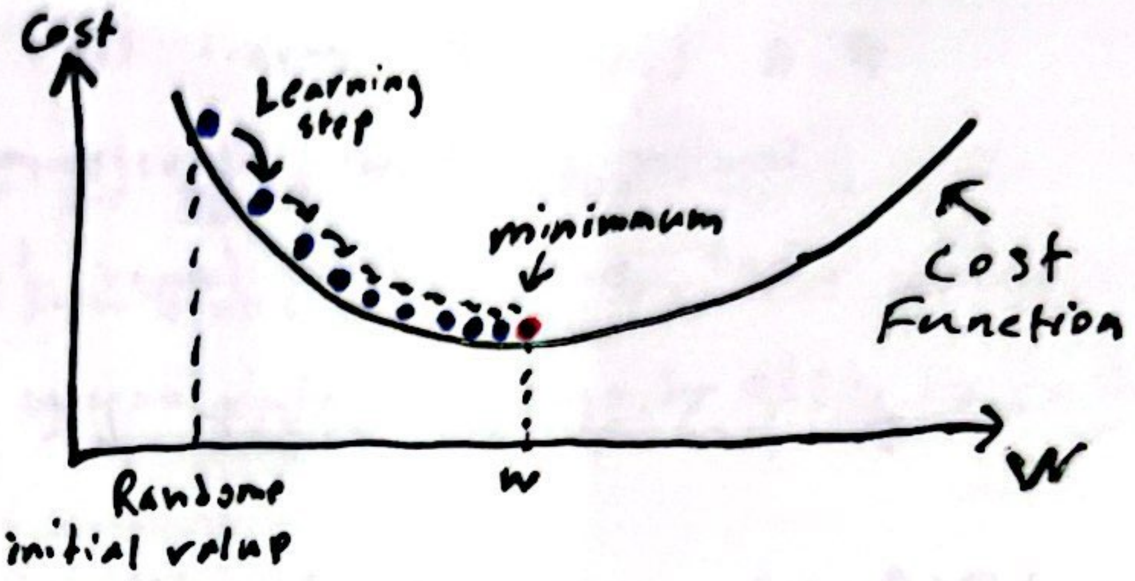
Gradient Descent (Basic)

- a very important optimization algorithm used to train ML models by iteritvly adjusting model parameters to minimize errors

★ it is one of the main methods of minimizing cost functions

Like a hiker trying to find the lowest point in a valley by always stepping in the steepest downhill direction

★ Gradient Decent calculates the Direction in which the models error decreases most rapidly and updates parameters accordingly



★ How: for each step it computes the gradient, essentially the slope of the error with respect to each parameter then adjust those parameters in the opposite direction of the gradient using the Learning Rate to determine the step size this process goes on until the model reaches a minimum error or stops improving significatly.

Momentum-Based Gradient decent

- just like a ball rolling down a hill gradient decent can get stuck in a local minimum. But a real ball has momentum which can help it shoot past local minima helping it find a lower local minimum or the global minimum

★ Since the actual space is large and high dimentional its very rare to find a global min but we find a local min thats good enough for the Model

★ Backpropigation is used to calculate gradients which GD then uses to update Model params