

[GitHub Link](#)

Neural Network: A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates

Types of NN:

1- Supervised Learning:

- **Regression:** Predict an output depends on the inputs data, for example: Predict house price.
- **Classification:** Classify the inputs into different classes, and start predicting the type depends on the classes, for example: online adverting, click on add? (0/1).
- The difference between regression and classification is: once we have a class the problem will defined as classification problem.

Input (x)	Output (y)	Application
Home Features	Price	Real Estate (Standard)
Ad, user info	Click on ad? 0/1	Online Adverting (Standard)
Image	Object (1...,1000)	Photo Tagging (CNN)
Audio	Text Transcript	Speech Recognition (RNN)
English	Arabic	Machine Translation (RNN)
Image, Radar info	Position of other cars	Autonomous Driving

2- Unsupervised Learning.

3- Semi-supervised Learning.

Types of Data:

1. **Structured Data:** Labels with well defined meaning “Database”
2. **Unstructured Data:** Data that's not actively managed in a transactional system “Database”, for example: Audio data.

1. **Training Data:** Training data is the data you use to train an algorithm model to predict the outcome you design your model to predict. (Inputs data)
2. **Validation Data:** A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model's hyperparameters.
3. **Test Data:** A test dataset, the same thing as validation data. But the difference is that this data was not used before in the model.

Activation Function: The activation function is a mathematical equation inside the neuron, and every layer has a fixed function. But can be a different function in other layers.

The activation function must be non-linear function. And the last layer can be linear activation function.

Types of Activation Functions:

1. **ReLU Activation Function:** piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. Range: $(0, \infty)$

$$f(x) = a = \max(0, x)$$

2. **Sigmoid Activation Function:** is used in binary classification, if the output either (0/1). Range: $(0, 1)$

$$S = \alpha(z) = \frac{1}{1 + e^{-z}}$$

3. **Tanh Activation Function:** The most used function, the one exception is for the output layer because if y is either zero or one, so we use sigmoid function. Range: $(-1, 1)$

$$\alpha(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

Functions:

1. **Loss Function 'L':** For one training example.
2. **Cost Function 'J':** For average training examples.

Gradient descent: Optimization algorithm which is commonly-used to train machine learning models and neural networks. Training data helps these models learn over time, and the cost function within gradient descent specifically acts as a barometer, gauging its accuracy with each iteration of parameter updates.

Backpropagation: is the essence of neural network training. It is the method of fine-tuning the weights of a neural network based on the error rate obtained in the previous epoch (i.e., iteration). Proper tuning of the weights allows you to reduce error rates and make the model reliable by increasing its generalization.

The backpropagation algorithm works by computing the gradient of the loss function with respect to each weight by the chain rule, computing the gradient one layer at a time, iterating backward from the last layer to avoid redundant calculations of intermediate terms in the chain rule; this is an example of dynamic programming

Neural networks – representation:

