Hochschule München University of Applied Sciences

Prüfung Maschinellen Lernens – PROBEKLAUSUR

WS 2022/23

Prof. Dr. David Spieler

Bearbeitungzeit: 90 min

Hilfsmittel: Taschenrechner

Das Aufgabenblatt ist wieder abzugeben.

Tragen Sie Ihre Rechnungen und Ergebnisse auf diesem Aufgabenblatt ein.

Bitte kontrollieren Sie, ob Sie eine vollständige Aufgabenstellung mit 7 Seiten erhalten haben.

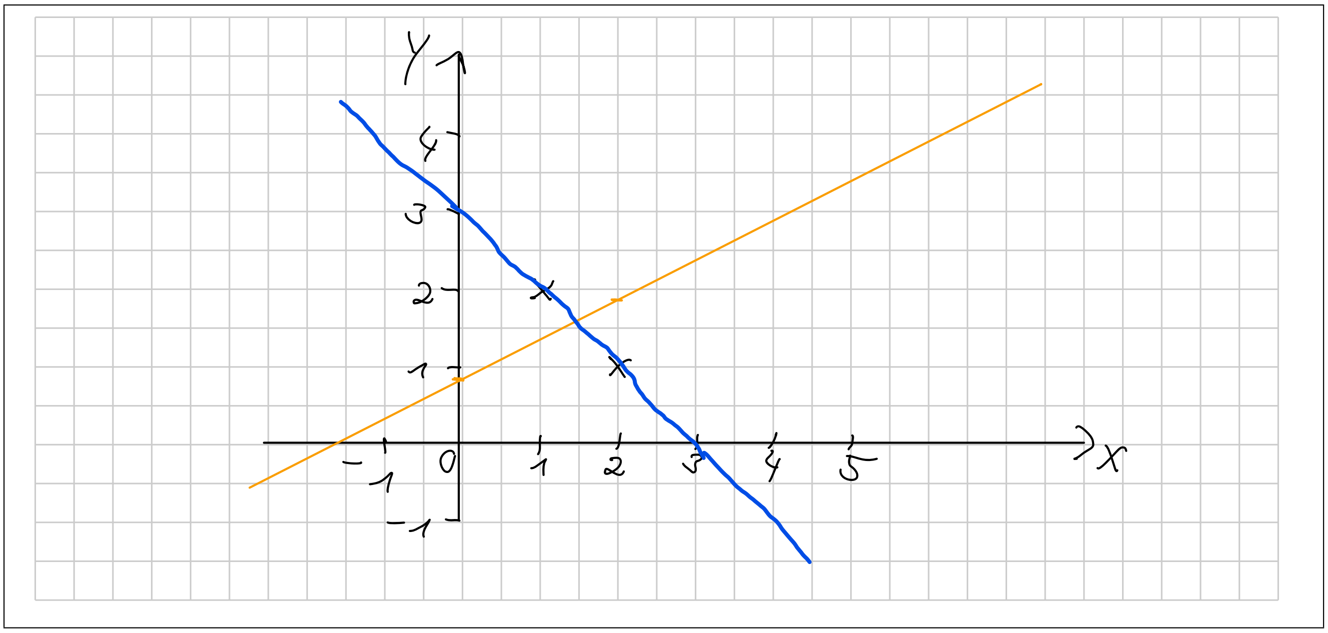
Name:	Vorname:
Matrikelnummer:	Platznummer:
Studiengang: ◯ Informatik ◯ Data Science & Scientific Computing ◯ Wirtschaftsinformatik	

erreichte Punktzahl:	Note:
Unterschrift Prüfer:	Zweitprüfer:

Aufgabe	1	2	3	4	5	$\sum$
Punkte	15	15	15	15	15	75
Erreichte Punkte						

Gegeben sei der Datensatz $\mathcal{D} = \{[1,2]^T, [2,1]^T\}$ für ein lineares Regressionsmodell $f : \mathbb{R} \to \mathbb{R}$ parametrisiert durch den Gewichtsvektor $\mathbf{w}$, also $f(x) = \mathbf{w}_1 x + \mathbf{w}_0$. Als Fehlerfunktion $E(\mathbf{w})$ verwenden Sie den mittleren quadratischen Fehler (MSE).

a) Zeichnen Sie die Punkte in ein Koordinatensystem.



b) Setzen Sie die Datenpunkte in die allgemeine Definition der Fehlerfunktion $E(\mathbf{w})$ ein. *Hinweis: Sie erhalten eine Funktion, welche nur von* $\mathbf{w}$ *abhängt.*

$$E(w) = \frac{1}{n} \sum_{i=1}^{2} \left(y^{(i)} - f_w(x^{(i)})\right)^2$$

$$E(w) = \frac{1}{2}\left[\left(2 - (1w_1 + w_0)\right)^2 + \left(1 - (2w_1 + w_0)\right)^2\right]$$

c) Berechnen Sie die partiellen Ableitungen der Fehlerfunktion, also $\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}_0}$ und $\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}_1}$.

$$\frac{\partial E(w)}{\partial w_0} = \frac{1}{2}\left[2 \cdot (2 - w_1 - w_0) \cdot (-1) + 2 \cdot (1 - 2w_1 - w_0) \cdot (-1)\right]$$

$$= \frac{1}{2}\left[-4 + 2w_1 + 2w_0 - 2 + 4w_1 + 2w_0\right]$$

$$= \frac{1}{2}\left[-6 + 6w_1 + 4w_0\right] = \underline{-3 + 3w_1 + 2w_0}$$

$$\frac{\partial E(w)}{\partial w_1} = \frac{1}{2}\left[2 \cdot (2 - w_1 - w_0) \cdot (-1) + 2 \cdot (1 - 2w_1 - w_0) \cdot (-2)\right]$$

4 w0

$$= \frac{1}{2}\left[-4 + 2w_1 + 2w_0 - 4 + 8w_1 + 2w_0\right]$$

$$= \frac{1}{2}\left[-8 + 10w_1 + 4w_0\right] = \underline{-4 + 5w_1 + 2w_0}$$

5w1 + 3w0 -4

d) Setzen die Gradienten $\nabla E(\mathbf{w}) = 0$ und lösen Sie das entstehende lineare Gleichungssystem.

$-3 + 3w_1 + 2w_0 = 0$

$-4 + 5w_1 + 2w_0 = 0$

$5w_1 = 4 - 2w_0$

$w_1 = 0{,}8 - 0{,}4\, w_0$

$2w_0 = 3 - 3w_1$

$w_0 = 1{,}5 - 1{,}5w_1$

$w_1 = 0{,}8 - 0{,}4 \cdot 0{,}75$

$w_0 = 1{,}5 - 1{,}5(0{,}8 - 0{,}4w_0)$

$w_1 = 0{,}5$

$w_0 = 1{,}5 - 1{,}2 + 0{,}6w_0$

$w_0 = 0{,}3 + 0{,}6\, w_0$

$0{,}4w_0 = 0{,}3$

$w_0 = 0{,}75$

e) Zeichnen Sie die Funktion f mit den Parametern aus Teil d) in das Koordinatensystem aus Teil a) ein.

wenn in teilaufgabe d) alles richtig berechnet wurde, dann kommt da
f(x) = -x+3 raus

Aufgabe 2: **(15 Punkte)**

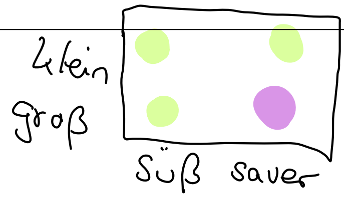
Gegeben sei der Datensatz

schaut gut aus

$$\mathcal{D} = \{([\text{klein}, \text{süß}]^T, \text{ja}), ([\text{groß}, \text{süß}]^T, \text{ja}), ([\text{klein}, \text{sauer}]^T, \text{ja}), ([\text{groß}, \text{sauer}]^T, \text{nein})\}.$$

a) Erstellen Sie einen Entscheidungsbaum mit der Entropie als Unreinheitsmaß und berechnen Sie für jeden Knoten $i(N)$ und für jedes Splitting $\Delta i(N)$.

$i(N) = -\sum_{i=1}^{n} y_i \log_2 y_i$ $\Delta i(N) = i(N) - P_L - P_R$



$i(N) = -\frac{3}{4}\log_2\left(\frac{3}{4}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) = 0,811$

Betrachtung Größe:

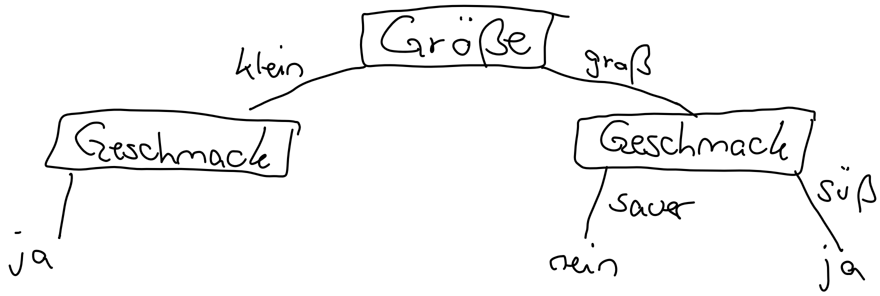
klein $= -\frac{2}{2}\log_2\left(\frac{2}{2}\right) - \frac{0}{2}\log_2\left(\frac{0}{2}\right) = 0$ $\Delta i(N) = 0,811 - \frac{2}{4}\cdot 0 - \frac{2}{4}\cdot 1 = 0,311$

groß $= -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$

Betrachtung Geschmack:

süß $= -\frac{2}{2}\log_2\left(\frac{2}{2}\right) - \frac{0}{2}\log_2\left(\frac{0}{2}\right) = 0$ $\Delta i(N) = 0,811 - \frac{2}{4}\cdot 0 - \frac{2}{4}\cdot 1 = 0,311$

sauer $= -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$



b) Angenommen Sie möchten Overfitting reduzieren. Erklären Sie kurz eine Möglichkeit dies im Fall eines Decision Trees zu tun?

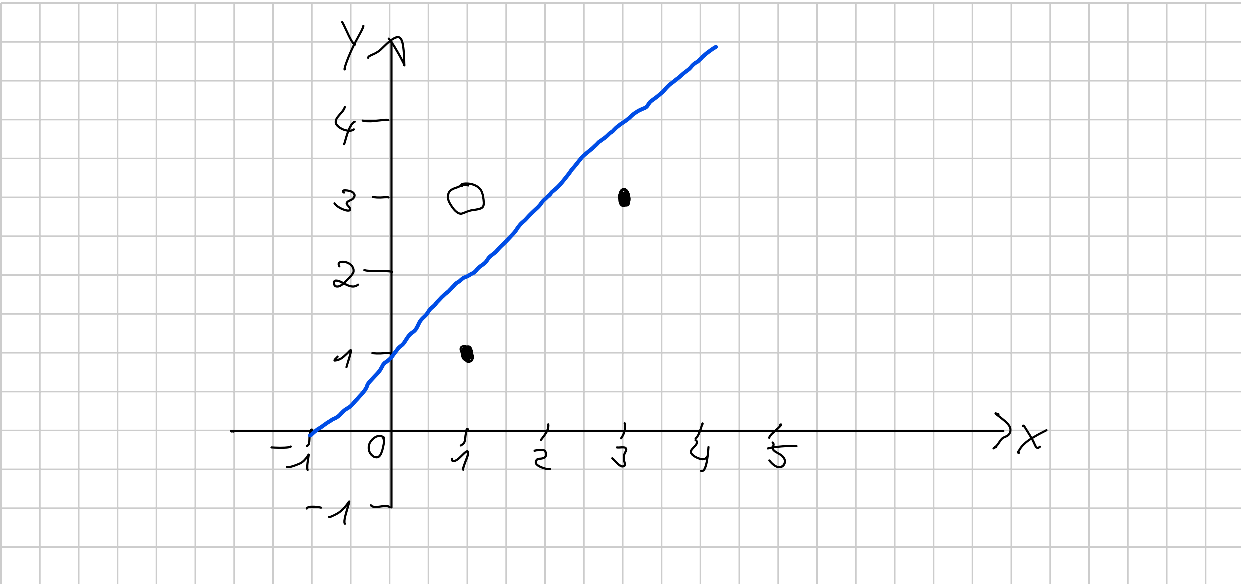
Stutzen des Baumes

Regeln einführen

Aufgabe 3: **(15 Punkte)**

Gegeben sei der Datensatz $\mathcal{D} = \{([1,1]^T, 1), ([3,3]^T, 1), ([1,3]^T, 0)\}$ und die Anfangsgewichte $\mathbf{w} = [\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2]^T = [1, -1, 1]^T$ eines Perzeptron-Klassifikators $f : \mathbb{R}^2 \to \{0, 1\}$.

a) Zeichnen Sie die Punkte in ein Koordinatensystem (Klasse 0: ○, Klasse 1: ●).



b) Wenden Sie den Perzeptron-Lernalgorithmus mit den oben stehenden Anfangsgewichten an, bis alle Trainingsdatenpunkte korrekt klassifiziert werden.

$P_1 = ([1,1]^T, 1)$, $P_2 = ([3,3]^T, 1)$, $P_3 = ([1,3]^T, 0)$ $w = [\overset{w_0}{1}, \overset{w_1}{-1}, \overset{w_2}{1}]^T$ | $w_1 \cdot x_1 + w_2 \cdot y_1 + w_0$

$P_1 = -1 \cdot 1 + 1 \cdot 1 + 1 = 1$ ✓
$P_2 = -1 \cdot 3 + 1 \cdot 3 + 1 = 1$ ✓
$P_3 = -1 \cdot 1 + 1 \cdot 3 + 1 = 3$ ↯

$w_{neu} = [1, -1, 1]^T - [1, 1, 3]^T = [\overset{w_0}{0}, \overset{w_1}{-2}, \overset{w_2}{-2}]^T$

$P_1 = -2 \cdot 1 + (-2) \cdot 1 + 0 = -4$ ↯
$P_2 = -2 \cdot 3 + (-2) \cdot 3 + 0 = -12$ ↯
$P_3 = -2 \cdot 1 + (-2) \cdot 3 + 0 = -8$ ✓

$w_{neu} = [0, -2, -2]^T + [1, 1, 1]^T + [1, 3, 3]^T = [\overset{w_0}{2}, \overset{w_1}{2}, \overset{w_2}{2}]^T$

$P_1 = 2 \cdot 1 + 2 \cdot 1 + 2 = 6$ ✓
$P_2 = 2 \cdot 3 + 2 \cdot 3 + 2 = 14$ ✓
$P_3 = 2 \cdot 1 + 2 \cdot 3 + 2 = 10$ ↯

$w_{neu} = [2, 2, 2]^T - [1, 1, 3]^T = [\overset{w_0}{1}, \overset{w_1}{1}, \overset{w_2}{-1}]^T$

$P_1 = 1 \cdot 1 + (-1) \cdot 1 + 1 = 1$ ✓
$P_2 = 1 \cdot 3 + (-1) \cdot 3 + 1 = 1$ ✓
$P_3 = 1 \cdot 1 + (-1) \cdot 3 + 1 = -1$ ✓

gesuchte Gewichtung $= [1, 1, -1]^T$

neue gewichte w = (1,1,-1) einsetzen in 0 = w1x1 + w2x2 + w0
0 = 1x1 -1 x2 + 1 -> nach x2 auflösen
x2 = x1 +1
f(x) = x +1
=> einzeichnen

c) Zeichnen Sie die Entscheidungsgrenze mit den Parametern aus Teil b) in das Koordinatensystem von Teil a) und schraffieren Sie den Bereich mit $f(\mathbf{x}) = 1$.

?

Aufgabe 4: **(15 Punkte)**

a) Ein Klassifikator klassifiziert die Testdatenpunkte $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ und $\mathbf{x}^{(3)}$ zu den Klassen -1, $+1$ und $+1$, wobei die echten Klassen $+1$, -1 und $+1$ sind. Berechnen Sie die Genauigkeit, Fehlerrate, Präzision und Trefferquote dieses Klassifikators bzgl. dieser Testdaten.

	x^1	x^2	x^3
klassi	-1	$+1$	$+1$
echt	$+1$	-1	$+1$

$$\text{Genauigkeit} = \frac{tp + tn}{tp + tn + fp + fn} = \frac{1 + 0}{1 + 0 + 1 + 1} = \frac{1}{3}$$

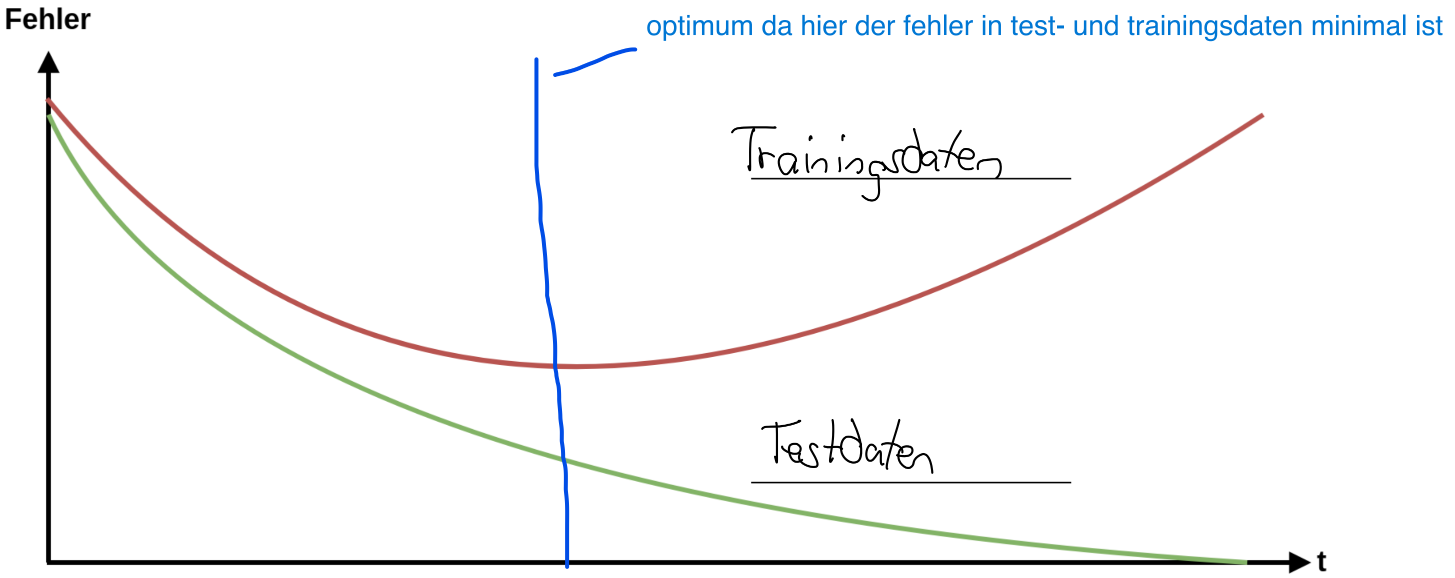
$$\text{Fehlerrate} = 1 - \text{Genauigkeit} = \frac{fp + fn}{tp + tn + fp + fn} = \frac{1 + 1}{1 + 0 + 1 + 1} = \frac{2}{3}$$

$$\text{Präzision:} \quad \frac{tp}{tp + fp} = \frac{1}{1 + 1} = \frac{1}{2}$$

Schaut gut aus

$$\text{Trefferquote:} \quad \frac{tp}{tp + fn} = \frac{1}{1 + 1} = \frac{1}{2}$$

b) Die folgende Abbildung zeigt eine typische Lernkurve mit dem Fehler über die Zeit. Tragen Sie die fehlenden Beschriftungen ein (zwei Arten von Daten, zwei Arten von Anpassungsproblemen an die Daten) und markieren Sie, wann Sie idealer Weise mit dem Training aufhören.



Überanpassung

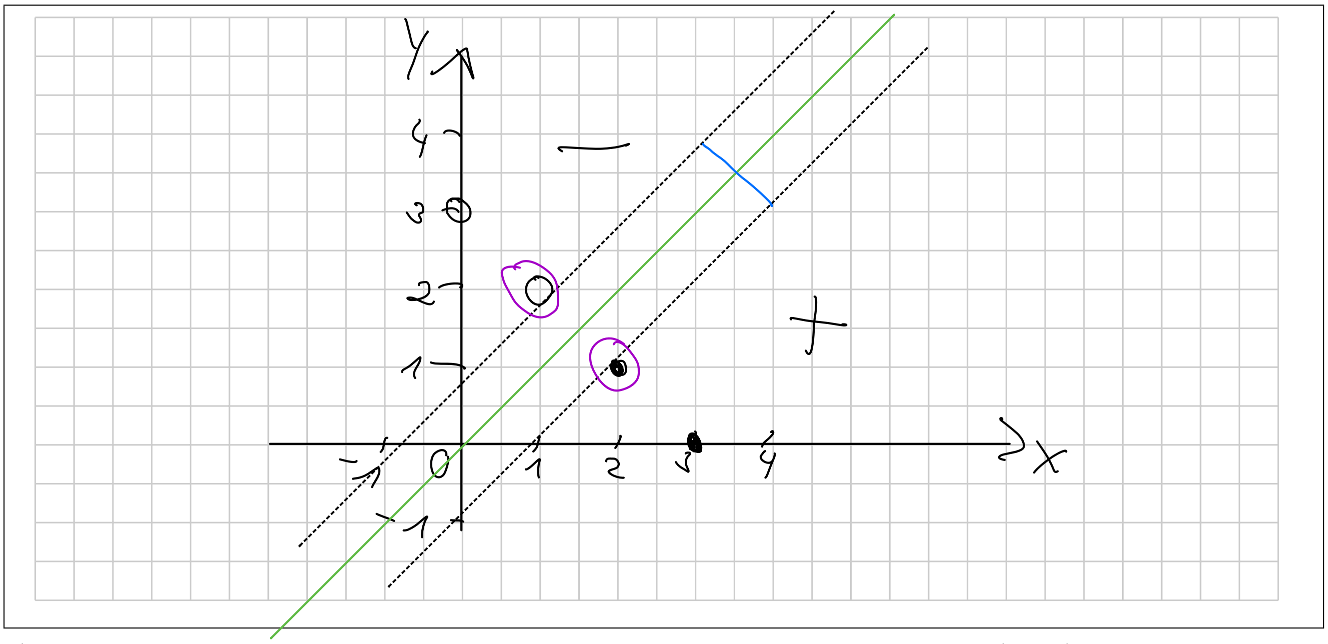
Unteranpassung

wenn das model zu sehr angepasst ist an die testdaten (also der fehler der testdaten gegen 0 geht), dann ist das überanpassung weil das model nicht mehr die testdaten repräsentieren kann

Aufgabe 5: **(15 Punkte)**

Gegeben sei der Datensatz $\mathcal{D} = \{([0,3]^T, -1), ([1,2]^T, -1), ([2,1]^T, 1), ([3,0]^T, 1)\}$.

a) Zeichnen Sie die Punkte in ein Koordinatensystem (Klasse -1: $\circ$, Klasse 1: $\bullet$).



b) Zeichnen Sie die Entscheidungsgrenze des Support-Vektor-Klassifikators (SVC) $f : \mathbb{R}^2 \to \{-1, 0, 1\}$ trainiert auf diesen Trainingsdaten ein. Kennzeichnen Sie die Region, die positiv (1) klassifiziert wird mit einem +-Zeichen und die Region, die negativ (-1) klassifiziert wird mit einem $-$-Zeichen.

c) Zeichnen Sie den Margin in der Abbildung von Teilaufgabe a) ein und umranden Sie die Support-Vektoren.

d) Bestimmen Sie eine Gleichung, welche alle Punkte $[x_1, x_2] \in \mathbb{R}^2$ auf der Entscheidungsgrenze beschreibt und entwickeln Sie daraus den SVC $f : \mathbb{R}^2 \to \{-1, 0, 1\}$.

die funktionsgleichung kann man hier einfach aus der Zeichnung entnehmen
-> f(x) = x ,umgeschrieben x2 = x1
0 = x1 - x2

einsetzen eines punktes....
f(2,1) = 2-1 = 1
f(1, 2) = 1-2 = -1
=> korrekt klassifiziert

zumindest glaube ich dass die Aufgabe so funktioniert.