

✓ Congratulations! You passed!

Grade received **100%** To pass 80% or higher

Go to next item

1. Fill in the blanks with the correct answer according to the descriptions in the boxes below:

1 / 1 point

Before... when it was all about 1

- Domain experts selected features
- Designed feature transforms
- Small number of more relevant features were enough

Now... 2 is about integrating everything

- Data generation and storage is less of a problem
- Squeeze out the best from data
- More high-dimensional data having more features

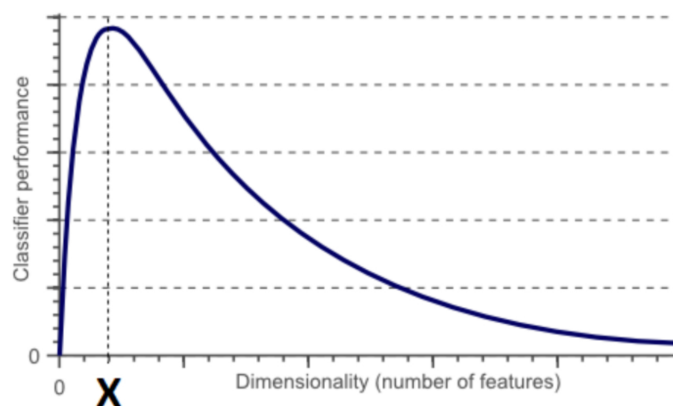
- ☒ 1. Data mining. 2. Data Science.
- ☐ 1. Dimensionality reduction. 2. Data Science.
- ☐ 1. Data mining. 2. Dimensionality reduction.
- ☐ 1. Data Science. 2. Data mining.

✓ Correct

That's right! The "before" and "now" of performance and resource requirements are represented respectively by the Data Mining and Data Science concepts.

2. What does the X value represent?

1 / 1 point



- ☐ The cursed number of dimensions.
- ☒ The optimal number of features.
- ☐ The number of features that reaches the maximum classification error.
- ☐ The worst number of features for making predictions.

✓ Correct

Exactly! The x-axis coordinate of this critical point represents the number of features required by the classifier to work at its best.

3. Which of the following are problems of high dimensionality in model performance? (Select all that apply)

1 / 1 point

- ☒ Solutions take longer to reach global optimum

✓ Correct

Right on track! Very often, reaching a global optimum is a more difficult task when dealing with high-dimensional problems.

- ☒ Higher runtimes and system requirements

✓ Correct

Correct! The more dimensions, the higher the system requirements. Therefore, dimensionality reduction

Correct: The more dimensions, the higher the system requirements. Therefore, dimensionality reduction helps optimize the system's performance.

- ☐ Smaller hypothesis space.
- ☒ The possibility of more correlated features is greater.



Correct

You've got it! When having more dimensions, it is possible to have more correlated features making the selection of the most relevant features a more difficult task.

4. What does the following line of code refer to? `count_params(model_n.trainable_variables)`

1 / 1 point

- ☐ The number of testing parameters for Model n.
- ☒ The number of training parameters for Model n.
- ☐ The number of dimensions for Model n.
- ☐ The number of classes for Model n.



Correct

That's right! This code line allows to count the number of training parameters for the input model.

5. The amount of training data available, the complexity of the decision surface, and the classifier type define the number of _____ to be used

1 / 1 point

- ☒ Features
- ☐ Spaces
- ☐ Models
- ☐ Datasets



Correct

That's right! These three aspects define the amount of features that will be used in a machine learning problem.

6. True Or False: Classification subspaces allow to minimize separation among classes, while regression subspaces are used for maximizing correlation between projected data and response variable.

1 / 1 point

- ☒ False
- ☐ True



Correct

That's right! Classification subspaces maximize the separation among classes, while regression intends to maximize the correlation between two variables.