

✓ Congratulations! You passed!

Grade received 100% To pass 80% or higher

Go to next item

## Dimensionality Reduction

Total points 6

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

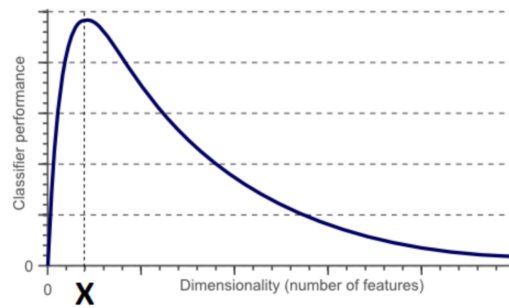
- ☐ 1. Data mining, 2. Dimensionality reduction.
- ☐ 1. Data Science, 2. Data mining.
- ☒ 1. Data mining, 2. Data Science.
- ☐ 1. Dimensionality reduction, 2. Data Science.

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

✓ 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. One of the following is not considered as a high-dimensionality impact:

1 / 1 point

- ☐ Solutions take longer to reach global optimum
- ☐ The possibility of more correlated features is greater.
- ☐ Higher runtimes and system requirements
- ☒ Smaller hypothesis space.

✓ Correct

That's correct! Indeed, increasing dimensionality produces a larger hypothesis space. This problem is called the Hughes effect.

4. What is the output of the code line: `count_params(model_n.trainable_variables)`

1 / 1 point

- ☐ Number of classes for Model n.

- ☐ Number of dimensions for Model n.
- ☒ Number of training parameters for Model n.
- ☐ Number of testing parameters 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 decision surface and the classifier type define the number of \_\_\_\_\_ to be used

1 / 1 point

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



**Correct**

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

6. Classification subspaces allows to minimize separation among classes, while regression subspaces are used for maximizing correlation between projected data and response variable

1 / 1 point

- ☐ True
- ☒ False



**Correct**

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