

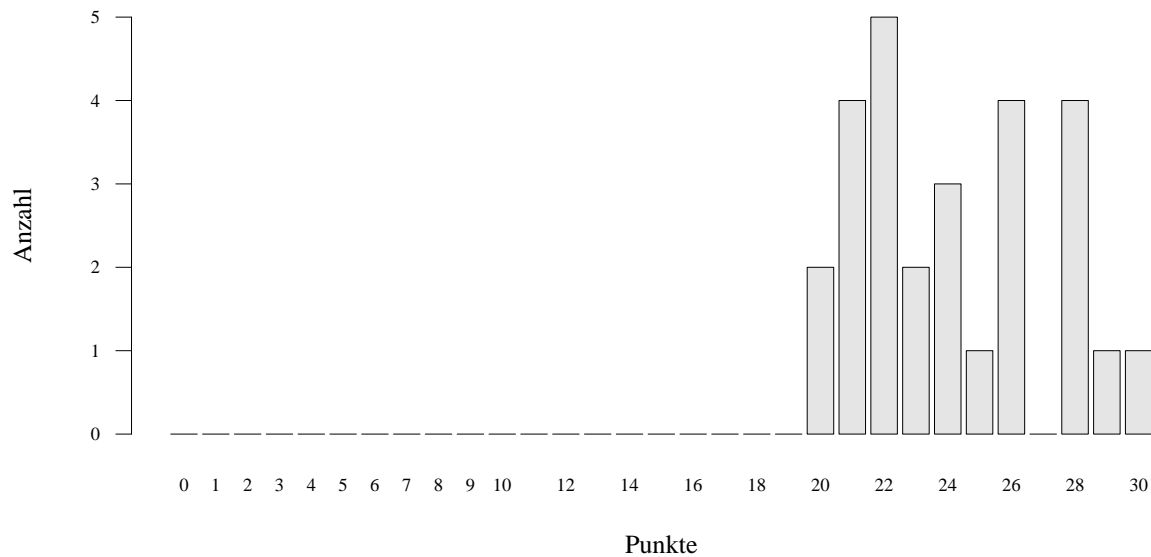
Klausurreport »Multivariate Verfahren«, SoSe 2025

Die Teilklausur “Multivariate Verfahren” im Rahmen der Modulabschlussklausur “Forschungsmethoden” (M.Sc. Psychologie, Modul A) im Sommersemester 2025 wurde am 18.07.2025 von 14:00 bis 16:00 Uhr in den Computer-Pools des Universitätsrechenzentrums (URZ) als elektronische Klausur (E-Klausur) mit 27 Teilnehmenden durchgeführt. Sie bestand aus 30 Multiple-Choice-Aufgaben mit jeweils vier Antwortmöglichkeiten und jeweils genau einer richtigen Antwort. Die Klausur ist diesem Bericht beigelegt, richtige Antworten sind auf der letzten Seite angegeben.

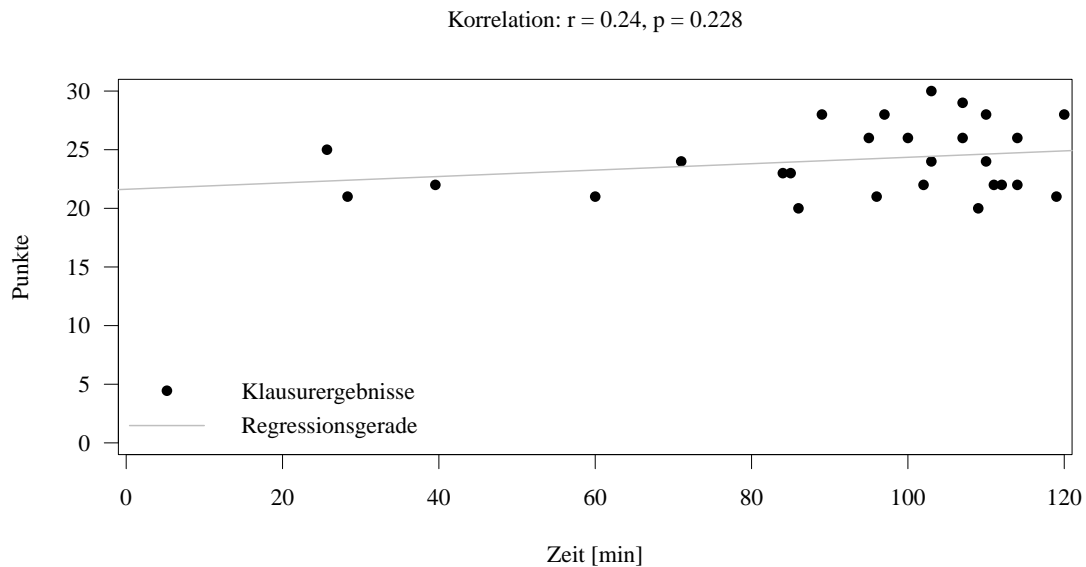
Ergebnisse

Die nachfolgende Abbildung zeigt die absolute Häufigkeitsverteilung der erzielten Punkte.

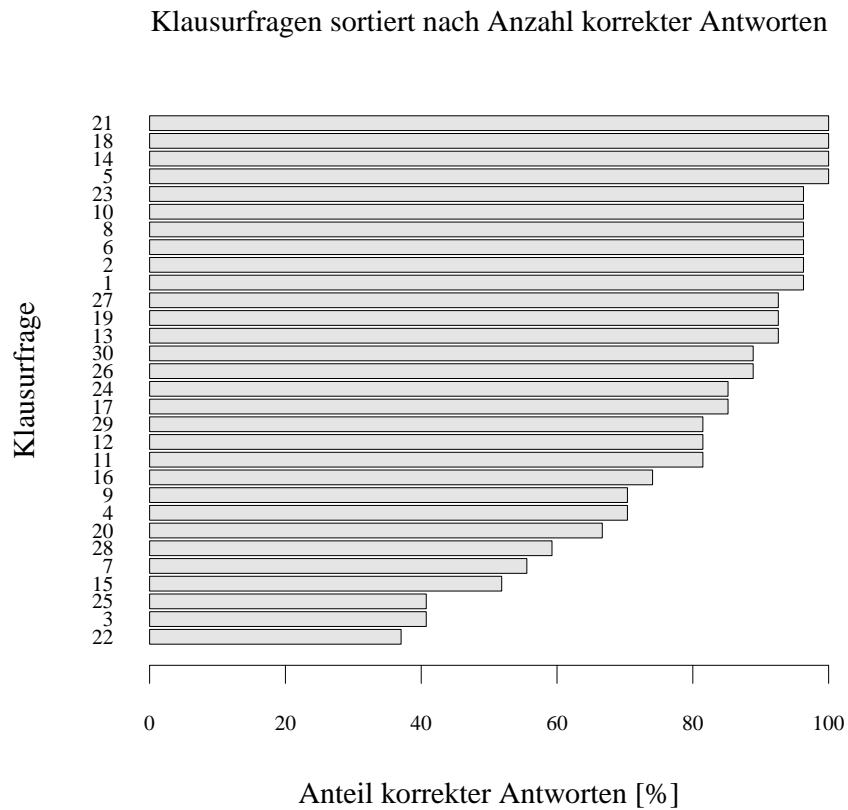
Median: 24.0, Mittelwert: 24.1, Gleitklauselgrenze: 18.8, $n = 27$



Die nachfolgende Abbildung zeigt die Korrelation zwischen verbrauchter Zeit (= Dauer von Start der Klausur bis zu elektronischer Abgabe) und erreichter Punktzahl (= Anzahl richtig beantworteter Fragen).



Die nachfolgende Abbildung zeigt die Reihenfolge der Klausurfragen, wenn man sie nach Anteil korrekter Antworten über alle Teilnehmenden hinweg sortiert. Dieser Anteil kann annäherungsweise als umgekehrter Schweregrad einer Frage interpretiert werden.



OTTO-VON-GUERICKE-UNIVERSITÄT MAGDEBURG
Fakultät für Naturwissenschaften
Institut für Psychologie
Lehrstuhl Methodenlehre I
Dr. rer. nat. Joram Soch

Modulabschlussklausur "Forschungsmethoden"
Teilklausur "Multivariate Verfahren"
Termin: 18.07.2025

Name, Vorname: _____

Matrikelnummer: _____

Bearbeitungshinweise:

- Die Klausur besteht aus **30 Aufgaben**.
- Bei jeder Aufgabe sind jeweils **vier Antwortmöglichkeiten** vorgegeben.
- Es trifft **immer genau eine** Antwort zu.
- Bitte kreuzen Sie bei jeder Aufgabe nur die **zutreffende Antwort** an.
- Für jede **richtig gelöste Aufgabe** erhalten Sie einen Punkt.

Viel Erfolg!

1. Was versteht man in der psychologischen Forschung unter einer abhängigen Variable?
 - a) eine Größe, die Experimentierende typischerweise unter ihrer Kontrolle haben und die im Verlaufe des Experiments variiert wird, um ihren Einfluss auf andere Größen zu ermitteln
 - b) eine Größe, die Experimentierende typischerweise messen und die das primäre Ergebnismaß des Experiments darstellt, das ggf. von anderen Größen beeinflusst wird
 - c) eine grundsätzlich eindimensionale Zufallsvariable
 - d) eine grundsätzlich mehrdimensionale Zufallsvariable

2. Welche Größe würde in einer kognitiv-neurowissenschaftlichen Studie, in der visuelle Stimuli unterschiedlicher Objektkategorie, Wahrnehmbarkeit (Kohärenz) und Präsentationsdauer gezeigt werden, während mittels Elektroenzephalographie (EEG) die Hirnaktivität erfasst wird, am ehesten als abhängige Variable betrachtet werden?
 - a) Objektkategorie
 - b) Wahrnehmbarkeit
 - c) Präsentationsdauer
 - d) EEG-Messergebnisse

3. Wieviele unabhängige Variablen hat eine klinisch-psychologische Studie, in der die Veränderung von BDI-Score (Beck Depression Inventory), VLMT-Testergebnis (Verbaler Lern- und Merkfähigkeitstest) und Glukokortikoid-Konzentration zwischen zwei verschiedenen Studiengruppen mit unterschiedlicher Psychotherapieform verglichen werden?
 - a) 0
 - b) 1
 - c) 2
 - d) 3

4. Welche dieser Matrixoperationen ist **nicht** elementweise definiert?
 - a) Matrixaddition
 - b) Matrixsubtraktion
 - c) Skalarmultiplikation
 - d) Matrixmultiplikation

5. Es seien $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$, $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $b = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$. Berechnen Sie $Ax + b$!
 - a) $Ax + b = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$
 - b) $Ax + b = \begin{pmatrix} 8 \\ 5 \end{pmatrix}$
 - c) $Ax + b = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$
 - d) $Ax + b = \begin{pmatrix} 11 \\ 6 \end{pmatrix}$

6. Gegeben sei die Matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Welche Aussage trifft dann zu?
- A ist eine symmetrische Matrix.
 - A ist eine Diagonalmatrix.
 - Die Determinante von A ist 5.
 - Die Spur von A ist 6.
7. Welche Aussage über die n -dimensionale Einheitsmatrix I_n ist **nicht** korrekt?
- Die Spur der Einheitsmatrix ist n .
 - Die Determinante der Einheitsmatrix ist n .
 - Die Inverse der Einheitsmatrix ist I_n .
 - Die Einheitsmatrix ist symmetrisch.
8. $A \in \mathbb{R}^{n \times n}$ sei eine quadratische Matrix und v sei ein Eigenvektor von A zum Eigenwert λ . Wie ist dieser Zusammenhang mathematisch definiert?
- $Av = \lambda v$
 - $\lambda^2 = Av$
 - $v^T v = A\lambda$
 - $A + v = \lambda + v$
9. ξ sei ein n -dimensionaler Zufallsvektor. Wie ist die Kovarianzmatrix $\mathbb{C}(\xi) \in \mathbb{R}^{n \times n}$ definiert?
- $\mathbb{C}(\xi) := \mathbb{E}((\xi - \mathbb{E}(\xi))^T(\xi - \mathbb{E}(\xi)))$
 - $\mathbb{C}(\xi) := \mathbb{E}((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T)$
 - $\mathbb{C}(\xi) := \mathbb{E}((\xi - \mathbb{V}(\xi))(\xi - \mathbb{V}(\xi))^T)$
 - $\mathbb{C}(\xi) := I_n + \mathbb{E}(\xi)\mathbb{E}(\xi)^T$
10. Was besagt das Theorem zur linear-affinen Transformation multivariater Normalverteilungen?
- Die marginalen Verteilungen einer multivariaten Normalverteilung sind ebenfalls multivariate Normalverteilungen.
 - Die bedingten Verteilungen einer multivariaten Normalverteilung sind ebenfalls multivariate Normalverteilungen.
 - Die Parameter einer gemeinsamen Normalverteilungen ergeben sich als linear-affine Transformation der Parameter der sie induzierenden marginalen und bedingten Normalverteilungen.
 - Jede linear-affine Transformation eines multivariat normalverteilten Zufallsvektors ist ebenfalls multivariat normalverteilt.

11. Gegeben seien eine univariate unabhängige Variable und eine multivariate abhängige Variable. Welche Methode ist ein passendes Datenanalyseverfahren?
- einfache lineare Regression
 - multiple lineare Regression
 - einfaktorielle multivariate Varianzanalyse (MANOVA)
 - Kanonische Korrelationsanalyse
12. Gegeben sei ein Datensatz y_1, \dots, y_n mit $y_i \in \mathbb{R}^m$ für $i = 1, \dots, n$. Wie ist das multivariate Stichprobenmittel der y_1, \dots, y_n definiert?
- $\bar{y} := \sum_{i=1}^n y_i$
 - $\bar{y} := \frac{1}{n} \sum_{i=1}^n y_i$
 - $\bar{y} := \frac{1}{n-1} \sum_{i=1}^n y_i$
 - $\bar{y} := \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})$
13. Gegeben sei das Szenario der einfaktoriellen multivariaten Varianzanalyse mit $p = 3$ Gruppen, $k = 12$ experimentellen Einheiten pro Gruppe und Datendimension $m = 2$. Wieviele Einträge hat in diesem Fall die Datenmatrix Y , d.h. aus wie vielen skalaren Zahlen besteht der Gesamtdatensatz?
- 36
 - 72
 - 360
 - 720
14. Im Modell der einfaktoriellen multivariaten Varianzanalyse werden die gruppenspezifischen Erwartungswertparameter mithilfe m -dimensionaler Vektoren $\mu_i \in \mathbb{R}^m$ mit $i = 1, \dots, p$ modelliert. Diese Vektoren haben dabei so viele Komponenten, wie es Datendimensionen gibt (m). Die Nullhypothese für dieses Modell besagt, dass die Erwartungswertparameter aller Gruppen identisch sind: $H_0 : \mu_1 = \dots = \mu_p$. Wie lautet die Alternativhypothese?
- Die Erwartungswertparameter mindestens zweier Gruppen unterscheiden sich in mindestens einer Komponente.
 - Die Erwartungswertparameter aller Gruppen unterscheiden sich paarweise in mindestens einer Komponente.
 - Die Erwartungswertparameter mindestens zweier Gruppen unterscheiden sich in allen Komponenten.
 - Die Erwartungswertparameter aller Gruppen unterscheiden sich paarweise in allen Komponenten.
15. Wie verhalten sich Kovarianz und Korrelation zweier Zufallsvariablen im Allgemeinen bei linear-affiner Transformation dieser Zufallsvariablen?
- Sowohl Kovarianz als auch Korrelation ändern sich.
 - Die Kovarianz bleibt gleich, die Korrelation ändert sich.
 - Die Kovarianz ändert sich, die Korrelation bleibt gleich.
 - Sowohl Kovarianz als auch Korrelation bleiben gleich.

16. Worin besteht das Ziel der Kanonischen Korrelationsanalyse (KKA) im Hinblick auf die Arbeit mit unabhängigen Variablen (UV) und abhängigen Variablen (AV)?
- Das Ziel einer KKA besteht darin, Linearkombinationen der UVs und der AVs zu finden, sodass die Korrelation dieser Linearkombinationen maximal ist.
 - Das Ziel einer KKA besteht darin, Linearkombinationen der UVs und der AVs zu finden, sodass die Varianz dieser Linearkombinationen maximal ist.
 - Das Ziel einer KKA besteht darin, beobachtete Werte der UVs und AVs zu Mittelwert-zentrieren, um Erwartungswert-unabhängige ("kanonische") Korrelationen dieser berechnen zu können.
 - Das Ziel einer KKA besteht darin, zwischen einer Menge von UVs und einer Menge von AVs alle paarweisen Korrelationen zu berechnen, um das am stärksten korrelierte Paar von UV und AV zu identifizieren.
17. Was ist das Prinzip einer leave-one-out cross-validation (LOO-CV)?
- Ein prädiktives Modell wird wiederholt an einem Datensatz trainiert und getestet, wobei in jeder Wiederholung der Trainingsdatensatz aus allen experimentellen Einheiten bis auf eine besteht und die ausgelassene experimentelle Einheit den Testdatensatz bildet.
 - Ein prädiktives Modell wird wiederholt an einem Datensatz trainiert und getestet, wobei in jeder Wiederholung der Trainingsdatensatz aus 9/10 der experimentellen Einheiten besteht und das restliche 1/10 der experimentellen Einheiten den Testdatensatz bildet.
 - LOO-CV bedeutet, dass prädiktive Modelle wiederholt unter Auslassung eines Features trainiert werden, wobei in jeder Wiederholung das geschätzte prädiktive Modell auf dem jeweils ausgelassenen Feature getestet wird.
 - LOO-CV bedeutet, dass bei der Berechnung von Performanzmetriken jeweils ein Exemplar der true negative (TN), false positive (FP), false negative (FN) und false positive (FP) ausgelassen wird, um unverzerrte Ergebnisse zu erhalten.
18. Die Einträge der Konfusionsmatrix für den Fall binärer Klassifikation heißen true negative (TN), false positive (FP), false negative (FN) und false positive (FP). Welche Aussage über diese Einträge ist korrekt?
- TN ist die Anzahl der negativen Fälle, die als positiv klassifiziert wurden.
 - FP ist die Anzahl der positiven Fälle, die als negativ klassifiziert wurden.
 - FN ist die Anzahl der negativen Fälle, die als negativ klassifiziert wurden.
 - TP ist die Anzahl der positiven Fälle, die als positiv klassifiziert wurden.
19. Welche Definitionsmenge und Zielmenge hat eine multivariate reellwertige Funktion?
- $f : \mathbb{R} \rightarrow \mathbb{R}$
 - $f : \mathbb{R} \rightarrow \mathbb{R}^m, m > 1$
 - $f : \mathbb{R}^n \rightarrow \mathbb{R}, n > 1$
 - $f : \mathbb{R}^n \rightarrow \mathbb{R}^m, n > 1, m > 1$

20. Gegeben sei eine bivariate reellwertige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Wie ist der Gradient dieser Funktion an der Stelle $x = (x_1 \ x_2)^T$ definiert?

a) $\nabla f(x) := \begin{pmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \end{pmatrix}$

b) $\nabla f(x) := \begin{pmatrix} \frac{\partial}{\partial x_2} f(x) \\ \frac{\partial}{\partial x_1} f(x) \end{pmatrix}$

c) $\nabla f(x) := \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_2} f(x) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(x) \end{pmatrix}$

d) $\nabla f(x) := \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x) \\ \frac{\partial^2}{\partial x_2 \partial x_2} f(x) \end{pmatrix}$

21. Welche der folgenden Mengen von Vektoren ist die kanonische Basis von \mathbb{R}^2 ?

a) $B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

b) $B = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\}$

c) $B = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$

d) $B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

22. $\tilde{X} = Q^T X$ sei der aus einer Hauptkomponentenanalyse eines Datensatzes $X \in \mathbb{R}^{m \times n}$ hervorgegangene transformierte Datensatz. Welche Aussage trifft **nicht** zu?

a) Die Dimensionalität des transformierten Datensatzes \tilde{X} ist $n \times m$.

b) Die paarweisen Stichprobenkorrelationen der Features in \tilde{X} sind Null.

c) Die Gesamtvarianz der Features in \tilde{X} entspricht der Gesamtvarianz der Features in X .

d) Die erste Zeile von \tilde{X} entspricht dem transformierten Feature mit der größten Varianz.

23. Welchen Wertebereich hat der Parameter μ einer Bernoulli-Zufallsvariable $y \sim \text{Bern}(\mu)$?

a) $0 \leq \mu \leq 1$

b) $0 \leq \mu < +\infty$

c) $-\infty < \mu \leq 0$

d) $-\infty < \mu < +\infty$

24. Gegeben sei die der Klassifikationsregel der linearen Diskriminanzanalyse zugrundeliegende bedingte Wahrscheinlichkeit $p(y = 1|x) = 1 / (1 + \exp(-\tilde{x}^T \beta))$? Was bezeichnet \tilde{x} in dieser Gleichung?

a) das Klassenlabel

b) den Featurevektor

c) den erweiterten Featurevektor

d) den Inferenzparametervektor

25. Was ist **kein** Bestandteil eines Generalisierten Linearen Modells für einen erweiterten Featurevektor $x \in \mathbb{R}^{m+1}$ und das mit ihm assoziierte Label y ?
- ein Parametervektor $\beta \in \mathbb{R}^{m+1}$
 - ein linearer Prädiktor $\eta = x^T \beta$
 - eine Link-Funktion $g(\mathbb{E}(y)) = \eta$
 - eine Wahrscheinlichkeitsverteilung $\mathbb{P}(\beta)$
26. Welche Funktion ist die Link-Funktion im Modell der Logistischen Regression?
- $g(\mu) = \mu$
 - $g(\mu) = \ln\left(\frac{\mu}{1-\mu}\right)$
 - $g(\mu) = \frac{1}{1+\exp(-\mu)}$
 - $g(\mu) = (x - \mu)^T \Sigma^{-1}(x - \mu)$
27. Welche Aussage über die Parameter w und w_0 einer Hyperebene mit der Ebenengleichung $w^T x + w_0 = 0$ für $x \in \mathbb{R}^m$ ist **nicht** korrekt?
- Die Wahl von w legt die Orientierung der Hyperebene im Featureraum fest.
 - Die Wahl von w_0 legt die Lage der Hyperebene im Featureraum fest.
 - Die Parameter w und w_0 beschreiben eine Ebene im m -dimensionalen Raum.
 - Sowohl w als auch w_0 ist ein m -dimensionaler Vektor.
28. Was versteht man, im Kontext von Support-Vektor-Maschinen, unter einem Supportvektor?
- Der Supportvektor ist der Nullvektor 0_m .
 - Der Supportvektor ist der Gewichtsvektor w der linearen Diskriminanzfunktion.
 - Ein Supportvektor ist jeder Featurevektor, der mittels einer linearen Diskriminanzfunktion richtig zugeordnet wird und damit die Klassifikation unterstützt.
 - Ein Supportvektor ist ein Featurevektor, der unter allen Featurevektoren den kleinsten Abstand zur durch die lineare Diskriminanzfunktion induzierten Hyperebene hat.
29. Was besagen universelle Approximationstheoreme?
- Neuronale Netze können eine Vielzahl von Funktionen beliebig genau approximieren, wenn die Anzahl der Neurone gegen Unendlich geht oder die Anzahl der Neuronenschichten gegen Unendlich geht.
 - Neuronale Netze können alle beliebigen Datensätze linear separieren.
 - Neuronale Netze klassifizieren besser als Support-Vektor-Maschinen.
 - Parameter neuronaler Netze müssen durch Gradientenverfahren gelernt werden.
30. Die Formel für das Potential von Neuron i in Schicht l lautet: $z_i^l = \sum_{j=1}^{n_{l-1}} w_{ij}^l a_j^{l-1} + w_{i, n_{l-1}+1}^l$. Wofür steht die Variable w_{ij}^l in dieser Formel?
- für die Aktivierung von Neuron j in Schicht $l - 1$
 - für den Bias von Neuron i in Schicht l
 - für das Gewicht der Verbindung von Neuron j in Schicht $l - 1$ und Neuron i in Schicht l
 - für das Gewicht der Verbindung von Neuron i in Schicht $l - 1$ und Neuron j in Schicht l

Lösungen:

1. b)
2. d)
3. b)
4. d)
5. b)
6. a)
7. b)
8. a)
9. b)
10. d)
11. c)
12. b)
13. b)
14. a)
15. c)
16. a)
17. a)
18. d)
19. c)
20. a)
21. a)
22. a)
23. a)
24. c)
25. d)
26. b)
27. d)
28. d)
29. a)
30. c)