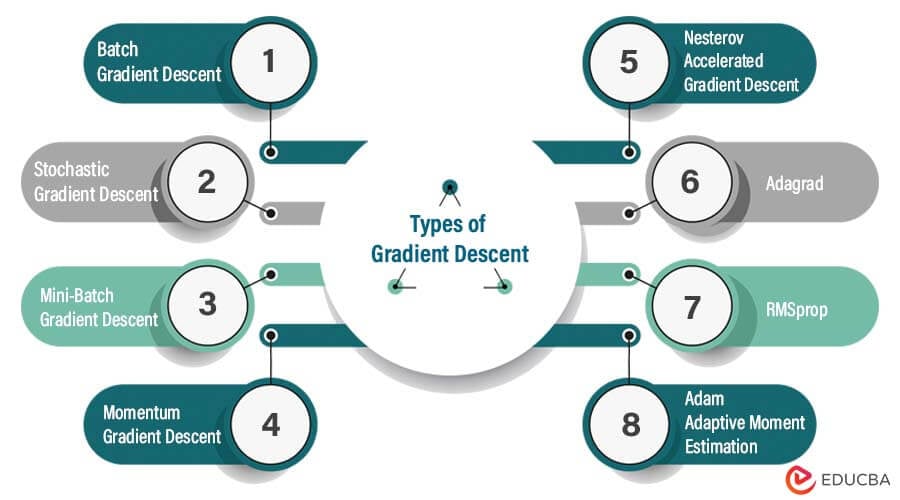
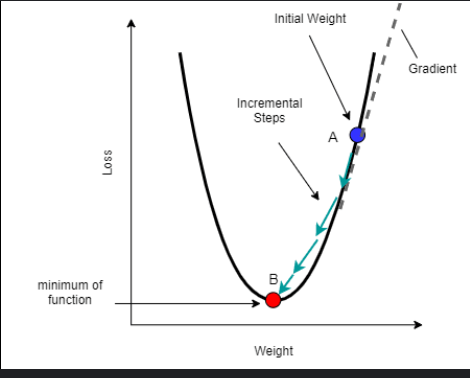
DIFFERENT TYPES OF OPTIMIZERS IN ANN

Optimizers play a crucial role in training artificial neural networks by adjusting the weights to minimize the loss function. Here are some common types of optimizers used in ANNs:



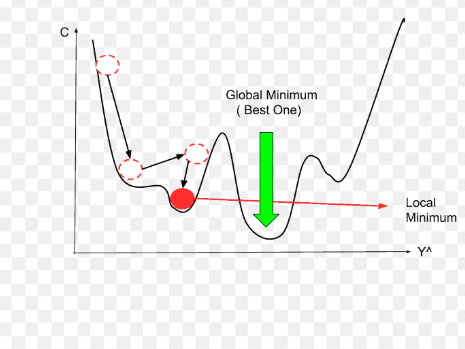
**1. Gradient Descent (GD)**

The basic optimization algorithm that updates weights by subtracting the gradient of the loss function multiplied by a learning rate.



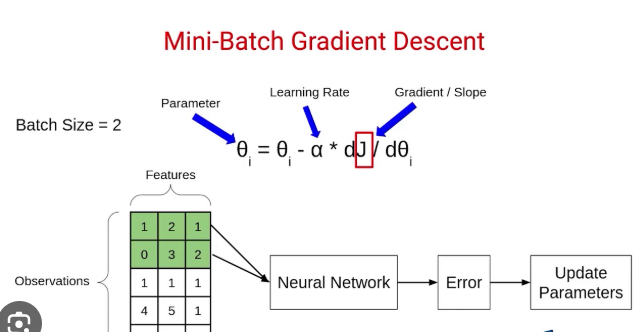
### 2. Stochastic Gradient Descent (SGD)

A variation of gradient descent that updates the weights using a single randomly chosen data point (or a small batch) rather than the entire dataset.



### 3. Mini-Batch Gradient Descent

Combines the benefits of both batch and stochastic gradient descent by updating weights using small batches of data.



**4. Momentum**

An extension of SGD that accumulates a velocity vector in directions of persistent reduction in the loss function, helping to accelerate convergence.

**5. Nesterov Accelerated Gradient (NAG)**

An improvement over momentum by adding a lookahead step, computing the gradient not just for the current position but for the anticipated position.

**6. Adagrad**

Adagrad adapts the learning rate for each parameter individually, scaling it inversely proportional to the square root of the sum of past gradients.

### 7. Adadelta

An extension of Adagrad that seeks to reduce its aggressive, monotonically decreasing learning rate.