Internship proposal late-spring/summer 2023

Scale invariant neural networks

Key words: Machine Learning, Symmetries, Extrapolation, Scale Invariance, Critical Phenomena, Phase Transition, Equivariance

1 Context

One of the main ingredients in the success of deep learning has been the advent of Convolution Neural Networks (CNNs)[1] thanks to Yann Lecun, whose original motivation was to enforce translation invariance in neural networks (NNs). As a byproduct he obtained a reduction of the number of parameters to learn, from quadratic scaling with input dimension to linear, and a natural hierachical layer structure, each layer being interpreted as a representation on the input layer at some scale. These networks have been extended in many ways, in particular to Graph Neural Networks (GNNS) [2], to adapt the idea of CNNs to non-regular lattices.

In addition, the way that NNs deal with symmetries have been generalized and formalized by the notion of equivariance [3], such that one can maintain in all hidden layers a representation that fulfills the desired symmetry (present in the input space), like rotations, or permuations. The symmetries enforced in NN architectures represent additional information, a kind of "expert-knowledge" or “inductive bias”, which comes in addition to the training data and ensures a better generalization when used on unseen data. Some of us are currently actively working on this topic [4].

As a matter of fact neural networks are known to be good at interpolating data, thanks to their ability to find smooth-ish functions that interpolate the input-output data relationship even in very complex, high dimensional spaces. Concerning extrapolation, i.e. the ability to regress a function on a large region of the input space outside of the training region, standard NNs are doomed to fail, as they are not designed for it.

2 Internship’s goals and details

We would like to investigate the possibility of extrapolation with neural networks, by exploring a common situation in statistical physics, where a symmetry with respect to a change of scale holds, and see how this symmetry could be inserted in the design of the NN. In a sense, this goes in the line of bringing new symmetries in NNs, here in a way that can be seen as allowing the NN to extrapolate.

Concretley, to start, the goal would be to build a model able to generalize its predictions to scales not present in the training set, or present with only very few data points. For this we would consider a test-bed which consists of avalanche process, typically a 2-d sandpile model, presenting self-organized criticality [5], i.e. for which fluctuations have no scale, but the model has memory, so that formally prediction are known to be possible.

This is clearly of interest for earthquakes or solar storms predictions [6], where from observing small and frequent eruptions, we want to be able to anticipate rare and cataclysmic ones, generally not present in the training data.

3 Further reading

[1] J.Wu, “*Introduction to convolution neural networks*”, (2017)

[2] Battaglia, ..., Pascanu, “*Relational inductive biases, deep learning, and graph networks*”, (2018)

[3] Cohen, Welling, "*Group equivariant convolutional networks*", ICML (2016)

[4] Pezzicoli, Charpiat, Landes, "*SE (3)-equivariant Graph Neural Networks for Learning Glassy Liquids Representations*" (2022)

[5] Paczuski, Maslov, Bak, "*Avalanche dynamics in evolution, growth, and depinning models*", Phys.Rev. E (1996)

[6] Charbonneau, McIntosh, Liu and Bogdan, "*Avalanche models for solar flares*", Solar physics (2001)

Expected abilities: Good knowledge of Machine Learning, ideally Deep Learning too, is required. Some knowledge of critical phenomena (phase transitions) is expected, at least from an M2 Physics student. What matters too is motivation!

Learned Skills: Deepening of DL understanding, progress in pytorch, discovering sandpile models, understanding of equivariance in general and scale-invariance in particular too.

Duration: The preferred duration would be of at least 4 months (paid internship). There is no PhD funding already connected to this offer, but PhD can be considered.

Lab: LISN, Université Paris-Saclay

Team: A & O (Algorithmes et Optimisation), INRIA team: TAU

Advisors:
Cyril Furtlehner (Cyril.Furtlehner@inria.fr), CR INRIA
François Landes (francois.landes@inria.fr), MCF U-P-Saclay.

Head of Lab: Sophie Rosset

Location: France, Gif-sur-yvette (Plateau de Saclay), Batiment 660 (Digiteo building)

Website: http://lptms.u-psud.fr/francois-landes/internships-phd/