Quiz M2.02

Note: For each question **make sure you select all of the correct options—** there may be more than one! Don't forget to use the sandbox notebook if you need.

Question 1 (1/1 point)

A model is overfitting when:

- a) both the train and test errors are high
- b) train error is low but test error is high ✔
- c) train error is high but the test error is low
- d) both train and test errors are low

EXPLANATION

solution: b)

The model has enough flexibility to make good predictions on the training set (including the noise) and therefore the train error is low. But such overfitting models fail to see the repeatable pattern that is useful to also make correct predictions on the test set.

You have used 1 of 1 submissions

Question 2 (1/1 point)

Assuming that we have a dataset with little noise, a model is underfitting when:

- b) train error is low but test error is high
- c) train error is high but the test error is low
- d) both train and test errors are low

EXPLANATION

solution: a)

Underfitting models are too constrained, even to make make correct predictions even on the training samples. Therefore their training error is high.

The test error is, on average, bounded below by the train error (it's harder to make good predictions that the model has never seen before). Therefore their test error is also high.

You have used 1 of 1 submissions

Question 3 (1/1 point)

For a fixed training set, by sequentially adding parameters to give more flexibility to the model, we are more likely to observe:

- [x] a) a wider difference between train and test errors ✔
- [ ] b) a reduction in the difference between train and test errors
- [ ] c) an increased or steady train error
- [x] d) a decrease in the train error ✔

✔

EXPLANATION

solution: a) d)

By giving more flexibility to the model, we can reduce underfitting (and therefore lower the train error) but we also risk overfitting and the test error might increase (or at least not decrease as as the train error) as result: the difference between the two kinds of errors will likely increase.

In the end, remember that what is the most important is to find a the best value of the parameter to get the best test error, and we do not care that much about the training error.

You have used 1 of 2 submissions

Question 4 (1/1 point)

For a fixed choice of model parameters, if we increase the number of labeled observations in the training set, are we more likely to observe:

- [ ] a) a wider difference between train and test errors
- [x] b) a reduction in the difference between train and test errors ✔
- [x] c) an increased or steady train error ✔
- [ ] d) a decrease in the train error

✔

Select all answers that apply

EXPLANATION

solution: b) c)

some underfitting).

You have used 1 of 2 submissions

Question 5 (1/1 point)

Polynomial models with a high degree parameter:

- [ ] a) always have the best test error (but can be slow to train)
- [ ] b) underfit more than linear regression models
- [x] c) get lower training error than lower degree polynomial models ✔
- [x] d) are more likely to overfit than lower degree polynomial models ✔

✔

Select all answers that apply

EXPLANATION

solution: c) d)

Polynomial models with a high degree parameters are more flexible and therefore likely to overift.

Linear models are polynomial models with degree 1 and are therefore less flexible than polynomials with higher degrees.

You have used 1 of 2 submissions

Question 6 (1/1 point)

a) True

b) False

EXPLANATION

solution: b)

It is almost always impossible to build machine learning models that are guaranteed to reach zero test error.

It is often the case that the input feature features are available are incomplete and do not allow us to make decisions that are accurate 100% of the time. Another way to frame this is that the target variable often depends on external factors not completely described by the observed variables, in which case one cannot make the correct prediction 100% of the time: there are some irreducible errors.

For the very simple deterministic tasks where the target variable is guaranteed to be fully determined by the input variables, then machine learning is usually not needed: one can often directly write the optimal decision rule directly as a Python program for instance.

You have used 1 of 1 submissions

YOUR EXPERIENCE

According to you, this whole 'Validation and learning curves' lesson was:

Too easy, I got bored

Adapted to my skills

Difficult but I was able to follow

Too difficult

To follow this lesson, I spent:

- **less than 30 minutes**
- **30 min to 1 hour**
- **1 to 2 hours**
- **2 to 4 hours**
- **more than 4 hours**
- **I don't know**

Submit

FORUM (EXTERNAL RESOURCE)

About the M2. Quiz M2.02 category

brospars Pedagogical team 27 Oct

(Replace this first paragraph with a brief description of your new category. This guidance will appear in the category selection area, so try to keep it below 200 characters.)

Use the following paragraphs for a longer description, or to establish category guidelines or rules:

- Why should people use this category? What is it for?
- How exactly is this different than the other categories we already have?
- What should topics in this category generally contain?
- Do we need this category? Can we merge with another category, or subcategory?