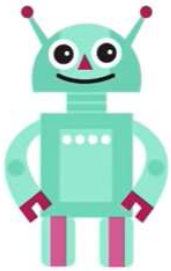
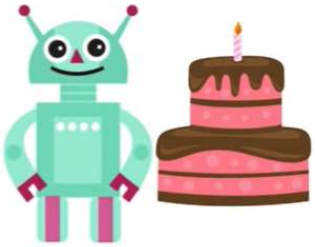
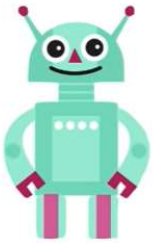
INTRODUCTION TO DEEP LEARNING:



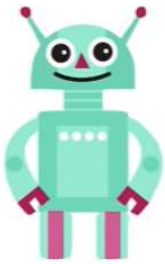
- turn right
- go 10 steps
- turn left
- go 4 steps
- grab cake



- calculate distance to cake
- move around, pick a direction where the distance decreases

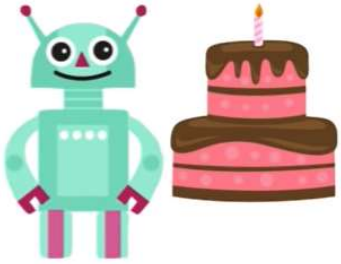


- calculate distance to cake
- move around, pick a direction where the distance decreases
- take a step in that direction

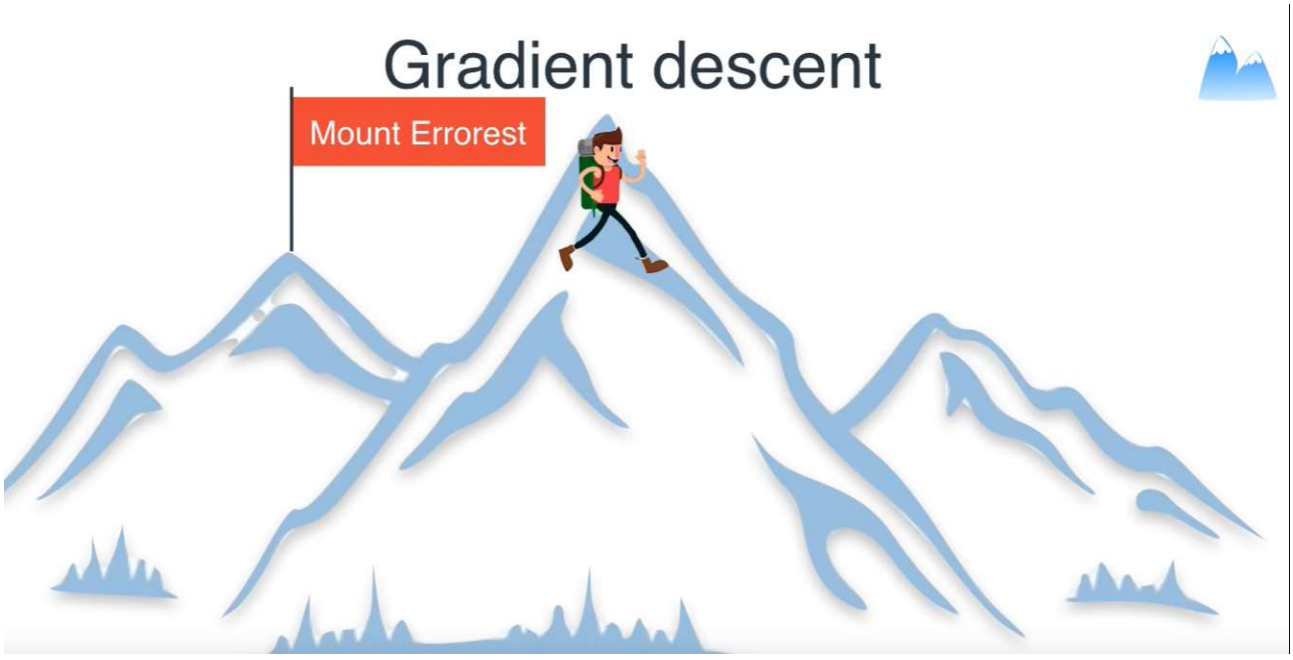


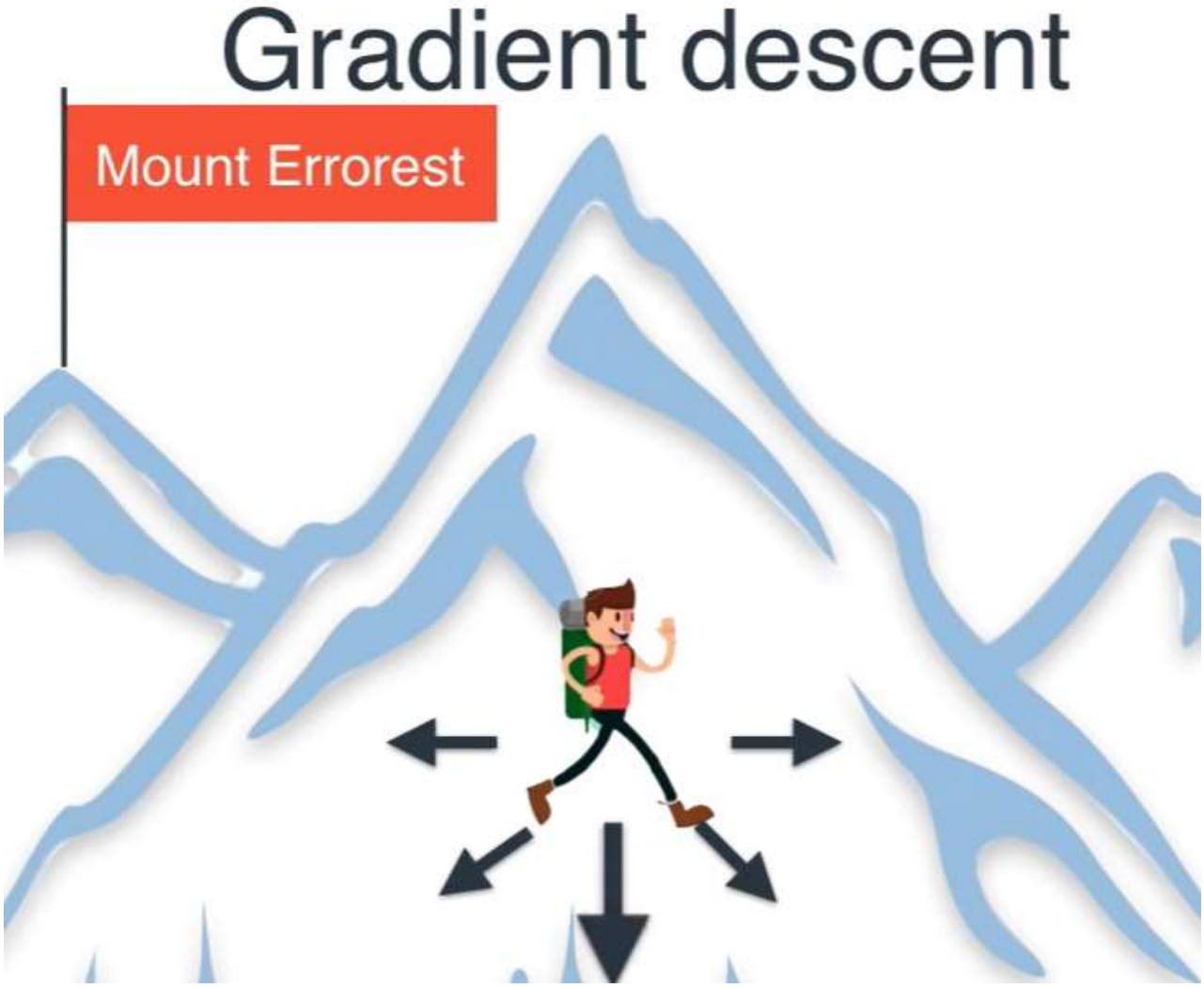
4.5

- calculate distance to cake
- move around, pick a direction where the distance decreases
- take a step in that direction
- repeat



Gradient descent





WE TOOK THE GRADIENT – THE DERIVATIVE

Gradient descent

Get cake

Minimize distance to cake

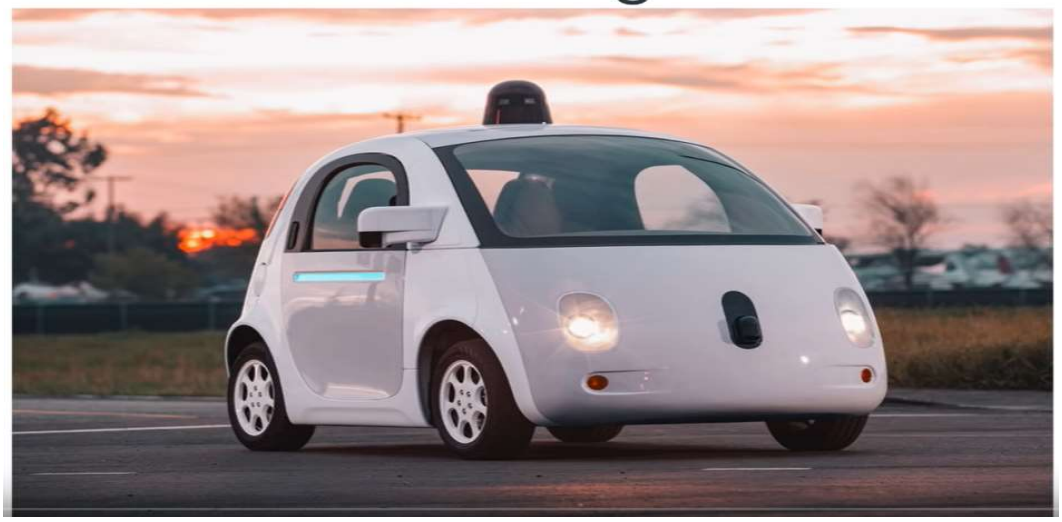
Descend from mountain

Minimize height

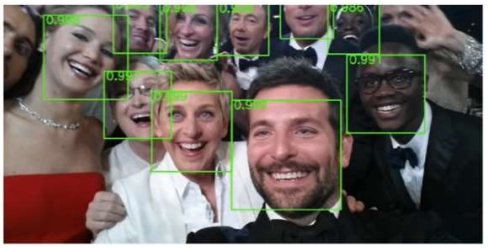
Solve any problem

Minimize error

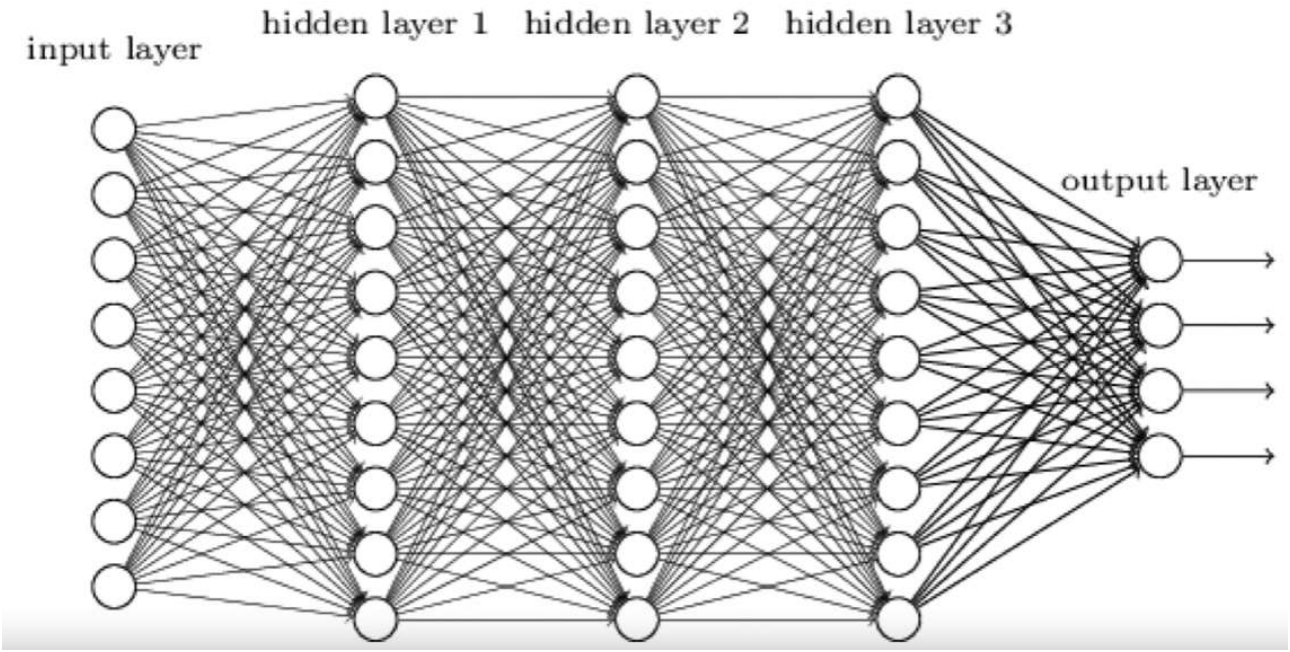
Self Driving Car



Many more things

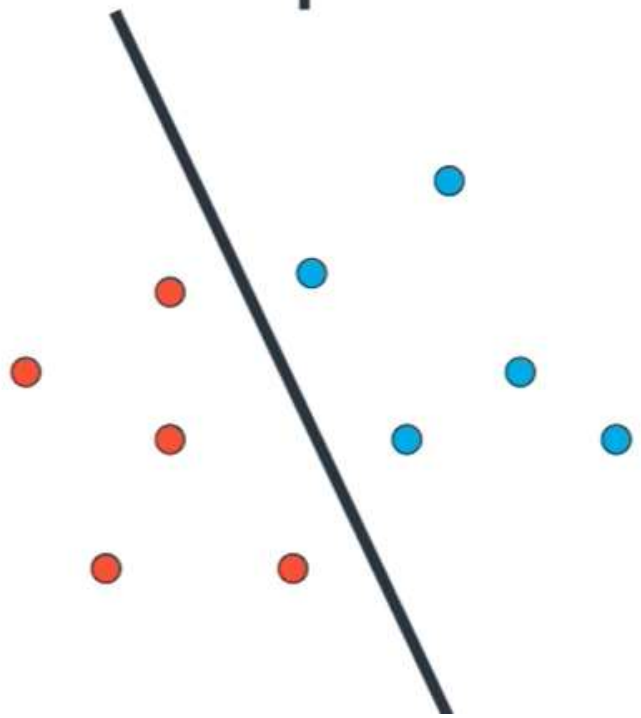


Neural Networks

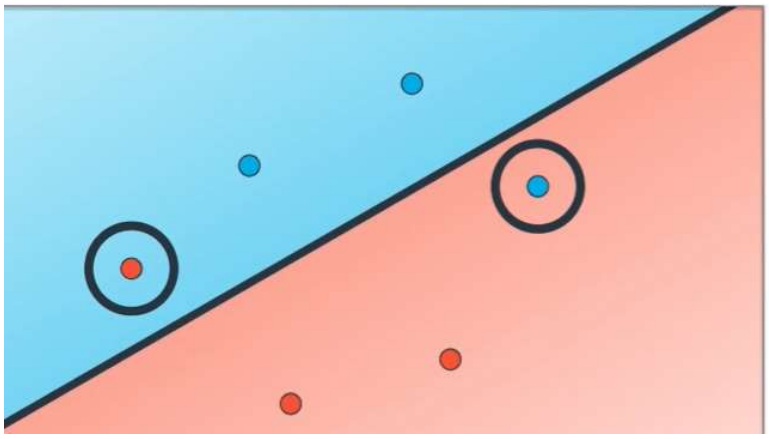


Neural Networks

Goal: Split Data

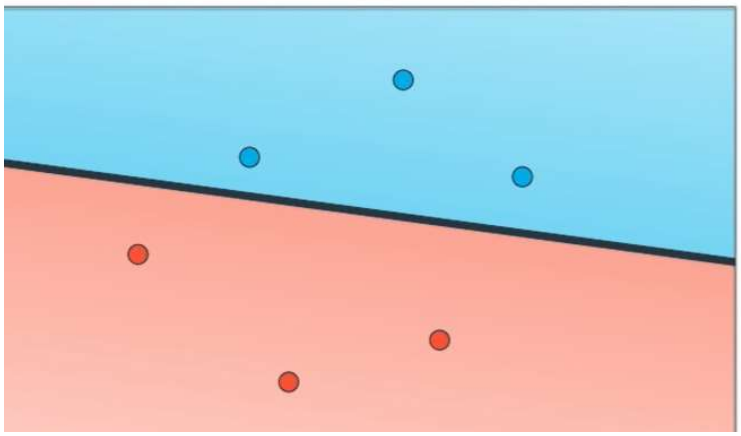


Goal: Split Data



2 errors

Goal: Split Data



0 errors

Hot!

Gradient descent

Get cake

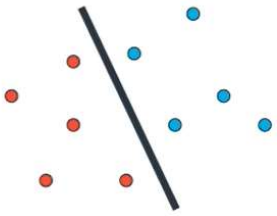
Distance to cake

continuous function

Descend from mountain

Height

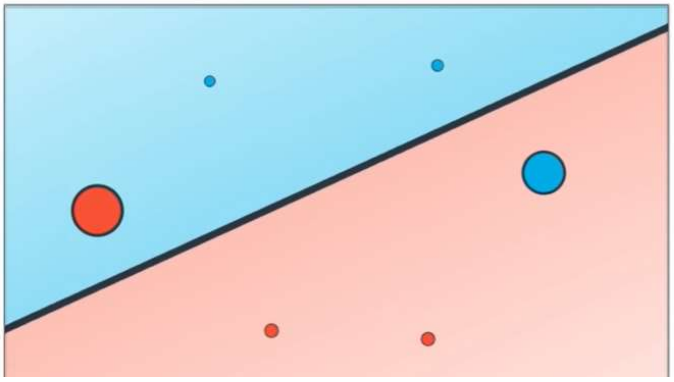
continuous function

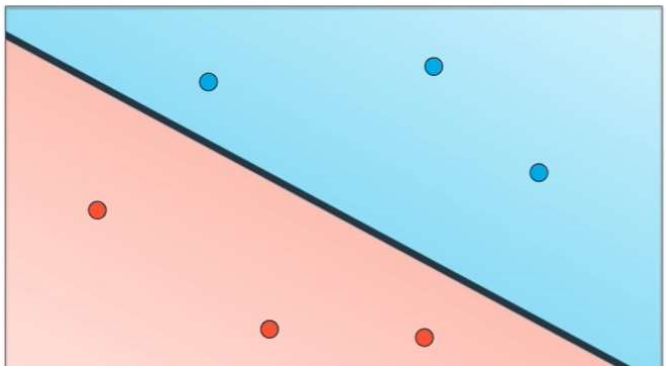


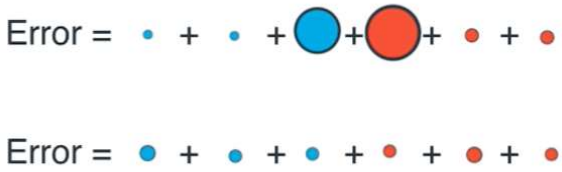
Split data

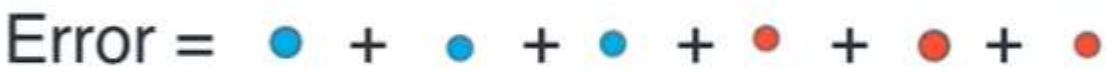
Number of errors

discrete function





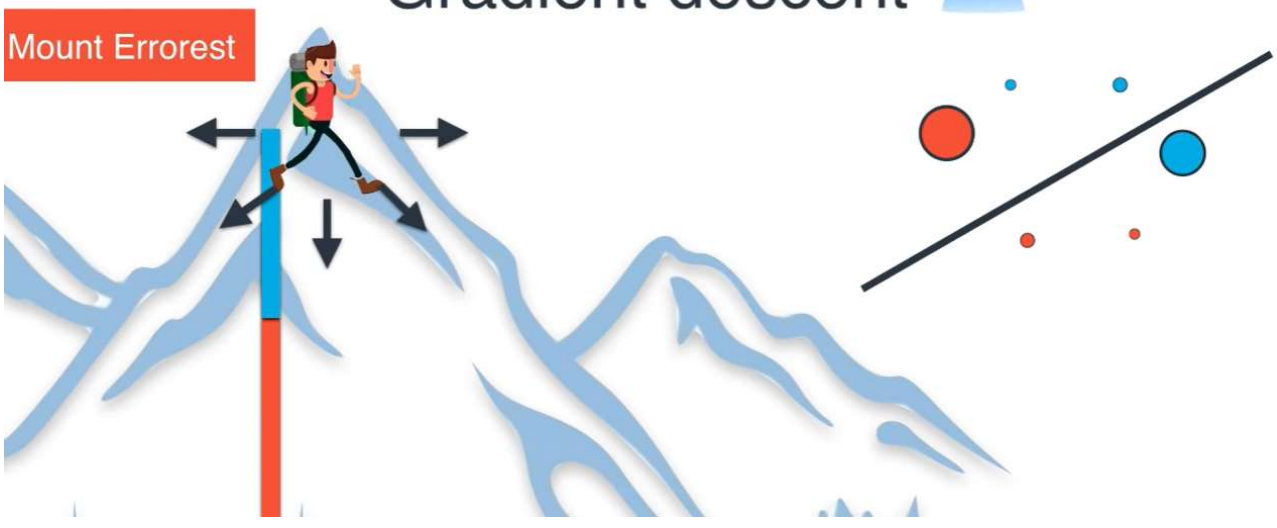




Minimize error

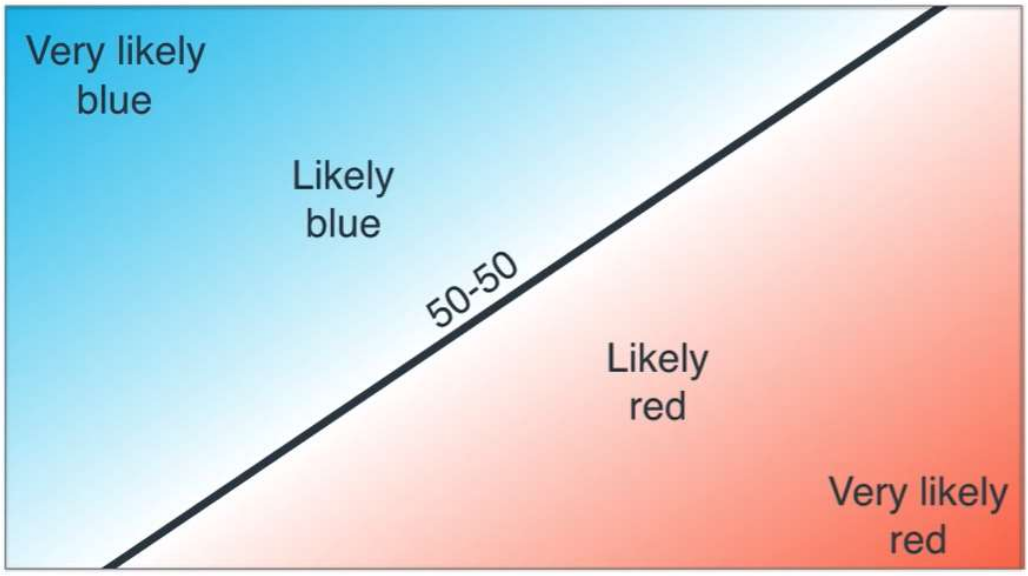
Gradient descent

Gradient descent

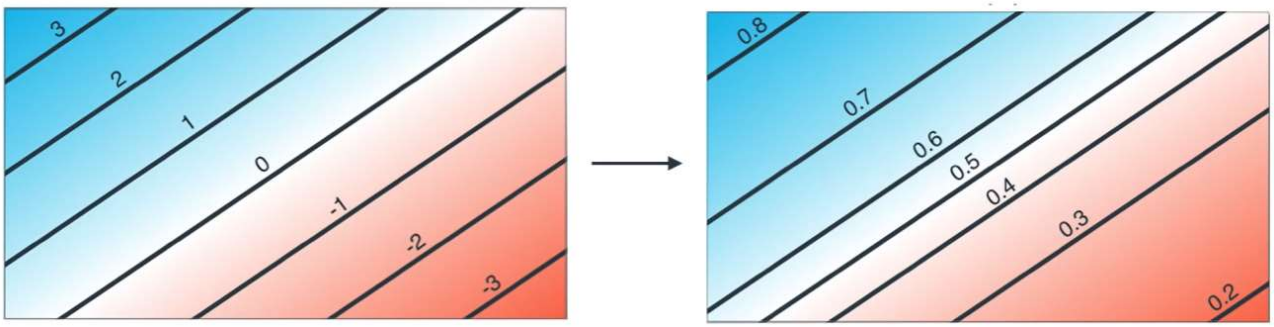


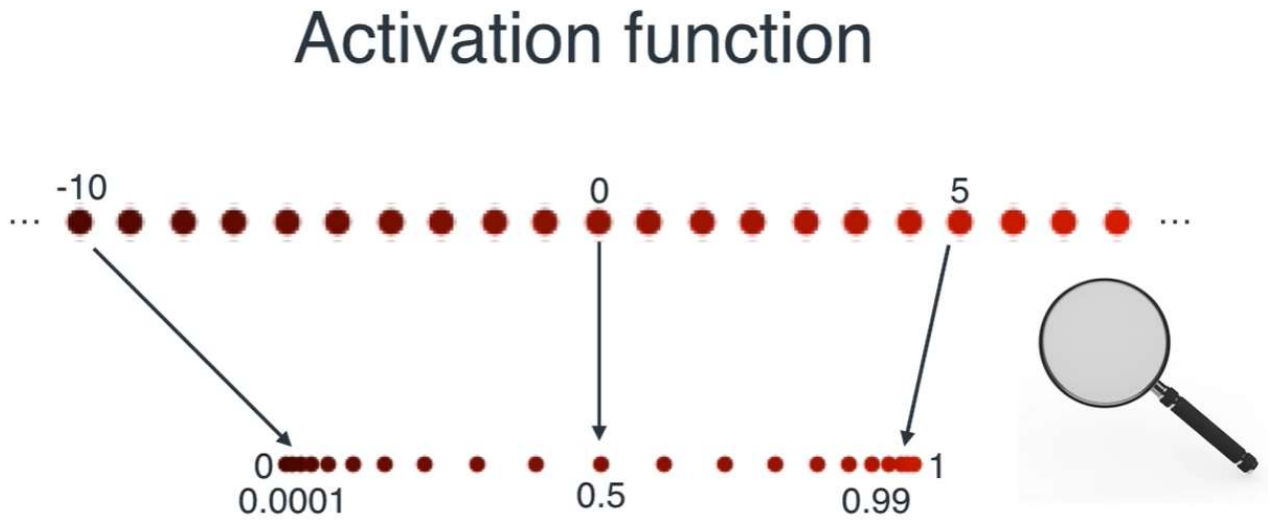
Gradient descent

Probability

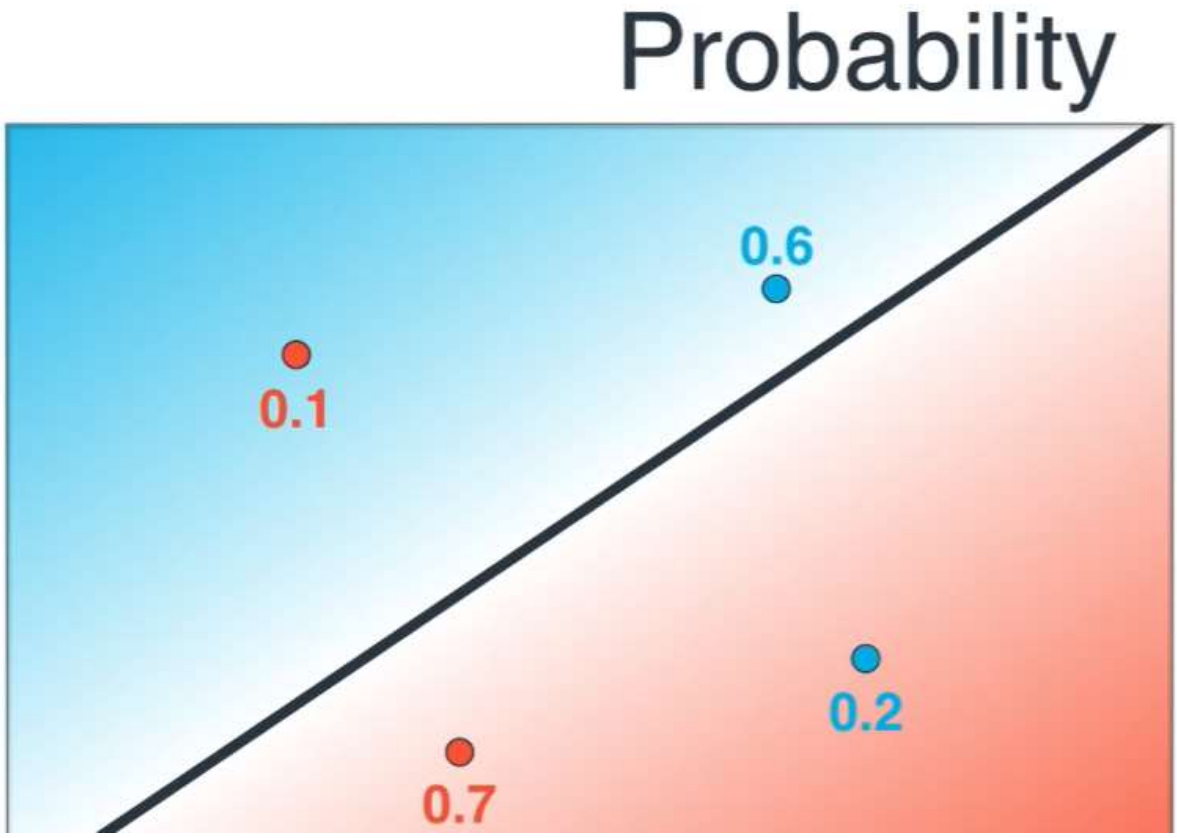


Probability

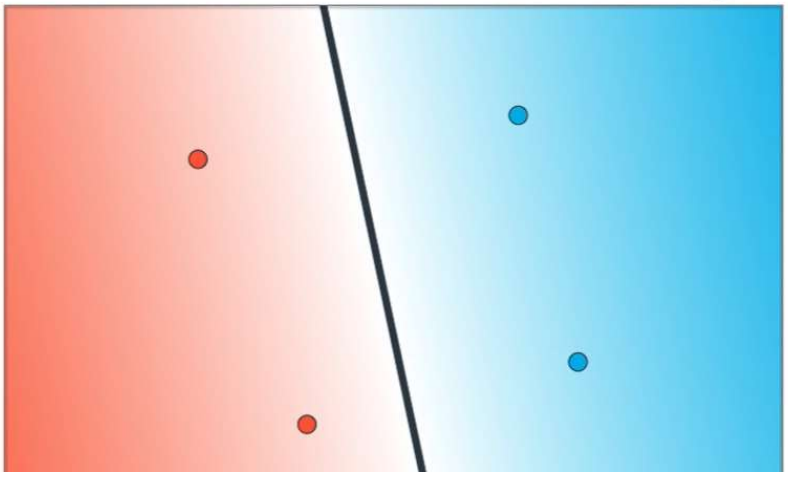




IF WE THINK THESE ARE INDIPENDENT EVENTS, THE PROBABILITY ALL FOUR HAPPENING, IS THE PRODUCT



Probability

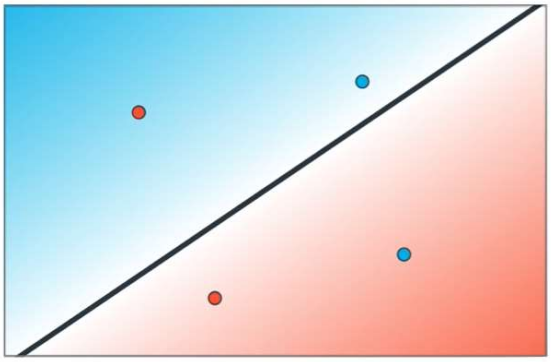


0.6*0.2*0.1*0.7 = 0.0084

0.7*0.9*0.8*0.6 = 0.3024

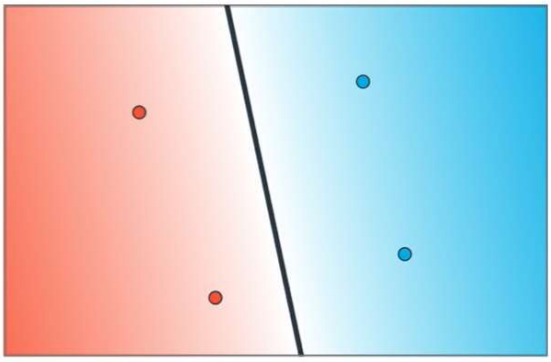
Maximum Likelihood

Error function



0.6*0.2*0.1*0.7 = 0.0084

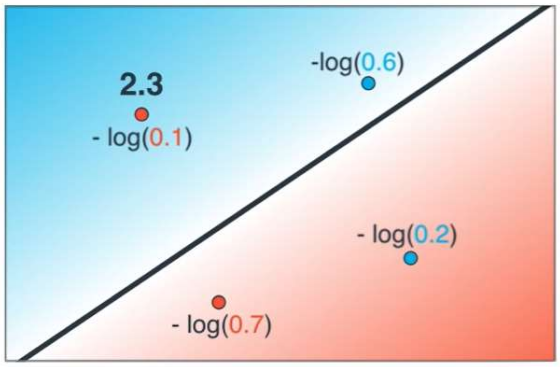
-log(0.6) - log(0.2) - log(0.1) - log(0.7) = 4.8



0.7*0.9*0.8*0.6 = 0.3024

-log(0.7) - log(0.9) - log(0.8) - log(0.6) = 1.2

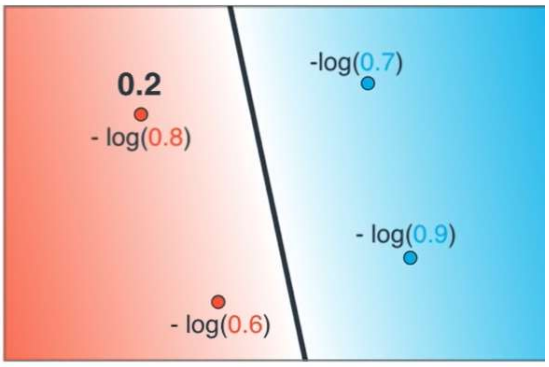
Error function



0.6*0.2*0.1*0.7 = 0.0084

-log(0.6) - log(0.2) - log(0.1) - log(0.7) = 4.8

2.3



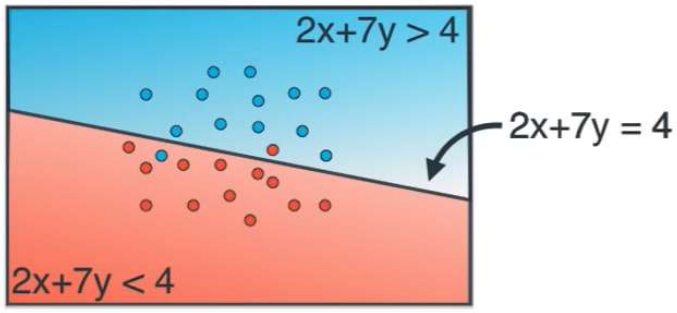
0.7*0.9*0.8*0.6 = 0.3024

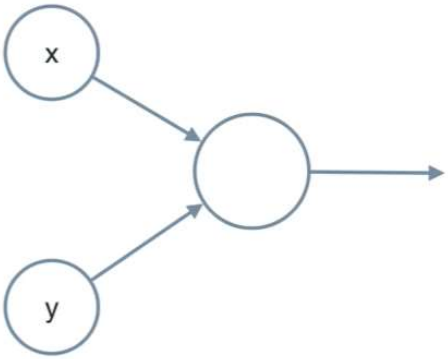
-log(0.7) - log(0.9) - log(0.8) - log(0.6) = 1.2

0.2

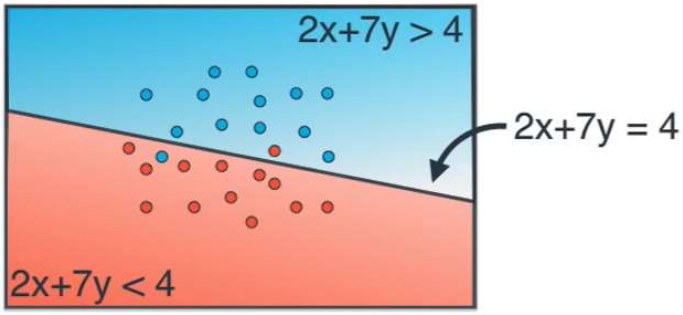
ERROR FUNCTION AS A PENALTY FOR EVERY POINT

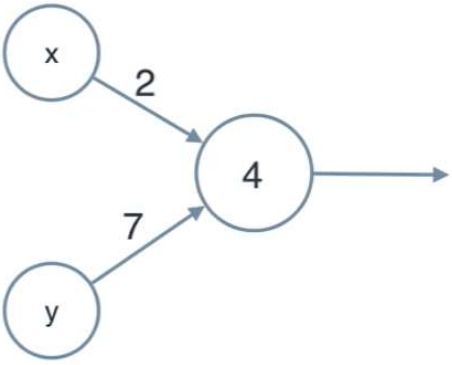
Neuron



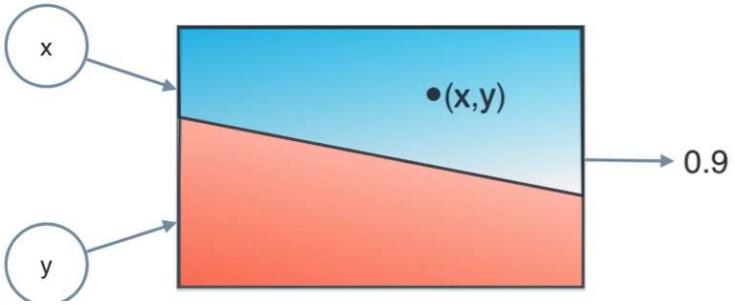


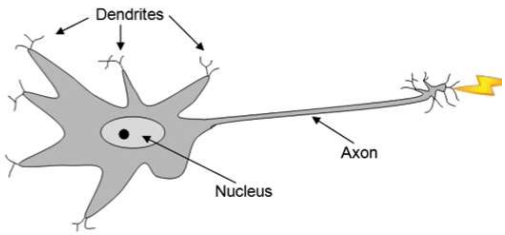
Neuron



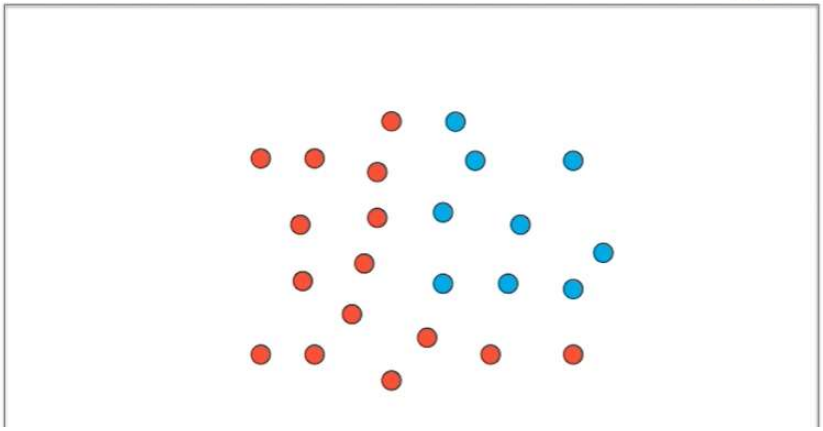


Neuron

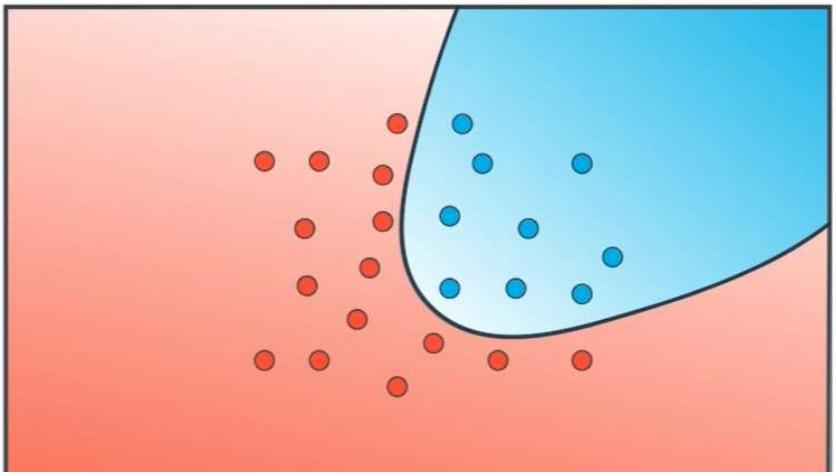




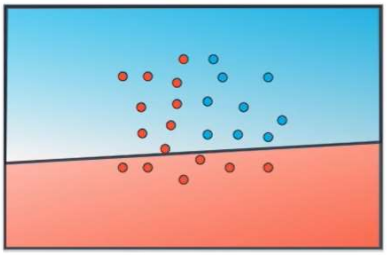
Non-linear regions

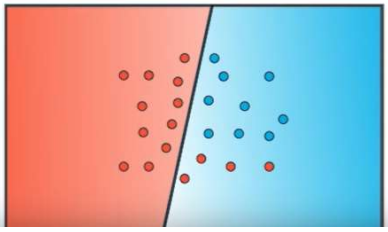


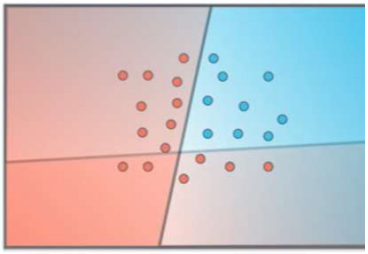
Non-linear regions

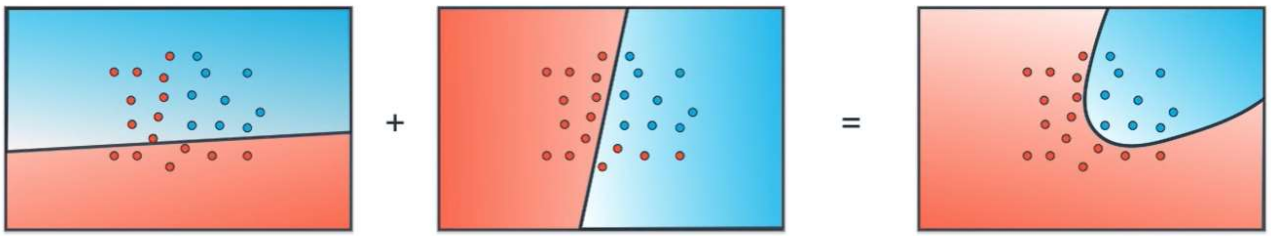


Combining Regions

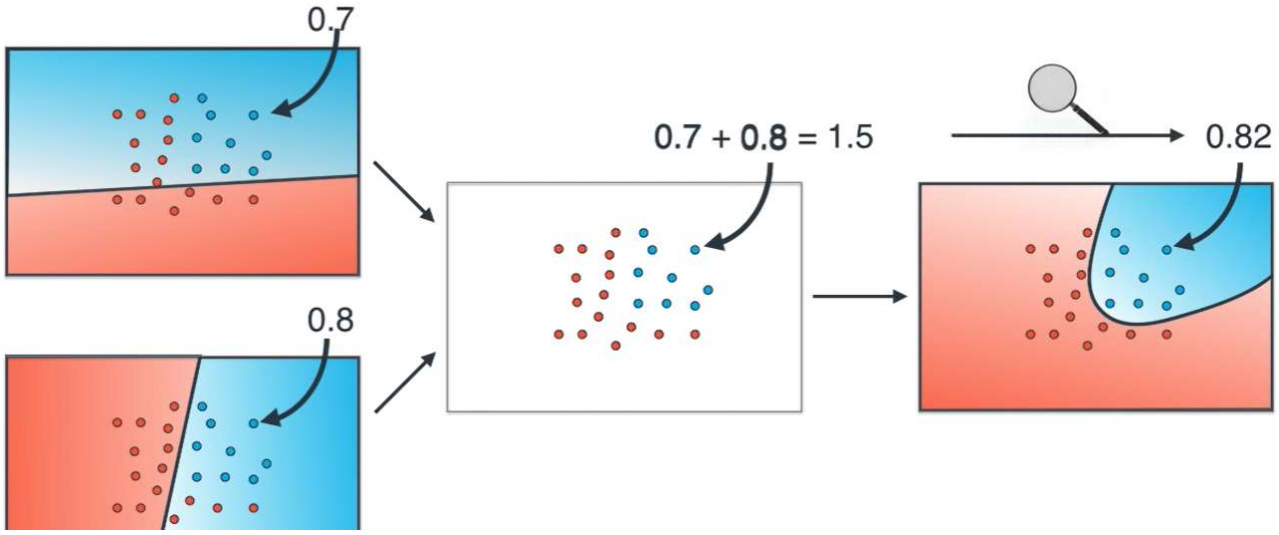




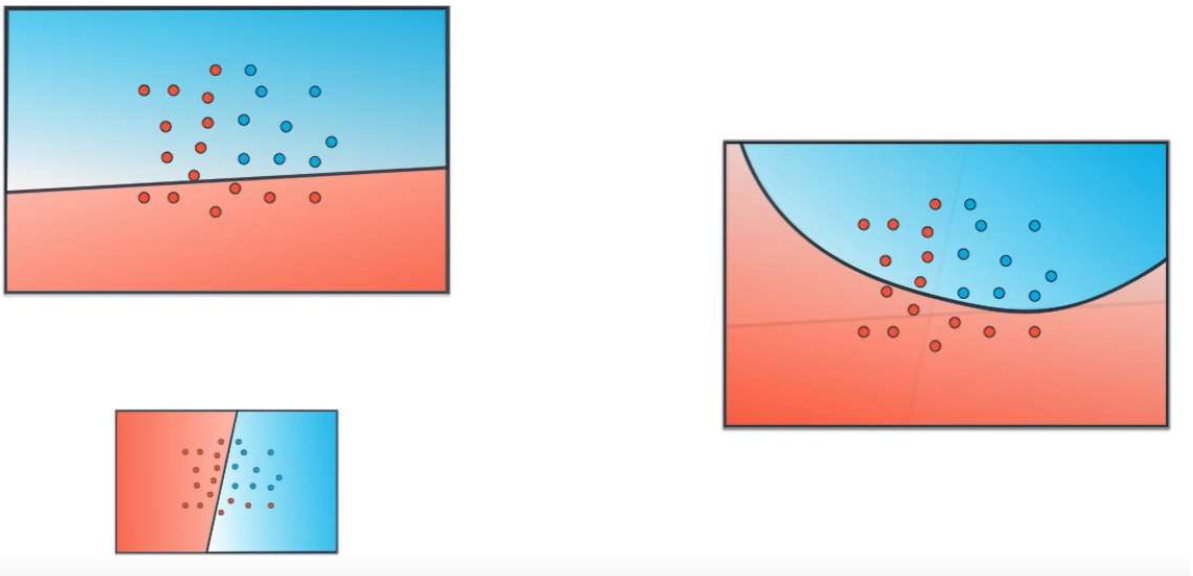


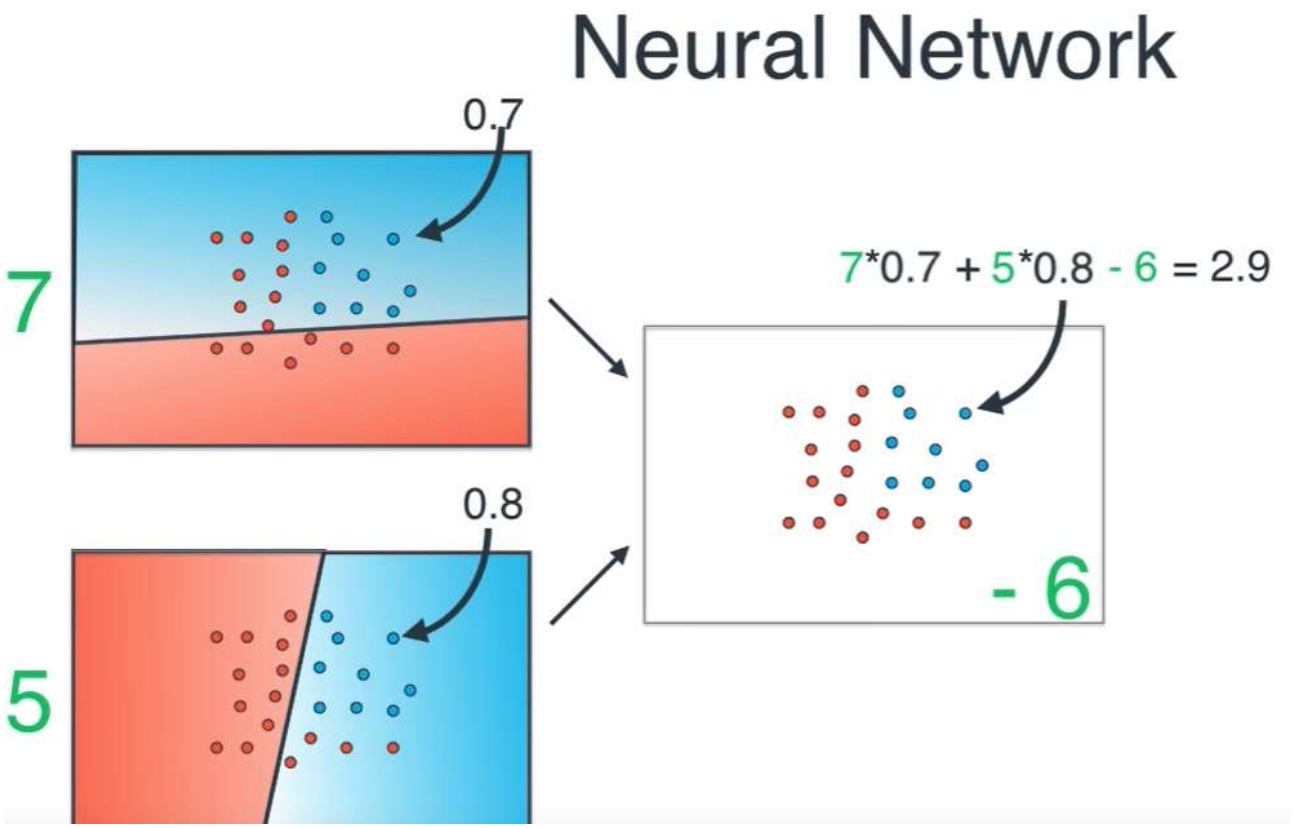


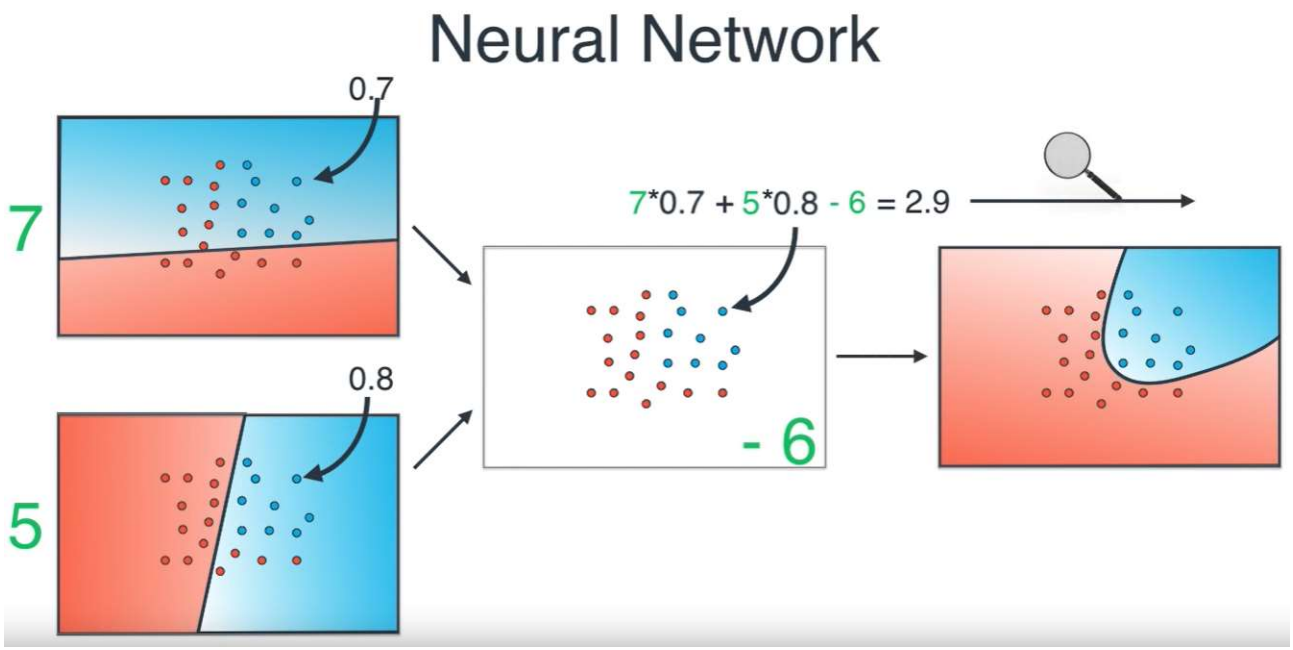
Neural Network



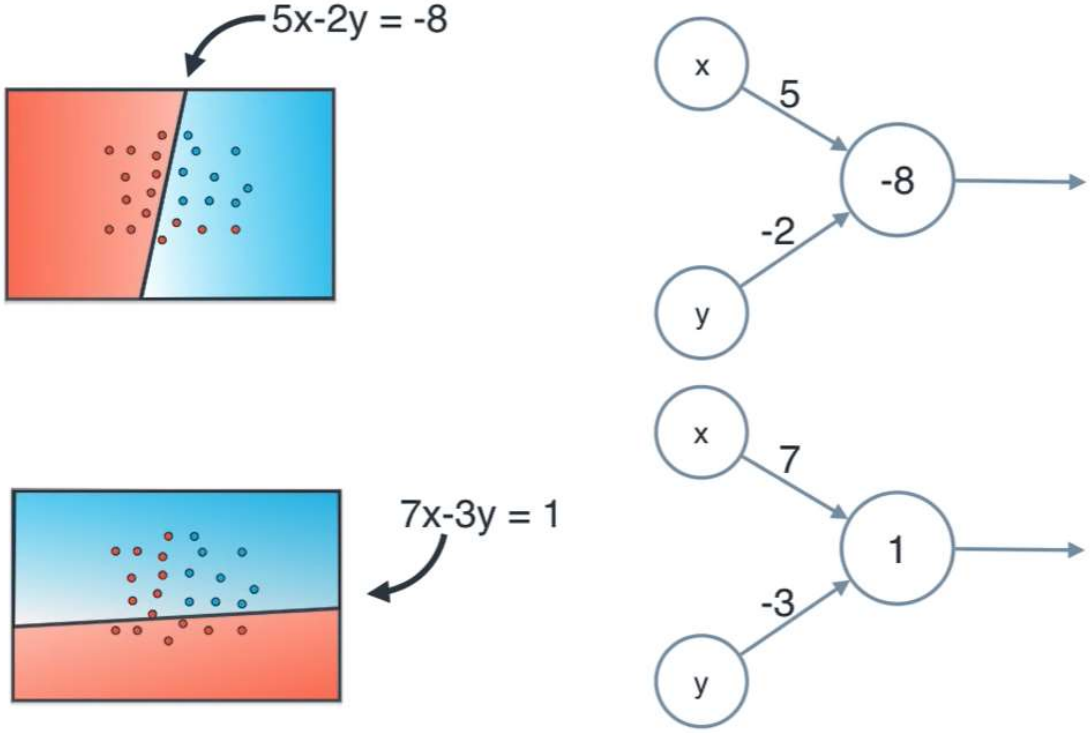
Combining Regions



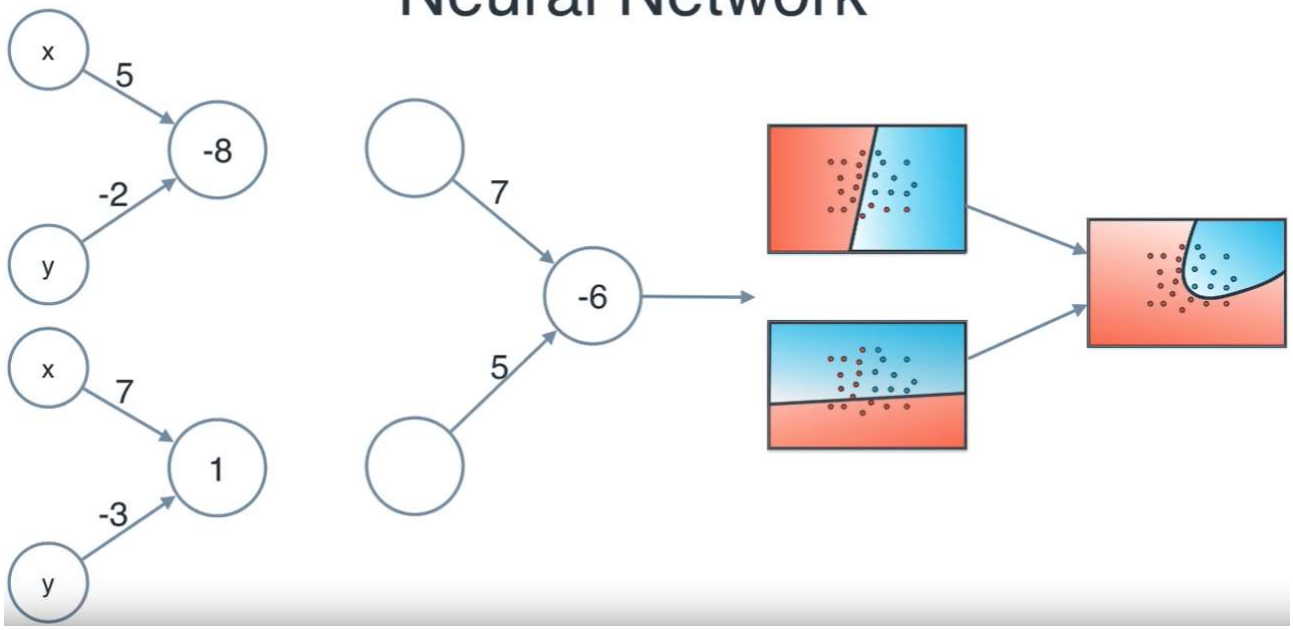

Neural Network
0.7
7
7*0.7 + 5*0.8 - 6 = 2.9
0.8
- 6
5


Neural Network
0.7
7
7*0.7 + 5*0.8 - 6 = 2.9
0.8
- 6
5

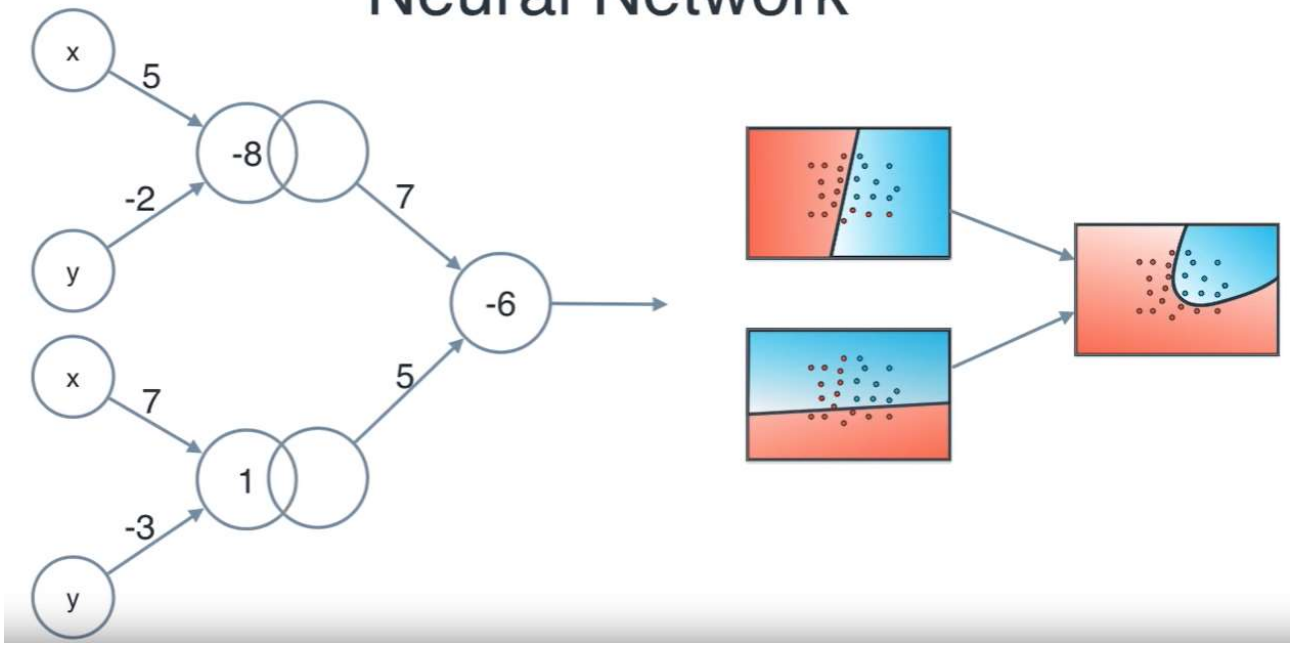
Neural Network



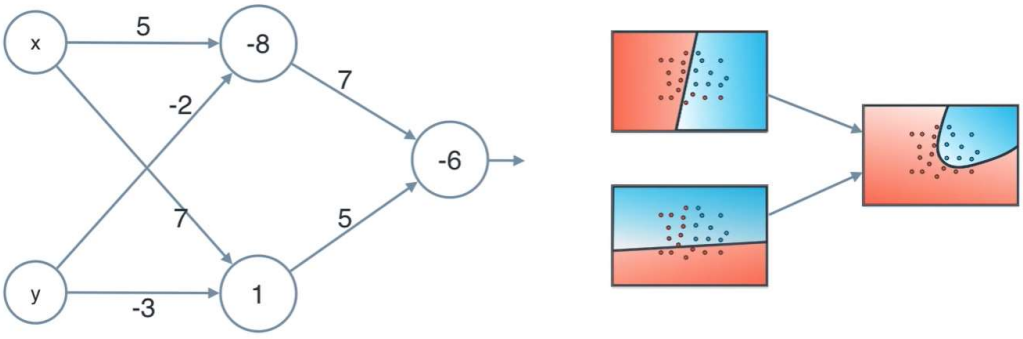
Neural Network



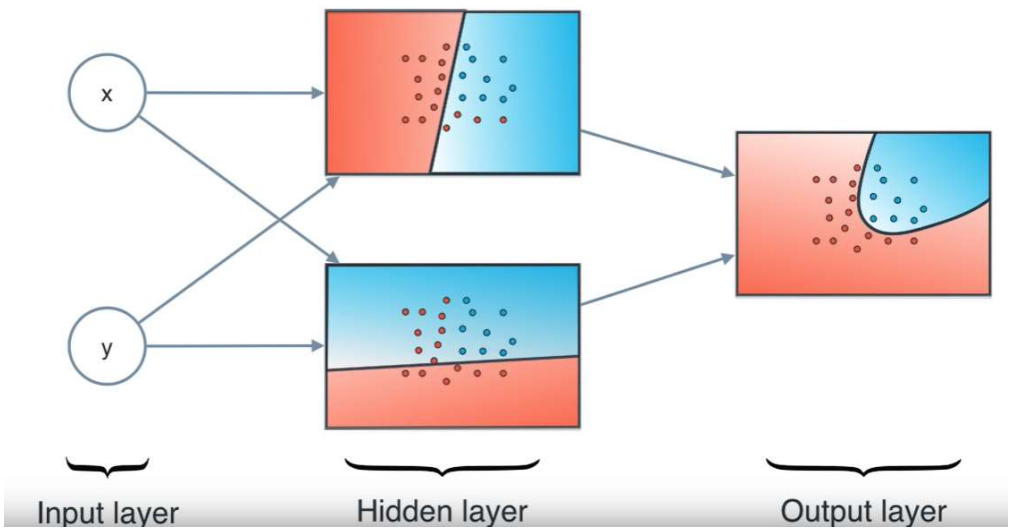
Neural Network

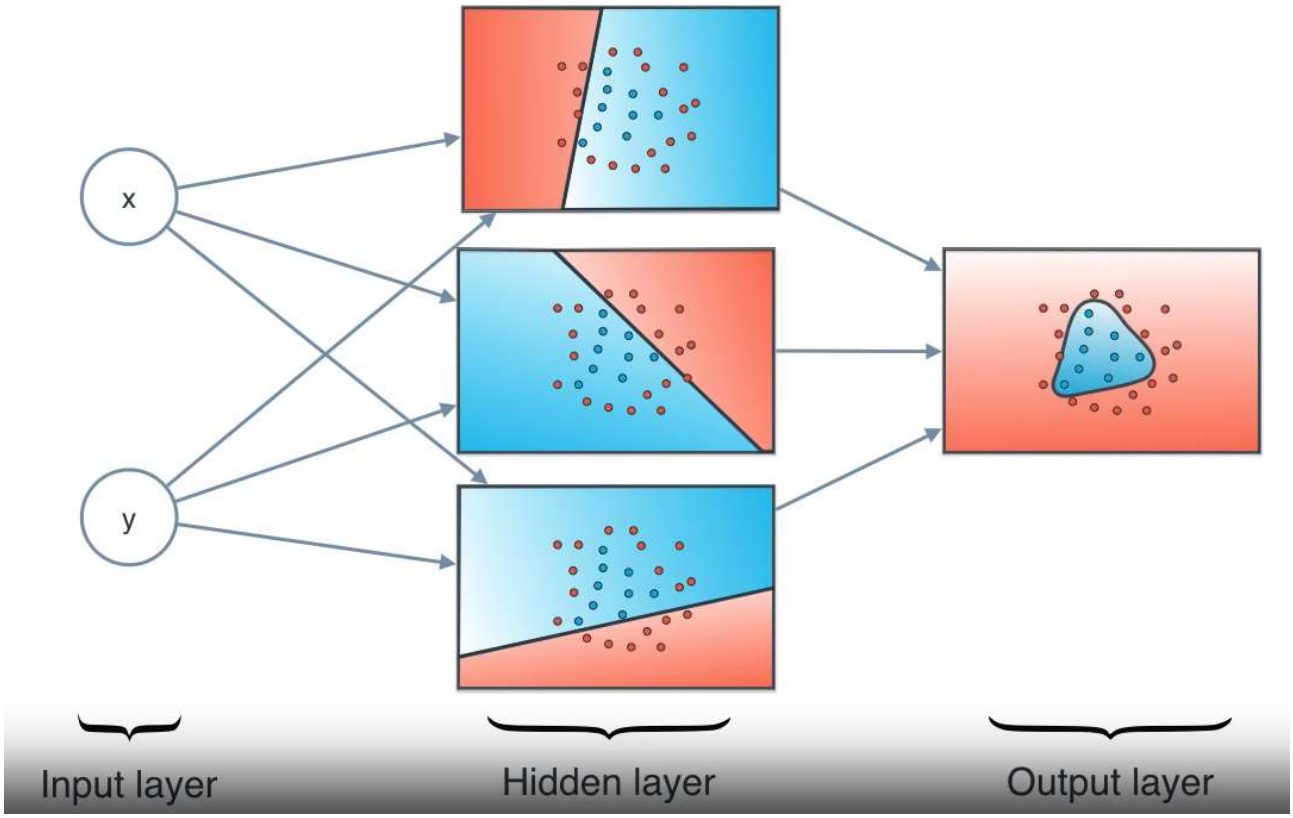


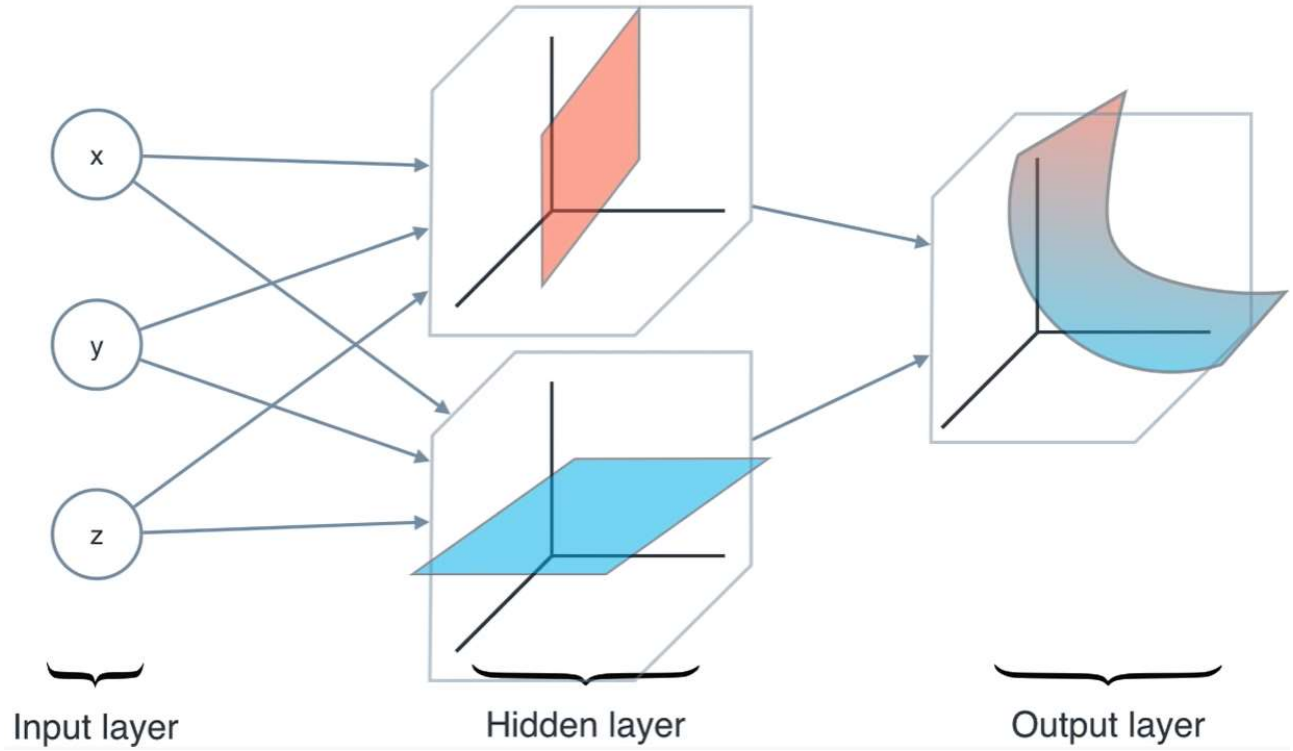
Neural Network



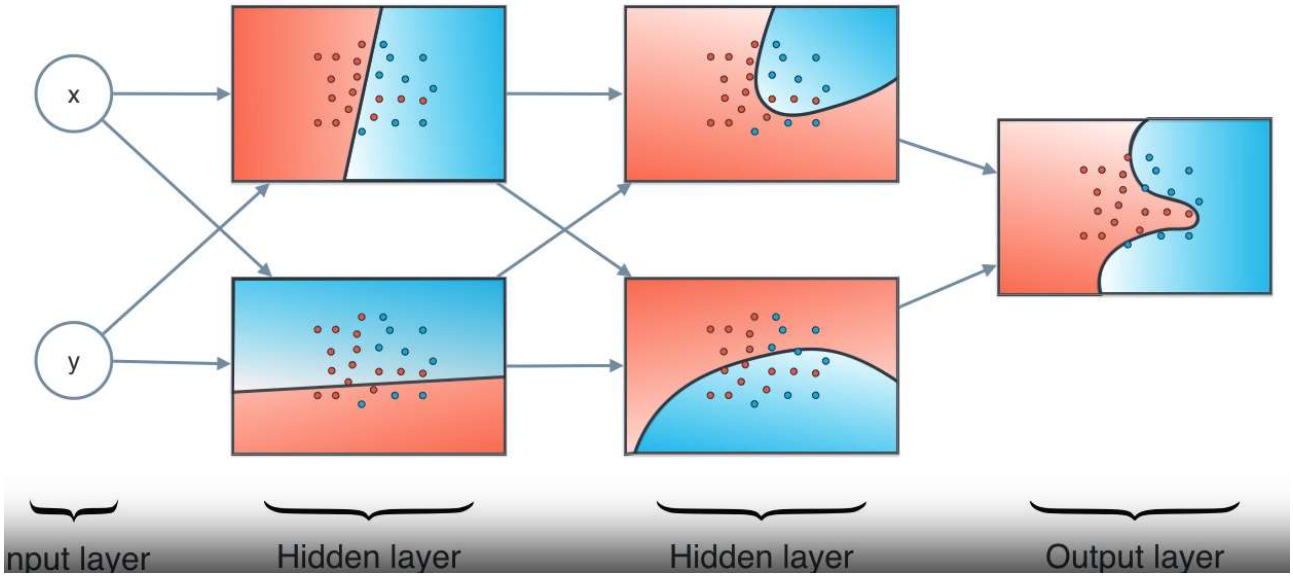
Neural Network




x
y
Input layer
Hidden layer
Output layer


x
y
z
Input layer
Hidden layer
Output layer

Deep Neural Network



Self Driving Car