

# Module overview

## Contents

- [What you will learn](#)
- [Before getting started](#)
- [Objectives and time schedule](#)

## What you will learn

This module will go into details regarding algorithms that are combining several models together, also called ensemble of models. We will present two families of such techniques: (i) based on bootstrapping and (ii) based on boosting. We will present bagging and random forest that belong to the former strategy and AdaBoost and gradient boosting decision tree that belong to the later strategy. Finally, we will go into details regarding the hyperparameters allowing to tune these models and compare them between models.

## Before getting started

The required technical skills to carry on this module are:

- skills acquired during the “The Predictive Modeling Pipeline” module with basic usage of scikit-learn;
- skills acquired during the “Selecting The Best Model” module, mainly around the concept of underfit/overfit and the usage of cross-validation in scikit-learn;
- skills acquired during the modules “Linear Models” and “Decision Tree Models”.

## Objectives and time schedule

The objective in the module are the following:

- understanding the principles behind bootstrapping and boosting;
- get intuitions with specific models such as random forest and gradient boosting;
- identify the important hyperparameters of random forest and gradient boosting decision trees as well as their typical values.

The estimated time to go through this module is about 6 hours.

---

By scikit-learn developers

© Copyright 2022.

[Join the full MOOC for better learning!](#)

Brought to you under a [CC-BY License](#) by [Inria Learning Lab](#), [scikit-learn @ La Fondation Inria](#), [Inria Academy](#), with many thanks to the [scikit-learn](#) community as a whole!