## Regression Cost Fn

1) Mean Squared Error

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad y_i = \text{Actual}, \ \hat{y}_i = \text{predicted}$$

- used for regression tasks
- Sensitive to outliers due to squared term (i.e sensitive to very small / large values)
- Average squared difference

2) Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

- Average absolute difference
- used for regression tasks, equal weights to all errors
- less sensitive to large errors compared to MSE

3) Huber loss

- Combination of MSE & MAE
- quadratic for small errors, linear for large errors (robust)

$$L(a) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2, & |y_i - \hat{y}_i| < \delta \quad : \text{small errors} \\ \delta|y_i - \hat{y}_i| - \frac{1}{2}\delta^2, & |y_i - \hat{y}_i| \geq \delta \quad : \text{large errors} \end{cases}$$

$\delta$ = threshold

- Good for Regression with moderate outliers.

---

## Classification Cost fn

1) Binary Cross Entropy loss (Log loss)

$$L = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log_2(\hat{y}_i) + (1-y_i)\log_2(1-\hat{y}_i)\right]$$

2) Categorical Cross Entropy

$$L = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K} y_{ij}\log_2(\hat{y}_{ij})$$ , K = No. of classes

$y_{ij} = 1$ if class $j$ is true label for instance $i$ else 0

3) Hinge Loss

- used in binary classification, esp. SVMs.
- Encourages correct classification with margin of separation

$$L = \frac{1}{n}\sum_{i=1}^{n}\max(0, 1 - y_i\hat{y}_i)$$

## Custom Loss Fn

1) Focal loss

- used in imbalance classification tasks.
- Modifies cross entropy to focus on minority classes

$$L = -\alpha(1-\hat{y}_i)^{\gamma}\log(\hat{y}_i)$$

$\alpha$: Weight assign to class, High $\alpha$ = More contribution

$\gamma$: Focusing parameter → focus on minority classes

Set $\alpha \geq 0.5$ for minority class

set $\gamma = 2$ to focus on difficult eg. like rare true positives.