

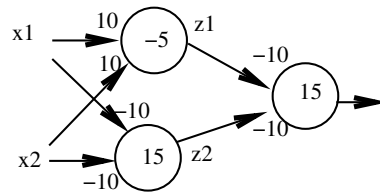
Übungen zu Ideen der Informatik

<https://www.mpi-inf.mpg.de/departments/algorithms-complexity/teaching/winter19/ideen/>

Blatt 12

Abgabeschluss: 20.1.2020

Aufgabe 1 (5 Punkte) Vervollständigen Sie die Tabelle und geben Sie an welche logische Funktion das abgebildete Netzwerk berechnet wird?



x_1	x_2	$z_1 =$	$z_1 \approx$	$z_2 =$	$z_2 \approx$	$o =$	$o \approx$
0	0	$g(-5)$	0				
0	1						
1	0						
1	1						

Lösung:

x_1	x_2	$z_1 =$	$z_1 \approx$	$z_2 =$	$z_2 \approx$	$o =$	$o \approx$
0	0	$g(-5)$	0	$g(15)$	1	$g(5)$	1
0	1	$g(5)$	1	$g(5)$	1	$g(-5)$	0
1	0	$g(5)$	1	$g(5)$	1	$g(-5)$	0
1	1	$g(15)$	1	$g(-5)$	0	$g(5)$	1

Das Netz berechnet die Funktion $x_1 \equiv x_2$. (Die Antworten $x_1 = x_2$ oder x_1 gleich x_2 sind auch OK.)

Aufgabe 2 (10 Punkte) Wir haben zwei Gruppen A und B von je 10.000 Personen. Bei der Gruppe A trifft das Merkmal Y in 50% der Fälle zu, bei der Gruppe B nur in 10% der Fälle. Wir haben eine Vorhersagemethode, die für beide Populationen die folgenden Fehlerraten hat: falsch-positiv-Rate = 10%, falsch-negativ-Rate = 0%.

a) Füllen sie die fehlenden Felder der folgenden Tabellen ein.

		Vorhersage		
		H = 0	H = 1	
Wahrheit	Y = 0	TN	FP	5000
	Y = 1	FN	TP	5000
		$\#(H = 0)$	$\#(H = 1)$	10000

Population A

		Vorhersage		
		H = 0	H = 1	
Wahrheit	Y = 0	TN	FP	9000
	Y = 1	FN	TP	1000
		$\#(H = 0)$	$\#(H = 1)$	10000

Population B

Lösung:

		Vorhersage		
		H = 0	H = 1	
Wahrheit	Y = 0	4500	500	5000
	Y = 1	0	5000	5000
		4500	5500	10000

Population A

		Vorhersage		
		H = 0	H = 1	
Wahrheit	Y = 0	8100	900	9000
	Y = 1	0	1000	1000
		8100	1900	10000

Population B

b) Was ist die positive Vorhersagequalität der Methode bei beiden Populationen?

$$\text{positive Vorhersagequalität} = \frac{TP}{\#(H = 1)}.$$

Lösung: Bei Population A ist die positive Vorhersagequalität gleich $5000/5500 = 10/11 \approx 0.91$. Bei Population B ist die positive Vorhersagequalität gleich $1000/1900 = 10/19 \approx 0.526$.

c) Argumentieren Sie:

- Das Verfahren hat eine vorherragende Qualität und diskriminiert keine der Populationen.
- Das Verfahren hat eine miserable Qualität und diskriminiert die Population B.

Lösung:

- Die Qualität des Verfahrens ist hoch, weil die Fehlerraten gering sind. Es diskriminiert nicht, weil die Fehlerraten für beide Populationen die gleiche sind.
- Die Qualität des Verfahrens ist miserabel, weil die positive Vorhersagequalität bei der Population B nur wenig besser ist als Raten. Es diskriminiert, weil die positive Vorhersagequalität bei der Population A gut ist, aber bei B miserabel.

Aufgabe 3 (5 Punkte) Betrachten Sie die Funktion $z = z(x, y) = x^2 + 4y^2$.

- a) Zeichnen Sie die Höhenlinien $z = 1$, $z = 2$ und $z = 4$ in ein Koordinatensystem ein. Auf welcher Höhenlinie liegen die Punkte $(x, y) = (2, 0)$, $(x, y) = (0, 1)$, und $(x, y) = (1, \sqrt{3/4})$?
- b) Die Ableitungen von z nach x und y sind $\partial z / \partial x = 2x$ und $\partial z / \partial y = 8y$. Der Gradient ∇z von z ist der Vektor bestehend aus den beiden Ableitungen. Daher $\nabla z = \begin{pmatrix} 2x \\ 8y \end{pmatrix}$. Zeichnen Sie die Gradienten an den drei Punkten des ersten Unterpunktes. Was ist der Zusammenhang zwischen Höhenlinien und Gradient?

Lösung: Die Höhenlinie $z = c$ ist eine Ellipse mit dem Halbachsen \sqrt{c} und $\sqrt{c/2}$. Der Gradient steht senkrecht auf der Höhenlinie.

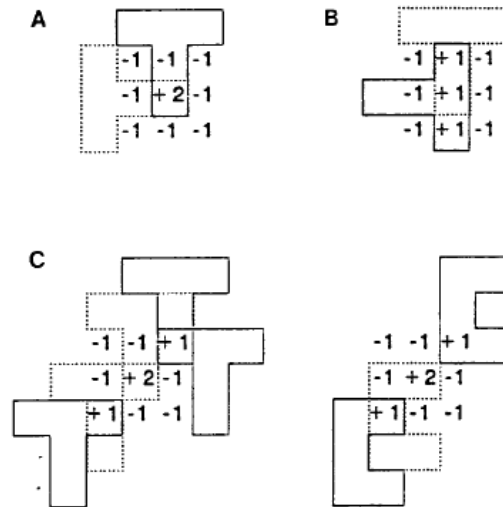
- c) Gradientenabstieg: Wir beginnen mit einem Punkt (x_0, y_0) und definieren dann eine Folge $(x_i, y_i), i \geq 1$, durch $(x_{i+1}, y_{i+1}) = (x_i, y_i) - h \nabla z(x_i, y_i) = (x_i - 2hx_i, y_i - 8hy_i)$. Dabei ist h die Schrittweite. Starten sie mit $(x_0, y_0) = (2, 3)$ und bestimmen sie die ersten vier Schritte bei Verwendung der Schrittweite $h = 1/8$. Das Minimum ist der Punkt $(0, 0)$. Wie nahe kommen sie ihm in 10 Schritten?

Lösung: We have $x_{i+1} = (1 - 2h)x_i = 3/4x_i$. Also $x_0 = 2, x_1 = 3/2, x_2 = 9/8, x_3 = 27/32$. Allgemein $x_i = 2 \cdot (3/4)^i$. Also $x_{10} = 2 \cdot (3/4)^{10}$. Für y erhalten wir: $y_1 = (1 - 8h)y_0 = 0$ und dann $y_3 = y_2 = y_1 = 0$.

d) Was passiert, wenn sie die Schrittweite $h = 1$ wählen?

Lösung: $x_{i+1} = (1 - 2h)x_i = -x_i$ und $y_{i+1} = (1 - 8h)y_i = -7y_i$. Also alterniert der x -Wert zwischen $+2$ und -2 . Der y -Wert explodiert.

Aufgabe 4 (10 Punkte) In der Vorlesung haben wir das Netz gesehen, das C und T unterscheiden kann. Ich habe in der Vorlesung erklärt, wie die Filter A und D funktionieren. Erklären Sie, wie die Filter B und C funktionieren.



- Welche Werte können die Filter B und C liefern bei Eingabe C bzw. T.
- Was muss das Ausgabeneuron leisten?

Lösung:

Filter B: Beim T liefert mindestens ein Neuron der Eingabeschicht einen Wert ≥ 2 . Bei Eingabe C ist der Wert immer ≤ 1 .

Wenn das T normal oder auf dem Kopf steht, die mittlere Spalte mit 2 Kästchen überlappt, und der Balken außerhalb des Filters liegt, bekommt man den Wert 2. Wenn das T liegt und der Balken des T mit der mittleren Spalte übereinstimmt, bekommt man den Wert 2.

Wenn das C die mittlere Spalte nicht oder nur in einem Quadrat überlappt, dann ist der Gesamtwert sicher ≤ 1 . Wenn das C die mittlere Spalte in genau 2 Quadranten überlappt, dann muss es auch eines der Felder mit Wert -1 überlappen. Also ist der Gesamtwert ≤ 1 . Wenn das C die mittlere Spalte in 3 Quadranten überlappt, dann steht es aufrecht und überlappt auch zwei Felder mit Wert -1 . Also ist der Gesamtwert ≤ 1 .

Das Ausgabeneuron sagt T, wenn mindestens ein Neuron der ersten Schicht den Wert 2 liefert.

Filter C: Beim C liefert mindestens ein Filter den Wert -3 . Beim T sind die Werte immer ≥ -2 .

Wenn das C normal steht und mit dem linken Rand des Filters aligniert ist, dann ist der Wert -3 . Analog für die drei anderen Lagen des C.

Nehmen wir an, das T steht normal. Wenn es drei Kästchen -1 überlappt, dann muss es sowohl der Stamm als auch der Balken des T den Filter überlappen. Also liegt der Stamm entweder in der linken Spalte des Filters (dann Gesamtwert -2) oder in der mittleren Spalte (dann Gesamtwert 0) oder in der rechten Spalte (Gesamtwert -2). Analog argumentiert bei den anderen Lagen des T.

Das Ausgabeneuron muss also nur entscheiden, ob es ein Eingabeneuron gibt mit Wert -3 gibt.

Maschinelles Lernen war spannend okay langweilig
 schwierig okay einfach