

Improving Performance

1. Improving performance

00:00 - 00:00

Now you know how to choose between classification, regression or unsupervised learning techniques and how to evaluate your results.

2. Machine learning workflow

00:00 - 00:00

After we evaluate our model, we have to decide if the performance is good enough. This video will cover how you can improve your model if you're unsatisfied with its performance.

3. Several options

00:00 - 00:00

There are several options, but we will focus on three of them. They all have scary names, like dimensionality reduction, hyperparameter tuning, ensemble methods, but are actually pretty straightforward.

4. Dimensionality reduction

00:00 - 00:00

Let's start with dimensionality reduction. A dimension denotes the number of features in your data, so dimensionality reduction essentially means reducing the number of features.

5. Dimensionality reduction: example

00:00 - 00:00

It may seem counter-intuitive to remove features. The more information we have, the better our predictions, right? Well, no. The truth is, some of these features might be completely irrelevant. For example, if we're trying to predict how long it will take us to go to the office, the time of day and the weather are interesting, but how many glasses of water we drank yesterday is unlikely to be very useful.

6. Dimensionality reduction: example

00:00 - 00:00

Some features might be highly correlated and carry similar information. We could keep only one feature and still have most information. For example, height and shoe size are highly correlated. Tall people are very likely to have a large shoe size. We could keep only the height without losing much additional information. We could also collapse multiple features into just one underlying feature. If we have two features, height and weight, we can calculate one Body Mass Index feature instead.

7. Hyperparameter tuning

00:00 - 00:00

Let's move on to the second option for improving model performance, hyperparameter tuning. A machine learning model is like a music production console. Depending on whether you're mixing

8. Hyperparameter tuning

00:00 - 00:00

pop,

9. Hyperparameter tuning

00:00 - 00:00

hip hop,

10. Hyperparameter tuning

00:00 - 00:00

or heavy metal, you don't apply the same settings to the guitar, the bass, the drums and the vocals. Some settings will sound better with certain genres than others.

11. Hyperparameter tuning

00:00 - 00:00

Well, in our case the genre is the dataset and the instruments settings are the hyperparameters. Depending on your dataset, different hyperparameter values will give better or inferior results.

12. Hyperparameter tuning: example

00:00 - 00:00

Let's look at an example. When we trained the SVM in the first lesson of this chapter, we switched from a straight line to a curved one. That's because we tuned the hyperparameter "kernel" from "linear" to "polynomial". There are many more hyperparameters that can be tuned in the SVM model. Playing with different values for each hyperparameter will impact our model's performance. You don't have to just guess combinations at random, there are structured ways to find the optimal values but that is outside the scope of this course.

13. Ensemble methods

00:00 - 00:00

The last option we want to cover is ensemble methods. This is a technique that combines several models in order to produce one optimal model.

14. Ensemble methods: classification

00:00 - 00:00

Imagine we're using three different models: models A, B and C. In a classification setting, we would use voting. If Model A and C say the student is accepted, and model B says she's not, then the observation is accepted, as it was the most common prediction amongst all models.

15. Ensemble methods: regression

00:00 - 00:00

In a regression setting, we use averaging. If Model A predicts a temperature of 5, Model B a temperature of 8, and Model C a temperature of 4, the observation gets assigned the average value, 5.67 degrees.

16. Let's practice!

00:00 - 00:00

And that's it! I hope you won't be scared next time you hear about dimensionality reduction, hyperparameter tuning or ensemble methods. Let's do a few more exercises, and onwards to Chapter 3!