Quiz M1.02

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)

Why do we need two sets: a train set and a test set?

- [ ] a) to train the model faster
- [x] b) to validate the model on unseen data
- [ ] c) to improve the accuracy of the model

✔

Select all answers that apply

You have used 1 of 2 submissions

Question 2 (1/1 point)

The generalization performance of a scikit-learn model can be evaluated by:

- [x] a) calling `fit` to train the model on the **training set**, `predict` on the **test set** to get the predictions, and compute the score by passing the predictions and the true target values to some metric function
- [x] b) calling `fit` to train the model on the **training set** and `score` to compute the score on the **test set**

☐ d) calling `fit_transform` on the data and then `score` to compute the score on the **test set**

✔

Select all answers that apply

You have used 2 of 2 submissions

Question 3 (1/1 point)

When calling `cross_validate(estimator, X, y, cv=5)`, the following happens:

☑ a) `X` and `y` are internally split five times with non-overlapping test sets

☐ b) `estimator.fit` is called 5 times on the full `X` and `y`

☑ c) `estimator.fit` is called 5 times, each time on a different training set

☐ d) a Python dictionary is returned containing a key/value containing a NumPy array with 5 scores computed on the **train sets**

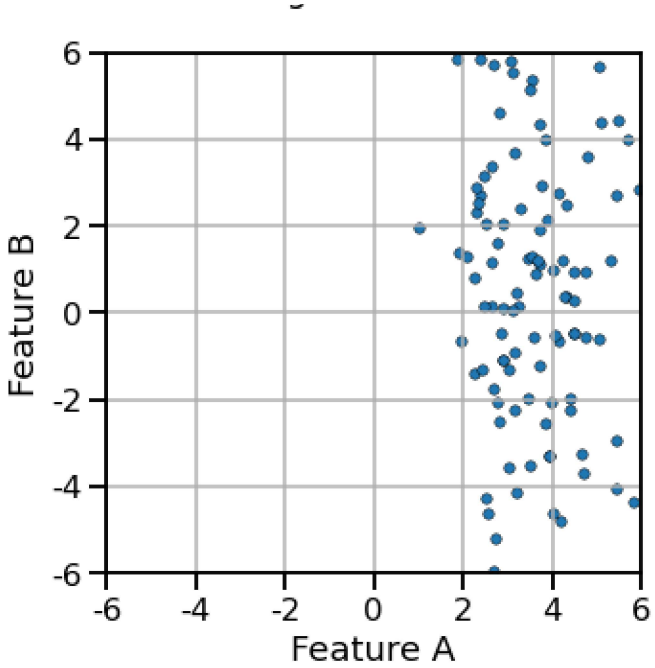
☑ e) a Python dictionary is returned containing a key/value containing a NumPy array with 5 scores computed on the **test sets**

✔

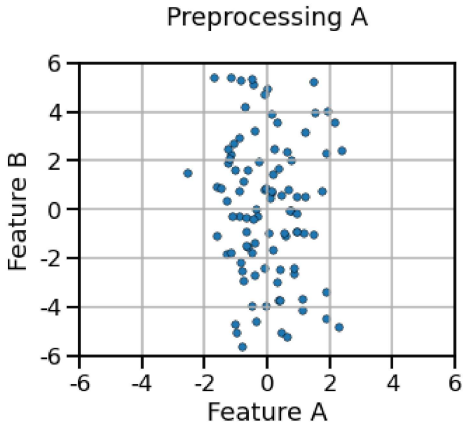
Select all answers that apply

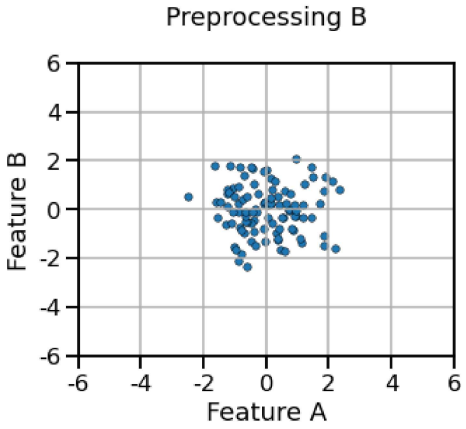
You have used 1 of 2 submissions

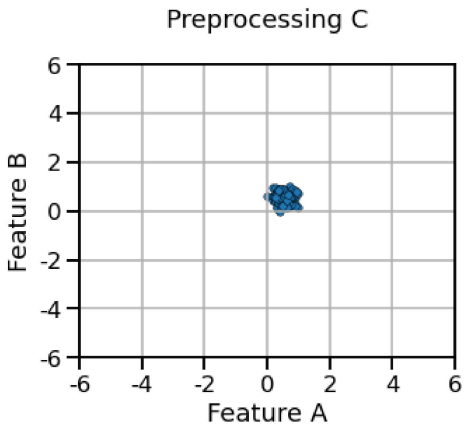
Question 4 (1/1 point)

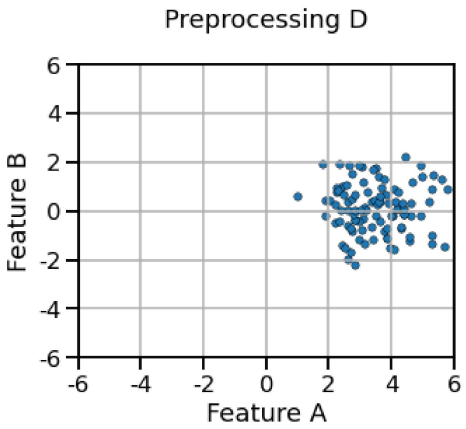


If we process the dataset using a `StandardScaler` with the default parameters, which of the following results do you expect:









- [x] b) Preprocessing B ✔
- [ ] c) Preprocessing C
- [ ] d) Preprocessing D

You have used 1 of 1 submissions

Question 5 (1/1 point)

Look at the plots and the answers of the previous question.
A `StandardScaler` transformer with the default parameter:

- [x] a) transforms the features so that they have similar ranges
- [ ] b) transforms the features to lie in the [0.0, 1.0] range
- [x] c) transforms feature values that were originally positive-only into values that can be negative or positive
- [x] d) can help logistic regression converge faster (fewer iterations)

✔

Select all answers that apply

You have used 1 of 2 submissions

Question 6 (1/1 point)

Cross-validation allows us to:

- [ ] a) train the model faster

- [x] c) estimate the variability of the generalization score

✔

Select all answers that apply

You have used 1 of 2 submissions

Question 7 (1/1 point)

`make_pipeline` (as well as `Pipeline`):

- [ ] a) runs a cross-validation using the transformers and predictor given as parameters
- [x] b) combines one or several transformers and a predictor
- [ ] c) tries several models at the same time
- [ ] d) plots feature histogram automatically

✔

Select all answers that apply

You have used 1 of 2 submissions

YOUR EXPERIENCE

According to you, this whole 'Fitting a scikit-learn model on numerical data' lesson of the course 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)

Topic	Replies	Last reply
About the M1. Quiz M1.02 category brospars 27 Oct	0	27 Oct brospars

There are no more M1. Quiz M1.02 topics. Ready to start a new conversation?

Terms and conditions