school of ai Algiers

Introduction to Neural Networks

AI2E

Samir Si-Mohammed

fs_si_mohammed@esi.dz

Abderaouf Gacem

fa_gacem@esi.dz

school of ai Algiers

AI2E

school of ai Algiers

01 INTRODUCTION

02 NN Definition

03 Learning Process

04 NN Architectures

AI2E

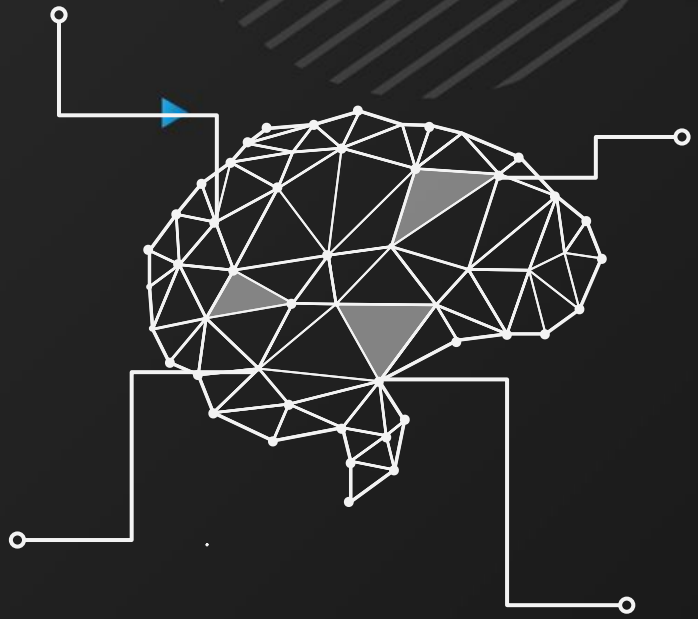
01. INTRODUCTION

school of ai Algiers

Non-exclusive MACHINE LEARNING MAP

Neural Networks & Deep LEARNING

- Perceptrons MLP
- Convolutional NN
- RNN
- GAN
- Auto-Encoders



CLASSICAL LEARNING

Supervised

- Regression (LinReg, ...)
- Classification (LogReg, SVM, ...)

Unsupervised

- Clustering (K-means, ..).

Ensemble Methods

Bagging

- Random forests

Boosting

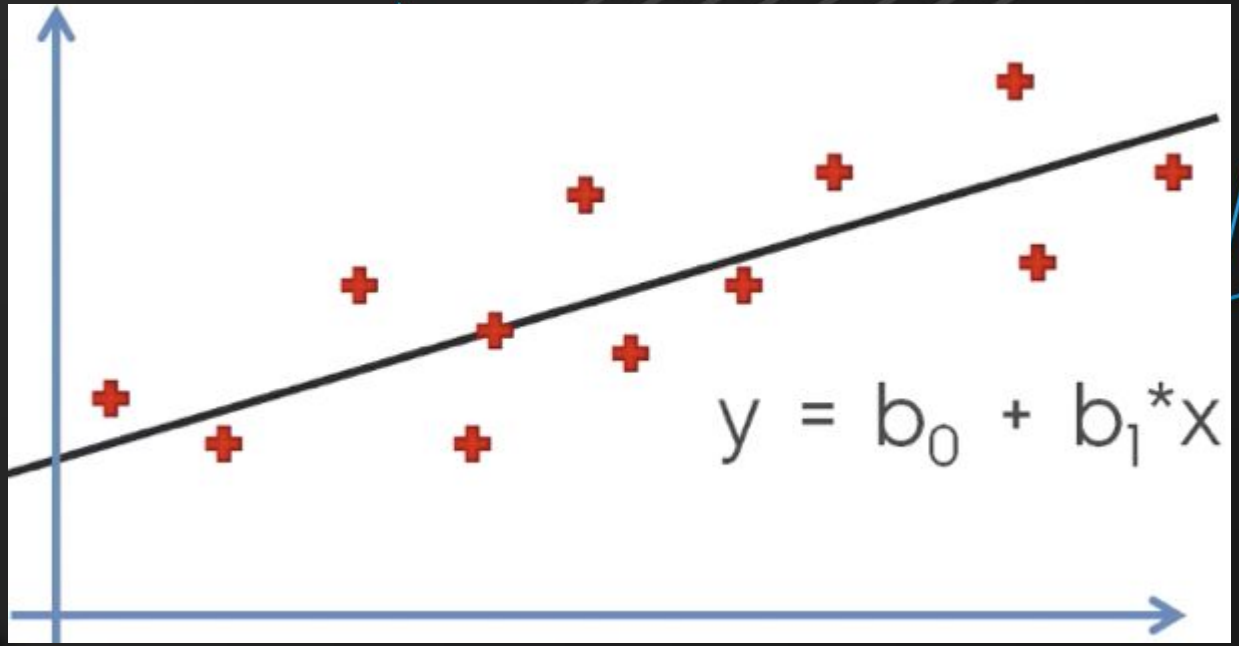
- XGBoost, LightGBM, ...

Reinforcement LEARNING

- Q-Learning
- Deep RL
- A3C

AI2E

Linear Regression

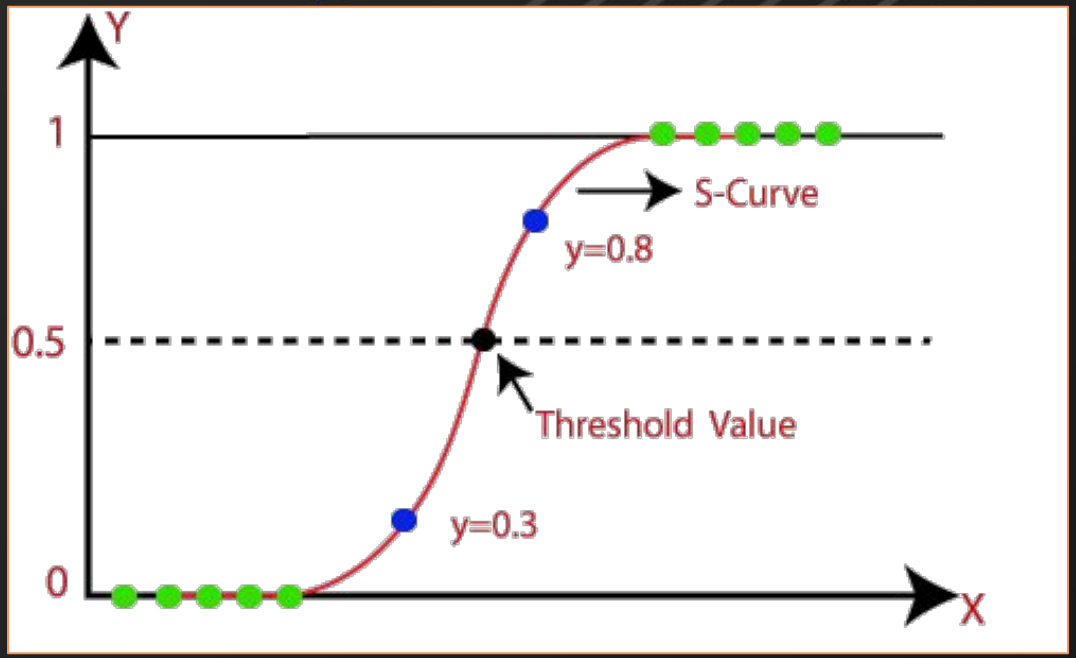


Linear Regression Plot (Clare Liu, towardsdatascience.com)

school of ai Algiers

AI2E

Logistic Regression

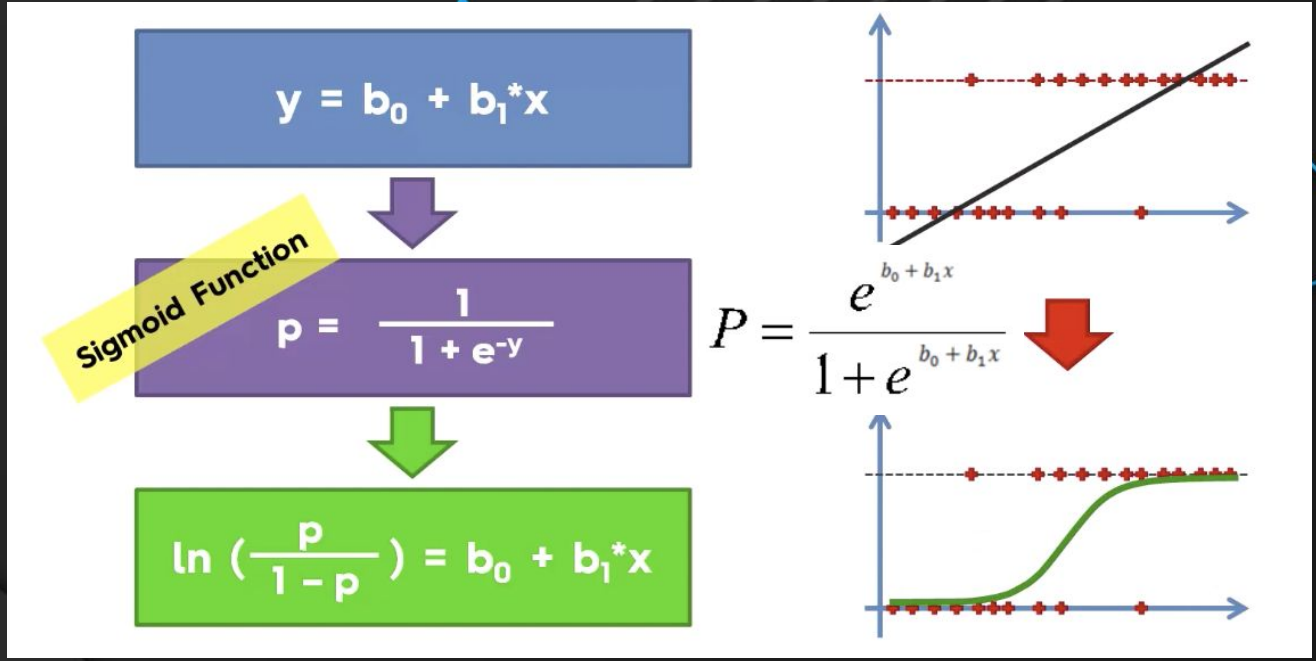


Logistic Regression Plot (JavatPoint.com)

school of ai Algiers

AI2E

Logistic Regression

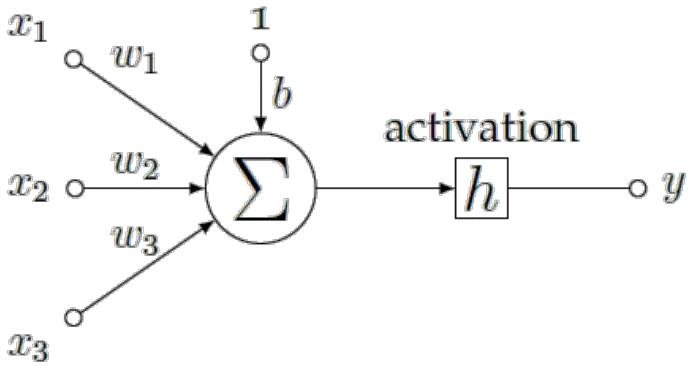


Linear to Logistic Regression (Clare Liu, towardsdatascience.com)

school of ai Algiers

AI2E

Logistic Regression



$$y = f(x) = h \circ g(x)$$

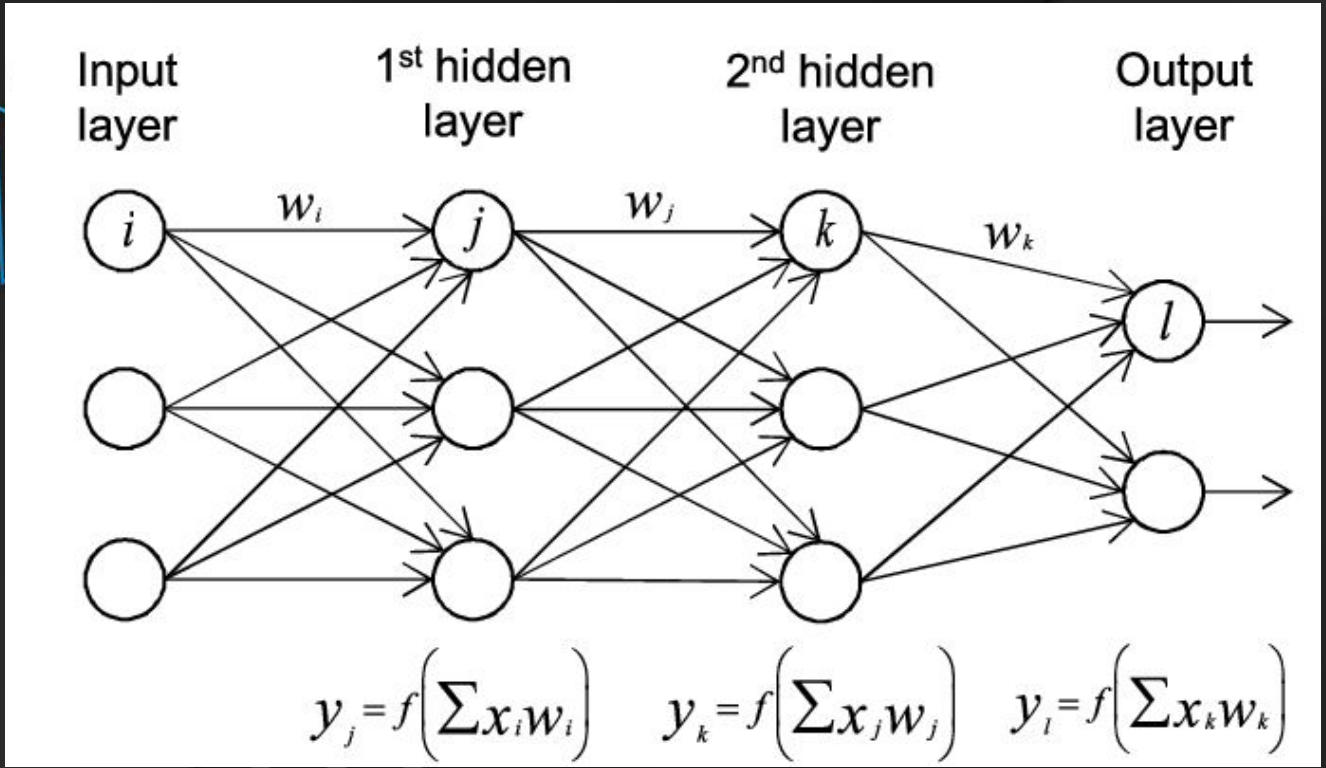
$$g(x) = W.x^T = \sum_{i=1}^{3} w_i.x_i + b$$

$$W = [w_1, w_2, w_3, b]$$

$$x = [x_1, x_2, x_3, 1]$$

02. NN Definition

Neural Network



NN graphical representation (Vieira et al, 2017)

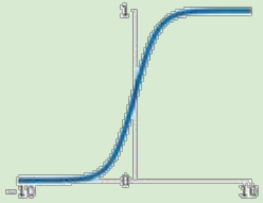
Neural Network Architecture

- A Neural Network is represented by a set of connected layers
- A layer is a set neurons formally defined by :
  - A weights matrix
  - A biases vector
  - A forward function
  - An activation Function
- A connection is associated to a weight, a neural network can have multiple types of connections.
- Different combinations of the highlighted terms yield to different neural network architectures.

Activation functions

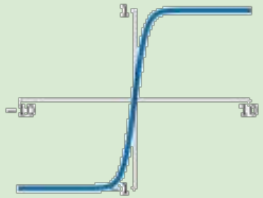
Sigmoid

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



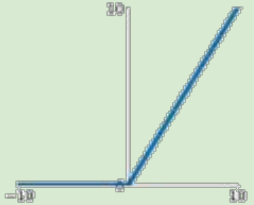
tanh

$$\tanh(x)$$



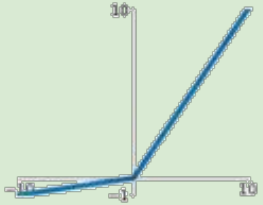
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

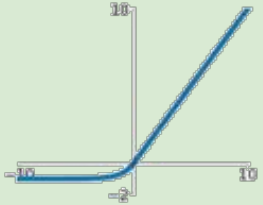


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



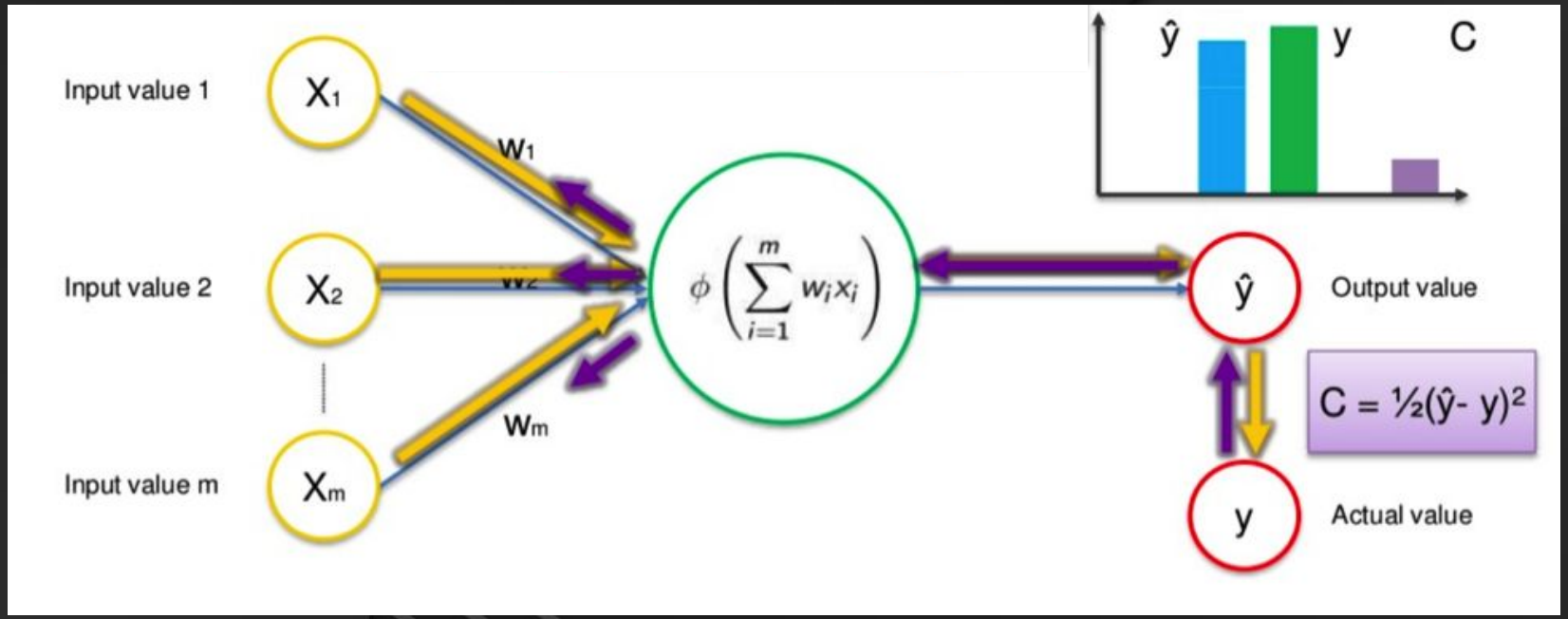
AI2E

03. Learning Process

school of ai Algiers

AI2E

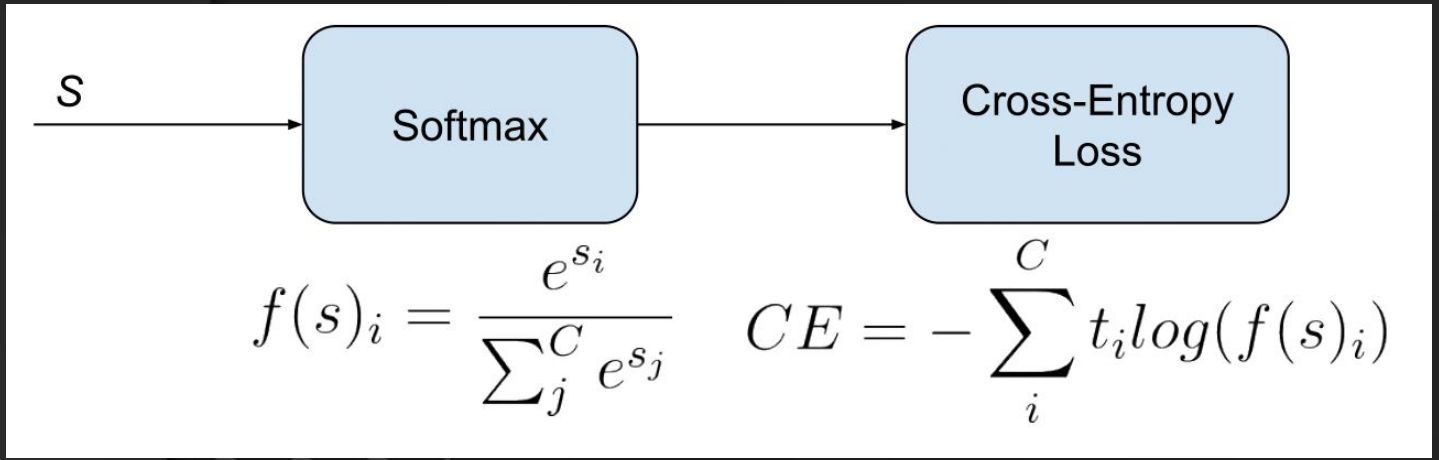
Learning Process



Learning with Logistic Regression (superdatascience.com)

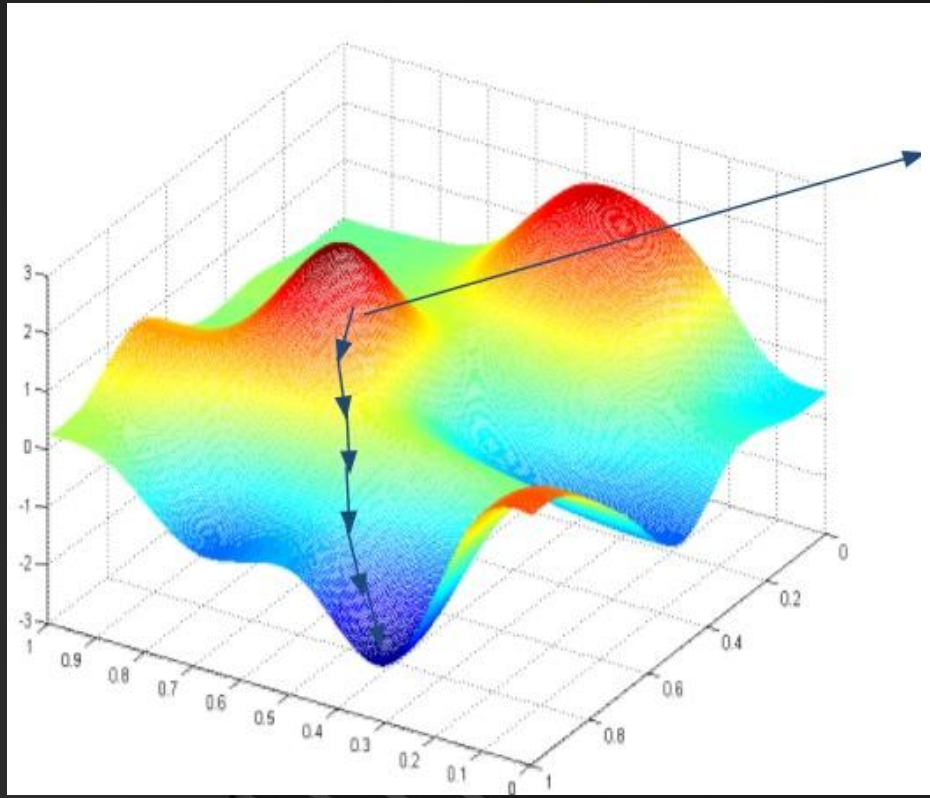
There are a multitude of usual functions to evaluate the result of the neural network in relation to what is expected, among others, we can quote :

- The quadratic error.
- Cross entropy.
- Kullback-Leibler divergence.
- Hellinger's distance.



Cross Entropy Formula (gombru)

Learning Process



$$w_i \leftarrow w_i + \Delta w_i$$

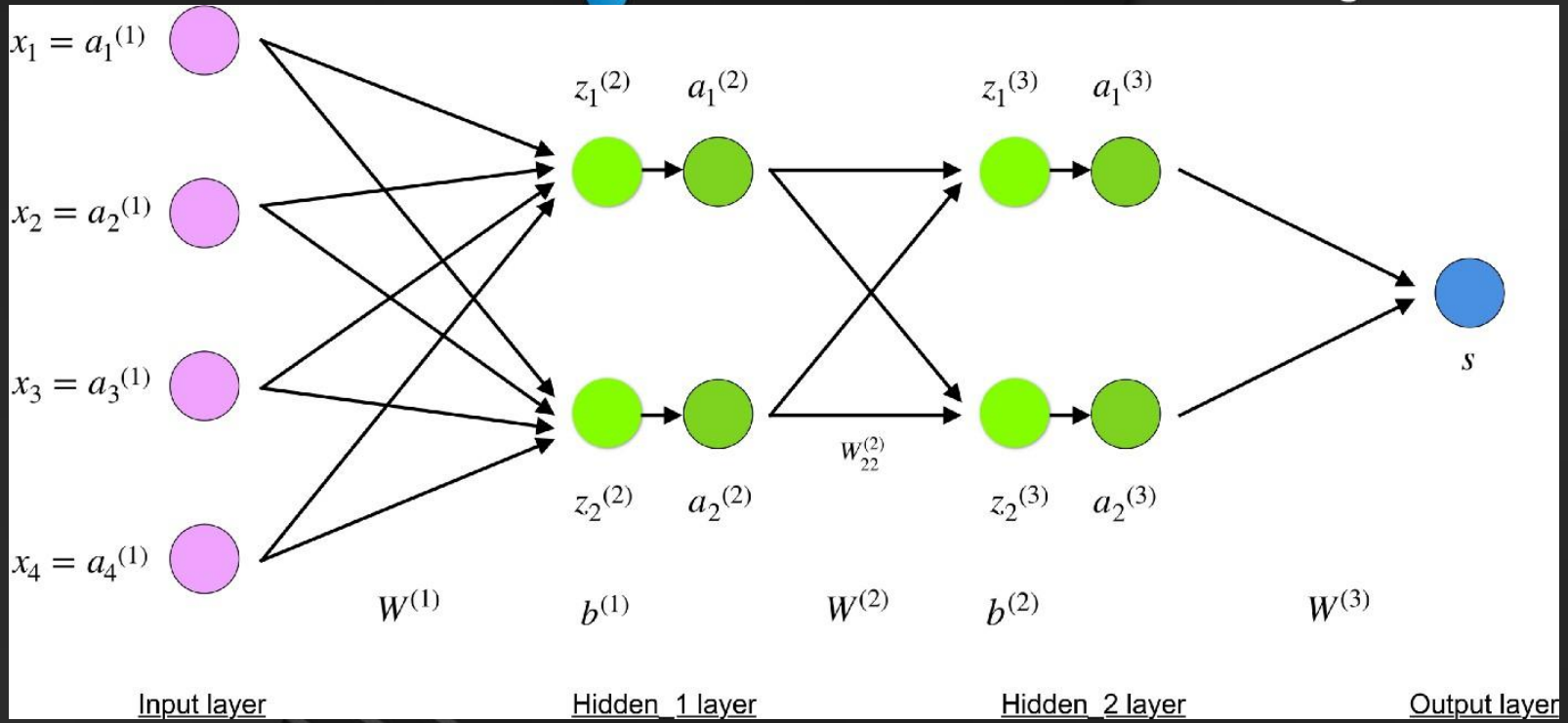
Where,

$$\Delta w_i = -\eta \frac{\partial C}{\partial w_i}$$

and

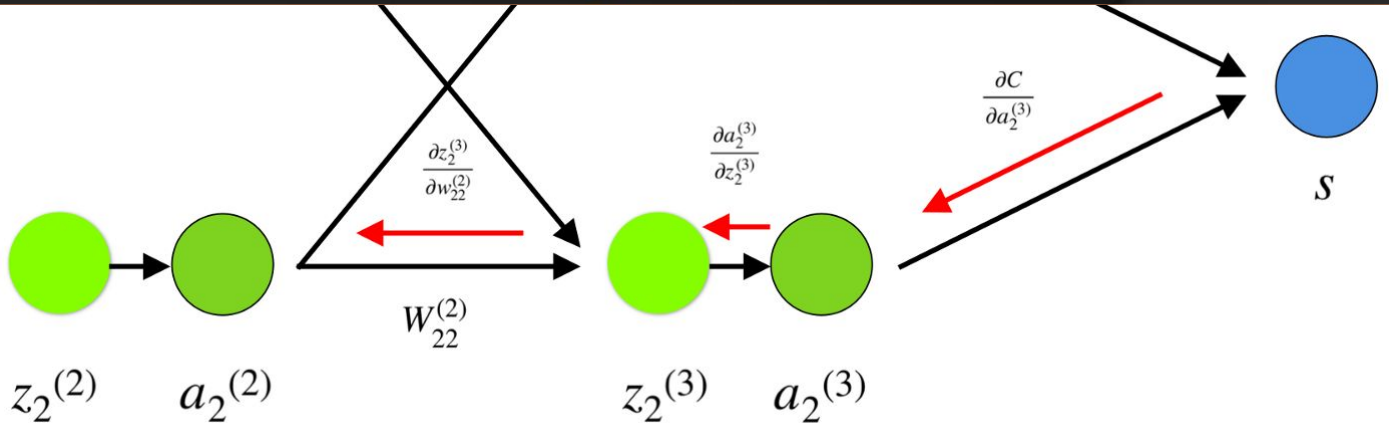
$$\frac{\partial C}{\partial w_i} = \frac{\partial}{\partial w_i} \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2$$

Learning Process



Simple 4-layer neural network illustration (Simeon Kostadinov, towardsdatascience.com)

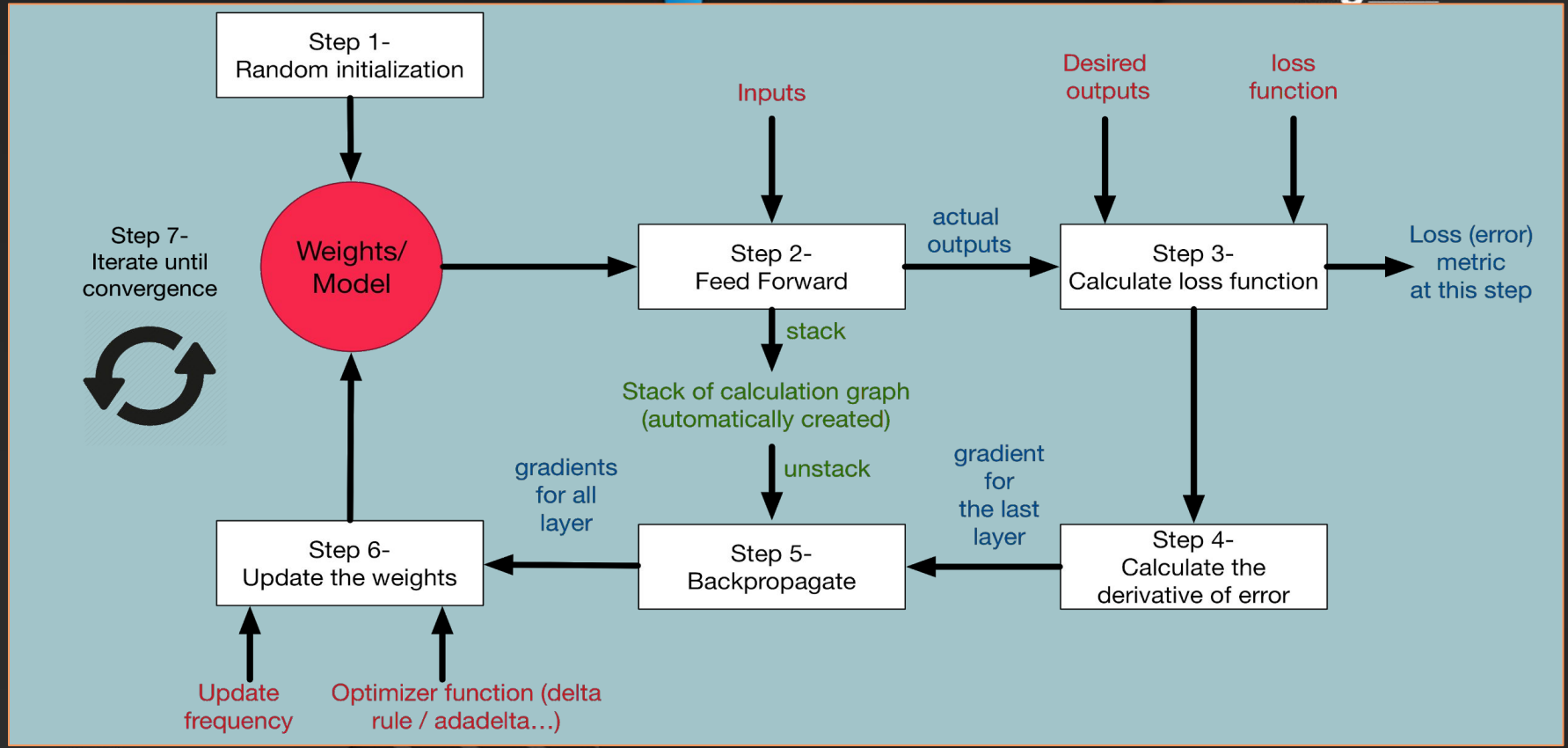
Learning Process



$$\frac{\partial C}{\partial w_{22}^{(2)}} = \frac{\partial C}{\partial z_2^{(3)}} \cdot \frac{\partial z_2^{(3)}}{\partial w_{22}^{(2)}} = \frac{\partial C}{\partial a_2^{(3)}} \cdot \frac{\partial a_2^{(3)}}{\partial z_2^{(3)}} \cdot a_2^{(2)} = \frac{\partial C}{\partial a_2^{(3)}} \cdot f'(z_2^{(3)}) \cdot a_2^{(2)}$$

backpropagation in a neural network (Simeon Kostadinov, towardsdatascience.com)

Learning process Summary

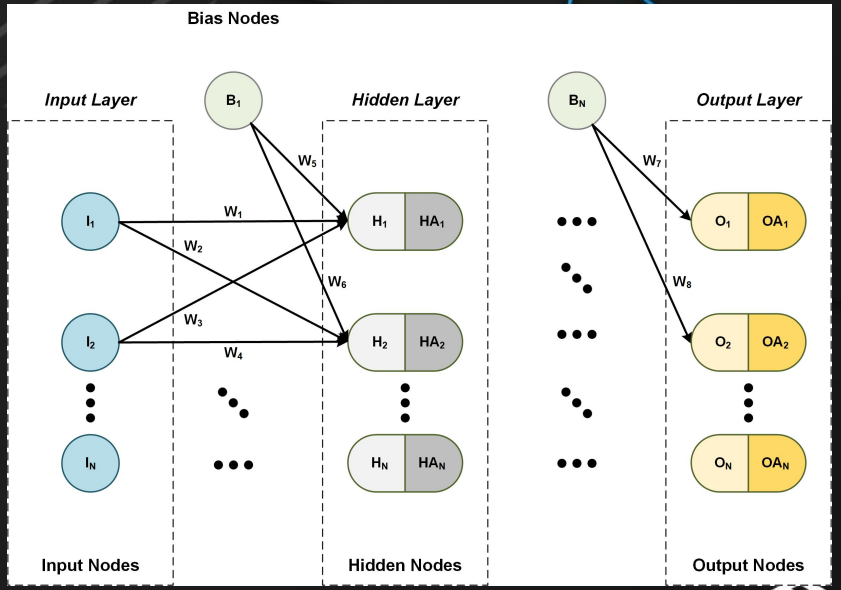


Neural networks step-by-step (Assaad MOAWAD, medium.com)

04. **NN Architectures**

school of ai Algiers

Architecturally, feed-forward neural network are a sequential stack of layers; each layer is a vertical stack of artificial neurons. Feed-forward neural networks are fully-connected.

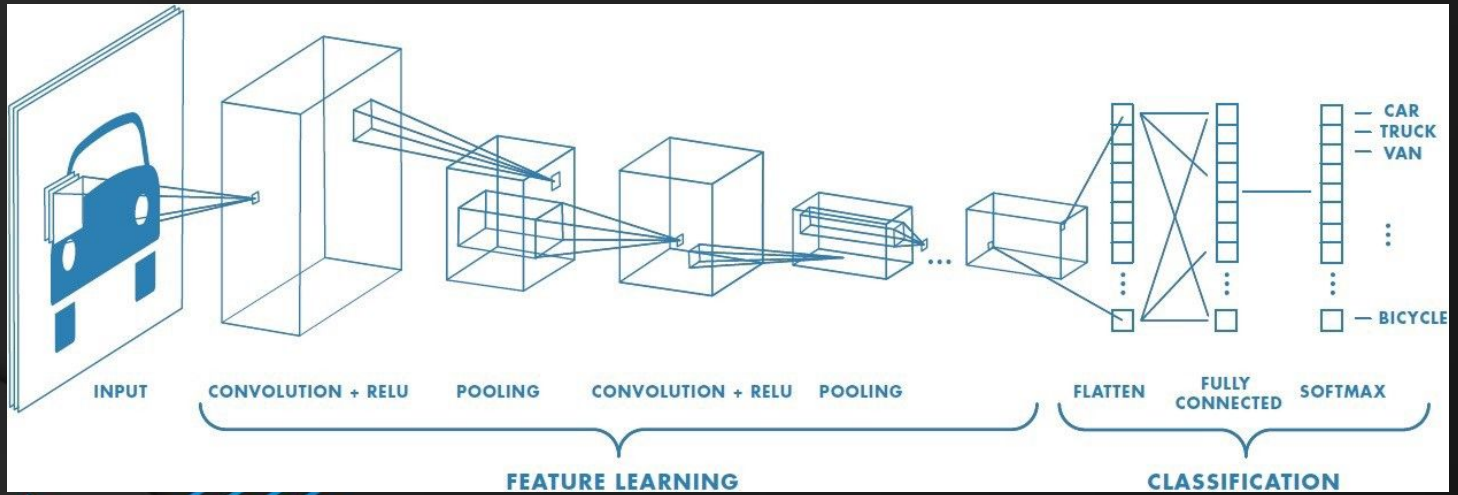


(Feed Forward NN Architecture, Edvin Beqari, 2018)

AI2E

school of ai Algiers

Convolutional NNs are a specialized kind of neural networks for processing data with grid-like topology, such as (time-)series (1D grid) and images (2D/3D grid).

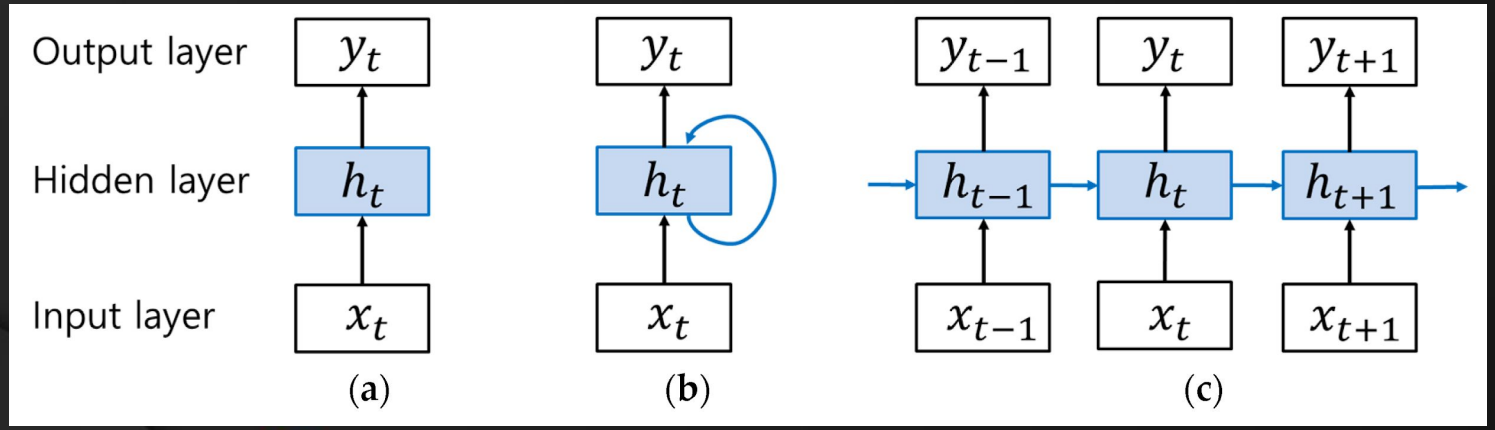


(CNN Architecture, Saha, 2018)

AI2E

school of ai Algiers

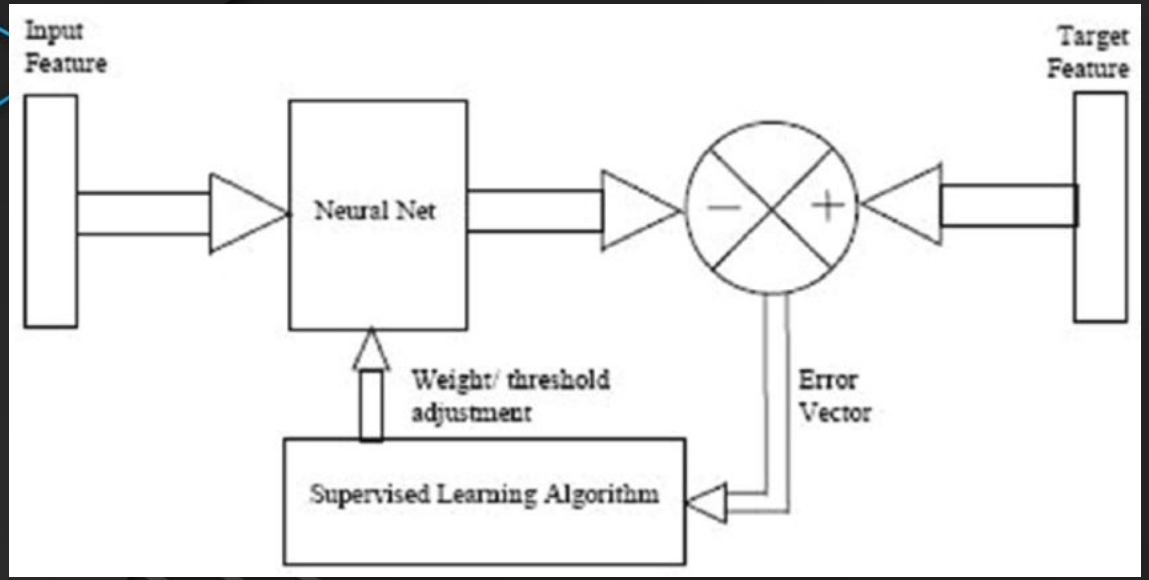
Recurrent neural networks are networks specialised for processing sequences of values x(1), ..., x(T). The crucial aspect of RNNs is that each unit uses information from previous unit to compute activations



(RNN Architecture, (Yangseon Kim et al., 2017))

AI2E

Supervised Learning



(NN for Supervised Learning, Elmolla & Chaudhuri, 2011)

school of ai Algiers

Applications

Structured Data :

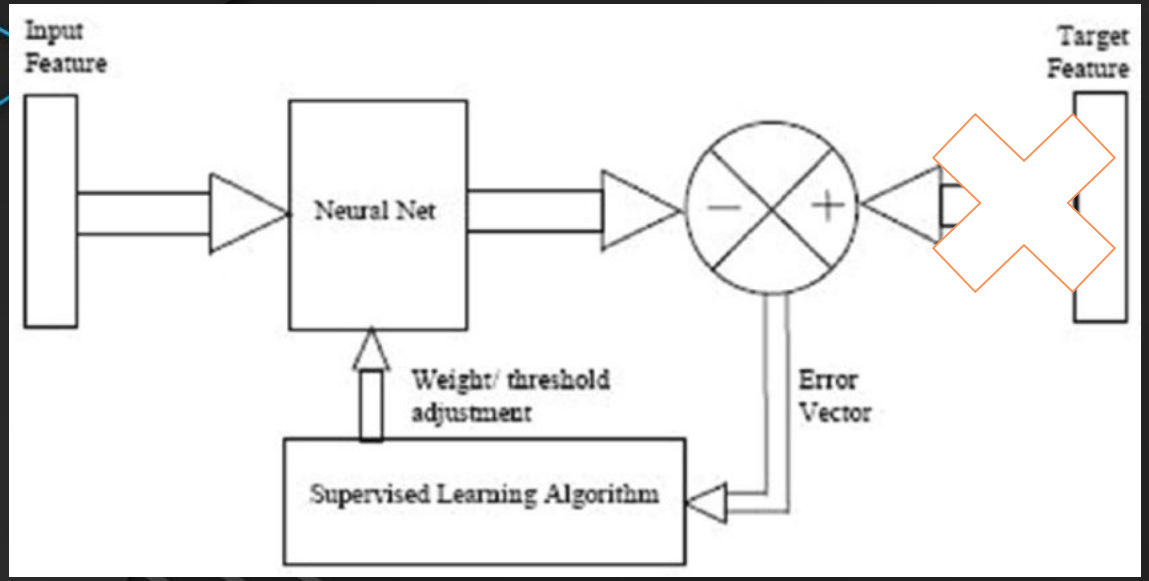
- Spam Message
- Fraud Detection

Unstructured Data :

- Facial Recognition
- Transcription
- Written Text

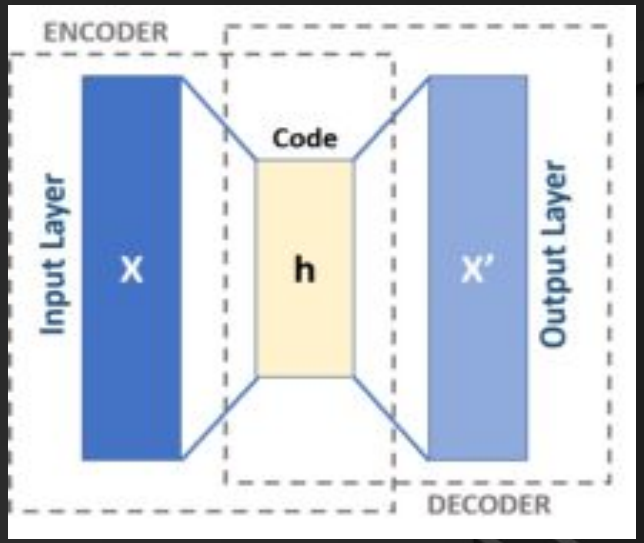
AI2E

Unsupervised Learning

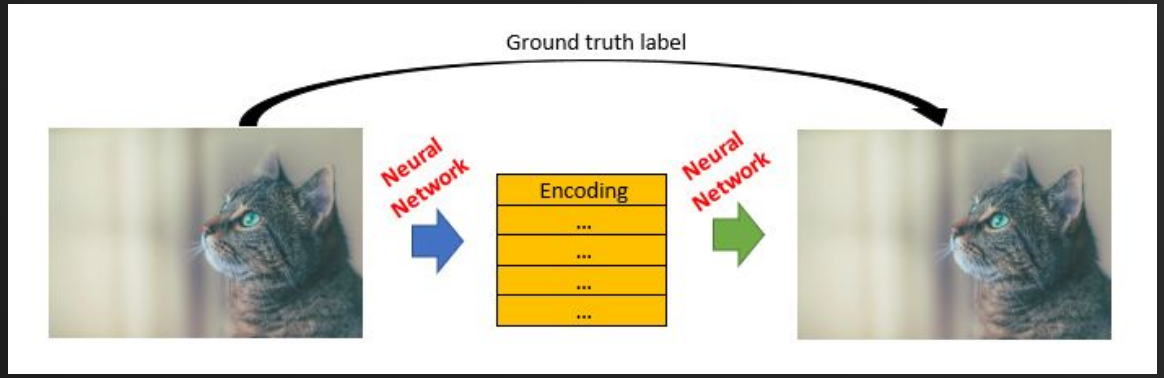


(NN for Supervised Learning, Elmolla & Chaudhuri, 2011)

Auto-encoders



(Autoencoder Architecture, Wikipedia)



(Joseph Lee Wei En, 2019)

References

PhD Davide Fiacconi (2019), Introduction to Deep Learning, Disruptive Summer School, Viterbo, Italy

Joseph Lee Wei En (2018), Autoencoders: Neural Networks for Unsupervised Learning, Medium

school of ai Algiers

AI2E

Let’s jump to coding

school of ai Algiers

AI2E