

## Huiswerk Wiskunde voor Neurale Netwerken

Inleveren 24 september om 11:00

1. Gegeven is een neurale netwerk met 2 inputs en 1 output. Deze output wordt gegeven door een threshold neuron. Dit houdt in dat de output van het neuron als volgt wordt bepaald:

$$y = f(a) = \begin{cases} 1, & \text{als } a \geq \theta \\ 0, & \text{anders} \end{cases}$$

Hier is  $a$  de totale input voor het neuron. Deze input wordt als volgt opgebouwd:

$$a = \sum_{i=1}^n w_i x_i = w_1 x_1 + w_2 x_2$$

Hier is  $x_1$  de eerste input en  $x_2$  de tweede input.  $w_i$  is het gewicht op de verbinding van input  $i$  met het threshold neuron ( $i$  is 1 of 2). Het netwerk zal dus op een bepaalde input een 1 als output hebben als  $w_1 x_1 + w_2 x_2$  groter (of gelijk) is dan  $\theta$ .

- (a) Geef  $\theta, w_1, w_2$  zodanig dat het netwerk precies de functie  $(x_1 \vee x_2)$  modelleert.
  - (b) Gegeven is nu  $\theta = 0.5$ . Kun je  $w_1, w_2$  geven zodanig dat het netwerk  $(x_1 \wedge x_2)$  modelleert? Beantwoord ook dezelfde vraag voor  $\theta = -0.5$ .
  - (c) Geef een beschrijving van een netwerk dat  $n$ -ary AND modelleert:  $(x_1 \wedge x_2 \wedge \dots \wedge x_n)$ . Dit wil zeggen: een netwerk met  $n$  inputs, waar de  $y = 1$  is als alle inputs 1 zijn, en anders  $y = 0$ . (Het is niet de bedoeling een  $n$  te kiezen voor het antwoord op deze vraag, hoewel het voor inzicht wel handig kan zijn de vraag eerst voor bijvoorbeeld  $n = 3$  op te lossen.)
2. Gegeven zijn de vectoren  $v_1, v_2$ :

$$\vec{v}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

- (a) Verifieer dat de lengte van  $\vec{v}_1$  gelijk is aan 1.
- (b) Bereken het inproduct  $\langle \vec{v}_1, \vec{v}_2 \rangle$ .
- (c) Geef een vector (deze noemen we  $\vec{v}_3$ ) met dezelfde lengte als  $\vec{v}_2$  waarvoor geldt dat  $\langle \vec{v}_2, \vec{v}_3 \rangle$  maximaal is. (Dat wil zeggen, er is geen andere vector te vinden waarvoor het inproduct met  $\vec{v}_2$  hoger kan uitvallen).
- (d) Geef nu een vector ( $\vec{v}_4$ ) waarvoor  $\langle \vec{v}_2, \vec{v}_4 \rangle$  minimaal is.
- (e) Meetkundig gezien: wanneer is het inproduct tussen twee vectoren met lengte 1 het grootst? En wanneer is deze het kleinst? (Hint: teken twee zulke vectoren met twee elementen in het platte vlak).

3. Gegeven zijn de matrix  $A$  en de vectoren  $x$  en  $y$ :

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 4 & 0 & 0 \end{pmatrix}, \vec{x} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \vec{y} = \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$$

- (a) Bepaal  $A^T$ .
  - (b) Bereken  $A\vec{y}$
  - (c) Bereken  $\langle \vec{x}, A\vec{y} \rangle$ .
  - (d) Laat zien voor alle vectoren  $\vec{u}, \vec{v}$ :  $\langle \vec{u}, A\vec{v} \rangle = \langle A^T \vec{u}, \vec{v} \rangle$ . (Hint: je kan vectoren als matrices zien, waarbij het aantal kolommen precies 1 is. Er geldt dan:  $\langle \vec{a}, \vec{b} \rangle = \vec{a}^T \vec{b}$ ).
  - (e) Bepaal  $A^{-1}$ .
  - (f) Bereken  $AA^{-1}$  en geef  $(AA^{-1})^{-1}$  en  $(AA^{-1})^T$
4. Een neurale netwerk met een enkel neuron (zoals in opgave 1) kan gezien worden als een inproduct met daarover een functie:

$$y = f(\langle \vec{w}, \vec{x} \rangle) = \begin{cases} 1, & \text{als } \langle \vec{w}, \vec{x} \rangle \geq \theta \\ 0, & \text{anders} \end{cases}$$

Het is natuurlijk ook mogelijk een netwerk te maken met meerdere neuronen. De output kan dan ook weer een vector van waarden zijn. Het meest gebruikelijk is dan om van iedere input een verbinding te

laten lopen naar iedere output. Dit betekent wel dat er meer gewichten zijn, om precies te zijn  $nm$ , waar  $n$  het aantal inputs is en  $m$  het aantal outputs. Deze gewichten kun je dan in een  $n \times m$  matrix zetten.

Gegeven is een netwerk met twee inputs en twee outputs. Laat  $W$  de  $(2 \times 2)$  gewichten matrix zijn,  $\vec{x}$  de input en  $\vec{y}$  de output. Er geldt:  $g(W\vec{x}) = \vec{y}$ , als we  $g$  analoog aan  $f$  uit opgave 1 definiëren, maar dan voor vectoren:

$$g(\vec{a}) = \begin{pmatrix} f(a_1) \\ f(a_2) \\ \vdots \\ f(a_n) \end{pmatrix}$$

Hier is  $f$  gewoon de threshold functie uit opgave 1.

- (a) Verifieer dat het eerste element van de output overeenkomt met de output van een netwerk met maar 1 output, waar de gewichten bestaan uit de eerste rij uit de matrix  $W$ . (In feite heb je dus gewoon evenveel afzonderlijke netwerken als outputs, alleen delen de netwerken dezelfde inputs.)
- (b) Als je meerdere outputs hebt, kunnen deze natuurlijk ook weer gebruikt worden als input voor een volgende laag outputs. De originele outputs worden dan *hidden* genoemd (omdat je van buitenaf alleen de input en de uiteindelijke output kan waarnemen). Gegeven is een netwerk met 2 inputs, 2 hidden neuronen en 1 output. Geef gewichten (zowel tussen alle inputs en alle hidden neuronen als tussen alle hidden neuronen en de output) en thresholds (van zowel de hidden neuronen als de output) zodanig dat het netwerk de XOR functie modelleert:

$$output = \begin{cases} 1, & \text{als } x_1 \neq x_2 \\ 0, & \text{anders} \end{cases}$$

(Hint: maak gebruik van de netwerken voor  $\wedge$  en/of  $\vee$ . Bedenk dat gewichten ook negatief kunnen zijn.)

- (c) Is het noodzakelijk om thresholds voor de hidden neurons en de output te hebben? Of kun je eventueel ook gewoon waardes doorgeven zonder er een functie overheen te gooien?