

Neural Network cheat sheet

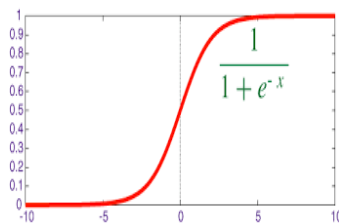
Week-1

Neural Network:

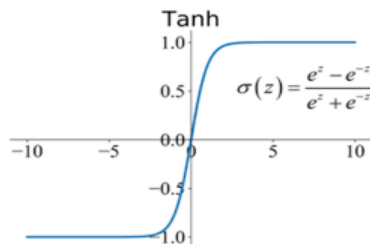
A neural network is a network or circuit of biological neurons, or, in a modern sense, an artificial neural network, composed of artificial neurons or nodes. Thus, a neural network is either a biological neural network, made up of biological neurons, or an artificial neural network, used for solving artificial intelligence (AI) problems.

Activation functions:

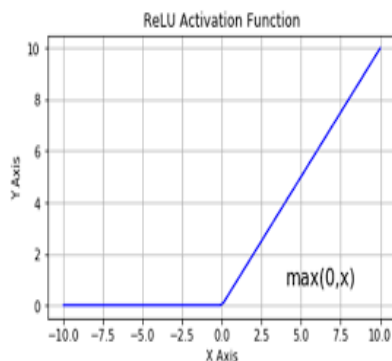
1) Sigmoid



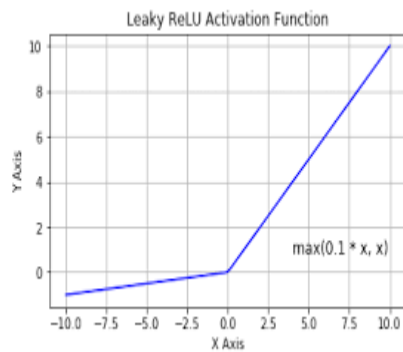
2) Tanh



3) ReLU

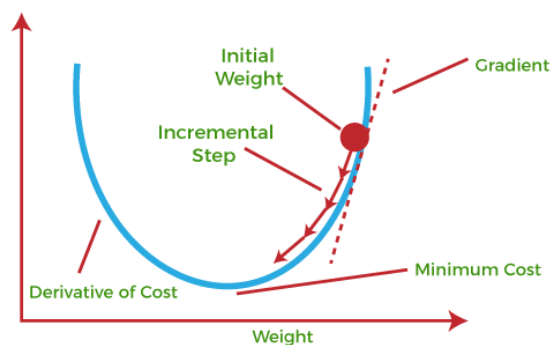


4) Leaky ReLU



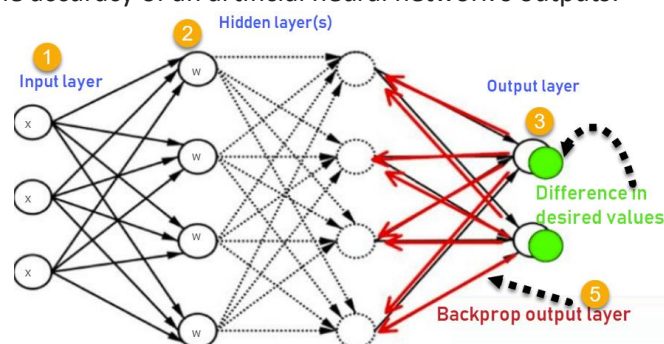
Gradient descent:

Gradient descent (GD) is an iterative first-order optimisation algorithm used to find a local minimum/maximum of a given function



Backpropagation:

Backpropagation is an algorithm used in artificial intelligence (AI) to fine-tune mathematical weight functions and improve the accuracy of an artificial neural network's outputs.



Learning rate:

The learning rate is a hyperparameter which determines to what extent newly acquired weights overrides old weights. In general, it lies between 0 and 1.

Momentum:

Momentum is used to decide the weight on nodes from previous iterations. It helps in improving training speed and also in avoiding local minimas.

Week -2

Chain Rule:

$$\frac{da}{dx} a(b(c(d(e(f(g(x)))))))$$

$$\frac{da}{dx} = \frac{da}{db} \times \frac{db}{dc} \times \frac{dc}{dd} \times \frac{dd}{de} \times \frac{de}{df} \times \frac{df}{dg} \times \frac{dg}{dx}$$

Softmax Layer:

The softmax function is used as the activation function in the output layer of neural network models that predict a multinomial probability distribution.

Cross Entropy Loss:

Cross-entropy loss, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1

$$L(z,y) = -[y \log(z) + (1-y) \log(1-z)]$$

Week-3

Data Pre-processing:

Pre-processing refers to all the transformations on the raw data before it is fed to the deep learning algorithm.

Data Augmentation:

Data augmentation is a set of techniques to artificially increase the amount of data by generating new data points from existing data. This includes making small changes to data or using deep learning models to generate new data points.

Weight initialization:

Weight initialization is used to define the initial values for the parameters in neural network models prior to training the models on a dataset.

Regularization:

Regularization is a set of techniques that can prevent overfitting in neural networks and thus improve the accuracy of a Deep Learning model when facing completely new data from the problem domain

Batch Normalization:

Batch Normalization reduce the internal covariate shift and instability in distributions of layer activations in Deeper networks can reduce the effect of overfitting.

Dropout:

Dropout is a technique used to prevent a model from overfitting. Dropout works by randomly setting the outgoing edges of hidden units (neurons that make up hidden layers) to 0 at each update of the training phase.