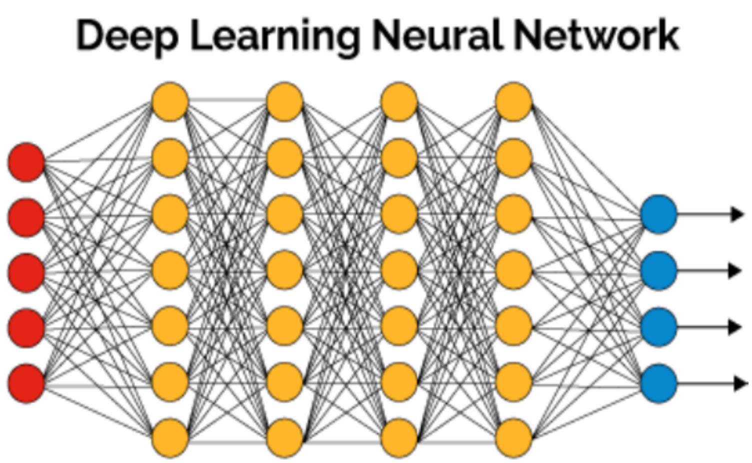
# Postdoc position on Machine Learning & Statistical Physics

Machine learning is a fascinating subject both from the theoretical and practical point of view. Although deep neural networks were introduced decades ago, only recently their implementation became possible thanks to an increased computational power combined with a huge amount of data (used as learning set). This lead to a real revolution that is still bursting right now [*Deep Learning* Y. LeCun, Y. Bengio, G. Hinton Nature 521 (2015)]. Almost every day a new ground-breaking application of machine learning--thought to be impossible just a few years ago--comes out. Despite their success, the theoretical understanding of deep neural networks is quite poor. Practitioners have developed recipes to construct and train them but fundamental questions remain opened. Answering them not only has the potentiality of leading to great improvements as well as avoiding dramatic pitfalls but it is a fascinating scientific subject at the boundary between high-dimensional statistics, theoretical physics, mathematics and computer science. The aim of this project is to theoretically analyze deep neural networks by studying simplified models that retain the essential ingredients and that are amenable to a full theoretical analysis.
In a nutshell deep neural networks use a very non-linear function of a very large number (e.g. $10^8$) of variables, called weights, to learn from a very large (e.g. $10^9$) number of data. A standard example is image recognition. The learning process is performed by minimizing a loss function: one has to find the weights such that the sum over the learning data set of the prediction errors is minimized.




*Example of an artificial neural network with four hidden layers in yellow, one input layer in red and one output layer in blue.*

Understanding the reason for the exceptional performance of DNNs is still a very open and central question at the cross-road between several fields: computer science, math and statistical physics. In fact, the problem encountered in the learning process of a deep neural network has very strong similarity with problems central in statistical physics. The loss function is very much like an energy function of a disordered system: the quenched disorder is encoded in the learning data set, the degrees of freedom are the weights, the learning process is like finding a good minimum of the energy landscape.

It is conjectured that one of the main reasons of the success of DNNs is the interplay between the properties of the data that are to be learnt and the architecture (deep and hierarchical) of DNNs. Another very important piece of the puzzle is the optimization process: how stochastic gradient descent finds good minima and avoid spurious and bad ones? These are the main issues we wish to study during the postdoc both in simple models solvable by statistical physics methods and in numerical experiments on realistic systems.

**Profile:** The postdoc position is for strong, creative individuals with a background in statistical physics and/or machine learning.

**Application:** Candidates must send a CV, a motivation letter and have two recommendation letters sent to giulio.biroli@ens.fr

**Deadline:** January 1st 2021

The postdoc will work at the physics department of the Ecole Normale Supérieure (Paris) and will be involved in the activities of the Paris Research Institute on Artificial Intelligence (PRAIRIE), the CFM-ENS Data Science Chair and the data science center @ ENS.

Some references representative of G. Biroli's research work on Machine Learning

Baity-Jesi, Marco, et al. "Comparing dynamics: Deep neural networks versus glassy systems." *arXiv preprint arXiv:1803.06969* (2018), ICML 2018.
d'Ascoli, Stéphane, et al. "Finding the Needle in the Haystack with Convolutions: on the benefits of architectural bias." *arXiv preprint arXiv:1906.06766, Advances in Neural Information Processing Systems*. 2019.
Geiger, Mario, et al. "Scaling description of generalization with number of parameters in deep learning." *arXiv preprint arXiv:1901.01608* (2019).
Mannelli, Stefano Sarao, et al. "Who is Afraid of Big Bad Minima? Analysis of gradient-flow in spiked matrix-tensor models." *arXiv preprint arXiv:1907.08226, Advances in Neural Information Processing Systems*. 2019.
D'Ascoli et al. "*Double Trouble in Double Descent : Bias and Variance(s) in the Lazy Regime", https://arxiv.org/abs/2003.01054, ICML 2020*
F. Pellegrini and G. Biroli, *"An analytic theory of shallow networks dynamics for hinge loss classification*",
https://proceedings.neurips.cc/paper/2020/file/3a01fc0853ebeba94fde4d1cc6fb842a-Paper.pdf,
*Advances in Neural Information Processing Systems*. 2020