

Introduction to Deep Learning

Recitation 1a
Spring 2021

Intelligence

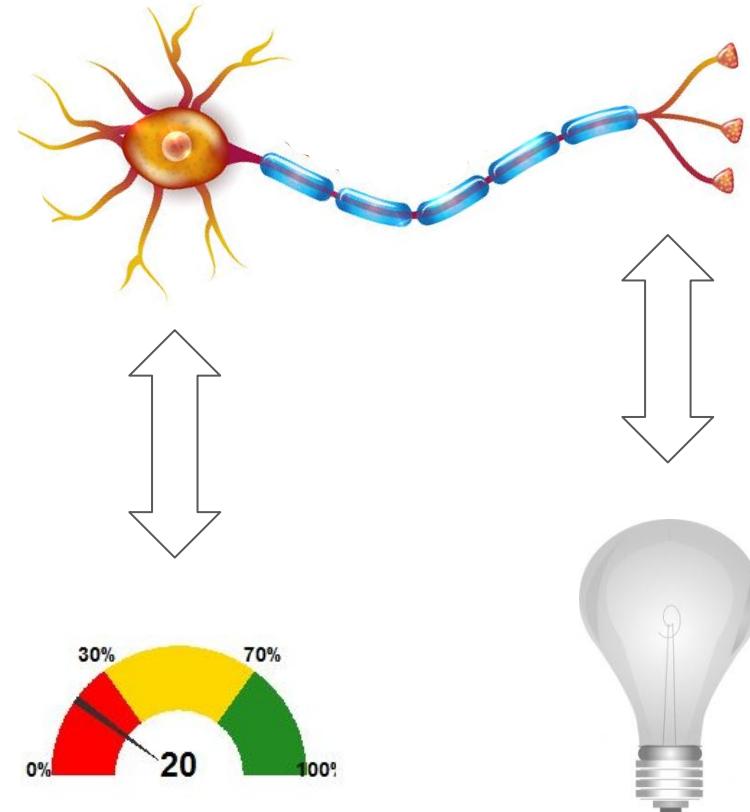
Tasks that are easy for humans can be difficult for machines.

Likewise, tasks that are easy for machines can be difficult for humans.



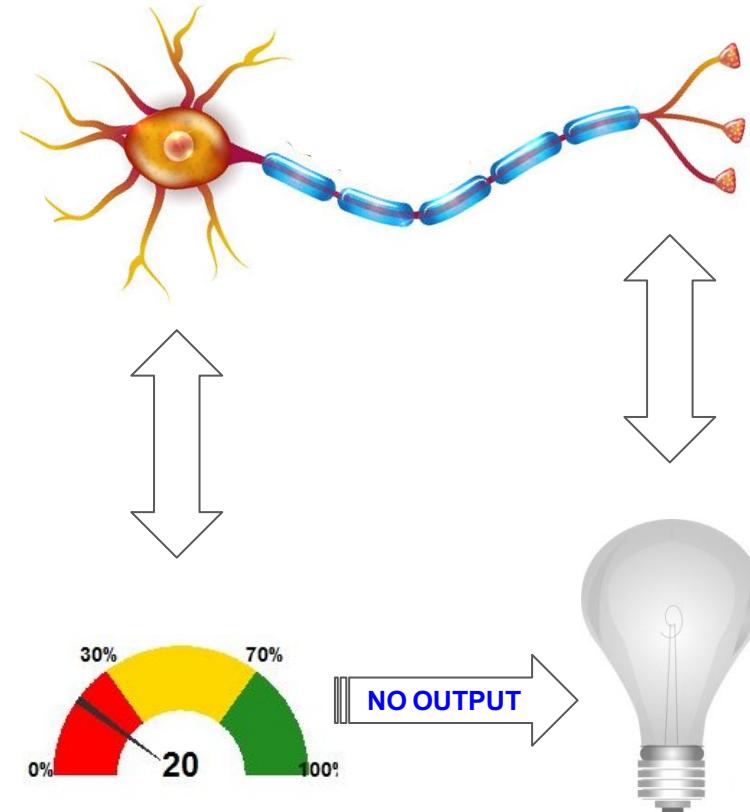
Neurons and Thresholds

Biological brains, made of connected neurons, are the inspiration for artificial neural networks.



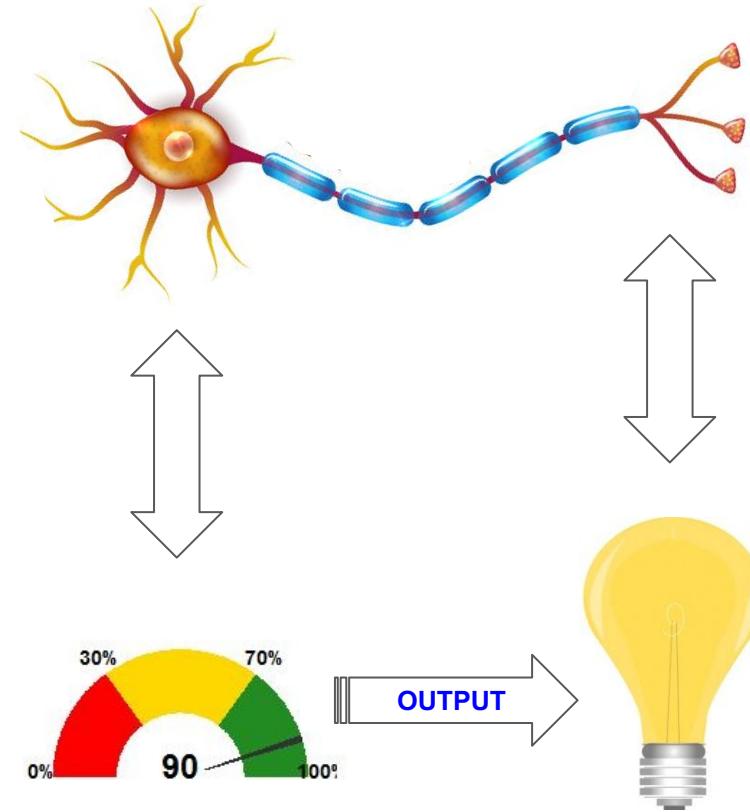
Neurons and Thresholds

Biological brains, made of connected neurons, are the inspiration for artificial neural networks.



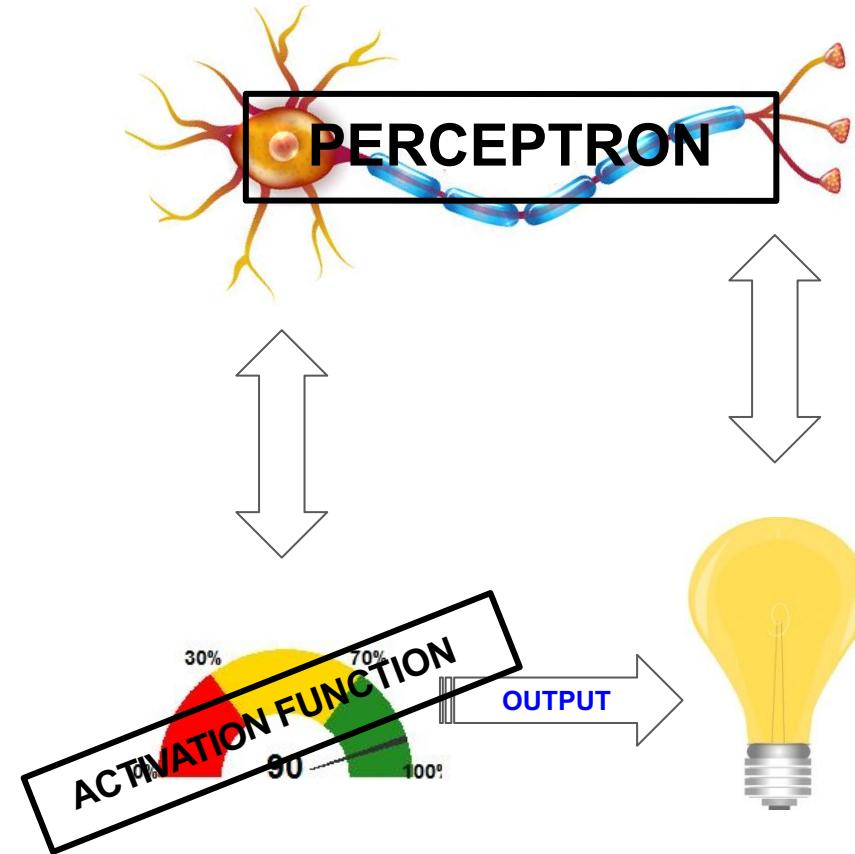
Neurons and Thresholds

Biological brains, made of connected neurons, are the inspiration for artificial neural networks.



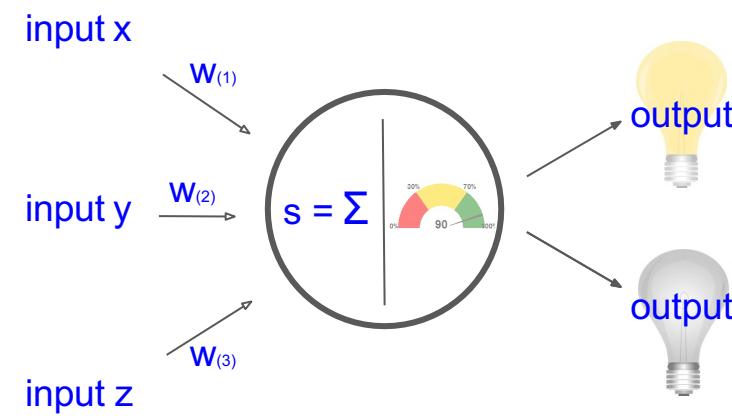
Neurons and Thresholds

Biological brains, made of connected neurons, are the inspiration for artificial neural networks.



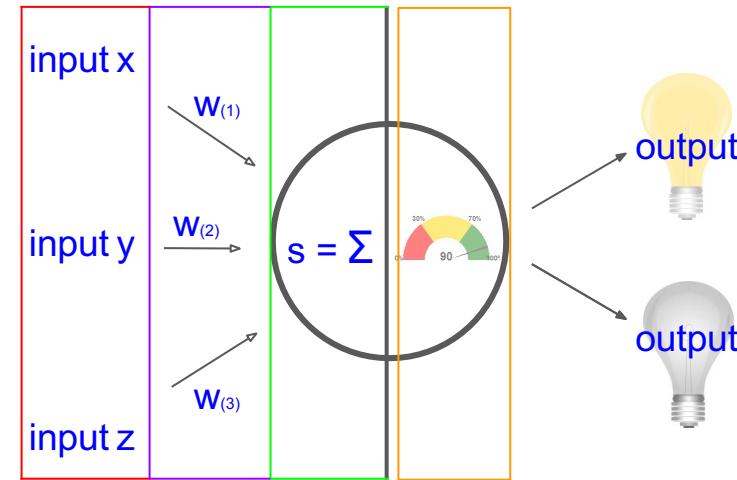
Perceptrons

- ... are single layer neural networks



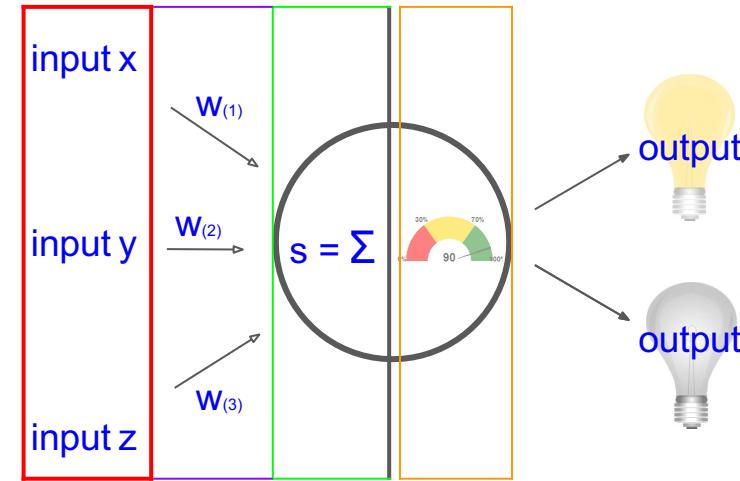
Perceptrons

- ... are single layer neural networks
- ... have four parts



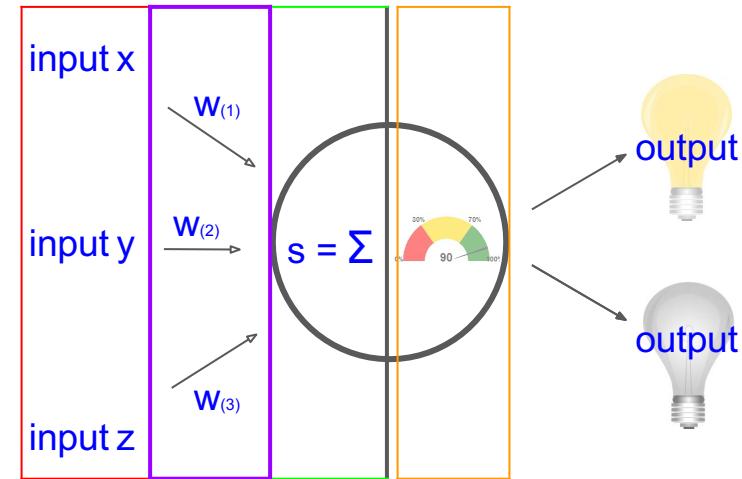
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs



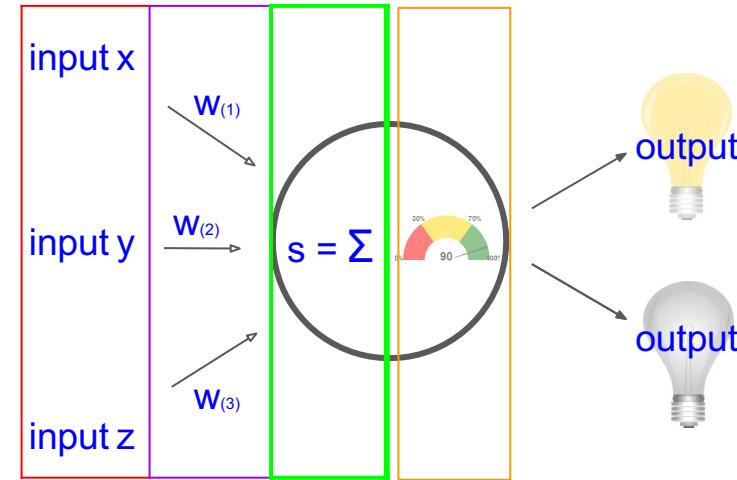
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights



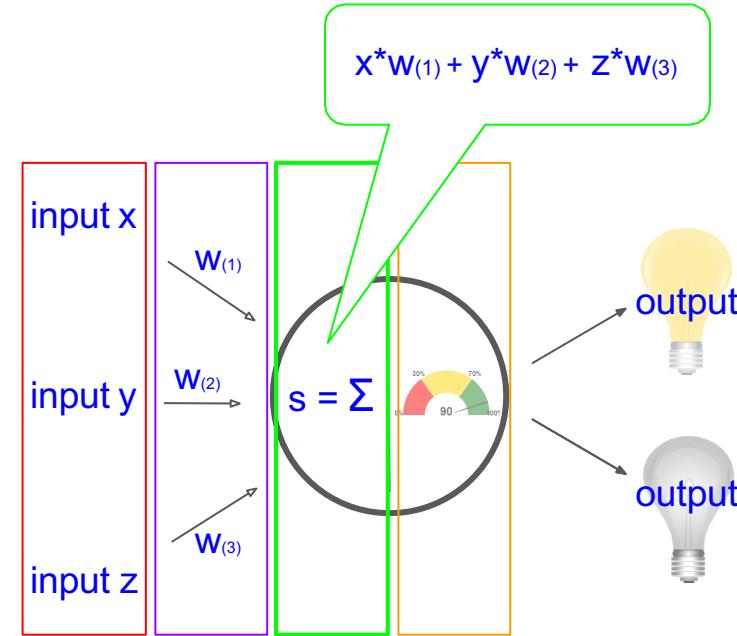
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum



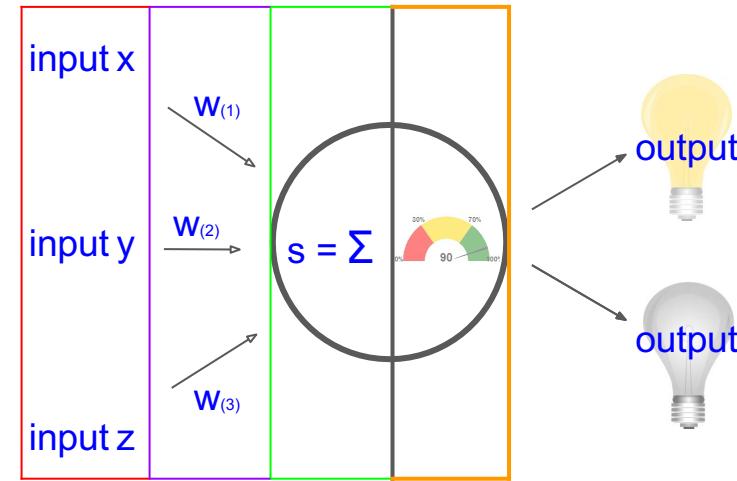
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum



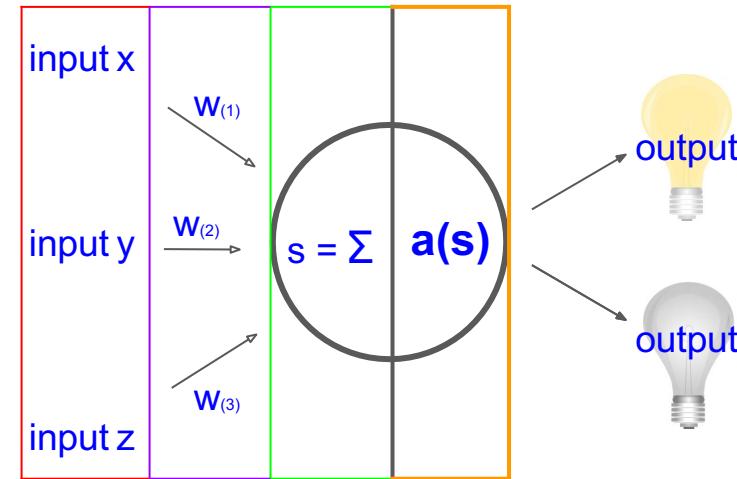
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum
 - Threshold



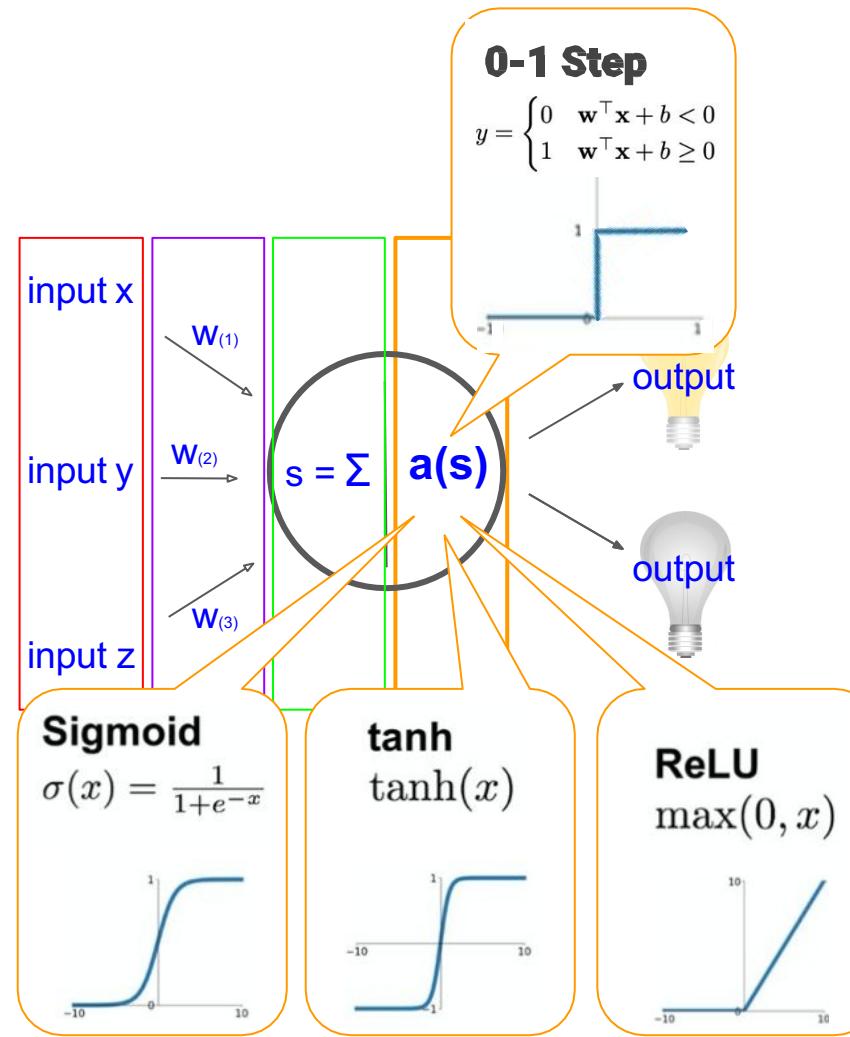
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum
 - Threshold Activation Functions



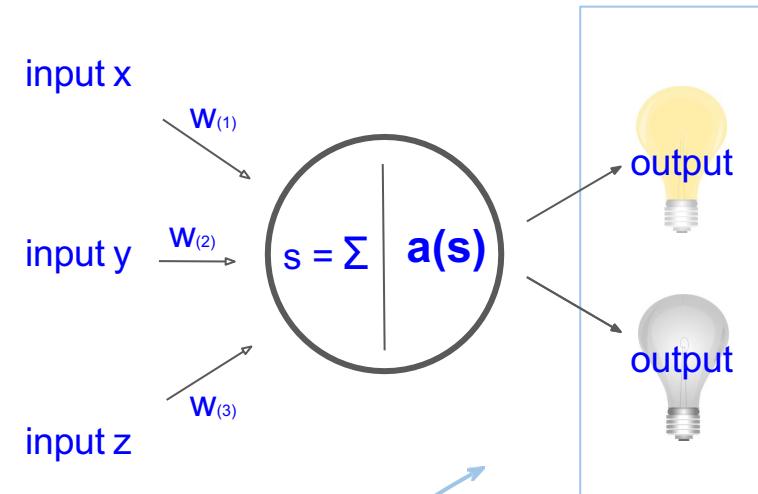
Perceptrons

- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum
 - Threshold Activation Functions

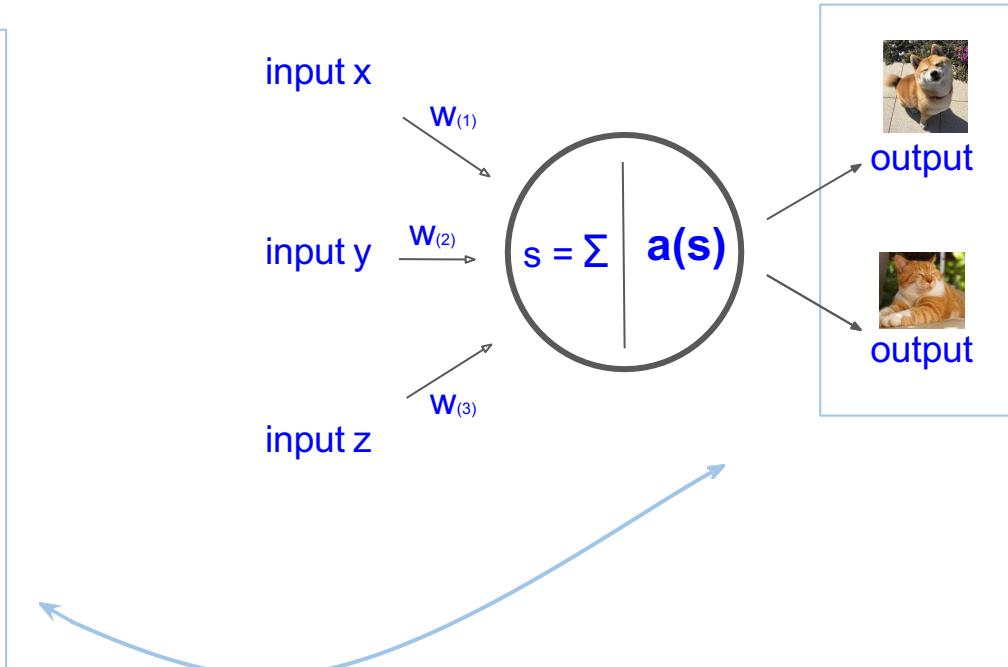
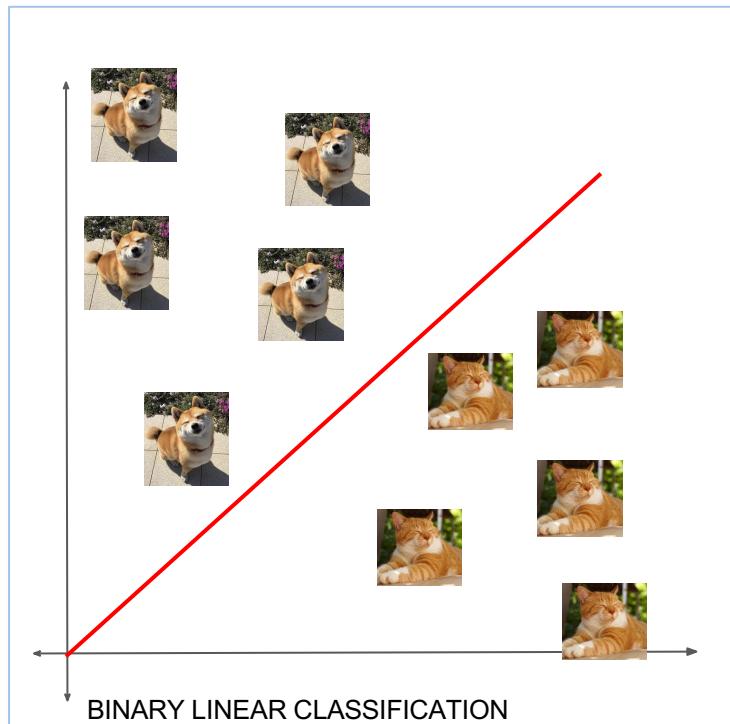


Perceptrons

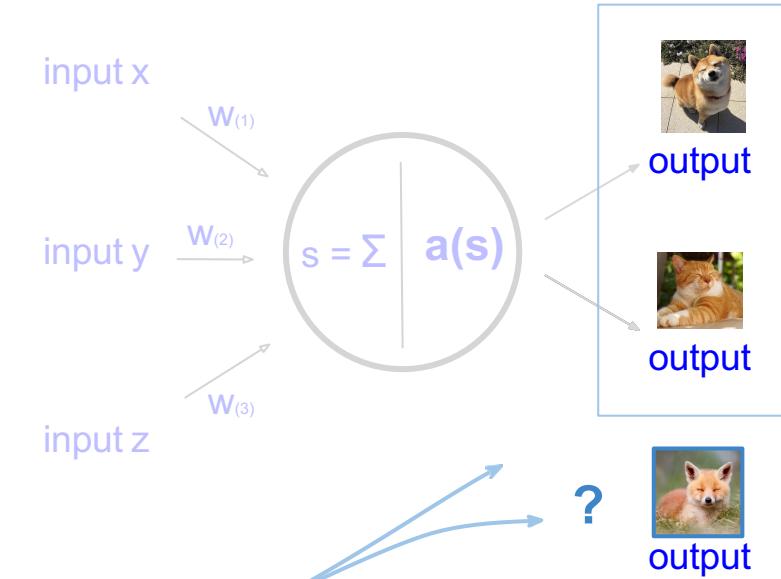
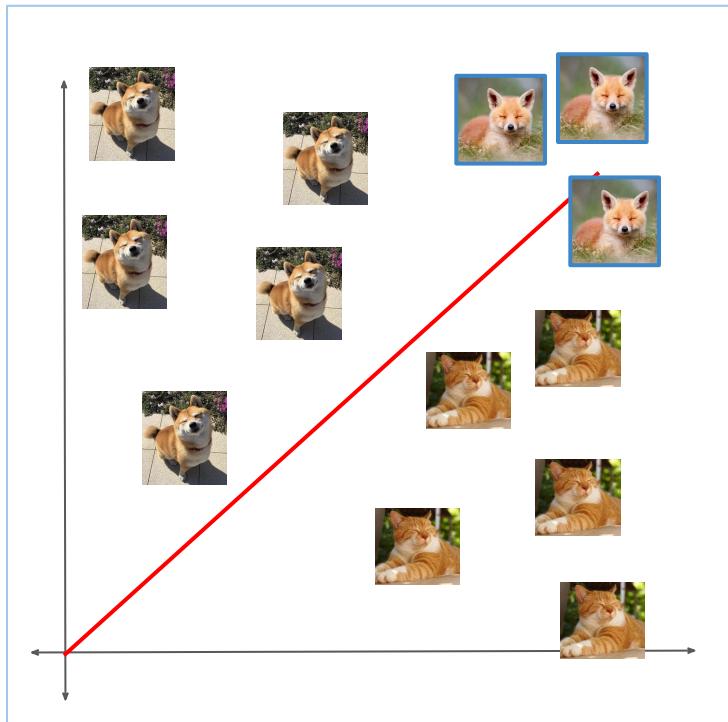
- ... are single layer neural networks
- ... have four parts
 - Inputs
 - Weights
 - Weighted Sum
 - Threshold Activation Functions
- ... are binary linear classifiers



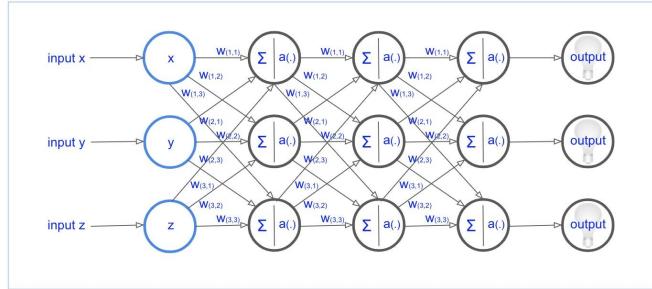
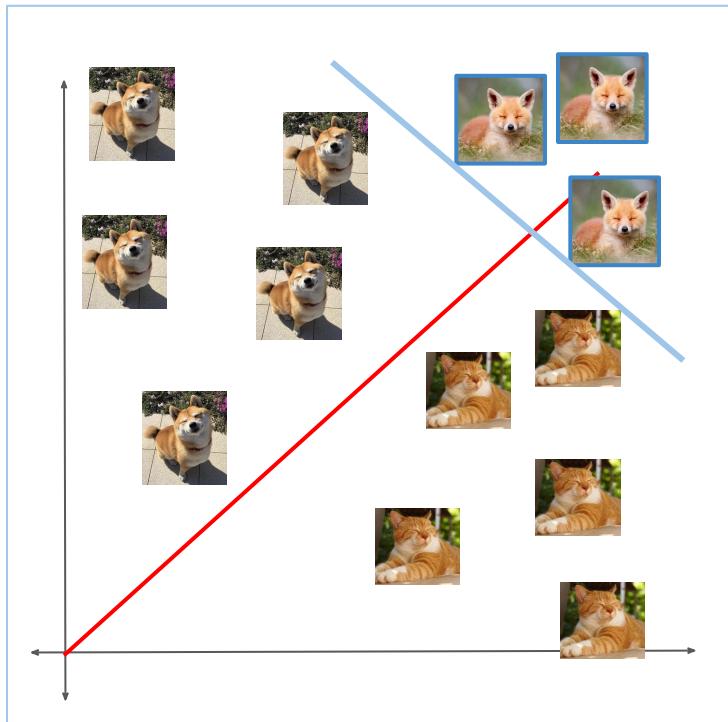
Perceptrons



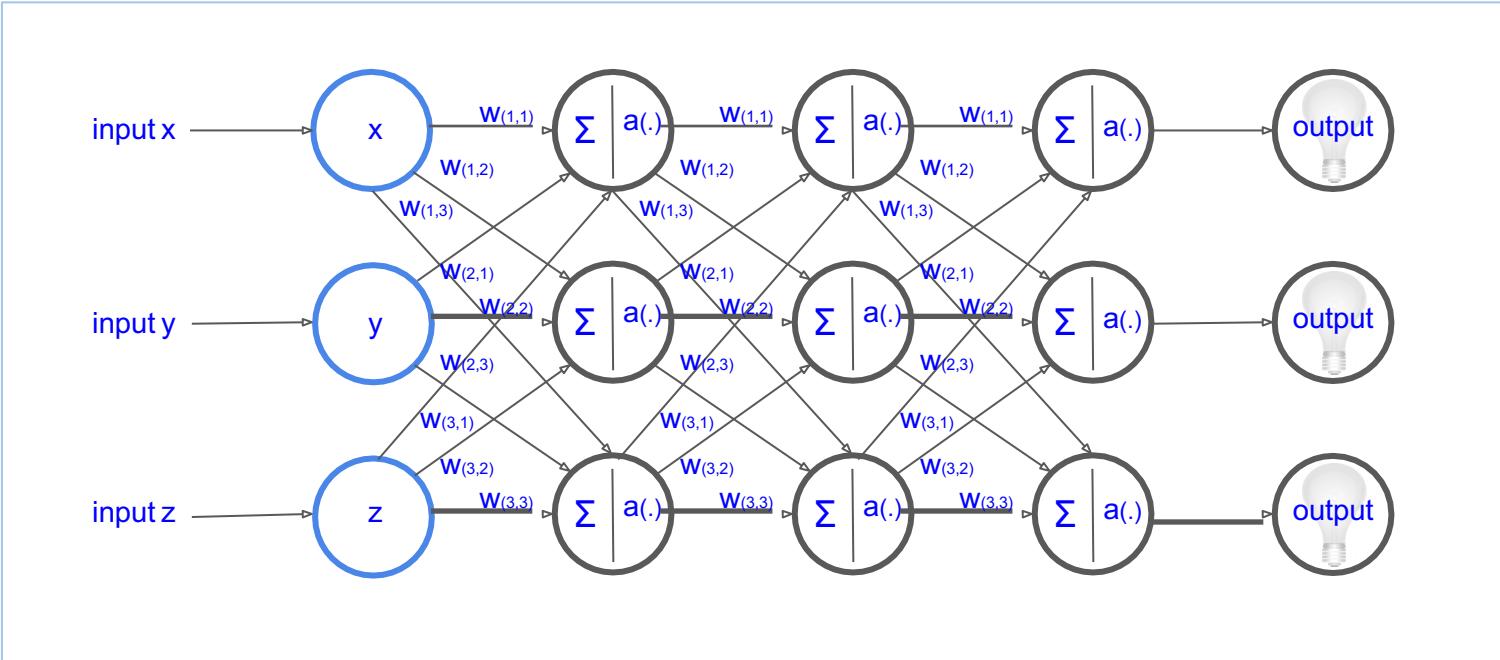
Multi-Layer Perceptrons



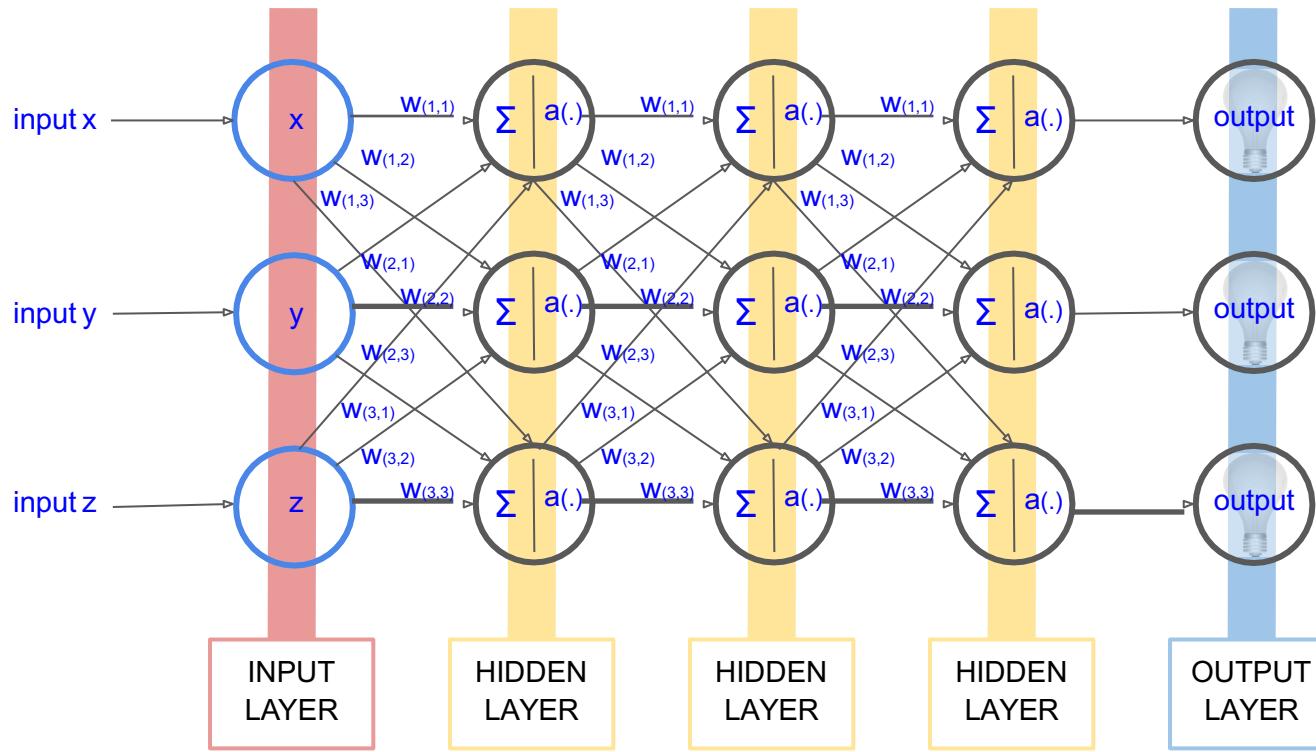
Multi-Layer Perceptrons



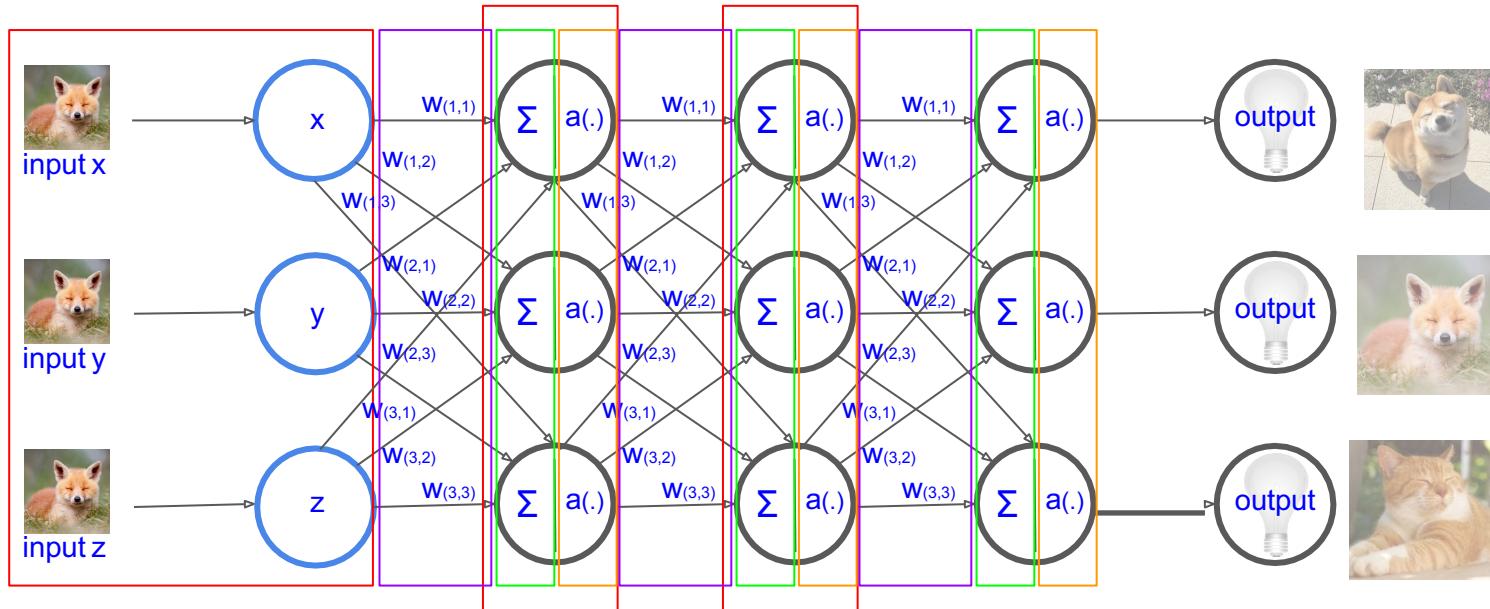
Multi-Layer Perceptrons



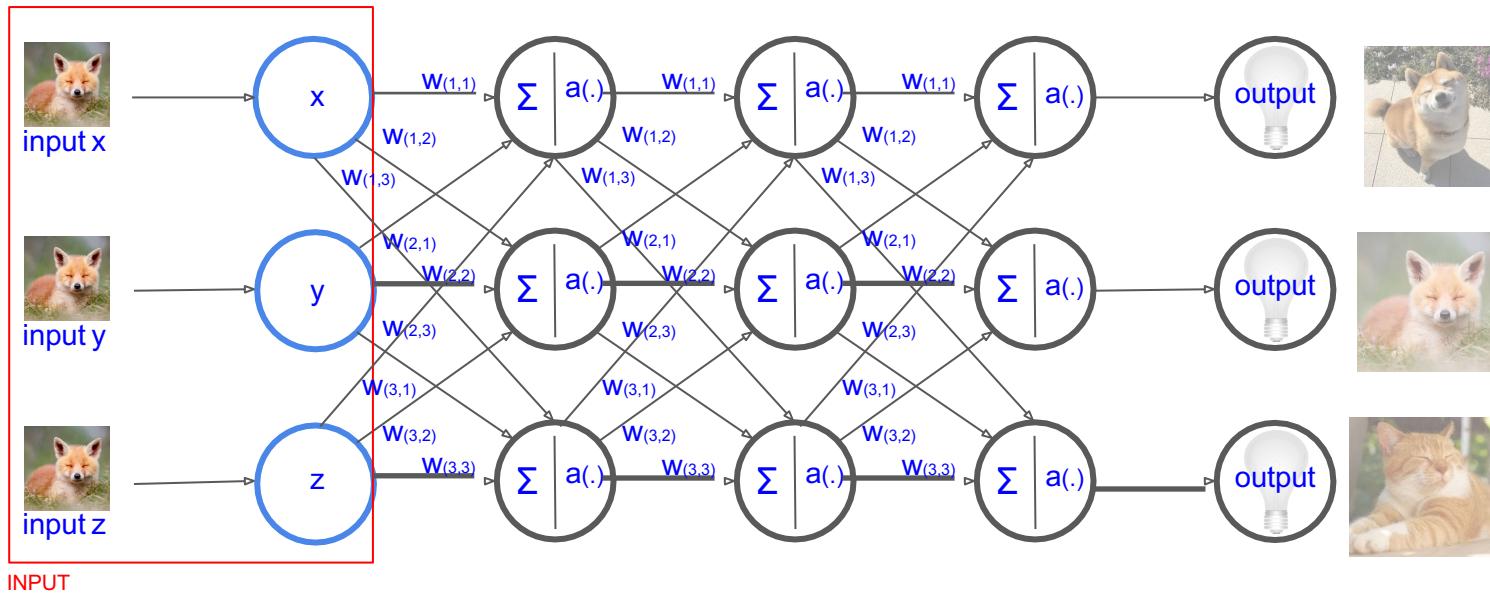
Multi-Layer Perceptrons



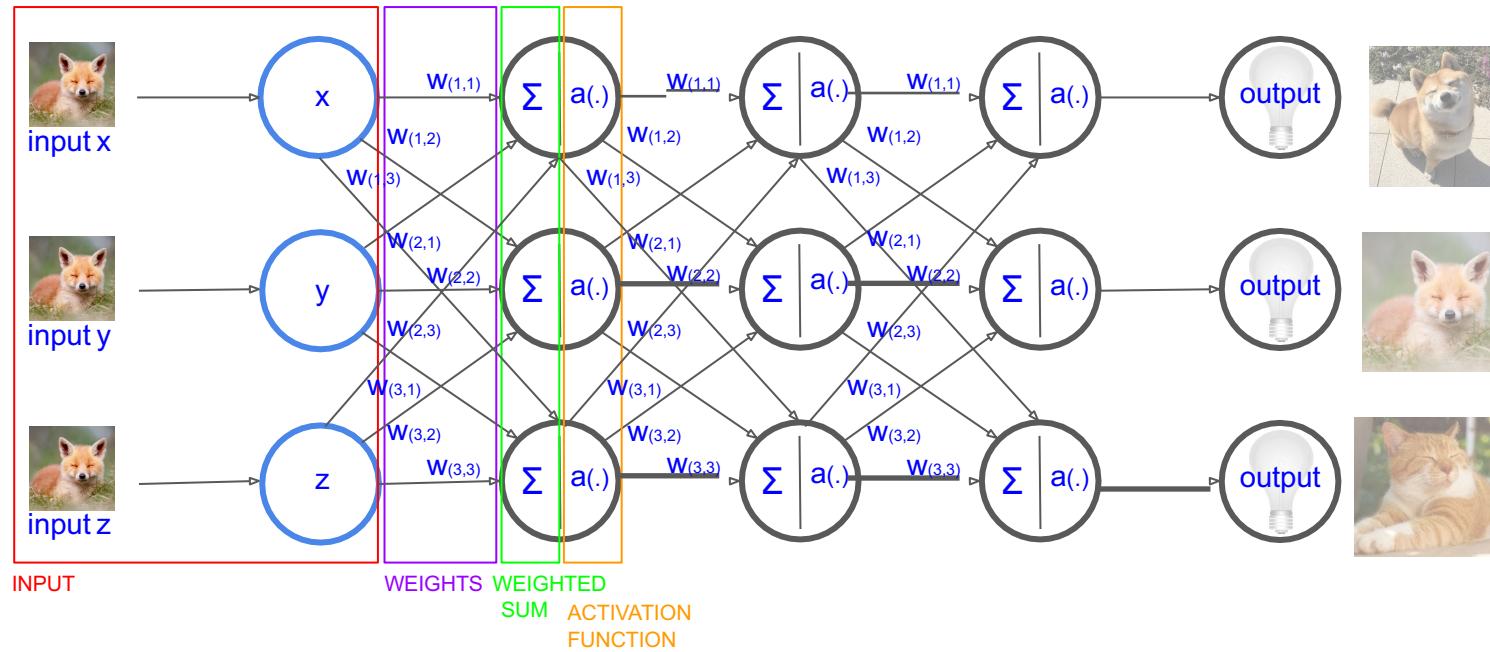
Forward Signal



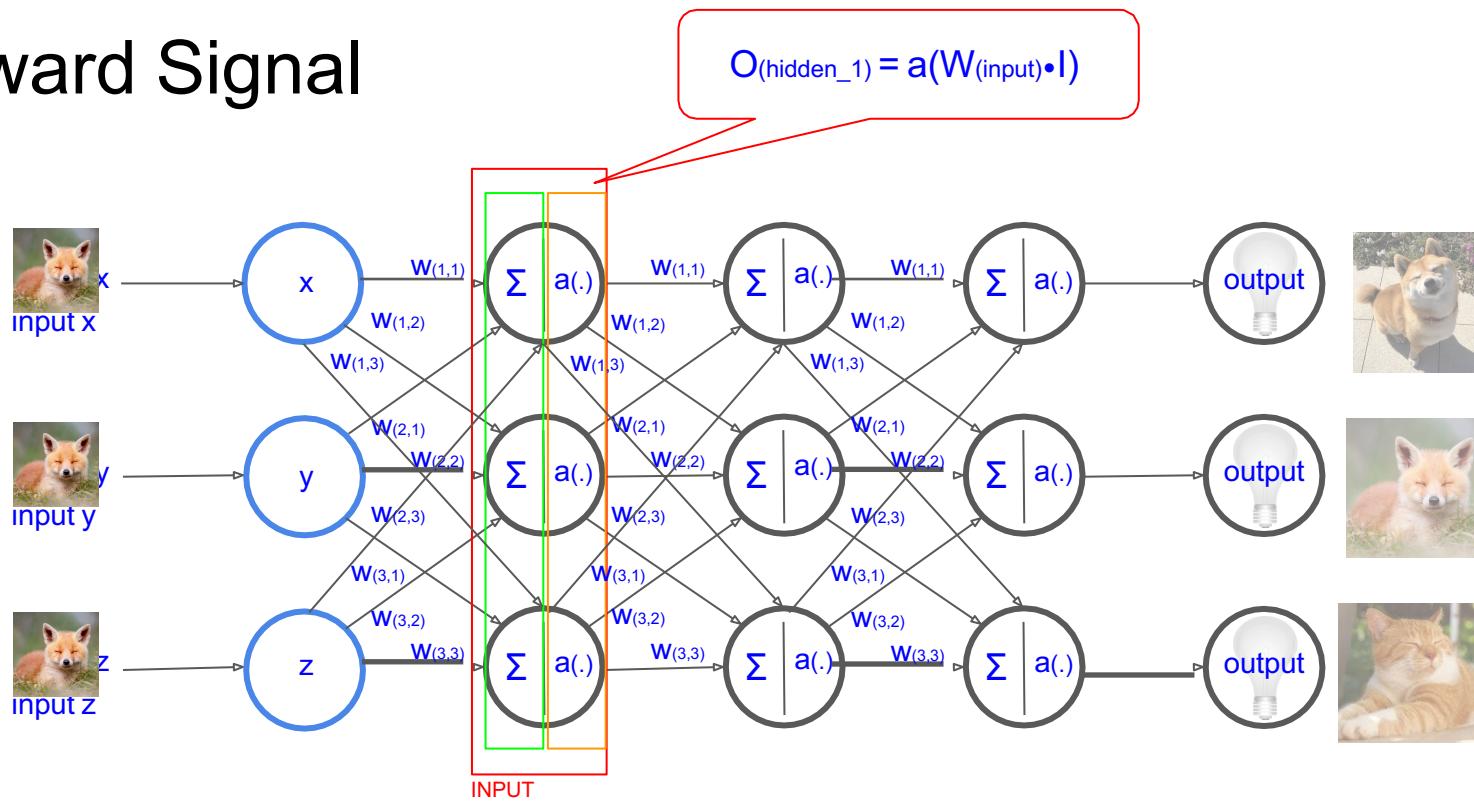
Forward Signal



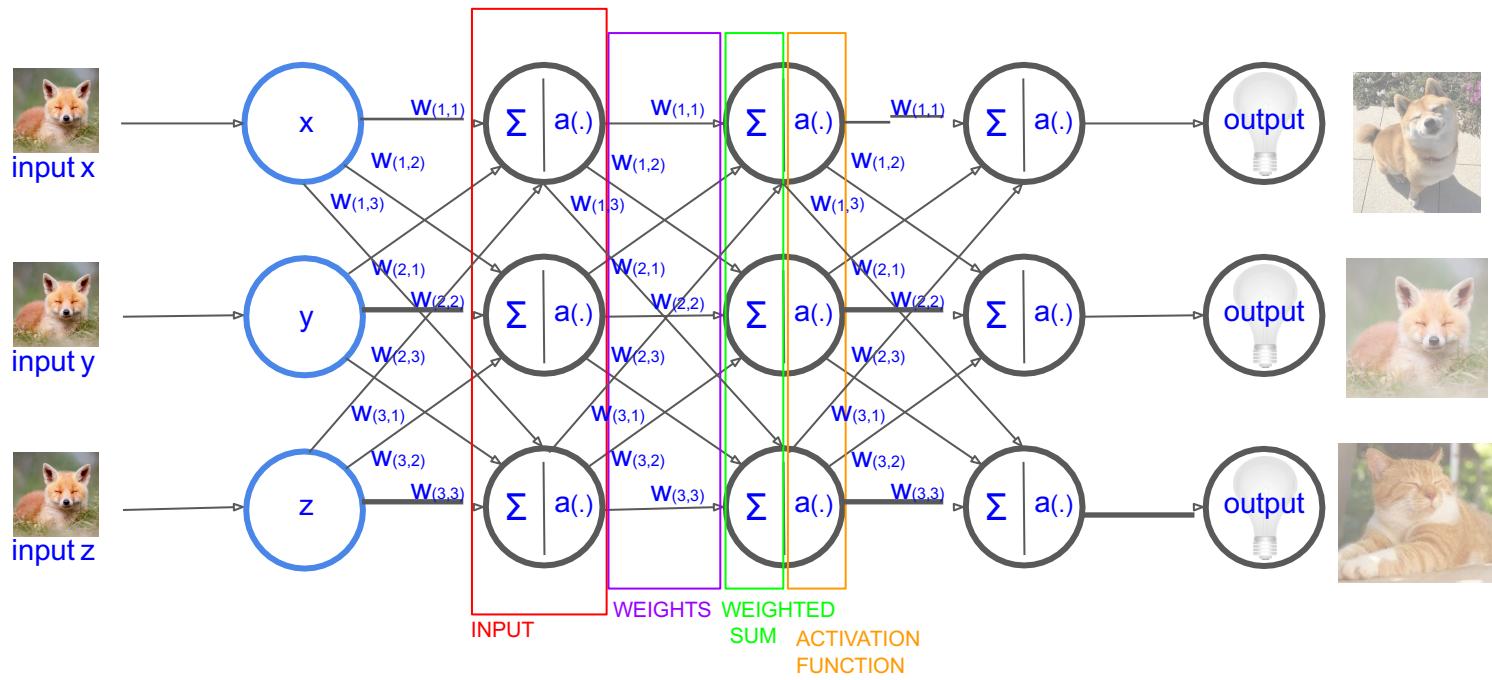
Forward Signal



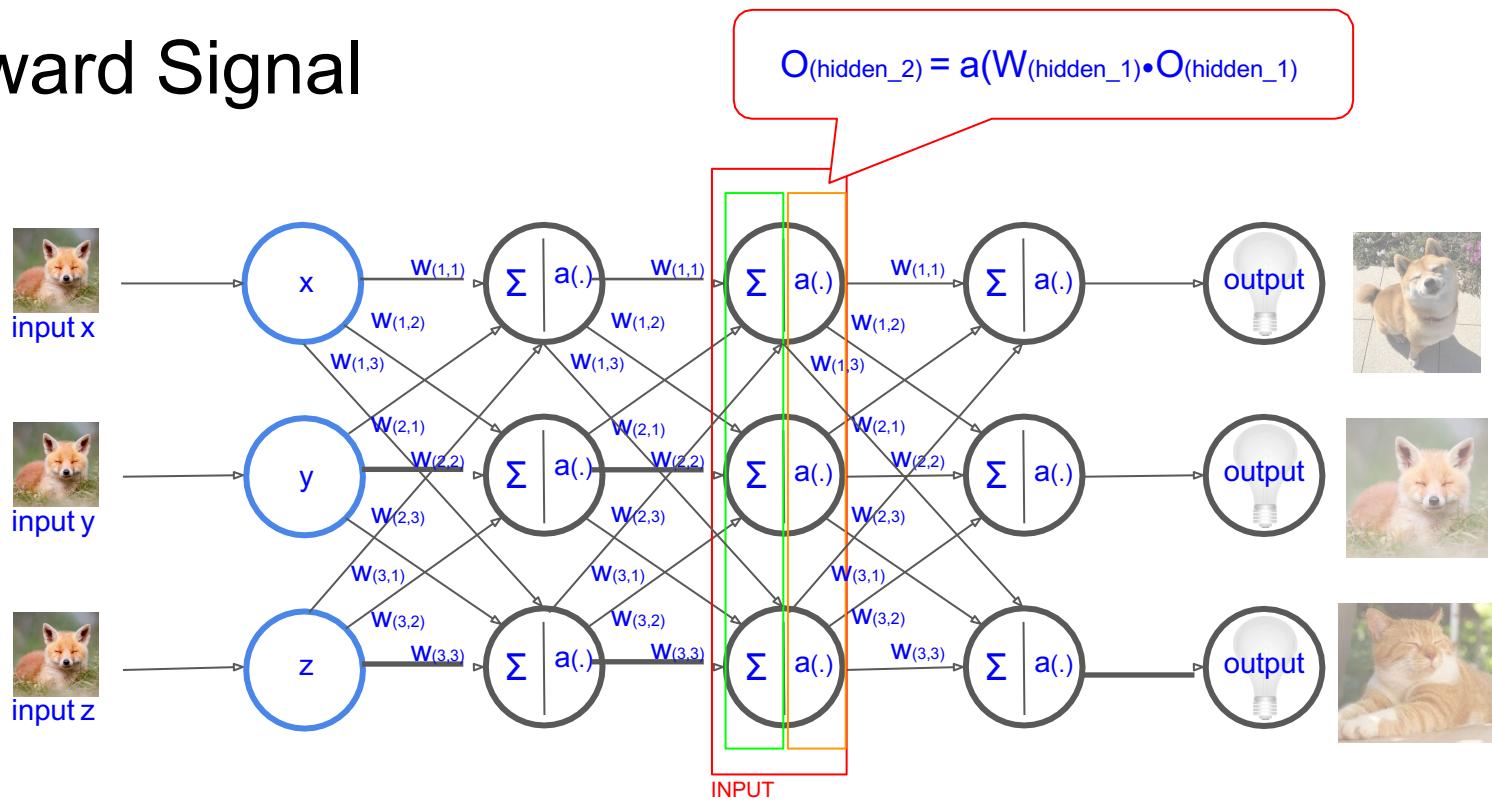
Forward Signal



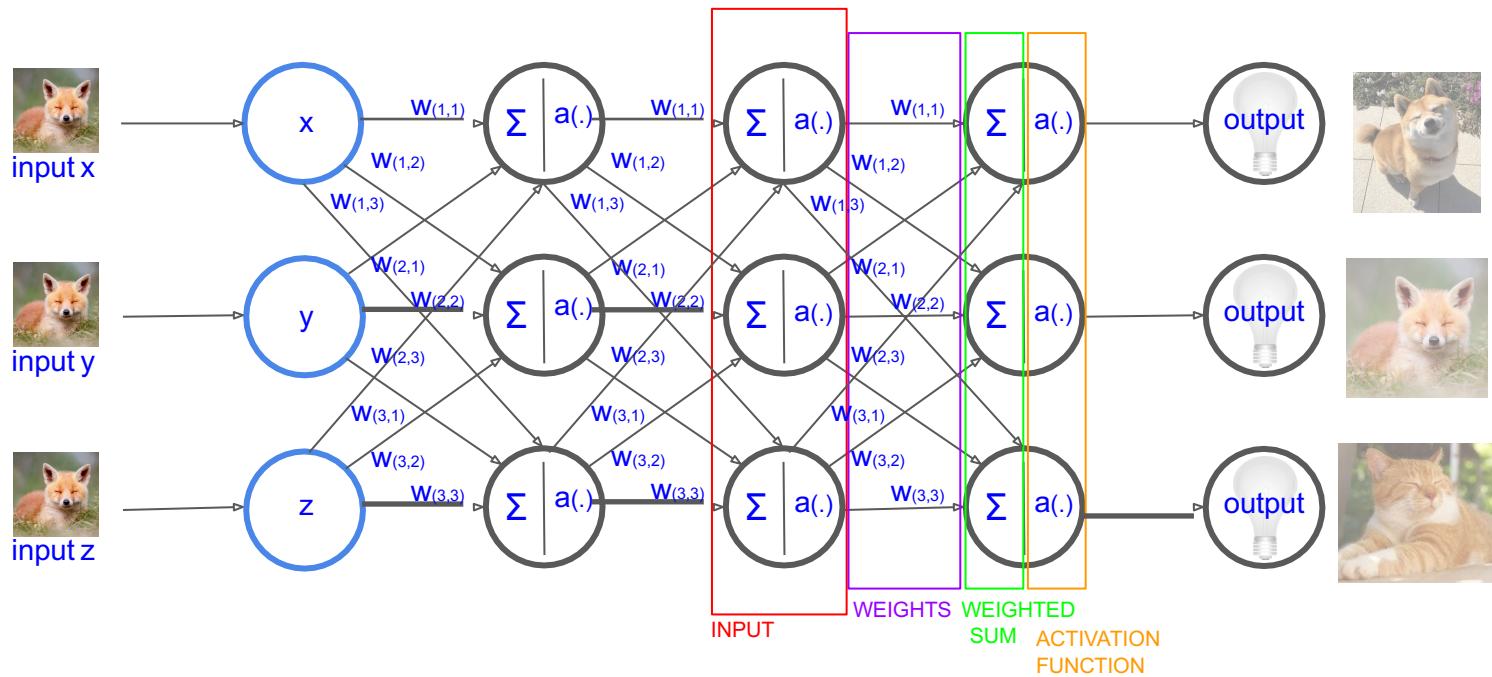
Forward Signal



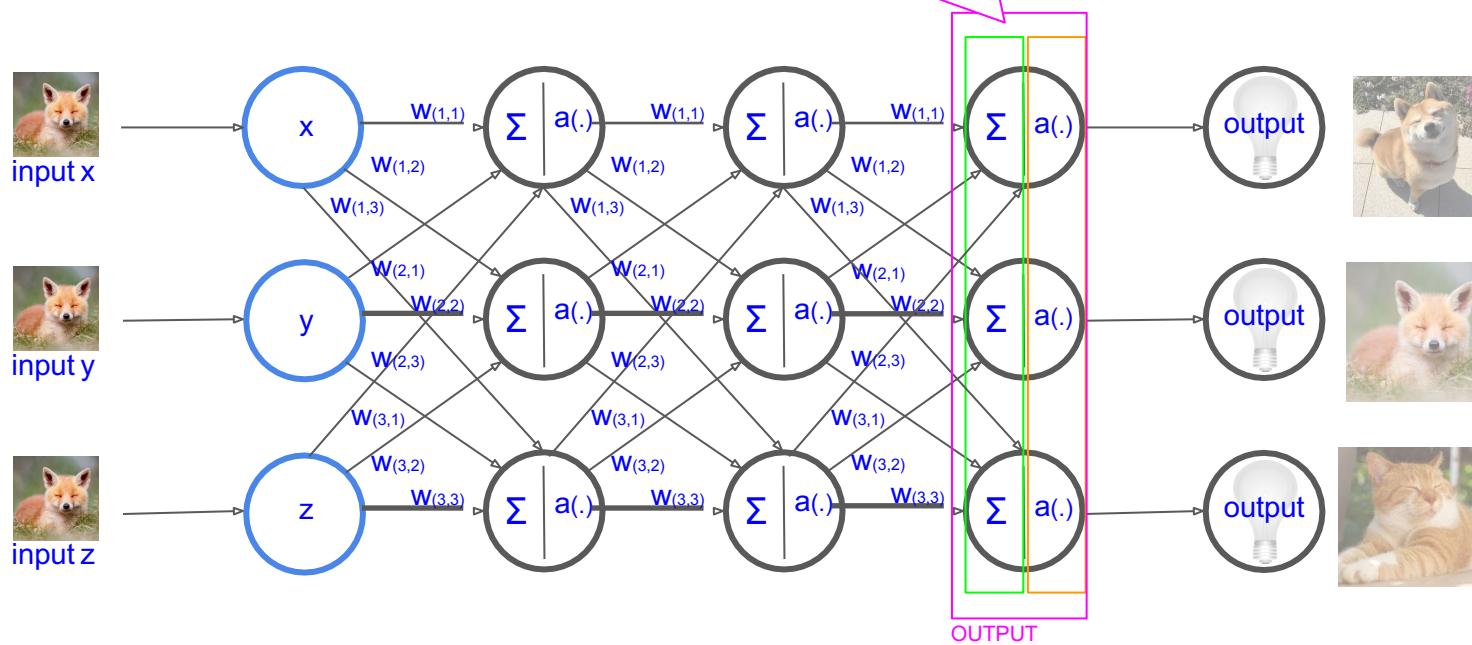
Forward Signal



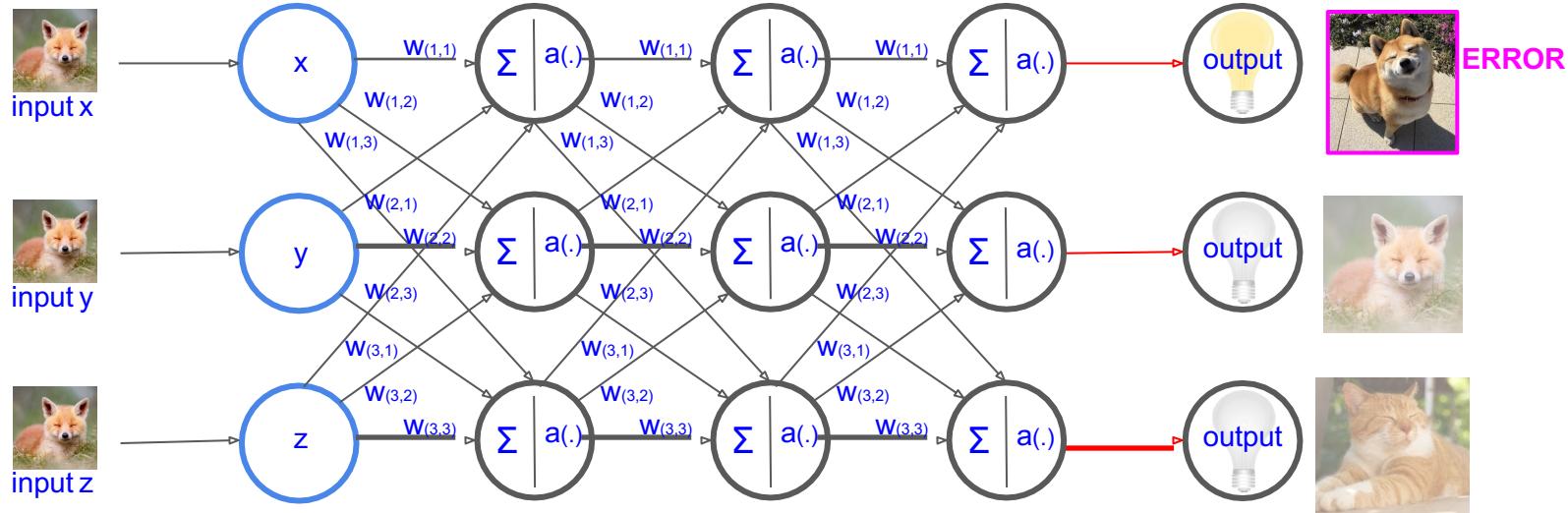
Forward Signal



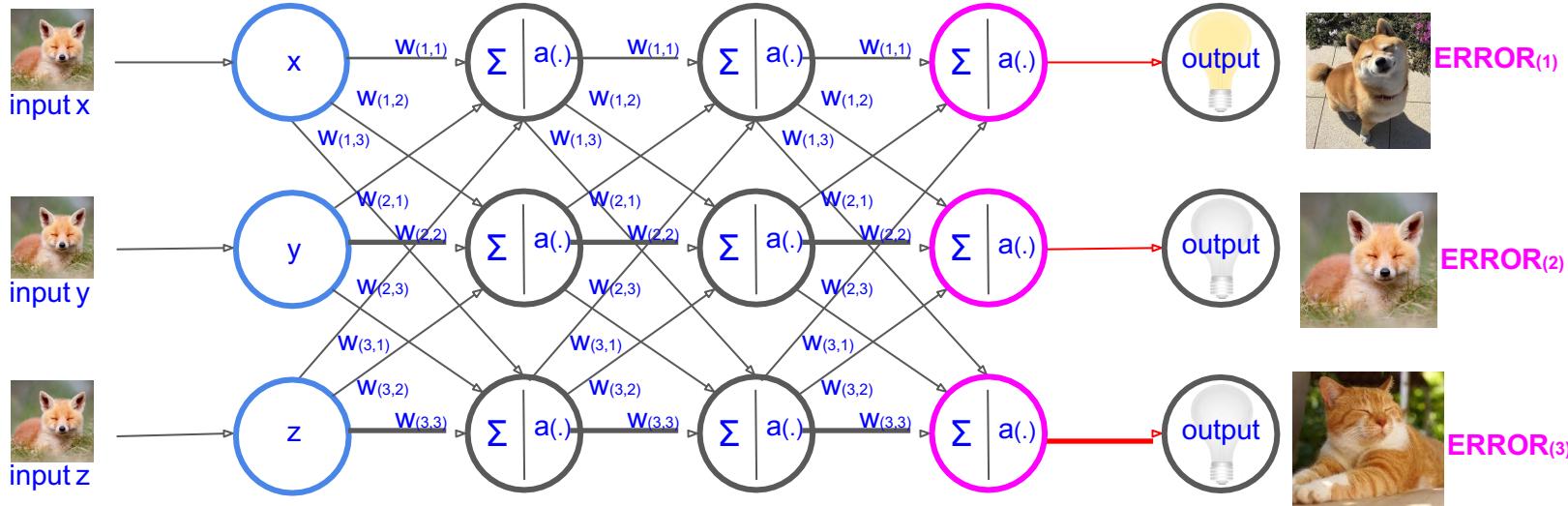
Forward Signal



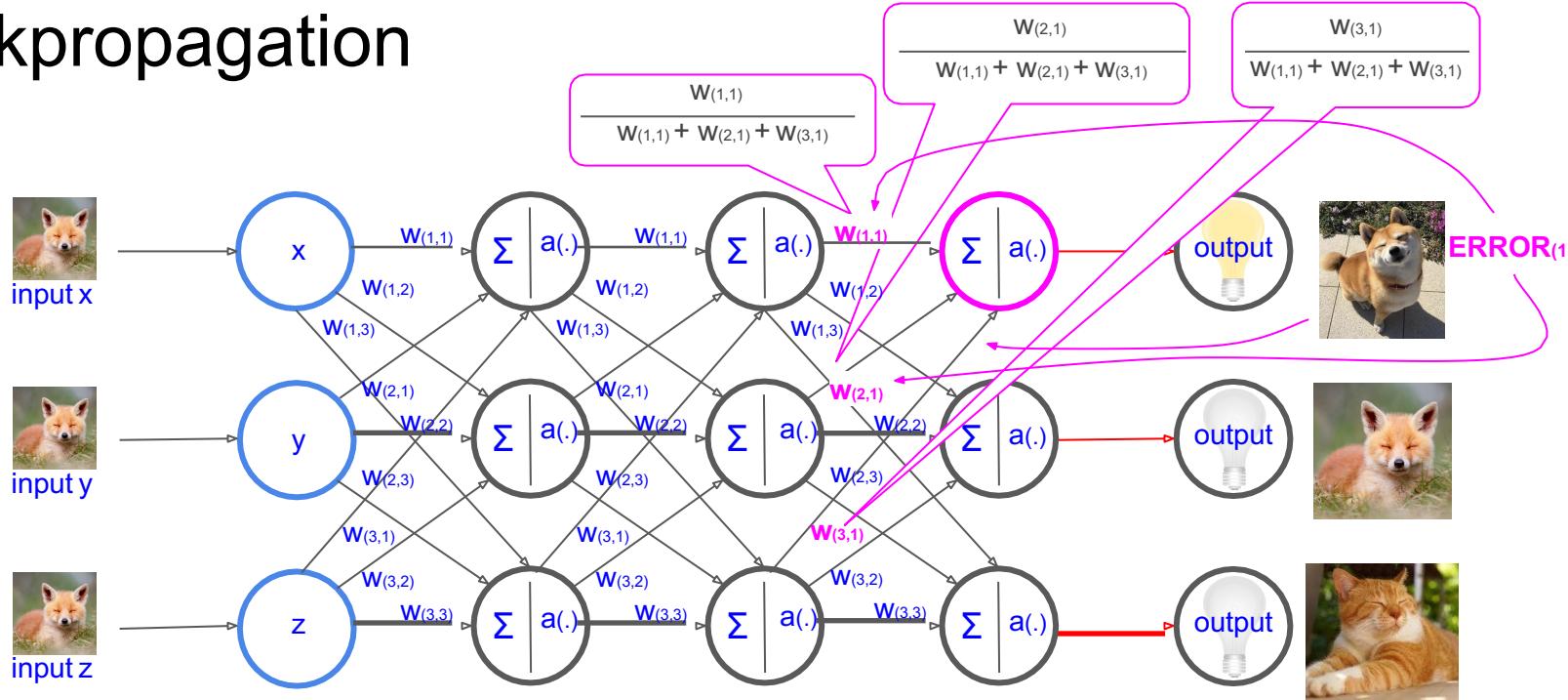
Error



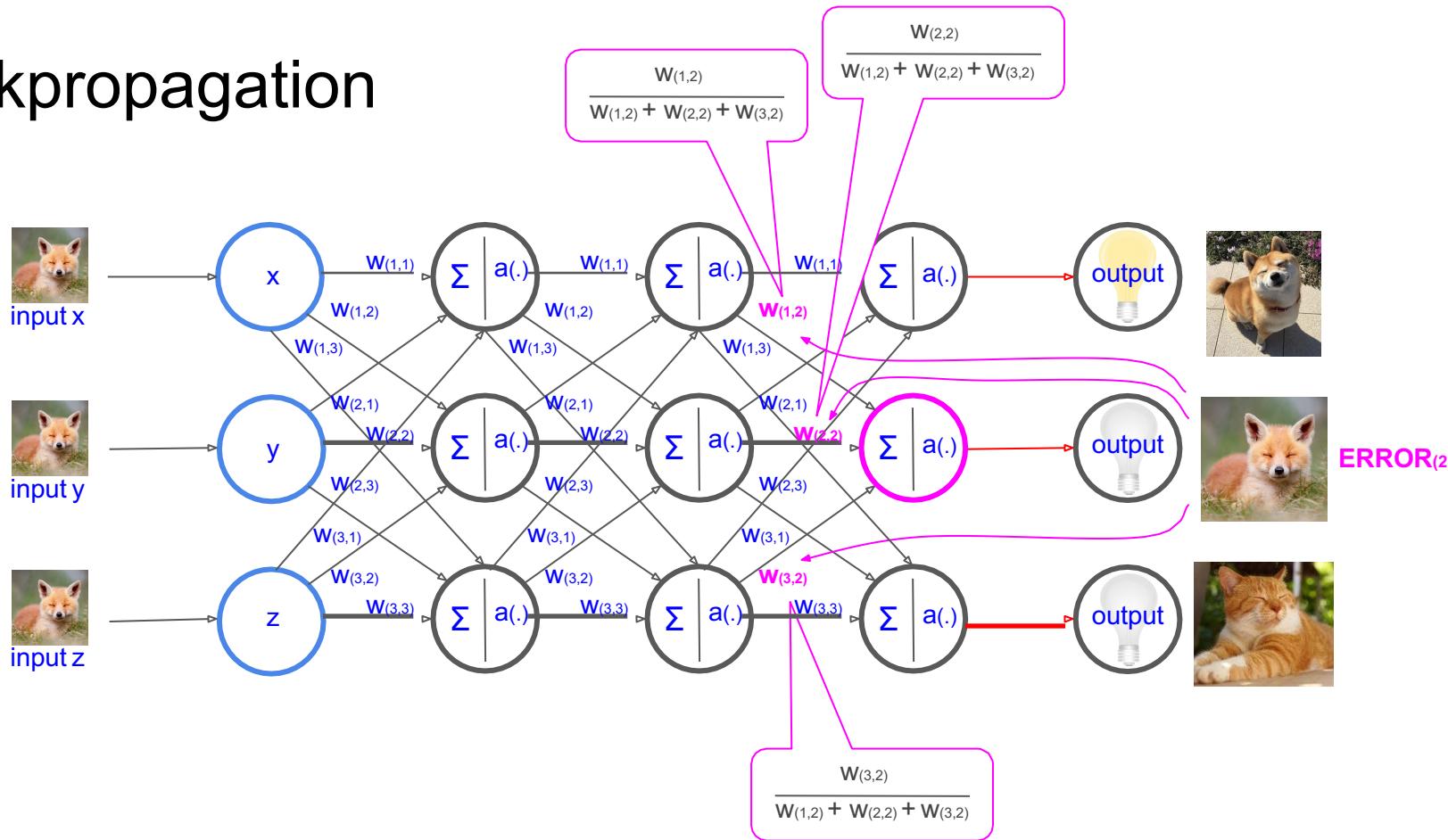
Error



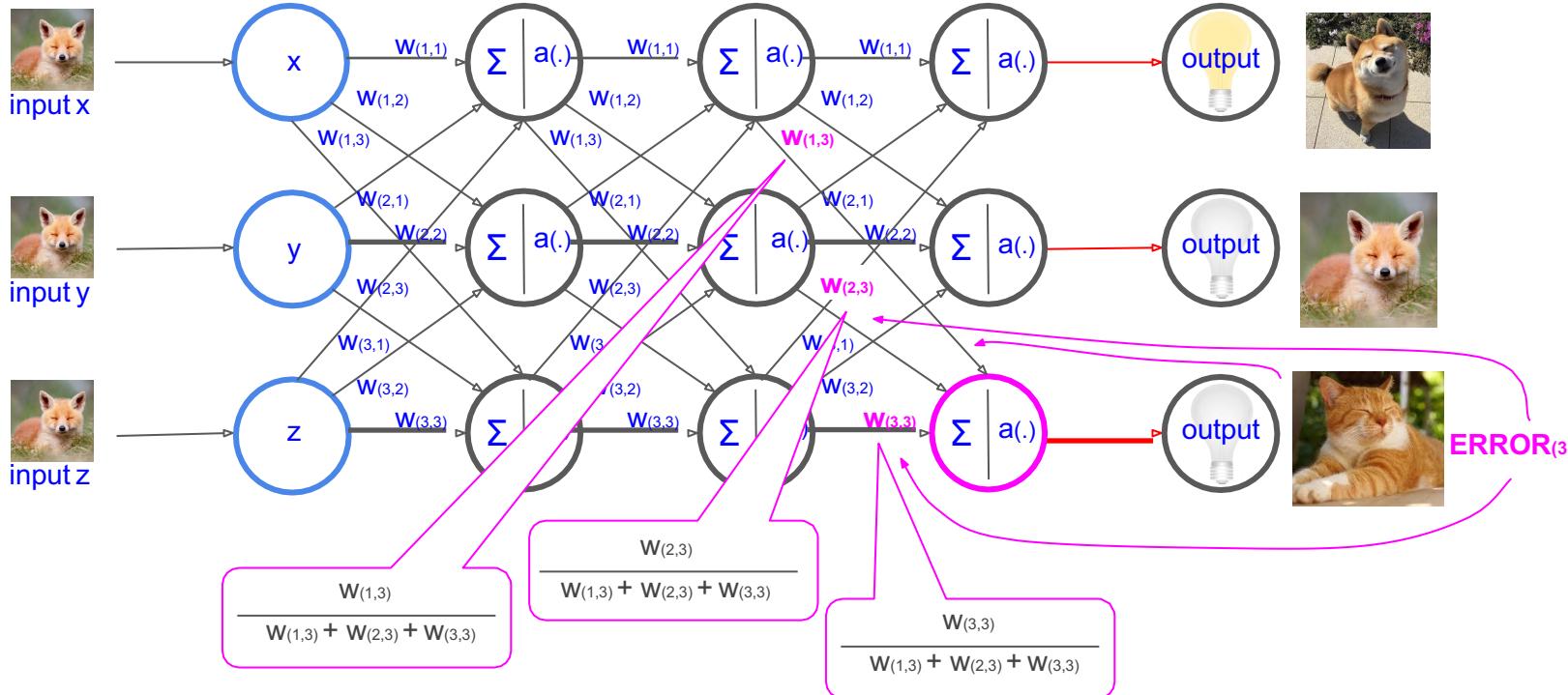
Backpropagation



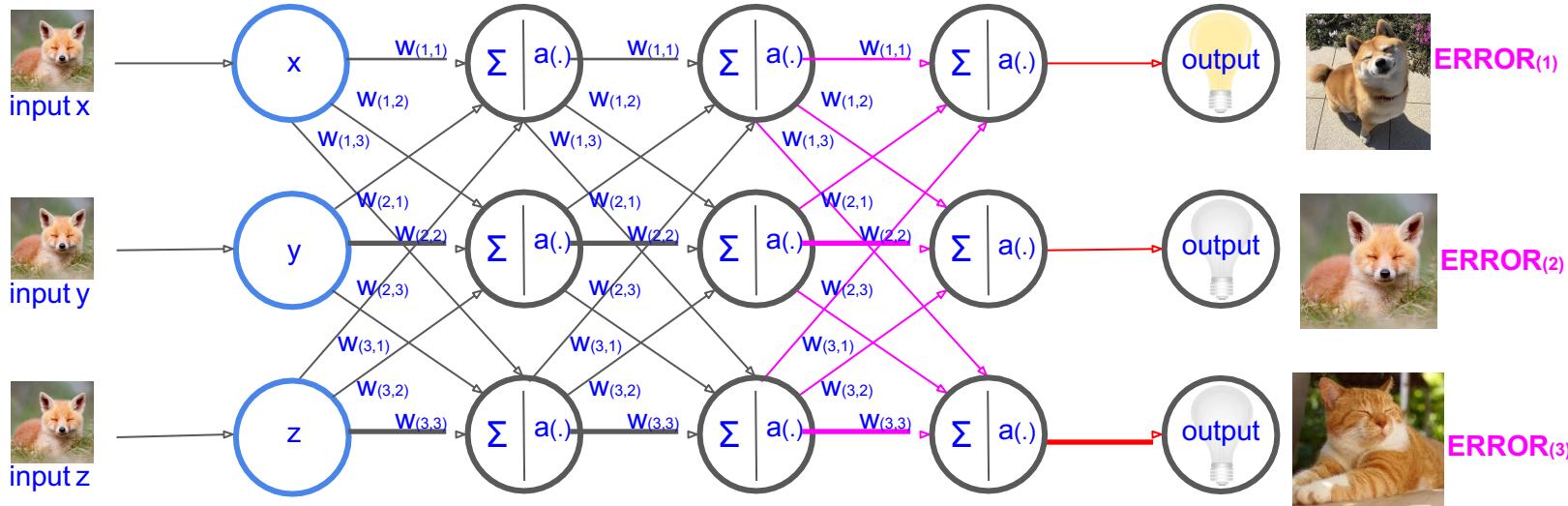
Backpropagation



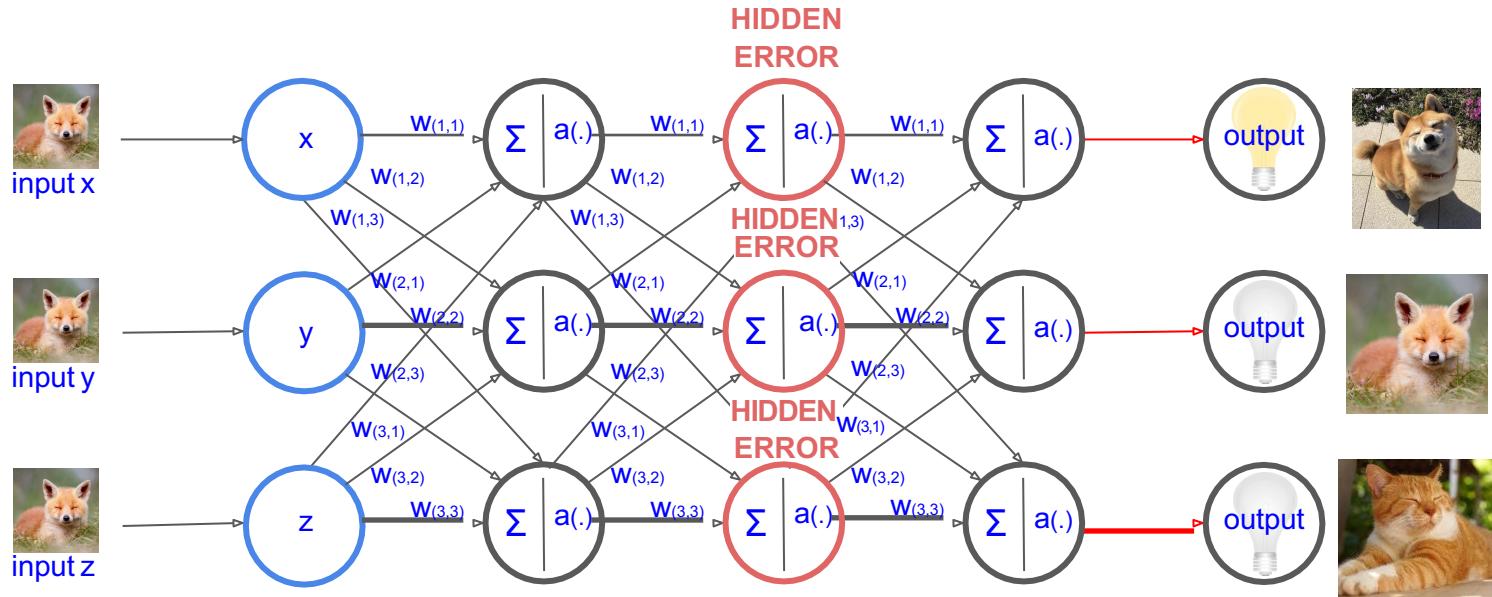
Backpropagation



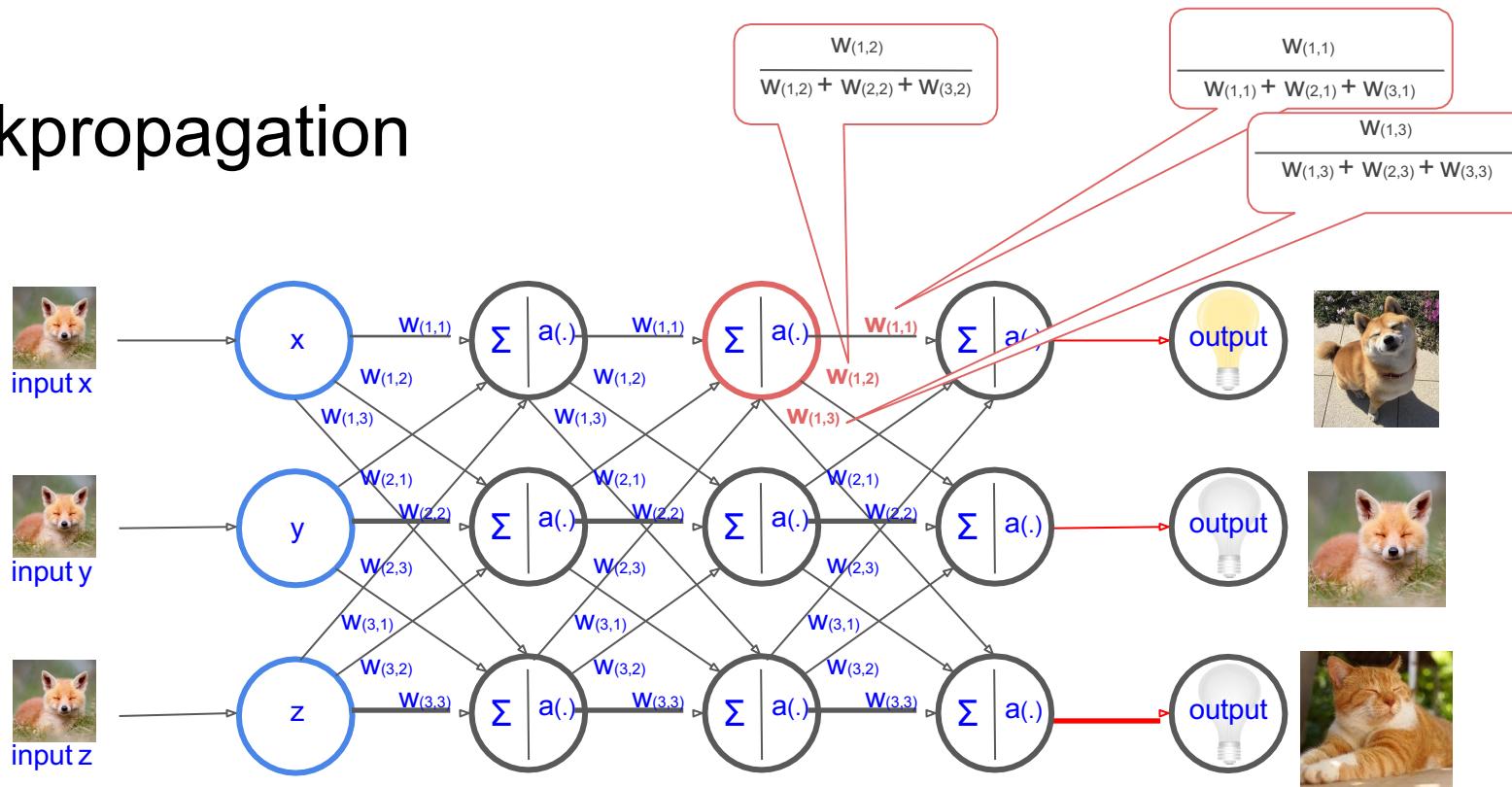
Backpropagation



Backpropagation

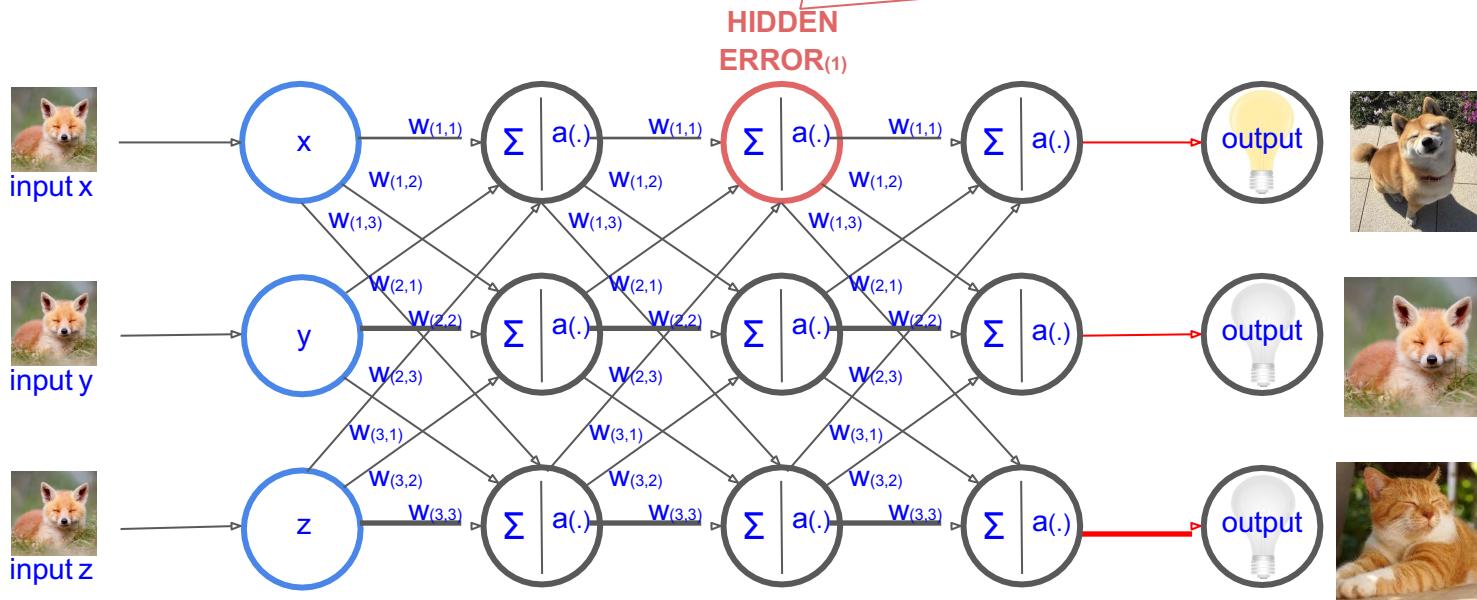


Backpropagation

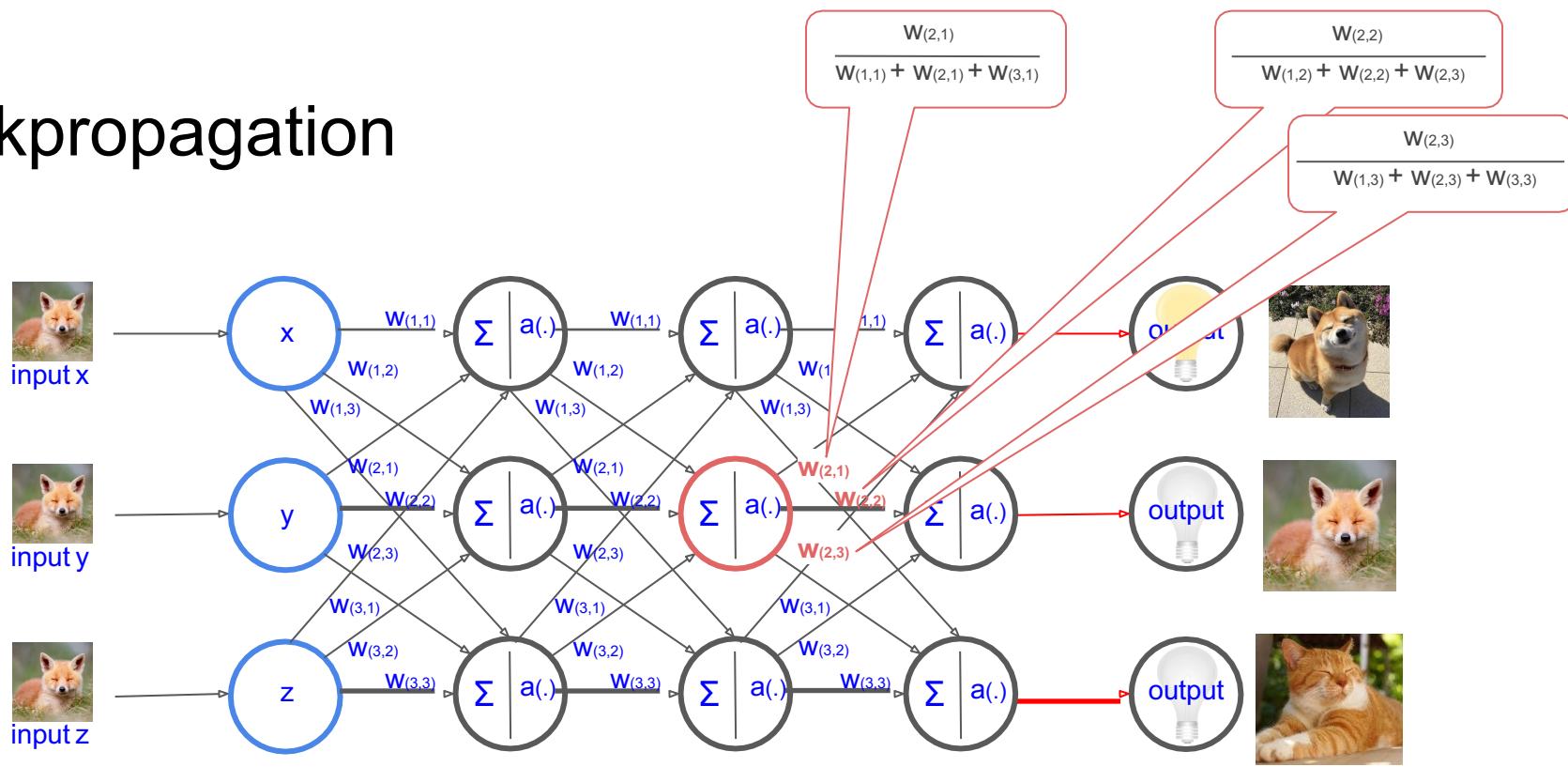


Backpropagation

$$\text{error}(2)^* \frac{W_{(1,2)}}{W_{(1,2)} + W_{(2,2)} + W_{(3,2)}} + \text{error}(1)^* \frac{W_{(1,1)}}{W_{(1,1)} + W_{(2,1)} + W_{(3,1)}} + \text{error}(3)^* \frac{W_{(1,3)}}{W_{(1,3)} + W_{(2,3)} + W_{(3,3)}}$$

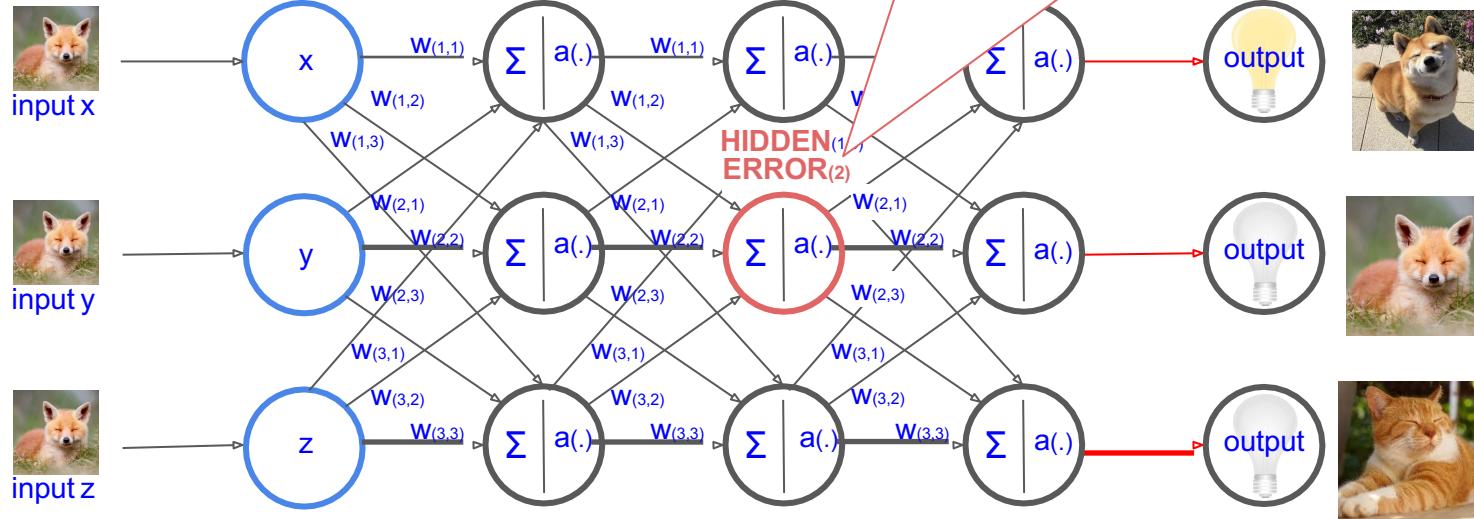


Backpropagation

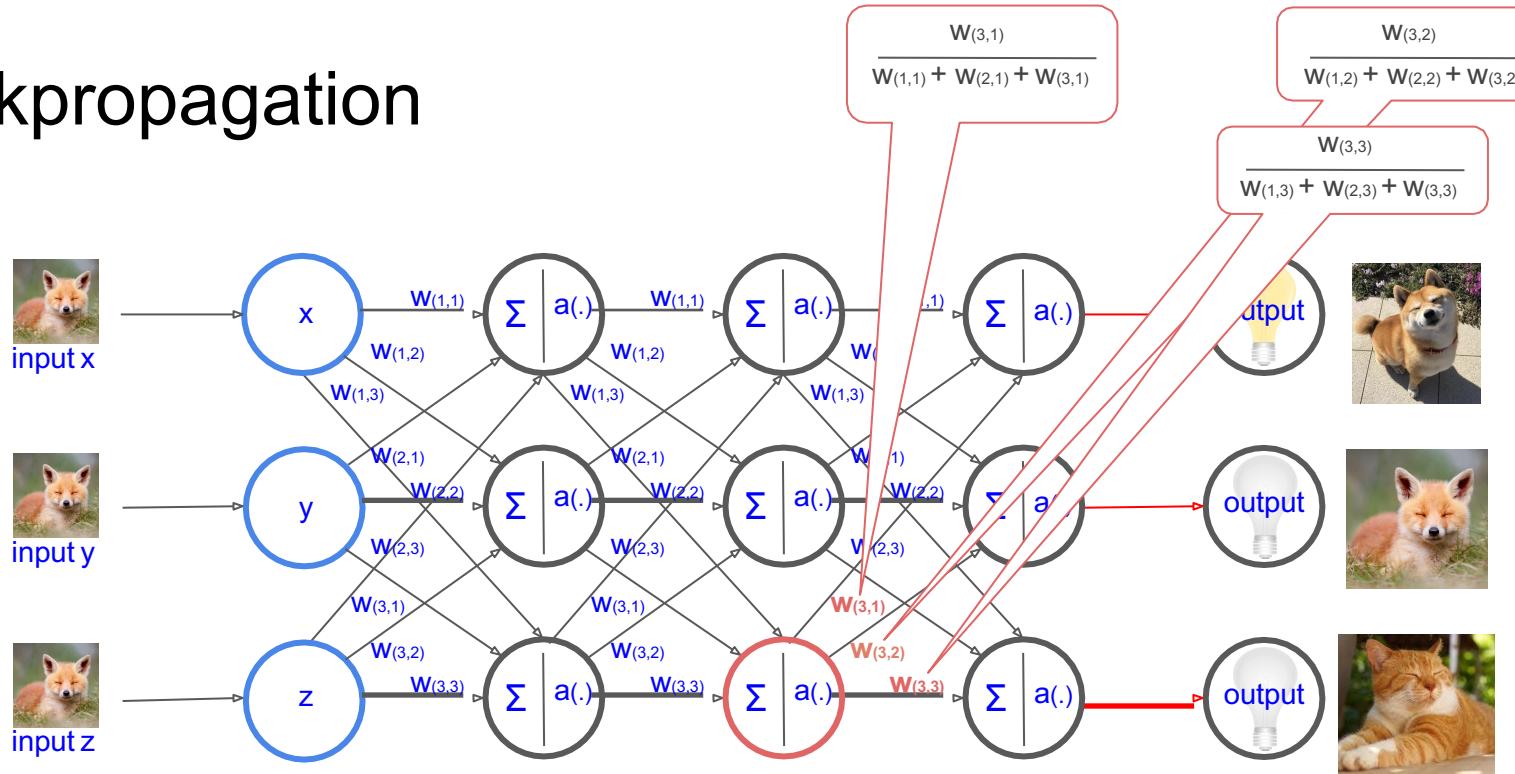


Backpropagation

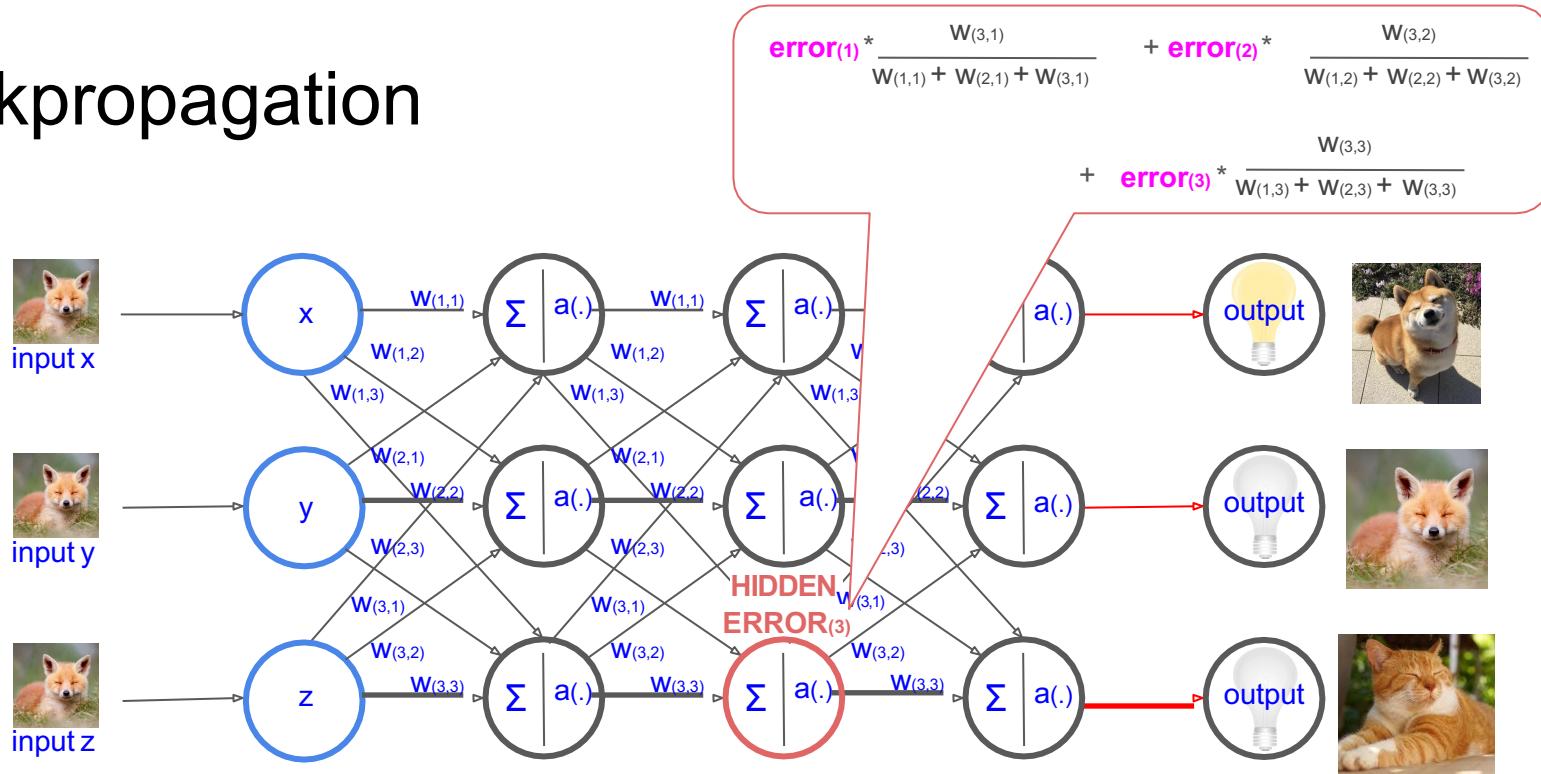
$$\text{error}_{(1)} * \frac{W_{(2,1)}}{W_{(1,1)} + W_{(2,1)} + W_{(3,1)}} + \text{error}_{(2)} * \frac{W_{(2,2)}}{W_{(1,2)} + W_{(2,2)} + W_{(3,2)}} + \text{error}_{(3)} * \frac{W_{(2,3)}}{W_{(1,3)} + W_{(2,3)} + W_{(3,3)}}$$



Backpropagation

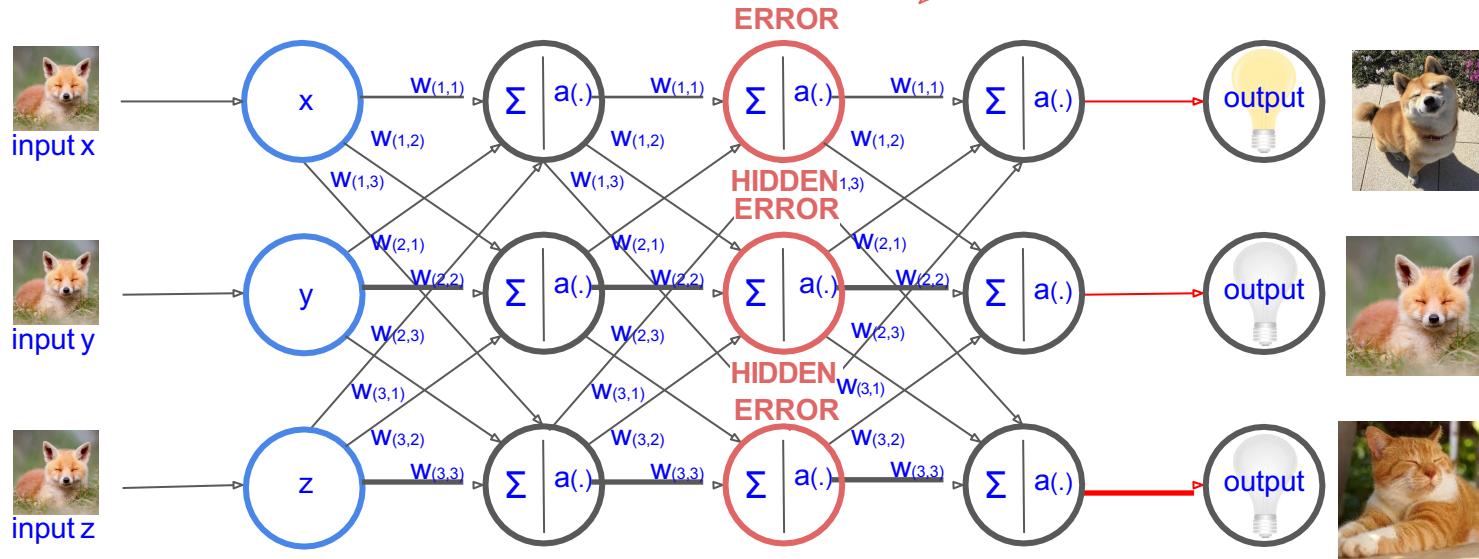


Backpropagation



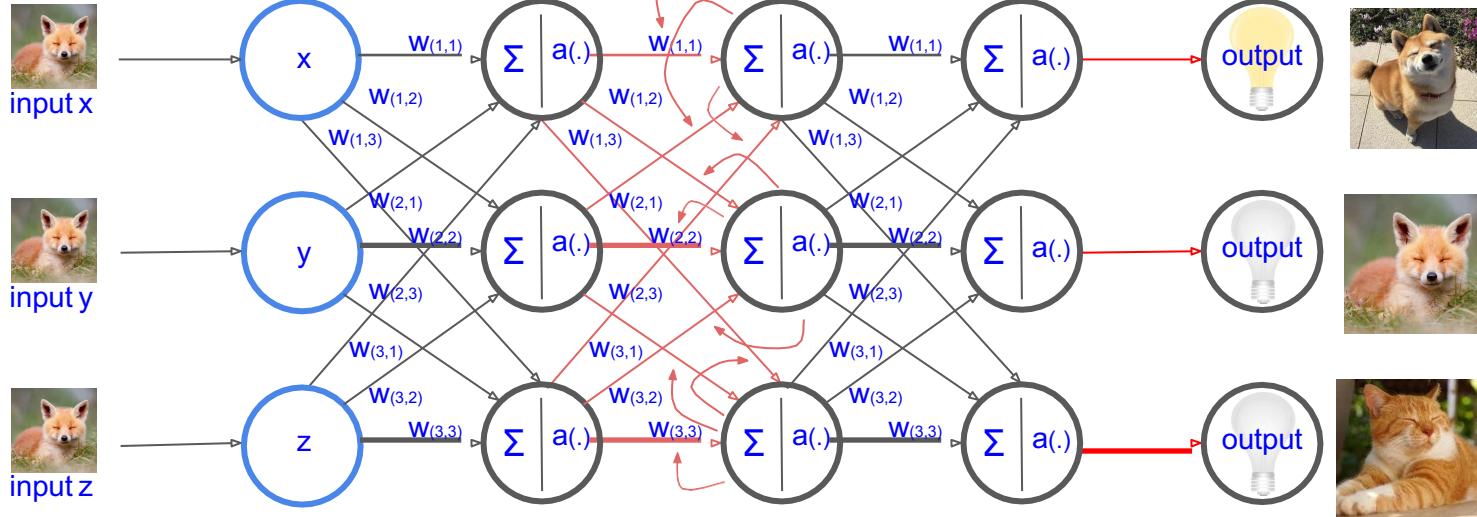
Backpropagation

$$E_{\text{hidden}} = W^T_{\text{hidden_output}} \cdot E_{\text{output}}$$



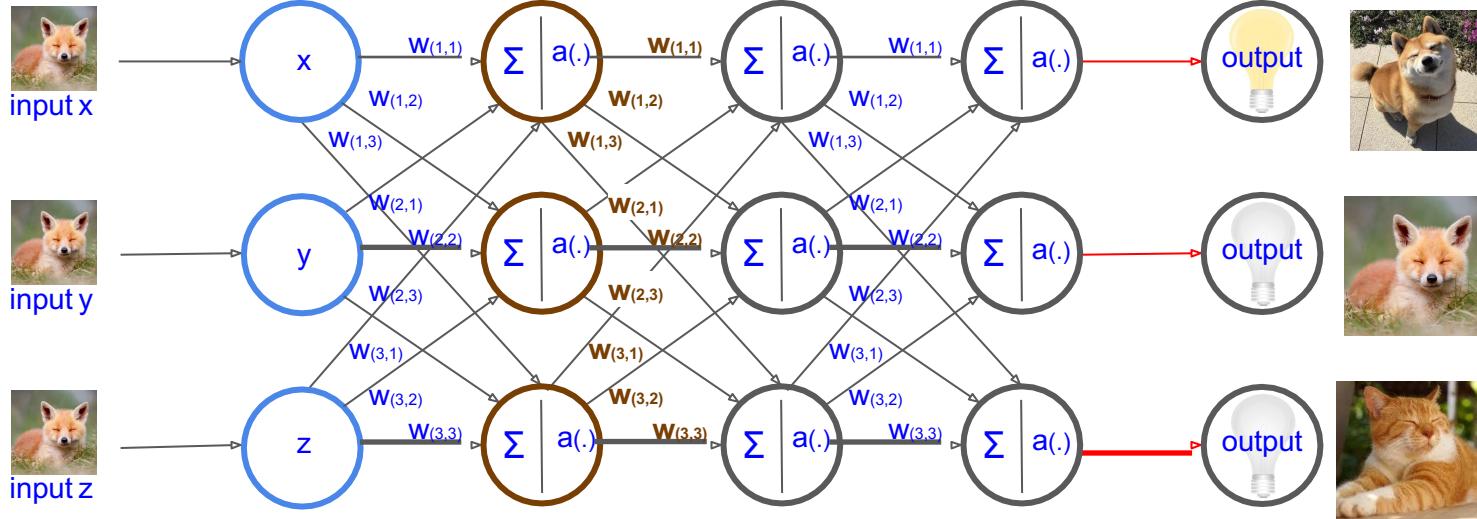
Backpropagation

$$E_{\text{hidden}} = W^T_{\text{hidden_output}} \cdot E_{\text{output}}$$



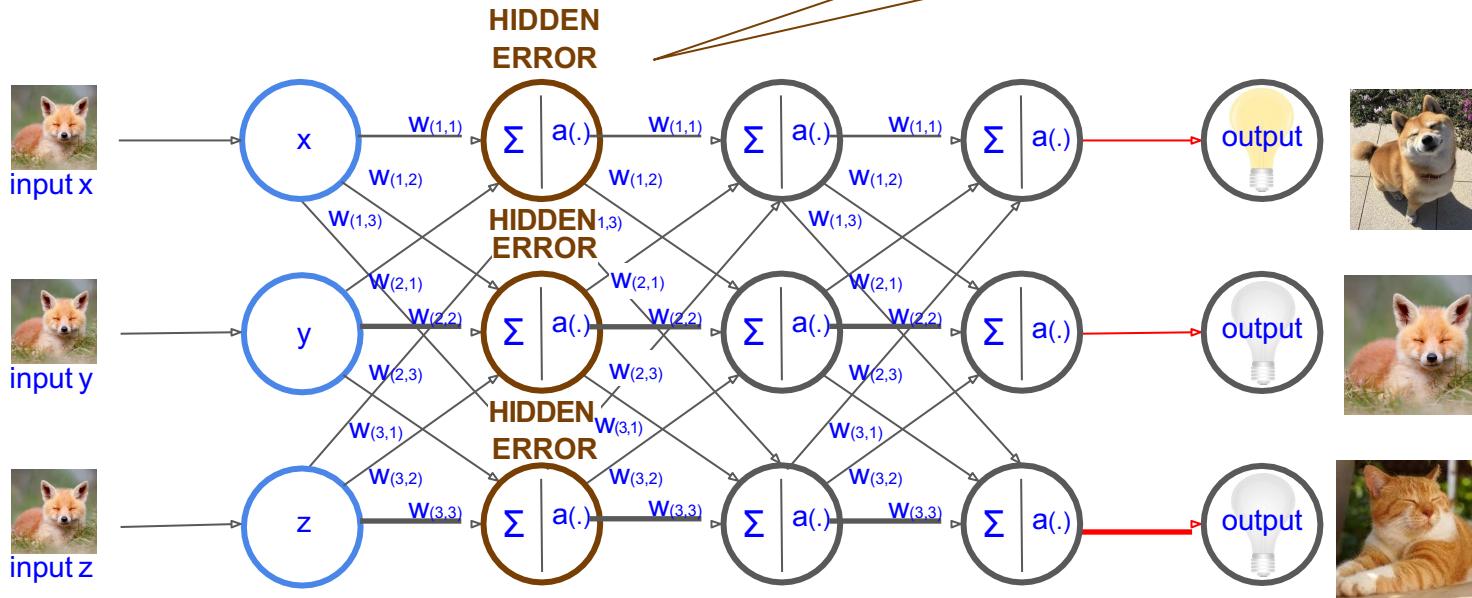
Backpropagation

$$E_{\text{hidden}} = W^T_{\text{hidden_output}} \cdot E_{\text{output}}$$



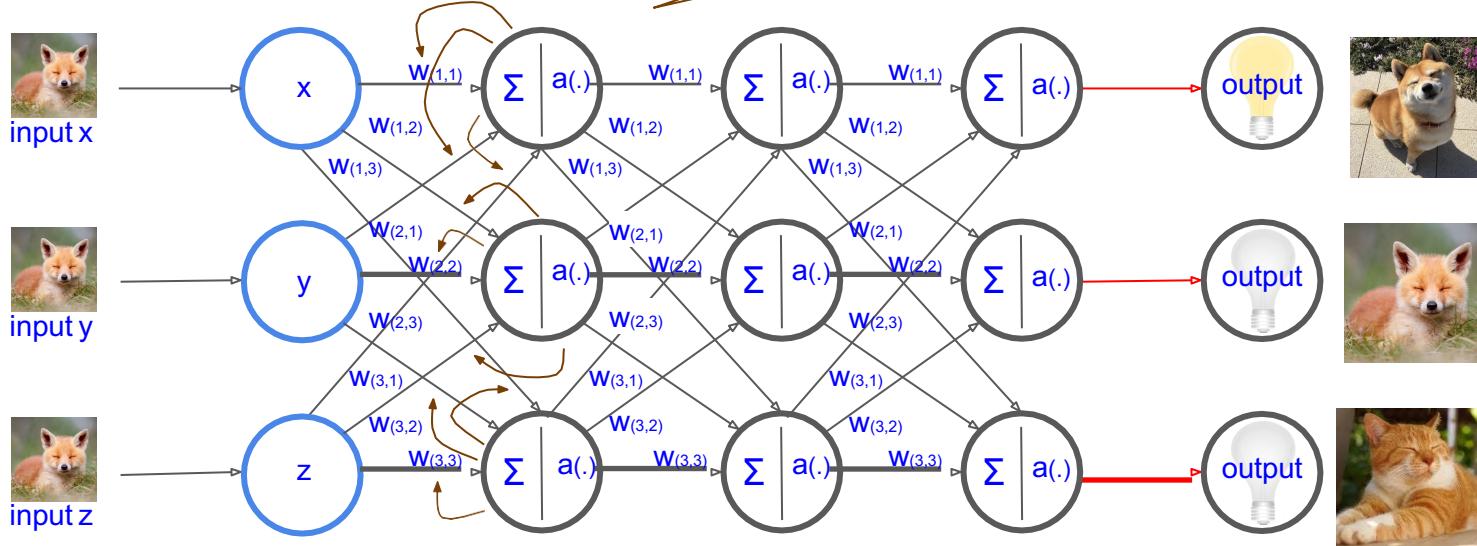
Backpropagation

$$E_{\text{hidden}} = W^T_{\text{hidden_output}} \cdot E_{\text{output}}$$

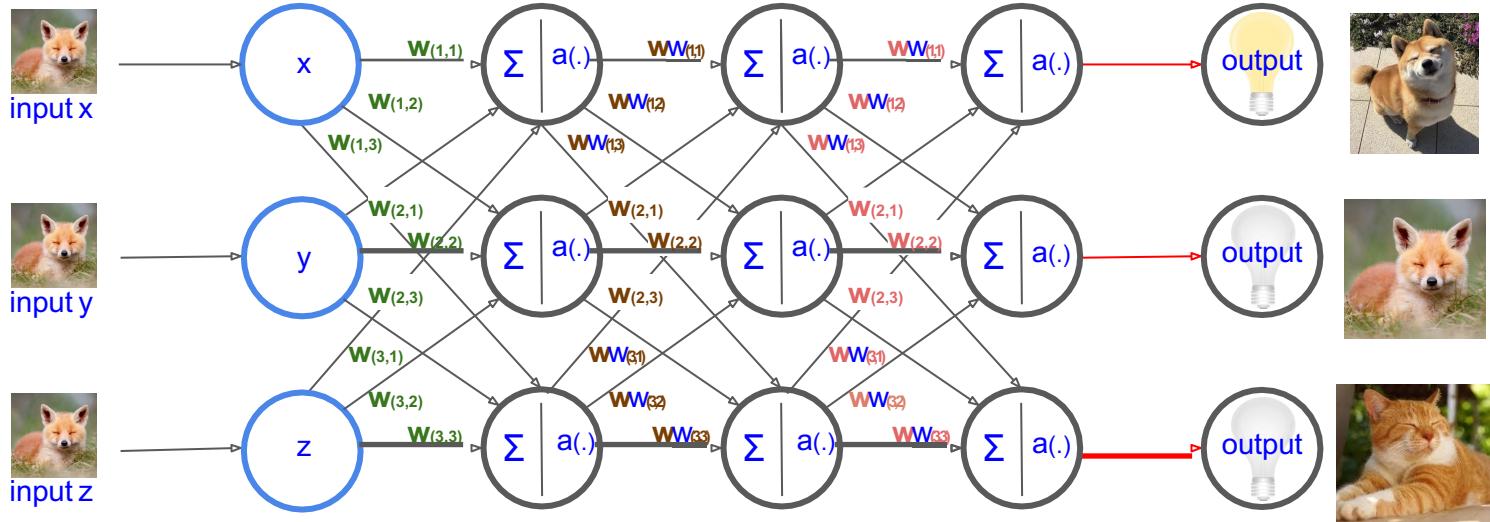


Backpropagation

$$E_{\text{hidden}} = W^T_{\text{hidden_output}} \cdot E_{\text{output}}$$

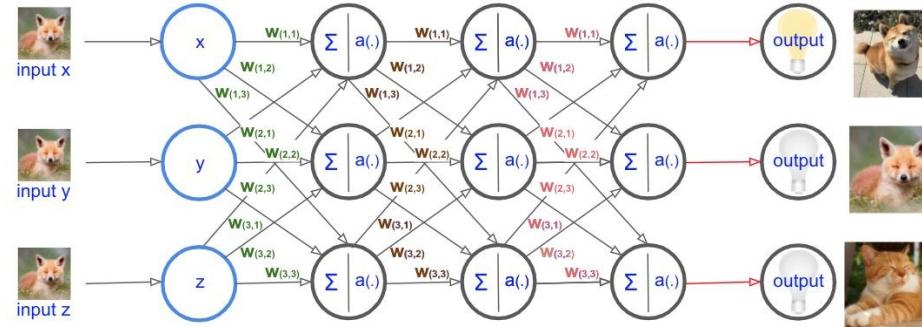


Updating Weights



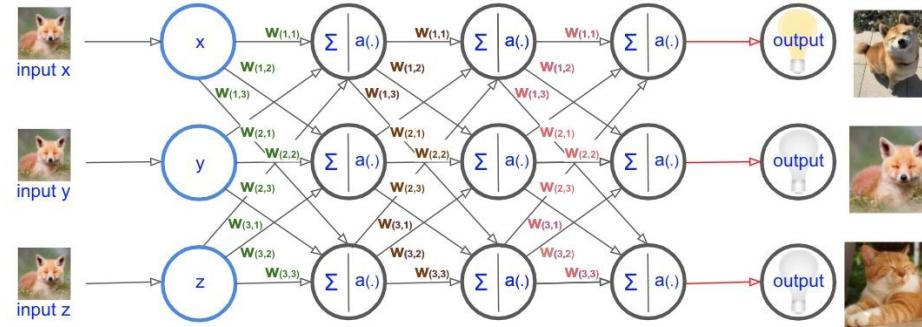
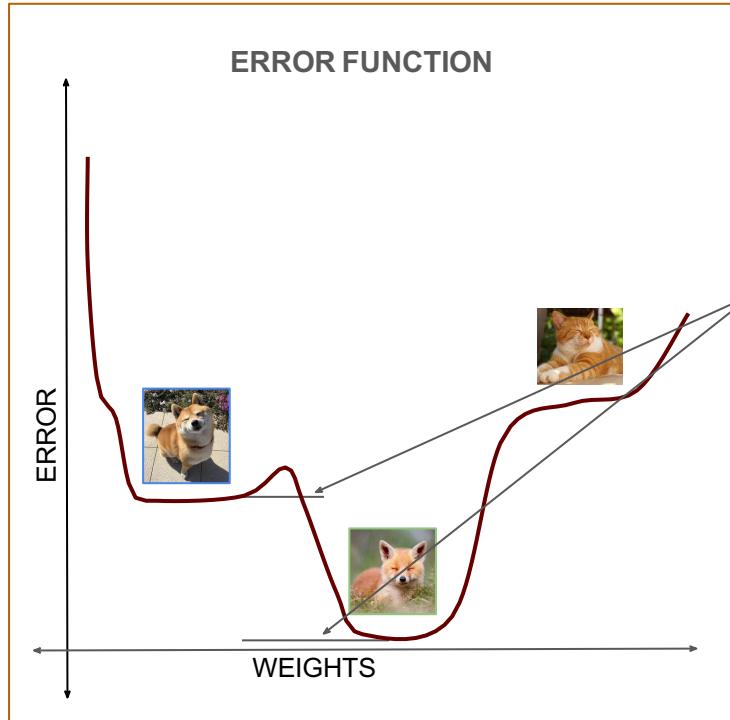
Updating Weights

error = desired target - actual output



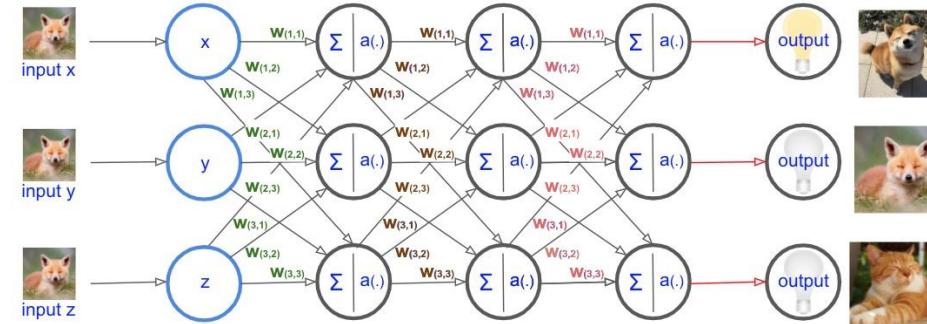
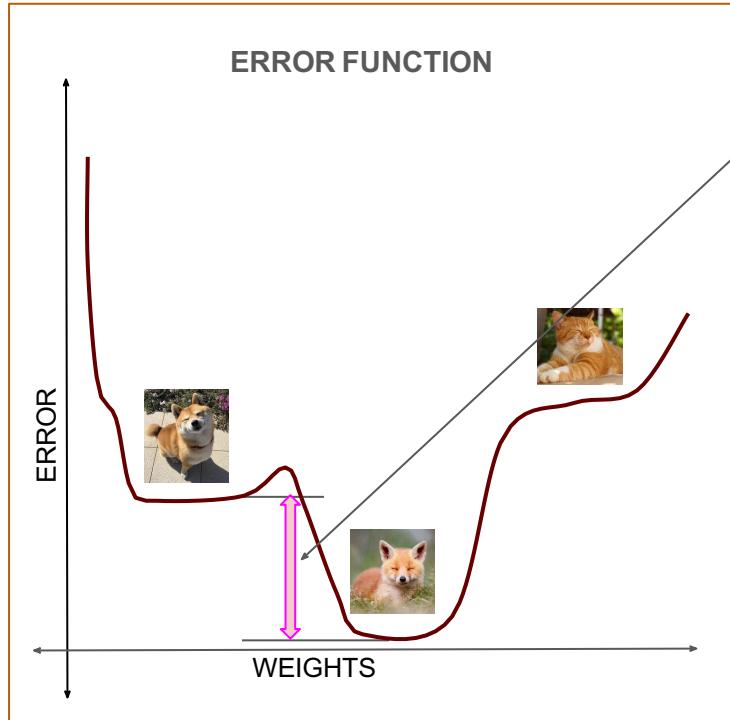
Updating Weights

$$\text{error} = \text{desired target} - \text{actual output}$$



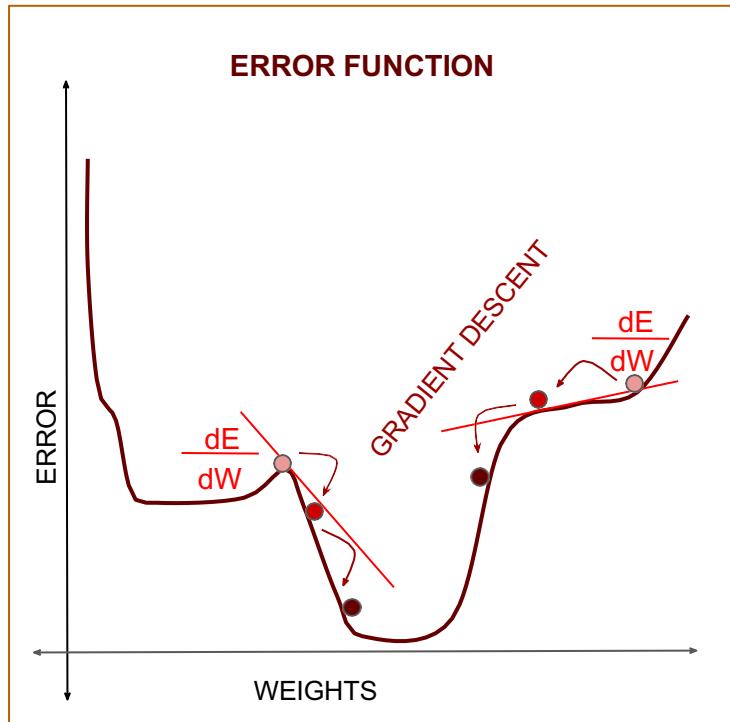
Updating Weights

error = (desired target - actual output)

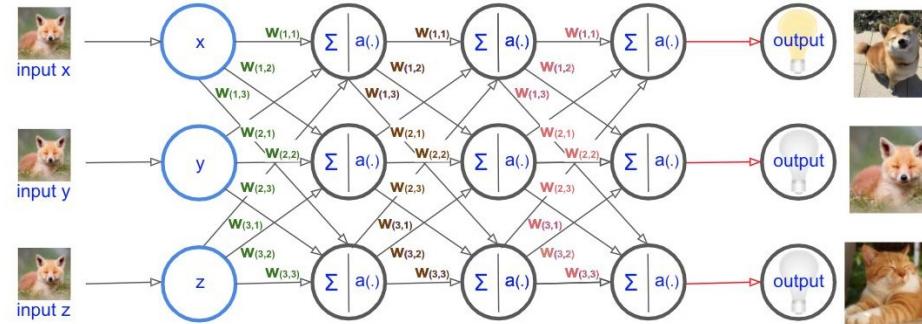


Updating Weights

error = (desired target - actual output)

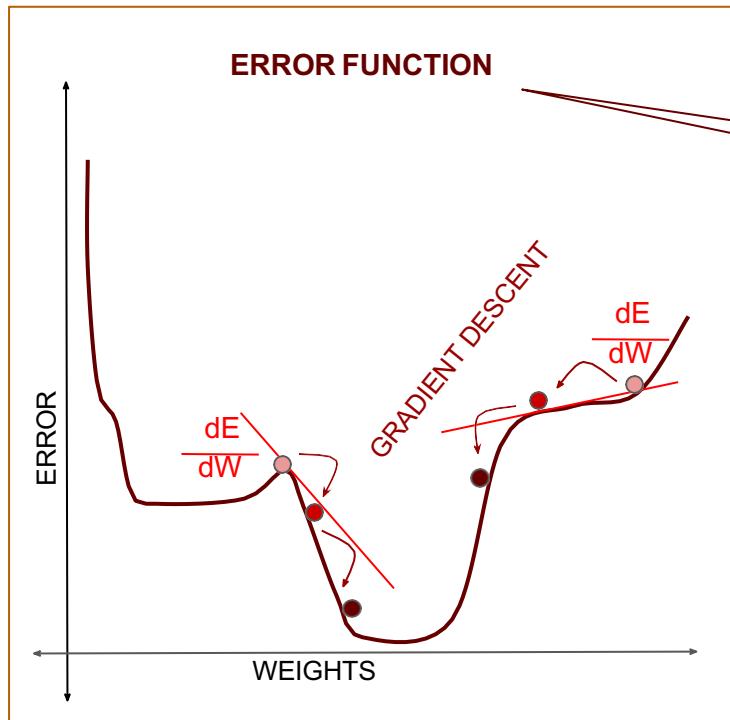


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$



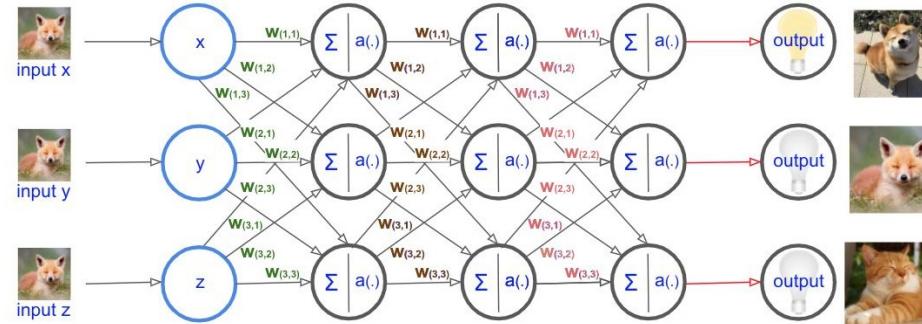
Updating Weights

error = (desired target - actual output)

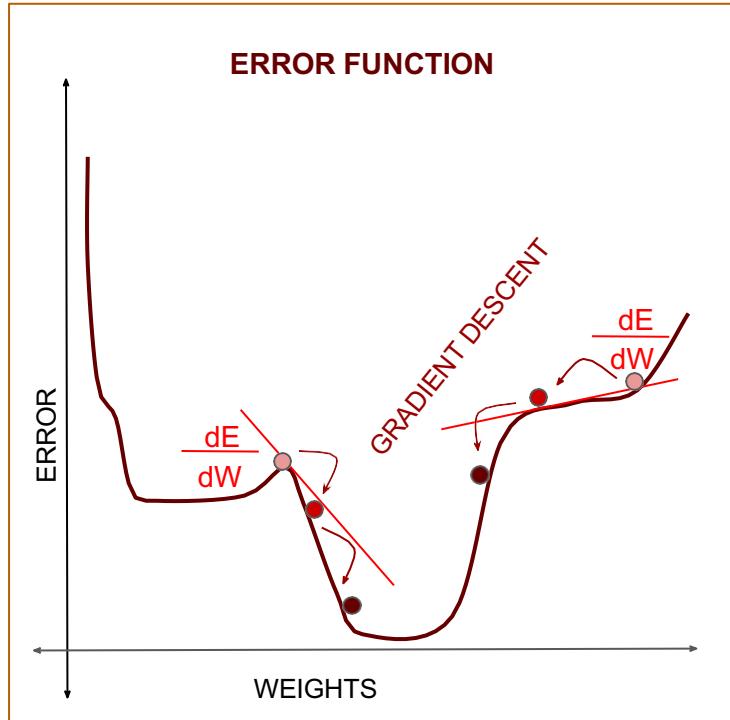


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

(desired target - actual output)²

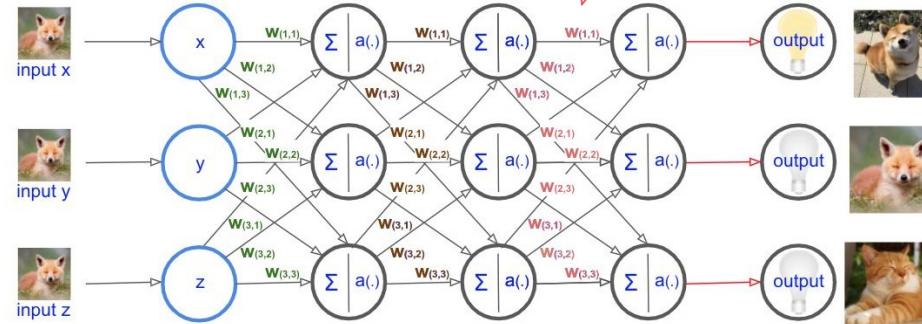


Updating Weights

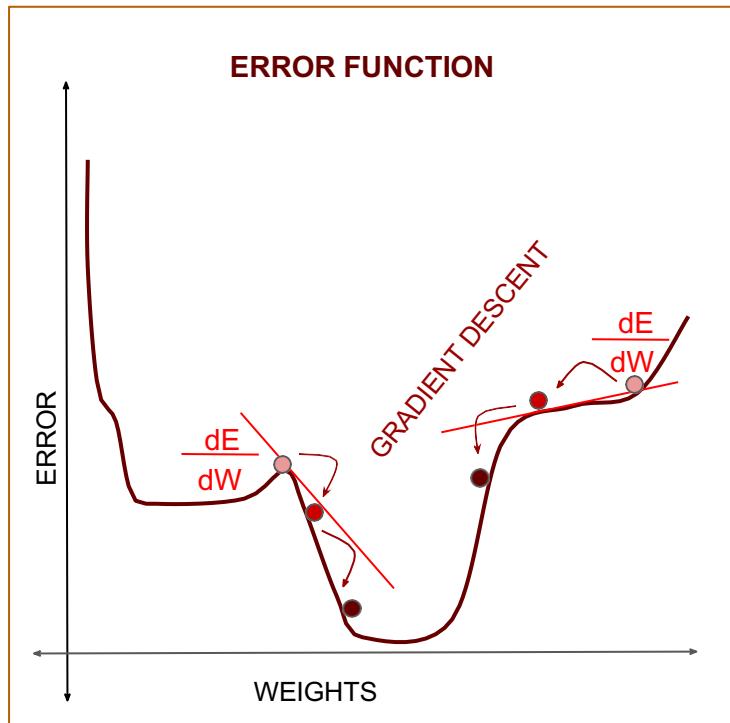


error = (desired target - actual output)

$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

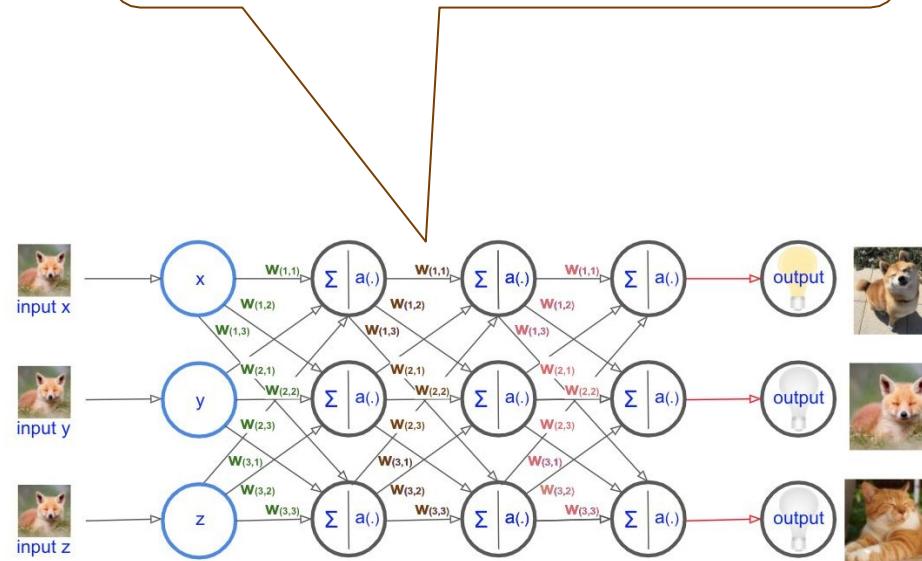


Updating Weights

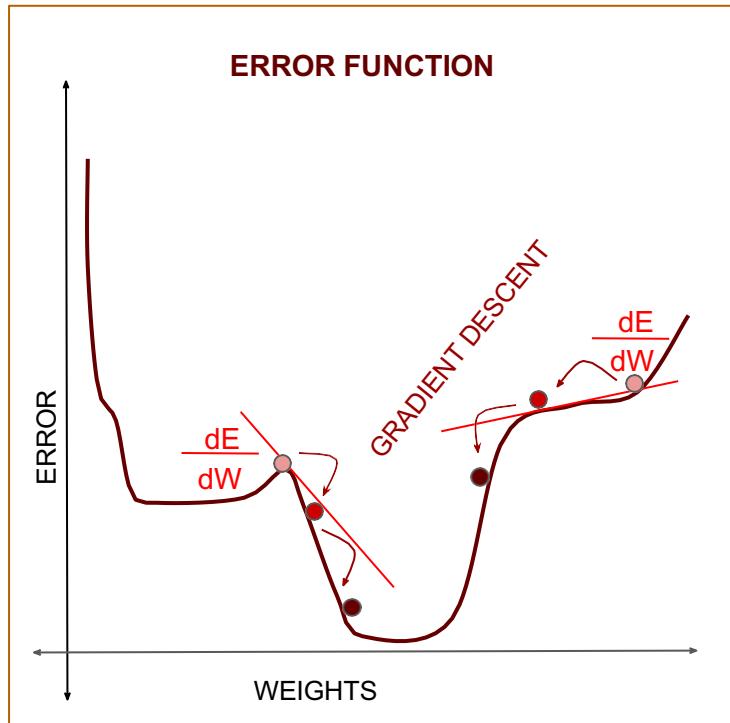


error = (desired target - actual output)

$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

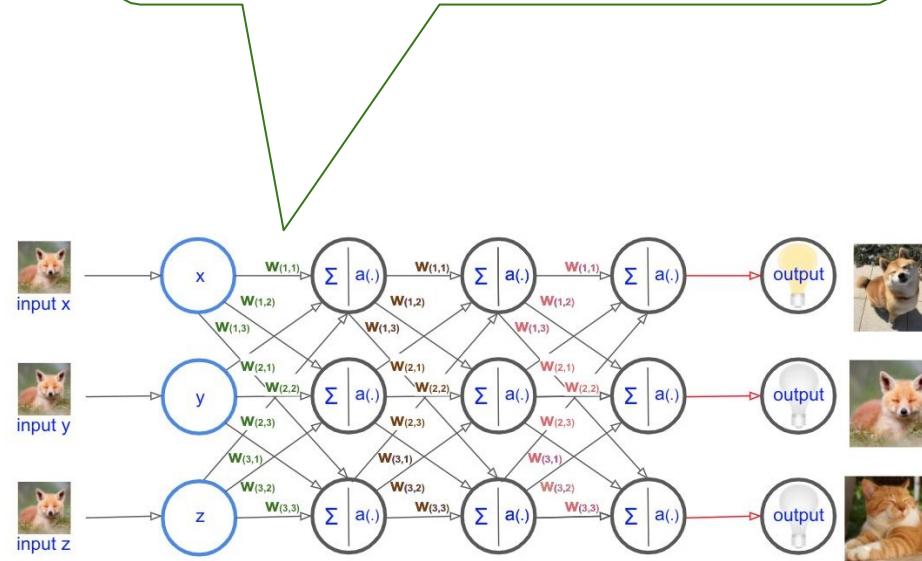


Updating Weights



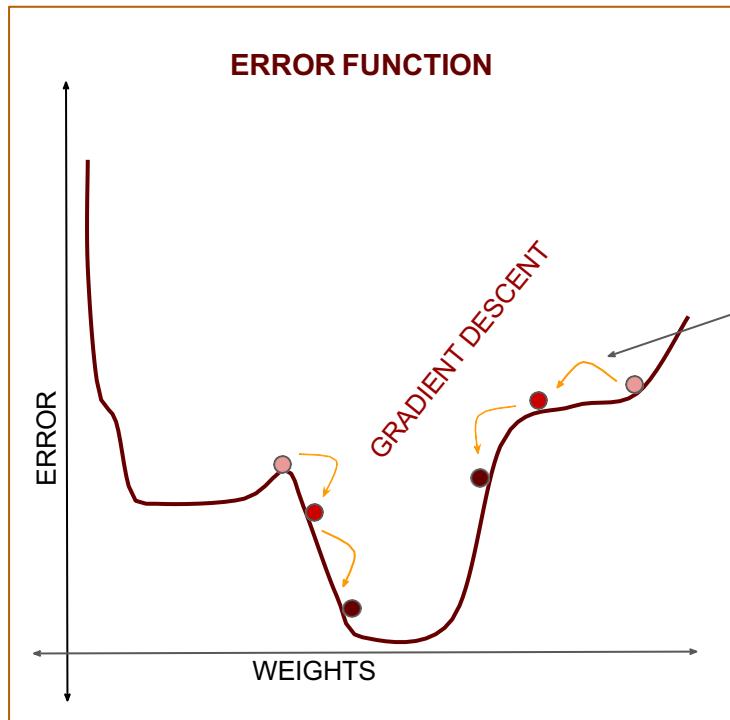
error = (desired target - actual output)

$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

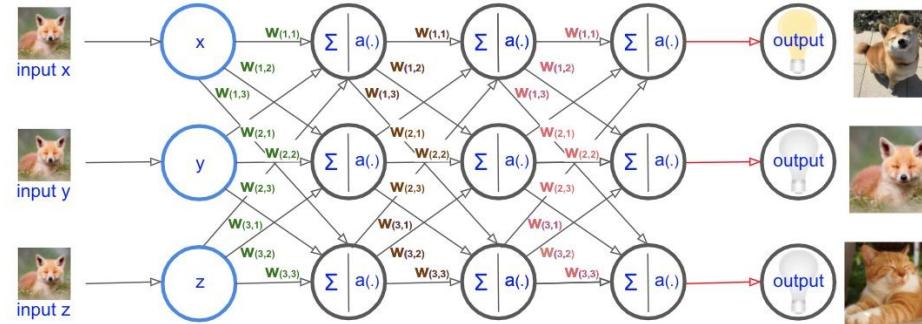


Updating Weights

error = (desired target - actual output)

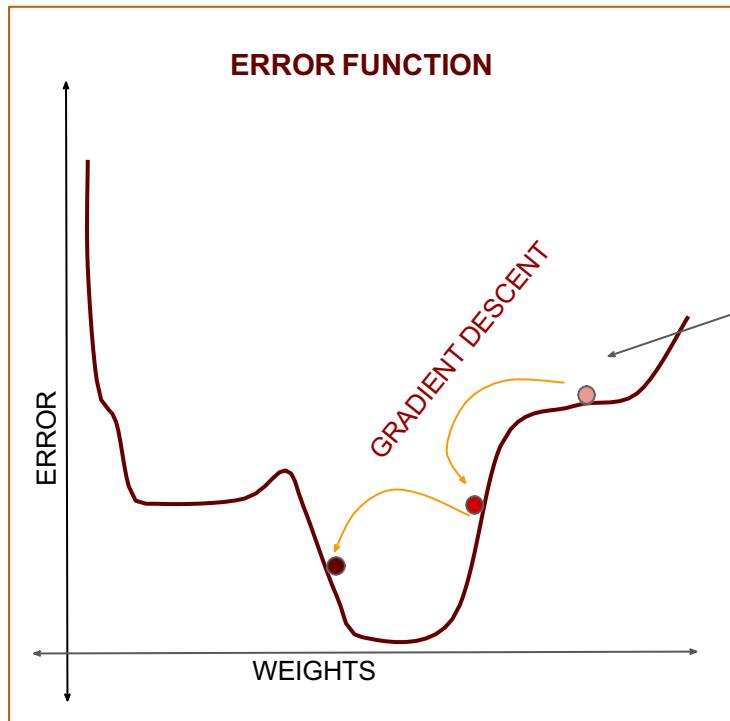


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

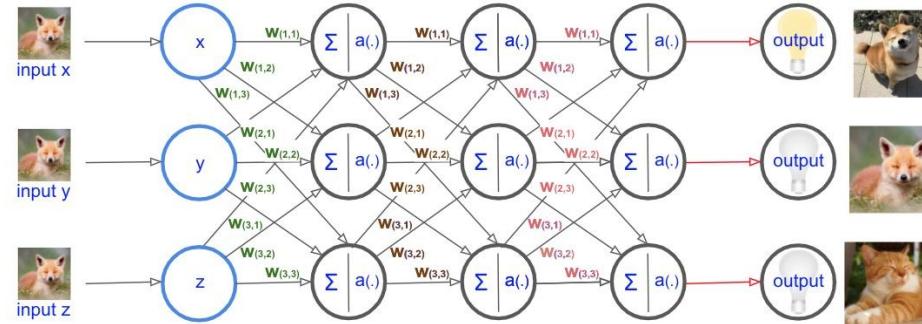


Updating Weights

error = (desired target - actual output)

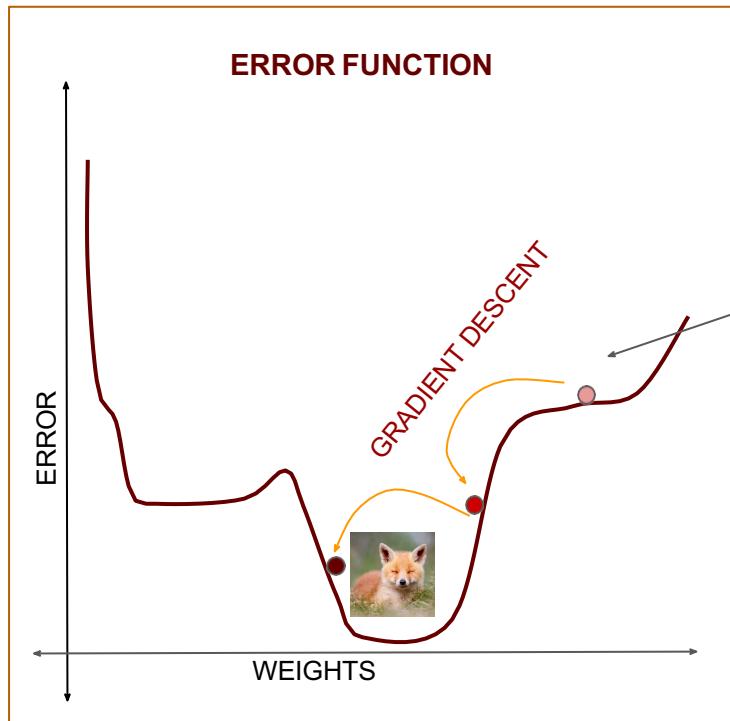


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

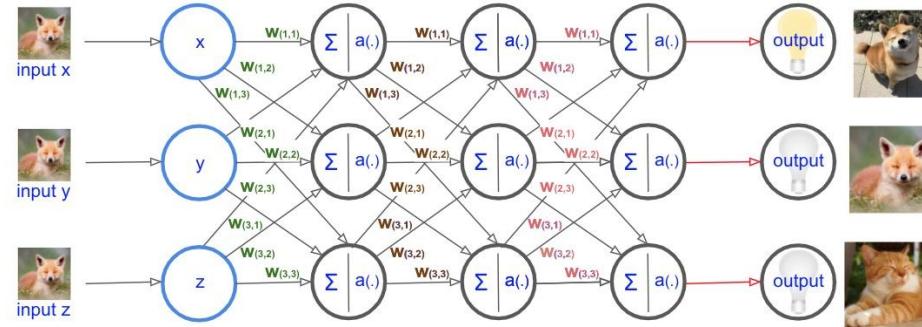


Updating Weights

error = (desired target - actual output)

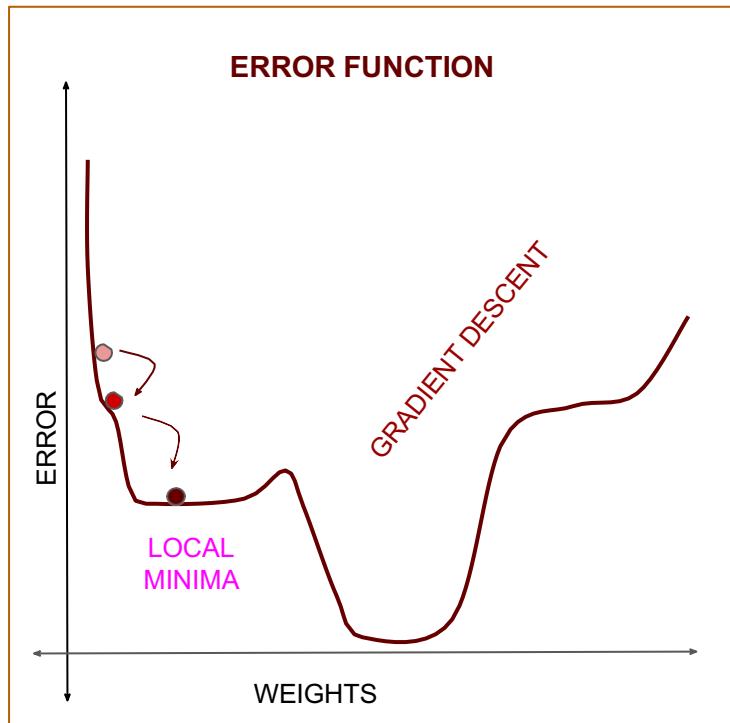


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$

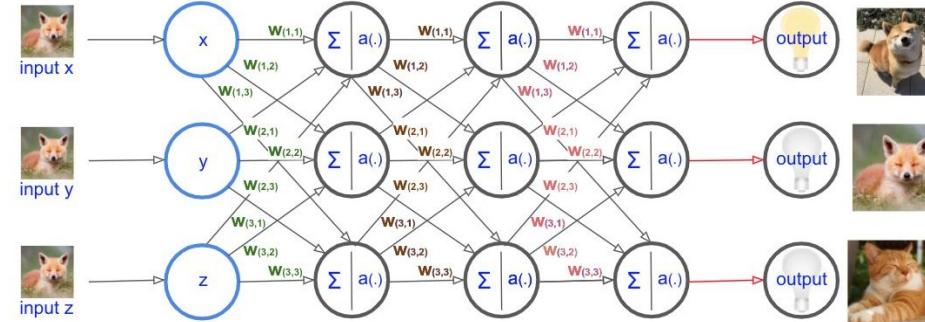


Updating Weights

error = (desired target - actual output)

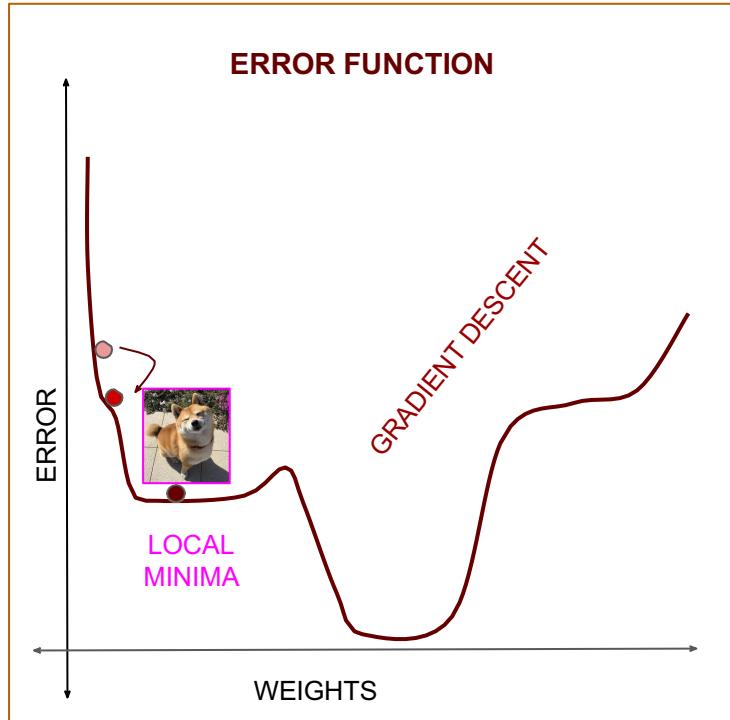


$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{dE}{d(\text{old_W})})$$



Updating Weights

$$\text{error} = (\text{desired target} - \text{actual output})$$



$$\text{new_W} = (\text{old_W} - \text{lr} * \frac{\text{dE}}{\text{d(old_W)}})$$

