

# Aufgaben

---

Annika Liebgott

December 15, 2022

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1. tanh
2. Sigmoid
3. ReLU
4. Softmax
5. Identität

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1.  $\tanh \Rightarrow$  hidden Layer
2. Sigmoid
3. ReLU
4. Softmax
5. Identität

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1.  $\tanh \Rightarrow$  hidden Layer
2. Sigmoid  $\Rightarrow$  Klassifikationsschicht
3. ReLU
4. Softmax
5. Identität

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1.  $\tanh \Rightarrow$  hidden Layer
2. Sigmoid  $\Rightarrow$  Klassifikationsschicht
3. ReLU  $\Rightarrow$  hidden Layer
4. Softmax
5. Identität

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1.  $\tanh \Rightarrow$  hidden Layer
2. Sigmoid  $\Rightarrow$  Klassifikationsschicht
3. ReLU  $\Rightarrow$  hidden Layer
4. Softmax  $\Rightarrow$  Klassifikationsschicht
5. Identität

## Aktivierungsfunktionen

Geben Sie für die folgenden Aktivierungsfunktionen an, in welchem Typ von Layer eines neuronalen Netzes sie typischerweise verwendet werden:

1.  $\tanh \Rightarrow$  hidden Layer
2. Sigmoid  $\Rightarrow$  Klassifikationsschicht
3. ReLU  $\Rightarrow$  hidden Layer
4. Softmax  $\Rightarrow$  Klassifikationsschicht
5. Identität  $\Rightarrow$  Regressionsschicht

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt  $\Rightarrow$  falsch
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt  $\Rightarrow$  falsch
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale  $\Rightarrow$  falsch
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt  $\Rightarrow$  falsch
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale  $\Rightarrow$  falsch
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation  $\Rightarrow$  wahr
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt  $\Rightarrow$  falsch
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale  $\Rightarrow$  falsch
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation  $\Rightarrow$  wahr
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren  $\Rightarrow$  wahr
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl

## Merkmalsreduktion

Welche Aussagen zur Merkmalsreduktion im maschinellen Lernen sind wahr, welche falsch?

1. Principal Component Analysis wird zur Merkmalselektion genutzt ⇒ falsch
2. Merkmalstransformationen erhalten die Bedeutung einzelner Merkmale ⇒ falsch
3. Beim Stochastic Proximity Embedding handelt es sich um eine Merkmalstransformation ⇒ wahr
4. Neuronale Netze können trainiert werden, um die besten Merkmale aus einer Merkmalsmatrix zu identifizieren ⇒ wahr
5. Bei einer sehr großen Anzahl an Merkmalen ist Brute-Force-Selektion das Mittel der Wahl ⇒ falsch

## Entscheidungstheorie nach Bayes

Geben Sie an, welche Bedingungen erfüllt werden müssen, damit aus einem Minimum-Bayesian-Risk-Klassifikator ein Maximum-Likelihood-Klassifikator wird.

## Entscheidungstheorie nach Bayes

Geben Sie an, welche Bedingungen erfüllt werden müssen, damit aus einem Minimum-Bayesian-Risk-Klassifikator ein Maximum-Likelihood-Klassifikator wird.

Antwort: 0/1-Loss und  $P(\omega_j) = \frac{1}{c}$

## **$k$ -Means**

- a) Erläutern Sie kurz die Funktionsweise des Algorithmus
  
  
  
  
  
  
- b) Wie funktioniert die Klassifikator-Initialisierung bei  $k$ -Means++?
  
  
  
  
  
  
- c) Was ist der Unterschied zwischen k-Means und dem Fuzzy- $c$ -Means-Klassifikator?

# Beispielaufgabe Klausur (F2022)

## *k*-Means

- a) Erläutern Sie kurz die Funktionsweise des Algorithmus
  1. Wähle zufällig  $k$  Samples als initiale Clustermittelpunkte.
  2. Berechne Abstand aller Samples zu den Mittelpunkten und ordne sie dem jeweils nächsten zu.
  3. Berechne die Mittelpunkte der neuen Cluster
  4. Wiederhole 2./3. bis sich nichts mehr ändert.
- b) Wie funktioniert die Klassifikator-Initialisierung bei  $k$ -Means++?
- c) Was ist der Unterschied zwischen k-Means und dem Fuzzy- $c$ -Means-Klassifikator?

# Beispielaufgabe Klausur (F2022)

## **$k$ -Means**

- a) Erläutern Sie kurz die Funktionsweise des Algorithmus
1. Wähle zufällig  $k$  Samples als initiale Clustermittelpunkte.
  2. Berechne Abstand aller Samples zu den Mittelpunkten und ordne sie dem jeweils nächsten zu.
  3. Berechne die Mittelpunkte der neuen Cluster
  4. Wiederhole 2./3. bis sich nichts mehr ändert.
- b) Wie funktioniert die Klassifikator-Initialisierung bei  $k$ -Means++?
- Statt alle Mittelpunkte zufällig zu wählen, wird nur der 1. komplett zufällig gewählt. Für den Rest wird die Auswahlwahrscheinlichkeit mit dem Abstand zum 1. gewichtet.
- c) Was ist der Unterschied zwischen  $k$ -Means und dem Fuzzy- $c$ -Means-Klassifikator?

# Beispielaufgabe Klausur (F2022)

## **$k$ -Means**

a) Erläutern Sie kurz die Funktionsweise des Algorithmus

1. Wähle zufällig  $k$  Samples als initiale Clustermittelpunkte. 2. Berechne Abstand aller Samples zu den Mittelpunkten und ordne sie dem jeweils nächsten zu. 3. Berechne die Mittelpunkte der neuen Cluster 4. Wiederhole 2./3. bis sich nichts mehr ändert.

b) Wie funktioniert die Klassifikator-Initialisierung bei  $k$ -Means++?

Statt alle Mittelpunkte zufällig zu wählen, wird nur der 1. komplett zufällig gewählt. Für den Rest wird die Auswahlwahrscheinlichkeit mit dem Abstand zum 1. gewichtet.

c) Was ist der Unterschied zwischen  $k$ -Means und dem Fuzzy- $c$ -Means-Klassifikator?

Statt einer harten Entscheidung für eine Klasse gibt es beim Fuzzy- $c$ -Means als Ergebnis eine Zugehörigkeitswahrscheinlichkeit zu allen Klassen.