

A comment on the loss functions

- ▶ Note that MSE and MAE are both a function of the error, $(y - \hat{f})$
- ▶ If we are predicting, say, the height of a person, each of these loss functions will penalize an error of 1cm by the same amount for any actual person's height
- ▶ This might be OK, since 1cm error is roughly equally acceptable for the tallest and shortest person, as well as for anyone in between
- ▶ If, however, we are predicting, e.g., distances to astronomical objects, an error of 5 light years is not equally fine for a galaxy 10 billion light years away and, say, Proxima Centauri, which is just about 4.25 light years away
 - So, these loss functions would probably not be applicable for such a task

$$\text{MSE} = \frac{1}{N} \sum_{i=1 \dots N} (y_i - \hat{f}_\theta(x_i))^2$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1 \dots N} |y_i - \hat{f}_\theta(x_i)|$$

- ▶ Try to make a similar analysis for MAPE and MSLE, to see what kind of errors would be penalized the same with them.