-anniversary-edition/>

Le blog de Martin Fowler — <https://martinfowler.com>

Le manifeste agile — <https://agilemanifesto.org>

Extreme programming: a gentle introduction — <http://www.extremeprogramming.org>

The Scrum Guide — <https://www.scrum.org/resources/scrum-guide>

Négociation

ECTS : 3

Volume horaire : 30

Description du contenu de l'enseignement :

L'objectif du cours est une initiation théorique et pratique à la négociation.

La méthode du cours est de relier constamment les résultats des simulations auxquelles participent activement les étudiants avec les éléments fondamentaux de toutes négociations, objets d'acquisition du cours. Le séminaire est aussi rythmé par un ensemble de vidéos d'acteurs de la sphère publique et privée qui témoignent de leurs expériences et proposent des idées. Ces vidéos sont aussi des objets d'analyses et de débats. Le séminaire se termine par une grande négociation multilatérale qui permet la mise en pratique immédiate des théories et techniques vues, tout en permettant de prendre mieux conscience de ses forces et faiblesses individuelles en négociation. Un focus est proposé sur les conditions particulières d'une négociation de recrutement, notamment en ce qui concerne le droit du travail, les clauses d'un contrat et l'ensemble des informations nécessaires afin de renforcer sa capacité à négocier.

Compétence à acquérir :

- Définir les différents types de négociations et les stratégies afférentes.
- Connaître et mettre en pratique une vigilance particulière quant aux fondamentaux de la négociation : tensions compétition/coopération, assertivité/empathie et intérêts/valeurs.
- Reconnaître les tactiques liées à une négociation gagnant-perdant et celles associées au gagnant-gagnant.
- Appliquer une méthode de préparation à la négociation à travers les phases de Diagnostics, de Buts et de Stratégies (DBS).
- Acquérir un réflexe d'amélioration permanente de sa capacité à négocier, avec ses atouts et ses axes de progrès.

Bibliographie, lectures recommandées :

Bibliographie sommaire d'ouvrages récents sur la négociation :

Audebert, P., La négociation, Editions d'Organisation, 1999. *Ouvrage utile abordant de manière précise les nombreuses tactiques de négociations.*

Alter, N., Donner et prendre, la coopération en entreprise, Editions La Découverte, 2010. *Une approche sociologique de la difficulté des entreprises à gérer l'approche coopérative*

Axelrod R., Théorie du comportement coopératif : comment réussir dans un monde d'égoïstes, Ed. Odile Jacob, édition de poche, 2006. *Sans rejet des modélisations mathématiques, une lecture nécessaire...*

Bellenger L., La négociation, Ed. PUF « Que sais-je ? », 1984. *Des approches synthétiques des théories de la négociation.*

Bourque R., Thuderoz C., Sociologie de la négociation, Ed. La Découverte, 2002. *Un panorama intelligent et large pour un texte compact sur l'état des théories de la négociation.*

Cialdini R., Influence et manipulation, Comprendre et maîtriser les techniques de persuasion. *Ouvrage de référence construisant une théorie des arguments à partir de pratiques professionnelles.*

Combalbert L. et Mery M., Négociator, la référence de toutes les négociations, Ed. Dunod, 2019. *Un ouvrage de promotion de la méthode PACIFICAT réalisé par des praticiens et formateurs.*

Demoulin, S., Psychologie de la négociation, Ed. Mardaga, 2014. *Du contrat de travail au choix des vacances...*

Dupont C., La négociation. Conduite, théorie, applications, Ed Dalloz, 1994. *Un ouvrage de référence en la matière.*

Dupont C., La Négociation post-moderne, Bilan des connaissances, acquis et lacunes, perspectives, Ed. Publibook, 2006. *Le complément du précédent...*

Fisher R., Ury W. & Patton B., Comment réussir une négociation, Ed. Seuil, 1982. *Le texte de référence, facile d'accès, fondateur de la théorie de la négociation raisonnée.*

Fisher R., Shapiro D., Gérer ses émotions, Ed. Robert Laffont 2008. *Complément important de l'ouvrage précédent, mais peut se lire indépendamment.*

Joule R-V. & Beauvois JL., Petit traité de manipulation à l'usage des honnêtes gens, Ed. PUG, 2004. *Une actualisation de Cialdini par deux psycho-sociologues français combinant théories et pratiques expérimentales.*

Pekar Lempereur A. & Colson A., Méthode de négociation, Ed. Dunod, 2004. *Ouvrage pédagogique centré sur les méthodes : la négociation raisonnée à la française...*

Rojot, J., Négociation, Ed. Vuibert, 2006. *Ouvrage théorique et pratique d'un spécialiste des organisations à la fois solide et complet dans ses approches.*

Stimec, A., La Négociation, Ed. Dunod, 2011. *Une très bonne synthèse de la « méthode Harvard » (négociation raisonnée) et de ses développements.*

Schopenhauer A., L'art d'avoir toujours raison, Ed. Mille et une nuits, 2000. *Tout est dans le titre... mais attention, négociation n'est pas seulement persuasion ou rhétorique !*

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

L'optimisation mathématique est un outil très puissant pour modéliser la prise de la meilleure décision possible, que l'on retrouve dans de nombreuses problématiques industrielles. Le but de ce cours est de présenter des outils modernes d'optimisation sur les plans théorique, algorithmique et expérimental. On s'intéresse tout particulièrement aux problèmes issus de la science des données, qui portent actuellement la recherche et le développement en optimisation, ainsi qu'aux formulations classiques utilisées en recherche opérationnelle.

La première partie de ce cours pose les bases de l'optimisation, en termes de modélisation et d'étude théorique d'un problème : on y présente également les algorithmes efficaces en optimisation de grande taille, notamment dans le contexte de l'optimisation convexe. Dans la seconde partie du cours, on se concentre sur les problèmes typiques de sciences des données, pour lesquels on étudie à la fois des reformulations ainsi que les algorithmes récents développés spécifiquement pour ce contexte.

Compétence à acquérir :

Proposer une boîte à outils permettant de formuler correctement un problème d'optimisation, de choisir un algorithme adapté à sa résolution et de travailler sur une implémentation (personnelle ou propriétaire) d'une méthode donnée.

Bibliographie, lectures recommandées :

D. P. Bertsekas. *Nonlinear programming*, 2016.

G. C. Calafiore and L. El Ghaoui, *Optimization Models*, 2014.

S. J. Wright and B. Recht. *Optimization for Data Analysis*, 2022.

Projet en Business Intelligence

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Mener de bout en bout un projet d'informatique décisionnelle sur la plateforme SAS

Conceptualiser la chaîne de création de valeur de l'exploitation des données afin d'extraire des informations stratégiques.

Conceptualiser l'industrialisation de l'analytique des Big Data dans la mise en production de projets.

Analyser des données et construire des rapports avec un outil de Data Viz.

Réaliser des modélisations de Machine Learning et de recherche opérationnelle et les mettre en production.

Créer le lien entre la théorie et la pratique, l'informatique et l'analytique, les autres cours, afin de gérer une orchestration opérationnelle.

Présentation de différents points de vue des Big Data sur plusieurs secteurs d'activité, avec des retours d'expérience, afin de prendre du recul sur des définitions parfois contradictoires.

Prise en main de SAS Viya sur des petits cas simples afin d'avoir tous les éléments pour réaliser un projet transversale, intégrant Data Viz, Machine Learning et recherche opérationnelle.

Qualité des données

ECTS : 2

Volume horaire : 18

Description du contenu de l'enseignement :

La qualité de l'information a de graves répercussions sur l'efficacité et l'efficacités des organisations et des entreprises. Par exemple, dans les organisations privées, telles que les sociétés de marketing ou les banques, il n'est pas surprenant d'avoir plusieurs registres de clients, mis à jour par différentes procédures organisationnelles, ce qui entraîne des informations contradictoires et en double. Ce cours vise à présenter les problèmes liés à la qualité de données, et de présenter les solutions qui peuvent être utilisées pour les résoudre. En particulier, il a pour objectif d'examiner les trois points suivants :

- Dimensions de la qualité de données. Le terme « qualité de données » est assez vague et englobe plusieurs facettes. Dans cette partie, nous allons examiner les différents dimensions de la qualité de données, allant de l'exactitude, complétude, jusqu'au la consistance et la confiance.

- Algorithme et solutions théoriques. Dans cette partie nous allons examiner en détails quelques solutions théoriques proposées pour résoudre certains problèmes de qualité de données, en particulier celles liées à l'identification d'objet.

· Outils et solutions pratiques. Dans cette partie, nous allons utiliser des outils, et montrer comment ils peuvent être utilisés pour améliorer la qualité de données.

Compétence à acquérir :

Le cours a pour objectif d'introduire les caractéristiques principaux liés à la qualité des données ainsi que de présenter les algorithmes pour leur traitement.

Stage

ECTS : 6

Systèmes, Langages et Paradigmes pour le Big Data

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Le cours s'articule en trois parties.

Dans la première, l'accent est sur le paradigme MapReduce et le système Hadoop, avec un focus sur son système de fichiers HDFS. Le cours illustrera les mécanismes de base de Hadoop pour le support de l'exécution parallèle de 'dataflow' MapReduce sur des clusters de machines. Une attention particulière sera donnée aux aspects algorithmiques et d'optimisation de dataflow MapReduce.

La deuxième partie présentera des langages de requête et d'analyse de données caractérisés par des mécanismes de haut niveau et qui sont compilés sur MapReduce. Le focus sera sur les langages Pig Latin et Hive, des langages incluant des mécanismes à la SQL. Les techniques de compilation vers MapReduce seront présentées.

La troisième partie sera consacrée à des évolutions de Hadoop, et en particulier au système Spark et au langage de support Scala. Le focus sera sur l'architecture de Spark, la notion de RDD, l'évaluation lazy de transformations et actions sur des collections distribuées RDD.

Les notions apprises seront mises en pratique dans un projet où les étudiants devront concevoir un dataflow pour l'analyse de grands volumes de données. L'implémentation sera faite tant en MapReduce qu'en Spark, et une analyse expérimentale sera effectuée pour comparer les performances des deux implémentations.

Compétence à acquérir :

Apprentissage des aspects fondamentaux des technologies Big Data pour la gestion et analyse de données massives.

Sécurité dans les SI

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

La plupart des politiques de sécurité des systèmes d'information reposent sur une vision négativiste de l'action humaine : dissuader, bloquer, détecter, etc. Ce cours vise à installer les bases de la sécurité pour permettre à chacun de poser les bonnes questions et de devenir un acteur engagé de la politique de sécurité des systèmes d'information de son entreprise. Les fondamentaux aussi bien qu'une vision prospective et innovante de la sécurité des systèmes d'information sont abordés dans ce cours.

- Techniques d'attaque, de défense, de persuasion
- Contexte législatif et réglementaire (CNIL, RGPD, NIST, etc.)
- Éthique sécurisée, lanceurs d'alerte
- Menaces ciblant l'humain, ingénierie sociale et manipulation

Compétence à acquérir :

Initiation à la sécurité des systèmes d'information : menaces et recommandations

Transparence des algorithmes pour la décision

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Les algorithmes jouent un rôle de plus en plus fondamental dans tous les aspects de notre vie. Une des principales craintes soulevées par les algorithmes est leur manque de transparence de leur fonctionnement, ce qui pourrait amener à l'avènement d'une « société boîte noire ». De plus, cette transparence est un prérequis pour pouvoir analyser et expliquer les décisions prises par les utilisateurs ou des entreprises.

Le but de ce cours est de discuter et de sensibiliser les étudiants aux questions soulevées par la transparence et la responsabilité des algorithmes telles que par exemple :

Quelles sont les situations où la transparence et l'explicabilité sont nécessaires et comment les implémenter en pratique ?

Comment réconcilier le besoin des entreprises de protéger leur propriété intellectuelle avec le besoin des citoyens de pouvoir comprendre la logique derrière les algorithmes qu'ils utilisent ou qui produisent un effet sur leurs vies ?

Comment améliorer la transparence, la responsabilité ainsi que l'équité des décisions prises par des systèmes algorithmiques sans pour autant compromettre leur utilité.

On s'intéressera ici à une catégorie particulière d'algorithmes, ceux qui sont utilisés pour l'aide à la décision ou dans des traitements qui ont des incidences sur les comportements individuels, qui ont donc un effet normatif. A titre d'exemples, on peut citer :

Les algorithmes de classement, qui établissent des priorités, des recommandations : on pense évidemment aux algorithmes de présentation des résultats des moteurs de recherche, mais aussi à ceux qui sont utilisés pour classer les candidats à un poste, etc.

Les algorithmes de catégorisation, de classification, de profilage comme ceux qu'on met en œuvre pour détecter des profils de potentiels terroristes, de fraudeurs, etc.

Ce cours abordera les points suivants :

Rappels de notions du langage Python

Transparence, explicabilité et responsabilité des algorithmes : définitions, liens, avantages et limites.

Le bien-fondé de la transparence des algorithmes pour l'informatique décisionnelle. Aspects juridiques.

Cas d'étude 1 : Transparence des algorithmes dans le cadre des systèmes de recommandation : le cas du filtrage collaboratif

Cas d'étude 2 : Transparence des algorithmes dans le cadre de la classification : classification ordonnée et arbre de décision.

Cas d'étude 3 : Transparence des algorithmes dans le cadre de l'élaboration des classements ou palmarès : prise en compte simultanée des préférences des utilisateurs et des données statistiques

Mise en œuvre en langage python, à travers un mini-projet, d'un algorithme transparent pour la résolution d'un problème, à partir de données réelles disponibles sur des plateformes comme <https://www.data.gouv.fr> ou <https://www.kaggle.com/>

Compétence à acquérir :

Rappels de notions du langage Python

Transparence, explicabilité et responsabilité des algorithmes : définitions, liens, avantages et limites.

Le bien-fondé de la transparence des algorithmes pour l'informatique décisionnelle. Aspects juridiques.

Visualisation de données

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Représenter des données de manière visuelle permet de tirer parti des capacités humaines en termes de perception et de cognition pour identifier des motifs particuliers, des éléments singuliers ou encore des anomalies dans ces données. . La visualisation permet aux utilisateurs de systèmes d'information d'explorer leurs données de manière interactive, d'en avoir une vue d'ensemble, d'en extraire de l'information qui pourrait difficilement être obtenue à travers des processus d'analyse automatique (fouille, apprentissage), de formuler de nouvelles hypothèses qui pourront ensuite être vérifiées, par exemple, au moyen de tests statistiques. Au-delà de ces aspects exploratoires, la visualisation de données sert aussi d'aide à la prise de décisions, et de support à la communication entre individus des résultats d'analyses effectuées sur ces données. Le but de ce cours est de donner une vision d'ensemble du domaine, des principes fondamentaux de la perception visuelle humaine aux techniques de visualisation adaptées aux différentes structures de données (données multivariées, arbres, graphes, séries temporelles, etc.). Les travaux pratiques seront réalisés au moyen de la bibliothèque D3 (Data-Driven Documents).

Compétence à acquérir :

- Vue d'ensemble du domaine de la visualisation de données
 - Perception visuelle
 - Visualisation de données multivariées
 - Visualisation de structures arborescentes
 - Visualisation de graphes et réseaux
 - Visualisation de données temporelles
 - Visualisation multi-échelle
 - Systèmes et boîtes à outils pour la visualisation
-

Université Paris Dauphine - PSL - Place du Maréchal de Lattre de Tassigny - 75775 PARIS Cedex 16 - 07/04/2025