University Grenoble-Alpes **GIPSA-lab, GSTATS Data Science Chair**
INTERNSHIP PROJECT **Romain COUILLET & Malik TIOMOKO**

2021 **Office D1178**

A Random Matrix Study of Transfer Learning

Project context and application

With the recent compelling showcases of deep learning in performing advanced classfication tasks, deep neural networks are given more and more attention these days. One particularly interesting task performed by these networks is called "transfer learning", also related to federated learning and multi-task learning, consisting in learning from a given task for which a large labelled database is available to learn and train another data-scarce task, or alternatively using an already trained network on a given task to learn another task.

Thanks to advanced techniques from random matrix theory, it has recently become clear that transfer learning can be smartly performed, fully understood and dramatically improved without resorting to these heavily complex and computationally intensive neural network structures [1]. Yet, the subject studied in [1] only concerns the trivial application of a linear regression-based transfer learning scheme, for both source and target data of the same size, and for identically source and target tasks.

The subject of this internship will be to extend the aforementioned analysis to the more challenging transfer learning of different tasks and different database types. The first step will consist in implementing existing solutions, including deep learning approaches, as a reference, before developing a novel random matrix analysis of the transfer learning framework. To this end, the student will exploit methods, techniques, and results developed in the recent years in the random matrix domain, such as [2-4].

The internship may lead to a PhD position opening as of October 2021.

Main steps

- Review of the literature on transfer and multi-task learning.
- Implementation of the main documented solutions.
- Theoretical analysis and development of improved solutions and guarantees using random matrix theory.

Associated domains: Random matrix theory, time series, signal processing, neural networks.

Requirements: Good coding skill in Matlab or Python, knowledge of the basics of random matrix theory, good understanding of general signal processing and machine learning concepts.

Location: The internship will take place at GIPSA-lab, University of Grenoble-Alpes, in the Grenoble area.

References

[1] M. Tiomoko, H. Tiomoko, R. Couillet, "Deciphering and Optimizing Multi-Task Learning: a Random Matrix Approach", (submitted to) International Conference on Learning Representation (ICLR'21), 2021.

[2] R. Couillet, F. Benaych-Georges, **"Kernel Spectral Clustering of Large Dimensional Data"**, Electronic Journal of Statistics, vol. 10, no. 1, pp. 1393-1454, 2016.
[3] C. Louart, Z. Liao, R. Couillet, **"A Random Matrix Approach to Neural Networks"**, The Annals of Applied Probability, vol. 28, no. 2, pp. 1190-1248, 2018.
[4] Z. Liao, R. Couillet, **"A Large Dimensional Analysis of Least Squares Support Vector Machines"**, IEEE Transactions on Signal Processing, vol. 67, no.4, pp. 1065-1074, 2018.