

Machine Learning

Aufgabe 1. *ReLU Aktivierungsfunktion*

10 P.

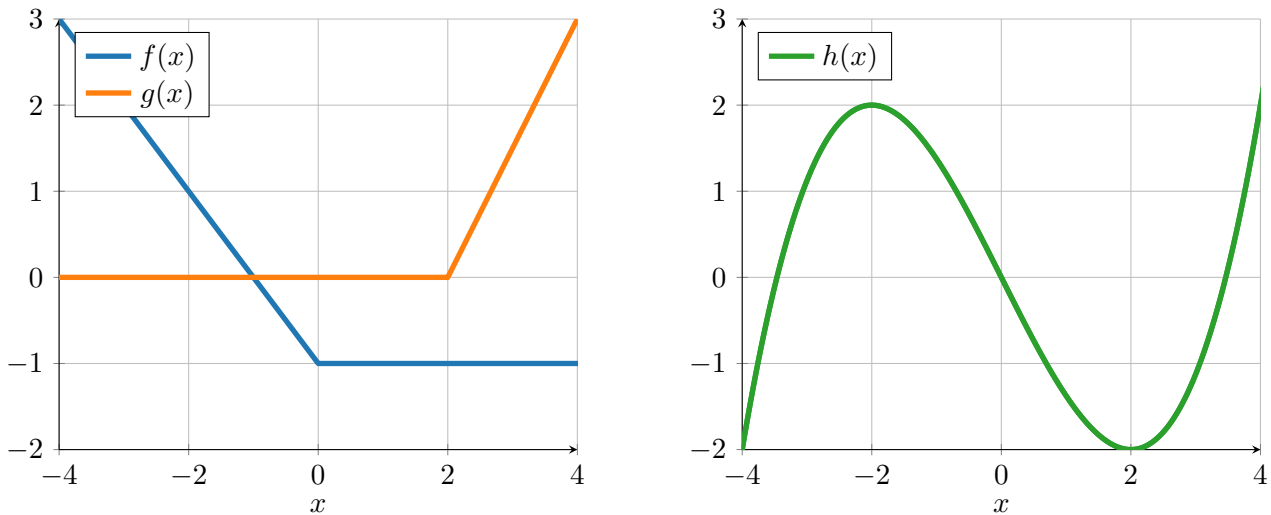


Abbildung 1:

- (a) Schreiben Sie die Funktionsgleichungen der in Abbildung 1 (links) dargestellten Funktionen mithilfe der ReLU-Funktion.
- (b) Approximieren Sie die in Abbildung 1 (rechts) dargestellte Funktion h durch eine affine Summe von 3 ReLU-Funktionen, also als

$$\tilde{h}(x) = \sum_{i=1}^3 a_i \cdot \text{ReLU}(b_i \cdot x + c_i) ,$$

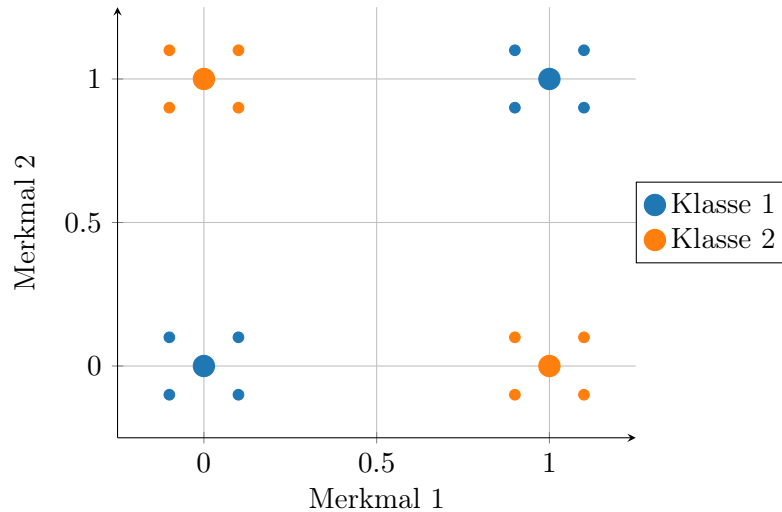
sodass der Fehler $\text{err}(\tilde{h}) := \max_{x \in [-4,4]} |h(x) - \tilde{h}(x)| < 1.5$ ist. Zeichnen Sie die Funktionen \tilde{h} in Abbildung 1 (rechts) ein und erklären Sie ihr Vorgehen.

- (c) Geben Sie die Architektur, sowie die Gewichtsmatrizen und Biasvektoren eines neuronalen Netzwerks an, das die Funktion \tilde{h} realisiert.
- (d) Gibt es ein neuronales Netzwerk der gleichen Architektur, aber mit anderen Gewichten und Biasvektoren, das die Funktion \tilde{h} ebenfalls realisiert? Begründen Sie Ihre Antwort.
- (e) Gibt es ein neuronales Netzwerk \tilde{h}_1 mit ReLU-Aktivierungsfunktionen, das die Funktion h exakt realisiert, also mit $\text{err}(\tilde{h}_1) = 0$? Begründen Sie Ihre Antwort.

Aufgabe 2. XOR Problem

10 P.

Gegeben sind die folgenden Daten.



(a) Geben Sie an, ob diese (Trainings-)Daten durch die folgenden Modelle zu 100% korrekt klassifiziert werden können. Begründen Sie ihre Antworten.

- (i) Lineare Diskriminanzanalyse
- (ii) Quadratische Diskriminanzanalyse
- (iii) Logistische Regression
- (iv) Logistische Regression mit quadratischen Merkmalen

Für Punkt (iv) genügt es die nur 4 groß gezeichneten Datenpunkte zu betrachten.

Wir betrachten nun ein neuronales Netz der Form

$$f(x) = \arg \max(W_2 \operatorname{ReLU}(W_1 x + b_1)) ,$$

das nur auf den 4 groß gezeichneten Datenpunkten trainiert wird. Dabei ist $b_1 \in \mathbb{R}^2$ und $W_1, W_2 \in \mathbb{R}^{2 \times 2}$.

- (b) Bestimmen Sie eine mögliche Wahl der Parameter W_1, W_2, b_1 , sodass das neuronale Netz die Daten zu 100% korrekt klassifiziert. Erklären Sie ihr Vorgehen.
- (c) Fertigen Sie eine handschriftliche Abbildung der Trainingsdaten und der gelernten Repräsentationen $\operatorname{ReLU}(W_1 x + b_1)$ an. Zeichnen Sie in beide Abbildungen die Entscheidungsregion ein.
- (d) Wiederholen Sie die Teilaufgaben (b) und (c) für eine zweite Wahl der Parameter. Insbesondere sollte ihre neue Wahl für W_1 nicht ein positives Vielfaches des vorherigen W_1 sein.