



ED 398 Géosciences, Ressources Naturelles et Environnement

Proposition de sujet de thèse pour la rentrée universitaire 2020-2021

1. Modalités d'encadrement

Unité(s) de recherche au sein de laquelle le doctorat est réalisé : MINES ParisTech, Centre de Géosciences, 35, rue Saint-Honoré, 77300 Fontainebleau, France

Directeur de l'unité : Lagneau, Vincent, professeur

Directeur(s)* de thèse (HDR ou équivalent) : Chauris, Hervé, professeur, Centre de Géosciences, MINES ParisTech

Co-directeur* de thèse (HDR ou équivalent) : Hachem, Elie, professeur, Groupe CFL, CEMEF, MINES ParisTech

Co-encadrant (non HDR) : Desassis, Nicolas, Centre de Géosciences, MINES ParisTech

Responsabilités spécifiques de chaque encadrant :

- *Hervé Chauris : imagerie sismique (domaine principal d'application)*
- *Elie Hachem : expérience en Machine Learning couplée à la physique*
- *Nicolas Desassis : équipe géostatistique, expérience en Machine Learning et inférence des paramètres*

*** Si un seul directeur de thèse est déclaré, il endosse 100% de la responsabilité de la thèse au regard de l'ED. Si 2 directeurs sont déclarés, ils partagent cette responsabilité à 50%. Le taux de responsabilité maximal est fixé à 300%. Les co-encadrants n'entrent pas dans le décompte, quel que soit leur rôle effectif**

2. Titre Machine Learning et respect des lois de la physique : application à l'imagerie sismique

Title : Physics-Informed Machine Learning in the context of seismic imaging

3. Adresse courriel du contact scientifique : herve.chauris@mines-paristech.fr

4. Description du projet de thèse [champ libre 1 page max).

Vue d'ensemble

Le domaine du Machine Learning est en pleine expansion, pour faire "parler" les données, après possiblement une phase d'apprentissage. Ces approches, trop souvent des boîtes noires, oublient les techniques plus traditionnelles, basées sur la physique, qui visent à expliquer et modéliser les phénomènes. L'objectif de la thèse est de développer, dans le contexte de l'imagerie sismique [1], une approche intermédiaire où les lois de la physique seront (partiellement) respectées. Les principales applications de Machine Learning en sismique concernent à ce jour des tâches de pré-traitement (débruitage, pointé, ...), mais très peu la partie imagerie (retrouver des paramètres du sous-sol, un problème non-linéaire). L'introduction de la physique dans le Machine Learning pourrait permettre de combler cette lacune.

Détails

En 2019, Raissi et al. ont montré comment il est possible de coupler des approches de Machine Learning avec le respect des lois de la physique, sous la dénomination de *Physics-informed Neural Networks (PINN)* [3]. Les applications concernent la résolution d'équations différentielles (i.e. le problème direct) et la résolution de problèmes inverses (retrouver les paramètres qui contrôlent le phénomène physique, par exemple la propagation des ondes, à partir d'observations). C'est cette deuxième approche qui sera développée dans la thèse proposée ici.

D'un côté, les réseaux de neurones profonds permettent en théorie de représenter toute fonction. L'apprentissage est souvent compliqué et dans les problèmes physiques les observables sont peu nombreuses et coûteuses à acquérir. D'un autre côté, les équations de la physique représentent une information cruciale que les approches de Machine Learning traditionnelles ne prennent pas en compte. L'approche de Raissi et al., 2019, comble cette lacune. La fonction objective (loss function) dans les réseaux de neurones comprend plusieurs termes pour s'assurer que les données prédisent bien les observables disponibles et que les lois de la physique sont également respectées [3]. Ce second terme peut être vu comme une régularisation, indispensable pour éviter de sur-exploiter des données bruitées (over-fitting). Les algorithmes d'auto-différentiation au sein des réseaux de neurones, permettent d'estimer les meilleurs paramètres qui minimisent la fonction objective.

Cette approche est très attractive et sera étendue et modifiée pour être applicable au domaine de l'imagerie sismique. Dans le cadre de l'acquisition sismique, une source active provoque des ondes acoustiques / élastiques qui se propagent dans le sous-sol. Des capteurs, souvent en surface, enregistrent la pression ou le déplacement des particules en fonction du temps. L'objectif est de déterminer des champs de vitesses de propagation des ondes (et tout autre champ qui contrôle la propagation des ondes comme la densité, l'anisotropie, ...).

Par rapport aux premiers cas illustrés dans la littérature autour du couplage entre Machine Learning et respect de la physique, l'imagerie sismique offre des particularités à prendre en compte :

- L'aspect propagatif des ondes fait que les champs d'onde ne sont pas lisses et que contrôler qu'un champ d'onde vérifie bien l'équation des ondes doit se faire a priori en un très grand nombre de points (contrairement à un problème diffusif avec une solution beaucoup plus régulière) ;
- La fonction coût traditionnelle en imagerie sismique contient des minima secondaires. Comment l'approche PINN peut-elle gérer cette difficulté ? Comment tirer avantage du contenu fréquentielle des données ? Dans les approches classiques, l'inférence utilise d'abord les basses fréquences, puis augmente le contenu fréquentiel. Comment le réseau de neurones peut-il tirer parti de cette approche (e.g. proxi pour la modélisation) ?
- Enfin, le nombre de paramètres à estimer est potentiellement très grand (des milliers ou largement plus, car les paramètres sont fonction de la position spatiale). Dans les premiers articles, seules quelques valeurs sont recherchées. Comment aborder cette thématique et jouer avec l'architecture des réseaux de neurones ? Dans ce contexte, les GAN (Generative Adversarial Networks) peuvent aider à déterminer la meilleure paramétrisation [2].

L'objectif de la thèse est donc de développer une approche de Machine Learning couplée aux lois de la physique, dans le cadre de l'imagerie sismique. Les validations se feront sur des données synthétiques et des données réelles.

Références principales

[1] Chauris, H. (2019). Full Waveform Inversion, in *Seismic Imaging, a practical approach*, J-L. Mari and M. Mendes (Eds.), EDP Sciences, chapter 5, 23 p., ISBN (ebook): 978-2-7598-2351-2, doi:10.1051/978-2-7598-2351-2.c007

[2] Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio (2014). Generative Adversarial Networks. *Proceedings of the International Conference on Neural Information*, arXiv:1406.2661

[3] Raissi, M., P. Perdikaris, G.E. Karniadakis (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378**, 686-707

5. Compétences et connaissances requises [champ libre 1/2 page max.]

Le/la candidat(e) doit avoir de très bonnes bases en mathématiques et en physique. Il/elle doit avoir un fort intérêt pour le Machine Learning et un goût pour les applications à l'imagerie géophysique du sous-sol, tout particulièrement avec les méthodes sismiques. Il/elle doit aussi avoir une très bonne expérience en programmation informatique. Un plus est d'avoir une expérience en calcul parallèle (HPC). Le/la candidat(e) doit maîtriser l'anglais, tant à l'écrit qu'à l'oral.

6. Conditions matérielles de réalisation du projet de recherche

Financement spécifiques obtenus pour le projet : Oui, si oui lesquels ? Contrat Doctoral

Financement des missions nécessaires pour la réalisation du projet : Oui, si oui lesquels et pour quelles missions ? Conférences

Accès à des bases de données spécifiques : Non, si oui lesquelles ?

Accès à des ressources documentaires spécifiques : Non, si oui lesquelles ?

Accès à des plateformes : Non, si oui lesquels ?

Accès à des grands instruments : Non, si oui lesquels ?

Autres :

7. Précisions sur les objectifs de valorisation des travaux issus du projet de recherche : Publications dans des revues de géophysique et de mathématiques appliquées (idem pour les conférences), participation active au projet MINDS.

Visa de la Direction de l'Unité

Commentaires éventuels :