

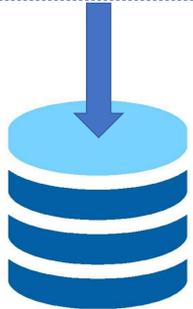
L'Intelligence Artificielle au service des Sites et Sols Pollués

Nous travaillons depuis 2019 sur l'application de l'Intelligence Artificielle aux problématiques des sites et sols pollués. Ces travaux sont découpés en trois grandes étapes: **(1) la constitution d'une base de données, (2) le développement d'algorithmes d'apprentissage automatique** et **(3) la validation des apprentissages.**

Création d'une base de données

L'objectif est de transformer les données disponibles en **informations exploitables numériquement.**

Problématique: Les données de départ sont sous différents formats.



Un travail important a été effectué pour faciliter l'extraction et la conversion des données : extraction de tableaux, reconnaissances d'écriture, mise en place de templates d'importation etc.

Base de Données (B.D.D.)
avec une architecture adaptée

900 000 analyses
13 000 ouvrages
30 000 échantillons

Quelques Résultats

	Algo entraîné	Polyg. Thiessen (X, Y, Z, [C10-C40])
Prédiction de [C10-C40] à partir des X,Y,Z	Précision (> L.D.) 39%	53%
	Précision (< L.D.) 37%	73%
	RMS (mg/kg MS) 82	74
Prédiction de [C10-C40] à partir des [HAP]	Précision (> L.D.) 68%	53%
	Précision (< L.D.) 81%	73%
	RMS (mg/kg MS) 49	74
Prédiction de [C10-C40] à partir de la géologie	Précision (> L.D.) 56%	53%
	Précision (< L.D.) 77%	73%
	RMS (mg/kg MS) 47	74
Prédiction de [SO ₄ ²⁻] sur éluat à partir de [SO ₄ ²⁻] sur brut	Précision (> L.D.) 100%	
	Précision (< L.D.) 100%	
	RMS (mg/kg MS) 18	

Limites de l'I.A.

I.A. = Boîte noire, pas d'accès direct aux modèles établis

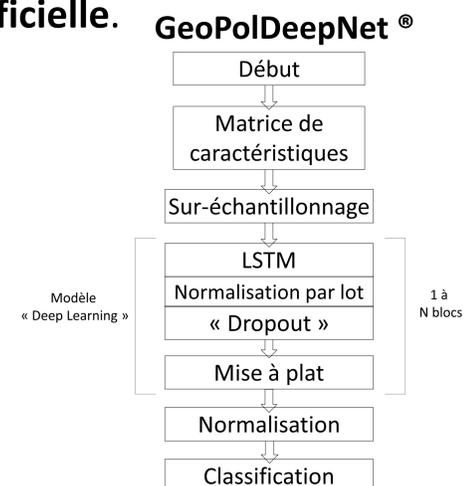


I.A. n'invente rien: ne peut apprendre que ce qui est possible d'apprendre avec les données utilisées pour l'entraînement. Une bonne connaissance de la B.D.D. est indispensable.

Développement des algorithmes

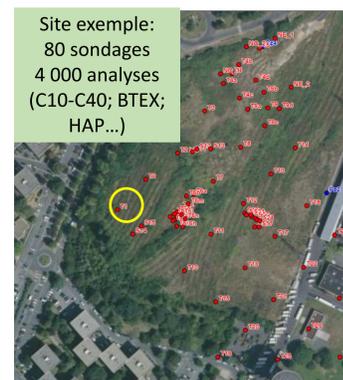
La Base de Données ainsi créée peut, entre autres, être valorisée grâce à des algorithmes d'apprentissage automatique (Machine Learning): **Intelligence Artificielle.**

Nous avons développé un algorithme d'apprentissage profond (**Deep-Learning**) capable de traiter et d'apprendre à partir de notre B.D.D. L'architecture de l'algorithme breveté (**Brevet FR3120260A1**) est donnée sur la figure ci-contre.



Apprentissage et validation des « acquis »

Les données et résultats d'analyses existants sont utilisés pour évaluer l'apprentissage des algorithmes.



Site exemple:
80 sondages
4 000 analyses
(C10-C40; BTEX;
HAP...)

Recommande 80X

- 1) Isole un sondage: Par exemple T1
- 2) Demande à l'algo. de prédire une concentration: [C10-C40] (0-1 m)
- 3) Compare le résultat avec la réalité (mesure en T1)

Trois indicateurs de performance sont utilisés:

- Précision sur les valeurs >L.D. (%)
- Précision sur les valeurs <L.D. (%)
- Moyenne quadratique des écarts (RMS) (mg/kg MS)

Comparaison des performances

I.A (GeoPolDeepNet)
vs
Polygone de Thiessen
(X, Y, Z, [])

Exemples d'application & Perspectives

Création automatique de modèles à partir de plusieurs types d'informations différentes:

- ✓ Géologie et hydrogéologie
- ✓ Liens entre plusieurs paramètres chimiques
- ✓ Informations complémentaires: Géophysiques, analyses MIP...

Développement d'outils de diagnostics et de suivis

« intelligents »: tests sur sites pilotes à réaliser (recherche de partenariats)

Combinaisons avec dispositifs de mesures sur site disponibles : Niton, XRF, PID...

Premières étapes vers la création de dispositifs autonomes

