The gradient descent method finds the minimum of cost functions while Stochastic gradient descent is computationally more efficient and optimizes functions by introducing regularization which, in turn, reduces noise. We can add constraints to the LS problem by adding a Ridge Regression or LASSO regularization term. Hinge Loss, compromises on complexity by using iterative approaches to mitigate the squared error loss during classification. It also solves non-separable data problems by implementing iterative algorithms on SVMs.