

Reproducing Kernel Banach Spaces pour réseaux profonds

Sujet de stage M2 Machine Learning

Informations générales

- **Niveau** : Master 2 (Machine Learning / Data Science)
- **Durée** : 4 mois
- **Lieu** : Laboratoire ERIC **Date de début** : Avril 2026
- **Encadrement** : Stéphane Chrétien et Jordan Patracone

1 Contexte scientifique

Les réseaux profonds sont des fonctions entrée-sortie, qui a une entrée, par exemple une image, associe un label en sortie, par exemple le type d'objet représenté dans l'image. Ces fonctions peuvent être étudiées à la lumière de la théorie de l'approximation, une discipline très utile pour comprendre les capacités de certaines constructions telles que les réseaux de neurones, à approximer une fonction donnée. Il a été récemment découvert que l'optimisation des réseaux de neurones à activations ReLU peuvent être interprétés comme la recherche de solutions de problèmes de *régression régularisée* où la variable est une fonction que l'on oblige, par pénalisation, à posséder des propriétés particulières. L'article [1] propose un cadre inspiré de la théorie des splines, introduit un espace fonctionnel compositionnel adapté aux architectures profondes, et établit un *propriété de représentation* : certaines solutions optimales admettent une représentation exacte sous forme de réseaux profonds ReLU.

L'analyse donnée dans l'article [1] met notamment en avant :

1. une norme à effet *parcimonieux* (sparsity-promoting),
2. l'apparition naturelle de *connexions résiduelles* (skip connections), et
3. des structures de *faible rang* (rank-bounded) pour les matrices de poids.

L'objectif du stage est double :

1. reproduire et évaluer expérimentalement les régularisations induites par ce cadre,
2. étudier dans quelle mesure l'approche permet de comprendre pourquoi les réseaux de neurones généralisent facilement à de nouvelles données.

2 Objectifs

1. Implémenter (ou approximer de manière fidèle) des régularisations inspirées de [1] et établir des baselines robustes (weight decay, éventuellement path-norm si disponible).
2. Approfondir la question de savoir si l'approche fonctionnelle proposée peut faire découvrir d'autres méthodes d'apprentissage, fondées sur le gradient stochastique fonctionnel par exemple.
3. Vérifier empiriquement l'impact de l'approche sur

- (a) performance et généralisation,
 - (b) parcimonie et structures effectives,
 - (c) rang numérique des matrices,
 - (d) robustesse et stabilité.
4. Réfléchir à l'étude de la généralisation dans le cadre fonctionnel proposé.

3 Compétences requises

- Très bonne maîtrise de **PyTorch** et des boucles d'entraînement.
- Connaissances solides en **optimisation** et **régularisation** (normes, contraintes, pénalités).
- Compréhension des architectures classiques en Deep Learning.
- Appétence pour la lecture d'articles et, si possible, un intérêt pour l'aspect théorique (non bloquant).

Références

- [1] Rahul Parhi and Robert D. Nowak. What kinds of functions do deep neural networks learn? insights from variational spline theory. *SIAM Journal on Mathematics of Data Science*, 4(2) :464–489, 2022.