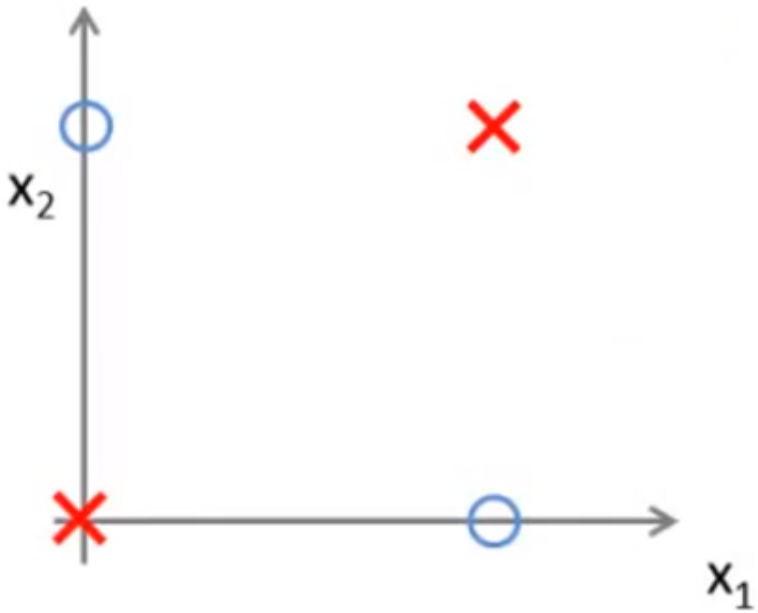
Examples and Intuitions 1

Applications

Neural Networks: Representation

Non-linear classification example: XOR/XNOR

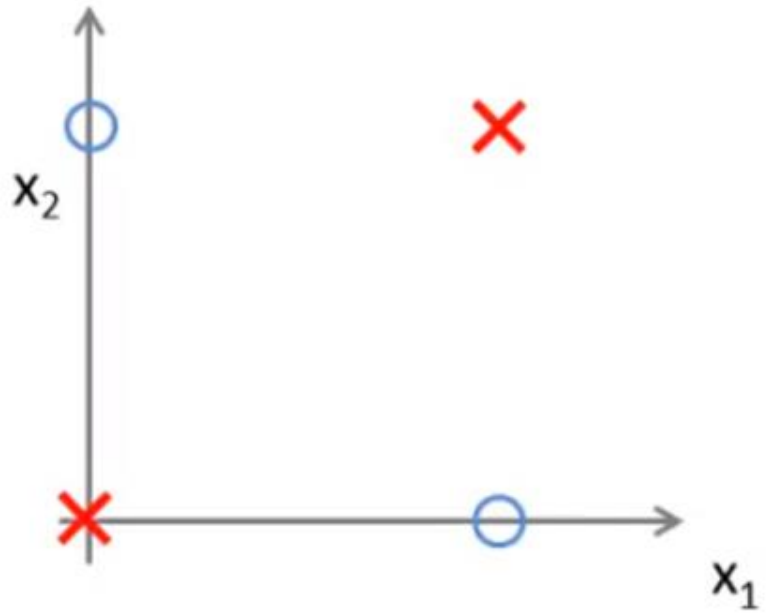
x_1, x_2 are binary (0 or 1).

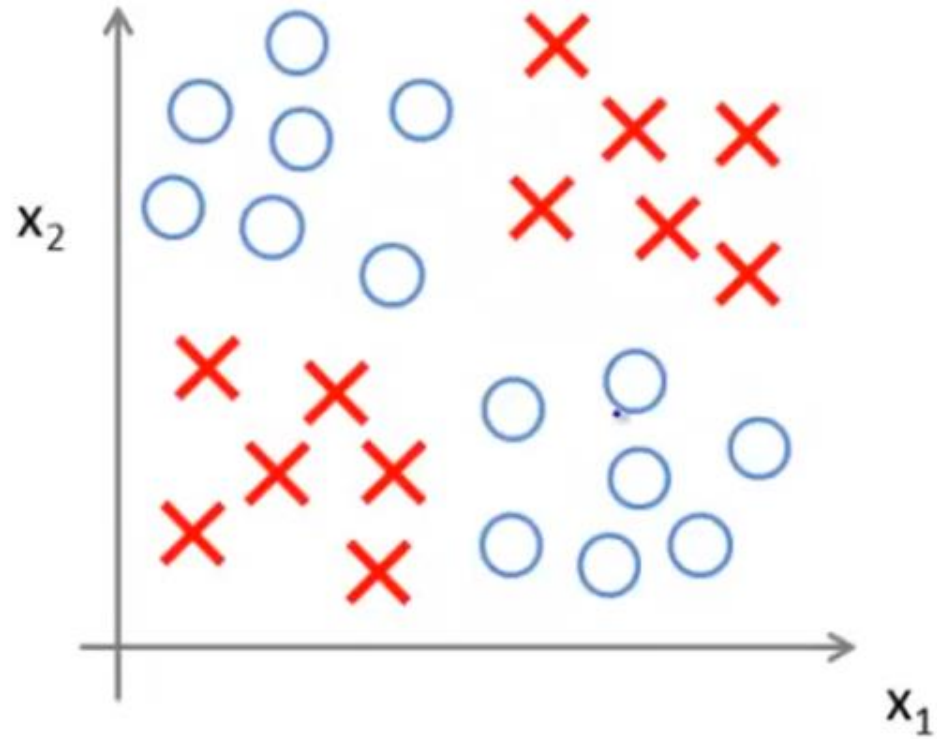


A simpler version of…

Non-linear classification example: XOR/XNOR

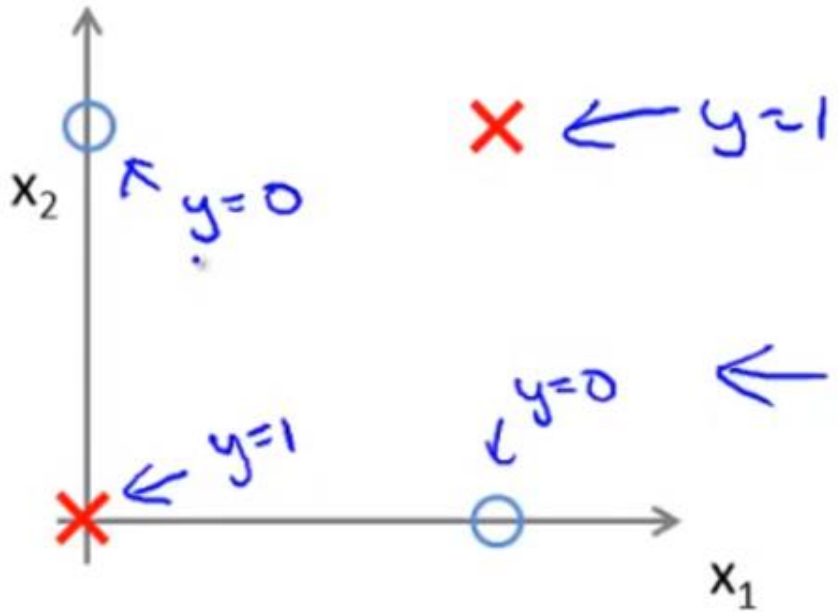
x_1, x_2 are binary (0 or 1).





Non-linear classification example: XOR/XNOR

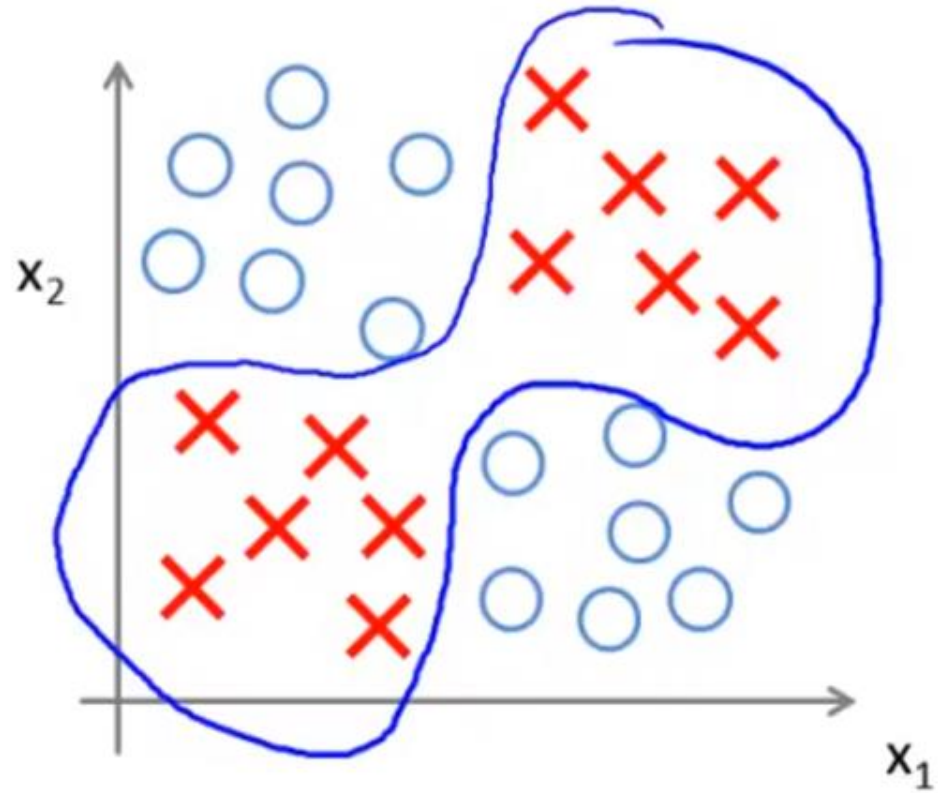
x_1, x_2 are binary (0 or 1).



$$y = x_1 \text{ XOR } x_2$$

$$x_1 \text{ XNOR } x_2$$

$$\text{NOT } (x_1 \text{ XOR } x_2)$$



Exclusive OR: utputs true only when inputs differ

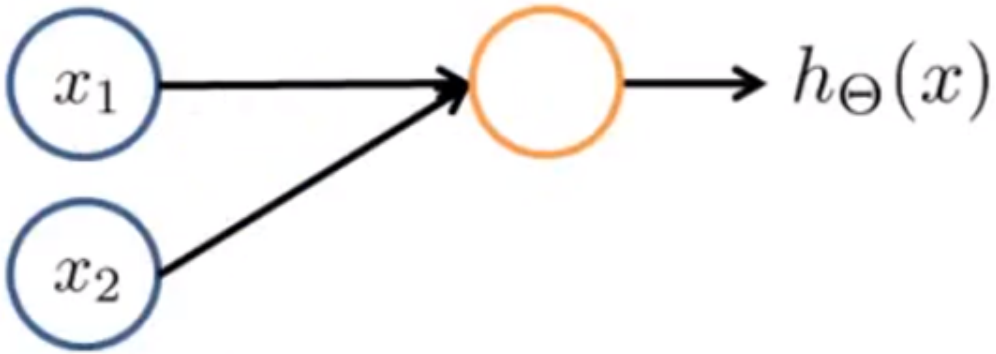
Building Networks

- We will try to find a network that fits the XOR example.
- But before, bu we're going to start with a slightly simpler one
- We will show a network that fits the **AND** function.

Simple example: AND

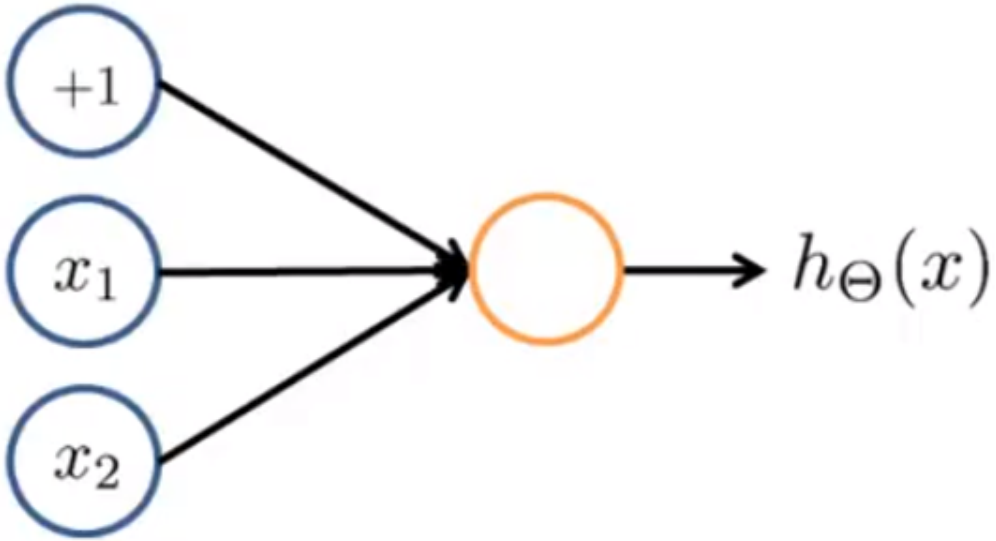
$x_1, x_2 \in \{0, 1\}$

$y = x_1 \text{ AND } x_2$



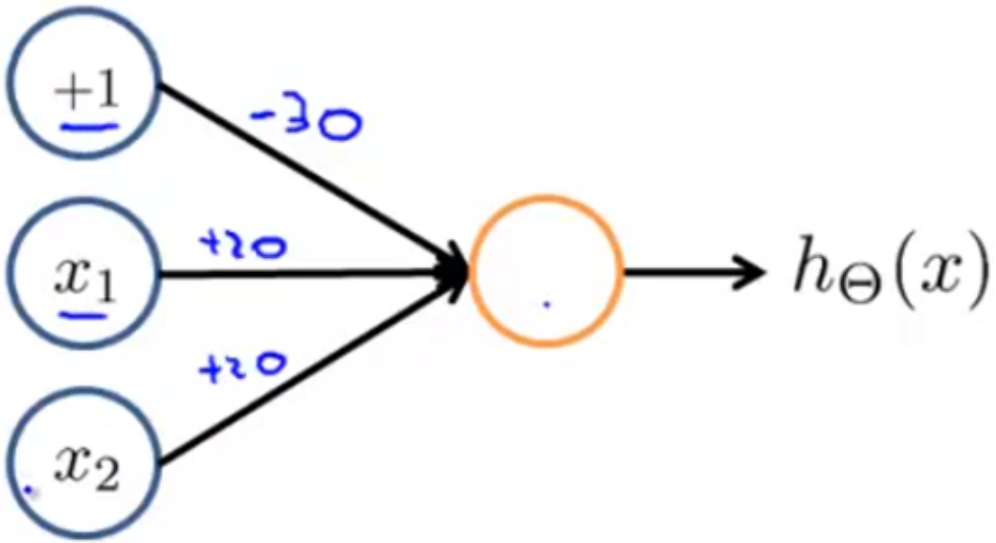
Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



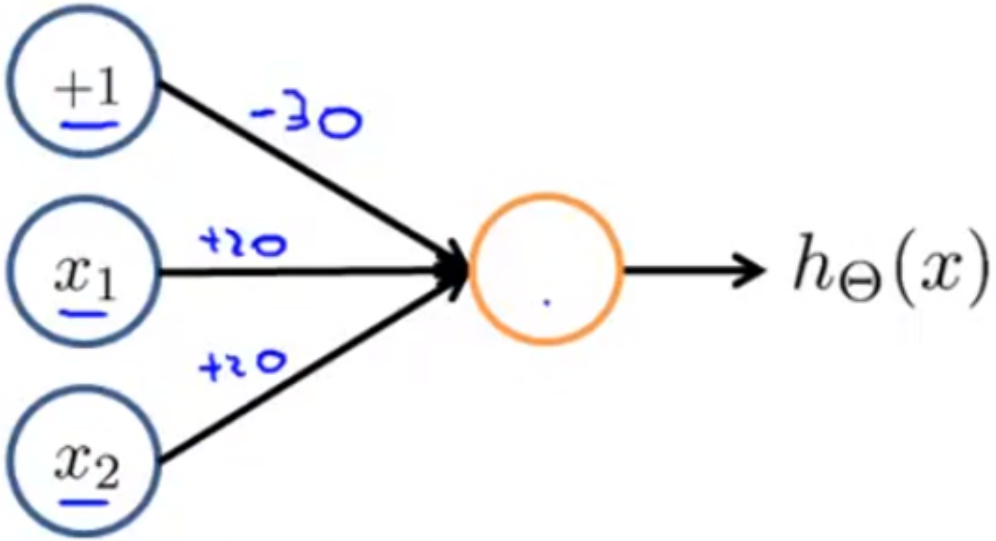
Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



Simple example: AND

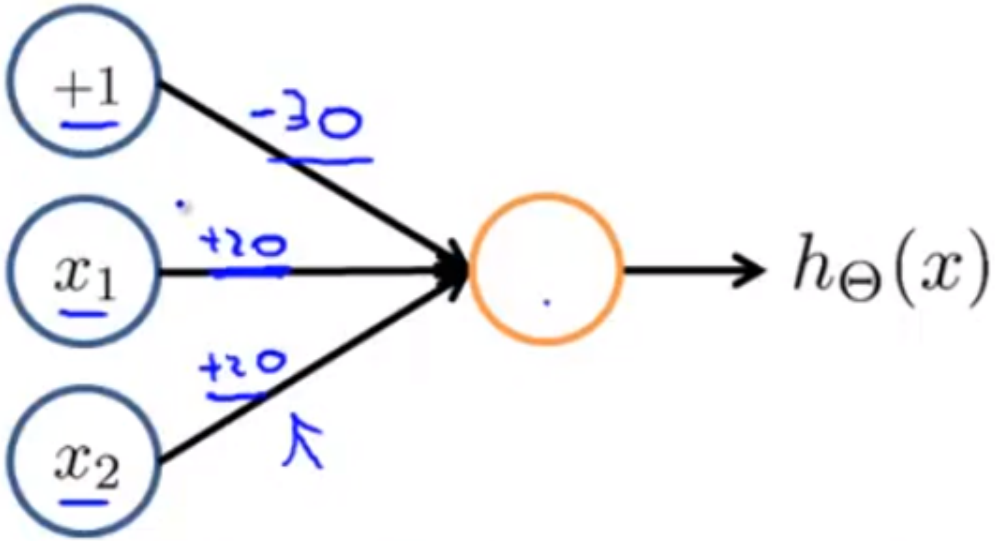
- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$

Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



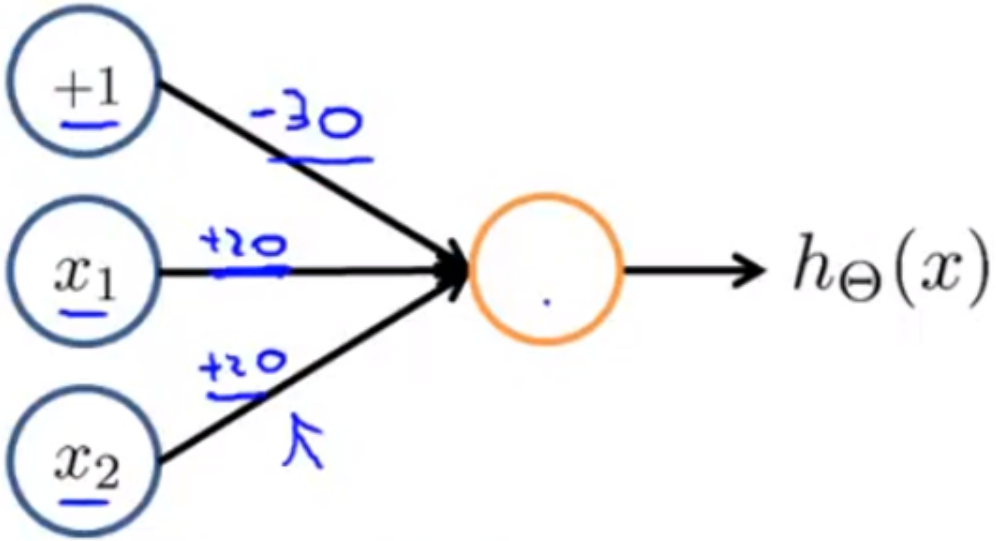
$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$

$\Theta_{10}^{(1)}$ $\Theta_{11}^{(1)}$ $\Theta_{12}^{(1)}$

4:00 / 7:15

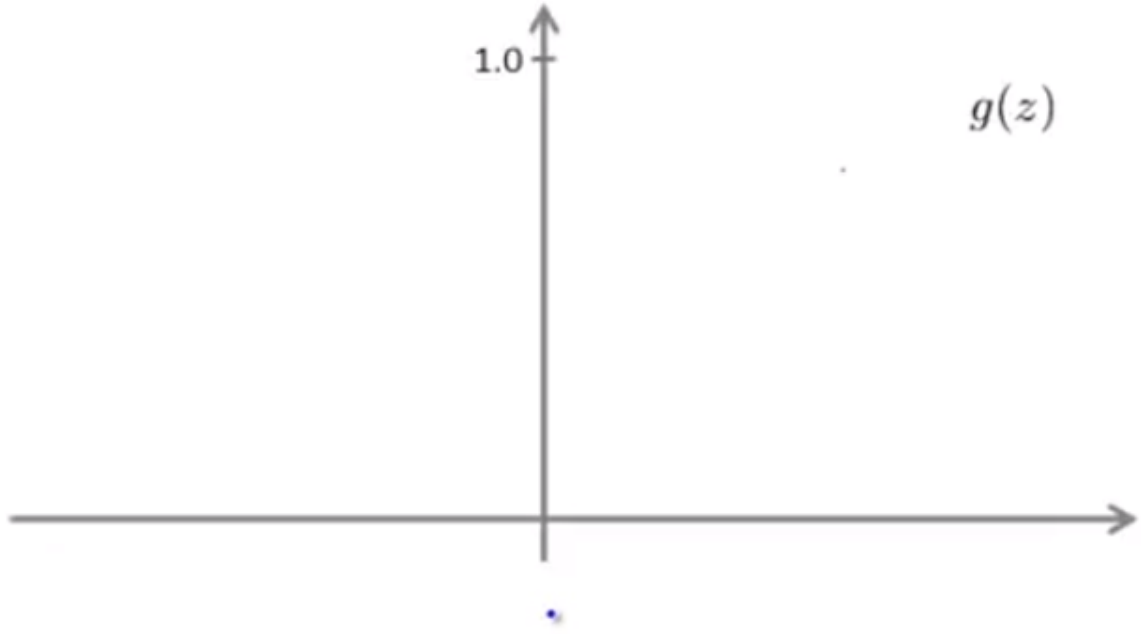
Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$

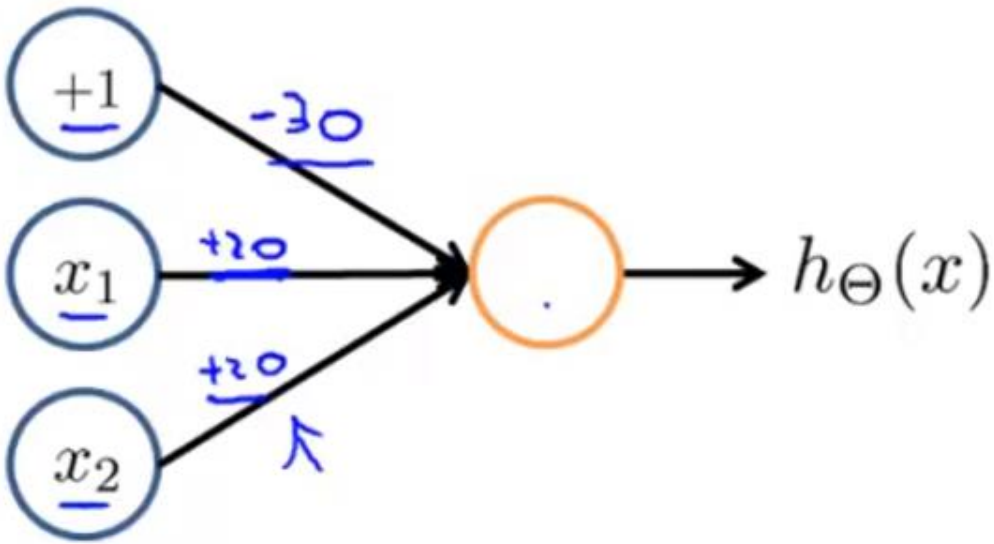
$\Theta^{(1)}_{10}$ $\Theta^{(1)}_{11}$ $\Theta^{(1)}_{12}$



x_1	x_2	$h_\Theta(x)$
0	0	
0	1	
1	0	
1	1	

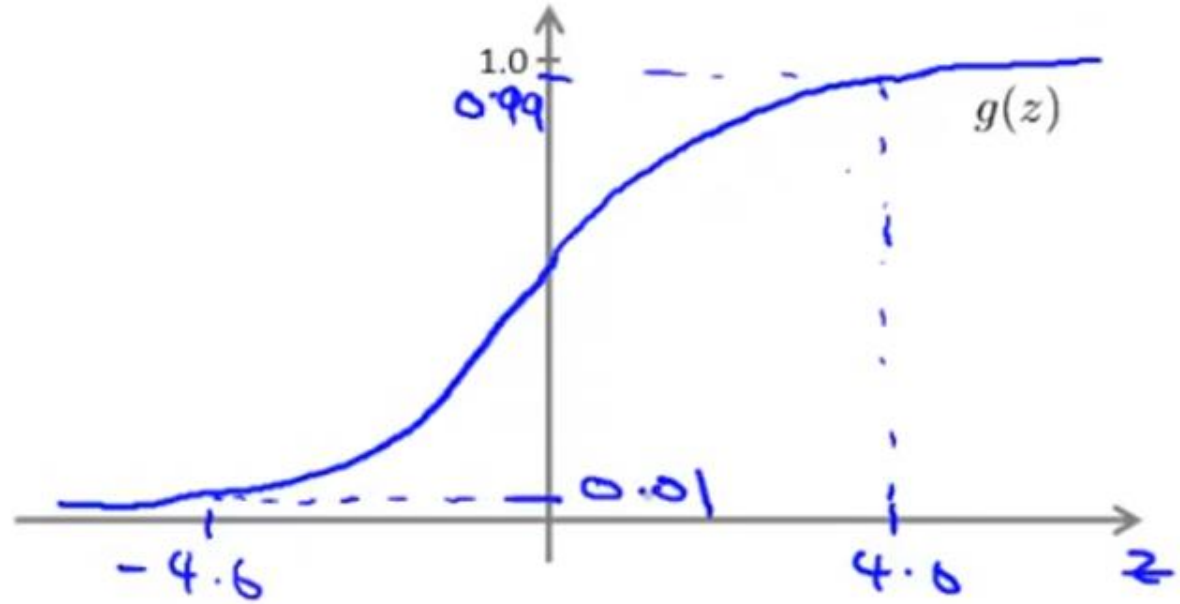
Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1$ AND x_2



$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$

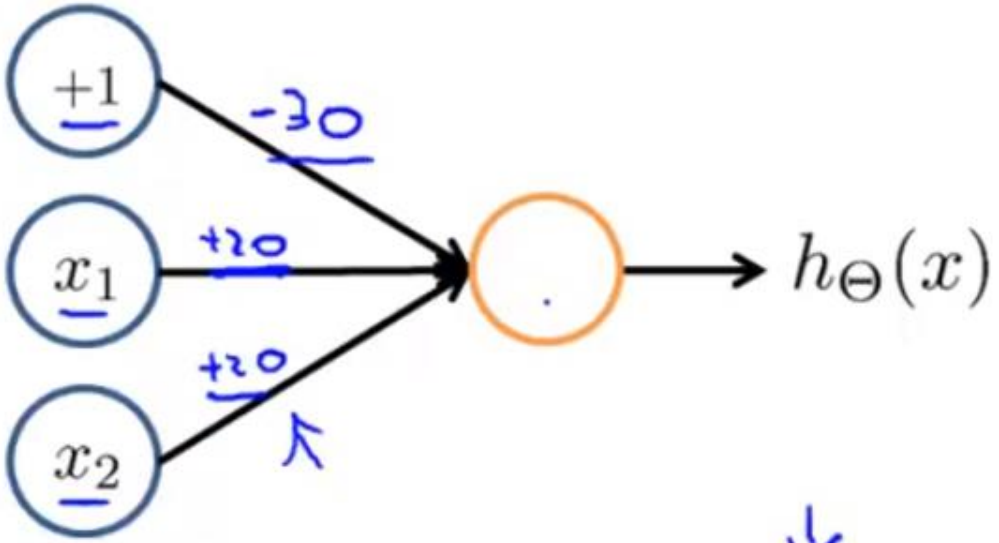
$\Theta^{(1)}_{10}$ $\Theta^{(1)}_{11}$ $\Theta^{(1)}_{12}$



x_1	x_2	$h_\Theta(x)$
0	0	
0	1	
1	0	
1	1	

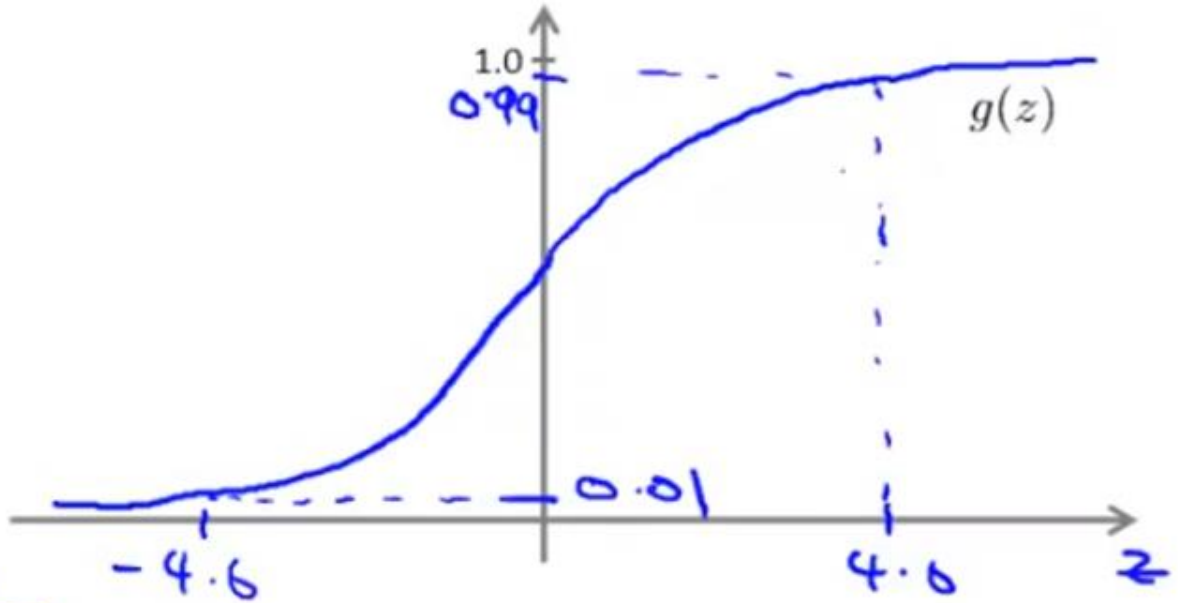
Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{ AND } x_2$



$$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$$

$\Theta^{(1)}_{10}$ $\Theta^{(1)}_{11}$ $\Theta^{(1)}_{12}$

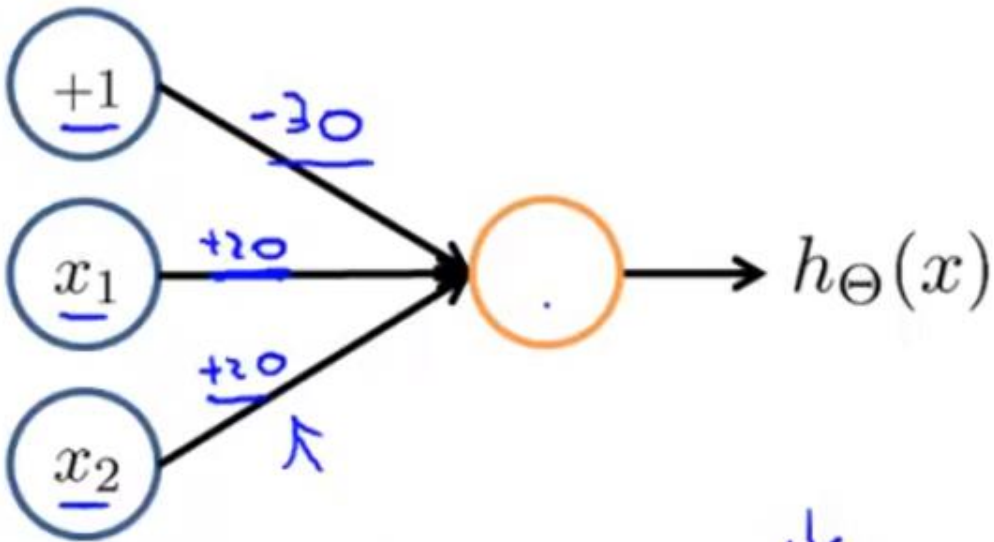


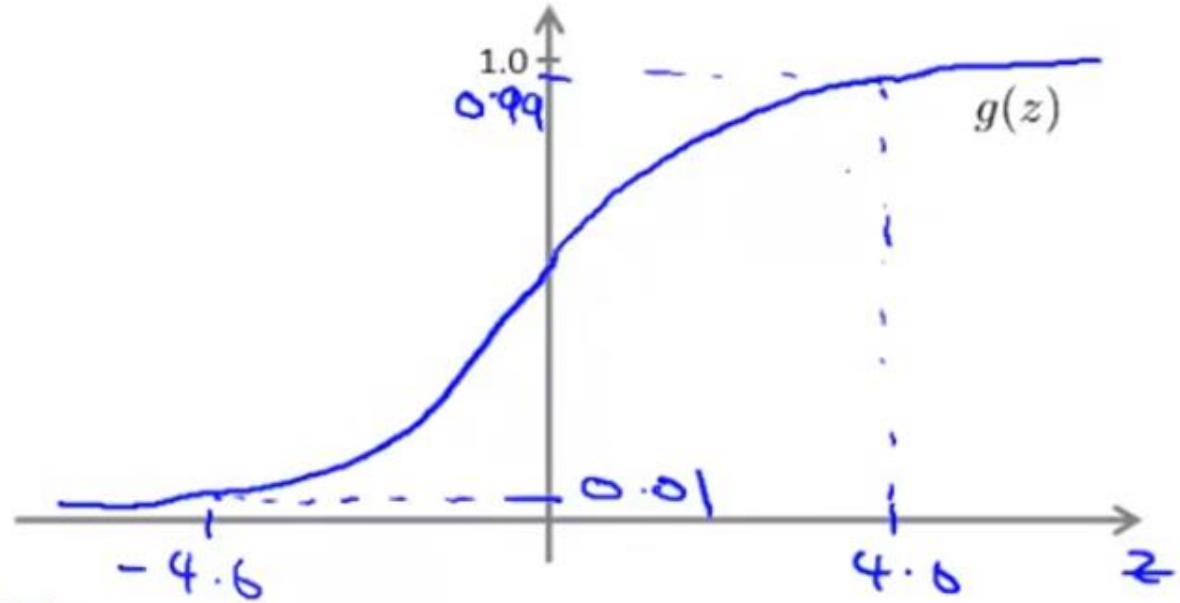
x_1	x_2	$h_\Theta(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

Andrew Ng

Simple example: AND

- $x_1, x_2 \in \{0, 1\}$
- $y = x_1$ AND x_2





$$h_\Theta(x) = g(-30 + 20x_1 + 20x_2)$$

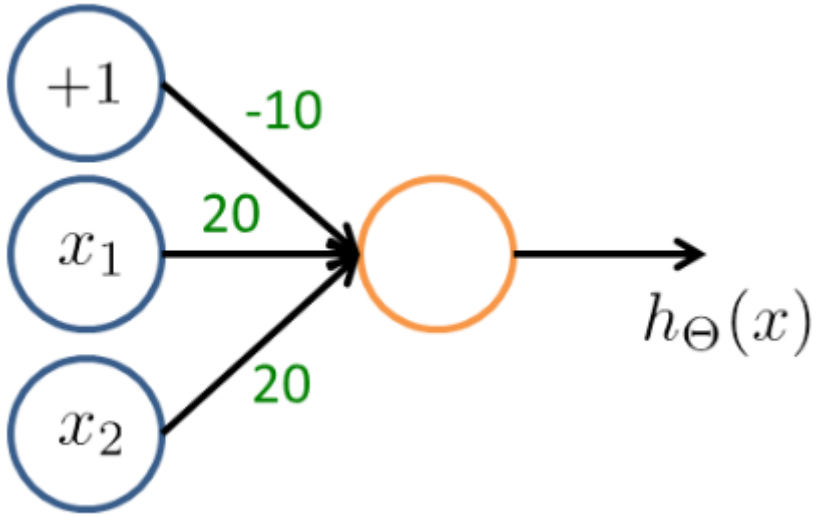
$\Theta_{10}^{(1)}$ $\Theta_{11}^{(1)}$ $\Theta_{12}^{(1)}$

x_1	x_2	$h_\Theta(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

$h_\Theta(x) \approx x_1$ AND x_2

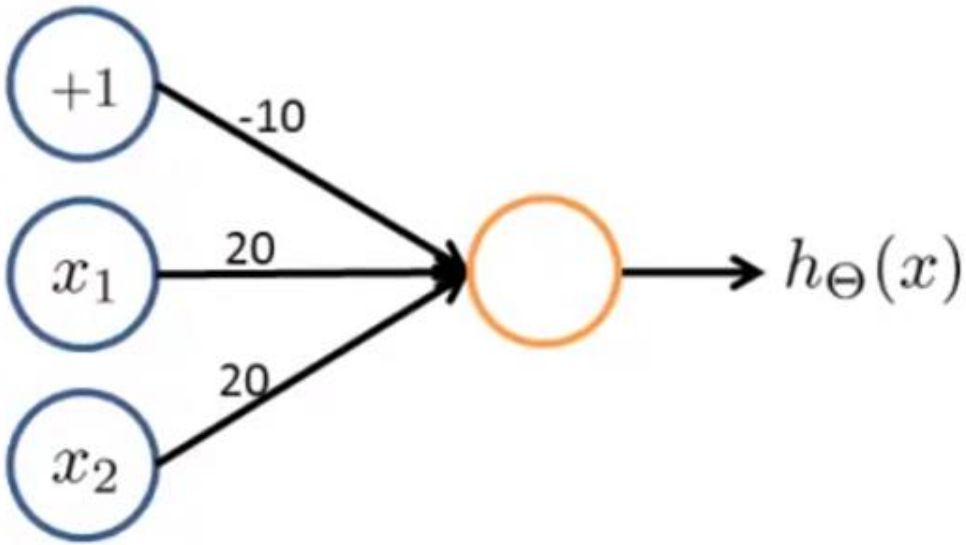
Exercise

- Suppose *x*1 and *x*2 are binary valued (0 or 1). What boolean function does the network shown below (approximately) compute?
- (Hint: One possible way to answer this is to draw out a truth table, similar to what we did in the previous slide).



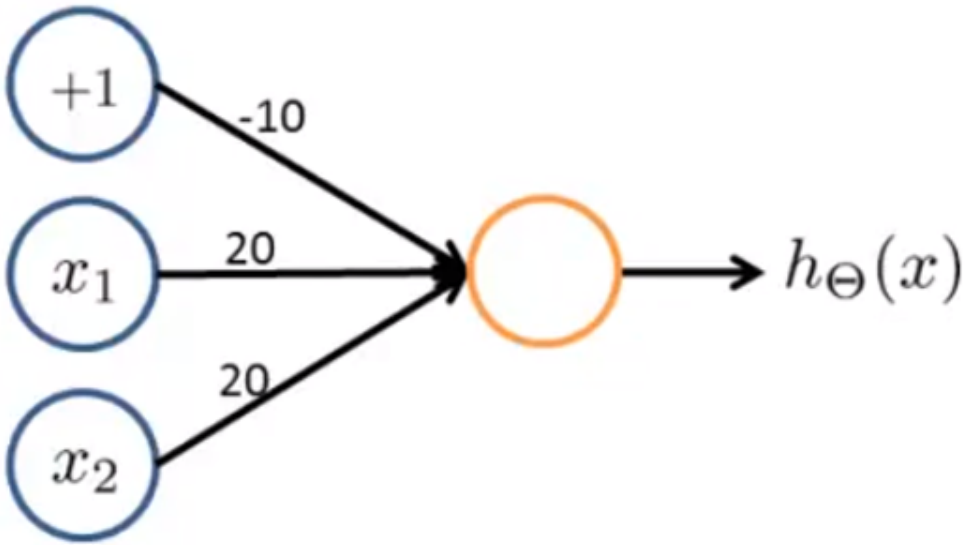
- x_1 AND x_2
- (NOT x_1) OR (NOT x_2)
- x_1 OR x_2
- (NOT x_1) AND (NOT x_2)

Example: OR function



x_1	x_2	$h_\Theta(x)$
0	0	
0	1	
1	0	
1	1	

Example: OR function



$g(-10 + 20x_1 + 20x_2)$

x_1	x_2	$h_\Theta(x)$
0	0	$g(-10) \approx 0$
0	1	$g(10) \approx 1$
1	0	≈ 1
1	1	≈ 1

Examples and Intuitions 2

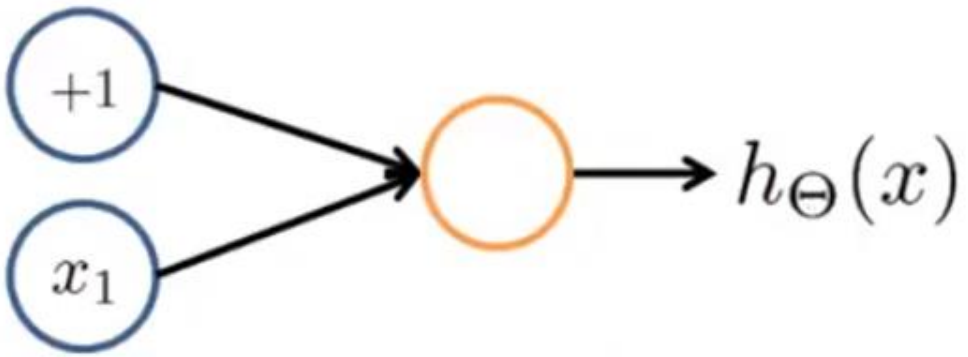
Applications

Neural Networks: Representation

→ x_1 AND x_2

→ x_1 OR x_2

Negation:



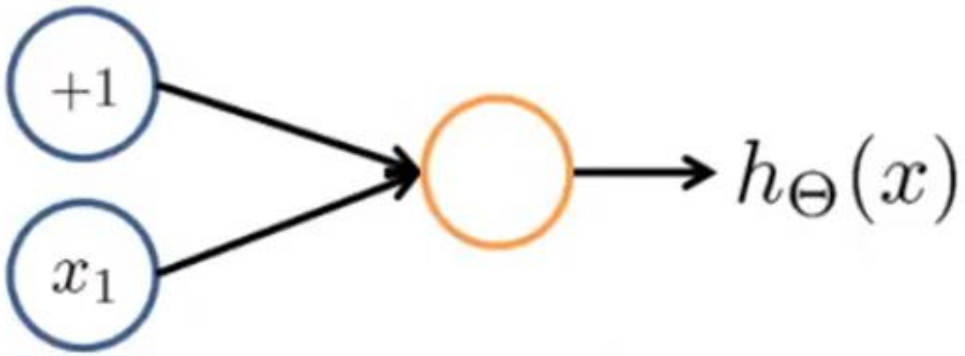
x_1	$h_\Theta(x)$
0	
1	

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

→ x_1 AND x_2 → x_1 OR x_2 {0,1}.

Negation:

NOT x_1

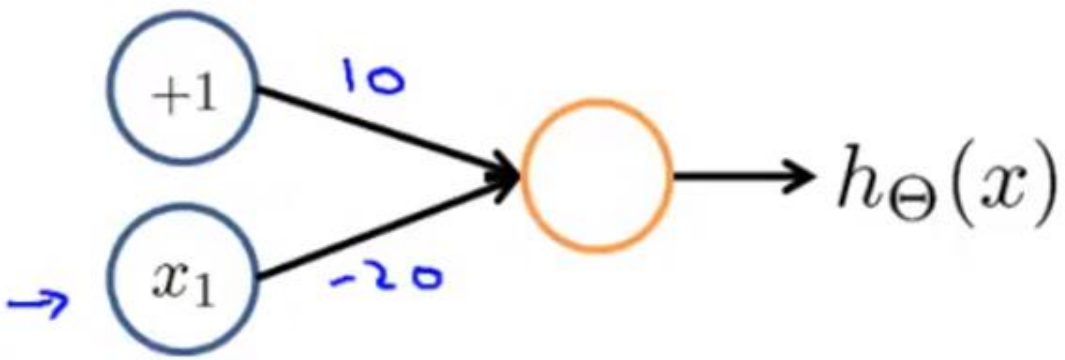


x_1	$h_\Theta(x)$
0	
1	

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

→ x_1 AND x_2 → x_1 OR x_2 {0,1}.

Negation: NOT x_1



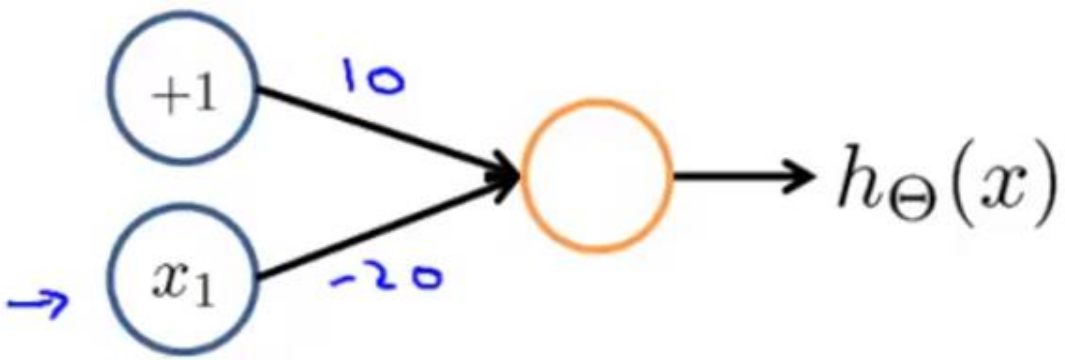
x_1	$h_\Theta(x)$
0	
1	

$$h_\Theta(x) = g(10 - 20x_1)$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

→ x_1 AND x_2 → x_1 OR x_2 {0,1}.

Negation: NOT x_1



x_1	$h_\Theta(x)$
0	$g(10) \approx 1$
1	$g(-10) \approx 0$

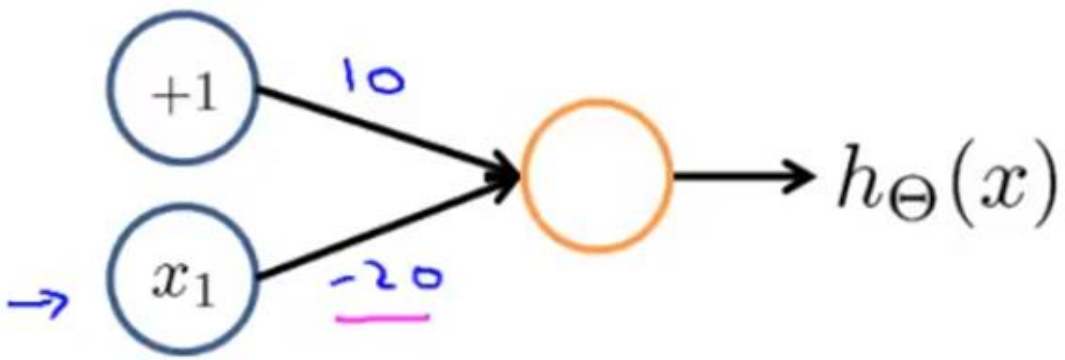
$$h_\Theta(x) = g(10 - 20x_1)$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

→ x_1 AND x_2 → x_1 OR x_2 {0,1}.

Negation:

NOT x_1



x_1	$h_\Theta(x)$
0	$g(10) \approx 1$
1	$g(-10) \approx 0$

$h_\Theta(x) = g(10 - 20x_1)$

→ (NOT x_1) AND (NOT x_2)

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

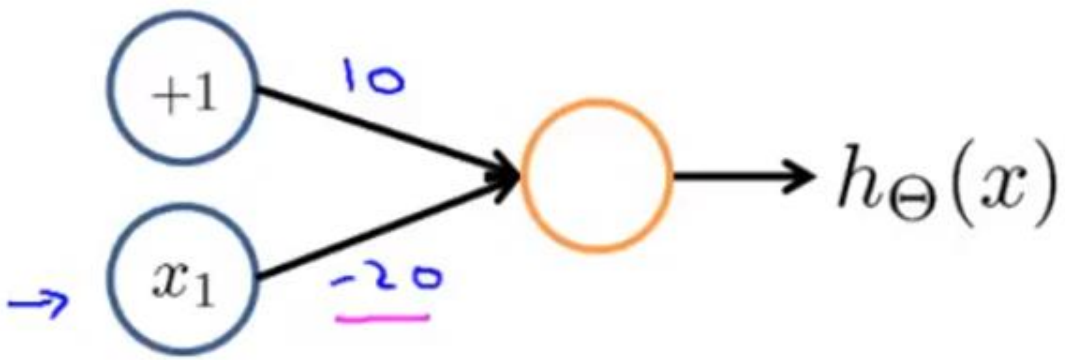
→ x_1 AND x_2

→ x_1 OR x_2

{0,1}.

Negation:

NOT x_1



x_1	$h_\Theta(x)$
0	$g(10) \approx 1$
1	$g(-10) \approx 0$

$h_\Theta(x) = g(10 - 20x_1)$

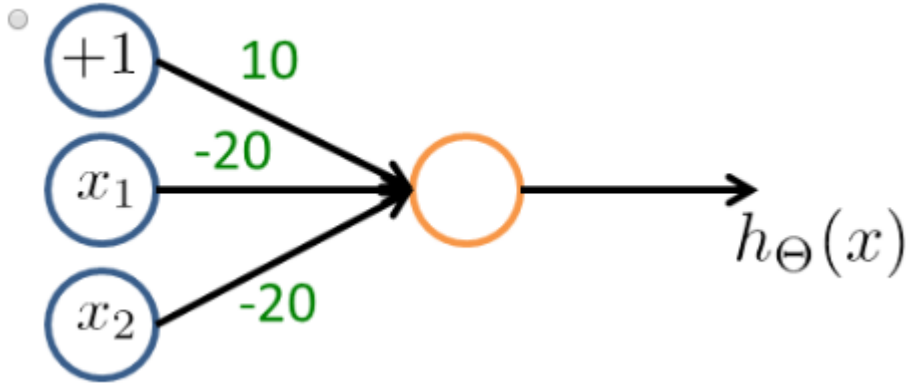
→ (NOT x_1) AND (NOT x_2)

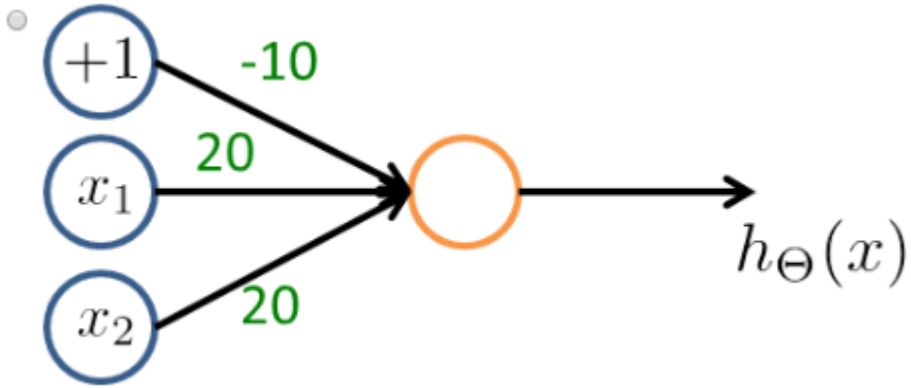
(=1 if and only if $x_1 = x_2 = 0$

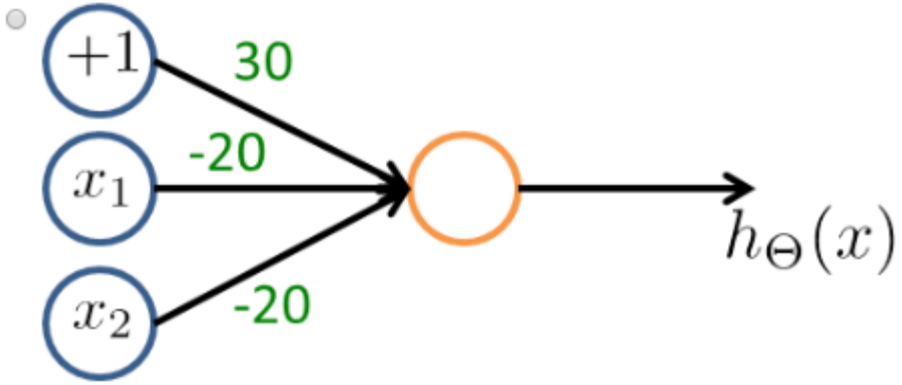
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

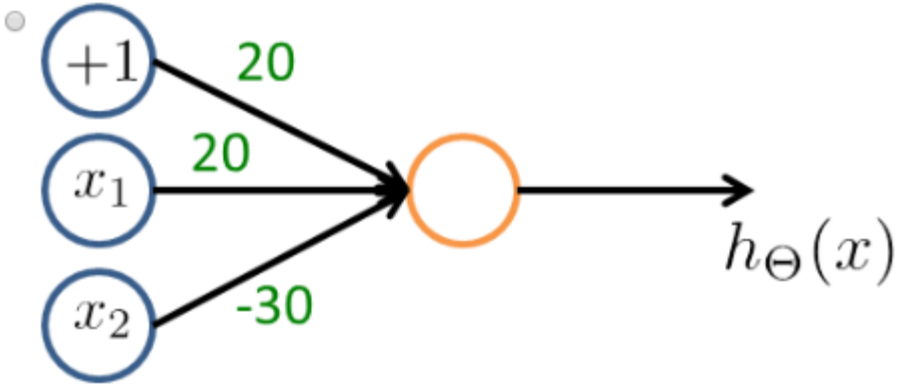
Exercise

- Suppose that $x1$ and $x2$ are binary valued (0 or 1). Which of the following networks (approximately) computes the boolean function (NOT $x1$) AND (NOT $x2$)?

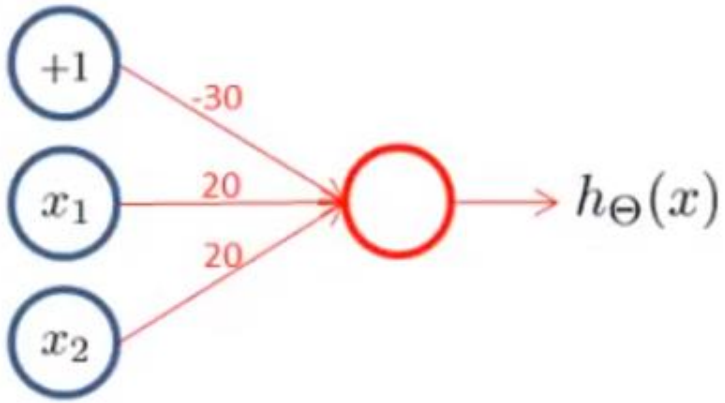




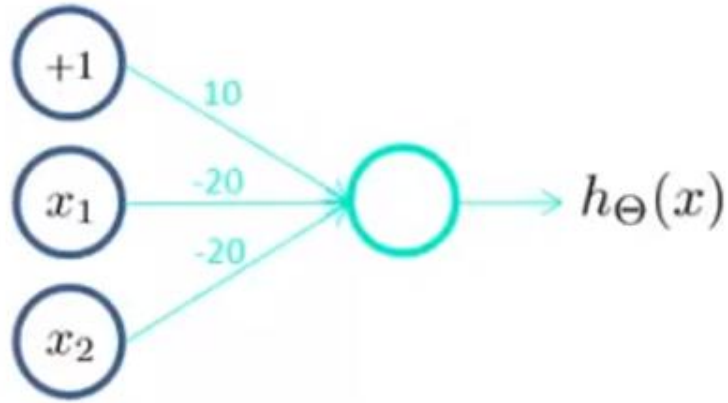




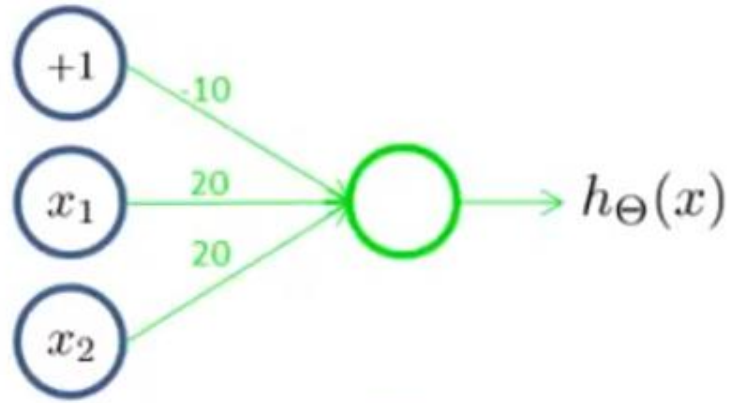
Putting it together: x_1 XNOR x_2



x_1 AND x_2



(NOT x_1) AND (NOT x_2)

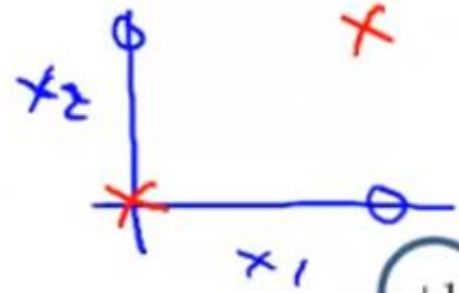


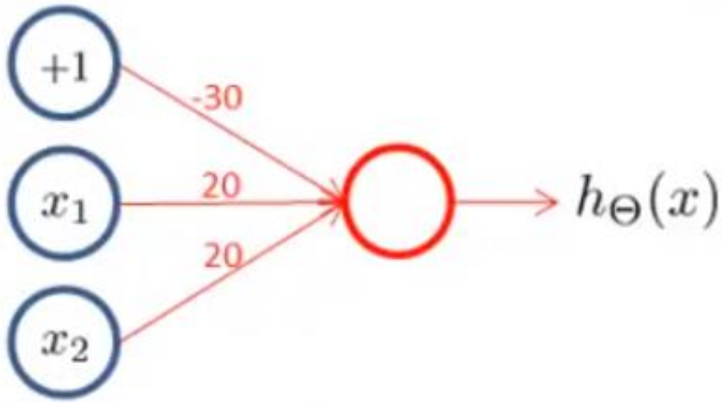
x_1 OR x_2

x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0			
0	1			
1	0			
1	1			

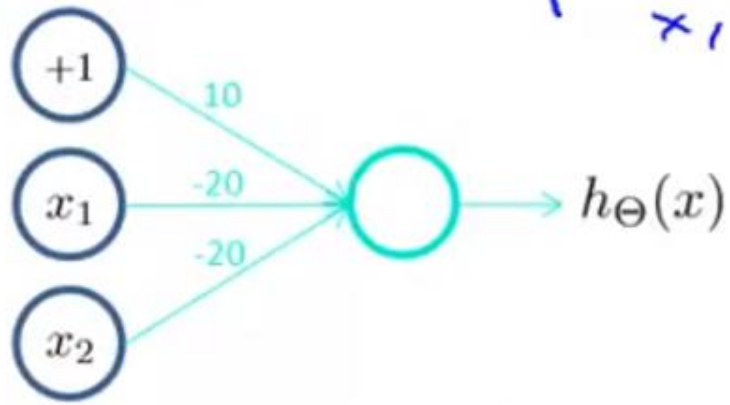
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Putting it together: x_1 XNOR x_2

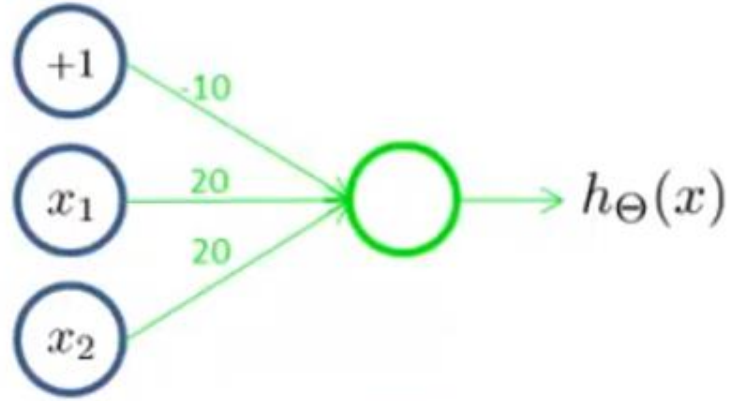




→ x_1 AND x_2



→ (NOT x_1) AND (NOT x_2)

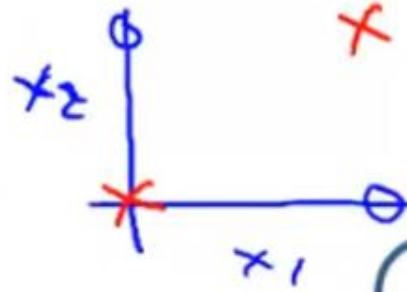


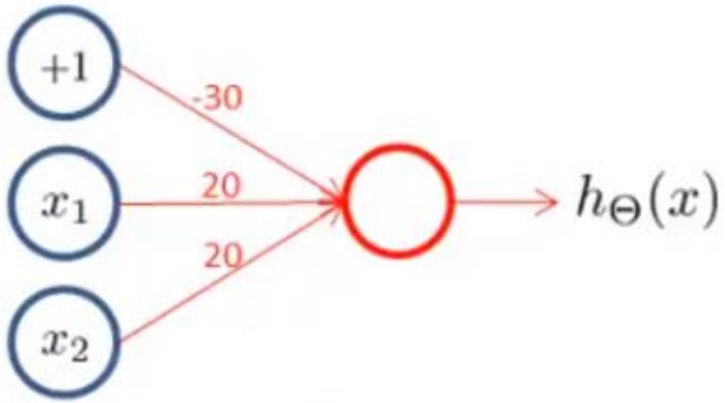
→ x_1 OR x_2

x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0			
0	1			
1	0			
1	1			

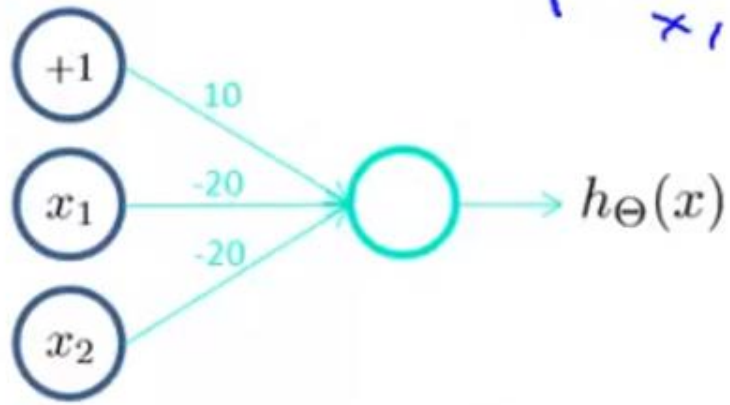
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Putting it together: x_1 XNOR x_2

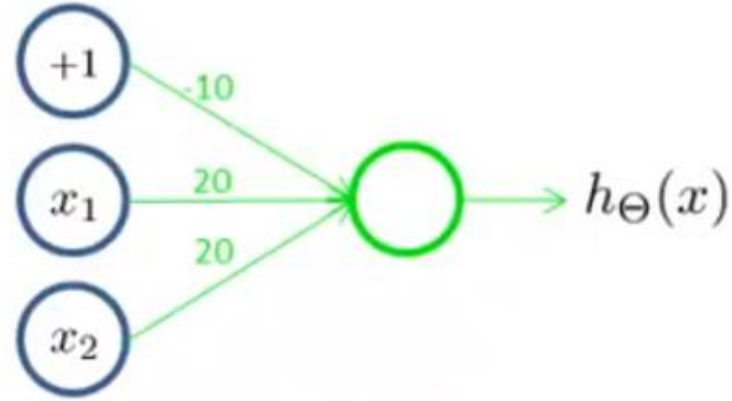




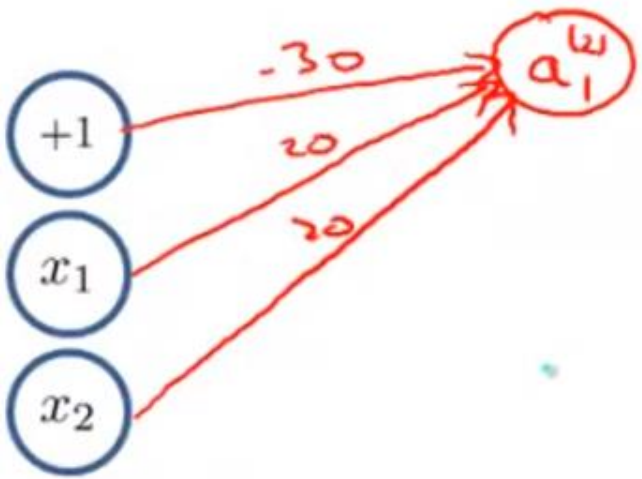
→ x_1 AND x_2



→ (NOT x_1) AND (NOT x_2)



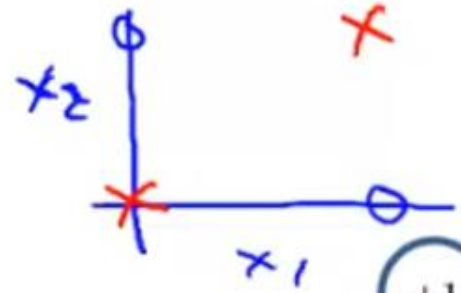
→ x_1 OR x_2

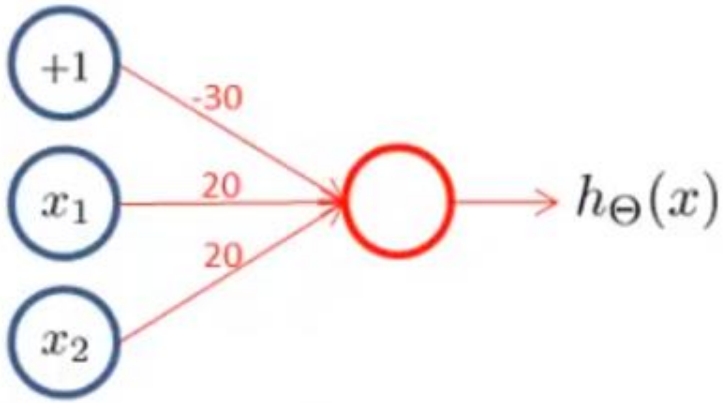


x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0			
0	1			
1	0			
1	1			

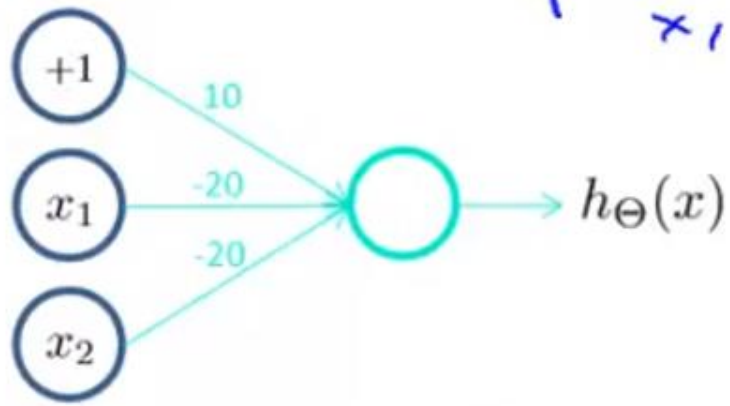
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Putting it together: x_1 XNOR x_2

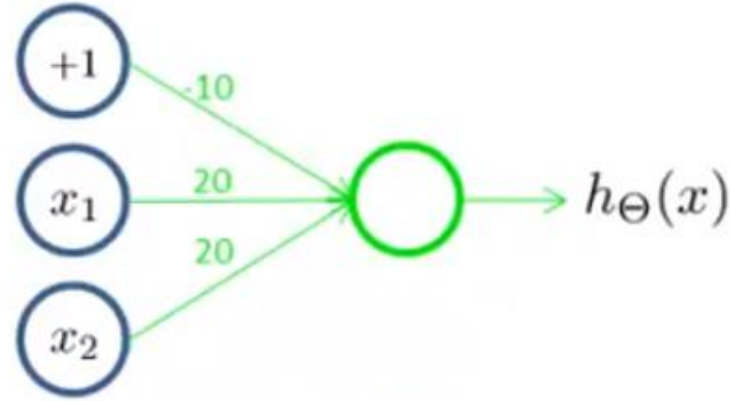




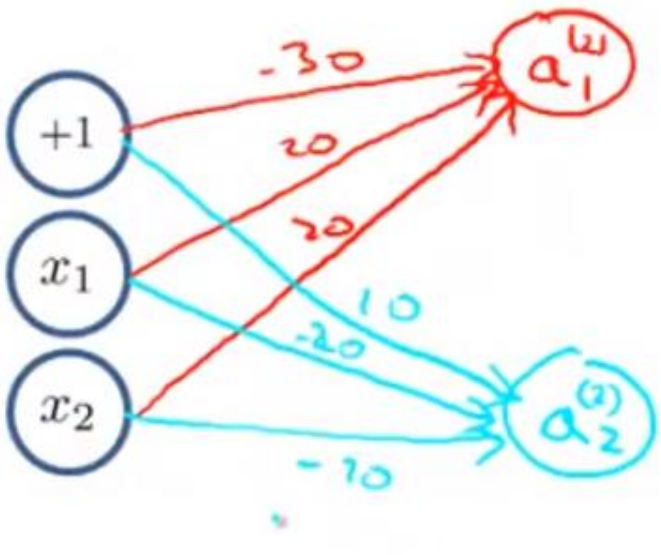
→ x_1 AND x_2



→ (NOT x_1) AND (NOT x_2)



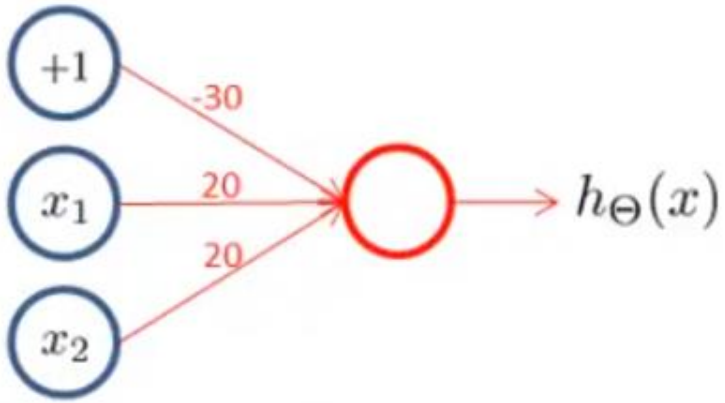
→ x_1 OR x_2



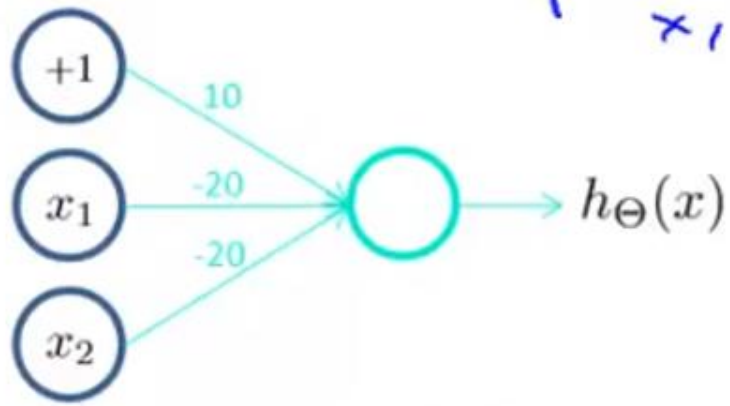
x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0			
0	1			
1	0			
1	1			

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

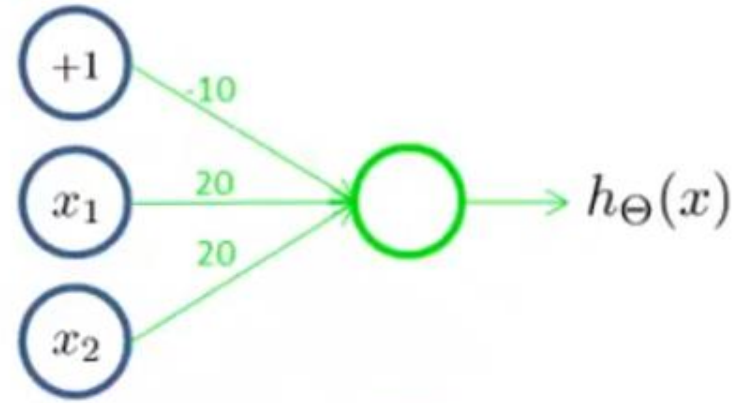
Putting it together: x_1 XNOR x_2



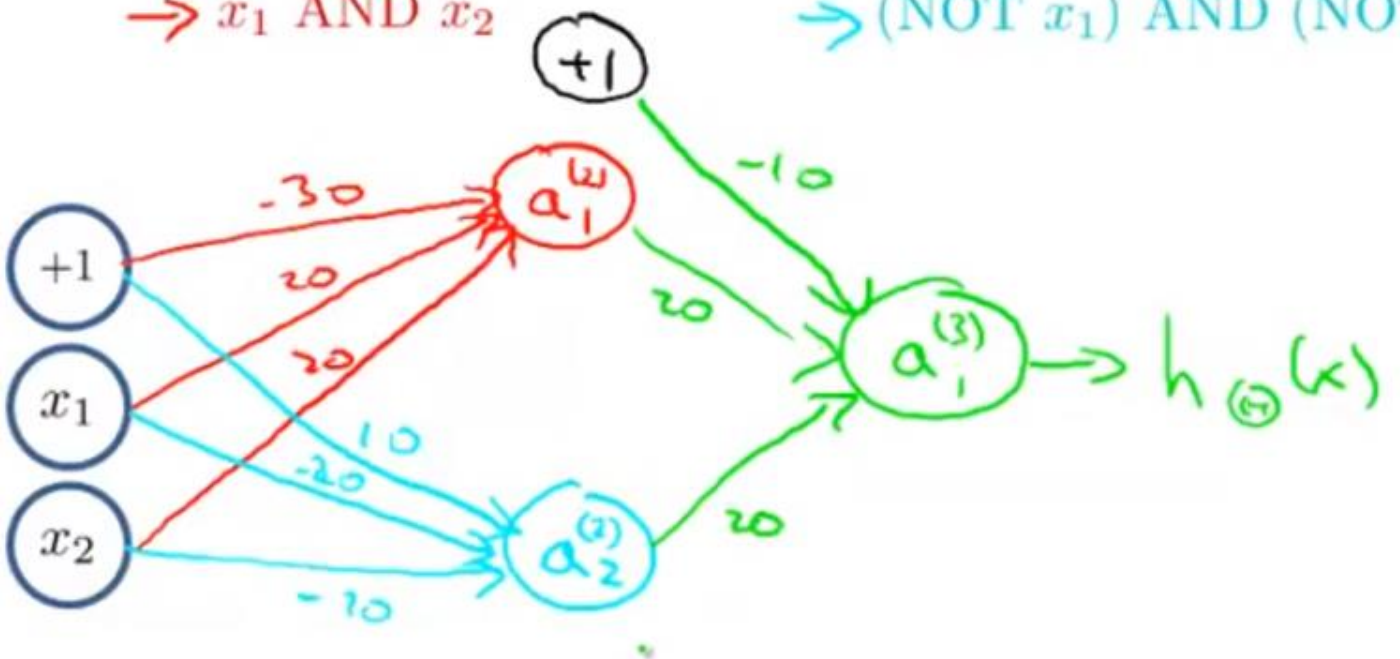
$\rightarrow x_1$ AND x_2



$\rightarrow$ (NOT x_1) AND (NOT x_2)



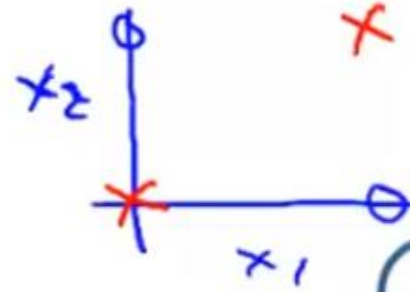
$\rightarrow x_1$ OR x_2

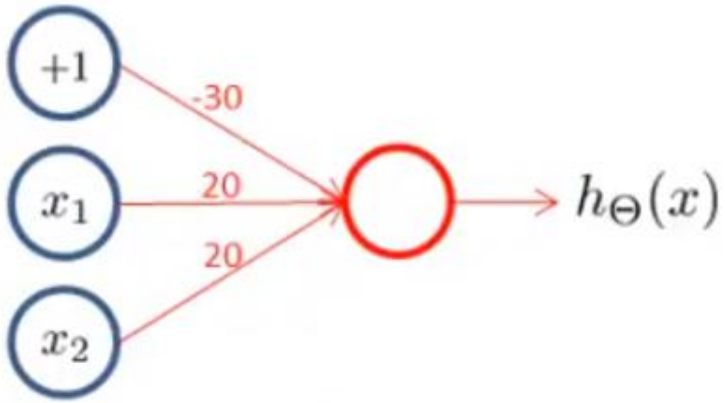


x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0	0	1	
0	1	0	0	
1	0	0	0	
1	1	1	0	

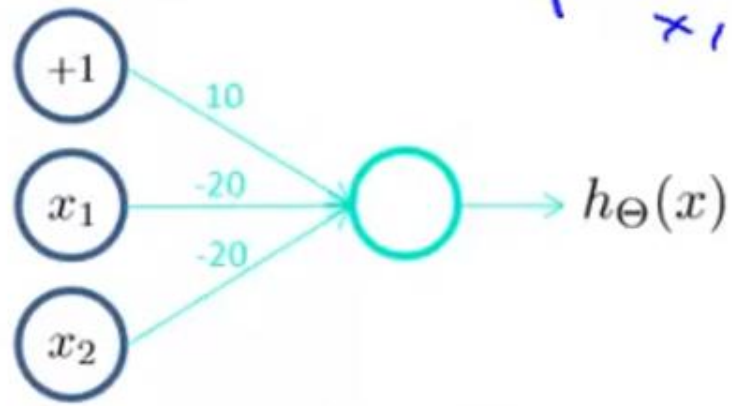
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Putting it together: x_1 XNOR x_2

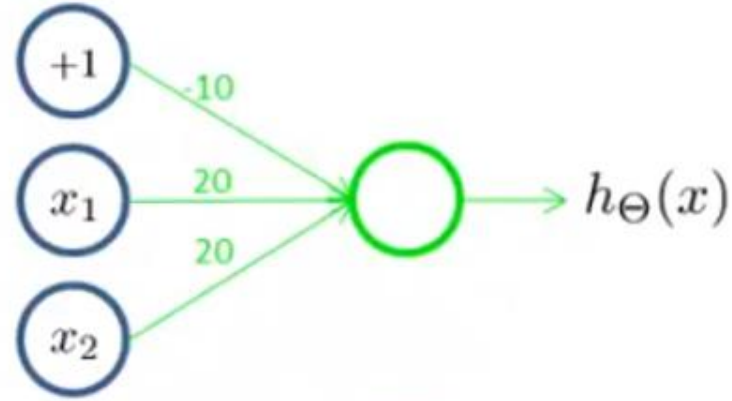




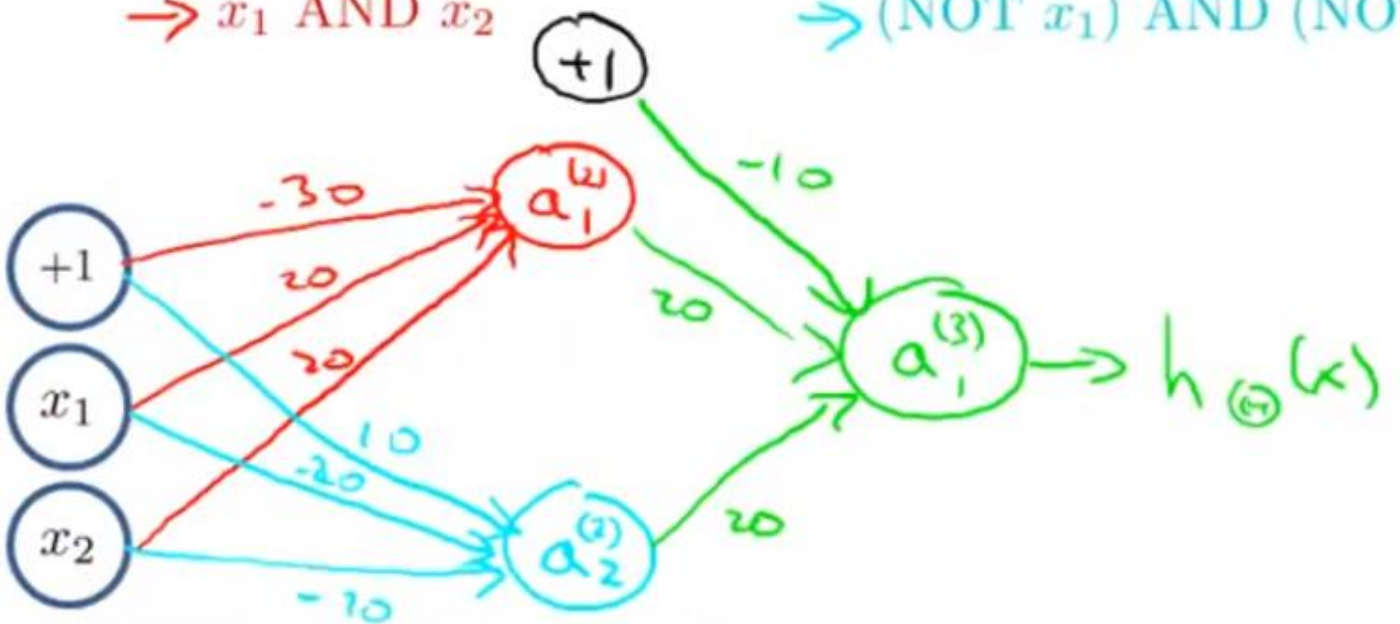
→ x_1 AND x_2



→ (NOT x_1) AND (NOT x_2)



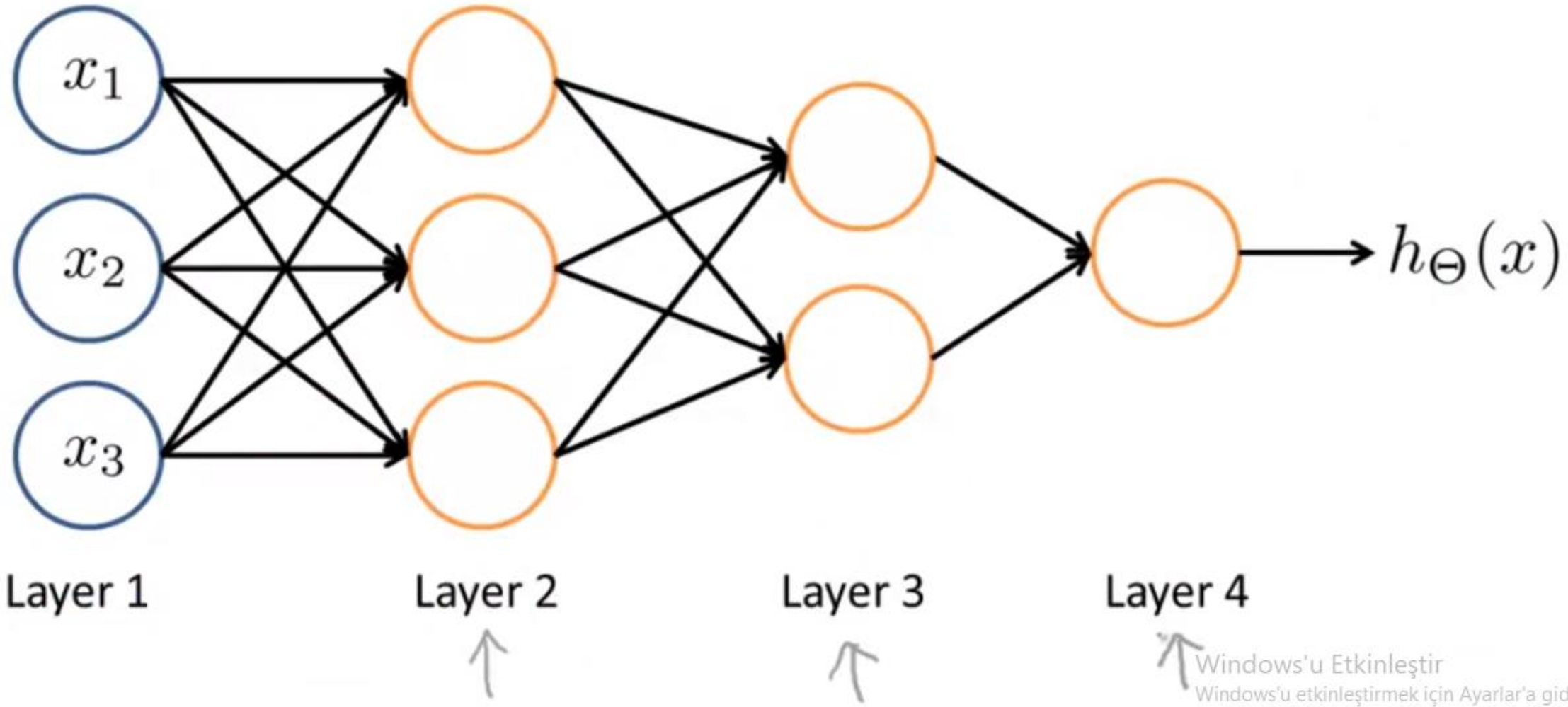
→ x_1 OR x_2



x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\Theta(x)$
0	0	0	1	1
0	1	0	0	0
1	0	0	0	0
1	1	1	0	1

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Neural Network intuition



Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiclass Classification

Applications

Neural Networks: Representation

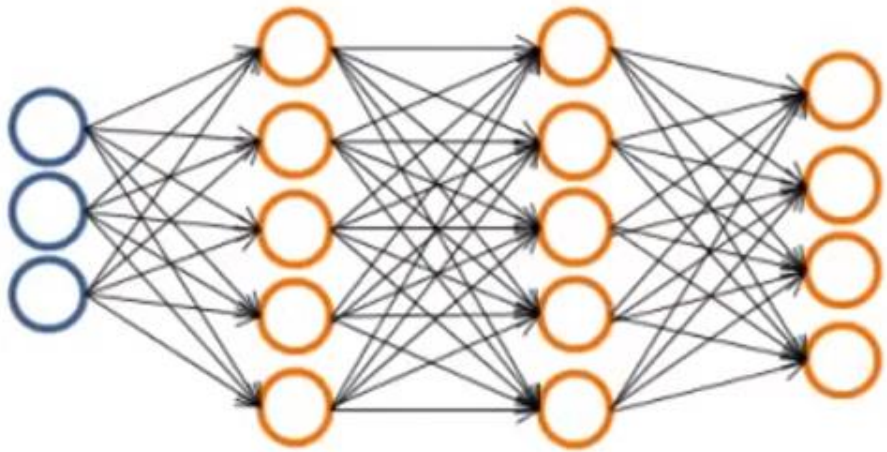
Multiple output units: One-vs-all.

Pedestrian

Car

Motorcycle

Truck



$$h_\Theta(x) \in \mathbb{R}^4$$

Want $h_\Theta(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

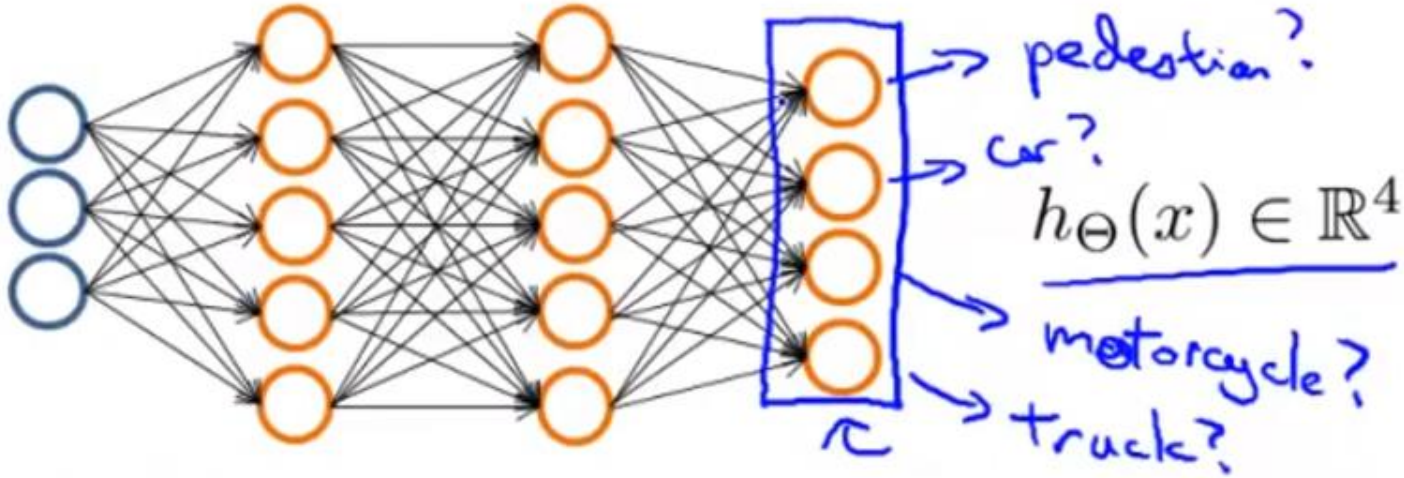
Multiple output units: One-vs-all.

Pedestrian

Car

Motorcycle

Truck

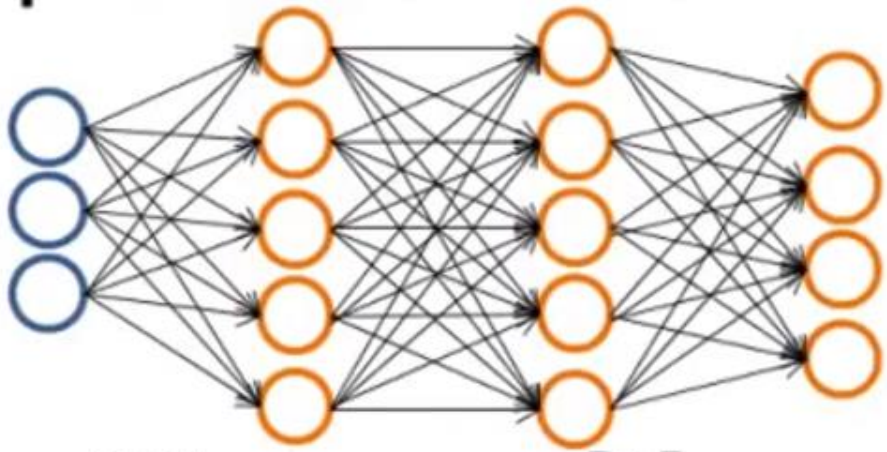


Want $h_\Theta(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiple output units: One-vs-all.



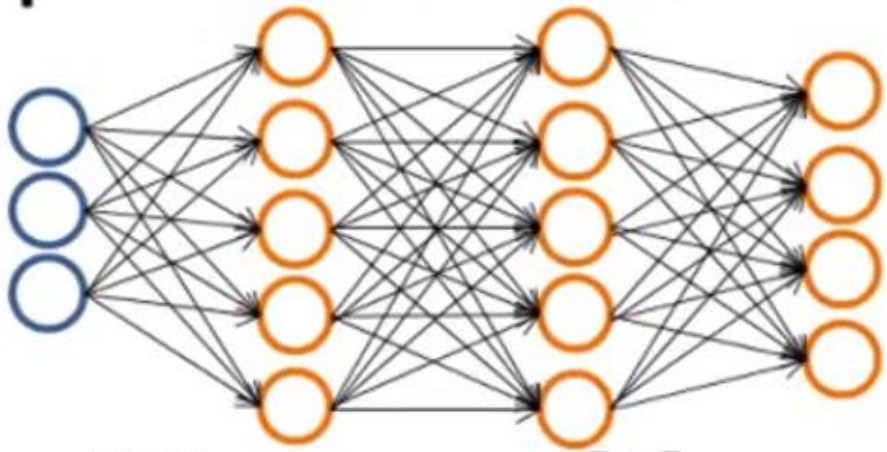
$h_\Theta(x) \in \mathbb{R}^4$

Want $h_\Theta(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiple output units: One-vs-all.



$h_\Theta(x) \in \mathbb{R}^4$

Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian when car when motorcycle

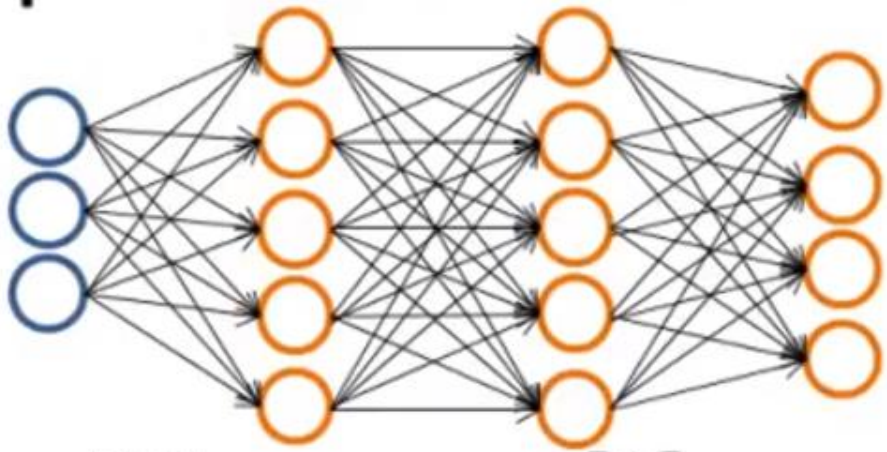
Training set: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$

$y^{(i)}$ one of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

pedestrian car motorcycle truck

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiple output units: One-vs-all.



$$h_\Theta(x) \in \mathbb{R}^4$$

Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Training set: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$

$y^{(i)}$ one of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

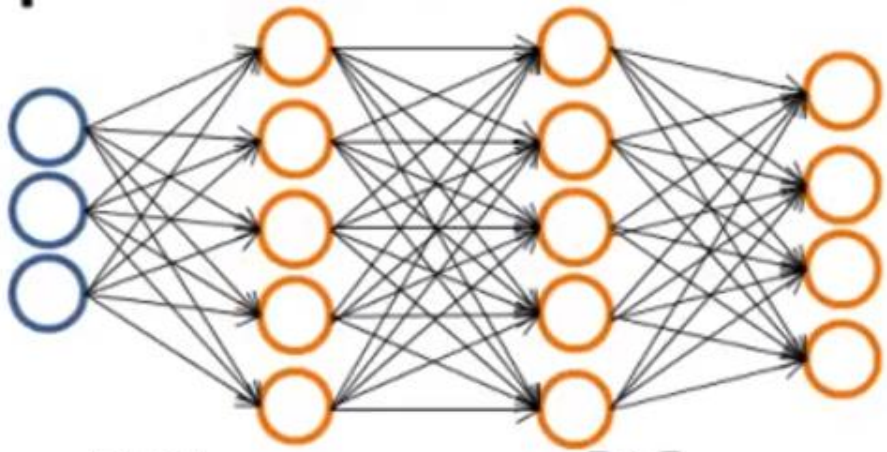
pedestrian car motorcycle truck

Previously

$y \in \{1, 2, 3, 4\}$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiple output units: One-vs-all.



$h_\Theta(x) \in \mathbb{R}^4$

Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Training set: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$

$y^{(i)}$ one of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

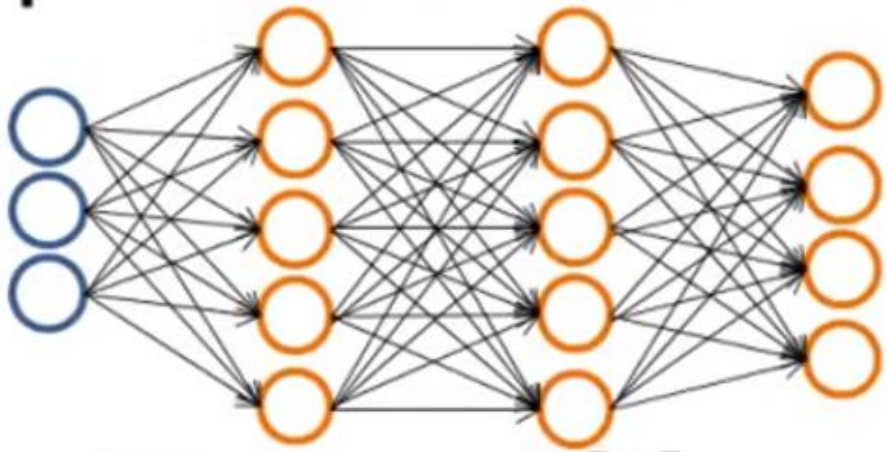
pedestrian car motorcycle truck

Previously $y \in \{1, 2, 3, 4\}$

$(x^{(i)}, y^{(i)})$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Multiple output units: One-vs-all.



$$h_\Theta(x) \in \mathbb{R}^4$$

Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Training set: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$

$y^{(i)}$ one of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

pedestrian car motorcycle truck

Previously

$y \in \{1, 2, 3, 4\}$

$(x^{(i)}, y^{(i)})$

$h_\Theta(x^{(i)}) \approx y^{(i)}$

$\in \mathbb{R}^4$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Exercise

- Suppose you have a multi-class classification problem with 10 classes.
- Your neural network has 3 layers, and the hidden layer (layer 2) has 5 units.
- Using the one-vs-all method described here, how many elements does $\theta^{(2)}$ have?