

Frage 2

Richtig

Erreichte Punkte 1,00 von 1,00

Assume you have the following input of size 4x4:

```
[[4 1 7 1]
 [1 1 3 7]
 [0 2 4 5]
 [6 0 4 3]]
```

What is the output after performing max pooling of size 2x2 with a stride of 2?

- a.

```
[[4 1 7 1]
 [1 1 3 7]
 [0 2 4 5]
 [6 0 4 3]]
```
- b.

```
[7]
```
- c.

```
[[4 7]
 [6 5]]
```
- d.

```
[[4 7 7]
 [2 4 7]
 [6 4 5]]
```

Die richtige Antwort ist:

```
[[4 7]
 [6 5]]
```

Frage 3

Richtig

Erreichte Punkte 1,00 von 1,00

Which aspects have to be taken into consideration when dealing with high-dimensional input data?

- a. More features might lead to longer model training times.
- b. More features take up more space in memory.
- c. Dimensionality reduction techniques might be useful.
- d. Often difficult to visualize.

Die richtigen Antworten sind: Often difficult to visualize., More features might lead to longer model training times., More features take up more space in memory., Dimensionality reduction techniques might be useful.

Frage 4

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about classification?

- a. In classification, the target values cannot be numbers.
- b. In classification, the target values are class labels.
- c. In classification, the target values are numerical values.
- d. In classification, there should be at least two different classes.

Die richtigen Antworten sind: In classification, the target values are class labels., In classification, there should be at least two different classes.

Frage 5

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about linear layers of a neural network?

- a. The layers' weights and biases are the trainable parameters of a neural network.
- b. The main principle of a neural network layer is multiplying its input values with weights and then adding a bias.
- c. Layers can be stacked.
- d. The number of output values of a layer must be equal to the number of its input values.

Die richtigen Antworten sind: The main principle of a neural network layer is multiplying its input values with weights and then adding a bias., The layers' weights and biases are the trainable parameters of a neural network., Layers can be stacked.

Frage 6

Teilweise richtig

Erreichte Punkte 0,50 von 1,00

Which of the following statements is/are true about algorithmic bias with respect to a well-working machine learning model?

- a. Algorithmic bias increases with higher model complexity.
- b. Algorithmic bias may occur when the target labels are unfair/incorrect.
- c. Algorithmic bias may occur when the data is incomplete.
- d. Algorithmic bias is part of the bias-variance trade-off.

Die richtigen Antworten sind: Algorithmic bias may occur when the data is incomplete., Algorithmic bias may occur when the target labels are unfair/incorrect.

Frage 7

Richtig

Erreichte Punkte 1,00 von 1,00

What is meant by the term "overfitting"?

- a. A model fits the test data (too) well but not the training data.
- b. A model fits the training data (too) well but not the test data.
- c. A model neither fits the training nor the test data well.
- d. A model with too many hyperparameters was selected.

Die richtige Antwort ist: A model fits the training data (too) well but not the test data.

Frage 8

Richtig

Erreichte Punkte 1,00 von 1,00

Assume you have a classification task where you want to detect whether an image contains a dog. Which of the following is/are potentially meaningful data augmentations with respect to this data?

- a. Adding slight noise.
- b. Rotating by 360 degrees.
- c. Zooming into the background.
- d. Flipping vertically.

Die richtigen Antworten sind: Flipping vertically., Adding slight noise.

Frage 9

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about regression?

- a. In regression, the input values are used to predict the corresponding target values.
- b. In regression, the target values must be between 0 and 1.
- c. In regression, the target values are class labels.
- d. In regression, the target values are numerical values.

Die richtigen Antworten sind: In regression, the target values are numerical values., In regression, the input values are used to predict the corresponding target values.

Frage 10

Richtig

Erreichte Punkte 1,00 von 1,00

Assume you have RGB images with **width=20** and **height=10**. What is the dimensionality when you want to train a model with such input data?

- a. 200
- b. 20
- c. 10
- d. 600

Die richtige Antwort ist: 600

Frage 11

Teilweise richtig

Erreichte Punkte 0,50 von 1,00

In a fully-connected neural network ...

- a. ... activation functions are not applicable.
- b. ... each hidden layer must have equally many nodes.
- c. ... all inputs are connected to all nodes of the following layer.
- d. ... the output layer is used for the final model prediction.

Die richtigen Antworten sind: ... all inputs are connected to all nodes of the following layer., ... the output layer is used for the final model prediction.

Frage 12

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about padding in convolutional neural networks?

- a. Padding is performed in the flat layer part of a CNN.
- b. Padding is used to keep the input size and output size the same.
- c. Padding cannot be used in combination with pooling.
- d. Padding is used to downsample the input.

Die richtige Antwort ist: Padding is used to keep the input size and output size the same.

Frage 13

Richtig

Erreichte Punkte 1,00 von 1,00

In the area of text processing, which of the following statements is/are true about one-hot encoding?

- a. The length of a one-hot encoded vector is the number of unique words (the dictionary/vocabulary size).
- b. Words can only be represented via one-hot encoding.
- c. For each word, there is a vector where all numbers are 0 but the element representing a specific word is set to 1.
- d. Words are represented with a user-specifiable-sized vector with a sum of 1.

Die richtigen Antworten sind: For each word, there is a vector where all numbers are 0 but the element representing a specific word is set to 1., The length of a one-hot encoded vector is the number of unique words (the dictionary/vocabulary size).

Frage 14

Teilweise richtig

Erreichte Punkte 0,67 von 1,00

Which of the following is/are useful loss functions for classification problems?

- a. Sigmoid.
- b. Mean-squared error.
- c. Softmax.
- d. Cross entropy.

Die richtige Antwort ist: Cross entropy.

Frage 15

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about the result of loss functions (= the loss)?

- a. The lower the loss, the better the prediction.
- b. The higher the loss, the better the prediction.
- c. The value range of every loss function is between [0, 1].
- d. Comparing the loss of two different loss functions is not meaningful.

Die richtigen Antworten sind: The lower the loss, the better the prediction., Comparing the loss of two different loss functions is not meaningful.

Frage 16

Richtig

Erreichte Punkte 1,00 von 1,00

Assume you have the following input text that you want to encode with one-hot-encoding:

the cat jumps across the table

Which of the following statements is/are true?

- a. The dictionary/vocabulary size is 5.
- b. There only exists one single one-hot encoding.
- c. The dictionary/vocabulary size is 6.
- d. There are multiple possible one-hot encodings.

Die richtigen Antworten sind: The dictionary/vocabulary size is 5., There are multiple possible one-hot encodings.

Frage 17

Falsch

Erreichte Punkte 0,00 von 1,00

Given the following dataset in tabular form, what is the dimensionality of this dataset?

age	ID	size
20	A	1.87
31	C	1.94
25	F	1.73
20	R	1.75

- a. 4
- b. 12
- c. 2
- d. 3

Die richtige Antwort ist: 3

Frage 18

Teilweise richtig

Erreichte Punkte 0,67 von 1,00

Why do convolutional neural networks (CNNs) generally perform well on image data?

- a. Because CNNs utilize fully-connected layers at the end of the model architecture.
- b. Because CNNs can deal with high-dimensional data.
- c. Because all CNNs are pretrained on image data.
- d. Because CNNs take advantage of the "local structure" in image data (neighboring pixels are often highly correlated).

Die richtige Antwort ist: Because CNNs take advantage of the "local structure" in image data (neighboring pixels are often highly correlated).

Frage 19

Richtig

Erreichte Punkte 1,00 von 1,00

Assume a multi-class classification problem with four classes [0, 1, 2, 3]. Further assume that you have a model with a softmax function at the end which produced the following output for some input sample (same order as class list):

(0.7, 0.03, 0.27, 0.0)

Which class should be chosen as the final classification prediction?

- a. Class 3
- b. Class 2
- c. Class 0
- d. Class 1

Die richtige Antwort ist: Class 0

Frage 20

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about the generalization error?

- a. The generalization error is the expected loss on future data.
- b. The generalization error is the difference between models that were trained with different input data.
- c. The generalization error can always be calculated.
- d. The generalization error is typically estimated.

Die richtigen Antworten sind: The generalization error is the expected loss on future data., The generalization error is typically estimated.

Frage 21

Richtig

Erreichte Punkte 1,00 von 1,00

What is weight sharing in convolutional neural networks?

- a. Applying the same kernel to all convolutional layers.
- b. Applying the same network architecture in different convolutional neural networks.
- c. Applying the same hyperparameters to all kernels in the network.
- d. Applying the same kernel to all input positions.

Die richtige Antwort ist: Applying the same kernel to all input positions.

Frage **22**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about stacking linear layers in a neural network?

- a. Stacking linear layers will always improve the model performance.
- b. Stacking linear layers increases the depth of the network.
- c. Stacking linear layers increases the width of the network.
- d. Stacking linear layers requires non-linear activation functions to actually increase the model complexity.

Die richtigen Antworten sind: Stacking linear layers increases the depth of the network., Stacking linear layers requires non-linear activation functions to actually increase the model complexity.

Frage **23**

Richtig

Erreichte Punkte 1,00 von 1,00

In a convolutional neural network, downsampling of an input may be achieved by ...

- a. ... 2x2 max-pooling.
- b. ... setting the stride to 2.
- c. ... setting the stride to 1.
- d. ... 1x1 max-pooling.

Die richtigen Antworten sind: ... 2x2 max-pooling., ... setting the stride to 2.

Frage **24**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about gradient descent?

- a. Gradient descent starts at a random position in the parameter space.
- b. Gradient descent is an iterative method.
- c. Gradient descent is guaranteed to find the global minimum.
- d. Gradient descent takes steps in the direction of the negative gradient of the function to minimize.

Die richtigen Antworten sind: Gradient descent is an iterative method., Gradient descent takes steps in the direction of the negative gradient of the function to minimize., Gradient descent starts at a random position in the parameter space.

Frage **25**

Richtig

Erreichte Punkte 1,00 von 1,00

Linear regression ...

- a. ... is suitable for regression.
- b. ... is suitable for classification.
- c. ... can be used for multi-dimensional data.
- ☒ d. ... is never a good model choice.

Die richtigen Antworten sind: ... is suitable for regression., ... can be used for multi-dimensional data.

Frage **26**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about loss functions?

- a. There are dedicated loss functions for regression and classification.
- b. The output of loss functions are the final model predictions.
- c. When comparing two different loss values calculated with two different loss functions, one should choose the loss function that yielded the lower loss value.
- d. Loss functions can be used to measure the difference between a model prediction and the true target.

Die richtigen Antworten sind: Loss functions can be used to measure the difference between a model prediction and the true target., There are dedicated loss functions for regression and classification.

Frage **27**

Teilweise richtig

Erreichte Punkte 0,50 von 1,00

Which of the following statements is/are true about an 8-bit RGB image?

- a. Every channel can encode 8 different values.
- b. Can be converted into a grayscale image without additional information.
- c. Every pixel is represented by 8 channels.
- d. Every channel can encode 2^8 (=256) different values.

Die richtigen Antworten sind: Every channel can encode 2^8 (=256) different values., Can be converted into a grayscale image without additional information.

Frage **28**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about data augmentation?

- a. Data augmentation can be used to create/generate new samples.
- b. Augmentation is restricted to the field of image processing.
- c. Useful augmentation techniques depend on the input data.
- d. The model performance might be worse than without augmentation.

Die richtigen Antworten sind: Useful augmentation techniques depend on the input data., Data augmentation can be used to create/generate new samples., The model performance might be worse than without augmentation.

Frage **29**

Richtig

Erreichte Punkte 1,00 von 1,00

A non-convex function ...

- a. ... sometimes has a closed-form solution.
- b. ... only has one minimum.
- c. ... usually occurs when training neural networks.
- d. ... never has a closed-form solution.

Die richtigen Antworten sind: ... usually occurs when training neural networks., ... never has a closed-form solution.

Frage **30**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following statements is/are true about the softmax function?

- a. The sum of all outputs equals 1.
- b. The output is always 1 for the predicted class and 0 for all others.
- c. It is suitable for multi-class classification problems.
- d. It is a generalization of the sigmoid function.

Die richtigen Antworten sind: It is a generalization of the sigmoid function., It is suitable for multi-class classification problems., The sum of all outputs equals 1.

Frage 31

Teilweise richtig

Erreichte Punkte 0,50 von 1,00

What is/are the main purposes of applying the logistic function (sigmoid) in logistic regression?

- a. To introduce non-linearity.
- b. To handle multi-dimensional input.
- c. To estimate the probability of class membership.
- d. To transform the loss into a closed-form solution.

Die richtigen Antworten sind: To introduce non-linearity., To estimate the probability of class membership.

Frage 32

Richtig

Erreichte Punkte 1,00 von 1,00

What implications does a non-closed-form solution of a function bring?

- a. The function cannot be used in machine learning.
- b. The minimum can be computed analytically.
- c. The minimum must be calculated with iterative methods.
- d. The function can either be non-convex or convex.

Die richtigen Antworten sind: The minimum must be calculated with iterative methods., The function can either be non-convex or convex.

Frage 33

Richtig

Erreichte Punkte 1,00 von 1,00

Considering tabular data, which of the following statements is/are true about the term "sample"?

- a. The data of a sample must be unique with respect to all other samples.
- b. A sample contains data.
- c. A collection of samples creates a data set.
- d. Samples might have a varying number of features.

Die richtigen Antworten sind: A sample contains data., A collection of samples creates a data set.

Frage 34

Teilweise richtig

Erreichte Punkte 0,50 von 1,00

Regarding continuous sound waves/signals, the sampling rate ...

- a. ... is a hyperparameter of the Fourier transform.
- b. ... defines for how long a signal can be sampled from.
- c. ... is typically measured in Hertz (Hz).
- d. ... is the rate at which discrete samples are retrieved from the signal.

Die richtigen Antworten sind: ... is the rate at which discrete samples are retrieved from the signal., ... is typically measured in Hertz (Hz).

Frage 35

Richtig

Erreichte Punkte 1,00 von 1,00

Given an input of size 9x9 and a kernel size of 3x3 with a stride of 1, what is the output size after the convolution operation?

- a. 9x9
- b. 7x7
- c. 8x8
- d. 6x6

Die richtige Antwort ist: 7x7

Frage 36

Richtig

Erreichte Punkte 1,00 von 1,00

Deep networks ...

- a. ... are only applicable in combination with pretrained networks.
- b. ... enable potentially more complex and powerful models.
- c. ... can suffer from the vanishing gradient problem.
- d. ... will always improve the prediction accuracy.

Die richtigen Antworten sind: ... enable potentially more complex and powerful models., ... can suffer from the vanishing gradient problem.

Frage 37

Falsch

Erreichte Punkte 0,00 von 1,00

k-nearest neighbors ...

- a. ... is a supervised learning model.
- b. ... is a subtype of neural networks.
- c. ... is a clustering method.
- d. ... incorporates randomness to reduce overfitting.

Die richtige Antwort ist: ... is a supervised learning model.

Frage 38

Richtig

Erreichte Punkte 1,00 von 1,00

Given the following labeled sample, which of the following statements is/are true?

 $x = (1.0, -0.9, -0.3), y = 1$

- a. y is called a target value.
- b. With only this single sample, we cannot tell whether it might be a classification or regression task.
- c. There cannot be another sample with the same data.
- d. There are two classes, 0 and 1.

Die richtigen Antworten sind: y is called a target value., With only this single sample, we cannot tell whether it might be a classification or regression task.

Frage 39

Richtig

Erreichte Punkte 1,00 von 1,00

Which techniques can be used to potentially improve a neural network model?

- a. Dropout
- b. Data augmentation
- c. Learning rate schedules
- d. Transfer learning

Die richtigen Antworten sind: Dropout, Data augmentation, Transfer learning, Learning rate schedules

Frage **40**

Richtig

Erreichte Punkte 1,00 von 1,00

Principal Component Analysis (PCA) ...

- a. ... is a dimensionality reduction method.
- b. ... is a data augmentation method.
- c. ... is supervised machine learning classification method.
- d. ... is a clustering method.

Die richtige Antwort ist: ... is a dimensionality reduction method.

Frage **41**

Richtig

Erreichte Punkte 1,00 von 1,00

Which of the following is/are typically used activation functions?

- a. ReLU
- b. Sigmoid
- c. Cross entropy
- d. Max pooling

Die richtigen Antworten sind: ReLU, Sigmoid

[◀ Dummy Exam](#)

Direkt zu: