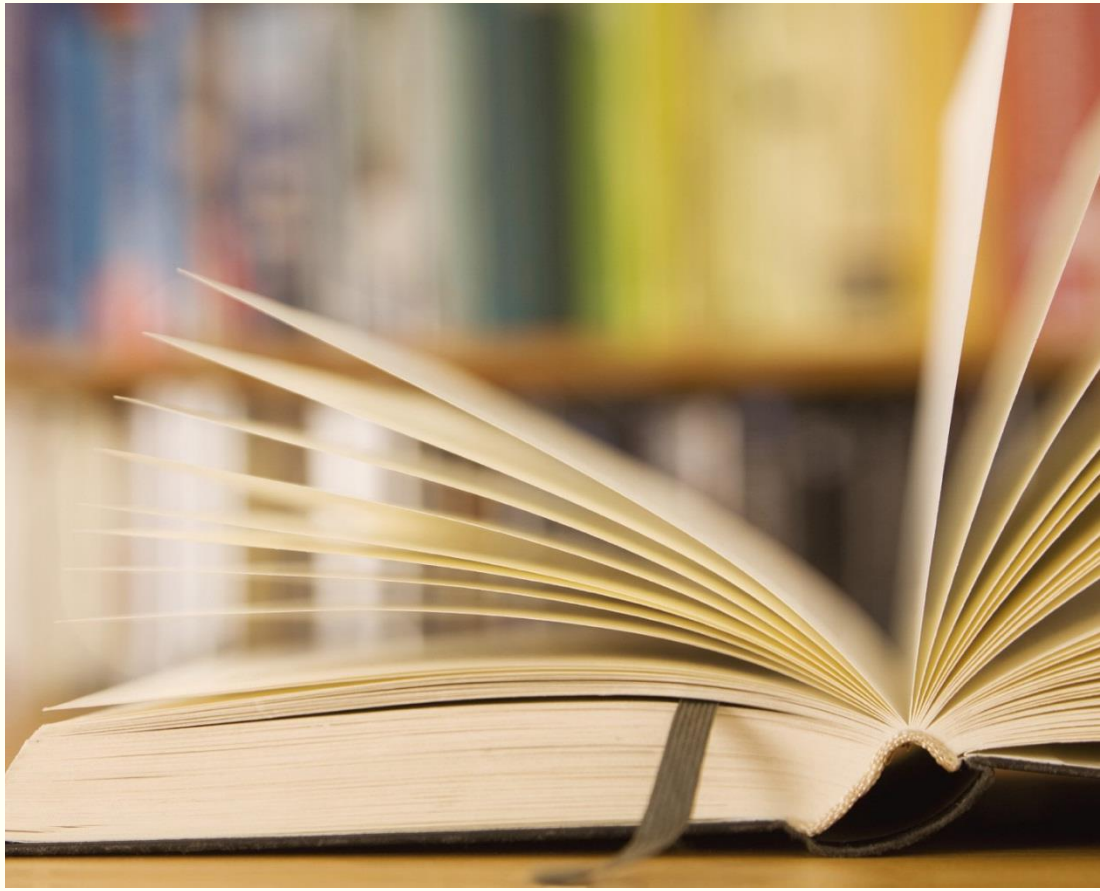
HUBER LOSS

Yongbing Anita Tao-Topinoja



Combination of MSE and MAE – Huber Loss function

- MSE (Mean Squared Error)

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

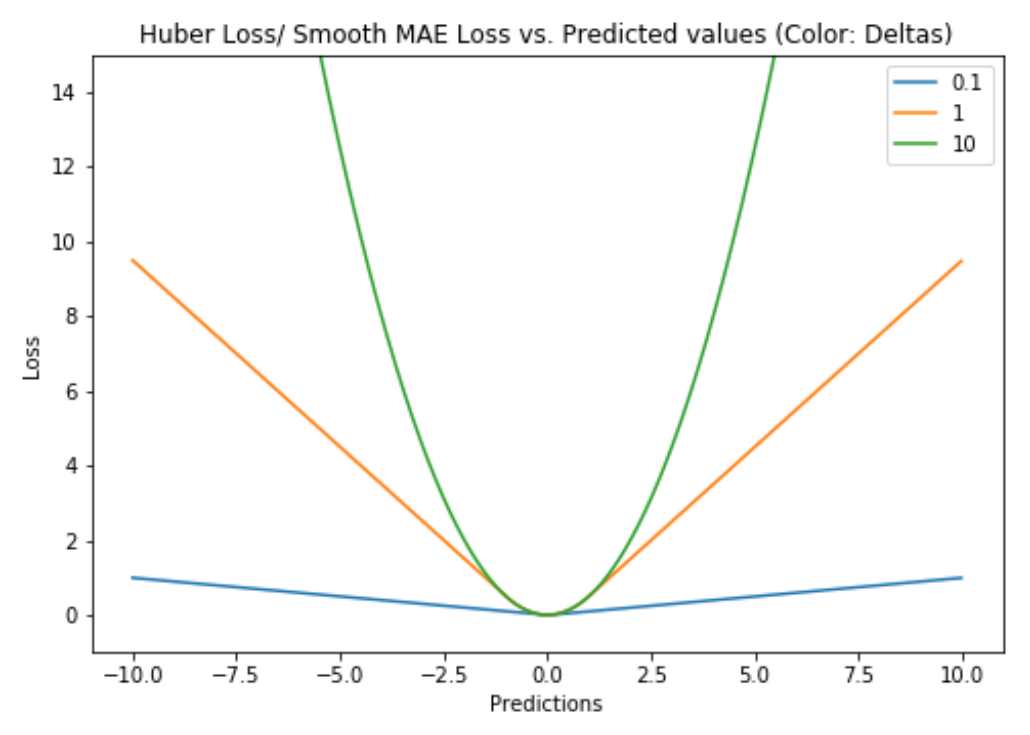
- Huber loss function

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & \text{for} |y - f(x)| \le \delta, \\ \delta\,|y - f(x)| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases}$$

- MAE (Mean Absolute Error)

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

Huber loss, MSE, MAE, and the parameter δ



- Huber loss is the transition between MSE and MAE. It is not as same as MSE to ignore the outliers, or not as same as MAE to exaggerate the outliers.
- Huber loss function is a regression function.
- Huber loss curved at the bottom, then straight lines to the top.
- The parameter δ of Huber loss in machine learning is default as 1.35, which means to achieve 95% statistical efficiency.
- Smaller parameter δ, more outliers will be considered; Larger parameter, less outliers will be considered. Neither of them are statistical efficient.

Thank you