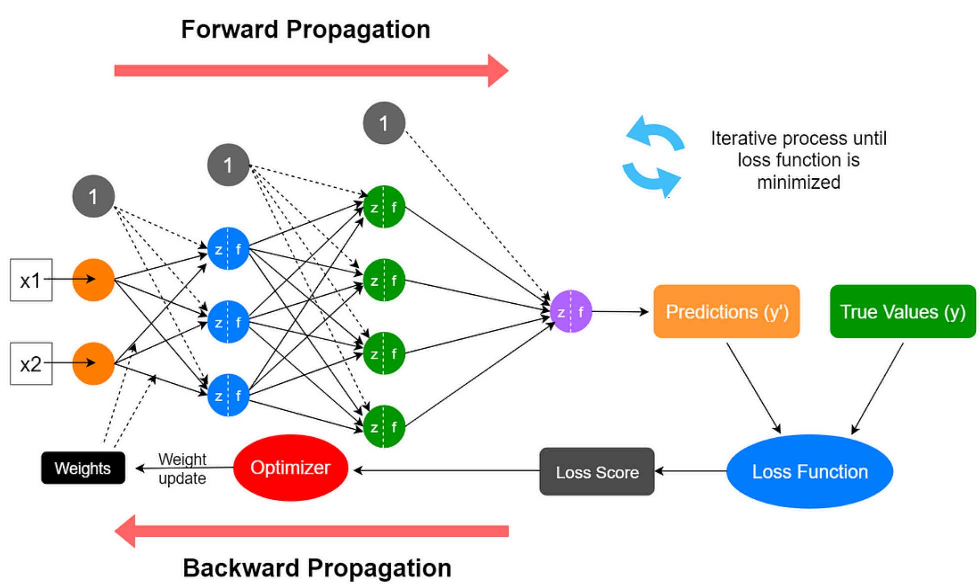
Neural network

A neural network is a deep learning model inspired by the structure and functioning of the human brain's neuron and neural networks. It's a type of learning algorithm that can learn to perform tasks by learning from mistakes by the use of forward propagation, backward propagation, weights, loss function and activation function, rather than being explicitly programmed to do so.



<u>Key components of a neural network:</u>

<u>1. Neurons</u>:

The basic building blocks of a neural network. In neural networks, neurons are mathematical functions that receive one or more inputs according to the number of datapoints. Each neuron typically has associated weights and biases that are adjusted during the learning process and produce an output.

<u>2. Layers</u>:

Neurons are organized into layers within a neural network. The three types of layers are:

Input Layer: This is the layer where the input data is fed into the network.

Hidden Layers: These are the layers between the input and output layers. They perform mathematical operations by using weight, biases and activation functions on the input data .

Output Layer: This is the final layer of the network. It produces the output of the network based on the mathematical operation performed by the hidden layers.

3. Weights and Biases:

Each neuron in adjacent layers is associated with a weight, which determines the strength of the connection. Biases are additional parameters that are added to the weighted sum of inputs before being passed through an activation function and this bias helps in adjusting the model in order to increase the accuracy. During training, the weights and biases are adjusted to minimize the difference between the predicted output and the actual output, basically we reduce the loss function in order to improve model accuracy.

4. Activation Functions:

Activation functions introduce non-linearity into the output of a neuron. They help neural networks learn complex patterns in the data by enabling them to approximate non-linear functions. Common activation functions include sigmoid, tanh, ReLU (Rectified Linear Unit), and leaky ReLU function.

Sigmoid function:

Formula: f(x) = 1/(1+e^{-x})

Range: (0, 1)

Characteristics:

- S-shaped curve,
- Smooth and continuous
- Maps to the input and give prediction

Tanh function:

Formula: tanh(x) = (e^{x}-e^{-x})/(e^{x}+e^{-x})

Range: (-1, 1)

Characteristics:

- Similar to sigmoid but centered at 0
- Symmetric around the origin

ReLU(Rectified Linear Unit):

Formula: relu(x) = max(0,x)

Range: [0, +inf)

Characteristics:
- Computational efficiency is good
- On priority function of deep learning

Leaky ReLU:
Formula: lrelu(x) = x if x≥ 0, ax if x<0
Range: (-inf, +inf)
Characteristics:
- Similar to relu but it allows small value too
- Non-zero gradient when the input is negative
- Prevent dying neuron
- Overcome vanishing gradient problem

ELU(Exponential linear unit):
Formula: elu(x) = x if x≥ 0, a(e^x-1) if x<0
Range: (-a, +inf)
Characteristics:
- Similar to relu but it operates at 0
- Makes learning rate better
- Prevent dying neuron
- Overcome vanishing gradient problem

5. Forward Propagation:
During the forward propagation phase, input data is passed through the network layer by layer from input layer to output layer, and mathematical formulations are applied at each layer until the output is produced.

6. Backpropagation:
Backpropagation is the process of updating the weights and biases of the network based on the loss function.

7. Loss Function:
The loss function measures the difference between the predicted output and the actual output of the network. The goal during training is to minimize this difference using techniques like gradient descent where we try to reach global minima.

Applications: Image classification, Natural language processing, Recommendation system, Speech recognition, Computer vision.