Einführung in die Neuroinformatik SoSe 2019

Institut für Neuroinformatik
PD Dr. F. Schwenker

2. Aufgabenblatt (Abgabe bis 21. Mai 2019 zur Vorlesung)

Aufgabe 1 (3 Punkte): Boole'sche Funktionen [Pen and Paper]

In diesem Übungsblatt wollen wir uns mit einfachen Neuronenmodellen, deren Transferfunktionen und dem Zusammenspiel von Neuronen in einfachen Netzen auseinandersetzen. In diesem ersten Teil konzentrieren wir uns auf einfache Boole'sche Funktionen, welche wir dann später noch benötigen werden.

Wir betrachten ein Schwellwertneuron, das drei binäre Variablen $\boldsymbol{x} = (x_1, x_2, x_3) \in \{0, 1\}^3$ als Input erhält und einen Output

$$y(\boldsymbol{x}) = \begin{cases} 1 & \langle \boldsymbol{x}, \boldsymbol{w} \rangle \geq \theta \\ 0 & \text{sonst} \end{cases}$$

erzeugt. Jedes Schwellwertneuron wird über drei Gewichte $\boldsymbol{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ sowie einen Schwellwert $\theta \in \mathbb{R}$ definiert.

1. Wie muss der Gewichtsvektor $\boldsymbol{w}$ sowie der Schwellwert θ definiert werden, so dass das Neuron zu einem
   a) logischen AND-Gatter wird.
   b) logischen OR-Gatter wird.
2. Ist die Darstellung in beiden Fällen eindeutig oder hätte der Gewichtsvektor $\boldsymbol{w}$ und der Schwellwert θ auch anders gewählt werden können?
3. Mit Hilfe von Schwellwertneuronen können wir Entscheidungen auf Basis von Eingangsvariablen treffen, indem wir die Variablen entsprechend gewichten. Dies wollen wir nun anhand eines konkreten Beispieles verdeutlichen. Stellt euch dazu bitte vor, dass es in eurer Stadt demnächst eine Veranstaltung geben soll und ihr euch fragt, ob ihr diese besuchen wollt oder nicht. Ihr habt euch dabei bereits für drei (binäre) Kriterien entschieden, die für euch relevant sind:
   - x_1: Wie ist das Wetter an dem Tag? (0 = schlecht, 1 = gut)
   - x_2: Geht ihr allein oder in Begleitung? (0 = allein, 1 = in Begleitung)
   - x_3: Gibt es etwas zu essen? (0 = hungern, 1 = nicht hungern)

   Das Wetter x_1 ist euch besonders wichtig: ihr verlasst das Haus nicht, wenn das Wetter schlecht ist. Bei den anderen beiden Kriterien seid ihr flexibler: hier reicht es, wenn eines der beiden zutrifft. In einer Boole'schen Funktion ausgedrückt heißt das also

   $$V(\boldsymbol{x}) = x_1 \wedge (x_2 \vee x_3)$$

   wobei ihr bei $V(\boldsymbol{x}) = 0$ nicht und bei $V(\boldsymbol{x}) = 1$ zu der Veranstaltung geht.

a) Stelle dieses Entscheidungsproblem mit einem Netz aus mindestens zwei Schwellwertneuronen dar. Greife dazu ggf. auf die AND- und OR-Neuronen von zuvor zurück. Hinweis: binäre Gewichte sind hier ausreichend.

b) Mit einer geschickten Wahl der (nicht mehr binären) Gewichte und des Schwellwertes lässt sich diese Situation auch mit einem einzigen Schwellwertneuron darstellen. Finde dazu eine passende Parameterkonstellation $(\boldsymbol{w}, \theta)$.

Aufgabe 2 (4 Punkte): Schwellwertneuronen [Pen and Paper]

Im Vergleich zur vorherigen Aufgabe wollen wir nun nicht mehr von binären Eingangsvariablen ausgehen. Stattdessen nehmen wir an, dass diese in einem Wertebereich von $[0; 1]$ liegen. In diesem Sinne können die Variablen als Wahrscheinlichkeit interpretiert werden, z. B. ist bei $x_1 = 0.4$ in 40 % der Fälle gutes Wetter zu erwarten.

1. Ihr wollt zur Veranstaltung gehen, wenn das Wetter mindestens mit einer Wahrscheinlichkeit von 50 % (≥ 0.5) gut wird und wenigstens eines der beiden anderen Kriterien zu 25 % (≥ 0.25) zutrifft. Stelle diese Situation mit einem Netzwerk aus Schwellwertneuronen dar.

2. Nehmen wir nun an, ihr wollt nur noch zur Veranstaltung gehen, wenn die Summe der Eingangsvariablen exakt 1 betragt, d. h.

$$V(\boldsymbol{x}) = \begin{cases} 1 & \sum_{i=1}^{3} x_i = 1 \\ 0 & \text{sonst} \end{cases}$$

   Grafisch wird dadurch eine Ebene im $\mathbb{R}^3$ aufgespannt (vgl. Abbildung 1). Konstruiere ein Netz aus Schwellwertneuronen, welches diese Ebene erzeugt, d. h. $V(\boldsymbol{x})$ berechnet. Tipp: es ist vielleicht hilfreich, zuerst die beiden anderen Volumenbereiche in Abbildung 1 zu berechnen und dann ein weiteres Neuron zu nehmen, das auf die Gemeinsamkeiten reduziert.

Aufgabe 3 (3 Punkte): Logistische Neuronen [Pen and Paper]

Schwellwertneuronen erzeugen immer eine „harte" Ausgabe in dem Sinne, dass das Ergebnis stets binär ist. Insbesondere können dadurch Sprünge und Diskontinuitäten entstehen. In vielen Fällen hätte man jedoch lieber eine kontinuierliche Ausgabe. Eine Möglichkeit besteht darin, auf logistische Neuronen zurückzugreifen, die als Aktivierungsfunktion die logistische Sigmoid-Funktion verwenden

$$f(x) = \frac{1}{1 + e^{-x}}.$$

Wir wollen das Verhalten dieser Funktion anhand des einfachen Beispielnetzwerkes in Abbildung 2 veranschaulichen. Das Netzwerk erhält nur eine Eingabe $x \in \mathbb{R}$, die an zwei Neuronen in der versteckten Schicht weitergereicht wird. Das Ergebnis dieser beiden Neuronen wird an ein Ausgabeneuron mit linearer Transferfunktion, definierten Gewichten von 1 und einem Bias von 0 weitergereicht. Die Gesamtausgabe des Netzwerks bezeichnen wir mit $y(x)$.

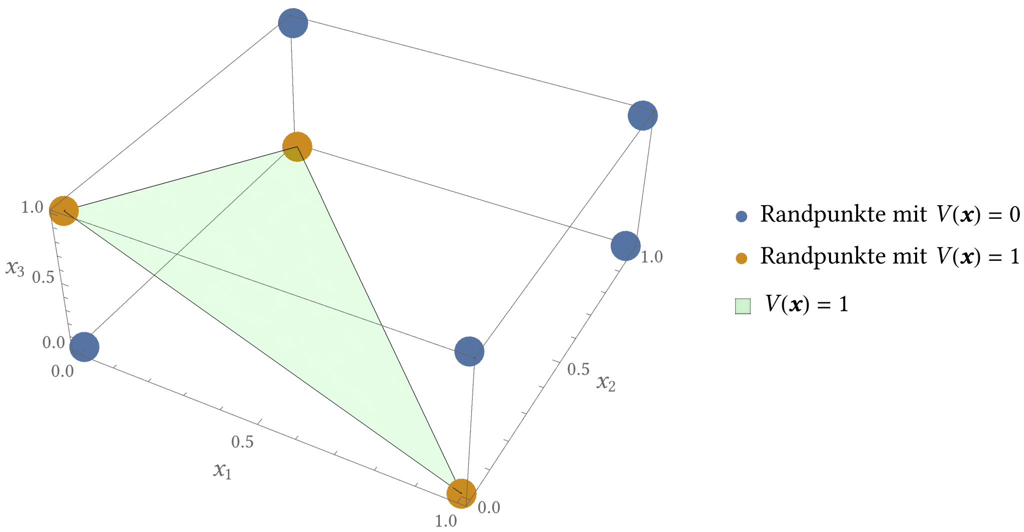

Abbildung 1: Ebene im $\mathbb{R}^3$, die durch die Funktion $V(\boldsymbol{x})$ aufgespannt wird.

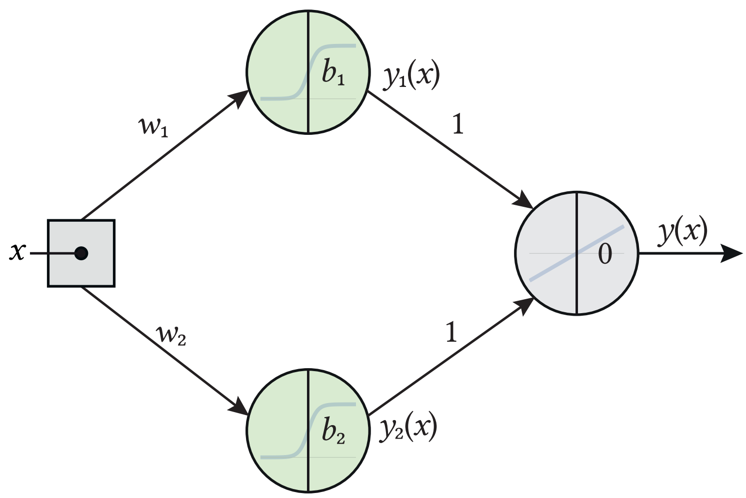

Abbildung 2: Beispielnetzwerk aus Aufgabe 3.

1. Betrachten wir zuerst nur die Ausgabe $y_1(x)$ des oberen Neurons in der versteckten Schicht

$$y_1(x) = f\left(w_1 \cdot x + b_1\right).$$

Wir wollen zuerst verstehen, wie das Gewicht w_1 und der Bias b_1 die Form der logistischen Funktion beeinflussen. Rufe dazu die Animation aus dem Buch Neural Networks and Deep Learning[1] von Michael A. Nielsen auf. Das dort dargestellte Netzwerk ähnelt dem hier verwendeten. Insbesondere lässt sich das Gewicht w_1 und der Bias b_1 beeinflussen (Zahlen mit der Maus nach links und rechts ziehen) und im rechten Teil die Ausgabe $y_1(x)$ ablesen.

   a) Wie wirkt sich eine Erhöhung von w_1 auf die Funktion $y_1(x)$ aus?

   b) Was passiert, wenn w_1 negativ wird?

   c) Mit Hilfe des Bias b_1 kann die Funktion $y_1(x)$ horizontal verschoben werden. Bestimme für Veränderungen des Bias (Erhöhung ↑ bzw. Verringerung ↓) wie sich die Verschiebung (nach rechts oder links) für unterschiedliche Vorzeichen von w_1 auswirkt:

	$w_1 > 0$	$w_1 < 0$
$b_1 \uparrow$		
$b_1 \downarrow$		

2. Die wahren Stärken von neuronalen Netzwerken entstehen durch das Zusammenspiel vieler einzelner Neuronen in unterschiedlichen Schichten. Abbildung 3 zeigt den Output

$$y(x) = y_1(x) + y_2(x) = f(w_1 \cdot x + b_1) + f(w_2 \cdot x + b_2)$$

für unser Netzwerk für eine bestimmte Parameterkonstellation. Welche Parameter w_1, b_1 für das erste und w_2, b_2 für das zweite Neuron könnten die Funktion erzeugen? Hinweis: alle Parameter haben entweder den Wert $+1$ oder -1, d. h. $w_1, b_1, w_2, b_2 \in \{-1, 1\}$.

[1] http://neuralnetworksanddeeplearning.com/chap4.html#basic_manipulation

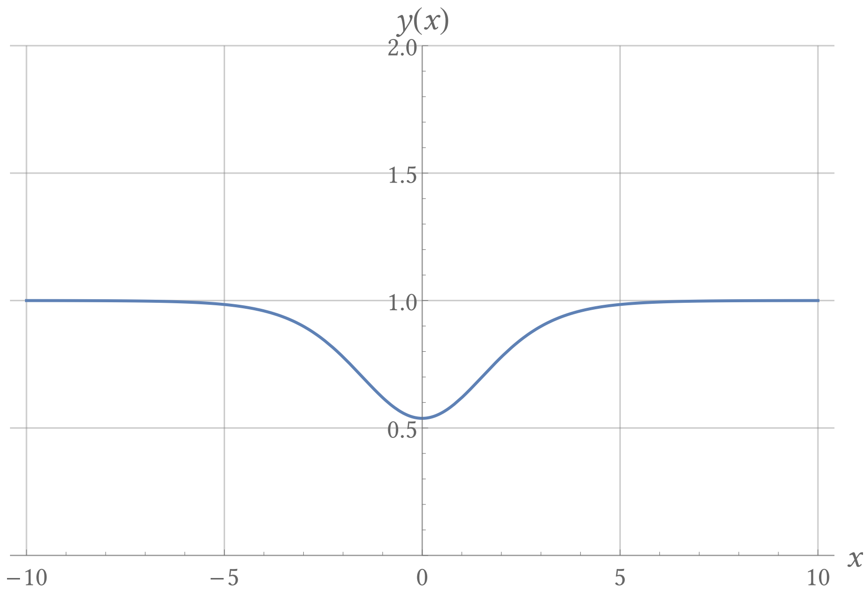

Abbildung 3: Ausgabe des Netzwerkes, welches in Aufgabe 3 verwendet wird.