**Gradient Descent:**

Gradient descent is an optimization algorithm which is commonly-used to train machine learning models and neural networks. It trains machine learning models by minimizing errors between predicted and actual results.

**Types of Gradient Descent**

1. Batch gradient descent

Batch gradient descent sums the error for each point in a training set, updating the model only after all training examples have been evaluated. This process referred to as a training epoch. It uses the entire training dataset to compute the gradient of the cost function in each iteration.

1. Stochastic gradient descent

Stochastic gradient descent (SGD) runs a training epoch for each example within the dataset and it updates each training example's parameters one at a time.

1. Mini-batch gradient descent

Mini-batch gradient descent combines concepts from both batch gradient descent and stochastic gradient descent. It splits the training dataset into small batch sizes and performs updates on each of those batches.

**Validation set**

A validation set is a subset of the data that is used to evaluate the model during training, but it is not used to train the model itself. The main purpose of the validation set is to provide an unbiased evaluation of the model's performance while it is being trained.

**Validation Loss**

Validation loss is the value of the loss function computed on the validation set. It measures how well the model performs on data that it has not been trained on.

The validation loss is a measure of how well the model generalizes to the validation set. It represents the error on unseen data.