

Module overview

Contents

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

What you will learn

In the previous module, we presented the general cross-validation framework and used it to evaluate models' performance. However, this is important to keep in mind that some elements in the cross-validation need to be decided depending on the nature of the problem: (i) the cross-validation strategy and (ii) the evaluation metrics. Besides, it is always good to compare the models' performance with some baseline model.

In this module, we present both aspects and give insights on when to use a specific cross-validation strategy and a metric. In addition, we will also give some insights regarding how to compare a model with some baseline.

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.

Objectives and time schedule

The objective in the module are the following:

- understand the necessity of using an appropriate cross-validation strategy depending on the data;
- get the intuitions behind comparing a model with some basic models that can be used as baseline;
- understand the principles behind using nested cross-validation when the model needs to be evaluated as well as optimized;
- understand the differences between regression and classification metrics;
- understand the differences between metrics.

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!