



Fédération
Informatique
de Lyon



**LABORATOIRE
HUBERT CURIEN**
UMR • CNRS • 5516 • SAINT-ETIENNE

MOLE - Model Optimization from Learning Examples

Stage de Recherche

4 à 6 mois

-----English Below-----

Dans le cadre du projet MOLE - Model Optimization from Learning Examples, réunissant les laboratoires ERIC et Hubert Curien, nous recrutons un.e stagiaire pour une durée de 4 à 6 mois.

Contexte du projet

L'approche du learning-to-learn[1,5], sous domaine du meta-apprentissage [6,7], consiste à apprendre à un modèle à optimiser un autre modèle. L'idée du projet MOLE est d'exploiter un paradigme récent : les modèles de diffusion [3], qui semblent bien adaptés à cette tâche. Le meta-learner, entraîné à partir d'historiques d'états de paramètres de réseaux de neurones, permettrait de prédire les poids d'un réseau en une passe. On obtient un procédé d'entraînement beaucoup plus frugal qu'un entraînement classique par descente de gradient stochastique. Les modèles de diffusion ont déjà été utilisés comme meta-learner[4], mais l'existant présente de nombreuses limitations, notamment vis-à-vis de la modélisation des paramètres à optimiser et de l'absence de conditionnement à l'étape de méta-entraînement. Le projet vise à répondre à ces différentes limitations.

Travail attendu

Le.a stagiaire devra travailler en étroite collaboration avec les deux membres du projet.

Plus spécifiquement les verrous scientifiques identifiés sont :

- déterminer le meilleur support de représentation des réseaux de neurones et adapter les modèles de diffusion à cette représentation
- Implémentation d'un modèle de diffusion pour la prédiction de paramètres de réseaux de neurones. Cette partie concerne les aspects plus techniques de développement/entraînement du modèle, d'abord reproduction de Peebles et al [4] puis implémentation du modèle proposé.

L'objectif de l'organisation des stages est de bénéficier au mieux des singularités disciplinaires des deux porteurs du projet : mêler les aspects théoriques de la modélisation et les problématiques techniques d'implémentation et de reproduction de l'existant.

Compétences requises

Le/la candidat.e doit posséder des compétences solides en Apprentissage Automatique (maîtrise des framework d'apprentissage deep tels que PyTorch/TensorFlow), mais aussi des compétences avancées en Python, et en modèles génératifs.

Salaire

4,05 euros par heure, suivant les grilles salariales universitaires

Informations supplémentaires

Le lieu d'accueil est au choix, le laboratoire ERIC ou le Laboratoire Hubert Curien.

Le Laboratoire Hubert Curien (<https://laboratoirehubertcurien.univ-st-etienne.fr>) est une unité mixte de recherche (UMR 5516) de l'Université Jean Monnet de Saint-Etienne, du Centre National de la Recherche Scientifique (CNRS) et de l'Institut d'Optique Graduate School. Les activités de recherche du laboratoire sont organisées selon deux départements scientifiques : Optique, photonique et surfaces et Informatique, sécurité, image. L'équipe Data Intelligence, au sein de laquelle la personne recrutée travaillera, est spécialisée dans le domaine du Machine Learning

Le laboratoire ERIC (Unité de Recherche des Universités Lyon 2 et Lyon 1) développe des recherches théoriques et appliquées dans les domaines de la science des données et de l'informatique décisionnelle. Elles visent à valoriser les grandes bases de données complexes, notamment dans les domaines des lettres, langues, sciences humaines et sociales (LLSHS) et se situent dans les domaines de la Science des données, de l'Informatique décisionnelle et des Humanités numériques.

Pour candidater, merci d'envoyer à antoine.gourru@univ-st-etienne.fr et jairo.cugliari@univ-lyon2.fr : un CV détaillé, une lettre de motivation et votre dernier relevé de notes.

[1] Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M. W., Pfau, D., Schaul, T., ... & De Freitas, N. (2016). Learning to learn by gradient descent by gradient descent. *Advances in neural information processing systems*, 29.

[2] Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. (2015, June). Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning* (pp. 2256-2265). PMLR.

[3] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840-6851.

[4] Peebles, W., Radosavovic, I., Brooks, T., Efros, A. A., & Malik, J. (2022). Learning to Learn with Generative Models of Neural Network Checkpoints. *arXiv e-prints*, arXiv-2209.

[5] Hochreiter, S., Younger, A. S., & Conwell, P. R. (2001). Learning to learn using gradient descent. In *Artificial Neural Networks—ICANN 2001: International Conference Vienna, Austria, August 21–25, 2001 Proceedings 11* (pp. 87-94). Springer Berlin Heidelberg.

[6] Schmidhuber, J. (1987). *Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook* (Doctoral dissertation, Technische Universität München).

[7] Finn, C., Abbeel, P., & Levine, S. (2017, July). Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning* (pp. 1126-1135). PMLR

--



MOLE - Model Optimization from Learning Examples

Stage de Recherche

4 à 6 mois

-----English Version -----

As part of the MOLE - Model Optimization from Learning Examples project, involving the ERIC and Hubert Curien laboratories, we are recruiting an intern for a period of 4 to 6 months.

Project background

The learning-to-learn approach [1,5], a sub-domain of meta-learning [6,7], consists in teaching a model to optimize another model. The idea of the MOLE project is to exploit a recent paradigm: diffusion models [3], which seem well suited to this task. The meta-learner, trained from state histories of neural network parameters, would predict the weights of a network in one pass. The result is a much more frugal training process than conventional stochastic gradient descent training. Diffusion models have already been used as meta-trainers[4], but there are a number of limitations, particularly with regard to the modeling of the parameters to be optimized and the lack of conditioning at the meta-training stage. The project aims to address these limitations.

Expected work

The intern will work in close collaboration with the two project members.

More specifically, the scientific challenges identified are :

- determining the best medium for representing neural networks and adapting diffusion models to this representation
- Implementing a diffusion model for neural network parameter prediction. This section deals with the more technical aspects of model development/training, first reproducing Peebles et al [4] and then implementing the proposed model.

The aim of the organization of the internships is to take full advantage of the disciplinary singularities of the two project leaders: to combine the theoretical aspects of modeling with the technical issues of implementing and reproducing the existing model.

Skills required

The candidate must have solid skills in Machine Learning (mastery of deep learning frameworks such as PyTorch/TensorFlow), as well as advanced skills in Python and generative models.

Salary

4.05 euros per hour, in accordance with university salary scales

Additional information

You can choose to work at the ERIC laboratory or the Laboratoire Hubert Curien.

The Laboratoire Hubert Curien (<https://laboratoirehubertcurien.univ-st-etienne.fr>) is a joint research unit (UMR 5516) of the Université Jean Monnet de Saint-Etienne, the Centre National de la Recherche Scientifique (CNRS) and the Institut d'Optique Graduate School. The laboratory's research activities are organized into two scientific departments: Optics, Photonics and Surfaces, and Computing, Security and Imaging. The Data Intelligence team, in which the new recruit will work, specializes in Machine Learning.

The ERIC laboratory (a research unit of the Universities of Lyon 2 and Lyon 1) develops theoretical and applied research in the fields of data science and business intelligence. It aims to make the most of large, complex databases, particularly in the fields of literature, languages, humanities and social sciences (LLSHS), in the areas of Data Science, Business Intelligence and Digital Humanities.

To apply, please send to antoine.gourru@univ-st-etienne.fr and jairo.cugliari@univ-lyon2. Please send a detailed CV, a cover letter and your most recent transcript.