PhD position in Engineering and Computer Science, Sorbonne Université, Paris, Fr

Foundation Models for Physics-Aware Deep Learning

Contact : Patrick Gallinari, patrick.gallinari@sorbonne-universite.fr

Location: Sorbonne Université, Pierre et Marie Curie Campus, 4 Place Jussieu, Paris, Fr. Machine Learning and Information Access team.

Candidate profile: Master degree in computer science or applied mathematics, Engineering school. Background and experience in machine learning. Good technical skills in programming.

How to apply: please send a cv, motivation letter, grades obtained in master, recommendation letters when possible to patrick.gallinari@sorbonne-universite.fr

Start date: October/November 2024 for three years

Note: The research topic is open and depending on the candidate profile could be oriented more on the theory or on the application side

Keywords: deep learning, physics-aware deep learning, fluid dynamics, AI4Science

---

Context

Physics-aware deep learning is an emerging research field aiming at investigating the potential of AI methods to advance scientific research for the modeling of complex natural phenomena. This is a fast-growing research topic with the potential to boost scientific progress and to change the way we develop research in a whole range of scientific domains. An area where this idea raises high hopes is the modeling of complex dynamics characterizing natural phenomena occurring in domains as diverse as climate science, earth science, biology, fluid dynamics. A diversity of approaches is being developed including data-driven techniques, methods that leverage first principles (physics) prior knowledge coupled with machine learning, neural solvers that directly solve differential equations. Despite significant advances, this remains an emerging topic that raises several open problems in machine learning and application domains. Among all the exploratory research directions, the idea of developing **foundation models** for **learning from multiple physics** is emerging as one of the **fundamental challenges in this field**. This PhD proposal is aimed at exploring different aspects of this new challenging topic.

Research Directions

Foundation models have become prominent in domains like natural language processing (GPT, Llama, Mistral, etc) or vision (CLIP, DALL-E, Flamingo, etc). Trained with large quantities of data using self-supervision, they may be used or adapted for downstream tasks while benefiting through pre-training from large amounts of training data. Initial attempts at replicating this framework in scientific domains is currently being investigated in fields as diverse as protein (Jumper et al. 2021), molecule (Zhou 2023), weather forecasting (Pathak 2022, Nguyen 2023, Kochkov 2024). **Is the paradigm of foundation models adaptable to more general physics modeling such as the complex behavior of dynamical systems?** Large initiatives are emerging on this fundamental topic (http://micde.umich.edu/SciFM24, https://iaifi.org/generative-ai-workshop**)**. Some preliminary attempts are currently being developed (McCabe 2023, Subramanian 2023**,** Hao 2024). They suggest that learning from multiple steady-state or time dependent partial differential equations (PDEs) could enhance the prediction performance on individual equations. This high stake, high gain setting might be the next big move in the domain of data-driven PDE modeling. **The objective of the PhD is to explore different directions pertaining to the topic of foundation models for physics, focused on the modeling of dynamical systems**.

Solving parametric PDEs

A first step is to consider solving parametric partial differential equations (PDEs), i.e. PDEs from one family with varying parameters including initial and boundary conditions, forcing functions, or coefficients. It is possible that different parameters values, give rise to very different dynamics. Current neural solvers operate either on fixed conditions or on a small range of parameters with training performed on a sample of the parameters. A first direction will be to analyze the potential of representative NN solvers to interpolate and extrapolate out of distribution to a large range of conditions when learning parametric solutions. A key issue is then the development of training techniques allowing for fast adaptation on new dynamics. We will investigate methods inspired from meta-learning for adaptive strategies (Yin 2021, Kirchmeyer 2022).

Tackling multiple physics

The foundation approach is particularly interesting in the case of scarce data, provided physics primitive could be learned from related but different PDE dynamics that are available in large quantities and then transferred to the case of interest. Learning from multiple PDEs raises algorithmic challenges since they operate on domains with different space and time resolutions, shapes and number of channels. We will consider an Encode-Process-Decode framework so that the commonalities between the dynamics are encoded and modeled in a shared latent space and the encoding-decoding process allows to project from and to the observation space for each PDE. As for the temporal variability of the observations, one will consider models that can operate on irregular series in the spirit of (Yin2023). This framework will be evaluated with selected backbones.

Generalization and few shot capabilities

Generalization to new dynamics is the core problem motivating the development of foundation models in science. This is a key issue for the adoption of data-driven methods in physics and more generally in any context were the data is scarce. We will consider the general framework of few shot learning aiming at fine tuning pre-trained models for downstream tasks. In this context the objective will be to develop frameworks for the fast adaptation of foundation models to target tasks. Different strategies will be analyzed and developed including parameters sampling, meta-learning for adaptation (Yin 2023) and strategies inspired from the developments in semantics and language applications like in-context learning (Chen 2024).

Position and Working Environment

The PhD studentship is a three years position starting in October/November 2024. It does not include teaching obligation, but it is possible to engage if desired. The PhD candidate will work at Sorbonne Université (S.U.), Pierre et Marie Campus in the center of Paris. He/She will integrate the MLIA team (Machine Learning and Deep Learning for Information Access) at ISIR (Institut des Systèmes Intelligents et de Robotique). MLIA is collaborating with fellow scientists from other disciplines such as climate or fluid mechanics. The PhD candidate will be encouraged to get involved in such collaborations.

References

Chen, W., Song, J., Ren, P., Subramanian, S., Morozov, D., & Mahoney, M. W. (2024). Data-Efficient Operator Learning via Unsupervised Pretraining and In-Context Learning. 1–21. http://arxiv.org/abs/2402.15734

Hao, Z., Su, C., Liu, S., Berner, J., Ying, C., Su, H., Anandkumar, A., Song, J., & Zhu, J. (2024). DPOT: Auto-Regressive Denoising Operator Transformer for Large-Scale PDE Pre-Training. Icml. http://arxiv.org/abs/2403.03542

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., ˇZ´ıdek, A., Potapenko, A., et al. Highly accurate protein structure prediction with alphafold. Nature, 596(7873):583–589, 2021.

Kirchmeyer, M., Yin, Y., Donà, J., Baskiotis, N., Rakotomamonjy, A., & Gallinari, P. (2022). Generalizing to New Physical Systems via Context-Informed Dynamics Model. ICML.

Kochkov D, Yuval J, Langmore I, et al. Neural General Circulation Models. In: ArXiv:2311.07222v2. ; 2024.

McCabe, M., Blancard, B. R.-S., Parker, L. H., Ohana, R., Cranmer, M., Bietti, A., Eickenberg, M., Golkar, S., Krawezik, G., Lanusse, F., Pettee, M., Tesileanu, T., Cho, K., & Ho, S. (2023). Multiple Physics Pretraining for Physical Surrogate Models. 1–25 http://arxiv.org/abs/2310.02994

Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K., and Grover, A. Climax: A foundation model for weather and climate. arXiv:2301.10343, 2023.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., et al. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators.

Subramanian, S., Harrington, P., Keutzer, K., Bhimji, W., Morozov, D., Mahoney, M., & Gholami, A. (2023). Towards Foundation Models for Scientific Machine Learning: Characterizing Scaling and Transfer Behavior. NeurIPS. http://arxiv.org/abs/2306.00258

Yin, Y., Ayed, I., de Bézenac, E., Baskiotis, N. and Gallinari, P. (2021). LEADS: Learning Dynamical Systems that Generalize Across Environments. *Neurips*.

Yin, Y., Kirchmeyer, M., Franceschi, J.-Y., Rakotomamonjy, A., & Gallinari, P. (2023). Continuous PDE Dynamics Forecasting with Implicit Neural Representations. *ICLR*.

Zhou, G., Gao, Z., Ding, Q., Zheng, H., Xu, H., Wei, Z., Zhang, L., and Ke, G. Uni-mol: a universal 3d molecular representation learning framework. 2023