Can Deep Neural Networks fail with non-smooth functions?

Keywords: Neural networks, supervised learning, approximation capacity, smoothness classes

Context: Deep neural networks are known to outperform classical machine learning algorithms in numerous situations where the data exhibit a complex structure (images, videos, natural language processing, ...). However for tabular data the situation is completely different. Classical strategies such as random forests or boosting are known to achieve state-of-the-art performance when applied to tabular data.
At this stage, it raises the need for better understanding the approximation properties of the deep neural networks (*DNNs*), especially in the interpolation regime. This would shed a new light on possible reasons for this suboptimal behavior.
Let us also emphasize that the approximation properties of DNNs are of utmost interest for analyzing generative models based on diffusion processes.

Objectives: We consider the usual supervised framework (classification or regression). A first step of the internship consists in reviewing the already existing literature on the approximation properties of DNNs. A particular focus must be given to the following papers [1] and the following books [2, 3] where chapters are dedicated to the DNN analysis.
A second step of the intership will be dedicated to the study, from both theoretical and numerical perspectives, of the approximation capacity of DNNs where the regression function exhibits various smoothness properties. In particular, the main goal of the internship is to identify properties of the regression function for which DNNs outperform classical strategies.
Specifically, the objectives of this internship are to

1. review the existing literature;
2. implement MLP on several toy examples and compare its performance with random forests;
3. study its theoretical properties on classes of smooth and non-smooth regression functions.

We insist on the fact that *the balance of the internship between theory and numerical aspects will depend on the intern skills.*

Supervisors: Alain Célisse (SAMM, Université Paris 1 Panthéon-Sorbonne),
and Christophe Denis (SAMM, Université Paris 1 Panthéon-Sorbonne).

Required skills: M2 level trainee in statistics/machine learning/optimization. Python programming.

Practical information: Applicants should send a CV and transcripts of the last two years to alain.celisse@universite-paris1.fr and christophe.denis1@univ-paris1.fr

- April-September 2025
- Location: SAMM, Université Paris 1 Panthéon-Sorbonne
- Grant from SAMM

References

[1] Schmidt-Hieber, Johannes Nonparametric regression using deep neural networks with ReLU activation function, The Annals of Statistics 2020

[2] Gyorfi Laszlo A Distribution-Free Theory of Nonparametric Regression. Springer (2002)

[3] Bach Francis Learning theory from first principles. MIT press (2024)