1. (2 Punkte) Ein Perzeptron hat die Parameter $w_1 = 2$, $w_2 = -1$, $b = -3$. Es feuert, wenn $z = w_1x_1 + w_2x_2 + b \geq 0$ ist. Berechne z und die Ausgabe a für die Punkte $(1, 1)$, $(3, 2)$, $(2, 1)$, $(0, -4)$.

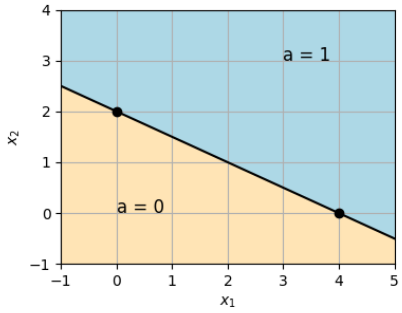
Lösung: $(1, 1)$: $z = -2$, $a = 0$
$(3, 2)$: $z = 1$, $a = 1$
$(2, 1)$: $z = 0$, $a = 1$
$(0, -4)$: $z = 1$, $a = 1$
Beim Punkt $(2, 1)$ ist $z = 0$, er liegt genau auf der Trenngeraden. Das Perzeptron feuert dort.

2. (3 Punkte) Bestimme Parameter w_1, w_2, b für ein Perzeptron, das für die Punkte $(2, 4)$ und $(6, 3)$ feuert, aber nicht für den Punkt $(6, 4)$. Zeichne die Punkte und deine Trenngerade. Überprüfe deine Parameter mit Python.

Lösung: Zum Beispiel $w_1 = -1$, $w_2 = -1$, $b = 9.5$. Dann ist $z = 9.5 - x_1 - x_2$, das Perzeptron feuert für $x_1 + x_2 \leq 9.5$. Die Trenngerade ist $x_2 = -x_1 + 9.5$.
$(2, 4)$: $z = 3.5$, $a = 1$ $(6, 3)$: $z = 0.5$, $a = 1$ $(6, 4)$: $z = -0.5$, $a = 0$
Es gibt viele weitere Lösungen, z.B. $w_1 = -1$, $w_2 = -2$, $b = 13.5$. Wichtig ist nur, dass die Gerade den Punkt $(6, 4)$ von den beiden anderen Punkten trennt.

3. (3 Punkte) Ein Perzeptron hat die Parameter $w_1 = 1$, $w_2 = 2$, $b = -4$.
   (a) Bestimme die Gleichung der Trenngeraden in der Form $x_2 = mx_1 + n$ und ihre Schnittpunkte mit den Achsen.
   (b) Zeichne die Gerade. Auf welcher Seite feuert das Perzeptron?
   (c) Was gibt das Perzeptron für die Punkte $(1, 1)$ und $(2, 1)$ aus?

Lösung:
(a) $x_1 + 2x_2 - 4 = 0$, also $x_2 = -0.5x_1 + 2$. Schnittpunkte $(0, 2)$ und $(4, 0)$.
(b) Oberhalb der Geraden (z.B. bei $(4, 4)$ ist $z = 8 \geq 0$).



(c) $(1, 1)$: $z = -1$, $a = 0$. $(2, 1)$: $z = 0$, $a = 1$, der Punkt liegt auf der Geraden.

4. (3 Punkte) Bestimme Parameter für Perzeptronen, die die folgenden logischen Funktionen berechnen. Die Inputs sind 0 oder 1.
   (a) AND mit zwei Inputs
   (b) NAND (NOT AND) mit zwei Inputs
   (c) NOT mit einem Input x (Parameter w und b)

Lösung:
(a) AND: $w_1 = 1$, $w_2 = 1$, $b = -1.5$. Die Summe $x_1 + x_2$ ist nur bei $(1, 1)$ mindestens 1.5.
(b) NAND: $w_1 = -1$, $w_2 = -1$, $b = 1.5$. Das ist AND mit umgedrehten Vorzeichen.
(c) NOT: $w = -1$, $b = 0.5$. Für $x = 0$ ist $z = 0.5$, für $x = 1$ ist $z = -0.5$.
Es gibt jeweils viele weitere Lösungen.

5. (3 Punkte) Gegeben sind zwei Perzeptronen. P hat die Parameter $w_1 = -1$, $w_2 = -1$, $b = 0.5$, Q hat die Parameter $w_1 = 1$, $w_2 = -1$, $b = -0.5$. Fülle die Tabelle aus. Welche logischen Funktionen berechnen P und Q?

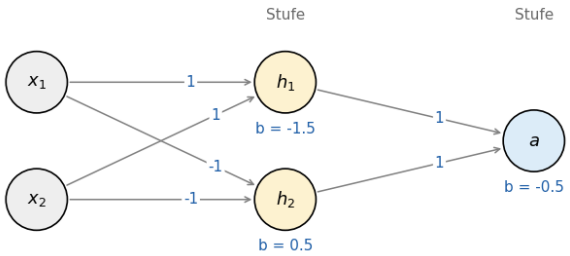
x_1	x_2	P	Q
0	0	1	0
0	1	0	0
1	0	0	1
1	1	0	0

Lösung: P ist NOR (NOT OR): Es feuert nur, wenn beide Inputs 0 sind. Q feuert nur bei $x_1 = 1$ und $x_2 = 0$, es berechnet x_1 AND NOT x_2.

6. (3 Punkte) Beweise, dass kein Perzeptron mit zwei Inputs die Funktion XOR berechnen kann. Schreibe dazu für jede der vier Zeilen der Wahrheitstafel eine Ungleichung für w_1, w_2 und b auf und führe sie zu einem Widerspruch.

Lösung: Das Perzeptron feuert bei $z \geq 0$. XOR verlangt:
$(0, 0) \to 0$: $b < 0$ $(0, 1) \to 1$: $w_2 + b \geq 0$ $(1, 0) \to 1$: $w_1 + b \geq 0$ $(1, 1) \to 0$: $w_1 + w_2 + b < 0$
Addiert man die zweite und dritte Ungleichung: $w_1 + w_2 + 2b \geq 0$, also $w_1 + w_2 + b \geq -b$. Wegen $b < 0$ ist $-b > 0$, also $w_1 + w_2 + b > 0$. Das widerspricht der vierten Ungleichung.
Geometrisch: Die Punkte $(0, 1)$ und $(1, 0)$ lassen sich nicht durch eine Gerade von $(0, 0)$ und $(1, 1)$ trennen. XOR ist nicht linear trennbar.

7. (4 Punkte) Drei Perzeptronen sind zu einem Netz verbunden. Alle feuern bei $z \geq 0$.



h_1 hat die Parameter $w_1 = 1$, $w_2 = 1$, $b = -1.5$, h_2 hat die Parameter $w_1 = -1$, $w_2 = -1$, $b = 0.5$, das Ausgabeperzeptron a hat die Parameter $v_1 = 1$, $v_2 = 1$, $c = -0.5$.

(a) Fülle die Tabelle aus.

x_1	x_2	h_1	h_2	a
0	0	0	1	1
0	1	0	0	0
1	0	0	0	0
1	1	1	0	1

(b) Welche logischen Funktionen berechnen h_1, h_2 und das ganze Netz?

(c) Warum kann ein einzelnes Perzeptron diese Funktion nicht berechnen?

Lösung:

(b) h_1 ist AND, h_2 ist NOR, a ist OR von h_1 und h_2. Das Netz gibt 1 aus, wenn beide Inputs gleich sind. Das ist das Gegenteil von XOR (XNOR).

(c) Wie bei XOR liegen die Punkte mit Ausgabe 1, $(0, 0)$ und $(1, 1)$, diagonal gegenüber. Keine Gerade trennt sie von $(0, 1)$ und $(1, 0)$.

8. (4 Punkte) Ein Perzeptron soll mit der Lernregel die Funktion AND lernen. Am Anfang ist $w_1 = w_2 = b = 0$, die Lernrate ist $lr = 1$. Das Perzeptron feuert bei $z \geq 0$. Die Lernregel lautet

$$w_1 \leftarrow w_1 + lr \cdot (y - a) \cdot x_1, \quad w_2 \leftarrow w_2 + lr \cdot (y - a) \cdot x_2, \quad b \leftarrow b + lr \cdot (y - a)$$

   (a) Führe zwei Epochen durch und trage die Werte ein. Die letzten drei Spalten enthalten die Parameter nach dem Schritt.

Epoche	(x_1, x_2)	y	z	a	w_1	w_2	b
1	$(0,0)$	0	0	1	0	0	-1
1	$(0,1)$	0	-1	0	0	0	-1
1	$(1,0)$	0	-1	0	0	0	-1
1	$(1,1)$	1	-1	0	1	1	0
2	$(0,0)$	0	0	1	1	1	-1
2	$(0,1)$	0	0	1	1	0	-2
2	$(1,0)$	0	-1	0	1	0	-2
2	$(1,1)$	1	-1	0	2	1	-1

   (b) Nach sechs Epochen ist $w_1 = 2$, $w_2 = 1$, $b = -3$, die sechste Epoche läuft ohne Fehler. Prüfe, dass das Perzeptron mit diesen Parametern AND berechnet.

Lösung:

(b) $z = -3, -2, -1, 0$ für $(0,0), (0,1), (1,0), (1,1)$. Nur bei $(1,1)$ ist $z \geq 0$, also ist das AND.

9. (3 Punkte) Ein Neuron mit der Sigmoidfunktion $\sigma(z) = \frac{1}{1+e^{-z}}$ hat die Parameter $w_1 = 2$, $w_2 = 2$, $b = -1$.

   (a) Berechne die Ausgabe a für die Punkte $(0,0)$, $(0,1)$, $(1,1)$.

   (b) Welcher Klasse ordnet das Neuron die Punkte zu, wenn es bei $a \geq 0.5$ die Klasse 1 wählt? Bei welchem Punkt ist es sich am sichersten?

   (c) Welche logische Funktion berechnet das Neuron?

Lösung:

(a) $(0,0)$: $\sigma(-1) = 0.2689$ $(0,1)$: $\sigma(1) = 0.7311$ $(1,1)$: $\sigma(3) = 0.9526$

(b) $(0,0)$ zur Klasse 0, $(0,1)$ und $(1,1)$ zur Klasse 1. Am sichersten ist es bei $(1,1)$, dort ist a am weitesten von 0.5 entfernt.

(c) OR: $(1,0)$ ergibt wie $(0,1)$ den Wert $\sigma(1)$.

10. (3 Punkte) Ein Perzeptron und ein Neuron haben beide die Parameter $w_1 = -1$, $w_2 = 1$, $b = -1.5$. Für den Punkt $(1, 0)$ soll die Ausgabe 1 sein.
    (a) Berechne die Ausgabe von Perzeptron und Neuron für $(1, 0)$.
    (b) Jetzt wird $w_1 = -0.5$ gesetzt. Berechne die Ausgaben erneut.
    (c) Warum ist das Neuron für das Lernen besser geeignet?

Lösung:

(a) $z = -2.5$. Perzeptron: $a = 0$. Neuron: $a = \sigma(-2.5) = 0.0759$.

(b) $z = -2$. Perzeptron: $a = 0$. Neuron: $a = \sigma(-2) = 0.1192$.

(c) Beim Perzeptron ändert sich die Ausgabe nicht, man sieht nicht, dass die Änderung in die richtige Richtung ging. Beim Neuron steigt die Ausgabe in Richtung 1. Man kann also erkennen, in welche Richtung man einen Parameter verändern muss.

11. (4 Punkte) (a) Welche Rolle spielen die Gewichte und der Bias bei einem Perzeptron?
    (b) Was bedeutet linear trennbar? Gib ein Beispiel für Daten, die nicht linear trennbar sind.
    (c) Warum funktioniert die Perzeptron-Lernregel nicht für ein Netz mit einer versteckten Schicht?
    (d) Was ist der Unterschied zwischen einem Perzeptron und einem Neuron?

Lösung:

(a) Die Gewichte geben an, wie stark jeder Input in die Summe eingeht, auch mit negativem Vorzeichen. Der Bias ist der negative Schwellwert: Je größer er ist, desto leichter feuert das Perzeptron. Geometrisch legen die Gewichte die Richtung der Trenngeraden fest, der Bias verschiebt sie.

(b) Die Punkte der beiden Klassen lassen sich durch eine Gerade trennen. Die vier Punkte von XOR sind nicht linear trennbar.

(c) Die Lernregel braucht für jedes Perzeptron die richtige Ausgabe y. Die kennt man nur für die Ausgabeschicht, nicht für die versteckten Perzeptronen.

(d) Das Perzeptron benutzt die Stufenfunktion und gibt nur 0 oder 1 aus. Das Neuron benutzt die Sigmoidfunktion, seine Ausgabe liegt stetig zwischen 0 und 1 und ändert sich bei kleinen Änderungen der Parameter ein wenig.

12. (4 Punkte) Die Mehrheitsfunktion mit drei Inputs $x_1, x_2, x_3 \in \{0, 1\}$ gibt 1 aus, wenn mindestens zwei Inputs 1 sind. Schreibe ein Programm, das mit der Perzeptron-Lernregel ein Perzeptron mit drei Inputs für die Mehrheitsfunktion trainiert. Speichere die Gewichte in einer Liste `w`. Starte mit allen Parametern 0 und $lr = 1$. Gib die gelernten Parameter und die Zahl der Epochen aus und prüfe das Ergebnis für alle acht Inputs.

Lösung:

```
X = [(x1, x2, x3) for x1 in [0, 1] for x2 in [0, 1] for x3 in [0, 1]]
Y = [1 if sum(x) >= 2 else 0 for x in X]

def perzeptron(x, w, b):
    z = sum(w[i] * x[i] for i in range(len(x))) + b
    return 1 if z >= 0 else 0

w = [0, 0, 0]
b = 0
lr = 1
for epoche in range(1, 101):
    fehler = 0
    for x, y in zip(X, Y):
        a = perzeptron(x, w, b)
        if a != y:
            fehler += 1
            w = [w[i] + lr * (y - a) * x[i] for i in range(3)]
            b = b + lr * (y - a)
    if fehler == 0:
        break

print(w, b, epoche)
print(all(perzeptron(x, w, b) == y for x, y in zip(X, Y)))
```

Ergebnis: `w = [1, 1, 1]`, `b = -2` nach 4 Epochen. Das Perzeptron berechnet die Mehrheitsfunktion für alle acht Inputs richtig.