

Revision questions for Chapter 9

Last updated: December 1, 2022

If you are asked to define some notion, you should explain carefully all notation (if any) that you use in your definition. [Answers to some questions are given in blue.](#) All other answers can be found in the module notes (lecture slides or lab worksheets) provided on the module's Moodle page.

1. Consider the following code for estimating the prediction accuracy of the SVM on a training set `X_train`, `y_train` using feature normalization and parameter selection:

```
scaler = MinMaxScaler().fit(X_train)
X_train_scaled = scaler.transform(X_train)
X_test_scaled = scaler.transform(X_test)
param_grid = {'C': [0.01, 0.1, 1, 10, 100],
              'gamma': [0.01, 0.1, 1, 10, 100]}
grid = GridSearchCV(SVC(), param_grid=param_grid)
grid.fit(X_train_scaled, y_train)
print(grid.score(X_test_scaled, y_test))
```

Locate data leakage in this code, and explain why it may lead to an overoptimistic estimate. [Chapter 9, slides 4–8.](#)

2. Describe the `scikit-learn` class `Pipeline` paying particular attention to its `fit`, `predict`, and `score` methods. [Chapter 9, slides 11–14 and 19.](#)
3. List two benefits of using the class `Pipeline`. [Chapter 9, slide 14.](#)
4. Describe the use of the `scikit-learn` class `Pipeline` for parameter selection using a grid search. [Chapter 9, slides 15–18.](#)
5. Define the cross-conformal predictor based on a given inductive conformity measure. [Chapter 9, slides 22–25.](#)
6. Explain why the requirement of validity may be violated for cross-conformal predictors. [Chapter 9, slide 28.](#)
7. Make sure you can solve the exercise implicit in the example on slide 29 of Chapter 9. [Chapter 9, slides 29–37.](#)
8. Define the *calibration curve* for a set predictor. [Chapter 9, slides 39–40.](#)
9. Describe the class `KFold` in `scikit-learn`. [Lab Worksheet 9, Section 2.](#)
10. Describe the method `decision_function` in the class `SVC` and explain how it can be useful in implementing inductive conformal predictors and cross-conformal predictors. [See Lab Worksheet 9, Section 3, and Assignment 3 \(task 7\).](#) The method `decision_function` computes the inductive conformity measure. It is used as described in Chapter 6, slide 41, and Chapter 9, slide 24.

Similar lists of questions will be produced for all chapters of the module to help students in revision. There is no guarantee that the actual exam questions will be in this list, or that they will be in any way similar.