Gradient Descent

Gradient Descent

- Gradient descent is an optimization algorithm used to minimize some function (cost function) by iteratively moving in the direction of steepest descent as defined by the negative of the gradient.

- In machine learning, gradient descent to update the parameters of the model.

Algorithm

Repeat until Convergence
{

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

}

Mathematics of Gradient Descent

Cost Function - Mean Squared Error (MSE)

Consider the Following Equation of a straight line

$$y = mx + c$$

The Cost function for the above Equation can be written as

$$f(m,c) = \frac{1}{n}\sum_{i=1}^{n}(y_i - y\_pred_i)$$

$$f(m,c) = \frac{1}{n}\sum_{i=1}^{n}(y_i - (mx_i + c))$$

Partial Derivate with respect to slope and intercept

$$\frac{\partial(f(m,c))}{\partial m} = -\frac{2}{n}\sum_{i=1}^{n} x_i(y_i - (mx_i + c))$$

$$\frac{\partial(f(m,c))}{\partial m} = -\frac{2}{n}\sum_{i=1}^{n}(y_i - (mx_i + c))$$

New Slope and Intercept Values

$$Slope_n = Slope_{n-1} - \alpha \frac{\partial(f(m,c))}{\partial m}$$

$$Intercept_n = Intercept_{n-1} - \alpha \frac{\partial(f(m,c))}{\partial c}$$

$$\alpha \rightarrow LearningRate$$

$$\frac{\partial}{\partial m}, \frac{\partial}{\partial c} \rightarrow PartialDerivative$$