**[Que-50.12] - Describe the difference between batch gradient descent and stochastic gradient descent.**

* **Batch Gradient Descent:** Uses the entire dataset to compute the gradient at each step.
  + **Pros:** Converges to the global minimum for convex functions.
  + **Cons:** Can be very slow and computationally expensive for large datasets.
* **Stochastic Gradient Descent (SGD):** Uses only one training example to compute the gradient at each step.
  + **Pros:** Faster convergence, less memory required.
  + **Cons:** Can be noisy and may not converge to the exact global minimum.