

Deep Learning and Neural Network

Marie Beurton-Aimar, LE Van Linh

December 1, 2020

Deep Learning

Deep learning is a class of machine learning that:

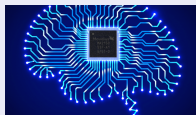
- use a cascade of multiple layers for extracting and transforming the features.
- learn in supervised or unsupervised mode
- learn with different levels of presentation



[ZDNet, 2018]



[Siecle Digital, 2016]

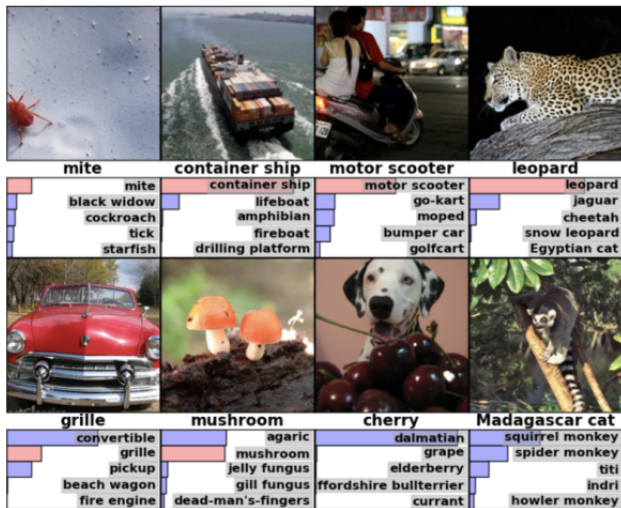


[Silicon ANGLE, 2016]



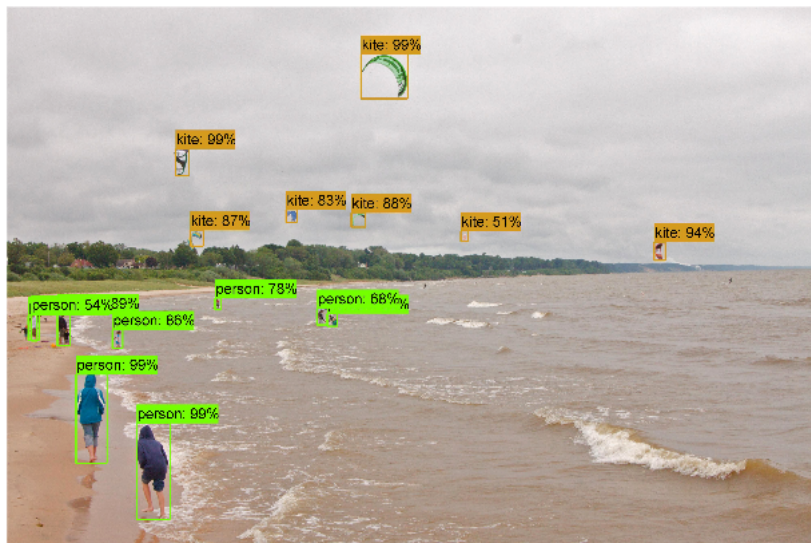
ROSS Intelligence, 2018

Deep Learning - Images classification



[Machine Learning Mastery, 2016]

Deep Learning - Object detection



[Medium, 2018]

Deep Learning - Machine translation



Traduction

Anglais

Français

Vietnamien

Détecter la langue



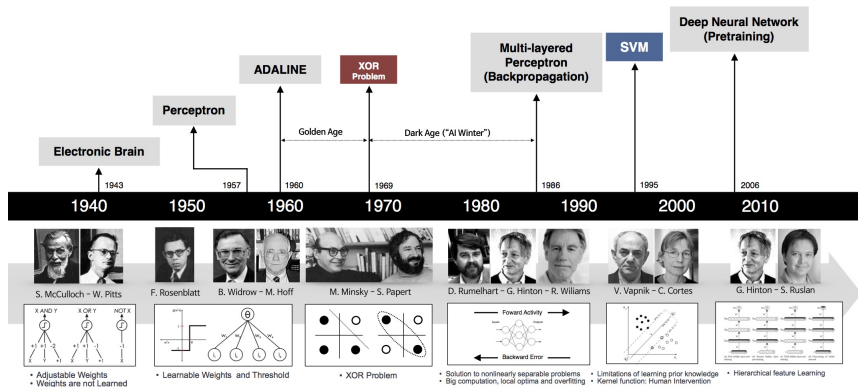
bonjour
Je m'appelle Linh



Deep Learning - Speech recognition



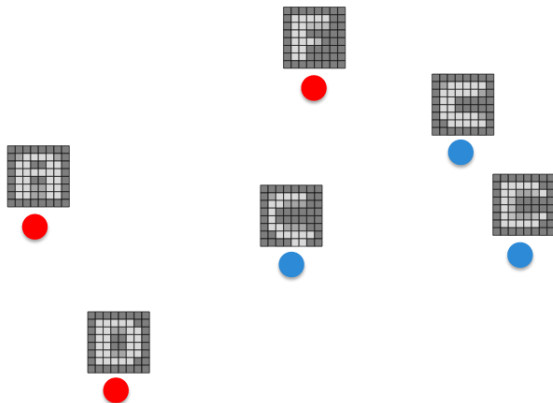
History



[beamandrew, 2017]

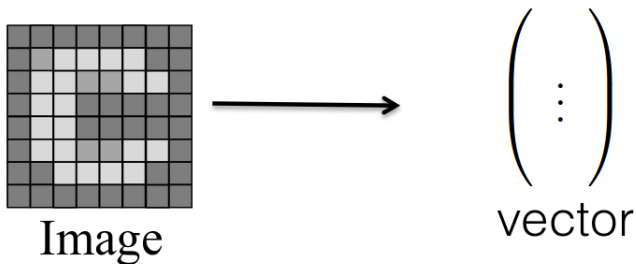
Perceptron idea

- Collect data for training: includes positive and negative examples



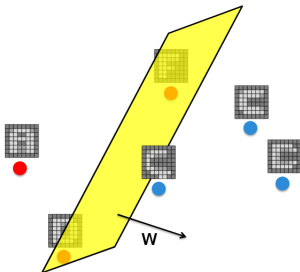
Perceptron idea

- Represent data as vectors:



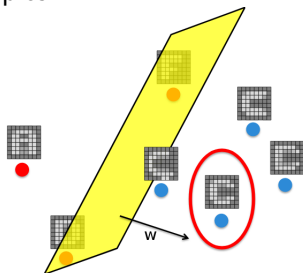
Perceptron idea

- Function: $f(x) = w^T x + b$
- Train the function on training data: find w and b so that
 - If sample x is positive, $f(x)$ is positive
 - If sample x is negative, $f(x)$ is negative

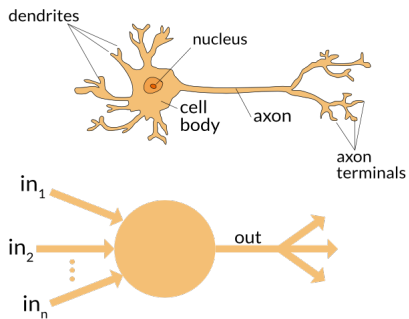


Perceptron idea

- Testing on new examples

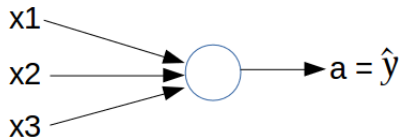


The Inspiration for a Perceptron

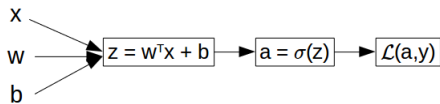
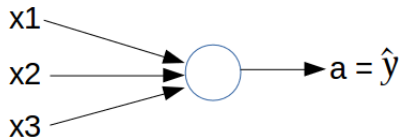


[appliedgo, 2016]

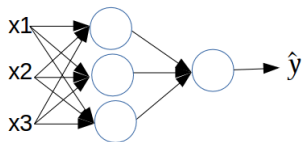
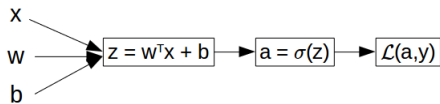
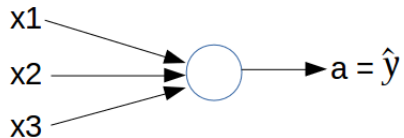
What is a neural network?



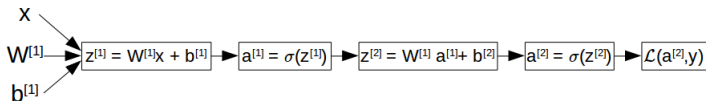
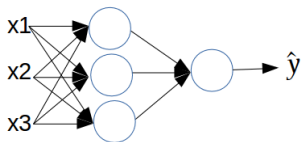
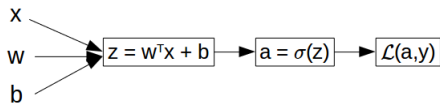
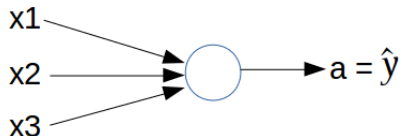
What is a neural network?



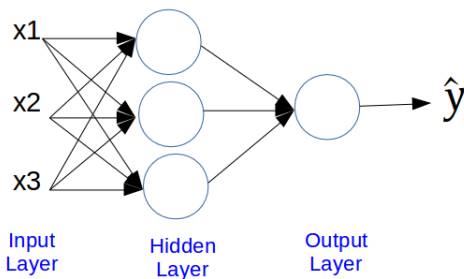
What is a neural network?



What is a neural network?



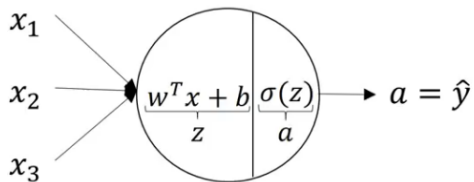
NN Representation



This is a:

- 2 layers neural network
- The hidden layers and output layer (sometime) will have the parameters (W,b)

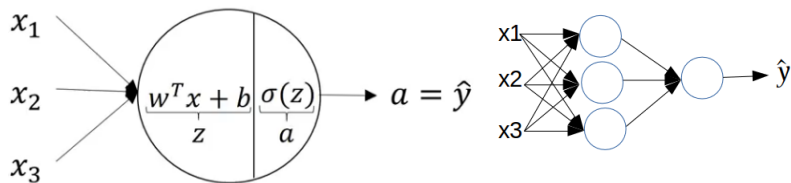
NN Computing



$$z = w^T x + b$$

$$a = \sigma(z)$$

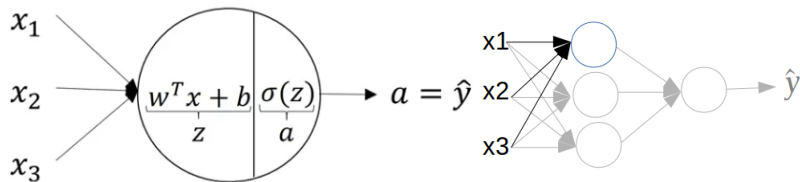
NN Computing



$$z = w^T x + b$$

$$a = \sigma(z)$$

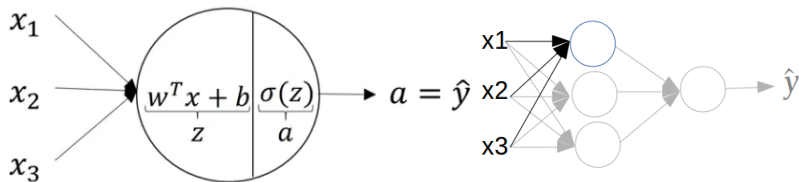
NN Computing



$$z = w^T x + b$$

$$a = \sigma(z)$$

NN Computing



$$z = w^T x + b$$

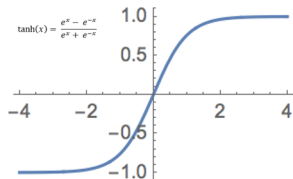
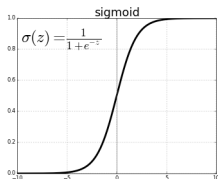
$$a = \sigma(z)$$

$$z_1^{[1]} = w_1^{[1]T} x + b_1^{[1]}$$

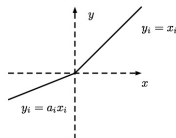
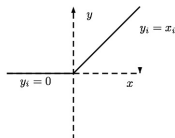
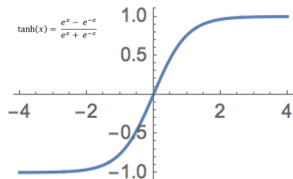
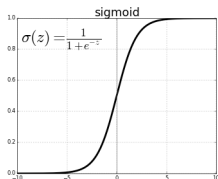
$$a_1^{[1]} = \sigma(z_1^{[1]})$$

Note: $a_i^{[l]}$ denotes node a_i at layer l .

Activation functions

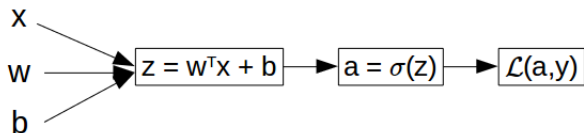


Activation functions



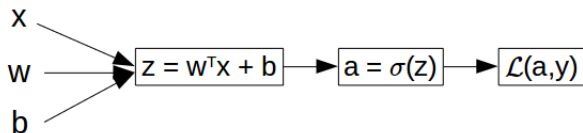
- For hidden units: use tanh, ReLU should be better than sigmoid

Backward propagation intuition

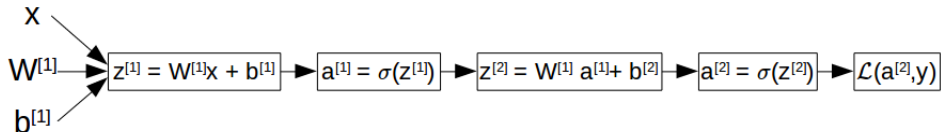


What is values of da , dz , dw , and db ?

Backward propagation intuition



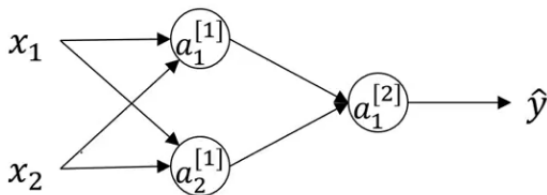
What is values of da , dz , dw , and db ?



What is values of $da^{[2]}$, $dz^{[2]}$, $dw^{[2]}$, $db^{[2]}$, $da^{[1]}$, $dz^{[1]}$, $dw^{[1]}$, and $db^{[1]}$?

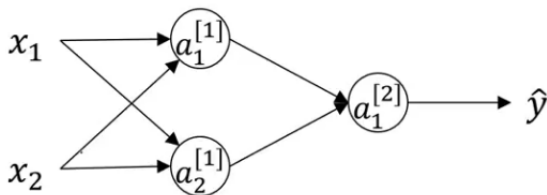
Parameters random initialization

If $W = [0]$ then the computing of neural network is exactly the same function at every layer?



Parameters random initialization

If $W = [0]$ then the computing of neural network is exactly the same function at every layer?



Solution: Randomly initialization the parameters

- $W = \text{np.random.randn}((2,2)) * 0.01$
- $b = \text{np.zeros}((2,1))$

Summary

- Overview about Neural Network with hidden layers
- Initialize the parameters
- Activation of neural networks