Programme de Formation

Data Science pour l'agriculture : méthodes de machine learning et utilisation de l'IA Générative

Organisation

Durée : 21 heures

Mode d'organisation : Mixte

Contenu pédagogique

Public visé

Ingénieurs et chercheurs travaillant dans les domaines de l'agriculture et de l'environnement.

Objectifs pédagogiques

- Acquérir les bases permettant de manipuler les principales méthodes de machine learning pour des objectifs de prédiction.
- Appréhender la pratique réelle de ces méthodes au travers d'exemples et de travaux pratiques avec le logiciel R.
- Savoir intégrer l'usage de ces méthodes dans un workflow opérationnel avec l'appui de l'intelligence artificielle générative (IAG).

Prérequis

- Avoir des connaissances de base du logiciel R
- Maîtriser les méthodes de base de régression

L'inscription à la formation est validée par un questionnaire de positionnement.

Description

Classe virtuelle 1

Présentation de la formation, tour de table et introduction à la datascience

Méthodes de régression (pénalisée, PLS,GAM, sélection)

+ ***Travail individuel*** « Méthodes de régression » à réaliser en asynchrone et à rendre avant 17h aux formateurs (durée estimée 1h)

Classe virtuelle 2

Retour sur le travail individuel « Méthodes de régression »

Méthodes d'évaluation de la qualité de prédiction

+ ***Travail individuel*** « Méthodes d'évaluation » à réaliser en asynchrone et à rendre avant 17h aux formateurs (durée estimée 1h)

Classe virtuelle 3

Retour sur le travail individuel « Méthodes d'évaluation »

Arbres, random forest et gradient boosting

+ ***Travail individuel*** « Arbres, random forest et gradient boosting» à réaliser en asynchrone et à rendre avant 17h aux formateurs (durée estimée 1h)

Classe virtuelle 4

Retour sur le travail individuel « Arbres, random forest et gradient boosting»

Interprétabilité des modèles de machine learning : Partial Dependence Plots et SHapley Additive exPlanations

Utiliser l'intelligence artificielle générative (IAG) pour coder et faire du machine learning

+ ***Travail individuel*** « Interprétabilité des modèles » et à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h)

+ ***Travail individuel*** « IAG pour coder sous R » et à réaliser en asynchrone et à rendre avant 17h aux formateurs (durée estimée 1h)

Classe virtuelle 5

Retour sur les travaux individuels « Interprétabilité des modèles » et « IAG pour coder sous R »

Principe des LLM et de l'IAG et intégration dans flux de travail scientifique complet

Quiz de fin, conclusion et clôture de la formation

Equipement nécessaire

Disposer d'un équipement pour participer à une classe virtuelle (ordinateur avec connexion internet et web caméra). Durant la formation il est conseillé de s'isoler.

Modalités pédagogiques

- Exposés, échanges et quiz
- Pour chaque cours, des apports théoriques et des illustrations pratiques seront proposés
- Travaux pratiques personnels asychrones à réaliser les après-midis et à rendre aux formateurs avant 17h

Supports pédagogiques

- PDF de la formation fourni en début de session
- Ouvrage collectif « Data science pour l'agriculture et l'environnement – Méthodes et applications avec R et Python » Editions Ellipses

Modalités d'évaluation

- Questionnaire de positionnement
- Évaluation des compétences acquises lors de la formation par des quiz, des exercices pratiques individuels et un quiz final
- Évaluation de la formation par un questionnaire de satisfaction

Informations complémentaires

Moyens de suivi et formalisation à l'issue de la formation

- Émargement par classe virtuelle
- Certificat de réalisation transmis par courriel à l'issue de la formation

Informations sur l'admission

Délais d'accès : vous pouvez vous inscrire jusqu'à 3 semaines avant la formation. Au-delà, merci de nous contacter

Informations sur l'accessibilité

Afin que cette formation soit accessible à tous, nous invitons les personnes qui ont besoin d'aménagements à nous le signaler en contactant directement notre référent handicap : nous contacter