PhD position in Paris, France

Integrating Deep Learning and Physics for Modeling Complex Dynamics, Applications to Climate Science

---

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

Location: Sorbonne Université, Pierre et Marie Curie Campus, 4 Place Jussieu, Paris, Fr

Candidate profile: Master in computer science or applied mathematics, Engineering school. Strong background and experience in machine learning and 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: September/ October 2021

Context

Deep Learning has found important success in many application fields. It is beginning to be explored 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. However, the application of state of the art DNN models often meets limited success in scientific applications. This is due to different factors: the complexity of the underlying physical phenomenon, the large data requirement of deep neural networks (DNNs), their inability to produce physically consistent results. The research community has 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 with a focus on climate modeling applications, and with the objective of combining model based physics (MB) and machine learning (ML) approaches.

Research directions

Combining Physics and Deep Learning

The integration of physics background and ML has recently motivated the interest of several communities (Willard 2020). This issue may been explored from different perspectives. We will focus here on the modeling of spatio-temporal dynamics 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. We are interested in solving two different problems. A first problem corresponds to the situation where the PDEs are too complex to run a full simulation and ones wants to reduce the simulation cost. A possible strategy here is to run a simulation at a coarse precision and use ML for complementing the physical simulation and reach high fidelity prediction. The second problem corresponds to the case where the PDE only provides partial information about the underlying physical phenomena and this physical knowledge it is to be complemented with ML by extracting the complementary information from data. Although they correspond to different objectives, the two problems share many similarities from a ML point of view. Initial attempts to solve similar problems can be found in recent work such as (de Bezenac 2018, Harlim 2020, Yin 2021). This will be further developed during the PhD project with the objective of analyzing and developing different integration frameworks.

Learning at Multiple Scales

Modeling dynamical physical processes often requires solving PDEs at different spatio-temporal scales. For example in climate, global phenomena are influenced by dynamics operating at a smaller scale. Global simulation models could not be run, due to their complexity, at fine discretization levels. This problem is known as "downscaling" and DNNs could help improve this multiscale problem. Similar problems occur e.g. in computational fluid dynamics. Learning at different scales is an open issue in ML. Most current DNN deployments for learning dynamics operate at a fixed spatio-temporal discretization. Recent advances (Sitzman 2020, Li 2021) allow us learning a function space instead of discrete flows and open the possibility for generalizing at different spatio-temporal resolutions. This will be used as starting point for learning at different scales with DNNs.

Uncertainty Quantification

Uncertainty quantification is of great importance in climate modeling. This requires characterizing the distribution $p(y|x)$ where y is the response and x the covariates of interest. Since Monte Carlo simulations are unfeasible for such applications, physics has developed solutions such as reduced order models for modeling uncertainty while ML often relies on Gaussian Processes for quantifying uncertainty in physical processes. However none of these approaches scales well to high dimensions. We will explore recent developments based on Neural Processes (Garnelo 2018, Norcliffe 2021) for modeling uncertainty.

Position and Working Environment

The PhD studentship is a three year position starting in September/ October 2021. 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 Machine Learning and Deep Learning for Information Accesss team (https://mlia.lip6.fr/) at S.U.

On the Climate side, the candidate will be co-supervised by M. Levy and S. Thiria from LOCEAN laboratory, https://www.locean-ipsl.upmc.fr/

References

Ayed, I., de Bezenac, Emmanuel , Pajot, A., Brajard, J. and Gallinari, P. 2019. Learning the hidden dynamics of ocean temperature with Neural Networks. *Climate Informatics* (2019).

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

Garnelo, M., Rosenbaum, D., Maddison, C.J., Ramalho, T., Saxton, D., Shanahan, M., Teh, Y.W., Rezende, D.J. and Ali Eslami, S.M. 2018. Conditional neural processes. *ICML* (2018), 1704–1713.

Harlim, J., Jiang, S.W., Liang, S. and Yang, H. 2021. Machine learning for prediction with missing dynamics. *Journal of Computational Physics*. 428, (2021), 109922.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A. and Anandkumar, A. Fourier Neural Operator for Parametric Partial Differential Equations. *ICLR* (2021), 1–16.

Norcliffe, A., Cristian, B., Day, B., Moss, J. and Liò, P. 2021. Neural ODE Processes. *ICLR* (2021), 1–14.

Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G. and University, S. 2020. Implicit Neural Representations with Periodic Activation Functions. *Neurips* (2020).

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

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).