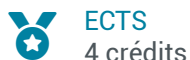


Outils Informatiques et Professionnels 3



Présentation

Description

Programme :

- I. Outils Professionnels :
Séminaires assurés par des intervenants extérieurs issus du monde de l'entreprise.
- II. Outils Informatiques :
 - A) HPC - Introduction au calcul parallèle
 - a) Calcul en mémoire partagée avec OpenMP
 - 1. Architecture à mémoire partagée
 - 2. Éléments de base
 - 3. Premier programme avec OpenMP
 - b) Introduction à la programmation parallèle en utilisant MPI (1)
 - 1. Architecture à mémoire distribuée
 - 2. Éléments de base
 - 3. Premier programme avec MPI
 - c) Introduction à la programmation parallèle en utilisant MPI (2)
 - 1. Communication de base en MPI
 - 2. Exemples d'applications MPI
 - 3. Débogage d'un programme parallèle
 - d) Introduction au calcul sur GPU avec Cuda
 - 1. Premier programme Cuda
 - 2. Exemples d'applications
 - e) Soumission et gestion des calculs sur un cluster
 - B) R pour l'analyse de données
 - a) Tidyverse pour l'analyse de données
 - b) Manipulation des données avec dplyr, tidyr et readr
 - 1. Préparation des données (filtrage, sélection, nettoyage)
 - 2. Agrégation, jointure et tri des données
 - 3. Calcul des statistiques descriptives d'un groupe de données
 - c) Visualisation des données avec ggplot2

1. Visualisation de données et personnalisation des graphiques
 2. Création de cartes géospatiales en utilisant des données de géolocalisation
- d) Shiny pour la réalisation d'interfaces statistiques dynamiques
1. Création d'applications web interactives
 2. Déploiement de modèles statistiques partagés

Logiciels :

Les travaux pratiques de ce module seront réalisés avec les langages C++ (sur un système Linux) et R

Objectifs

1. Comprendre le fonctionnement des ressources de calcul et connaître leurs limites
2. Identifier les algorithmes de calcul dont l'efficacité peut être améliorée par l'utilisation de processeurs multicœurs ou GPU
3. Analyser la performance et les résultats de codes de calcul
4. Manipuler des bases de données avec la librairie R tidyverse de manière efficace et intuitive
5. Déployer des modèles statistiques partagés avec R Shiny
6. Réaliser des interfaces statistiques dynamiques avec R Shiny

Pré-requis obligatoires

Cours de C++ du semestre 2 du Master 1 ou équivalent.

Bibliographie

1. <https://link.springer.com/book/10.1007/978-3-319-21903-5> Introduction to HPC with MPI for Data Science, Frank Nielsen. Springer. 2016.
2. <https://carpentries-incubator.github.io/hpc-intro/> Introduction to High-Performance Computing, carpentries incubator, 2023.
3. <https://www.tidyverse.org/Documentation> tidyverse
4. <https://shiny.posit.co/Documentation> shiny