

Société organisatrice

Depuis 2001, la SFdS organise des ateliers en science des données destinés avant tout aux ingénieurs, chercheurs, doctorants et post-doctorants. Ces Ateliers sont originaux car ils abordent des sujets porteurs et/ou nouveaux.

Cette année et compte-tenu d'une demande massive, la SFdS vous propose un atelier exceptionnel, portant sur **les méthodes d'apprentissage par (agrégation d') arbres de décision**. Ces dernières se révèlent très performantes pour la prévision en se classant régulièrement parmi les meilleurs algorithmes dans les nombreuses compétitions de machine learning. L'atelier débutera par une journée de **mise à niveau en Machine Learning**. La deuxième journée sera consacrée au **fondamentaux des méthodes d'apprentissage par arbres de décision incluant les forêts aléatoires**. Enfin, la troisième journée proposera une **introduction au boosting** appliqué aux méthodes d'apprentissage par arbres de décision. Les différents procédés d'agrégation d'arbres (forêts aléatoires et boosting) seront comparés en termes de propriétés statistiques et de performance sur données réelles. Chaque partie du cours sera illustrée par **des travaux pratiques** sous R.

L'atelier sera animé par **Christine Keribin** (Univ. Paris-Saclay), **Robin Genuer** (Université de Bordeaux) et **Laurent Rouvière** (Université Rennes 2).

**Les 30 novembre, 1 & 2
décembre 2022
9h00-12h30 et 14h00-17h30**

Afin d'assurer une plus grande convivialité à la formation et de favoriser les échanges entre les participants, **le nombre de places est limité à 20**.

Date limite d'inscription : **23 novembre 2022**

Organisation

Le cours se déroulera à l'IHP, 11 rue Pierre et Marie Curie 75005 Paris.

Accès : RER ligne B station : Luxembourg

Bus : 21, 27, 38, 84, 85, 89

Plan de la formation

Mercredi 30 novembre 2022

par **Christine Keribin**

Machine learning

- ✓ Positionner IA/ML/apprentissage machine/statistique
- ✓ Les bases de l'apprentissage / risque / évaluation de la performance
- ✓ Panorama des méthodes / de la régression linéaire aux réseaux de neurones
- ✓ Quels enjeux actuels ? taille et variété des données / grande dimension / données manquantes/ biais et fairness / apprentissage fédéré / respect de la vie privée
- ✓ Un cas d'étude d'apprentissage supervisé avec le "nouveau R" : ggplot2, dyplr et tidyverse

Jeudi 1^{er} décembre 2022

par **Robin Genuer**

Arbres CART et Forêts aléatoires

- ✓ Arbres CART : principe général, construction de l'arbre maximal et élagage
- ✓ Application sur données réelles avec R et outils disponibles dans le package rpart
- ✓ Forêts aléatoires : principe général, paramétrage et importance des variables
- ✓ Application sur données réelles avec R et aides à l'interprétation des résultats avec le package randomForest

Vendredi 2 décembre 2022

par **Laurent Rouvière**

Introduction au boosting

- ✓ Historique et premiers algorithmes boosting
- ✓ Minimisation de risque convexifié : gradient boosting
- ✓ Calibration des algorithmes boosting
- ✓ Extreme gradient boosting : xgboost
- ✓ Applications sur R

Bulletin d'inscription

A retourner au secrétariat de la SFdS, IHP, 11 rue Pierre et Marie Curie, 75231 Paris cedex 05.

Votre inscription ne sera définitive qu' à réception du règlement ou d' un bon de commande et la confirmation de celle-ci vous sera demandée le 1^{er} jour de la formation.

NOM.....Prénom.....

Fonction.....Organisme.....

Adresse complète.....

.....

Code Postal.....Ville.....

Pays.....Tél.....Fax.....E-mail.....

souhaite m' inscrire au cours : **Mise à niveau en machine learning & Méthodes d' apprentissage par arbres**

de décision

organisé les **30 novembre, 1^{er} et 2 décembre 2022** à l' IHP.

verse m'engage à verser par chèque par virement par bon de commande :

2 jours de formation : 700 € 400 € 350 € 150 € selon ma catégorie (cf. Droits d' inscription).

Choix des 2 jours : Mercredi 30 novembre Jeudi 1 décembre Vendredi 2 décembre

3 jours de formation : 1000 € 550 € 500 € 225 € selon ma catégorie (cf. Droits d' inscription).

DATE et SIGNATURE :

Droits d'inscription

Tarif général

- 3 jours de formation 1000 €*
2 jours de formation au choix 700 €*

Salariés (universités et des établissements publics d'enseignement ou de recherche)

- 3 jours de formation 550€*
2 jours de formation au choix 400€*

Doctorants

- 3 jours de formation 500€*
2 jours de formation au choix 350€*

Étudiants non-salariés (ex : master)

- 3 jours de formation 225€*
2 jours de formation au choix 150€*

*Pour des raisons juridiques, les ateliers doivent s'adresser uniquement aux membres de la SFdS. Les non-membres de l'association doivent s'acquitter de leur adhésion pour l'année en cours pour pouvoir participer aux formations. Merci dans ce cas de prendre contact avec le secrétariat.

Par suite de votre inscription, vous recevrez une confirmation par voie électronique : ce document sera à conserver et à présenter le premier jour de la formation.

Droits d'inscription

Les droits d'inscription au cours comprennent :

- la participation à la formation
- les documents de cours
- les pauses café
- les repas du midi

Annulation

Le remboursement des droits d'inscription sera possible sur demande écrite transmise **avant le 23 novembre 2022** et donnera lieu à une retenue de 150 € pour frais. Aucun remboursement ne sera effectué après cette date.

Modalités de règlement

Le règlement des droits d'inscription à la formation peut se faire par :

- **chèque bancaire** ou **chèque postal** émis sur une banque française, libellé à l'ordre de : *Société Française de Statistique*
- **virement bancaire**
- **bon de commande**

Numéro d'enregistrement Formation Continue :
11 75 53626 75

La SFdS n'est pas assujettie à TVA



Atelier Statistique

Machine Learning & Méthodes d'apprentissage par arbres de décision

PARIS (I.H.P.)

**30 novembre, 1 & 2 décembre
2022**

Tél : (33) 01 44 27 66 60

E-mail : sfds@ihp.fr
Site internet : <http://www.sfds.asso.fr>