Python avancé : programmation scientifique

Date : **10 au 13 juin 2024**

Lieu : Pythagore F.D-Paris-9ème

Nombre de stagiaires : 10

Objectifs

- Maitriser les bibliothèques de programmation scientifique en Python
- Explorer et analyser des données
- Programmer des algorithmes et calculer de manière efficace en Python

Public visé

Développeurs en programmation scientifique employés par le CNRS

Modalités pédagogiques

Alternance entre présentations des fonctionnalités des bibliothèques et de leurs objets. Les illustrations seront proposées sur des cas concrets. De nombreux exercices permettront les mises en pratique.

Programme

1/ Python avancé (1 jour)

- Rappels concepts fondamentaux Python (architecture objet du langage, type de données, structure de données)
- Bonnes pratiques / Design de code : PEP8, structuration script et modules, POO, notion de packaging
- Fonctionnalités avancées de Python : gestionnaire de contexte, liste par compréhension, itérateur, générateur, …
- Évaluation de performances : outils de profilage (timing et mémoire)
- Limite des performances du langage Python natif

2/ Manipulation de tableaux et opérations mathématiques avec NumPy (1 jour)

- Structure de donnée : les numpy array (différences avec les listes python, définition, type de données, allocation mémoire, shape, reshape)
- Création de tableau (random, à partir de données existantes, …)
- Accès aux éléments du tableau : indexation simple, slicing & striding, indexation avancée (par tableaux d'entiers ou de booléens)
- Opérations sur les tableaux : opérations basiques, broadcasting, méthodes spécifiques (min, max, mean, std…), travailler à partir de formules mathématiques
- Entrées/sorties avec Numpy
- Explication de l'optimisation des opérations sur les tableaux avec Numpy : Vectorisation, Allocation mémoire, binding C++, compilation sur architecture dédiée, optimisation des ressources du processeur (CPU/SIMD optimisation)
- Bannir les boucles for
- Comparaison de performances avec Python natif
- Profilage / Timing

3/ Calcul scientifique avec SciPy (0.5 jour)

- Présentation générale de librairie
- Quelques exemples : Fit de courbe, Algèbre linéaire, Fourrier, Traitement du signal, …

4/ Exploration et analyse de données avec Pandas (1 jour)

- Structure de données : DataFrame, Serie
- Fonctionnalités essentielles : Attributs, statistiques descriptives

Type de données, …

- Indexation et sélection de données : Basique, accès par attribut, slicing, selection par label, par position, par fonction callable, methodes de selection (isin, where), selection par masque
- Opérations mathématiques sur les DataFrames : opérations accélérées sur les données, application de fonction mathématiques
- Opérations sur les DataFrames : concaténation, fusion, regroupement
- Travailler avec des données textuelles
- Travailler avec des données manquantes
- Travailler avec des Série temporelles
- Outils d'entrée/sortie (csv, HDF5, …)

5/ Accélération de code Python (0,5 jour)

- Présentation générale et rapide des outils existant pour accélérer les codes python (Numba, Cython, Pythran)
- Passage à l'échelle et parallélisme avec Dask

Conditions d'inscription

Date limite d'inscription : 10/05/2024

Inscription : https://formation.ifsem.cnrs.fr/

Renseignements :
ifsem-formation.contact@cnrs.fr