

## Learning with Missing Values: Theoretical Insights and Application to Health Databases

*mercredi 3 avril 2024 11:20 (50 minutes)*

Missing values are ubiquitous in many fields such as health, business or social sciences. To date, much of the literature on missing values has focused on imputation as well as inference with incomplete data. In contrast, supervised learning in the presence of missing values has received little attention. In this talk I will explain the challenges posed by missing values in regression and classification tasks. In practice, a common solution consists in imputing the missing values prior to learning. I will show how different baseline methods for handling missing values compare on several large health databases with naturally occurring missing values. We will then examine the theoretical foundations of Impute-then-Regress approaches. Finally, I will present a neural network architecture for learning with missing values that goes beyond the two-stage Impute-then-Regress approaches.

**Orateur:** LE MORVAN, Marine (INRIA, Saclay)