



Thèse en science des données géostatistiques

Mots clés :

Apprentissage statistique ; apprentissage profond ; Processus gaussien ; Géostatistique ; Krigeage ; Ré-échantillonnage

Description de la thèse :

Récemment des approches géostatistiques adaptées aux particularités des sites et sols pollués ont été développées et sont de plus en plus utilisées pour cartographier les sols en fonction de seuils réglementaires ou estimer des quantités de matériaux contaminés, tout en quantifiant les incertitudes locales ou globales associées. Le couplage des deux outils, mesures sur site et méthodes géostatistiques, permet d'envisager des campagnes de reconnaissance optimisées où le nombre et la localisation de nouveaux points de mesure sont déterminés au fur et à mesure de l'acquisition des données dans l'objectif de réduire l'incertitude affectant la modélisation de la contamination.

L'objectif est la réduction des incertitudes liées aux coûts de la dépollution de sols de friches et sites pollués, dues au modèle géostatistique avec un échantillonnage souvent aléatoire peu représentatif. Pour ce faire, plusieurs projets ont été soutenus par l'ADEME au fil du temps. Les méthodes géostatistiques actuellement utilisées sont basées sur le calcul de variogrammes par krigeage (interpolation spatiale à variance minimale). Le krigeage en géostatistiques n'est autre que les processus gaussiens en apprentissage statistique.

La présente thèse vise à développer de nouvelles méthodes issues de la théorie de l'apprentissage statistique avec les réseaux de neurones à architectures profondes (Deep Learning en anglais). En effet, le domaine applicatif visé, et plus généralement les géostatistiques, n'ont pas encore profité de l'émergence des méthodes de l'apprentissage statistique avec les réseaux profonds. Deux axes de recherche sont envisagés au cours de cette thèse :

- une meilleure modélisation en couplant l'apprentissage profond aux processus gaussiens et
- le développement de méthode de ré-échantillonnage optimal.

Profil recherché :

- Étudiante ou étudiant en M2 ou 5ème année d'une école d'ingénieur, avec une motivation particulière pour les sciences des données et/ou les statistiques avancées.
- De solides compétences en programmation en Python et/ou C++.
- Un bon sens des relations humaines pour travailler en étroite collaboration avec la startup Tellux.

Localisation : Rouen

Statut : Salarié de l'ADEME

Date limite de candidature : 23 mars 2021

Contact :

- Antonin Van Exem, CEO Tellux : antonin.vanexem@tellux.fr
- Paul Honeine, Professeur au LITIS à Rouen : paul.honeine@univ-rouen.fr
- Maxime Berar, Maître de Conférences au LITIS à Rouen : maxime.berar@univ-rouen.fr