Internship Master or Engineering degree

AI for Science: Physics Based Deep Learning for Modeling Complex Dynamics. Application to Climate

https://mlia.lip6.fr/available-positions/

Informations

Contact : patrick.gallinari@sorbonne-universite.fr

Where : Machine Learning and Information Access team - MLIA - *https://mlia.lip6.fr*, Sorbonne University, Paris, Fr

Dates and duration : 6 months starting in spring 2022

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

Candidate profile: Master or engineering degree in computer science or applied mathematics. The candidate should have a strong scientific background with good technical skills in programming.
Note: The research topic is open and depending on the candidate profile could be oriented more on the theory or on the application side. The machine learning team collaborates with colleagues from Climate science at Sorbonne Universite.

Stipend : classical French academic internship gratification around 550 E/ mois

Context

AI for science is concerned with the exploration of machine learning for scientific computing in domains traditionally dominated by physics models (first principles) like earth science, climate science, biological science, etc. It is particularly promising in problems involving processes that are not completely understood, or computationally too complex to solve by running the physics inspired model. Researchers from multiple disciplines have started to explore how to integrate physics knowledge and data, a challenging direction. We consider here the modeling of complex dynamical systems characterizing natural phenomena, a recent and fast growing research topic (Willard et al. 2020, Thuerey et al. 2021), with a focus on climate modeling applications (de Bezenac 2018, Ayed 2020), and with the objective of combining model based physics (MB) and machine learning (ML) approaches.

Objective

The global objective is the development of new models integrating physics prior knowledge and deep learning (DL) for spatio-temporal dynamics characterizing physical phenomena such as those underlying earth-science and climate observations. The classical modeling tools for such dynamics in physics and applied mathematics rely on partial differential equations (PDE). We then consider situations where the physical prior background is provided by PDEs. Two main directions will be explored.

Hybrid systems - Interfacing Deep neural Networks (DNNs) and PDEs

The key issue addressed here is how to combine prior information expressed as PDEs and information extracted from data. ML comes as a complement to numerical models by allowing to take into account information not present in the model or to integrate information provided by observation data. From a DNN perspective, physical background constitutes prior knowledge that guides and constrains the learning process. A principled method for combining physics and DL has been proposed in (Yin et al.

2021a) and further refined for climate data. This will be a starting point for the internship. The objective will then be to extend this work towards modeling more complex situations closer to real data.

Domain generalization for deep learning as dynamical models

Depending on the progress on the first topic, one will consider the issue of domain generalization of hybrid models. Explicit physical models come with guarantees and can be used in any context (also called domain or environment) where the model is valid. This is not the case for DNNs, and we have no guarantee that they can be extrapolated to new physical environments. Generalization has been at the heart of ML for several years. We propose here to tackle the problem by drawing inspiration from recent ML frameworks developed for handling the new research topic of domain generalization, such as (Yin 2021b, Wang 2021).

Application to climate data

The application will target the modeling of the dynamics of ocean circulation, which is a component of climate models. It is now possible to observe by altimetry surface ocean currents. The launch of new satellite observation projects in 2021 will generate small-scale data. However, these are not directly exploitable by current physical modeling techniques. The objective is to instantiate the models combining deep learning and physics developed during the internship in order to handle complex climate simulation data close to the real observations.

References

Ayed, I., de Bézenac, E., Pajot, A., Brajard, J., and Gallinari, P. 2020. Learning Dynamical Systems from Partial Observations. ICASSP 2020

de Bezenac, E., Pajot, A., and Gallinari, P. 2018. Deep Learning For Physical Processes: Incorporating Prior Scientific Knowledge. ICLR.

Thuerey, N., Holl, P., Mueller, M., Schnell, P., Trost, F. and Um, K. 2021. Physics-based Deep Learning, arXiv:2109.05237

Wang, R., Walters, R. and Yu, R. 2021. Meta-Learning Dynamics Forecasting Using Task Inference. *http://arxiv.org/abs/2102.10271* (2021), 1–20.

Willard, J.D., Jia, X., Xu, S., Steinbach, M. and Kumar, V. 2020. Integrating physics-based modeling with machine learning: A survey. arXiv:2003.04919, 1–34.

(Yin et al. 2021a)Yin, Y., Le Guen, V., Dona, J., de Bezenac, E., Ayed, I., Thome, N. and Gallinari, P. 2021. Augmenting Physical Models with Deep Networks for Complex Dynamics Forecasting. *ICLR* (2021).

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