

OFFRE DE STAGE M1/M2

Développement d'un modèle de fondation en spectrométrie de masse

Date de publication : 10/01/2025.

Lieu : Verneuil-en-Halatte (60) - accessible en transports en commun, à 40 mn au Nord de Paris

Type de contrat : stage

Contact : jean-yves.chatelier@ineris.fr - Tél. : 06 19 23 43 84 pour plus d'informations.

DESCRIPTION DE L'OFFRE

Contexte :

Les grands modèles de langue ou les modèles de type Unet pour les images ont montré leur efficacité sur des tâches de classification ou de prédiction.

Même si ces modèles ont nécessité des sources de données d'entraînement volumineuses, il semble que le concept de modèle de fondation soit particulièrement bien adapté pour résoudre des problèmes pour lesquels nous ne disposons pas de telles quantités de données annotées.

A l'Ineris, nous recourons de plus en plus à des appareillages (spectrométrie de masse) qui permettent de détecter de plus en plus de substances dans des échantillons environnementaux sans pour autant que nous soyons capables de les caractériser humainement compte tenu de l'ampleur de la tâche que cela représenterait (nombre de substances existantes, quantité de données à comparer...).

Aussi le recours à des techniques d'apprentissage profond peut nous aider à mieux caractériser ces substances (formule chimique, familles, abondance...) et ainsi apporter des réponses aux questions que se posent les populations et les décideurs sur les risques d'exposition à des substances toujours plus nombreuses dans l'environnement et potentiellement dangereuses.

Objectifs du stage :

Après une étude bibliographique sur l'utilisation de modèles de fondation dans le domaine de l'analyse de données en spectrométrie de masse, l'objectif du stage consistera à mettre en œuvre et documenter un modèle de fondation soit à partir d'un modèle existant soit en le définissant ex nihilo après entraînement sur une base de données d'apprentissage type base de spectres de masse collaborative (Massbank).

Une fois ce modèle mis en place, il s'agira de le raffiner et le valider sur des données expérimentales disponibles à l'Ineris. Des comparaisons de performance pourront aussi être envisagées sur des résultats obtenus antérieurement au sein de l'Ineris par une autre approche.

Si le temps le permet, nous testerons l'intérêt du modèle de fondation développé sur un nouveau cas d'usage (priorisation de substances, cartes de contrôle, prédiction de caractéristiques...).

PROFIL

Bac +3/4 - Ce stage s'adresse à la spécialité : Informatique, Data Science...

Curieux/se et ouvert.e aux nouvelles technologies, vous êtes prêt.e à chercher l'information auprès de vos collègues mais aussi des acteurs du domaine en synthétisant, voire en vulgarisant les informations recueillies (exemples d'utilisation, sources documentaires).

Vos capacités rédactionnelles vous permettent en particulier de vulgariser les notions inhérentes aux concepts des méthodes à base de réseaux de neurones

Vous savez déjà manipuler des méthodes d'apprentissage profond en python. La pratique du lancement de codes sur une partition GPU d'un cluster de calcul haute performance serait un plus.

La connaissance, voire la pratique, en chimie analytique est vivement recommandée.

Vous êtes à l'aise avec l'anglais et comprenez les documentations techniques du domaine.

Vous possédez un sens du contact qui vous permet de vous adresser facilement à des experts métier de tous domaines et de tous niveaux.

DIVERS

Durée : 4 à 6 mois (à partir de mars/avril 2025)

Stage conventionné et indemnisé

Stage basé au siège de l'INERIS à Verneuil en Halatte (60550 - Oise), en présentiel

Indemnisation des frais de transport sous conditions

Envoyer votre candidature (CV et LM) par courriel

Ce poste est ouvert aux personnes en situation de handicap.