10 Types of Machine Learning Optimizers

1. Gradient Descent (GD)

Updates weights by computing the entire dataset's gradient to minimize loss.

2. Stochastic Gradient Descent (SGD)

Updates weights using one random sample per iteration for faster convergence.

3. Mini-Batch Gradient Descent

Uses small random batches of data to balance speed and stability.

4. Momentum Optimizer

Accelerates SGD by adding a fraction of the previous update to the current one.

5. Nesterov Accelerated Gradient (NAG)

Improves momentum by looking ahead before applying the gradient.

6. Adagrad

Adapts the learning rate individually for each parameter based on past gradients.

7. RMSProp

Uses a moving average of squared gradients to adjust the learning rate dynamically.

8. Adam (Adaptive Moment Estimation)

Combines momentum and RMSProp for efficient adaptive learning rates.

9. Adadelta

An extension of Adagrad that reduces its aggressive learning rate decay.

10. Nadam

Adam optimizer with Nesterov momentum for improved convergence speed.