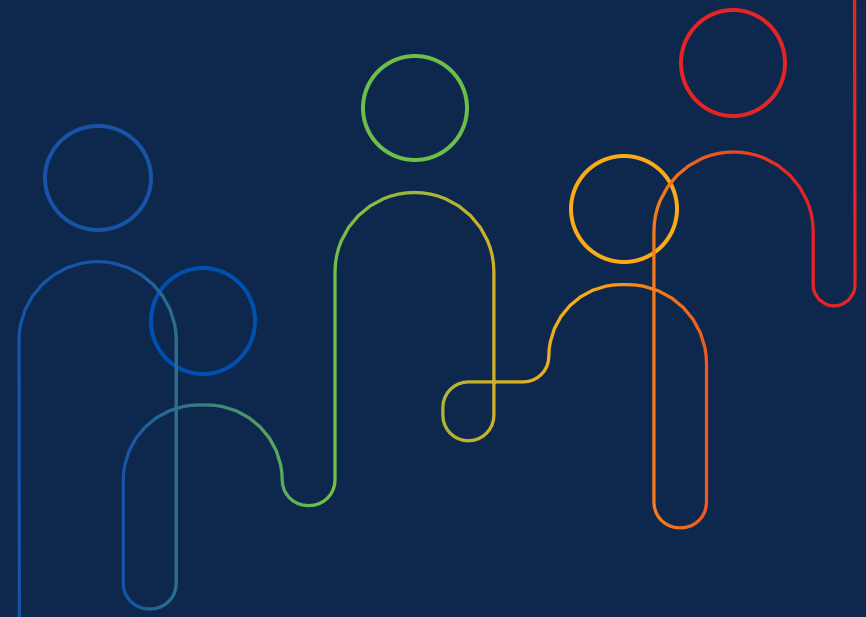
Cisco

Learning & Careers in Partnership with the Generative Artificial Intelligence (GAI) Team

Artificial Neural Networks Overview

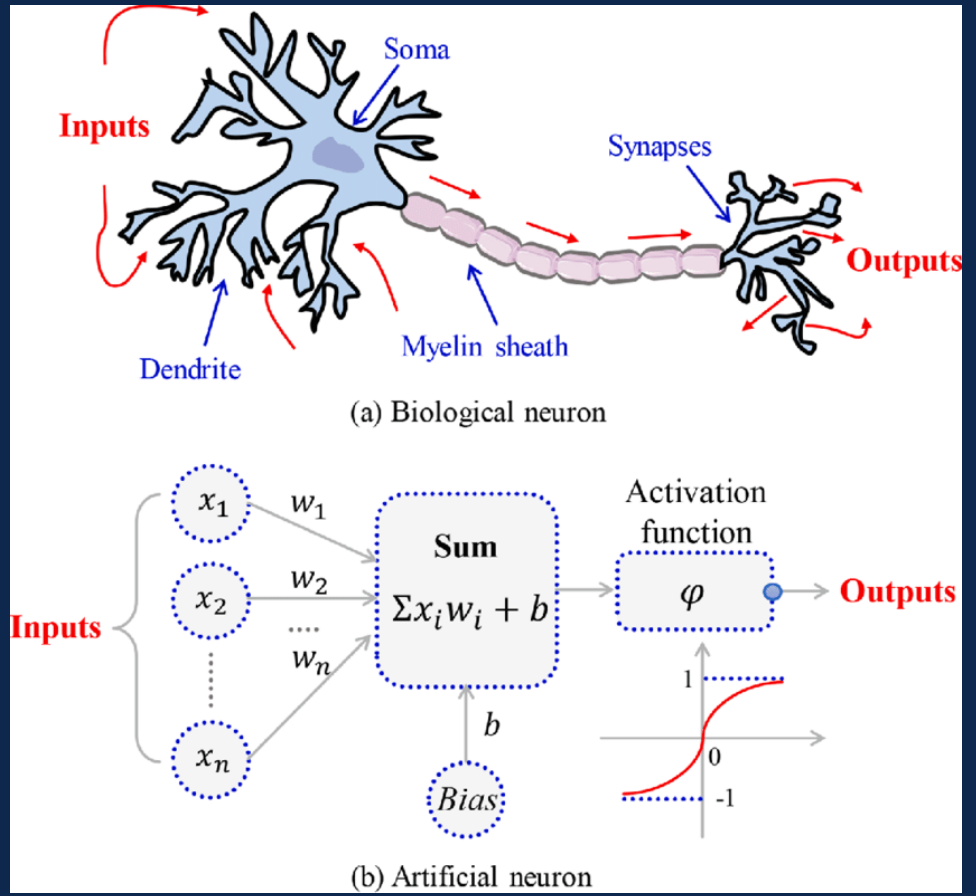


Mohammed Hamzeh, Principal Engineer
TAC Engine

1 Dec 2023

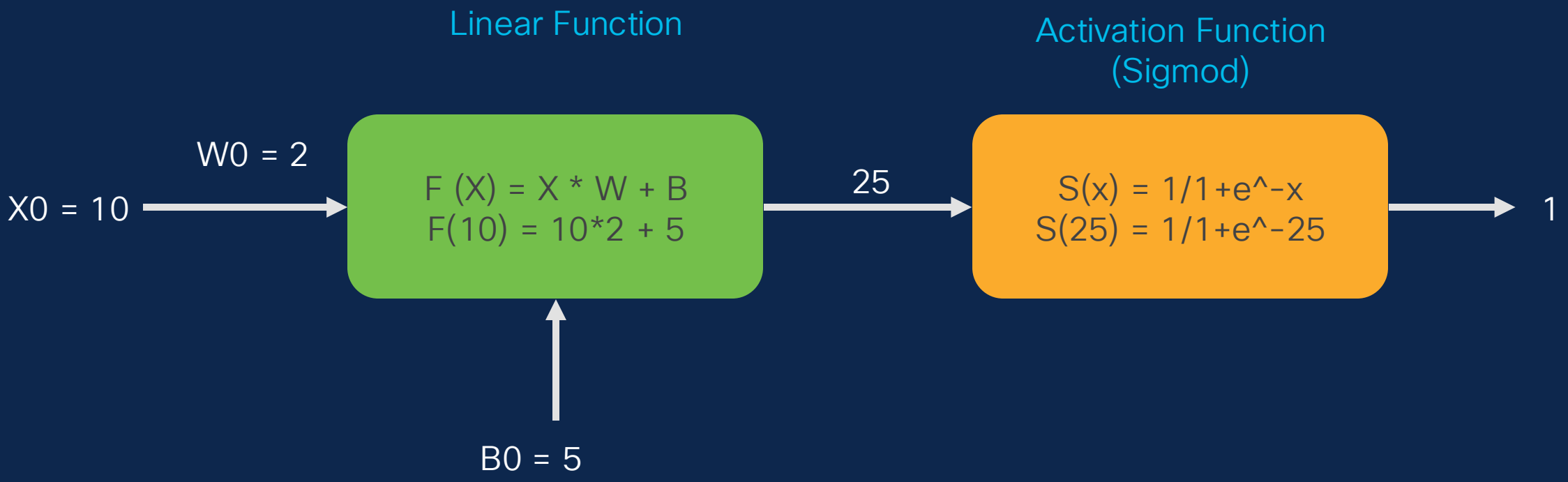
Neural Networks

- A neuron takes input (firings) from other neurons
- The inputs are summed in a weighted manner
- If it receives enough input, it “fires”
- Learning is through a modification of the weights



© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Simple Artificial Neuron Example

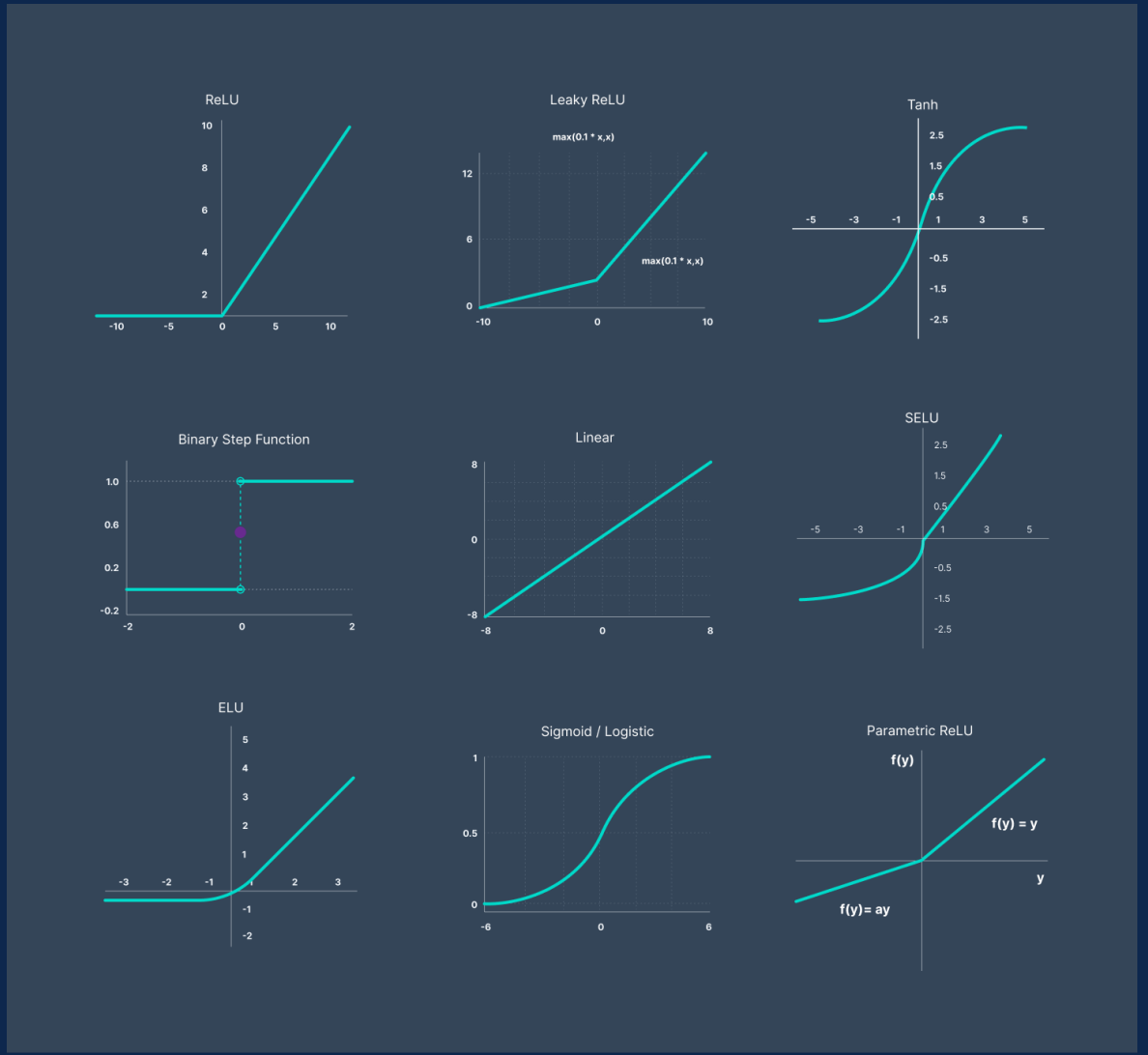


- The following function represent the neuro: Sigmod (X0*W0+B0)
- W0 and B0 are called trainable parameters.

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Activation Functions

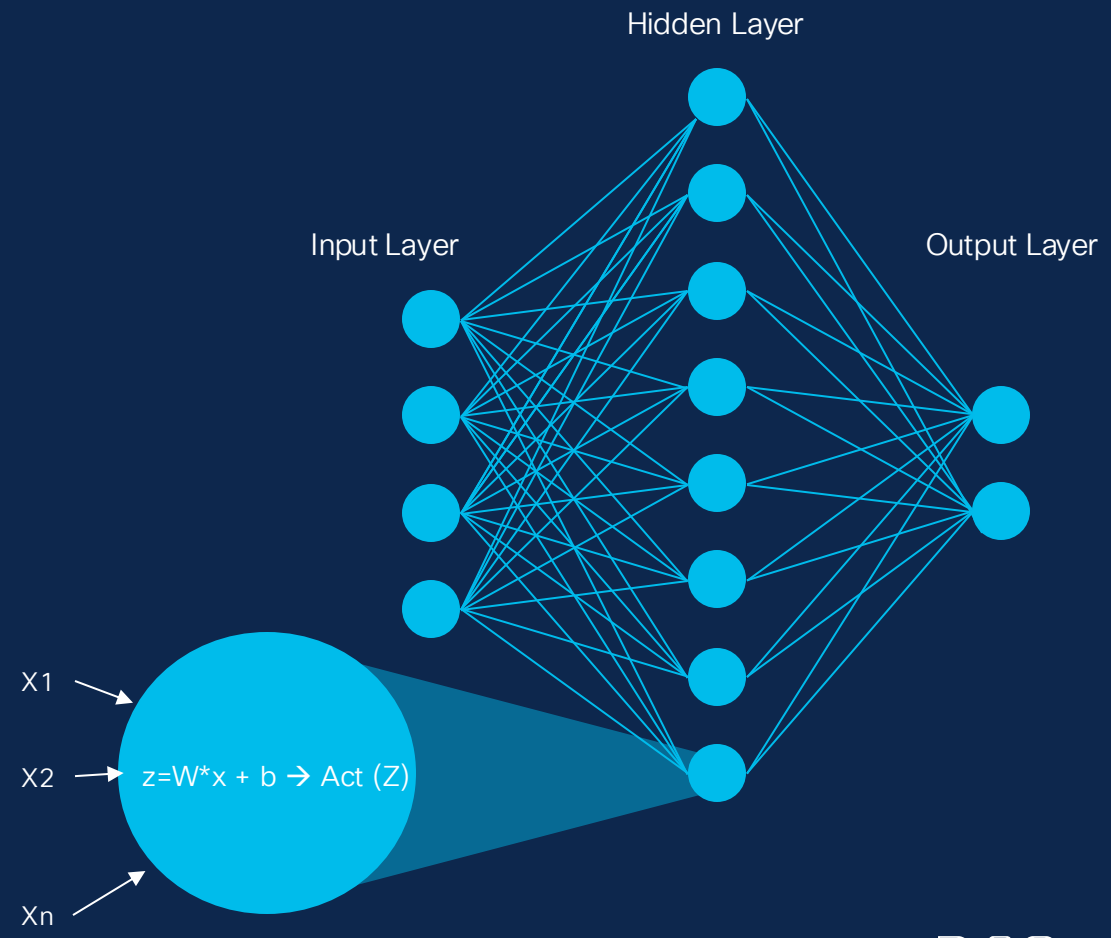
An Activation Function decides whether a neuron should be activated or not. This means that it will decide whether the neuron's input to the network is important or not in the process of prediction using simpler mathematical operations.



© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Shallow Neural Networks

- A Shallow network is a ANN which has only one or two hidden layers.
- The hidden layers are between the input and output layer.
- These Networks are used for simple tasks such as:
  - Linear egression
  - Binary classification
  - Feature Extraction (Embeddings)

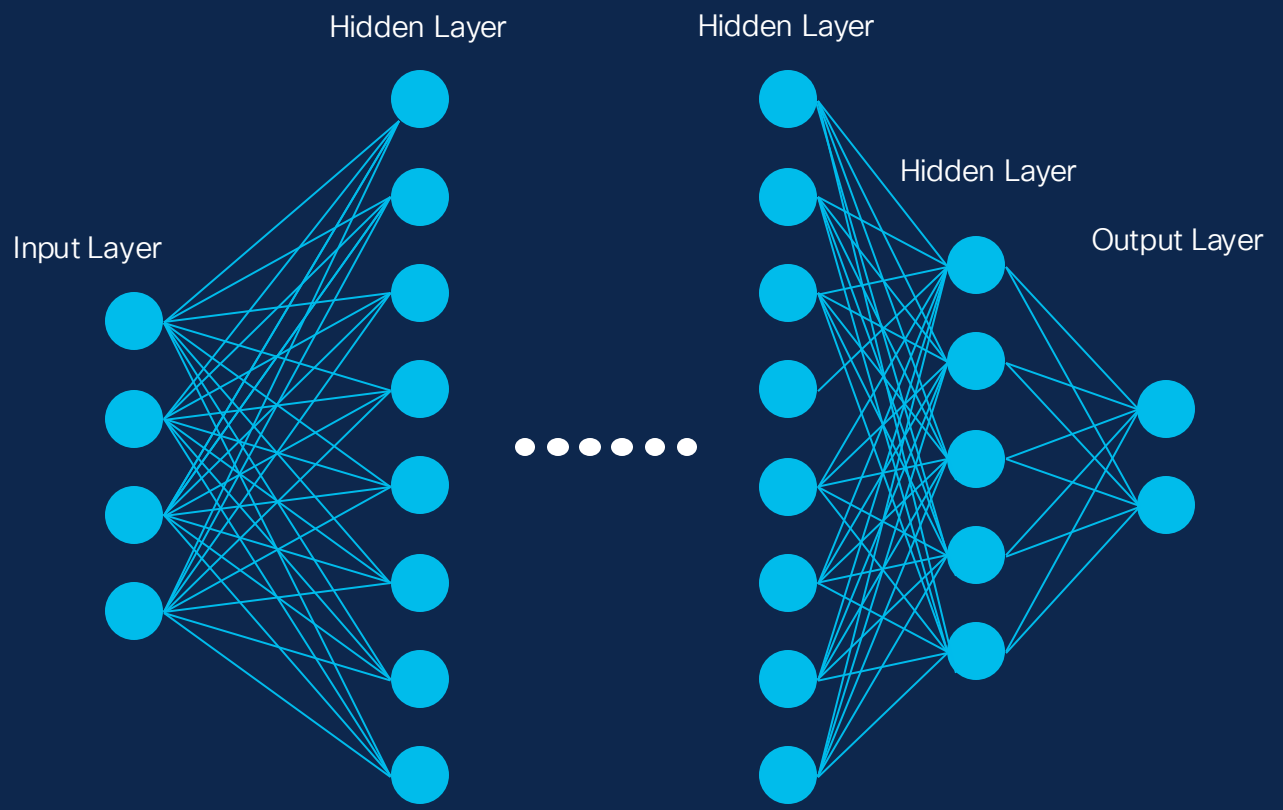


© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Cisco Confidential

Deep Neural Networks

- A Deep Neural Network can have many hidden layers.
- Used for more complex tasks like image segmentation or natural language processing.

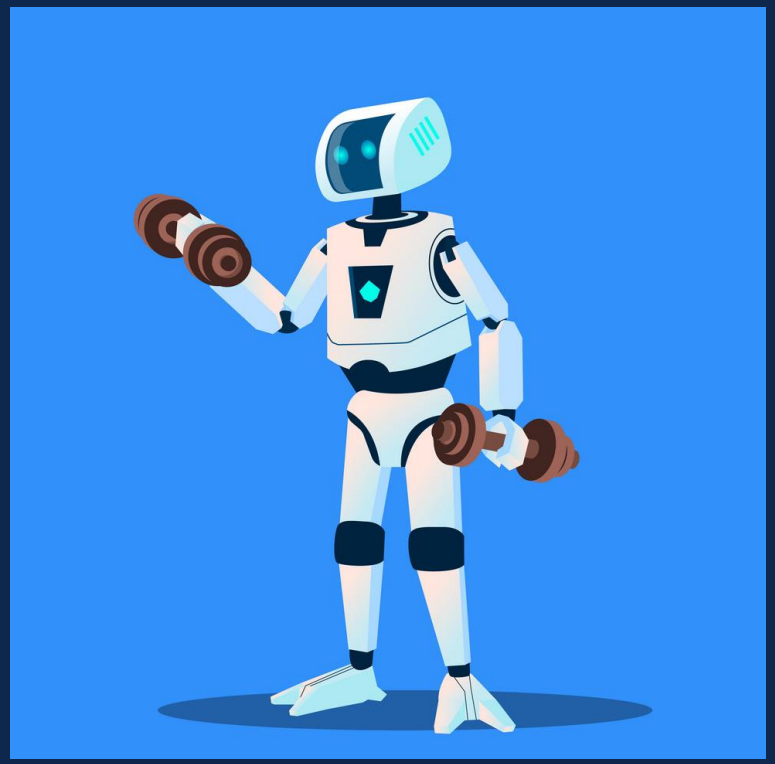


© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Cisco Confidential

Neural Network Training

- The learning (training) process is an iterative process in which the calculations through each layer with the goal of updating the network parameters.
- At the start we initialize all the parameters randomly.
- Then at each step:
  - We pass the input through the network (forward pass).
  - Measure its performance using a loss function
  - Optimize the network parameters to improve the loss (back propagation)



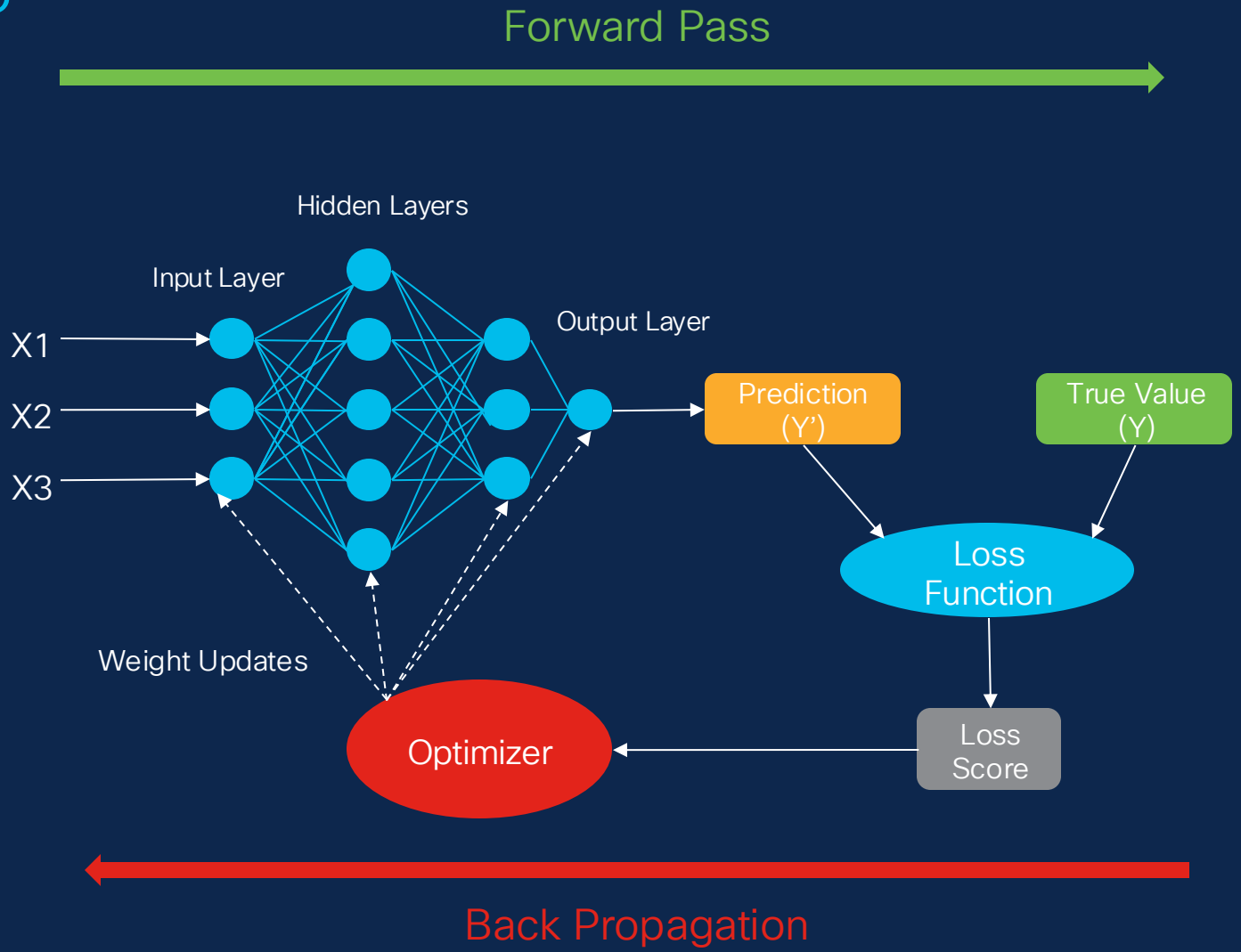
© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Cisco Confidential

Neural Network Training

Training a model involves the following steps:

- Initialize the weights:
  - This can be as simple as picking random values of the weights
- Training Loop, in which we do:
  1. Pass X number of training samples into the model
  2. Calculating the error in the model prediction using loss a function (MSE for example).
  3. Computing the gradients using back propagation.
  4. Use an optimizer to update the model weight.
  5. Repeat until done



© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Cisco Confidential

ANN Training Basic Terms

- **Batch Size:** The number of data points (like images or texts) the network looks at in one iteration. If you read 10 pages of a book at a time, then 10 pages is your batch size.
- **Epoch:** One complete cycle where the network has seen all the training data once. If you have 1,000 images and you've shown them all to the network, that's one epoch.
- **Iteration:** One step in the training process where the network updates its knowledge based on a subset of the data.
- **Forward Pass:** The process of inputting data into the network and getting an output. Think of it as asking the network a question and getting an answer.
- **Loss Function:** A measure of how wrong the network's predictions are. It's like a score where lower is better.

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Cisco Confidential

ANN Training Basic Terms

- **Gradients:** Directions that tell the network how to adjust its settings to get a better score. Imagine it as feedback on which way to go to improve.
- **Back Propagation:** A method used to calculate the gradients. It's like finding out which parts of the network contributed most to the error and how to fix them.
- **Optimizer:** Once we have the gradients from back propagation, we need a strategy or method to adjust the weights of the network. That's where the optimizer comes in.
- **Learning Rate:** How much the network changes its settings in each iteration. It's like deciding how big of a step to take when trying to reach a destination.

© 2023 Cisco and/or its affiliates. All rights reserved. Cisco Confidential


CISCO