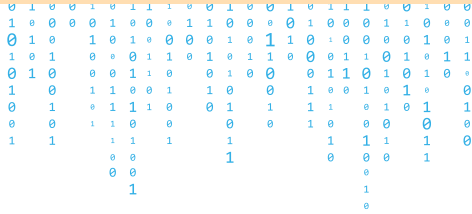
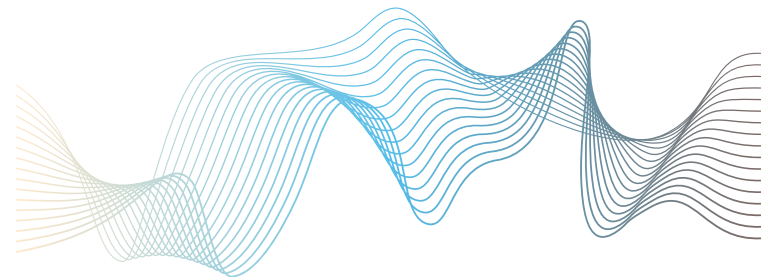


# Introduction to Random Forests and Application in High-Dimensional Data Analysis



This course delves into **the power of Random Forests**, a machine learning technique that excels at analyzing data with many features, often called **high-dimensional data**. Imagine a strong forest – that's the idea behind Random Forests. **They combine the predictions of multiple decision trees, like a simplified flow chart, to make accurate and reliable forecasts.** The course will likely explore how Random Forests are built and why they're particularly useful for high-dimensional data. **You'll learn how they overcome challenges like overfitting and identify the most important features influencing the outcome.** By the end, you'll gain valuable skills for applying Random Forests to real-world problems in various fields, from genomics and finance to image recognition.



## Prerequisites

This course is open to PhD students specializing in statistics, machine learning or epidemiology, with basic knowledge in statistics (inference principle, classical estimators), and in R (or in a close language like Python, matlab)



**BORDEAUX  
POPULATION  
HEALTH** | Research  
Center - U1219



**TSPED**  
SCHOOL OF PUBLIC HEALTH

université  
de **BORDEAUX**

**Inserm**



Sciences Po  
Bordeaux

*Inria*



Université  
**BORDEAUX  
MONTAIGNE**

# Introduction to Random Forests and Application in High-Dimensional Data Analysis

## PROGRAM

This is a **6-hour workshop** on **decision trees and random forests for analyzing high-dimensional data using R**.

**The morning session** will cover the fundamentals of decision trees and how to use them in R. **In the afternoon**, you'll learn about random forests, a powerful technique for high-dimensional data analysis, and get hands-on experience applying it to real genomic data using R. **This training course combines** a presentation of the general principles of the method, and practical work using the R software with application of the method to real dataset (genomic data)

## Training Objectives

**This training course aims to equip its participants with :**

- An understanding the general mechanism of random forests
- To apply random forests in order to analyze high-dimensional dataset

**Registration deadline :**

**Date :**

**Schedule :**

**Place :**