

# **Journée Statistique et Informatique pour la Science des Données à Paris Saclay**

## **Rapport sur les contributions**

ID de Contribution: 1

Type: **Non spécifié**

## Supervised Learning with Missing Values

*vendredi 5 février 2021 10:30 (40 minutes)*

Some data come with missing values. For instance, a survey's participant may ignore some questions. There is an abundant statistical literature on this topic, establishing for instance how to fit model without biases due to the missingness, and imputation strategies to provide practical solutions to the analyst. In machine learning, to build models that minimize a prediction risk, most work default to these practices. As we will see, these different settings lead to different theoretical and practical solutions.

I will outline some conditions under which machine-learning models yield the best-possible predictions in the presence of missing values. A striking result is that naive imputation strategies can be optimal, as the supervised-learning model does the hard work [1]. A challenge to fitting a machine-learning model is that there is a combinatorial explosion of possible missing-values patterns such that even when the output is a linear function of the fully-observed data, the optimal predictor is complex [2]. I will show how the same dedicated neural architecture can approximate well the optimal predictor for multiple missing-values mechanisms, including difficult missing-not-at-random settings [3].

[1] Josse, J., Prost, N., Scornet, E., & Varoquaux, G. (2019). On the consistency of supervised learning with missing values. arXiv preprint arXiv:1902.06931.

[2] Le Morvan, M., Prost, N., Josse, J., Scornet, E., & Varoquaux, G. (2020). Linear predictor on linearly-generated data with missing values: non consistency and solutions. AISTATS 2020.

**Orateur:** VAROQUAUX, Gaël (INRIA Parietal)

ID de Contribution: 2

Type: **Non spécifié**

## Analysis of Gradient Descent on Wide Two-Layer Neural Networks

*vendredi 5 février 2021 11:20 (40 minutes)*

Artificial neural networks are a class of “prediction” functions parameterized by a large number of parameters – called weights – that are used in various machine learning tasks (classification, regression, etc). Given a learning task, the weights are adjusted via a gradient-based algorithm so that the corresponding predictor achieves a good performance on a given training set. In this talk, we propose an analysis of gradient descent on wide two-layer ReLU neural networks for supervised machine learning tasks, that leads to sharp characterizations of the learned predictor. The main idea is to study the dynamics when the width of the hidden layer goes to infinity, which is a Wasserstein gradient flow. While this dynamics evolves on a non-convex landscape, we show that its limit is a global minimizer if initialized properly. We also study the “implicit bias” of this algorithm when the objective is the unregularized logistic loss: among the many global minimizers, we show that it selects a specific one which is a max-margin classifier in a certain functional space. We finally discuss what these results tell us about the generalization performance and the adaptivity to low dimensional structures of neural networks. This is based on joint work with Francis Bach.

**Orateur:** CHIZAT, Lenaïc (LMO)

ID de Contribution: 3

Type: **Non spécifié**

## Maximum Entropy Distributions for Image Synthesis under Statistical Constraints

*vendredi 5 février 2021 14:00 (40 minutes)*

The question of texture synthesis in image processing is a very challenging problem that can be stated as followed: given an exemplar image, sample a new image that has the same statistical features (empirical mean, empirical covariance, filter responses, neural network responses, etc.). Exponential models then naturally arise as distributions satisfying these constraints in expectation while being of maximum entropy. Now the parameters of these exponential models need to be estimated and samples have to be drawn. I will explain how these can be done simultaneously through the SOUL (Stochastic Optimization with Unadjusted Langevin) algorithm. This is based on a joint work with Valentin de Bortoli, Alain Durmus, Bruno Galerne and Arthur Leclaire.

**Orateur:** DESOLNEUX, Agnès (Centre Borelli)

ID de Contribution: 4

Type: **Non spécifié**

## Deep Unfolding of a Proximal Interior Point Method for Image Restoration

*vendredi 5 février 2021 12:00 (40 minutes)*

Variational methods have started to be widely applied to ill-posed inverse problems since they have the ability to embed prior knowledge about the solution. However, the level of performance of these methods significantly depends on a set of parameters, which can be estimated through computationally expensive and time-consuming processes. In contrast, deep learning offers very generic and efficient architectures, at the expense of explainability, since it is often used as a black-box, without any fine control over its output. Deep unfolding provides a convenient approach to combine variational-based and deep learning approaches. Starting from a variational formulation for image restoration, we develop iRestNet [1], a neural network architecture obtained by unfolding an interior point proximal algorithm. Hard constraints, encoding desirable properties for the restored image, are incorporated into the network thanks to a logarithmic barrier, while the barrier parameter, the stepsize, and the penalization weight are learned by the network. We derive explicit expressions for the gradient of the proximity operator for various choices of constraints, which allows training iRestNet with gradient descent and backpropagation. In addition, we provide theoretical results regarding the stability of the network. Numerical experiments on image deblurring problems show that the proposed approach outperforms both state-of-the-art variational and machine learning methods in terms of image quality.

[1] C. Bertocchi, E. Chouzenoux, M.-C. Corbineau, J.-C. Pesquet and M. Prato. Deep Unfolding of a Proximal Interior Point Method for Image Restoration. *Inverse Problems*, vol. 36, pp. 034005, 2020.

**Orateur:** CHOUZENOUX, Emilie (CVN)

ID de Contribution: 5

Type: **Non spécifié**

# Input Similarity from the Neural Network Perspective

*vendredi 5 février 2021 14:50 (40 minutes)*

Given a trained neural network, we aim at understanding how similar it considers any two samples. For this, we express a proper definition of similarity from the neural network perspective (i.e. we quantify how undissociable two inputs A and B are), by taking a machine learning viewpoint: how much a parameter variation designed to change the output for A would impact the output for B as well?

We study the mathematical properties of this similarity measure, and show how to estimate sample density with it, in low complexity, enabling new types of statistical analysis for neural networks. We also propose to use it during training, to enforce that examples known to be similar should also be seen as similar by the network.

We then study the self-denoising phenomenon encountered in regression tasks when training neural networks on datasets with noisy labels. We exhibit a multimodal image registration task where almost perfect accuracy is reached, far beyond label noise variance. Such an impressive self-denoising phenomenon can be explained as a noise averaging effect over the labels of similar examples. We analyze data by retrieving samples perceived as similar by the network, and are able to quantify the denoising effect without requiring true labels.

**Orateur:** CHARPIAT, Guillaume (LRI)

ID de Contribution: 6

Type: **Non spécifié**

## Deep Neural Network for Audio and Music Transformations

*vendredi 5 février 2021 15:30 (40 minutes)*

We will first discuss how deep learning techniques can be used for audio signals. To that aim, we will recall some of the important characteristics of an audio signal and review some of the main deep learning architectures and concepts used in audio signal analysis. We will then illustrate some of these concepts in more details with two applications, namely informed singing voice source separation and music style transfer.

**Orateur:** RICHARD, Gaël (Télécom Paris)