Module overview

Contents

What you will learn

In this module, will go further into details regarding models that use linear parametrization.
We will see how to use this family of models for both classification and regression problems.
Besides, we will explain how to fight over-fitting using regularization.
Finally, we will show how linear models can be used with data presenting non-linearity.

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

In this module, your objectives are to:

- understand the linear models parametrization;
- understand the implication of linear models in both regression and classification;
- get intuitions of linear models applied in higher dimensional dataset;
- understand the effect of regularization and how to set it;
- understand how linear models can be used even with data showing non-linear relationship with the target to be predicted.

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