Computing Foundations of Data Sciences

Titulaire
Dimitrios SACHARIDIS (Coordonnateur)

Mnémonique du cours
INFO-H600

Crédits ECTS
5 crédits

Langue(s) d'enseignement
Anglais

Période du cours
Premier quadrimestre

Campus
Solbosch

Contenu du cours

The course has three parts:

- **Introduction to Python**. The basic programming concepts are introduced using the Python languages.
- **Data Science with Python**. An introduction to data science concepts is made using Python modules. The focus is on manipulating and analyzing numerical and tabular data using numpy and pandas modules. The introduction also covers handling missing values, dealing with hierarchical data like json, and basic data visualization methods.
- **Introduction to Big Data Management**. This part introduces the challenges of big data, discusses computational architectures and programming abstractions like distributed/ parallel computing and map-reduce. These concepts are discussed in the context of the python framework Dask, and Apache Spark using python as the interface (pySpark).

Objectifs (et/ou acquis d'apprentissages spécifiques)

Upon completion of the course, the student is expected to:

- Obtain a good level of programming with Python.
- Understand the basic data science concepts, and be able to manage and analyze structured data.
- Be able to perform data science tasks using Python.
- Understand the challenges of working with Big Data, and understand the techniques that can be used to address these challenges.
- Be able to perform big data management in Python using the Dask and Apache Spark frameworks.

Pré-requis et co-requis

Connaissances et compétences pré-requises

A basic understanding of programming is assumed. Familiarity with Python is preferable, but the course will provide all necessary material.

Méthodes d'enseignement et activités d'apprentissages

The teaching of each part consists of a brief theoretical introduction and is followed by practical sessions.

There are two assignments that correspond to the topics of the first two parts.

There is a project, to be done in groups, that corresponds to the topic of the third part.

Support(s) de cours

Université virtuelle

Autres renseignements

Lieu(x) d'enseignement

Solbosch

Contact(s)

Dimitris Sacharidis dimitris.sacharidis@ulb.be

Méthode(s) d'évaluation

Projet et Travail pratique

Méthode(s) d'évaluation (complément)

Students are evaluated based on two assignments and one project.

Each assignment contributes 25% of the total mark (i.e., 5 points).

The project contributes 50% of the total mark (i.e., 10 points).

There is no final exam.

Langue(s) d'évaluation principale(s)

Anglais

Programmes

Programmes proposant ce cours à l'école polytechnique de Bruxelles

MS-BGDA | Master de spécialisation en science des données, Big data | bloc U

Programmes proposant ce cours à la Solvay Brussels School of Economics and Management

MS-BGDA | Master de spécialisation en science des données, Big data | bloc U

Programmes proposant ce cours à la faculté des Sciences

MS-BGDA | Master de spécialisation en science des données, Big data | bloc U