

USSI6G - Informatique S3

Présentation

Programme

Contenu

Positionnement : Semestre 3	UE : Informatique	ECTS : 7
		Nombre d'heures : 143,5h
		Modalité : Présentiel (100%)

L'UE Informatique est constituée de 4 éléments constitutifs d'UE :

- Cloud – Duplication – Hadoop – Map reduce
- Développement logiciel R
- Programmation avancée
- TP Informatique – Python (3)

La note finale associée à cette UE est calculée en pondérant les notes des 4 ECUE avec les coefficients associés.

ECUE 1 : Cloud – Duplication – Hadoop – Map reduce (Coef : 2) – 59,5h (Cours, TD,TP) - 15h (Estimation temps de travail personnel)

Objectifs : L'objectif de ce cours et ses applications est d'acquérir les connaissances générales sur les systèmes distribués et les intergiciels pour l'urbanisation des systèmes d'information distribués, pour la conception d'applications distribuées large échelle.

Compréhension des mécanismes sous-jacents du Cloud Computing.

Être capable de différencier les avantages et inconvénients des systèmes distribués vs les systèmes « centralisés » .

Contenu :

1. Systèmes distribués : principes théoriques de fonctionnement
 - a. Rappels sur les processus dans les systèmes centralisés : mécanismes de synchronisation, exclusion mutuelle, multi- threads, ordonnancement des processus, mémoire.
 - b. Exigences des systèmes distribués et principes généraux.
 - c. Gestion du temps : horloges logiques, ordre causal.
 - d. Algorithmes utilisés dans les systèmes distribués : exclusion, synchronisation, élection, terminaison, coupure, état global, ... (Quelques exemples)
 - e. Gestion partagée des données : désignation/ transparence, cohérence mémoire, localisation, P2P, ...
2. Architectures des systèmes distribués
 - a. Infrastructures : data centers, multi-coeurs, clustering, architectures mémoires NUMA, UMA, ...

Mis à jour le 21-02-2025



Code : USSI6G

Unité spécifique de type mixte
7 crédits

Responsabilité nationale :
EPN06 - Mathématique et statistique / 1

Contact national :

Cnam Nouvelle Aquitaine
2 Avenue Gustave Eiffel
Téléport 2
86960 Chasseneuil Futuroscope
05 49 49 61 20

naq_info@lecnam.net

- b. Méthodes de virtualisation et de containerisation.
 - c. Exemple de l'écosystème Hadoop et de HDFS
3. Tolérance aux pannes dans les systèmes distribués :
 - a. Définitions : pannes, erreurs, , sureté de fonctionnement, types de pannes, pannes byzantines, mécanismes de réplication : redondances matérielles, mécanismes de recouvrement, ...
 - b. Introductions aux bases NoSQL et à MongoDB
 - c. Etude de cas et exemple de tolérance aux pannes avec MongoDB et Hadoop-HDFS
4. Architectures de type Cloud computing
 - a. Notions de PaaS, SaaS ...
 - b. Exemples Google, Microsoft, AWS.
5. Traitements parallèles de masse - Map -reduce
 - a. Paradigme Map reduce
 - b. Exemple de programmation map reduce avec MongoDB et Hadoop HDFS -YARN
 - c. Evolution du modèle Map reduce – Exemple de Apache Spark

Compétences visées : A l'issue du cours l'étudiant est capable de :

1. utiliser Map reduce sur MongoDB,
2. gérer la tolérance aux pannes dans les systèmes distribués,
3. passer à l'échelle sur le Cloud (AWS).

Modalités d'évaluation : Contrôle continu (rendus de TP)

ECUE 2 : Développement logiciel R (Coef : 1) – 31,5h (Cours, TP, Projet) - 10h (Estimation temps de travail personnel)

Objectifs : Approfondir les connaissances sur le langage R via R Studio

Contenu :

1. Présentation de R Studio
2. Fondamentaux du langage R
 1. Vecteurs et facteurs
 2. Tableaux de données
 3. Matrices
 4. Listes
3. Manipulation de fichiers
 1. Excel et CSV, Tidy Data
 2. Texte, analyse textuelle, expressions régulières
 3. XML et langage XPath
 4. Fichiers de données géographiques
4. Découvertes des bibliothèques associées

L'accent est mis sur la découverte de ces sujets via des TP.

Compétences visées : A l'issue de la formation, l'apprenant sera capable de :

- utiliser R pour programmer et analyser des données,

- savoir les bibliothèques à utiliser.

Modalités d'évaluation : Contrôle continu sous forme de TP

ECUE 3 : Programmation avancée – (Coef : 1) – 31,5h (Cours, TP, Projet) - 10h (Estimation temps de travail personnel)

Objectifs : Ce cours est la suite de l'ECUE Informatique générale – Programmation algorithmique de première année.

Apprendre à mettre en place un projet JAVA complexe avec la gestion de dépendances correspondante.

Maîtriser les concepts avancés de la programmation et conception objet en JAVA.

Contenu :

1. Les bases
2. Les entrées/sorties
3. Gestion du cycle de vie des applications
4. Tests unitaires
5. Réflexion
6. Concurrency
7. Modélisation OO

Compétences visées :

- Maîtriser le langage Java et les concepts objets avancés cités dans le programme.
- Pouvoir analyser et faire évoluer un projet JAVA existant.

Modalités d'évaluation : Contrôle continu. Tous les TP (un par sujet) sont notés avec un coefficient différent selon la complexité.

ECUE 4 : TP Informatique – Python (3) – (Coef : 1) – 21h (TP) – 10h (Estimation temps de travail personnel)

Objectifs :

- Présentation des besoins d'intelligence artificielle dans le monde privé.
- Gérer un projet en Python sur un sujet d'IA.

Contenu :

1. Proposition de projet : un projet doit traiter au moins une des thématiques suivantes
 1. Machine Learning
 2. Deep Learning N
 3. LP (Natural Language Processing)
 4. Computer Vision

2. Identification des objectifs et problématiques correspondantes
3. Rédaction d'un plan de travail pendant 5 séances
 1. Préparation de dataset bruts
 2. Preprocessing de dataset
 3. Développement des méthodes d'IA
 4. Expérimentation

Compétences visées : Maitriser le langage Python et ses utilisations dans l'IA

Modalités d'évaluation : Travail en groupe de 3 ou 4 étudiants. Soutenance + démonstration (dernière séance). Notation du code et du rapport.

Modalités de validation

- Contrôle continu
- Projet(s)
- Mémoire
- Examen final