





## High Dimensional Nonstationary Time Series IRTG 1792 Short Course

## **Erwan Scornet**

## Several views on the consistency of random forests

Tree-based methods are among state-of-the-art methods to handle tabular data sets. Among them, random forests (Breiman, 2001) are appealing, notably due to their parallel nature and good predictive performances. The theory behind random forests remains obscure. Historically, attempts to understand their behavior first focus on stylized random forests, whose construction is independent of the data set. Even if these stylized versions depart greatly from that of the original algorithm, their analyses reveal some interesting properties, that are thought to hold for Breiman's forests. In this short course, we will go through several properties of random forests and illustrate them with some properly chosen random forest models. The connection with kernel methods and its impact on the algorithm consistency will also be discussed.



Erwan Scornet is an assistant professor at the Center for Applied Mathematics (CMAP) in Ecole Polytechnique near Paris. His research interests focus on theoretical statistics and Machine Learning with a particular emphasis on nonparametric estimates. He did his PhD thesis on a particular algorithm of Machine Learning called random forests, under the supervision of Gérard Biau (LSTA - Paris 6) and Jean-Philipe Vert (Institut Curie).

## February 10, 2022 | 10:00-13:00 (CET) | DOR1+Zoom

