

Syllabus Anglais version anglaise

Name of module :

- English

Objectives and learning outcomes :

Level C1/C2 of the CEFRL (Common European Framework of Reference for Languages)

- Develop the skills necessary to students to succeed in their future professional lives in a variety of cultural contexts.
- Acquire the necessary linguistic autonomy and improve the specialized language allowing professional integration and communication of scientific expertise in an international context.

Competences

Speak fluently in front of an audience, using registers adapted to different contexts and different interlocutors.

Use a foreign language other than French easily: written and oral comprehension and expression:

- Understand a scientific or professional article written in English on a subject related to their field.
- Write a scientific or technical paper in English that is adapted, of high quality and that respects the standards and usages of the English-speaking scientific community.
- Interact orally in English in order to succeed in formal and informal exchanges during lectures, meetings or professional interviews.

Course Content

- Develop skills related to the understanding of scientific or professional publications written in English as well as skills necessary to understand oral scientific communications.
- Acquire expression tools in order to master an oral and/or written presentation and to tackle a critical discussion in the scientific field, (e.g. rhetoric, linguistic elements, pronunciation...).
- Acquire the elements of critical argumentation in an oral and/or written scientific publication
- Develop a broader reflection on their place, integration and influence as scientists in society, addressing issues of current affairs, ethics, integrity, etc... .

Prerequisite : Level B2 CEFRL

Key words

Project - Scientific English - Writing - Publishing - Communications - scientific critical thinking - interculturality

Assessment methodology : Continuous assessment:

Syllabus version anglaise

Module Name :

- Big Data

Objectives and learning outcomes : (800 caractères)

- This course is intended to describe advanced methods of statistical learning such as the support vector machines and neural networks.

Course content : (1200 caractères)

- Reminders on supervised learning : sample, regression and classification, population risk, empirical risk, Bayesian estimation, cross validation, VC dimension.
- Support vector machines in classification : separable SVMs, non-separable SVMs, kernel trick.
- Support vector machines in regression .
- Neural networks : Definitions and a short zoology, simple properties of the prediction, backpropagation, batch normalization, weight decay, dropout, Lipschitz constraint, properties of the optimization landscape, approximation power, VC dimension.

Bibliography : (320 caractères)

- The elements of Statistical Learning : data mining, Inference and prediction, T. Hastie, R. Tibshirani, J. Friedman
- « Deep Learning », I. Goodfellow, Y. Bengio, A. Courville, MIT press. Available at <https://www.deeplearningbook.org/>

Key words : (160 caractères)

Machine learning, Statistical Learning, support vector machines, neural networks, deep learning

Assessment methodology : (240 caractères)

- The assessment is based on one examination and the evaluation of the hands-on sessions.

Machine Learning, 3 ECTS

18h de cours, 12h de TP, 27h de projet,

LEARNING OBJECTIVES

The goal of this course is to study the methods of statistical learning both from a theoretical and from a practical point of view. It will allow to understand the algorithms underlying machine learning and their properties in terms of generalisation errors. The usual supervised and non supervised algorithms will be studied and applied to practical cases.

COURSE CONTENT

Principles of supervised learning

- Probabilistic notions
- Study of generalisation errors
- Uniform limit theorems on functional classes and complexity theory

Supervised classification

- Bayes classifier and its optimality
- Generalisation errors
- Classical algorithms : study and application of K-NN, Regression Logit, Classification trees

Unsupervised Learning

- Measure quantification and clustering stability
- K-Means and convergence proof
- Extension to kernel spaces
- Clustering algorithms linked to density (DBSCAN, Mean Shift)
- Autoencoders

Boosting and Bagging principles

- Theory
- Random Forests
- XGBoost

High dimension and functional spaces

- The curse of dimension
- Principles of sparse constrained estimation
- Extension of the methods to the functional case
- Spaces in sub-manifolds (LLE and spectral manifolds)

Syllabus version française

Nom du module :

- IA pour la décision

Objectifs : (800 caractères)

- L'objectif de ce cours est l'acquisition des concepts essentiels des domaines de l'intelligence artificielle formelle liés à la décision et à l'optimisation à partir de préférences et de connaissances complexes, imprécises ou incomplètes
- Compétences et connaissances acquises après validation du module :
 - o savoir modéliser un problème de décision sous connaissances incertaines et préférences complexes
 - o savoir évaluer sa complexité
 - o choisir et mettre en œuvre algorithmes pertinents et efficaces pour le résoudre
 - o familiarisation avec les problématiques de recherche en Intelligence Artificielle

Contenu : (1200 caractères)

- Ce cours abordera tout d'abord les différents modèles proposés par la théorie de la décision, qu'il s'agisse de décision sous incertitude (théorie de la décision, utilité qualitative, utilité non additive), de décision simultanée (théorie des jeux), ou de décision collective (vote, partage équitable).
- On étudiera la question, centrale en IA, des langages de représentation et les modes de raisonnement associés: langages de représentation de préférences (réseaux de préférences "Ceteris Paribus", réseaux d'utilité additive ou non, langages logiques), langages pour la décision sous incertitude (diagrammes d'influence, arbres de décision)
 - Afin de pouvoir étudier la richesse d'expressivité et le coût computationnel de ces langages, certaines notions avancées de la théorie de la complexité seront présentées (hiérarchie polynomiale, complexité des problèmes de recherche, complexité spatiale et compilation de connaissances) et replacées dans le contexte des problématiques abordées par ce cours.
 - Enfin, le cours traitera d'aspects algorithmiques du traitement de décision sous incertitude, d'une part en ce qui concerne l'apprentissage des modèles (apprentissage de réseau de fonctions de coût par exemple), d'autre part en ce qui concerne leur utilisation (optimisation de préférences par exemple).
 - Le cours sera complété par des séances d'exercices et la présentation de cas d'étude (planification de prises de vues satellitaires, configuration de produit, apprentissage de préférences, partage de coût)

Pré-requis : (240 caractères)

- Ce cours s'appuie sur des bases généralement abordées en M1, à savoir : (i) les bases de la théorie de la complexité, (ii) la programmation par contraintes et/ou la PNLE et/ou la logique propositionnelle, (iii) la modélisation et le raisonnement à partir de connaissances incertaines

Références bibliographiques : (320 caractères)

- Decision Making Process. Bouyssou, Dubois, Prade, Pirlot Editors ; Wiley. 2009
- A Course on Game Theory, M. Osborne et A. Rubinstein, MIT Press, 1994
- Computational Complexity, Christos H. Papadimitriou. Addison-Wesley 1994
- CPLEX : <https://www.ibm.com/fr-fr/products/ilog-cplex-optimization-studio>
- ToolBar2 : <https://informatique-mia.inrae.fr/prodmia/devinf/toulbar2>

Mots-clés : (160 caractères)

- Intelligence Artificielle ; Décision sous incertitude; Connaissances incomplètes ; Préférences ; Raisonnement et inférence ; Optimisation combinatoire ; Théorie de la complexité

Modalités d'évaluation : (240 caractères)

- Analyse de cas d'études
- Lecture et synthèse d'articles scientifiques
- Examen écrit
- plus formellement :
 - o CC : 30%, R16 (abs \rightarrow 0), reportée en 2nde session avec le même coefficient
 - o CT : 70%, écrit 2h
 - o 2nde session : écrit 2h 70%+ note projet 30%
-

Syllabus version anglaise

Module Name :

- AI for decision

Objectives and learning outcomes : (800 caractères)

- The aim of this course is the acquisition of the essential concepts of formal Artificial Intelligence applied to decision making under imprecise or incomplete knowledge and complex preferences.
- Skills and knowledge acquired after validation of the module :
 - o AI modelling of decision problems under uncertainty/complex preferences
 - o Evaluation of the complexity of such problems
 - o Definition and implementation of relevant and efficient algorithms
 - o Familiarization with AI research issues.

Course content : (1200 caractères)

- This course will first address the different models proposed by Decision Theory for problems of decision under uncertainty (probabilistic decision theory, qualitative utility, non additive utility), simultaneous decision (game theory), or collective decision (social choice, fair sharing).
- We will then study the question, central in AI, of representation languages and the associated modes of reasoning, and in particular the frameworks devoted to the representation of preferences ("Ceteris Paribus" networks, utility networks, logical languages) and decision under uncertainty (influence diagrams, decision trees).
- In order to allow the study of the expressiveness and the computational cost of these languages, some advanced notions of complexity theory will be presented (polynomial hierarchy, complexity of search problems, spatial complexity and compilation of knowledge) and related to the decision/optimization problems addressed in this course.
- Finally, the course will deal with the algorithmic aspects of decision making under uncertainty, on the one hand with regard to the acquisition of the models (learning of cost function models for example), and on the other hand with regard to their use (optimisation of preferences for example).
- The course will be complemented by exercise sessions and by presentation of case studies (satellite planning, product configuration, learning preferences, cost sharing)

Prerequisites : (240 caractères)

- This course requires notions generally covered in M1, namely: (i) basic complexity theory, (ii) constraint programming and/or MILP and/or propositional logic, (iii) modelling and reasoning under uncertain knowledge.

Bibliography : (320 caractères)

- Decision Making Process. Bouyssou, Dubois, Prade, Pirlot Editors; Wiley. 2009
- A Course on Game Theory. M. Osborne et A. Rubinstein. MIT Press, 1994
- Computational Complexity, Christos H. Papadimitriou. Addison-Wesley ,1994
- CPLEX : <https://www.ibm.com/fr-fr/products/ilog-cplex-optimization-studio>
- ToolBar2 : <https://informatique-mia.inrae.fr/prodmia/devinf/toulbar2>

Keywords : (160 caractères)

- Artificial intelligence; Decision under uncertainty; Incomplete knowledge; Preferences; Reasoning and inference; Combinatorial optimisation; Complexity theory

Assessment methods : (240 caractères)

- Analysis of case studies
- Reading and summary of scientific articles
- Written exam

Syllabus version française

Nom du module : Intelligence Artificielle 2 : Apprentissage automatique 2

Objectifs :

Cette matière présente un ensemble de méthodes pour modéliser et traiter des problèmes par apprentissage automatique au-delà de la classification simple : apprentissage de structures, problèmes avec supervision partielle, apprentissage par renforcement. L'accent sera mis également sur la méthodologie de validation, et abordera les limites des techniques actuelles du point de vue robustesse, explicabilité et éthique.

Les compétences acquises permettront d'identifier les problèmes tombant dans ce périmètre, d'appliquer les méthodes adaptées en respectant les contraintes méthodologiques et éthiques.

Contenu :

- ! Théorie de l'apprentissage : Notion d'apprenabilité, PAC, VC dimensions
- ! Apprentissage structuré : Modèles séquentiels (HMM, CRF), méthodes supervisées et leurs variantes structurées (SVM, Random Forest)
- ! Alternatives au supervisé: semi-supervisé, supervision faible, apprentissage par transfert, self-learning, apprentissage de représentations et applications à l'image
- ! Apprentissage multiple: apprentissage joint, multi-tâches, méthodes d'ensemble
- ! Apprentissage par renforcement (MDP, Q-learning, Deep Reinforcement Learning)
- ! Aspects pratiques et méthodologiques (plate-formes, démarche expérimentale, évaluations)
- ! Limites de l'apprentissage : biais, justice et éthique, robustesse, explicabilité et présentation des méthodes existantes.

Pré-requis :

Compétences en algèbre linéaire, probabilités et statistiques, et avoir suivi une introduction à l'apprentissage automatique

Références bibliographiques :

- Apprentissage artificiel, Deep learning, concepts et algorithmes. Antoine Cornuéjols, Laurent Miclet, Vincent Barraen, 3^e édition, 2018.
- Reinforcement Learning: An Introduction. Richard S. Sutton & Andrew G. Barto, 2^e édition. MIT Press, Cambridge, MA, 2018

Environnement : PyTorch ; Scikit learn

Mots-clés : Apprentissage automatique, apprentissage structuré, renforcement, limites et problèmes liés à l'apprentissage automatique

Modalités d'évaluation :

Le module sera évalué avec le rendu d'un projet de TP, et une évaluation écrite terminale.

CCTP 30%

CT 70%

Règle d'absence R16

Report de la note CCTP en 2e session (coeff 30%)

Syllabus version anglaise

Module Name : Artificial Intelligence 2 : Machine learning 2

Objectives and learning outcomes : (800 caractères)

This course presents methods for modeling and handling of problems with machine learning, beyond simple supervised classification. It will address structure learning, partial supervision settings, and reinforcement learning. The course will also focus on methodological and evaluation issues, and the current limitations of machine learning with respect to robustness, explainability and ethical issues.

Content :

- ! Theory of machine learning : learnability, PAC, VC dimensions
- ! Structured learning : sequential models (HMM, CRF), supervised methods and their structural counterparts (SVM, Random Forest)
- ! Beyond supervision : semi-supervised learning, weak supervision, transfer learning, representation learning and their application to image processing.
- ! Joint learning, multi-task learning, ensemble learning
- ! Reinforcement learning : Markov Decision Problems, Q-learning, Deep Reinforcement Learning
- ! Practical and methodological aspects : frameworks, experimental methodology, evaluations issues
- ! Limitations : bias, justice and ethics ; robustness, explainability and presentation of existing methods to address these issues.

Prerequisite : Linear algebra ; Probability and statistics ; Introduction to machine learning

Bibliography :

- Apprentissage artificiel, Deep learning, concepts et algorithmes. Antoine Cornuéjols, Laurent Miclet, Vincent Barraen, 3^e édition, 2018.
- Reinforcement Learning: An Introduction. Richard S. Sutton & Andrew G. Barto, 2^e édition. MIT Press, Cambridge, MA, 2018
- Framework : Pytorch ; Scikit learn

Keywords : machine learning ; structure learning ; reinforcement learning ; limitations of machine learning

Assessment methodology :

Students will be evaluated with a written exam (70%) and a practical assignment (30%)

Syllabus version française

Nom du module :

Analyse du son, des images et vision par ordinateur (ASIVO)

Objectifs :

Ce cours doit permettre aux étudiants d'effectuer les traitements classiques sur les images et le son (parole et musique). Il permet également de découvrir le domaine de la vision par ordinateur.

Compétences

- Connaître les techniques de base de traitement d'images.
- Apprendre à caractériser la parole et la musique : modèles acoustiques et prosodiques de l'état de l'art.
- Savoir réaliser un système de traitement automatique de la parole et de la musique : de la paramétrisation à la reconnaissance.
- Savoir retrouver le relief d'une scène à partir de deux images.
- Appréhender le mouvement des objets à partir d'une séquence d'images.

Contenu :

Cours-TD

Commentaire à enlever dans la version finale : Son et musique

- Modélisation statistique de la parole et de la musique.
- Reconnaissance automatique de la parole : paramétrisation et modélisation acoustique.
- Introduction aux modèles de langage.
- Spécificités de la musique.

Commentaire à enlever dans la version finale : Image

- Traitements point à point des images :
 - o Opérations algébriques et géométriques sur les images, méthodes d'interpolation.
 - o Transformée log, gamma, exponentielle.
 - o Notion d'histogramme, égalisation d'histogramme.
- Rehaussement d'images
 - o Opérations et filtrage spatial.
 - o Filtrage fréquentiel .
- Introduction à la restauration d'images
 - o Modélisation élémentaire et notion de PSF, de flou.
 - o Méthodes classiques de restauration (filtre inverse, filtre de Wiener, introduction au filtrage avancé).

Commentaire à enlever dans la version finale : Vision

- Outils pour la vision par ordinateur.
- Modélisation et calibrage géométriques d'une caméra.
- Stéréovision binoculaire : géométrie du capteur, mise en correspondance de pixels.
- Éléments de vision dynamique.

Aspect pratique

- Mise en place d'un système complet de reconnaissance automatique audio (mots clés ou sons clés).
- Développement d'applications de traitement d'images basiques.
- Découverte de la vision par ordinateur à travers des exemples de stéréovision de mouvement traitement des images à travers les techniques d'amélioration.

Pré-requis :

Introduction au traitement du signal, aux signaux sonores et aux images (M1)
Machine Learning 1 (M1)

Références bibliographiques :

- J.P. Haton, C. Cerisera, D. Fohr, Y. Laprie. Reconnaissance automatique de la parole. Edition Dunod, 2006.
- Gareth Loy, Musimathics, Volume 1 : The Mathematical Foundations of Music, The MIT Press, 2006.
- Gonzalez, Woods, Digital Image Processing, 2008.
- Trucco, Verri, Introductory Techniques for 3-D Computer Vision, 1998.

Mots-clés :

Paramétrisation de parole et musique, modélisation sonore, reconnaissance automatique de la parole, traitement d'histogramme, filtrage spatial, calibrage, stéréovision.

Modalités d'évaluation :

Contrôle terminal écrit de la partie théorique et contrôle continu des travaux pratiques.

Syllabus version anglaise

Module Name :

- Indiquer le nom du module

Objectives and learning outcomes :

- Donner ici une description textuelle des objectifs de formation du module en une dizaine de lignes.
- Identifier la (ou les) compétence(s) attestées(s) après validation du module
- Identifier la (ou les) connaissances acquises(s) après validation du module

Course content :

- Donner le plan détaillé du cours et le déroulé des travaux dirigés associés
- Donner le plan des TPs
- En cas de projet dans le module, décrire le travail attendu pendant le projet

Prerequisite :

- Décrire les connaissances/compétences requises pour suivre cet enseignement

Bibliography :

- Donner 2 ou 3 références bibliographiques de livres pouvant être utilisé comme support et/ou complément pour le cours. Donner au moins 1 référence en anglais.
- Si les enseignement (à travers les TPs par exemple) reposent sur des outils/environnements logiciels particuliers, donner, si pertinent, le lien vers un site web officiel de l'outil/environnement.

Key words :

- Liste des mots clés (entre 3 et 6 ?) synthétisant le contenu du cours.

Assessment methodology :

- Décrire succinctement les modalités de contrôle des connaissances et de suivi de l'acquisition de compétences qui sont envisagées pour ce module.

Informatique Graphique 2

Syllabus version française

Nom du module : Informatique Graphique 2

Objectifs :

Ce cours d'informatique graphique a pour objectifs de comprendre et maîtriser les modèles et algorithmes en informatique graphique et de savoir développer et mettre en œuvre l'ensemble des outils logiciels nécessaires au développement d'une application de rendu 3D physiquement réaliste.

Le cours théorique sera illustré par des résultats de recherche récents. Les travaux pratiques, fondés sur des logiciels open source orientés recherche, permettront de mettre en application certains aspects ciblés abordés en cours.

Compétences

Savoir choisir la méthode de simulation de l'éclairage adaptée à un problème donné

Savoir mettre en œuvre un estimateur Monte Carlo robuste et performant pour la simulation de l'éclairage

Savoir définir et entraîner un réseau de neurone pour l'estimation de densité

Connaissances

Fondements théoriques de l'informatique graphique

Estimateurs de Monte-Carlo et leurs propriétés

Algorithmes pour la simulation de l'éclairage

Contenu : L'unité d'enseignement comprends 18 cours/td (environ 1h30 de cours et 30 min de td par séance), deux master classes, 6 séances de tp pour introduire un projet (correspondant à une trentaine d'heure de travail étudiant)

A. Rendu (8 cours de 2h)

1. Modélisation de l'éclairage et de l'apparence
 - 1.a. De l'équation des transferts radiatifs à l'équation du rendu
 - 1.b. Modélisation et paramétrisation de l'apparence
2. Méthodes de résolution de l'équation des transferts radiatifs
 - 2.a. Discrétisation et éléments finis
 - 2.b. Estimateurs de Monte-Carlo
 - 2.c. Contrôle de la convergence des estimateurs de Monte Carlo
3. Modélisation de l'éclairage dans l'espace des chemins
 - 3.a. Construction de l'espace des chemins
 - 3.b. Estimateurs de Monte Carlo dans l'espace des chemins
 - 3.c. Échantillonnage par importance dans l'espace des chemins
4. Algorithmes de simulation de l'éclairage dans l'espace des chemins
 - 4.a. Algorithme du tracé de chemins
 - 4.b. Algorithme du tracé de chemins bidirectionnel

- 4.c. Algorithme de guidage de chemins
- 5. Apprentissage automatique pour la synthèse d'images
 - 5.a. Fonction de régression de la luminance
 - 5.b. Rendu neuronal
- 6. Apprentissage de densité et réduction de la variance
 - 6.a. Propriétés des séquences aléatoires
 - 6.b. Réseaux de normalisation
 - 6.c. Descente de gradient pour l'optimisation de l'échantillonnage
- 7. Modèles de caméra pour le calcul d'images de synthèse
 - 7.a. Modèle de caméra sténopé
 - 7.b. Modélisation des caméras par lentilles minces
- 8. Harmoniques sphériques pour la simulation de l'éclairage
 - 8.a. Les harmoniques sphériques réelles
 - 8.b. Projection et convolution efficace par harmoniques sphériques

B. Géométrie (10 cours de 2h)

C1 Subdivision

C2-3 Simplification QEM

C4-5 Géométrie différentielle et paramétrisation

C6 Laplacien discret (sur le maillage), lissage, calcul de fonction Laplacienne sur le maillage (déformation poignée)

C7 Acquisition de nuages de points 3D et recalage.

C8-9 Reconstruction : HRBF, poisson, MLS, Marching cubes et Ball pivoting

C9-10 Estimation de paramètres et reconnaissance

Master classes 2x2h

TP:

Rendu

Prise en main d'un path tracer open source

Ajout de fonctionnalité dans ce path tracer (échantillonnage adaptatif)

Géométrie

Subdivision/simplification

Le projet correspond à la finalisation du travail introduit en TP, en autonomie.

Pré-requis : Informatique graphique 1 (et ses pré-requis)

Mots-clés : Informatique graphique, rendu hors ligne, lancer de rayons, maillage, subdivision, simplification.

Modalités d'évaluation : Examen terminal écrit (2h rendu/2h géométrie)
Evaluation de projet

Références bibliographiques :

Physically Based Rendering: From Theory To Implementation, Matt Pharr, Wenzel Jakob, and Greg Humphreys, <http://www.pbr-book.org/>

Curves and Surfaces for CAGD: A Practical Guide (Fifth Edition) by G. Farin

Polygon Mesh Processing by M. Botsch, L. Kobbelt, M. Pauly, P. Alliez, B. Levy

Syllabus version française

Nom du module : Traitement Automatiques des Langues

Objectifs :

Le langage naturel est le moyen par excellence de communiquer des masses d'informations complexes et nuancées à des milliards de personnes. Le succès des techniques de dissémination comme le web témoignent de l'ubiquité et de l'utilité d'encoder des informations en langage naturel. Cependant, l'extraction des informations du web, ou autre source langagière, de manière automatisée, reste un défi théorique et technique.

L'objectif de cette UE est de donner les bases linguistiques, les modèles courants et des exemples d'applications computationnelles dans le domaine du Traitement Automatique du Langage. Trois grands axes de l'analyse linguistique et computationnelle seront abordés : la syntaxe, la sémantique, et le niveau pragmatique/discursif.

Contenu :

- Principaux niveaux d'analyse et concepts linguistiques (morphologie, syntaxe, sémantique, pragmatique et discours)
- Modèles syntaxiques et analyse syntaxique
- Représentations computationnelles du sens; sémantique distributionnelle
- Analyse de la structure du texte, analyse discursive
- Apprentissage de modèles et architectures courantes: modèles séquentiels, arbres, graphes
- Génération automatique de texte
- Plate-formes, et chaînes de traitement courantes
- Applications: extraction d'information, analyse de sentiments, question-réponse, traduction, résumé

Pré-requis :

Bases en apprentissage automatique (UE Apprentissage automatique 1 et 2).

Références bibliographiques :

-Delip Rao and Brian McMahan. Natural Language Processing with PyTorch. O'Reilly Media, 2019

- Yoav Goldberg. Neural Network Methods for Natural Language Processing. [Synthesis Lectures on Human Language Technologies](#). 2017

-Dan Jurafsky and James H. Martin. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall, Inc. 2019

<https://web.stanford.edu/~jurafsky/slp3/>

- Environnement : Pytorch

Mots-clés : Traitement automatique des langues, sources textuelles, apprentissage automatique pour l'extraction d'information à partir de textes.

Modalités d'évaluation : Le module sera évalué avec le rendu d'un projet de TP, et une évaluation écrite terminale.

CCTP 30%

CT 70%

Règle d'absence R16

Report de la note CCTP en 2e session (coeff 30%)

Module Name : Natural Language Processing

Objectives and learning outcomes :

Natural language is used for communication of complex and massive information between billions of humans. The popularity of the web as a means of disseminating information shows how useful natural language is to encode a large part of this information. Extracting information from the web or any linguistic source in a completely automated manner is however still a conceptual and technical challenge. This course aims at presenting the conceptual issues related to human language processing, current models and their applications to practical cases.

Course content :

- Main linguistic levels of analysis and concepts (morphology, syntax, semantics, pragmatics and discourse)
- Syntactic models and parsing
- Computational approaches of meaning; distributional semantics
- Analysis of text structure, discourse analysis and parsing
- Natural language generation
- Current models and learning architectures: sequences, trees, graphs
- Common practical frameworks and pipe-lines
- Applications: information extraction, sentiment analysis, question answering, summarization

Prerequisite : Machine learning

Bibliography :

-Delip Rao and Brian McMahan. Natural Language Processing with PyTorch. O'Reilly Media, 2019

- Yoav Goldberg. Neural Network Methods for Natural Language Processing.

[Synthesis Lectures on Human Language Technologies](#). 2017

-Dan Jurafsky and James H. Martin. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall, Inc. 2019

<https://web.stanford.edu/~jurafsky/slp3/>

- Environment : Pytorch

Key words :

NLP, text processing, machine learning for information extraction

Assessment methodology :

Evaluation based on a practical session (30%) and a final exam (70%)

Nom du module : Graph Data *Management and Mining*

Objectifs :

Résumé : Les données dites “graphe” sont extrêmement présentes de nos jours, que cela soit grâce à l’avènement du Web, des réseaux sociaux ou à la multiplicité des bases de données biologiques pour ne citer que quelques exemples de sources de données. Face à ces données complexes par nature, il est essentiel de fournir des outils et méthodes pour les représenter, les stocker, les interroger et en extraire de la connaissance. Ce cours s’attachera à fournir des outils théoriques et pratiques pour la bonne réalisation de ces tâches.

Compétences :

- Je connais et sais différencier les modèles de bases de données graphe
- Je connais la typologie des requêtes exécutées sur des graphes
- Je peux vérifier si je manipule un réseau dit “complexe”
- Je sais caractériser un noeud par son importance dans un réseau
- Je connais un ensemble de techniques non-supervisées qui peuvent être menées sur un réseau (représentation latente, calcul de communautés, extraction de motifs caractéristiques)
- Je sais mettre en place une démarche d’apprentissage supervisé sur un réseau (prédiction de liens)

Contenu : (1200 caractères)

- *Graph Data Management (1ECTS)*
 - o *Graph Database Models*
 - *Labeled Graphs*
 - *Hypergraph Model*
 - *Hypernode Model*
 - *Property Graph Model*
 - o *Querying Graph Databases*
 - *Classification of Graph Queries (adjacency, pattern matching, reachability, analytical)*
 - *Overview of existing Graph Query Language (TD1) (TP1)*
 - o *Review of Graph Database Systems*
- *Graph Data Mining (2ECTS)*
 - o *Complex network properties (TD2)*
 - o *Node-centric metrics (TD3) (TP2)*
 - o *Node embeddings (TP3)*
 - o *Pattern mining (TP3)*
 - o *Community Detection (TD4) (TP3)*
 - o *Link prediction (TD4) (TP4)*

Projet : le sujet et les données varieront chaque année et permettra d’approfondir un des aspects du cours

Pré-requis : (240 caractères)

- Théorie des graphes
- Introduction au machine learning

Références bibliographiques : (320 caractères)

- *Managing and Mining Graph* by Charu C. and Haixun Wang
- *Data Mining, The Textbook* by Aggarwal, Charu C.
- *Results on Graph Theory and Machine Learning in 2019* by Wei Gao

Mots-clés : (160 caractères)

- graph management, graph mining

Modalités d'évaluation : (240 caractères)

- Contrôle continu : projet
- Contrôle terminal : examen

Module Name : Graph Data *Management and Mining*

Objectives and learning outcomes :

- So-called "graph" data are extremely present nowadays, be it thanks to the advent of the Web, social networks or the multiplicity of biological databases, to name just a few examples of data sources. Faced with this inherently complex data, it is essential to provide tools and methods for representing, storing, querying and extracting knowledge from it. This course will focus on providing theoretical and practical tools for the proper performance of these tasks.
- Learning outcomes :
 - o I know and know how to distinguish between database models and graphs.
 - o I know the typology of queries executed on graphs
 - o I can check to see if I'm manipulating a so-called "complex" network
 - o I know how to characterize a node by its importance in a network...
 - o I know a set of unsupervised techniques that can be carried out on a network (latent representation, community computation, extraction of characteristic patterns)
 - o I know how to set up a supervised learning process on a network (link prediction)

Course content :

- *Graph Data Management (1ECTS)*
 - o *Graph Database Models*
 - *Labeled Graphs*
 - *Hypergraph Model*
 - *Hypernode Model*
 - *Property Graph Model*
 - o *Querying Graph Databases*
 - *Classification of Graph Queries (adjacency, pattern matching, reachability, analytical)*
 - *Overview of existing Graph Query Language (TD1) (TP1)*
 - o *Review of Graph Database Systems*
- *Graph Data Mining (2ECTS)*
 - o *Complex network properties (TD2)*
 - o *Node-centric metrics (TD3) (TP2)*
 - o *Node embeddings (TP3)*
 - o *Pattern mining (TP3)*
 - o *Community Detection (TD4) (TP3)*

Project : the subject matter and data will vary each year and will deepen one of the aspects of the course.

Prerequisite :

- Graph theory, machine learning

Bibliography :

- *Managing and Mining Graph* by Charu C. and Haixun Wang
- *Data Mining, The Textbook* by Aggarwal, Charu C.
- *Results on Graph Theory and Machine Learning in 2019* by Wei Gao

Key words :

graph management, graph mining

Assessment methodology :

- Continuous control: project
- Terminal control: examination

Modalités de contrôle de connaissances

Coefficient de la matière dans l'UE : 50%

Coefficient Epreuve1 : 30

Conservation pluri-annuelle note Epreuve 1 : Non

Type Epreuve 1 : CCTP - Contrôle continu TP

Paramétrage des absences Epreuve 1 : Règle 16 - ABI -> note 0, ABJ -> coef 0

Epreuve 1 en condition d'examen : Oui

Durée Epreuve 1 : 2h00

Nombre d'évaluations min. Epreuve 1 : 1

Nombre d'évaluations max. Epreuve 1 : 3

Natures évaluations Epreuve 1 : Compte rendu TP

Natures évaluations Epreuve 1 si distanciel : Compte rendu TP

1ère Session Epreuve 2

Coefficient Epreuve2 : 0

1ère Session Controle Terminal

Coefficient CT session 1 : 70

Nature évaluation CT session 1 : Ecrit

Durée CT session 1 : 2h00

Durée CT session 1 si distanciel : 2h00

Session de rattrapage

Organisation session de rattrapage : Oui

Report Note Epreuve 1 en seconde session : Non

Report Note Epreuve 2 en seconde session : Non

Coefficient CT session 2 : 100

Nature évaluation CT session 2 : Ecrit

Durée CT session 2 : 2h00

Syllabus version anglaise

Module Name :

- Design of experiments and uncertainty quantification.

Objectives and learning outcomes :

- This course is intended to introduce the concepts of design of numerical experiments and uncertainty quantification. The motivation of applications to computer models will be emphasized.

Course content :

- Introduction to Gaussian processes : definition, covariance functions, mean square smoothness.
- Gaussian conditioning, prediction, predictive intervals.
- Maximum likelihood.
- ANOVA decomposition of a function.
- Sobol sensitivity indices.
- Estimation with the Pick and Freeze method.
- Reproducing kernel hilbert spaces (RKHS) .
- Kernel trick.
- Kernel methods.

Bibliography :

- Gaussian processes for machine learning (Book, Rasmussen and Williams, 2006).
- Sensitivity analysis in practice: a guide to assessing scientific models. (Book, Saltelli , Tarantola , Campolongo and Ratto, 2004).
- Learning with kernels: support vector machines, regularization, optimization, and beyond. (Book, Schölkopf, Smola, 2002).

Key words :

Computer experiments, Gaussian processes, uncertainty quantification, RKHS, kernel methods.

Assessment methodology :

- The assessment is based on one examination and the evaluation of the hands-on sessions.

Computer science, 3 ECTS

18h de cours, 12h de TP, 27h de projet,

LEARNING OBJECTIVES

The goal of this course is to introduce students to structured languages (Fortran, C), object oriented programming (C++) with a particular stress put on the High Performance Computing on modern architectures. Parallel computing in the multi-core environment, on distributed memory systems and finally GPU computing will be presented.

COURSE CONTENT

Object oriented programming with C++
High performance computing
Parallel architectures
OpenMP parallelization on shared memory systems
MPI parallelization on distributed memory systems
GPU computing (CUDA)
Hybrid parallelization – combining MPI, OpenMP and GPU computing together.

Labs in C++.

PREREQUISITES

Programming and Algorithmics

KEYWORDS

C++, OpenMP, MPI, High Performance Computing

BIBLIOGRAPHY

- Introduction to High Performance Scientific Computing, by V. Eijkhout et al. (Creative Commons, 2015)