

Name:

## Short Quiz 2B

22 September 2025

**Question 1:** What is a test dataset? From test data  $(x_1, y_1), \dots, (x_T, y_T)$ , how do we estimate the performance of a predictor  $f$  for a cost  $c$ ?

**Answer:** A test dataset is used to estimate the performance of a predictor, with the formula

$$\frac{1}{T} \sum_{t=1}^T c(y_t, f(x_t)).$$

**Question 2:** What is underfitting for a model?

**Answer:** A model underfits if it cannot achieve a small train error: all the functions it contains approximate poorly the objective function of the task.

**Question 3:** True or false? Circle the right answer.

- T      F      Cross-validation is computationally expensive for large datasets.
- T      F      For a fixed model, using more training data reduces overfitting.
- T      F      The approximation error increases with the size of the model  $\mathcal{F}$ .

**Answer:** True — True — False.