TECHNISCHE UNIVERSITEIT EINDHOVEN

Faculteit Biomedische Technologie en Faculteit Wiskunde en Informatica

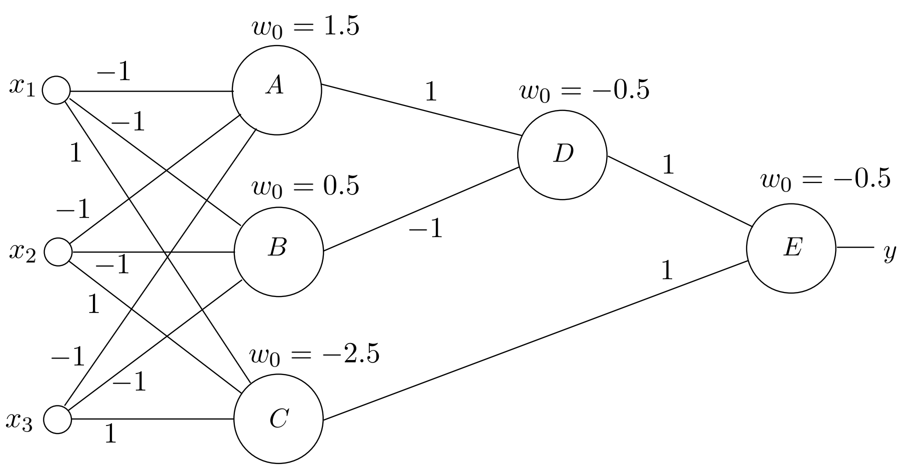
Tentamen Neurale Netwerken en Data Analyse (8C050) / Neurale Netwerken (2L490) op 25 maart 2003, 9.00-12.00 uur.

Alle antwoorden dienen duidelijk geformuleerd en gemotiveerd te worden.
Dit tentamen bestaat uit 4 bladzijden.

1) Beschouw een discreet één-laags perceptron met n inputs met gewichten $w_1, \dots, w_n$ en drempel $-w_0$.

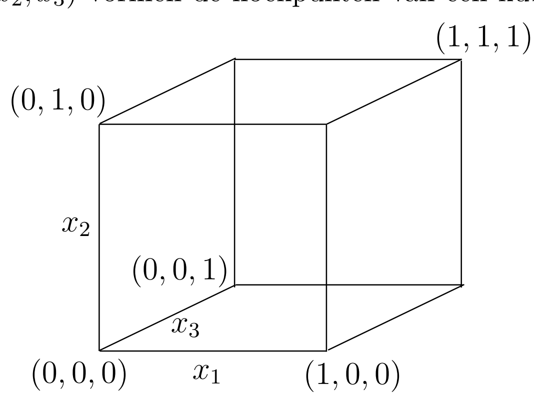
a) Geef de uitvoer y van zo'n perceptron bij gegeven input $\mathbf{x} \in \mathbb{R}^n$.

b) Beschouw het volgende netwerk, opgebouwd uit discrete neuronen. Ten overvloede, alle verbindingen van x_1, x_2 en x_3 naar de neuronen A en B hebben gewicht -1, en alle verbindingen van x_1, x_2 en x_3 naar neuron C hebben gewicht 1.



We gebruiken dit netwerk om de logische functie $y = f(x_1, x_2, x_3)$ te berekenen. We beperken ons dus tot x_1, x_2, x_3, y in de verzameling $\{0, 1\}$. Geef in tabelvorm de uitvoer van de neuronen A, B, C, D en E voor alle 8 mogelijke waarden van (x_1, x_2, x_3).

c) De mogelijke punten (x_1, x_2, x_3) vormen de hoekpunten van een kubus, zie onderstaande figuur.

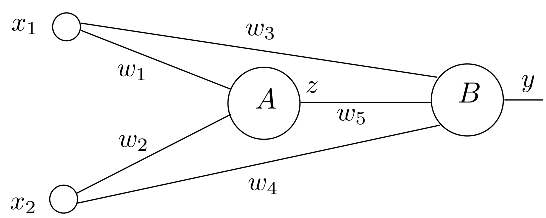


Teken voor ieder van de neuronen A, B, C, D en E in zo'n kubus de hoekpunten waar dit neuron uitvoer 1 heeft.

d) Geef een gelaagd netwerk met een minimaal aantal lagen dat dezelfde functie $f(x_1, x_2, x_3)$ berekent. Beargumenteer het door u gebruikte aantal lagen.

*****

2) Beschouw het volgende netwerk met invoer x_1 en x_2 en met één uitvoer y.



Invoer zowel als uitvoer zijn reëelwaardig. De neuronen in dit netwerk hebben geen drempel. Verder zijn het standaard neuronen met activatiefunctie f (met $f(z) = 1/(1 + \exp(-z)))$.

a) Geef de uitvoer y van dit netwerk als functie van de invoer x_1, x_2.

b) Geef de definitie van de kwadratische fout E bij een leervoorbeeld (x_1, x_2, t).

c) Leid de standaard gradient descent leerregels af voor de gewichten w_1, w_2, w_3, w_4 en w_5.

d) Stel dat we neuron A vervangen door een lineair neuron, zonder drempel. Laat zien dat het zo verkregen netwerk dezelfde klasse functies kan berekenen als een eenlaags netwerk bestaande uit één neuron met twee inputs.

*****

3) Beschouw een meerlaags feedforward netwerk, opgebouwd uit standaard neuronen, met een neuron in de uitvoerlaag. Laat E de kwadratische fout zijn bij een leervoorbeeld $(\mathbf{x}, t)$.

a) Geef de formule voor de gradient descent leerregel met momentum term.

b) Beschouw de gewichtscorrecties behorende bij de *gewone gradient descent leerregel* voor een gewicht w van het netwerk. Deze correcties Δw zijn op verschillende tijdstippen gegeven door:

t	Δw
0	1
1	$\frac{1}{2}$
2	$\frac{1}{4}$
3	$\frac{1}{8}$
4	$\frac{1}{8}$

Geef voor $t > 0$ de gewichtscorrecties behorende bij de gradient descent leerregel met momentum term, indien de in deze leerregel voorkomende extra parameter (β) gelijk is aan $\frac{1}{2}$.

c) Herhaal onderdeel b) voor gewichtscorrecties behorende bij de gewone gradient descent leerregel gegeven in de volgende tabel:

t	Δw
0	1
1	-1
2	1
3	-1
4	1

d) Wat is het beoogde doel van het toevoegen van een momentum term in de leerregel?

*****

4) Beschouw een algemeen Hopfield netwerk bestaande uit n neuronen, met onderlinge gewichten w_{ij} en drempels b_i.

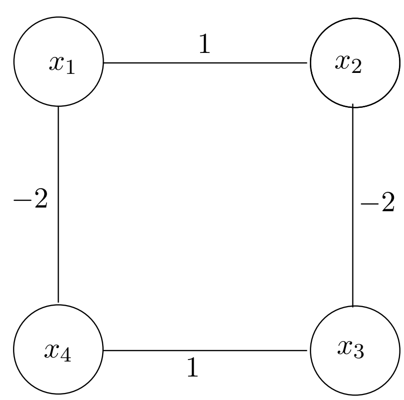
a) Geef voor een willekeurig Hopfield netwerk de asynchrone updateregel.

b) Geef de definitie van de consensus.

c) Geef de definitie van stabiele toestand.

d) Stel dat alle drempels gelijk aan 0 zijn, d.w.z. $b_i = 0$ voor $i = 1, \ldots, n$. Bewijs dat wanneer $\mathbf{x}$ een stabiele toestand is, dan ook $-\mathbf{x}$.

Beschouw nu het onderstaand Hopfield netwerk.



De gewichten staan bij de verbindingen tussen de neuronen. Niet getekende connecties hebben een gewicht 0. Alle neuronen hebben een drempel $b_i = 0$ voor $i = 1, 2, 3, 4$.

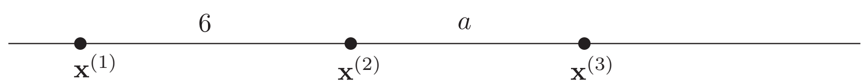
e) Teken de buurruimtegraaf en geef daarin van alle toestanden hun consensus. (D.w.z. teken een "rondje" voor iedere mogelijke toestand, en verbind de rondjes van toestanden die door één asynchrone stap in elkaar kunnen overgaan.) Geef aan welke toestanden locale maxima van de consensus zijn.

f) Stel dat bovenstaand Hopfield netwerk evolueert volgens de asynchrone update regel. Geef een mogelijk pad in de buurruimtegraaf vanuit de toestand $(-1, -1, -1, -1)$ naar een stabiele toestand.

*****

5a) Beschouw een aantal punten $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}$ in de n-dimensionale ruimte. Laat C een clustering van deze punten in k clusters beschrijven, d.w.z. een punt $\mathbf{x}^{(q)}$ zit dan in cluster $C(q)$. Geef de definitie van de bijbehorende fout E.

Beschouw nu drie punten $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}$ op een rechte lijn, punt $\mathbf{x}^{(2)}$ heeft afstand 6 tot $\mathbf{x}^{(1)}$, en punt $\mathbf{x}^{(3)}$ heeft afstand a tot $\mathbf{x}^{(2)}$.



We beschouwen twee mogelijke clusteringen van deze punten in twee clusters, namelijk de clustering C met $C(1) = C(2) = 1$ en $C(3) = 2$ en de clustering $\widetilde{C}$ met $\widetilde{C}(1) = 1$ en $\widetilde{C}(2) = \widetilde{C}(3) = 2$.

b) Bereken de fout E voor de clustering C en voor de clustering $\widetilde{C}$. Voor welke waarden van a is de clustering C beter dan de clustering $\widetilde{C}$?

c) Stel dat we het K-means algoritme uitvoeren met de clustering C als initiële clustering. Voor welke waarden van a wordt dan de stap naar de clustering $\widetilde{C}$ gemaakt?

d) Stel dat we het K-means algoritme uitvoeren met de clustering $\widetilde{C}$ als initiële clustering. Voor welke waarden van a wordt dan de stap naar de clustering C gemaakt?

e) Bestaan er waarden van a zodat zowel C als $\widetilde{C}$ niet veranderen als zij als initiële clustering bij het K-means algoritme worden gebruikt? Zo ja, geef die waarden, zo nee bewijs dat die waarden niet bestaan.

Waardering: (totaal 50 punten)

Opgave 1a:	2 punten	Opgave 2a:	1 punt	Opgave 3a:	2 punten
Opgave 1b:	2 punten	Opgave 2b:	1 punt	Opgave 3b:	3 punten
Opgave 1c:	2 punten	Opgave 2c:	5 punten	Opgave 3c:	3 punten
Opgave 1d:	4 punten	Opgave 2d:	3 punten	Opgave 3d:	2 punten
Opgave 4a:	1 punt	Opgave 5a:	1 punt		
Opgave 4b:	1 punt	Opgave 5b:	2 punten		
Opgave 4c:	1 punt	Opgave 5c:	3 punten		
Opgave 4d:	2 punten	Opgave 5d:	3 punten		
Opgave 4e:	3 punten	Opgave 5e:	1 punt		
Opgave 4f:	2 punten				