

Loss Function is a measure of how well the algorithm performs. It shows the Error in the model and how much it differs from expected optimal output.

### Difference between Cost Function and Loss Function

Cost Function is when a cost is associated with a decision. Total Aggregate of all the loss functions is the Cost Function. Loss Function is the loss per observation whereas Cost function is the loss for the entire dataset.

### Types of Loss Functions-Classification and Regression Loss

#### A)Regression Loss:

1)Mean Absolute Error:  $(\sum_{i=1}^n |Y_i - X_i|) / n$  where,  $Y_i$ =Prediction,  $X_i$ =Actual Value and  $n$ =total observations

2)Mean Squared Error:  $1/n (\sum_{i=1}^n (Y_i - Y'_i)^2)$  where,  $n$ =total observations,  $Y_i$ =actual observations,  $Y'_i$ =Predicted Values

3)Huber loss is defined as:  $\text{error}^2/2$ , if  $\text{error} < \text{delta}$  (ie, if it is a small error)  $\text{delta} * (|\text{error}| - \text{delta}/2)$ , otherwise (  $|\text{error}|$  means the absolute value error)

#### B)Categorical Loss

##### 1)Binary Cross Entropy

$$\text{Log Loss} = -\frac{1}{N} \sum_{i=1}^N y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)$$

- $y_i$  – actual values
- $\hat{y}_i$  – Neural Network prediction

##### 2)Categorical Cross Entropy

$$\text{Loss} = - \sum_{j=1}^K y_j \log(\hat{y}_j)$$

where k is number of classes in the data

$$\text{Cost} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k [y_{ij} \log(\hat{y}_{ij})]$$

where

- k is classes,
- y = actual value
- yhat – Neural Network prediction