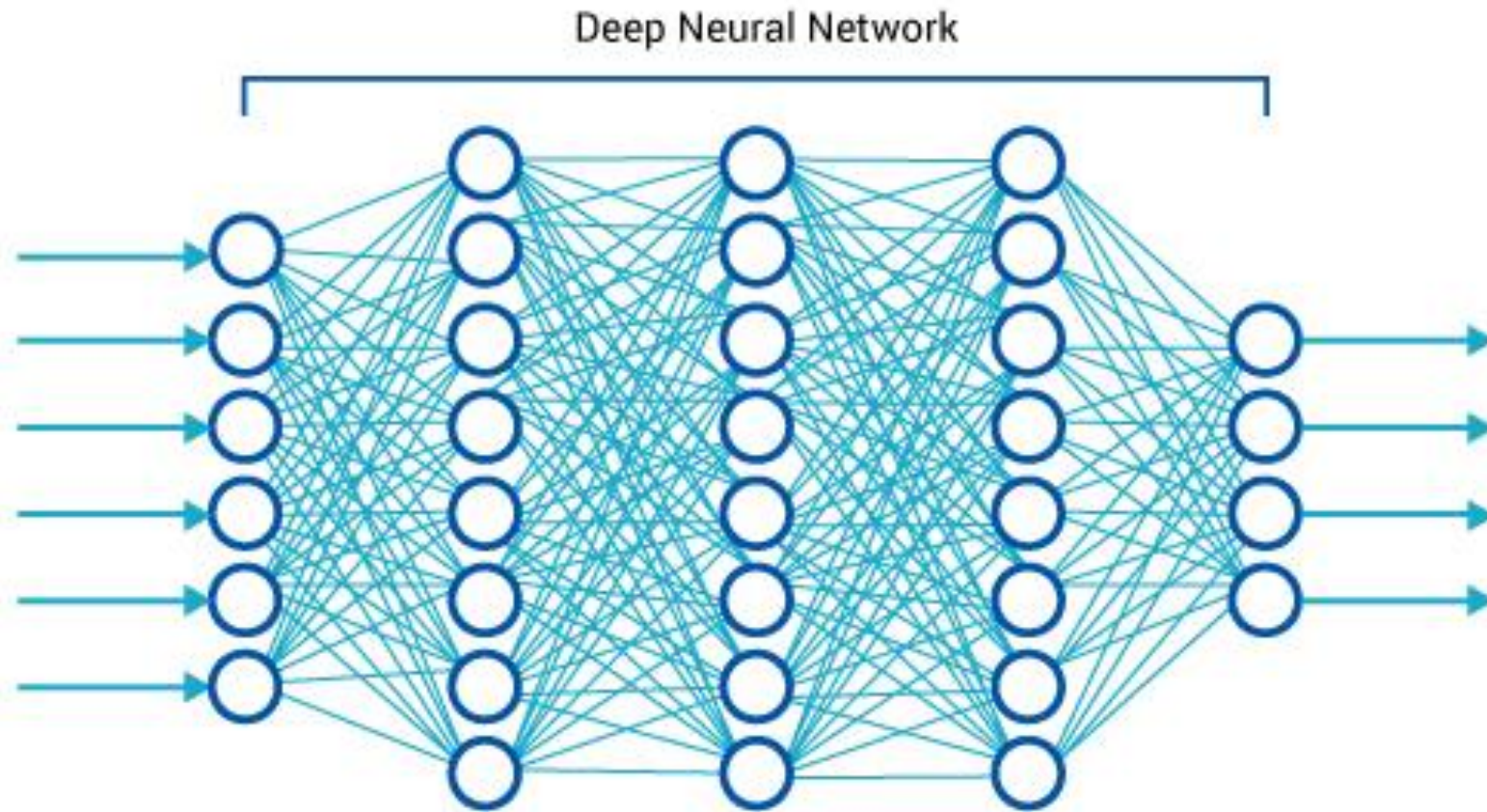
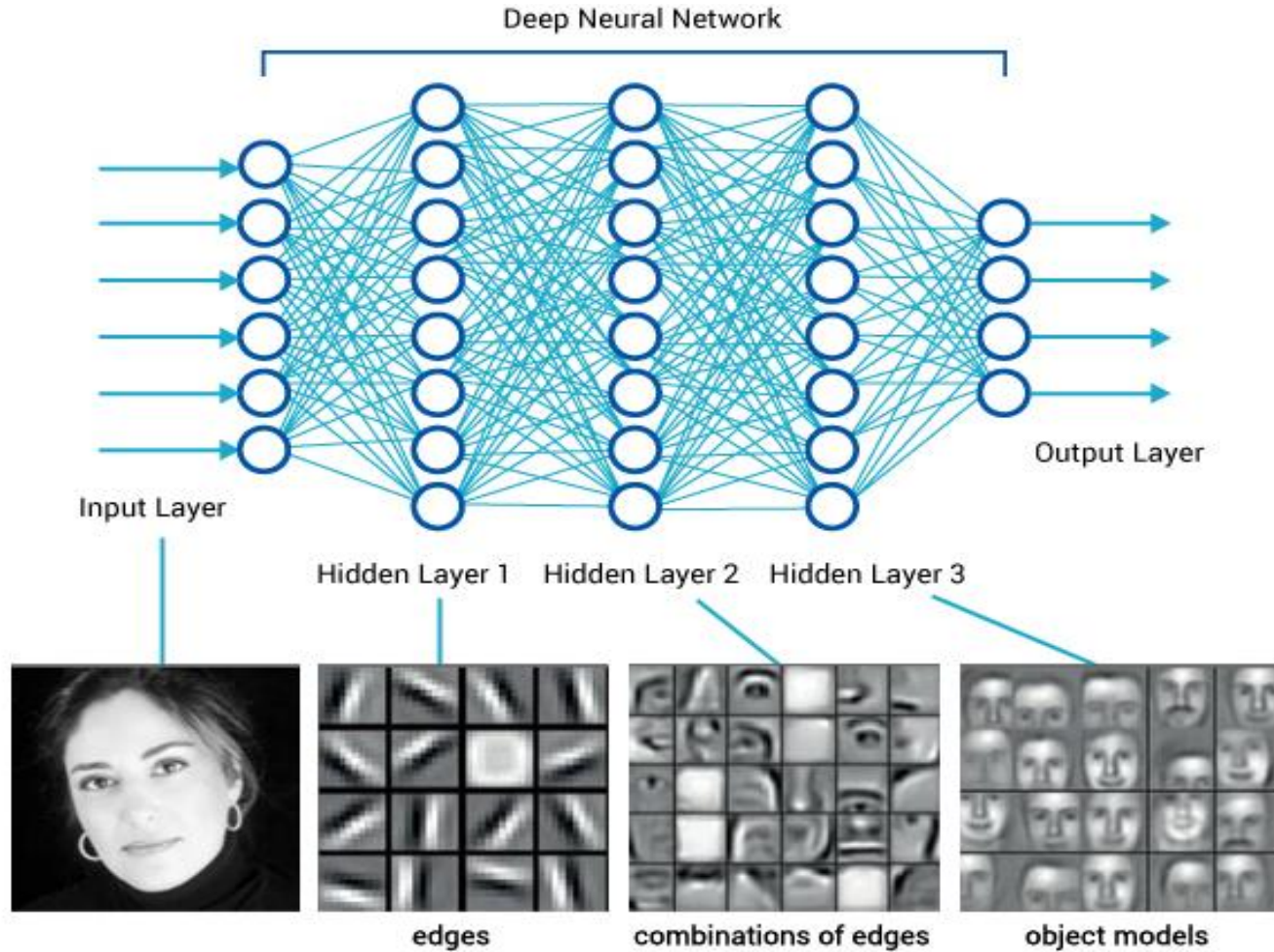


Deep Learning



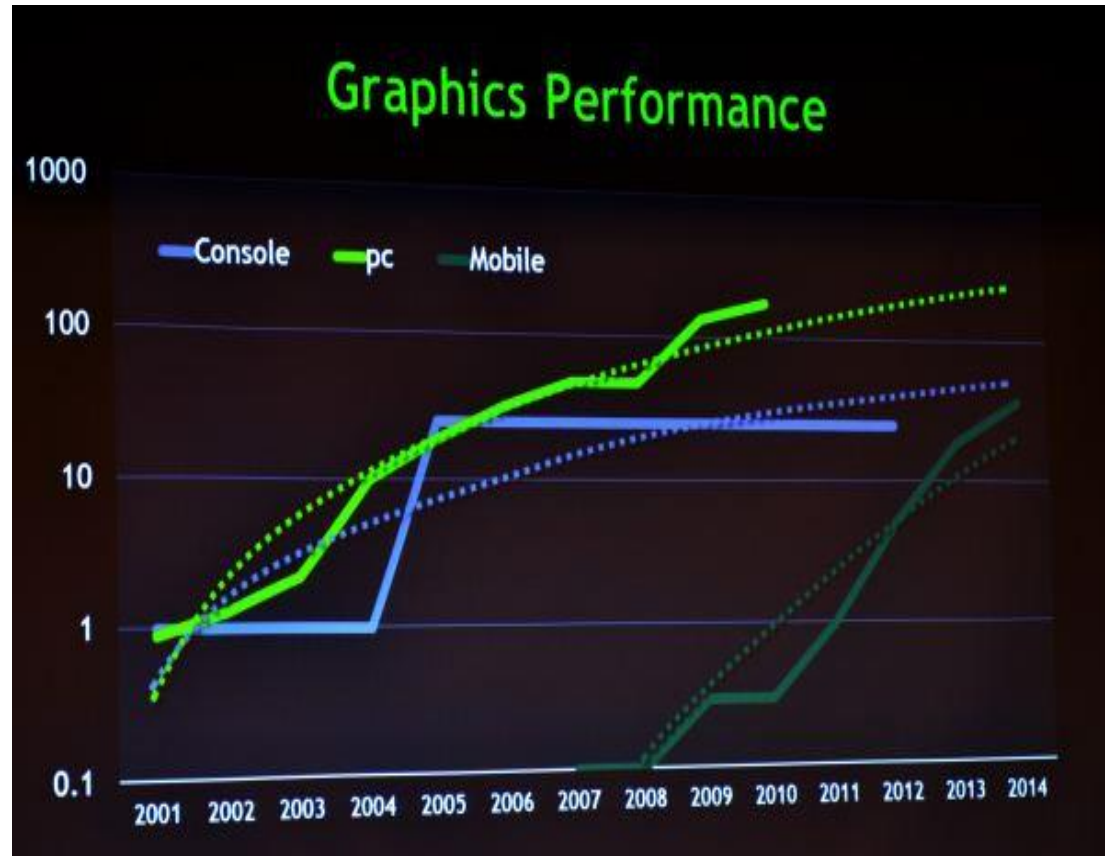
Deep Learning



Why is deep learning a growing trend?

- **few feature engineering**
- **state-of-the-art performance**

Deep Learning



Deep Learning Heroes



[Geoffrey Hinton](#)



[Yann LeCun](#)





[Andrew Ng](#)




[Yoshua Bengio](#)


Lectures for Neural Network



 Catalog

Search catalog



Institutions 

Overview

Syllabus

Creators

Ratings and Reviews

Neural Networks for Machine Learning

[Go to Course](#)

Already enrolled


Home > Data Science > Machine Learning


Neural Networks for Machine Learning

About this course: Learn about artificial neural networks and how they're being used for machine learning, as applied to speech and object recognition, image segmentation, modeling language and human motion, etc. We'll emphasize both the basic algorithms and the practical tricks needed to get them to work well.

[More](#)

Created by: University of Toronto




Taught by:  Geoffrey Hinton, Professor
Department of Computer Science

Financial Aid is available for learners who cannot afford the fee. [Learn more and apply.](#)

[Neural Networks for Machine Learning – University of Toronto](#)

Lectures for Neural Network



Catalog

Search catalog

Institutions

YC

Overview

Syllabus

Creators

Ratings and Reviews

Machine Learning

Go to Course

Already enrolled


Home > Data Science > Machine Learning


Machine Learning

About this course: Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many

More

Created by: Stanford University



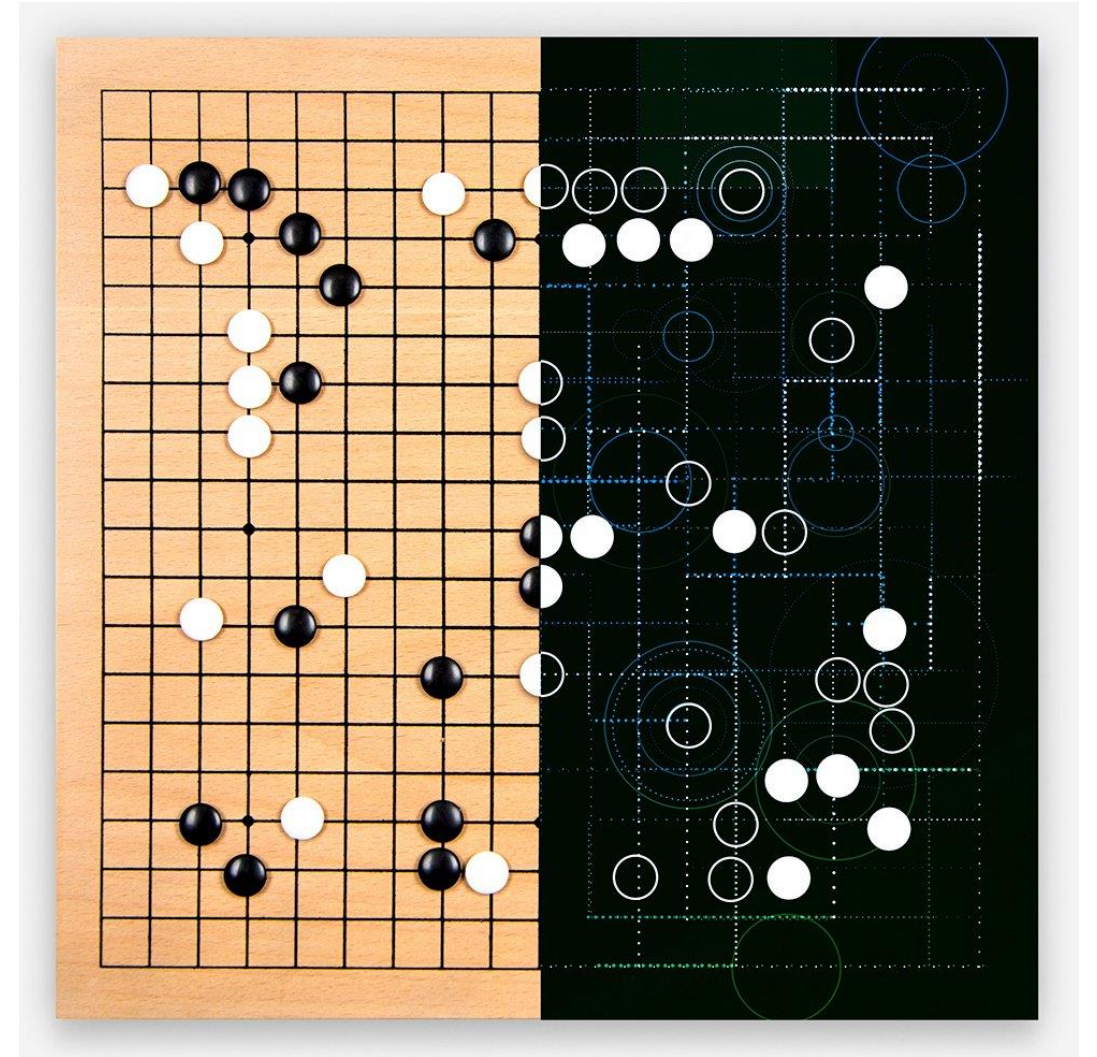


Taught by: Andrew Ng, Associate Professor, Stanford University; Chief Scientist, Baidu; Chairman and Co-founder, Coursera

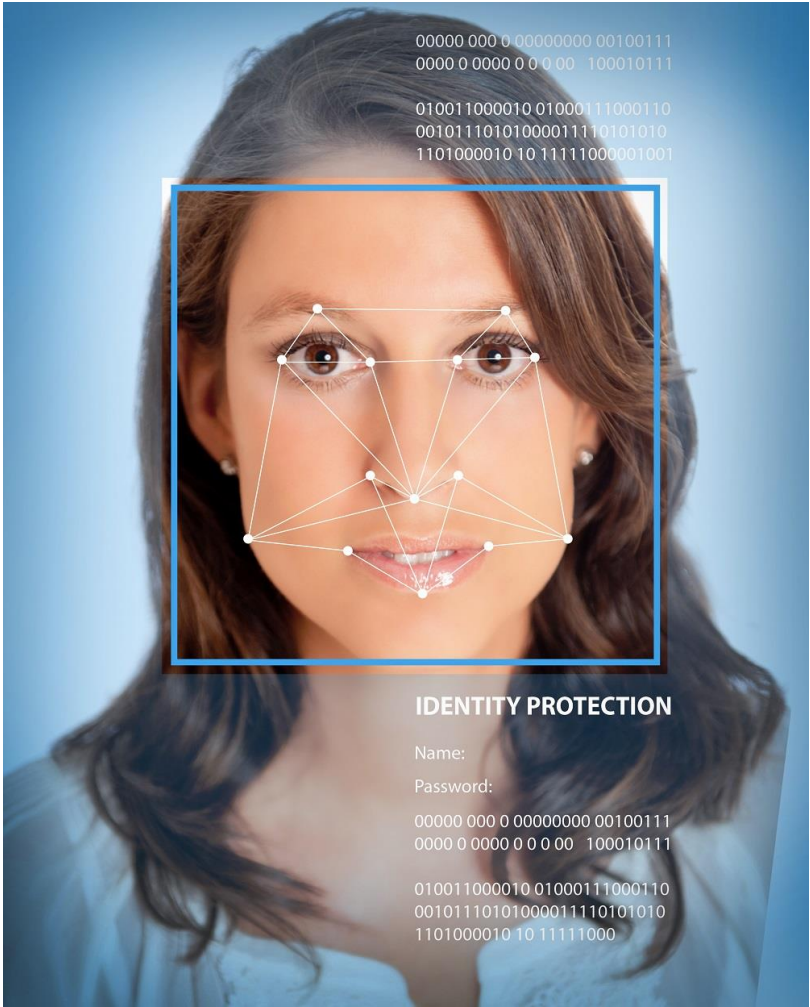
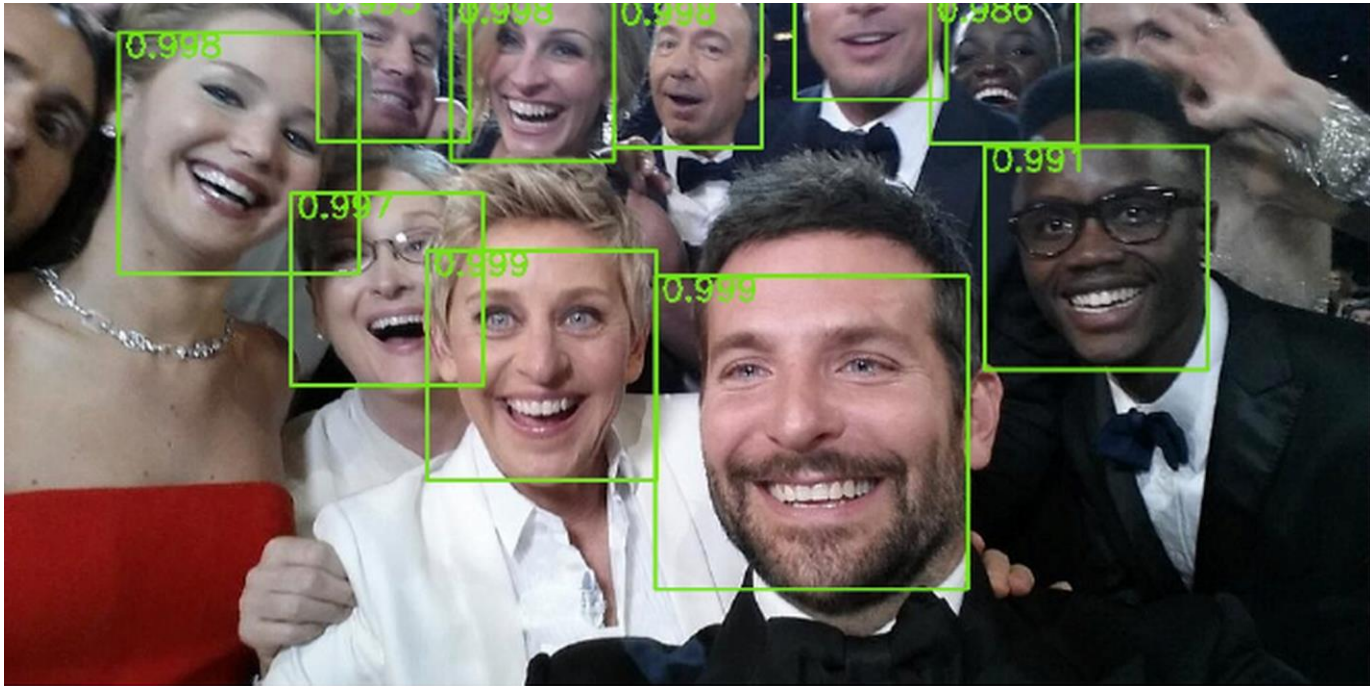
Financial Aid is available for learners who cannot afford the fee. [Learn more and apply.](#)

[Machine Learning – Stanford University](#)

Game AI



Face Detection & Recognition



Object Detection & Recognition

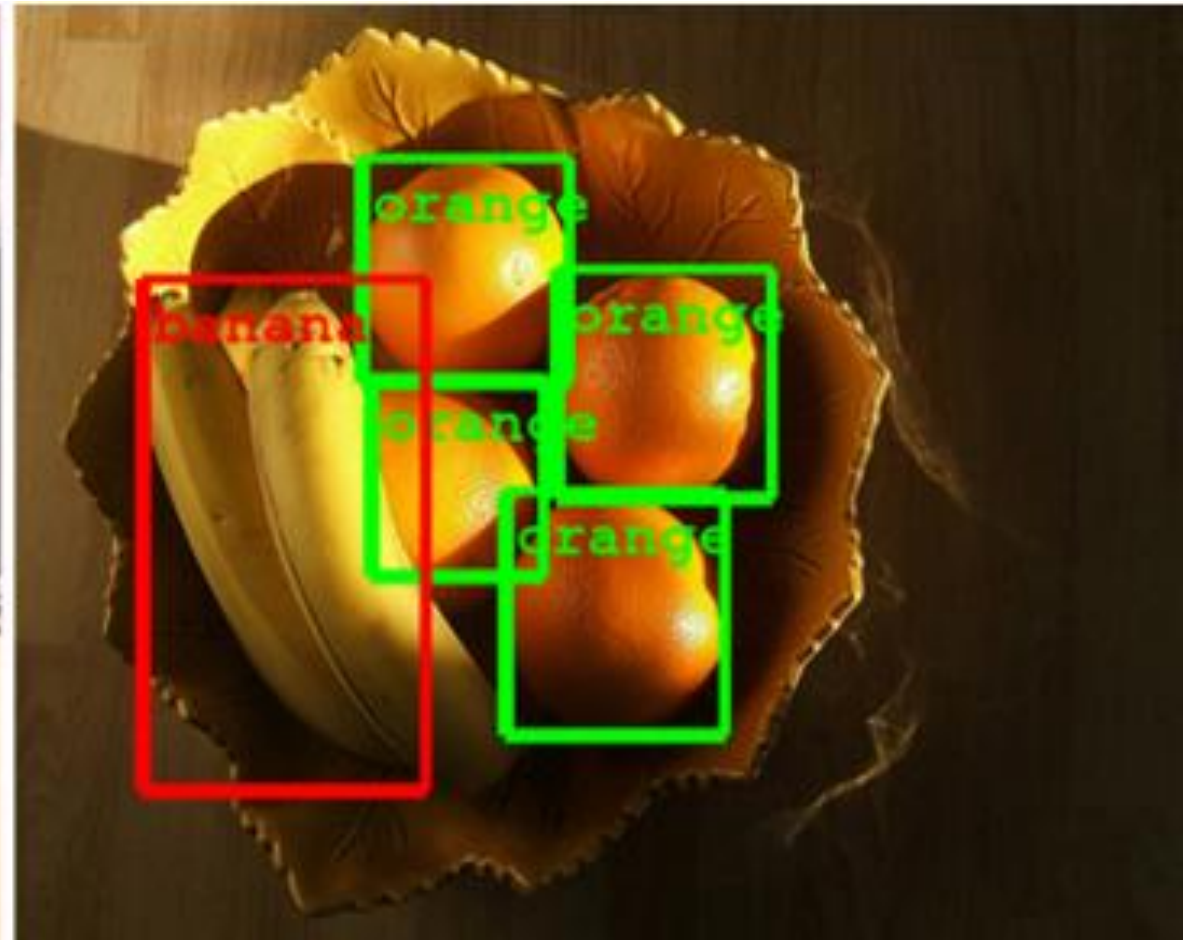
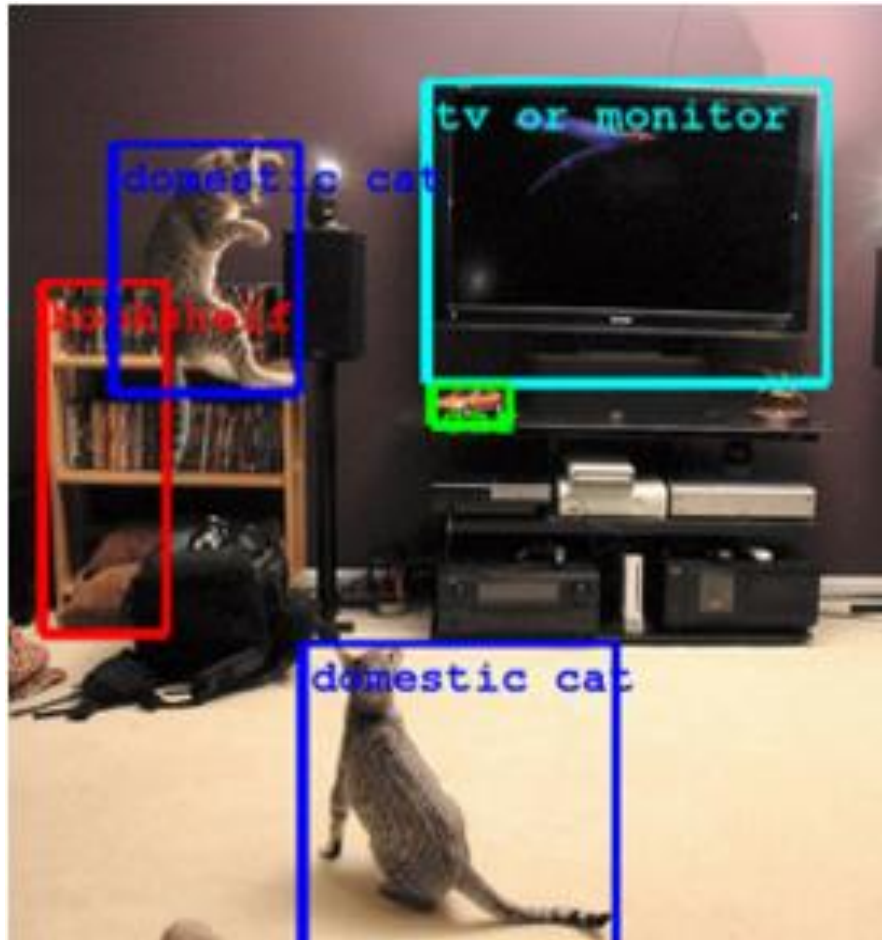
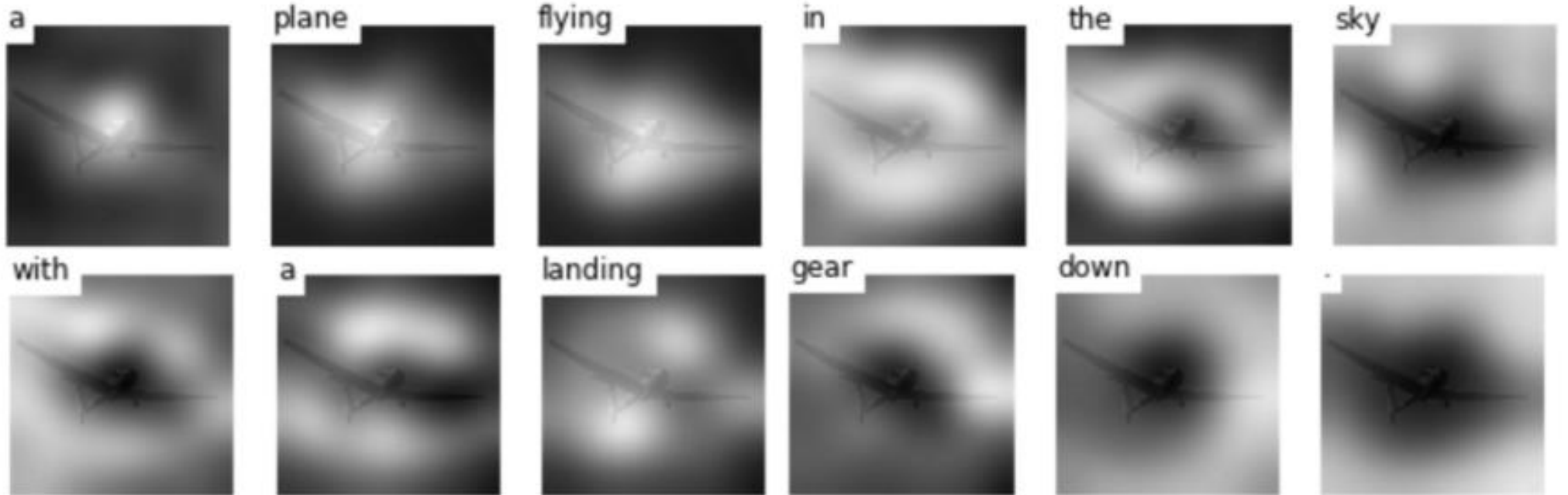


Image Captioning



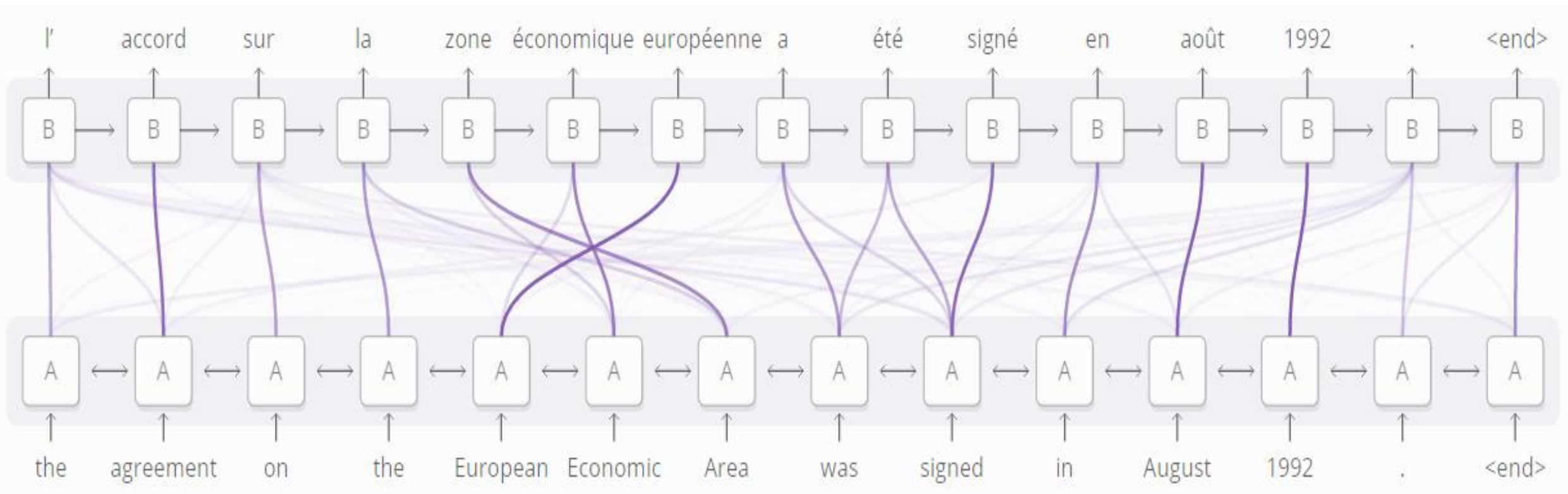
<https://github.com/yunjey/show-attend-and-tell>

Image Captioning

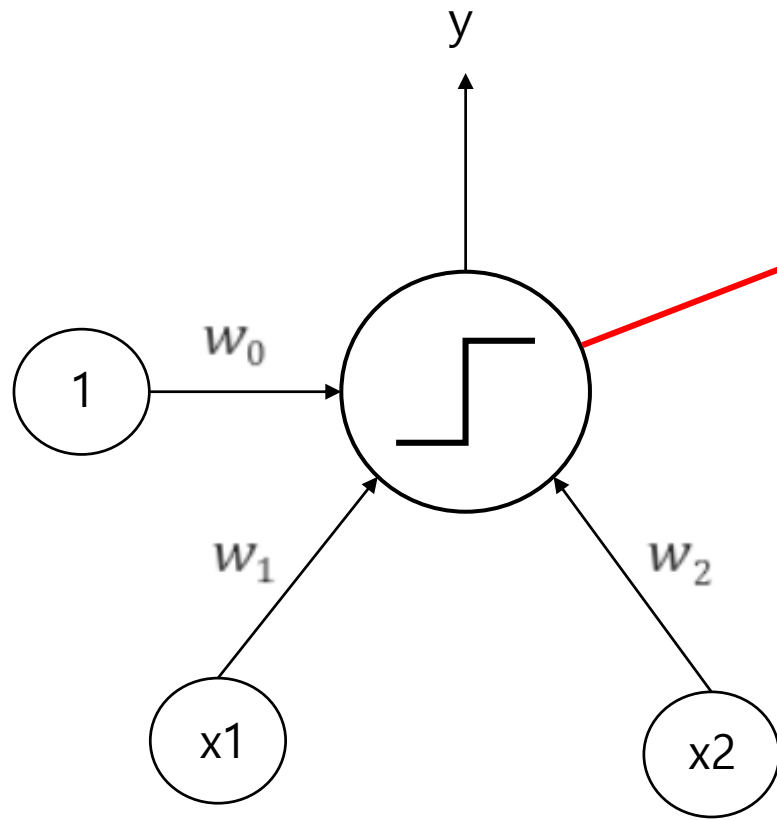


[DenseCap: Fully Convolutional Localization Networks for Dense Captioning](#)

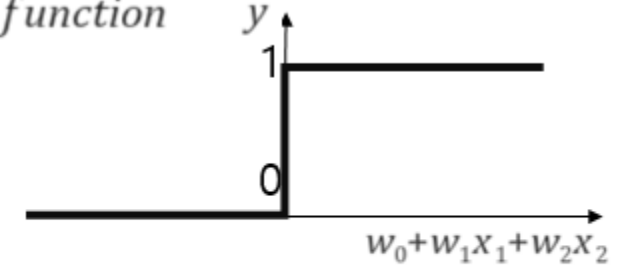
Machine Translation



Single Layer Perceptron



hard thresholding function



$$y = \begin{cases} 1 & w_0 + w_1x_1 + w_2x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

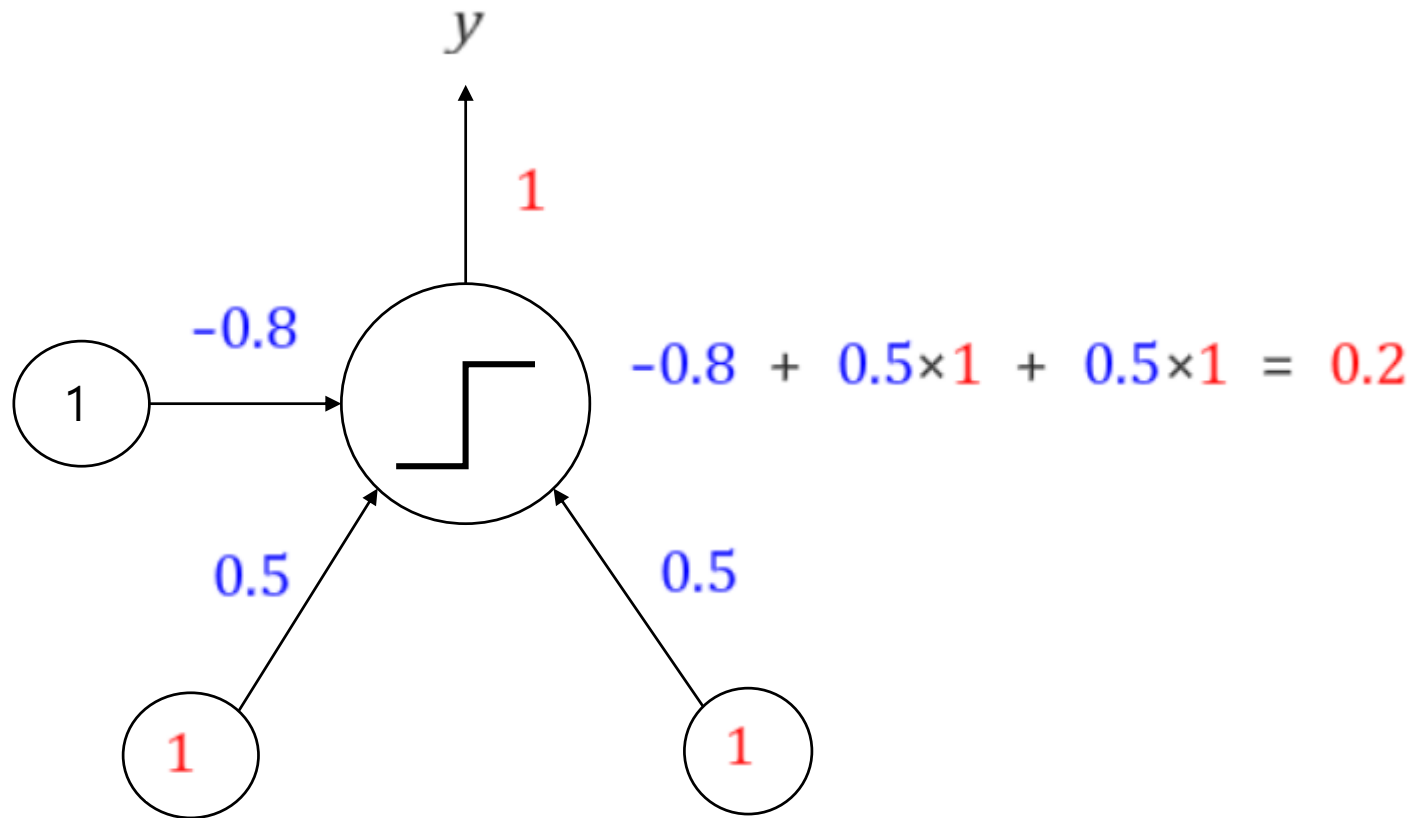
Inputs: x_1, x_2

Output: y

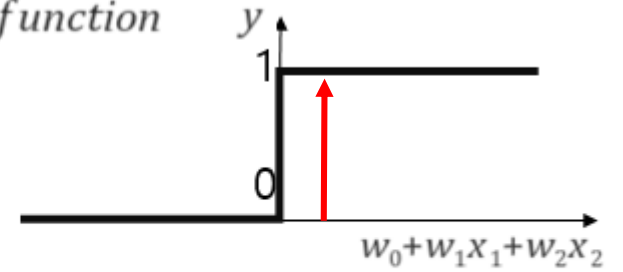
Weights: w_0, w_1, w_2

Bias: 1

Single Layer Perceptron



hard thresholding function

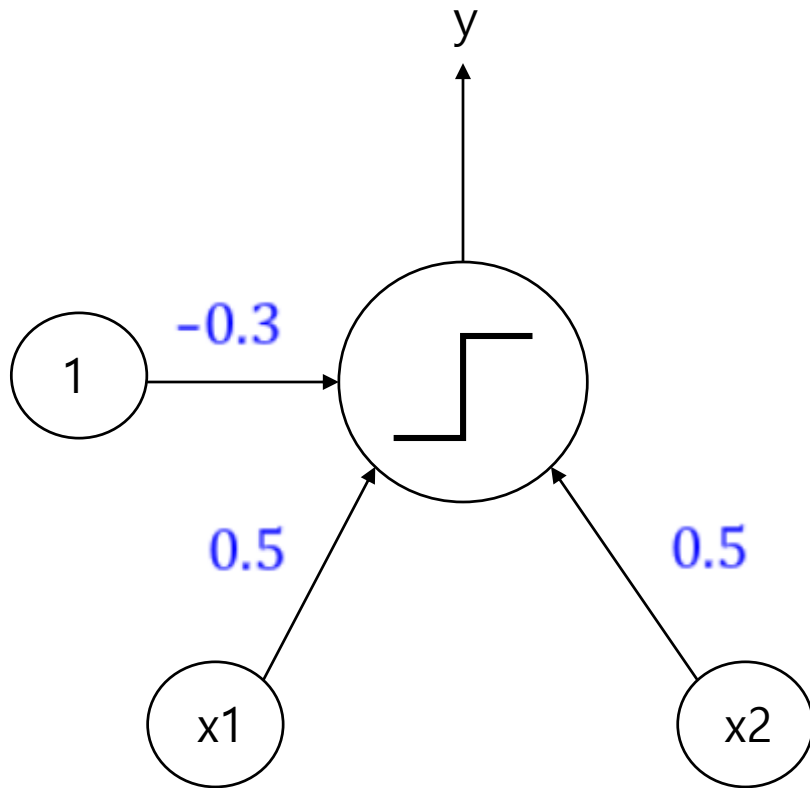


$$y = \begin{cases} 1 & w_0 + w_1x_1 + w_2x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

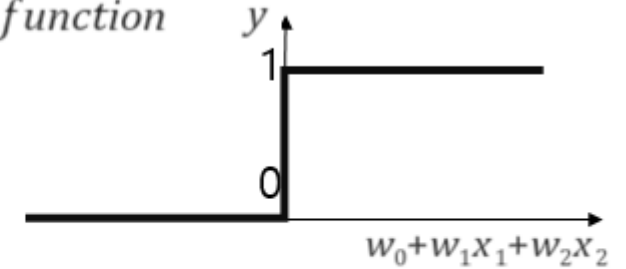
AND Gate

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

Single Layer Perceptron



hard thresholding function



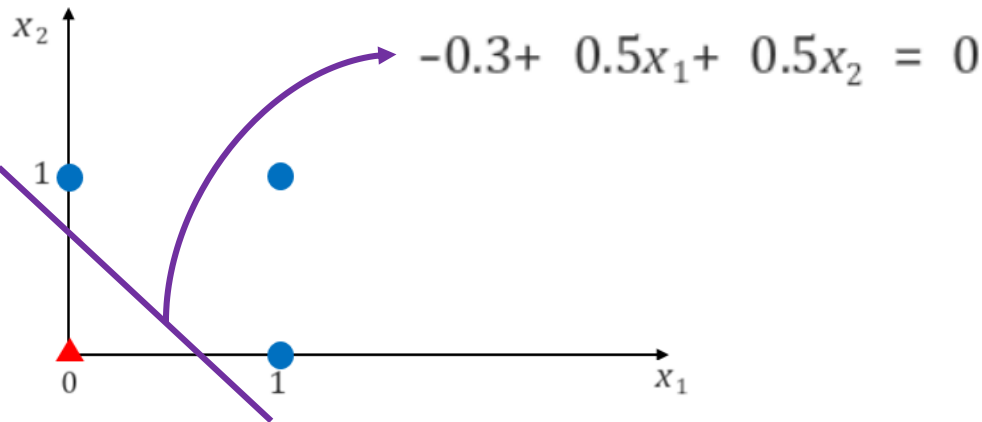
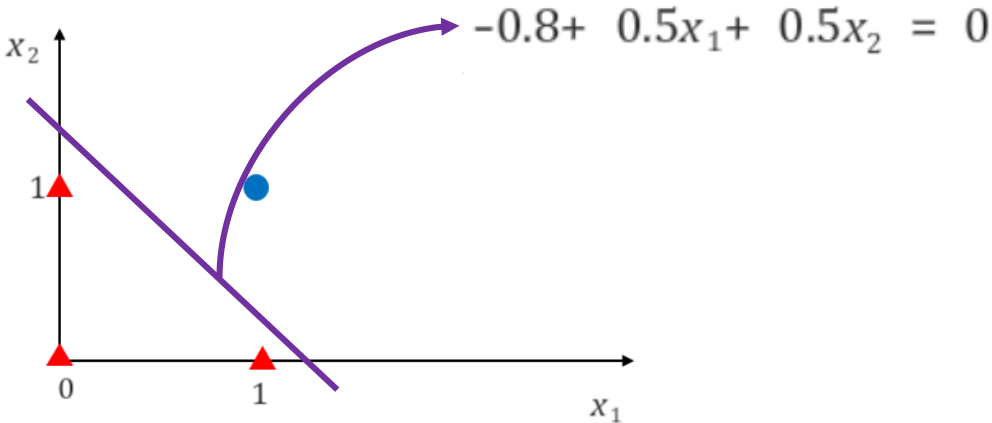
$$y = \begin{cases} 1 & w_0 + w_1x_1 + w_2x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

OR Gate

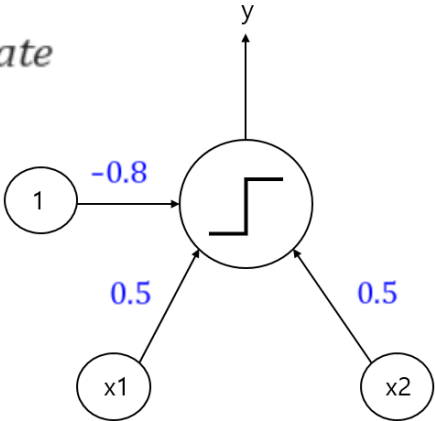
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

Single Layer Perceptron

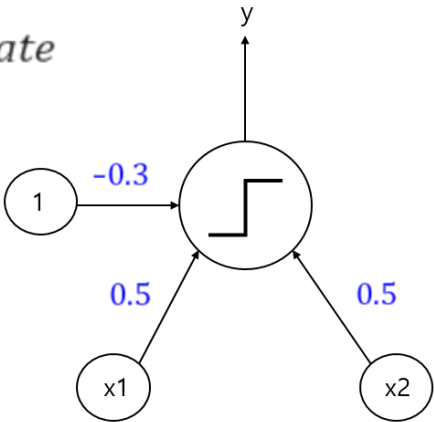
● : 1 ▲ : 0



AND Gate



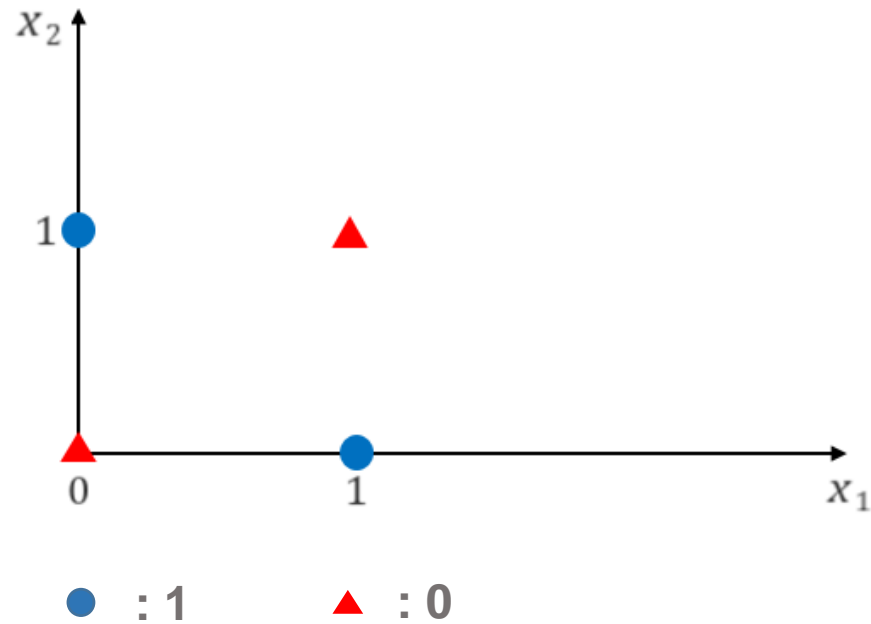
OR Gate



XOR Gate

Is it possible to solve XOR problem using a single layer perceptron?

No. Single layer perceptron can only solve linear problem. XOR problem is non-linear.

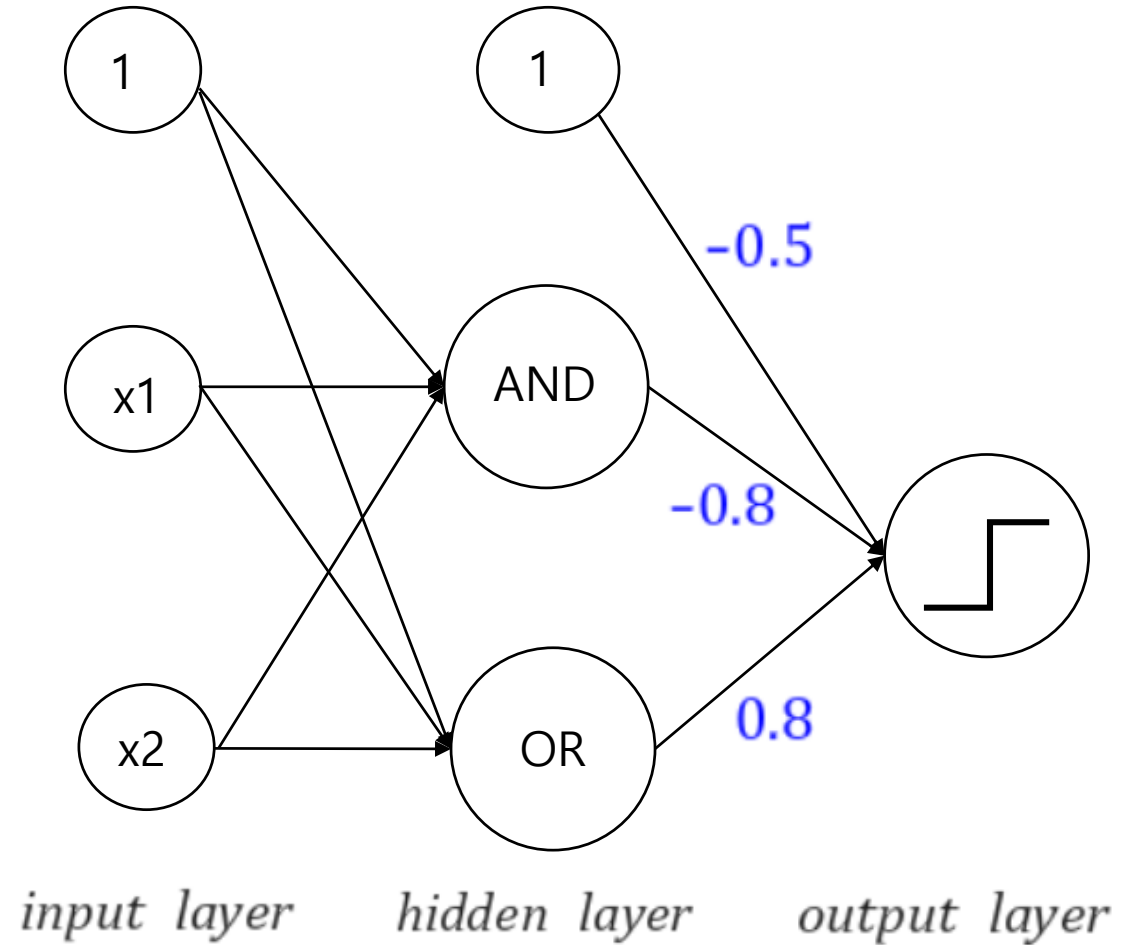
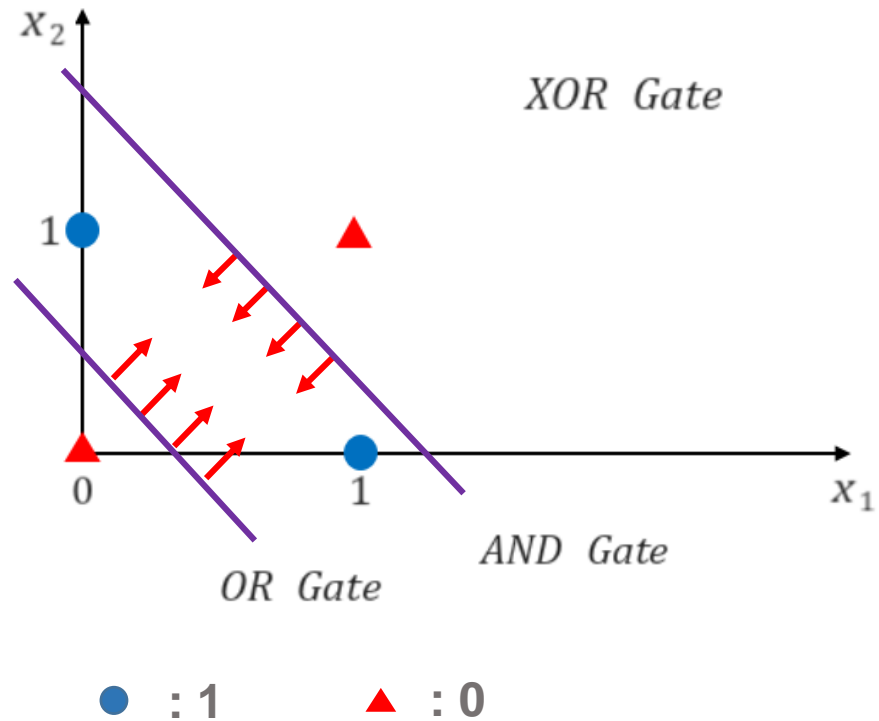


XOR Gate

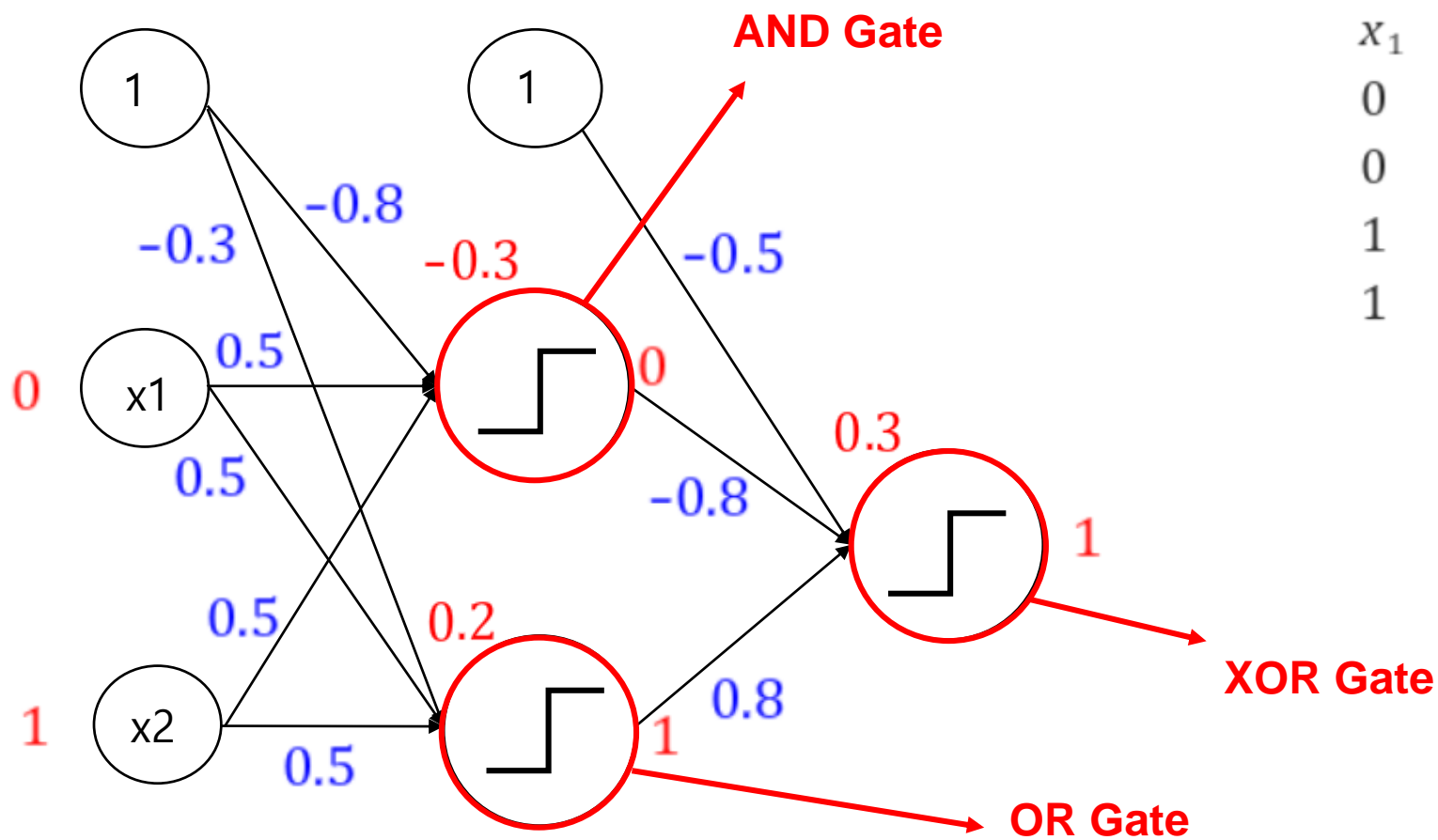
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

Multi Layer Perceptron

But if we use 2 single layer perceptron, we can solve XOR problem. This model is called multi layer perceptron.



Multi Layer Perceptron

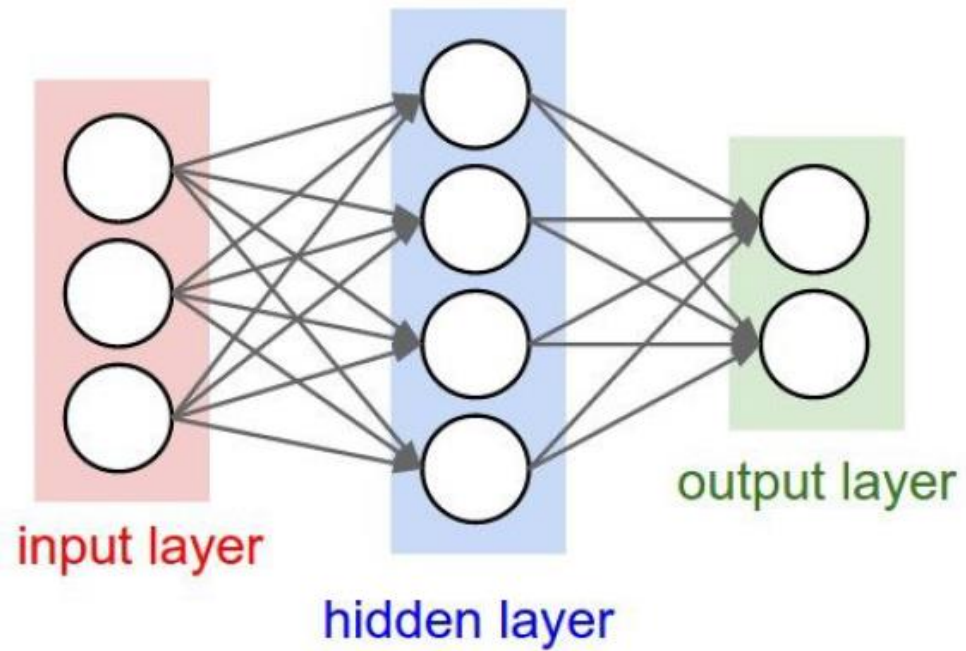


OR Gate		
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

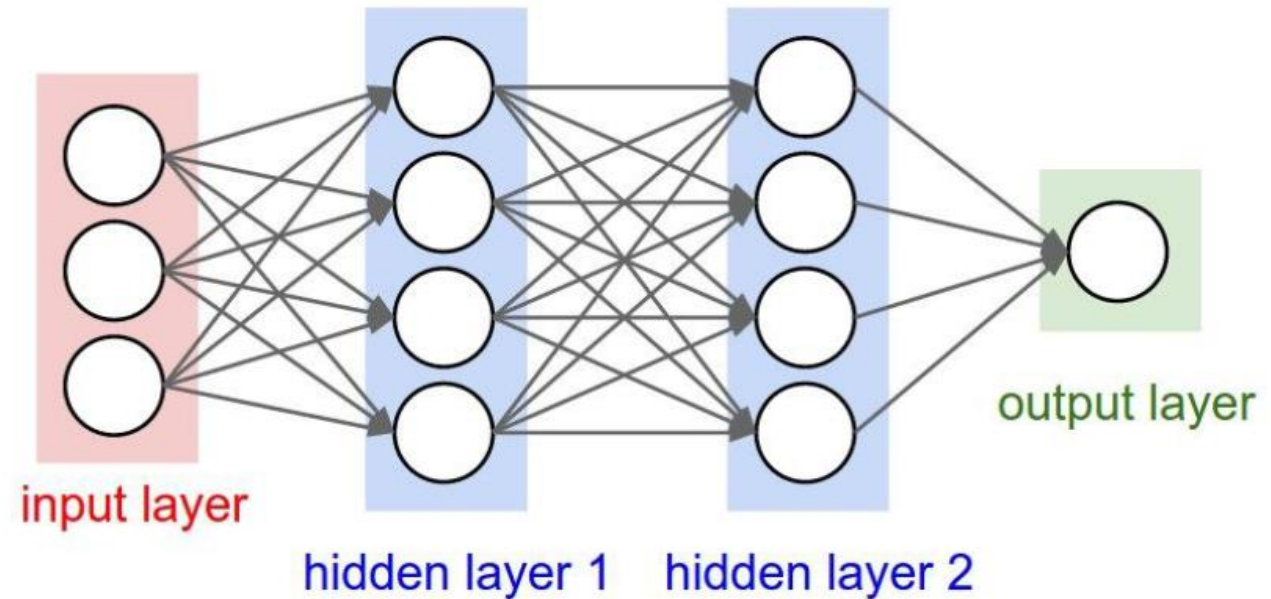
AND Gate		
x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

XOR Gate		
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

Neural Network Architecture

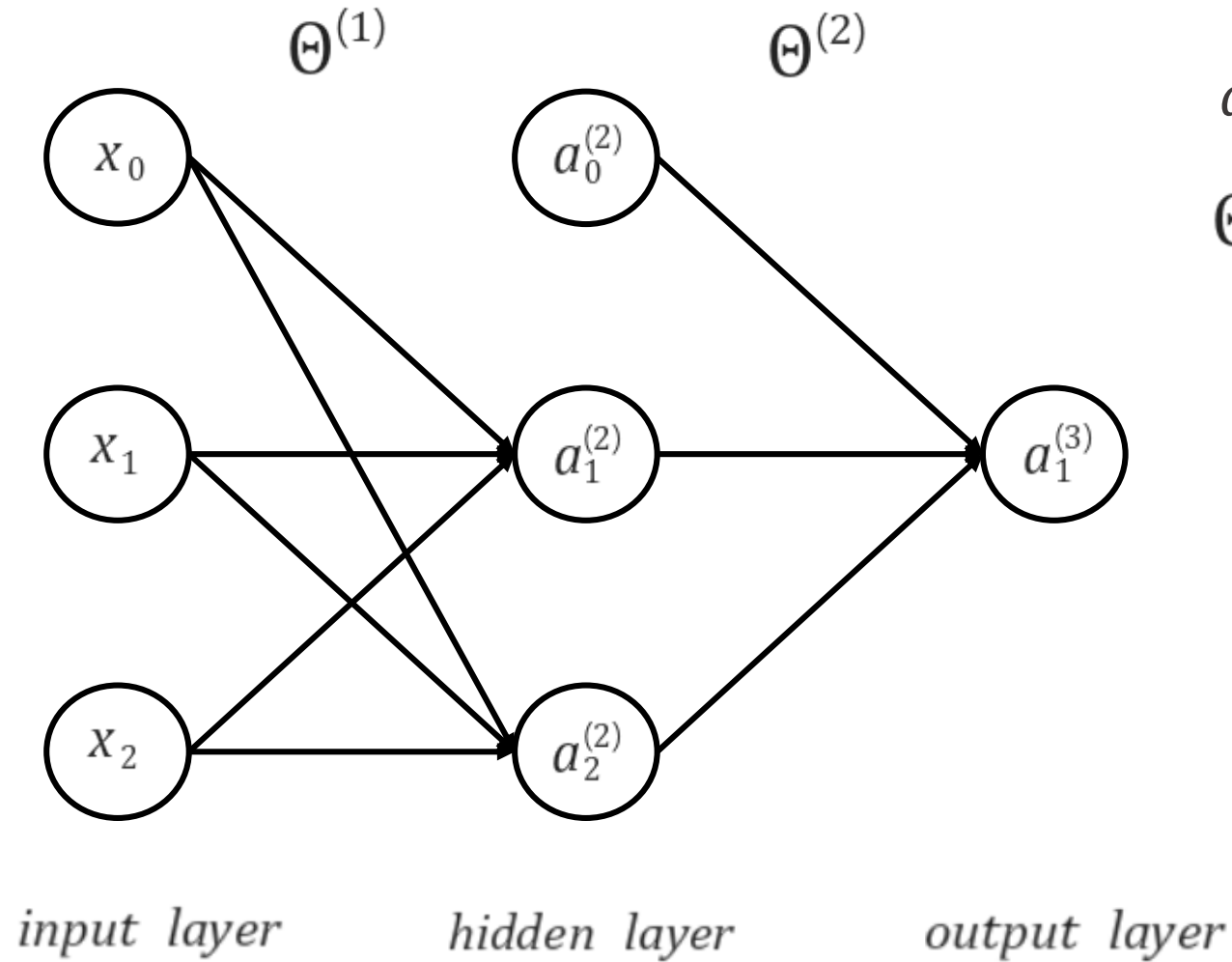


"2-layer Neural Network"
or
"1-hidden-layer Neural Network"



"3-layer Neural Network"
or
"2-hidden-layer Neural Network"

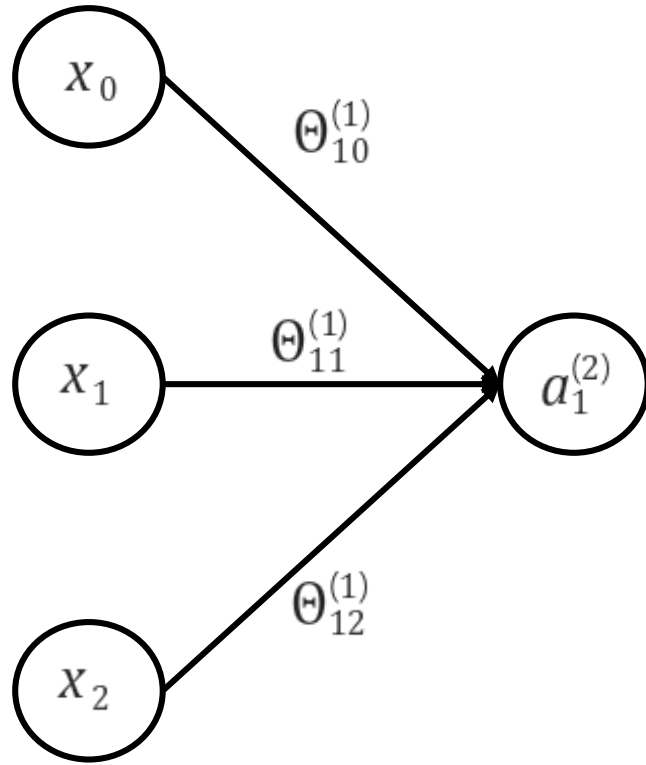
Forward Propagation



a_i^j : “activation” of i-th unit in j-th layer

$\Theta^{(j)}$: matrix of weights controlling function mapping from j-th layer to (j+1)-th layer

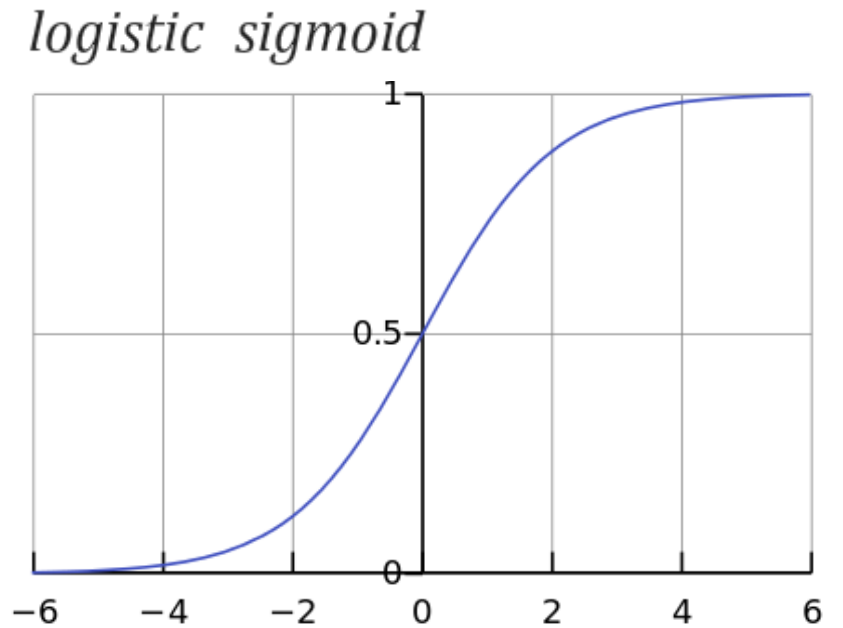
Forward Propagation



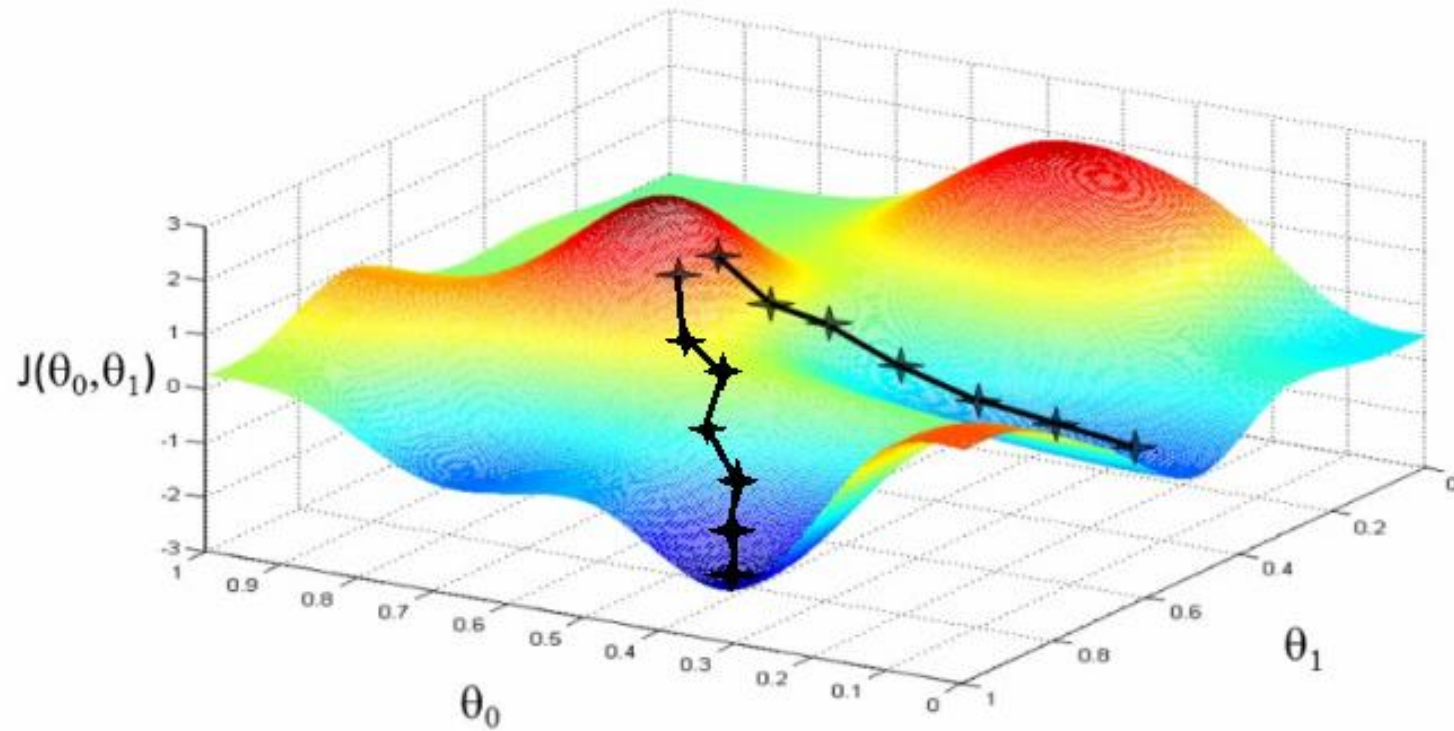
$$z_1^{(2)} = \Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2$$

$$a_1^{(2)} = g(z_1^{(2)})$$

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



Gradient Descent



compute $\frac{\partial J}{\partial \theta_0}$ and $\frac{\partial J}{\partial \theta_1}$

update weights with

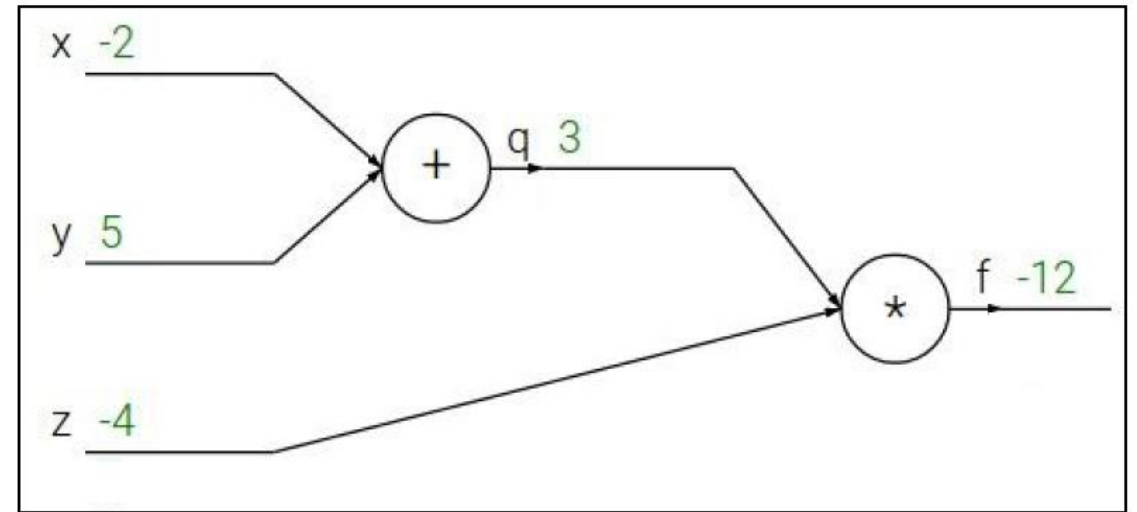
$$\begin{aligned}\theta_0 &:= \theta_0 - \alpha \cdot \frac{\partial J}{\partial \theta_0} \\ \theta_1 &:= \theta_1 - \alpha \cdot \frac{\partial J}{\partial \theta_1}\end{aligned}$$

Gradient “descent” optimization with learning rate ‘alpha’ (e.g. 0.01)

Computational Graph

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$



Computational Graph

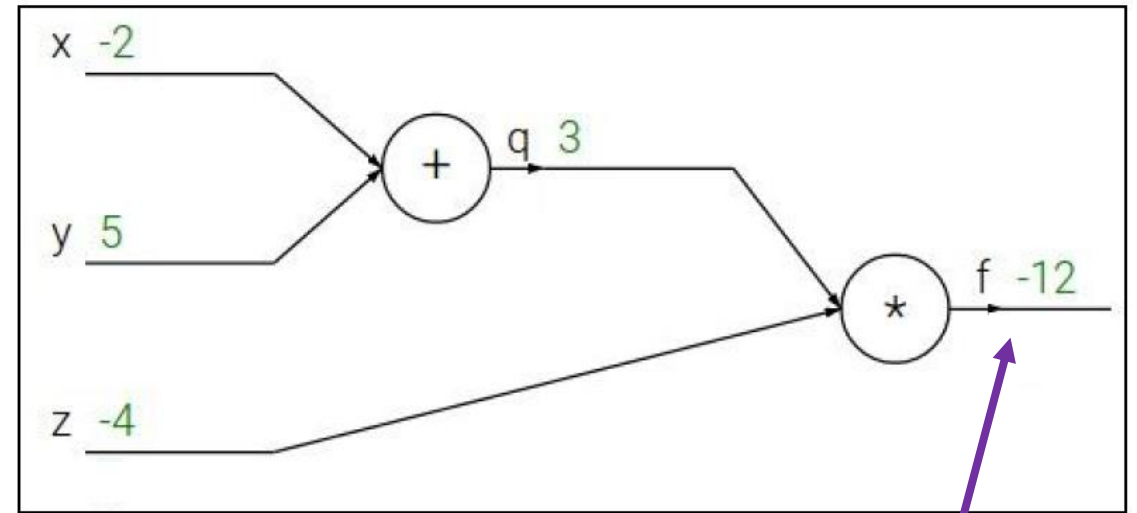
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial f}$$

Computational Graph

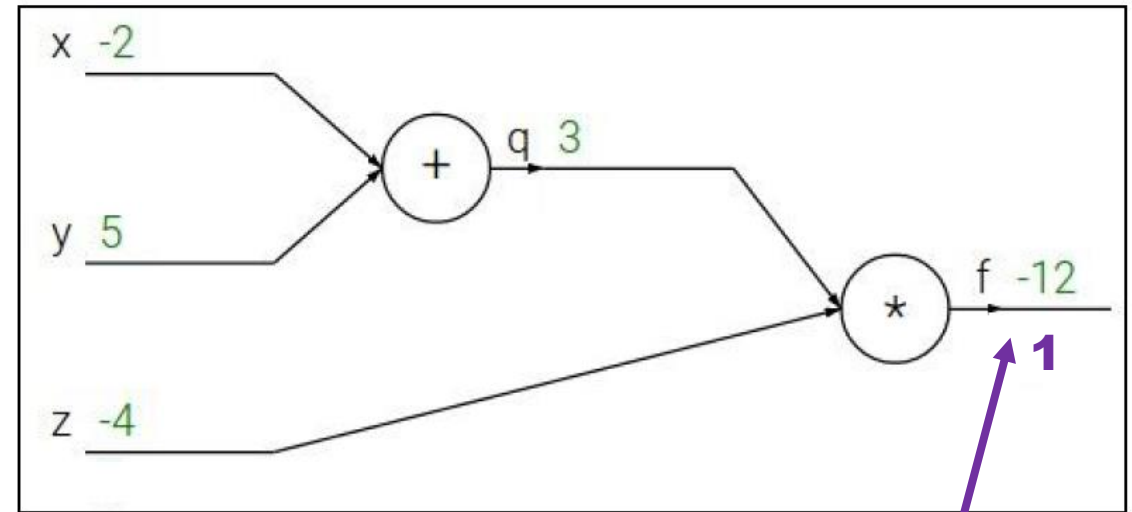
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial f}$$

Backpropagation using Chain Rule

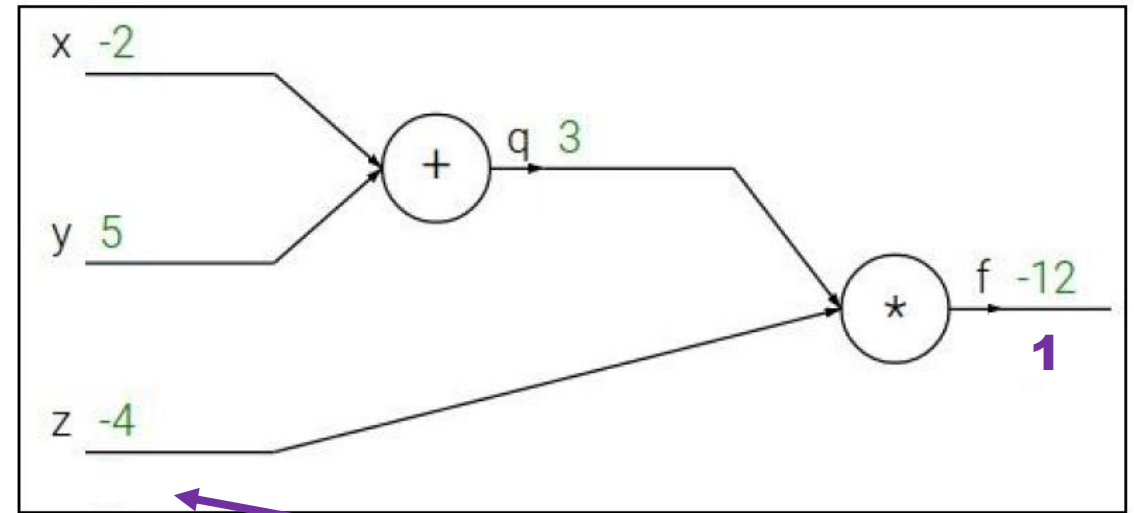
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial z}$$

A purple arrow points from this box to the input z of the multiplication node in the computational graph above.

Backpropagation using Chain Rule

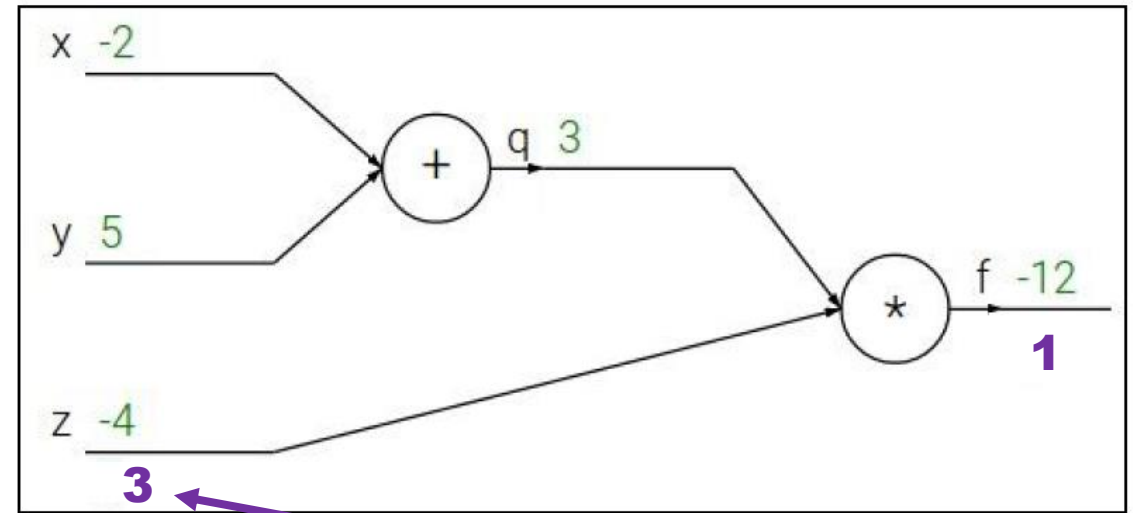
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial z}$$

Computational Graph

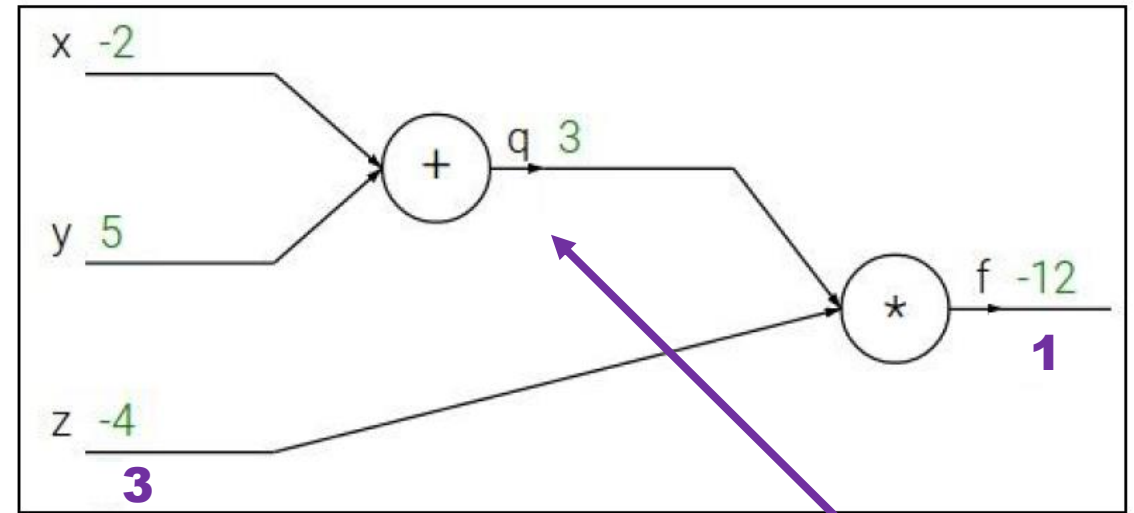
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial q}$$

Backpropagation using Chain Rule

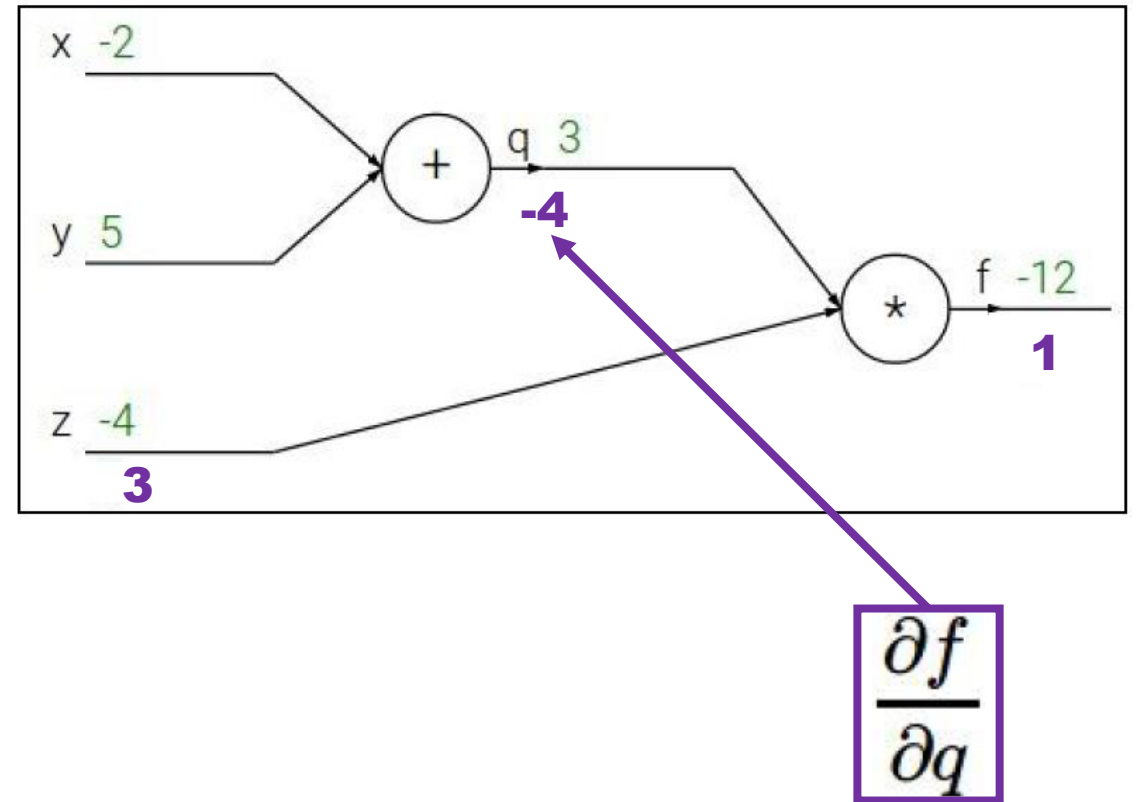
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



Backpropagation using Chain Rule

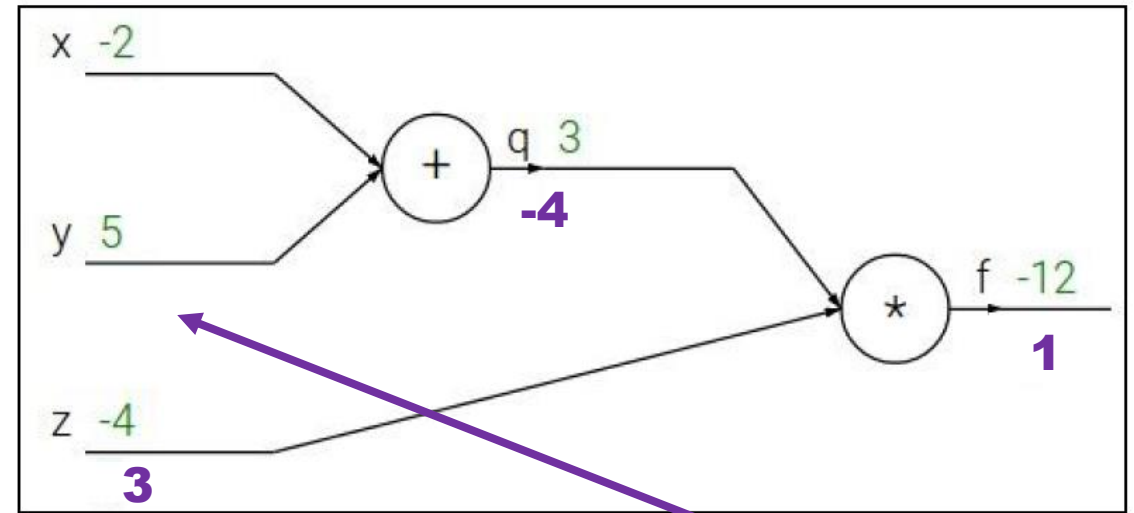
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial y}$$

Backpropagation using Chain Rule

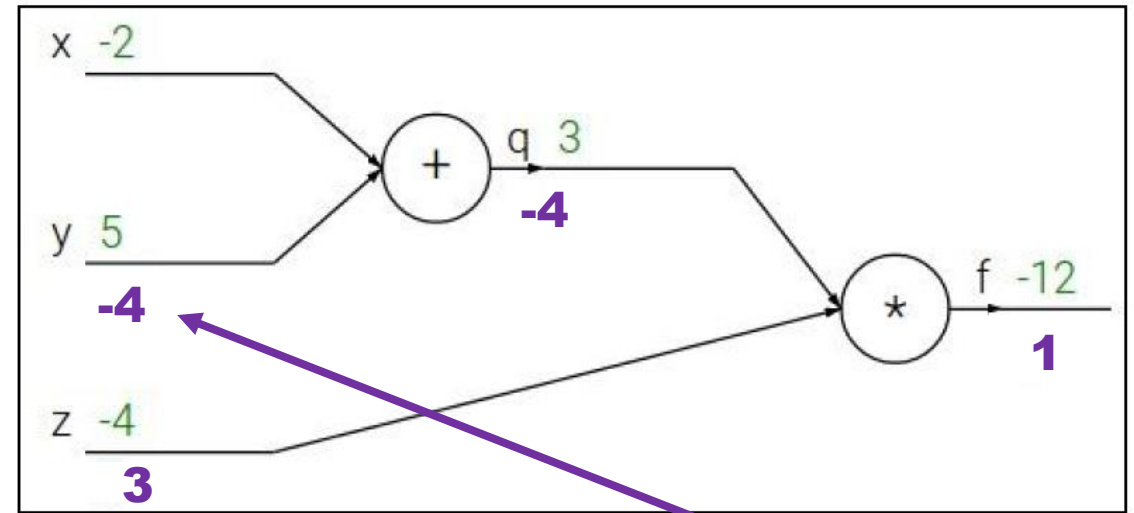
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



Chain rule:

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial y}$$

$$\frac{\partial f}{\partial y}$$

Backpropagation using Chain Rule

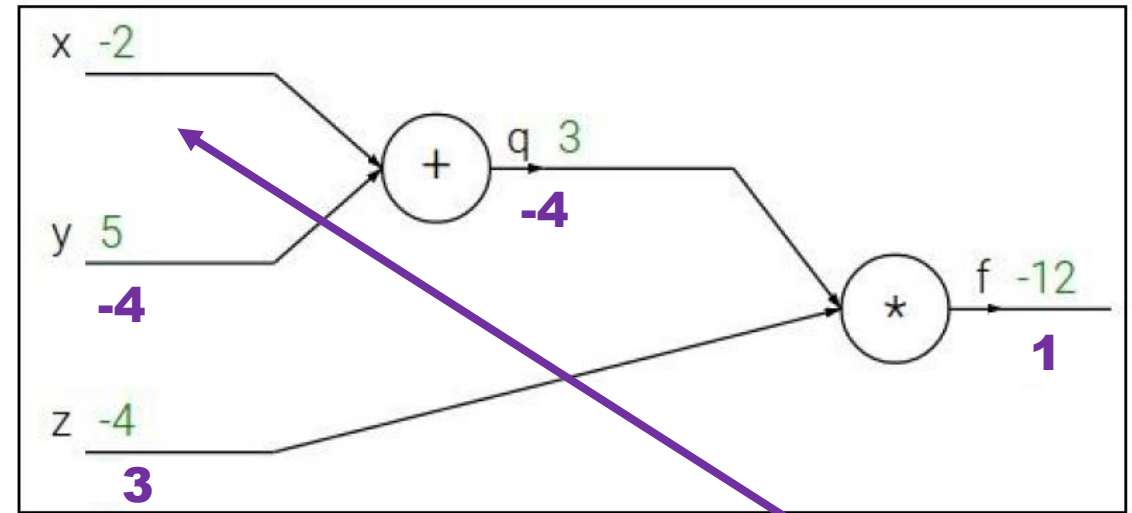
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$\frac{\partial f}{\partial x}$$

Backpropagation using Chain Rule

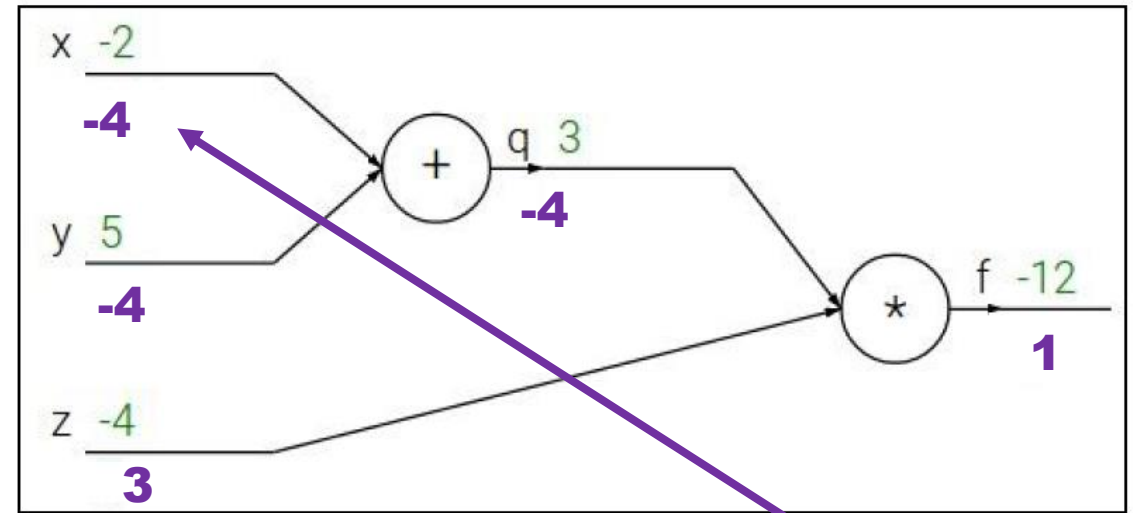
$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



Chain rule:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial x}$$

$$\frac{\partial f}{\partial x}$$