Call for a fully funded PhD position in Artificial Intelligence

DEEP LEARNING FOR MODELING PHYSICAL DYNAMICS

Location? The research will be mainly conducted at Sorbonne University of Abu Dhabi (SUAD) at UAE with regular visits to Paris, France. The diploma will be granted from the Sorbonne University of Paris, France.

Profile of the successful candidate? Master of science in engineering, mathematics or computer science

Salary? The position is fully funded including tuition exemption, competitive stipend and support for travelling.

How to apply? Send 1) CV, 2) motivation letter and 3) grades obtained in master to Prof. Abdenour Hadid (abdenour.hadid@sorbonne.ae) and Prof. Patrick Gallinari (patrick.gallinari@sorbonne-universite.fr) by email.

Deadline for applying: 30/10/2021 **Starting date**: 1 January 2022

Scope

Deep learning offers a new data driven approaches to the modeling of dynamical systems underlying natural observations (see Reichstein *et al.*, 2019 for an overview on that matter). This has recently given rise to new and prolific research topics focused on exploiting deep learning methods for modeling spatio-temporal dynamics. The use of deep learning and data-driven approaches for modeling natural phenomena from physical observations however suffers from limitations such as the lack of generalization, robustness and physical plausibility. Modeling natural phenomena from physical observations indeed raises new challenges for machine learning and deep learning. Being the result of multiple interacting physical processes, the observed phenomena can be extremely complex. The data are heterogeneous, noisy, and even when plentiful they usually represent scarce and partial information about the underlying process. So, the main question which arises is: how to build accurate and fast machine learning models from this limited training data.

Scientific Objectives

In this research, we plan to explore new solutions using deep learning to cope with the challenges in modeling dynamic systems. Our focus will be on Geosystems and geoscience applications (Reichstein *et al.*, 2019, Sun, A. Y., Yoon, H., Shih, C. Y., & Zhong, Z. 2021) and more particularly, fluid flows through porous media (Tang, M., Liu, Y., & Durlofsky, L. J. 2020). Our objective is to speed up the execution time of the expensive fluid flow simulations that usually take hours to run to forecast hydrocarbon recovery or to predict CO2 migration in the subsurface in CO2 storage. Speeding up the reservoir modelling workflow and simulation time, that usually take years to build and hours to run, can have dramatic effects on business decisions and risk analysis. Among the research directions:

- **Incorporating physical priors:** Instead of fully relying on data for capturing the underlying complex phenomena, it would be appealing to incorporate physical knowledge and priors into data driven deep learning systems. A related challenge is to develop physically meaningful statistical models.

- **Deep fusion**: In real word problems, data are heterogeneous, noisy, and often coming from multiple sensors at different resolutions. Traditional multimodal data fusion methods cannot properly capture these variabilities since these are shallow models that cannot learn well the intrinsic representation of data. So, deep models efficiently combining different sources of data are needed (Gao *et al*, 2020) and will be investigated.

- **Learning from synthetic data**. Physically-based complex simulators are used in many different fields. They allow us to generate plentiful of relatively realistic data that can be used for training models while covering a variety of operational conditions. A great challenge is to learn how to transfer deep learning models learned from these simulations to real world observations. Recent works have addressed this issue of learning from simulations to build a fast surrogate to address uncertainty quantification (Yinhao Zhu and Nicholas Zabaras, 2018) or for oil reservoirs history matching (Tang, M., Liu, Y., & Durlofsky, L. J. 2020).

- **On the use of transformers:** Recent studies (Lin *et al*, 2021) showed that Transforms can outperform conventional deep learning models in many applications such as natural language processing, computer vision, and audio processing. A transformer is basically a deep learning model that adopts the mechanism of attention, weighing the influence of different parts of the input data. It is hence appealing to adapt the mechanism of Transformers to our specific downstream tasks and applications.

Experiments and Validation

As a use case of complex physical phenomena, we will consider forecasts of hydrocarbons production and also forecast on CO2 migration in CO2 storage. The research team from TotalEnergies in France HQ will participate to the prioritization of the main research directions and will have the responsibility of data exchange andtesting and validation of techniques that will be developed during the project (particularly those involving geoscience applications). TotalEnergies will also provide datasets for the oil & gas applications and will help to scale up solutions to industrial case studies.

References

- Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat (2019), Deep learning and process understanding for data-driven Earth system science, *Nature*, *566*(7743), 195–204, doi:10.1038/s41586-019-0912-1
- Sun, A. Y., Yoon, H., Shih, C. Y., & Zhong, Z. (2021). Applications of physics-informed scientific machine learning in subsurface science: A survey. arXiv preprint arXiv:2104.04764.
- Zhu, Y. and Zabaras, N. 2018. Bayesian deep convolutional encoder decoder networks for surrogate modeling and uncertainty quantification. J. Comput. Phys. 366, C (August 2018), 415-447.
- Tang, M., Liu, Y., & Durlofsky, L. J. (2020). A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. Journal of Computational Physics, 413, 109456.
- Jing Gao, Peng Li, Zhikui Chen, Jianing Zhang, A Survey on Deep Learning for Multimodal Data Fusion, Neural Computation Volume 32 Issue 5 May 2020 pp 829-864
- Tianyang Lin, Yuxin Wang, Xiangyang Liu, Xipeng Qiu, A Survey of Transformers, https://arxiv.org/abs/2106.04554, 2021

Ph.D position specifications

- Selected applicant for this position will be full time student seeking a Ph.D. in Artificial Intelligence. The admission is contingent upon approval form the Sorbonne University, Paris. The diploma will also be granted from the Sorbonne University, Paris.
- This position is intended for applicants from any country holding a master's degree or equivalent, or an engineering degree conferring the master's degree.
- The Ph.D position is fully funded including tuition exemption, stipend and support for travelling.
- The Ph.D position should be started as soon as possible and not later the **1 January 2022.**
- The **deadline to apply is 30 October 2021**

Knowledge in the following areas is required:

- Strong background in mathematics and statistics
- Good background in AI techniques: Machine Learning: Linear regression, discriminant functions, supervised and unsupervised learning, deep learning, classification, pattern recognition, reinforcement learning, and neural networks.
- knowledge of geoscience will be an advantage.

Required technical skills:

- Ability to communicate effectively in English, both orally and in writing.
- Strong skills in programming languages such as Python, MATLAB, C++. Skills in Pytorch, TensorFlow, are a plus.
- Strong background in mathematical modeling and computer simulations.
- Ability to develop and deliver presentations.
- Ability to work effectively with a multiple diverse community.
- Strong commitment and perseverance to achieve assigned tasks and meet the deadlines.

Admission requirements:

- Minimum score of 550 in TOEFL, 6.5 in IELTS, or their equivalent.
- Minimum CGPA of 3.5 in the Master degree (or its equivalent). Major should be in Computer Science, Computer engineering, Mathematics or any closely related field.
- Minimum CGPA of 3.0 in the Bachelor degree (or its equivalent). Major should be in Computer Science, Computer engineering, Mathematics or any closely related field.
- Three recommendation letters from academic faculty or profession superiors.

Position and Working Environment

The PhD studentship is a three-year position starting in January 2022. The PhD candidate will work mainly at Sorbonne University of Abu Dhabi under the supervision of Prof. Abdenour Hadid. The candidate will be co-supervised by Prof. Patrick Gallinari from Sorbonne University of Paris and will work closely with a research team from TotalEnergies R&D in Pau.

Interested? **How to apply:** Send 1) CV, 2) motivation letter and 3) grades obtained in master to Prof. Abdenour Hadid (abdenour.hadid@sorbonne.ae) and Prof. Patrick Gallinari (patrick.gallinari@sorbonne-universite.fr) by email before **30 October 2021**.