Classifiers 2

Z. Straka, P. Švarný, J. Kostlivá

Today two examples:

1. Covid-19 testing example
2. Overfitting and underfitting

Covid-19 testing example

Covid-19 testing example

Source: Seznam zpravy
For details see a post from Jakub Steiner on Facebook.

Covid-19 testing example

Let's suppose that 0.5% of a population has already been infected by covid-19. Someone else bought covid-19 tests with specificity=0.9 (specificity $= \frac{TN}{TN+FP}$) and wants to test 2000 people from the population. How many of the tests will be false positive?

A: 10

B: 99

C: 199

D: 399

Covid-19 testing example

Let's suppose that 0.5% of a population has already been infected by covid-19. Someone else bought covid-19 tests with specificity=0.9 (specificity $= \frac{TN}{TN+FP}$) and wants to test 2000 people from the population. How many of the tests will be false positive?

C: $199 = (0.995 * 2000) * 0.1$

▶ $0.995 * 2000 = 1990$ people have not been infected
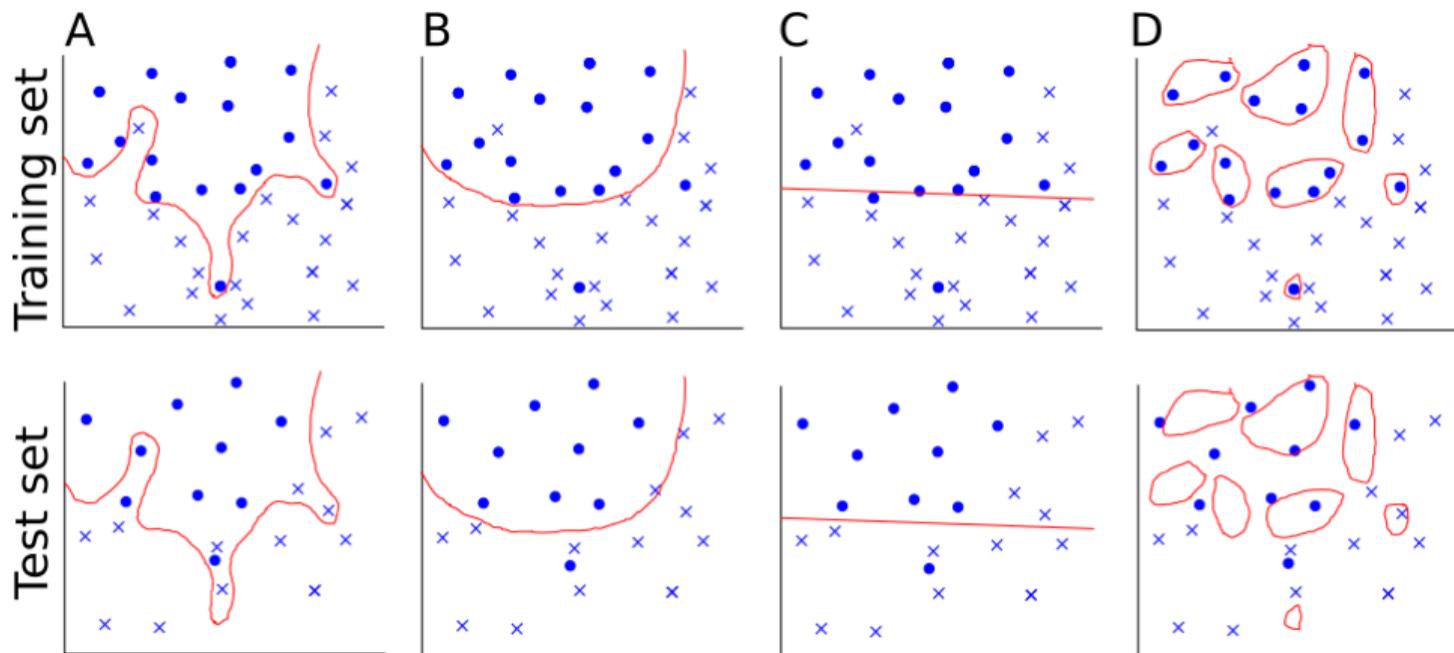▶ As specificity is 0.9, ten percent of all negative samples (i.e., TN+FP) are determined as false positive.

Overfitting and underfitting

Overfitting and underfitting
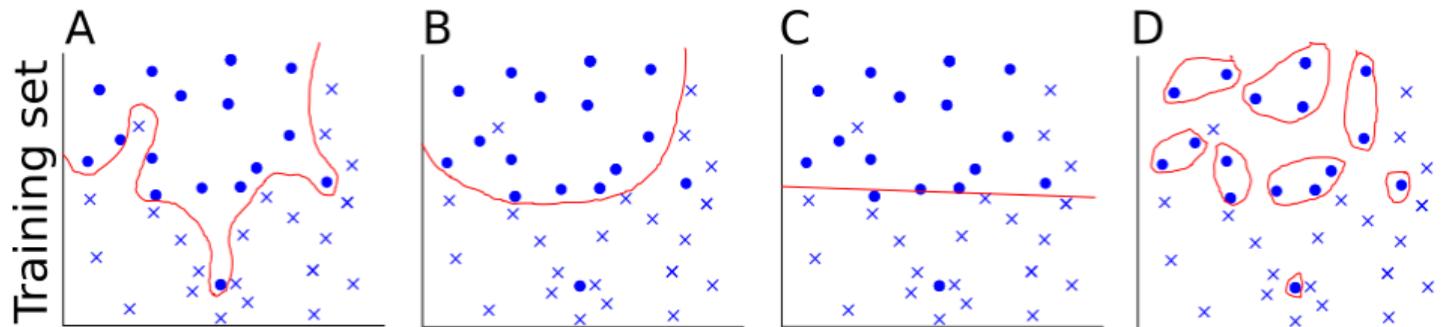
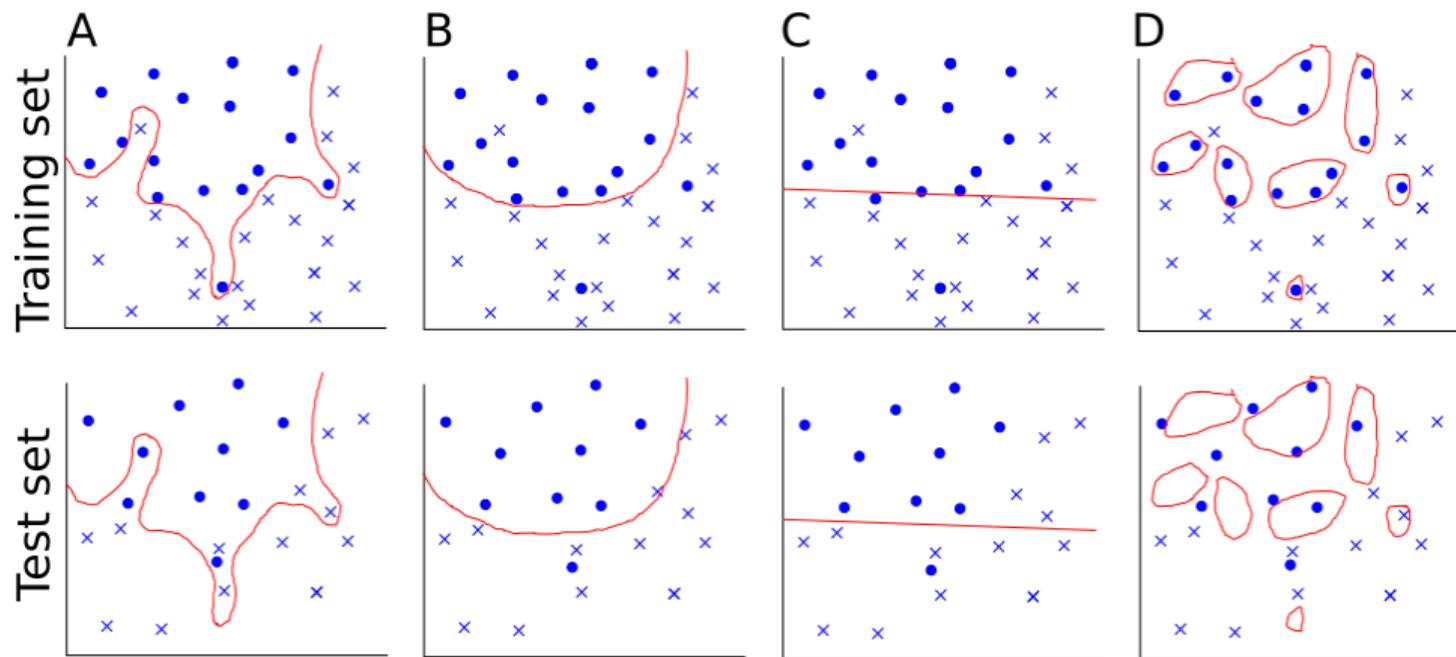Which classifier (A,B,C,D) is the best? Why?
Break-out rooms discussion..

Overfitting and underfitting
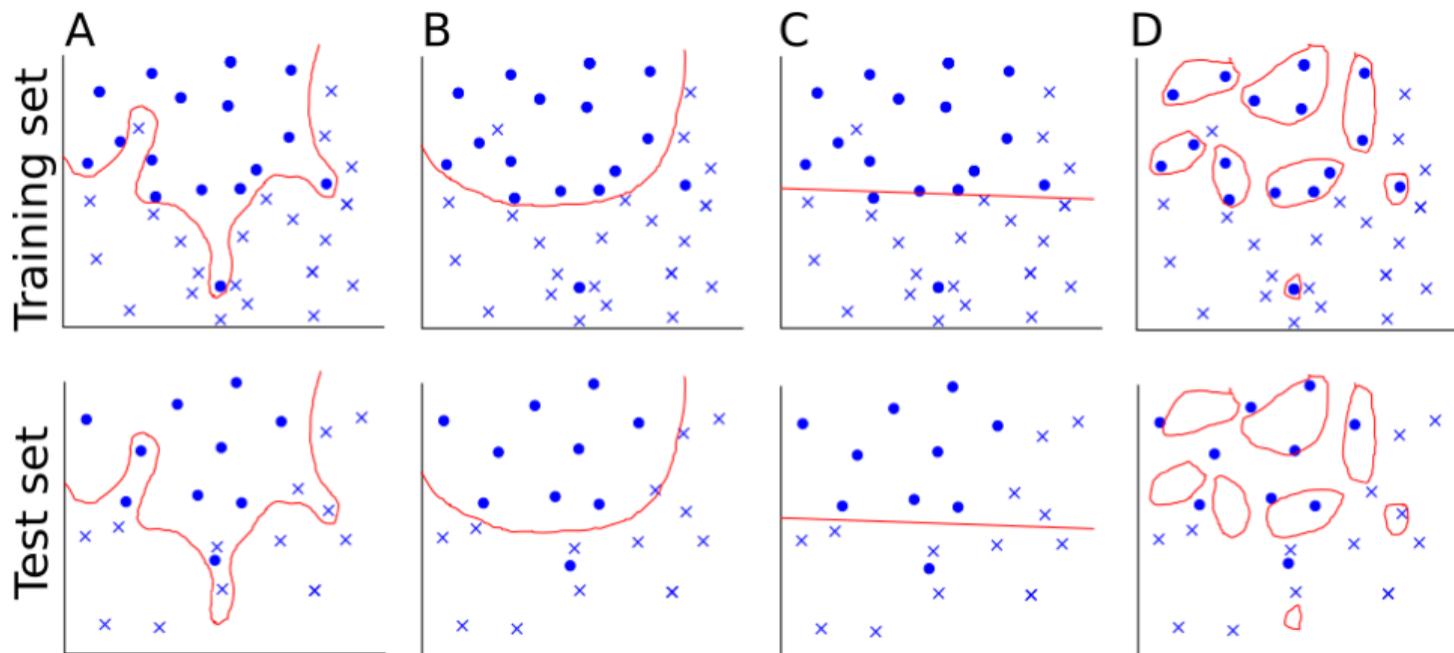
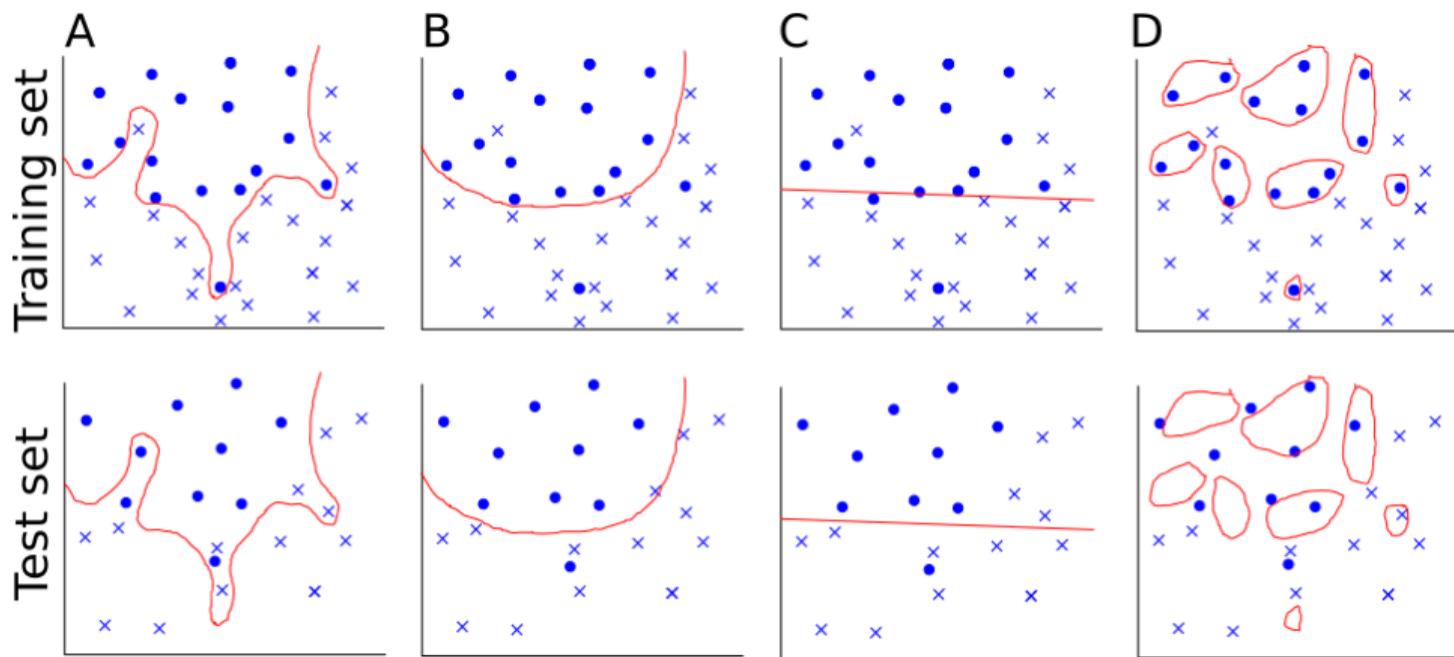

Which classifier (A,B,C,D) is the best? Why?
B

Overfitting and underfitting

Select a classifier which is underfitted.

Overfitting and underfitting

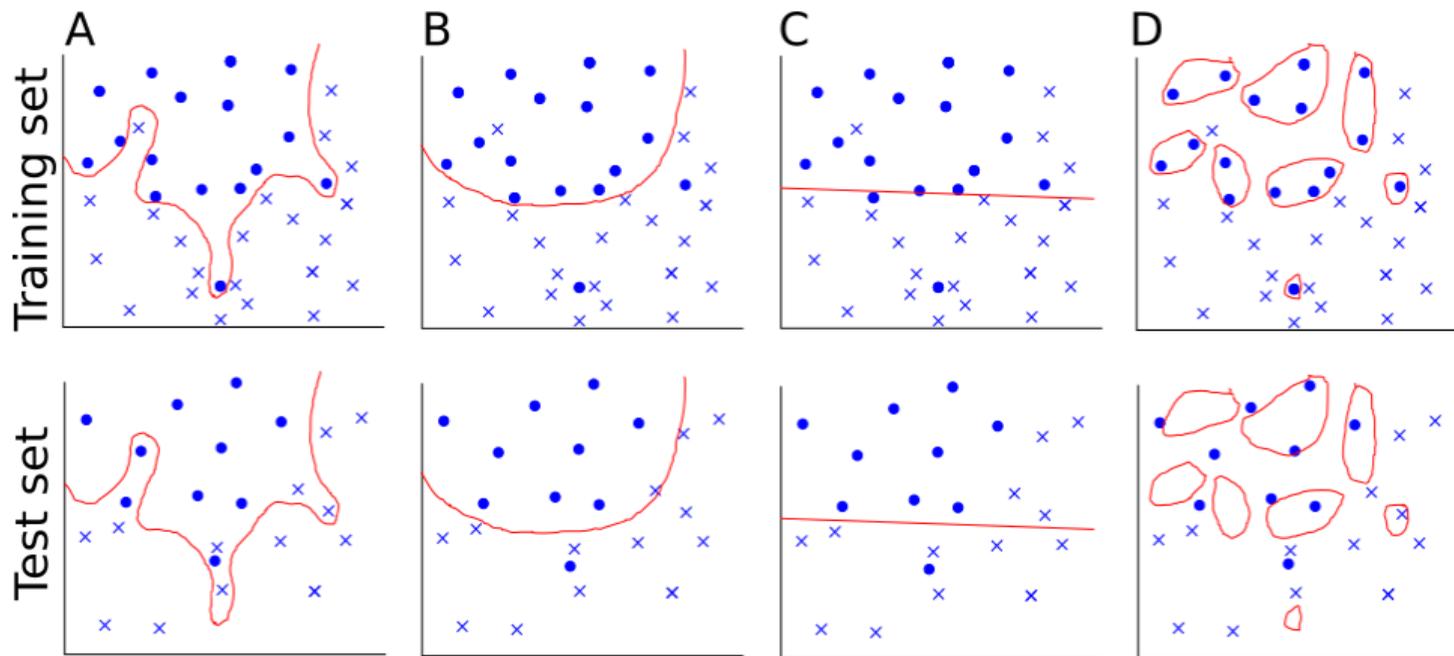Select a classifier which is underfitted.

C

Overfitting and underfitting

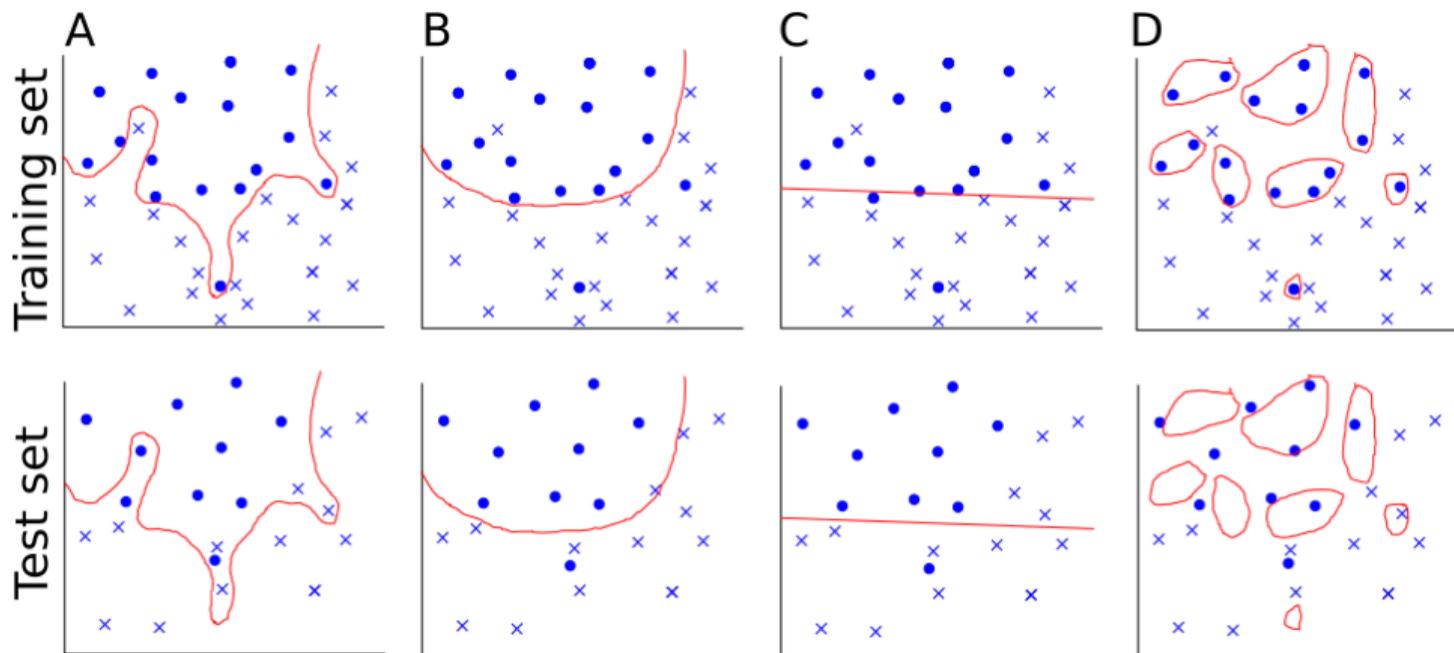Select a classifier which is underfitted.
Discuss how you recognize underfitting from training error.

Overfitting and underfitting

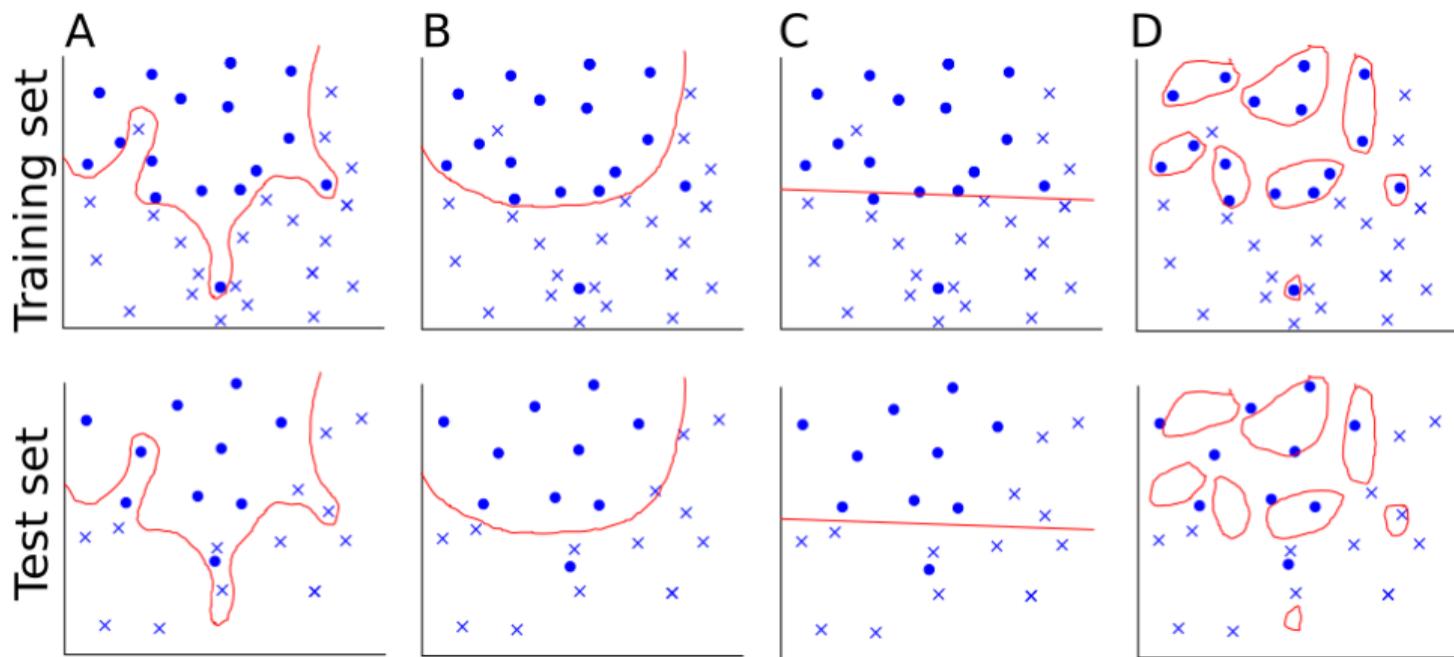Select a classifier which is overfitted.

Overfitting and underfitting

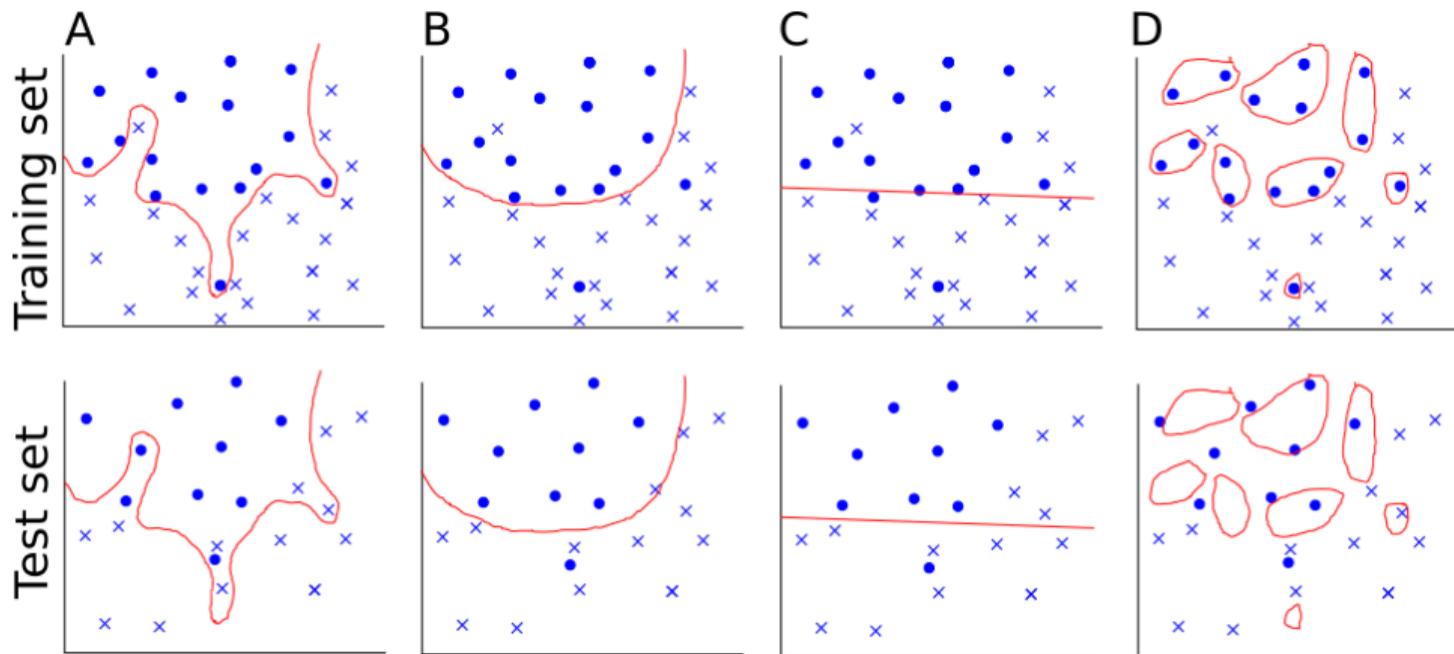Select a classifier which is overfitted.
A and **D**

Overfitting and underfitting

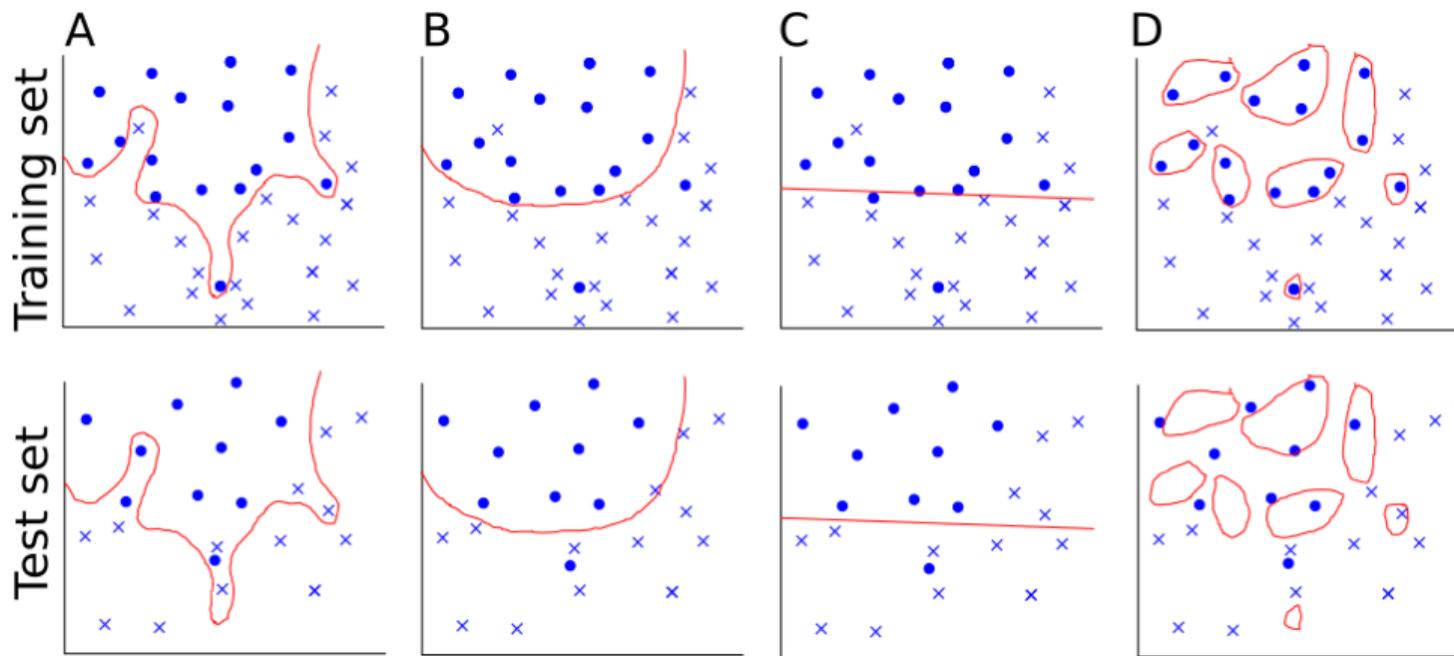Select a classifier which is overfitted.

Discuss how you recognize overfitting from training and validation error.

Overfitting and underfitting

Select a classifier with the worst generalization.

Overfitting and underfitting

Select a classifier with the worst generalization.
D

Overfitting and underfitting

What is most common relationship between training and validation error?

A: validation error $<$ training error

B: validation error $>$ training error

C: validation error $=$ training error

Overfitting and underfitting

What is most common relationship between training and validation error?

B: validation error $>$ training error