PhD position in Computer Science and Numerical Analysis
Deep Neural Networks and Differential Equations

Sorbonne Université, Paris, Fr

Laboratoire Jacques Louis Lions – INRIA Paris and Laboratoire d'informatique de Paris 6

Advisors and contacts: julien.salomon@inria.fr (Laboratoire Jacques-Louis Lions), patrick.gallinari@lip6.fr (Laboratoire d'informatique de Paris 6)

Starting date : October 2020

Keywords: Machine Learning, Deep Neural Networks, Numerical Analysis, Differential Equations

Objective

Differential equations form one of the bedrocks of scientific computing, while neural networks have emerged as the preferred tool of modern machine learning. They offer complementary strengths: the modelling power and interpretability of differential equations, and the approximation and generalization power of deep neural networks. The objective of the thesis is to develop links between DNNs and DEs in order to start answering central questions like: how could DNNs be used to solve PDEs, how the concepts of numerical analysis could be adapted to DNNs, how to develop hybrid models incorporating both NN modules and ODE/PDE solvers? On the application side, we will focus on PDEs arising from environmental applications. The PhD is at the interplay of machine learning and numerical analysis and will be co-supervised by specialists of the two domains. We envision three research directions.

Hybrid systems - Interfacing DNNs and ODE/PDE solvers

The key issue addressed here is how to combine prior information expressed as PDEs and information extracted from data. There are several instances for which this is a sensible approach. 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. For example, we have shown how to learn the motion field characterizing an advection-diffusion system [de Bezenac et al. 2018]. We will consider a general process where a source of physical knowledge under the form of a PDE representing partial information on the process is available. The objective is to develop schemes for completing this information with pure data based models.

Numerical analysis of DNNs for transport models

We will adapt concepts from the numerical analysis of hyperbolic PDEs to NNs solvers. A special emphasis will be given to the study of NNs when dealing with transport phenomena. Because of their hyperbolic nature, the transport equations have solutions which lose regularity with time and could therefore raise difficulties for NNs, although the loss of regularity is often mitigated by other phenomena (like viscosity) in real-life applications. We will investigate how the consistency property, which is expected of numerical schemes for a proper discretization of the continuous equation, can be extended to NNs. We will evaluate whether the NNs are subject to stability constraints, drawing a parallel with the Courant-Friedrichs-Lewy condition which applies to almost all numerical schemes.

DNNs and reduction for transport models

A third research direction will concern the reformulation of neural networks as reduction processes. The issue here is crucial: it is known that usual model reduction techniques fail in the case of transport models. This can be explained by the large Kolmogorov dimension of the samples produced over time by these models. DNNs could solve this problem since, by working with a number of parameters lower than the classical discretization, they are in fact a reduction algorithm. In this way, several recent papers analyze the approximation quality of neural networks. The results are encouraging since they show that their performance is equal or superior to classical approximation methods, such as linear methods of approximation, e.g. approximation by polynomials or by piecewise polynomials on prescribed partitions. The objective here is to analyze neural networks from the perspective of reduction methods, with a particular focus on verifying the properties usually considered in this context like the reproduction property or the construction of an error estimate.

Position w.r.t State of the Art

Solving DEs with DNNs dates back to the 90es [Lagaris et al. 1998], but it is only very recently (2017-2018) that this research topic fully emerged. This rapidly growing field motivates the interest of several communities. Two main directions are currently investigated. One is how to use DNNs for solving and identifying PDEs [Rudy et al. 2017, Raissi 2018, Long et al. 2018, Sirignano and Spiliopoulos 2018]. The other direction comes from the interpretation of successful DNNs like ResNET as ODE discretization schemes [Ruthotto and Haber 2018, Lu et al. 2018]. This motivates the development of new DNN architectures. The issues of stability and robustness for different schemes is currently being investigated [Haber et al. 2019, Chang et al. 2018]. [Chen et al. 2018] uses ODE solvers for supervised learning problems. Learning dynamics from incomplete information is addressed in [Ayed et al. 2020] and insertion of physical priors in DNNs in [de Bezenac et al. 2018].

Working Environment

The candidate will work at SCAI (Sorbonne Center for Artificial Intelligence) in Paris, under the supervision of Julien Salomon (numerical analysis) and Patrick Gallinari (Machine Learning).

Candidate profile

Master or engineering degree in computer science or applied mathematics. The topic is at the crossroad of machine learning and numerical analysis. The candidate should have a strong scientific background with specialization in one of the two domains, good technical skills in programming. Experience of project development with machine learning platforms such as PyTorch or TensorFlow is a plus.

Application

Send a CV, motivation letter and if possible recommendation letter or contact to julien.salomon@inria.fr and patrick.gallinari@lip6.fr

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*.

CHANG, B., MENG, L., HABER, E., RUTHOTTO, L., BEGERT, D., AND HOLTHAM, E. 2018. Reversible Architectures for Arbitrarily Deep Residual Neural Networks. *AAAI*, 2811–2818.

Chen, R.T.Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. 2018. Neural Ordinary Differential Equations. *NIPS*.

HABER, E., LENSINK, K., TRIESTER, E., AND RUTHOTTO, L. 2019. IMEXnet: A Forward Stable Deep Neural Network. *ICML*.

LAGARIS, I.E., LIKAS, A., AND FOTIADIS, D.I. 1998. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks 9*, 5, 987–1000.

LONG, Z., LU, Y., AND DONG, B. 2018. PDE-Net 2.0: Learning PDEs from Data with A Numeric-Symbolic Hybrid Deep Network.

LU, Y., ZHONG, A., LI, Q., AND DONG, B. 2018. Beyond Finite Layer Neural Networks: Bridging Deep Architectures and Numerical Differential Equations. *ICML*, 3282–3291.

RAISSI, M. 2018. Deep Hidden Physics Models: Deep Learning of Nonlinear Partial Differential Equations. *Journal of Machine Learning Research 19*, 1–24.

RUDY, S.H., BRUNTON, S.L., PROCTOR, J.L., AND KUTZ, J.N. 2017. Data-driven discovery of partial differential equations. *SCIENCE ADVANCES | 3* no. *4*, April.

RUTHOTTO, L. AND HABER, E. 2018. Deep Neural Networks Motivated by Partial Differential Equations. *https://arxiv.org/abs/1804.04272*, 2–10.

SIRIGNANO, J. AND SPILIOPOULOS, K. 2018. DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics 375*, Dms 1550918, 1339–1364.