**Gradient Descent:**

Gradient descent is an optimization algorithm used to minimize a function by iteratively moving towards the direction of the steepest decrease in the function. There are several types of gradient descent algorithms, including Batch Gradient Descent, Stochastic Gradient Descent, and Mini-Batch Gradient Descent. Here's some pseudo-code for each:

**Batch Gradient Descent:**

In Batch Gradient Descent, the entire training dataset is used to compute the gradient of the cost function with respect to the parameters in each iteration.

while not converged:

gradients = compute\_gradients(data, parameters)

parameters = parameters - learning\_rate \* gradients

**Stochastic Gradient Descent (SGD):**

In Stochastic Gradient Descent, only one randomly selected training example is used to compute the gradient in each iteration.

while not converged:

random\_example = select\_random\_example(data)

gradient = compute\_gradient(random\_example, parameters)

parameters = parameters - learning\_rate \* gradient

**Mini-Batch Gradient Descent:**

Mini-Batch Gradient Descent is a compromise between Batch GD and SGD. It uses a small random subset (mini-batch) of the training data in each iteration.

while not converged:

mini\_batch = select\_mini\_batch(data, batch\_size)

gradients = compute\_gradients(mini\_batch, parameters)

parameters = parameters - learning\_rate \* gradients