greatlearning
Learning for Life

Neural Networks

Introduction

Outline

- Introduction to neural networks
- Activation functions
- Feed Forward neural network
- Layer Details
- Training a neural network
- Error and Loss function
- Optimization, Backpropagation
- Gradient descent
- Early stopping and model saving
- Summary

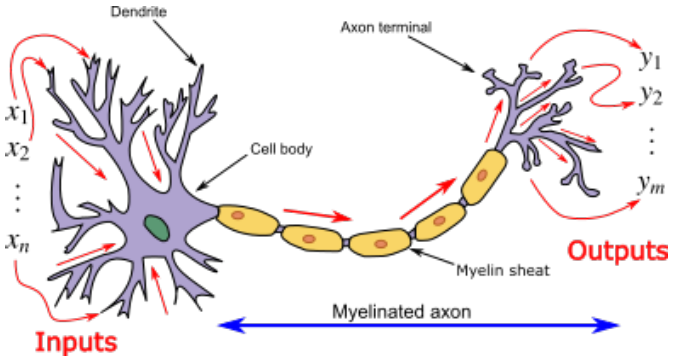
greatlearning
Learning for Life

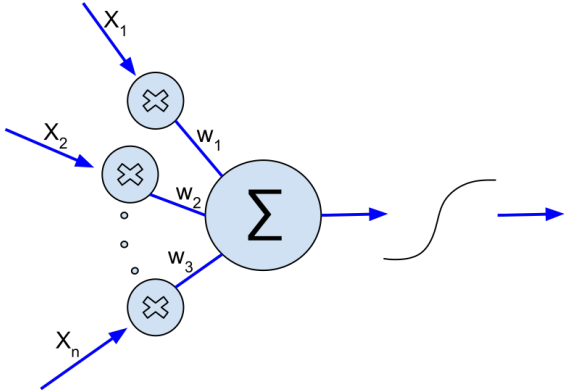
Neural Networks

Artificial Neural Networks are computing systems inspired from biological neuron

Neuron

Artificial neuron is inspired by biological neuron



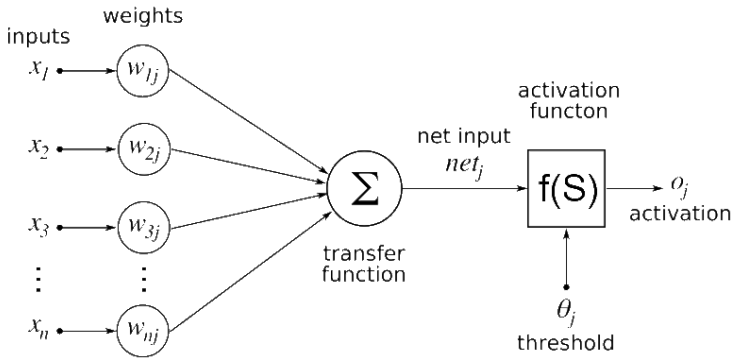


Activation function

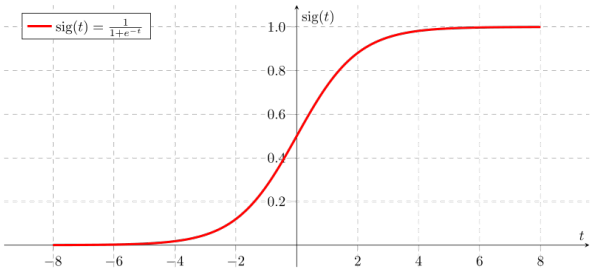
In artificial neural networks, helps in defining the output of a node when a input is given.

Different types of activations helps in different tasks.
Some examples are -

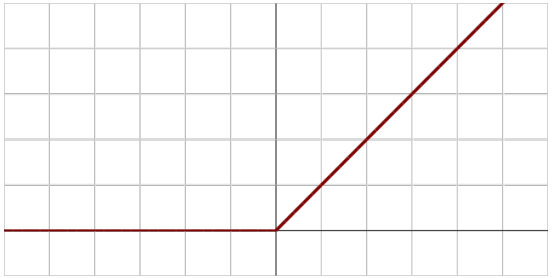
- Sigmoid
- Tanh
- ReLU
- Binary Step Function



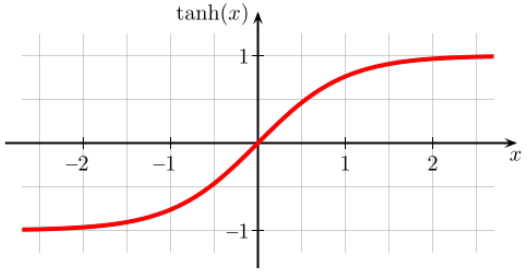
Types of activation function



Sigmoid



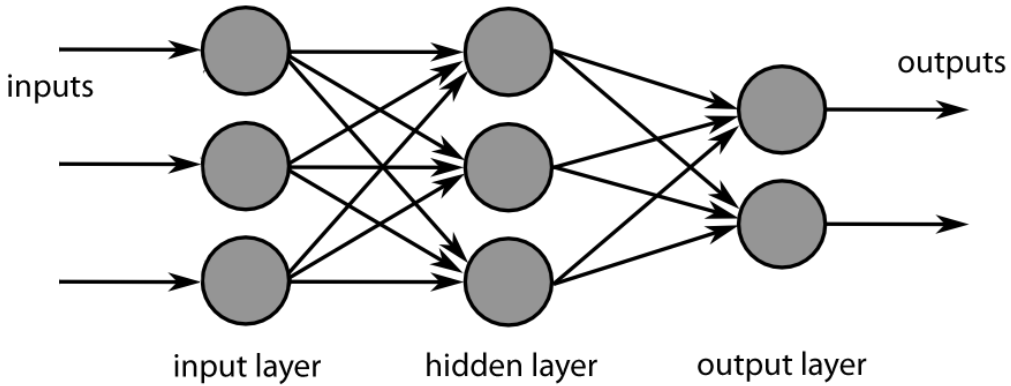
ReLU



Tanh

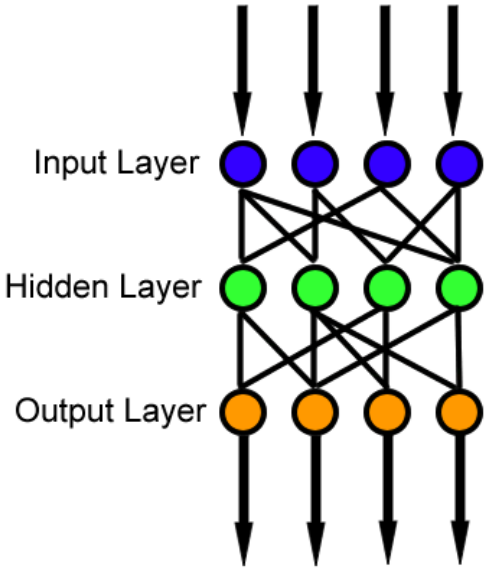
Activation function performs certain mathematical operations on it's input, which is a number

Feed forward neural network



This is a 2-layer neural network. One is the hidden layer (having 3 neurons) and the other is output layer (having 2 neurons).

Layer details



Output layer

- Represents the output of the neural network

Hidden layer(s)

- Represents the intermediary nodes.
- It takes in a set of weighted input and produces output through an activation function

Input layer

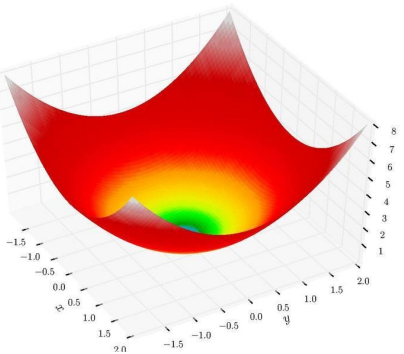
- Represents dimensions of the input vector (one node for each dimension)

Training a neural network

- Decide the structure of network
- Create a neural network
- Choose different hyper-parameters
- Calculate loss
- Reduce loss
- Repeat last three steps

Error and Loss function

- In general, error/loss for a neural network is difference between actual value and predicted value.
- The goal is to minimize the error/loss.
- Loss Function is a function that is used to calculate the error.
- You can choose loss function based on the problem you have at hand.
- Loss functions are different for classification and regression

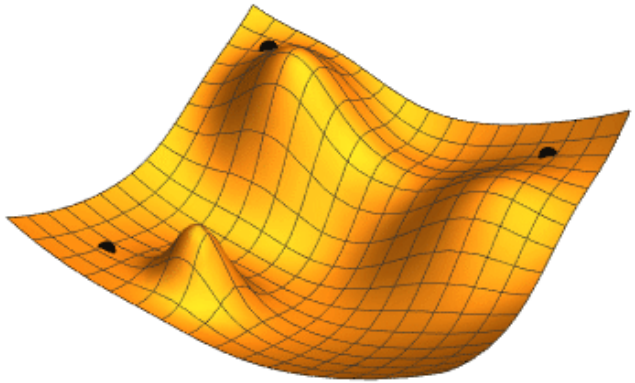


Optimization

In optimization, the main aim is to find weights that reduce loss

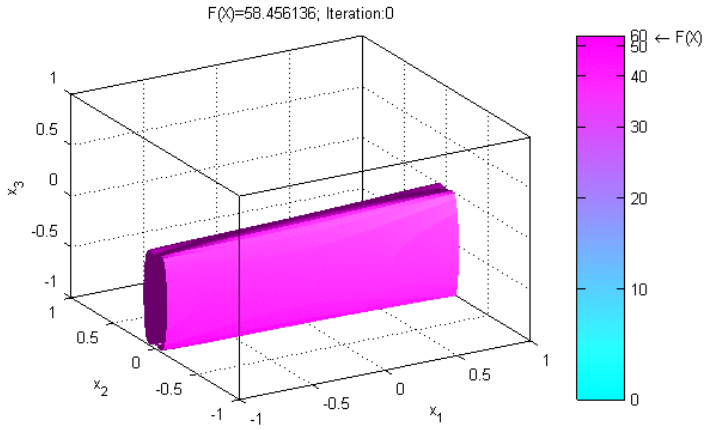
Gradient

- **Gradient** is calculated by optimization function
- Gradient is the change in loss with change in weights.
- The weights are modified according to the calculated gradient.
- Same process keep on repeating until the minima is reached



Gradient descent

Gradient descent is a method that defines a cost function of parameters and uses a systematic approach to optimize the values of parameters to get minimum cost function.



Gradient descent variations

Gradient descent has 3 variations, these differ in using data to calculate the gradient of the objective function

1. Batch gradient descent
   - Updates the parameter by calculating gradients of whole dataset
2. Stochastic gradient descent
   - Updates the parameters by calculating gradients for each training example
3. Mini-batch gradient descent
   - Updates the parameters by calculating gradients for every mini batch of "n" training example
   - Combination of batch and stochastic gradient descent

Backpropagation

- Backpropagation is used while training the feedforward networks
- It helps in efficiently calculating the gradient of the loss function w.r.t weights
- It helps in minimizing loss by updating the weights

Learning rate and Momentum

- 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.
- You can try different learning rates for a neural networks to improve results.
- Momentum is used to decide the weight on nodes from previous iterations. It helps in improving training speed and also in avoiding local minimas.

Early Stopping and Model Saving

- We need to monitor the error on validation set while training the model and if we feel that there is no change, then we should stop training.
- Save all the necessary details to reuse the model.

Summary of the Neural Network Process

greatlearning
Learning for Life

- A neural network is made of neurons
- These neurons are connected to each other
- Every neuron has an activation function that defines it's output
- Then we train our neural network to learn the parameter values i.e. weights and biases
- This process consists of forwardprop and backprop
- After forwardprop we calculate the loss using a loss function and propagate the information backwards, that's backprop
- This process is repeated layer by layer, until all the neurons receive a loss signal which describes their contribution to the total loss

greatlearning
Learning for Life

Thank you!

Happy Learning :)