Machine Learning

SIS 1.

1. Comparing MSE and RMSE

True values : $y_i = [500, 300, 800, 400, 6000]$

Predicted values : $\hat{y}_i = [450, 350, 780, 420, 910]$

Step 1: Compute the squared errors

$(y_i - \hat{y}_i)^2$:

1. $(500 - 450)^2 = 50^2 = 2500$
2. $(300 - 350)^2 = (-50)^2 = 2500$
3. $(800 - 780)^2 = (20)^2 = 400$
4. $(400 - 420)^2 = (-20)^2 = 400$
5. $(6000 - 910)^2 = (5090)^2 = 25908100$

2. Mean Squared Error

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

$$MSE = \frac{2500 + 2500 + 400 + 400 + 25908100}{5} =$$

$= 5182780$ MSE : 5182780

3. $$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} = \sqrt{MSE}$$

$$RMSE = \sqrt{5182780} = 2276,57$$

4. Interpretation of Results

The RMSE of 2276,57 indicates that, on average, the model's predictions are off by around 2276,57 units in the same scale as the data;

5. Difference between MSE and RMSE and why RMSE is preferred?

MSE is in squared units, making it harder to interpret.

RMSE is in the same units as the data, making it easier to understand

RMSE is preferred, because it is more interpretable and directly reflects the average error in real life terms, while MSÈ can be dominated by large errors

redictions
nits
RMSE
harder
data,
more
the
hile
errors

Exercise : Bias Variance decomposition of MSE

$$MSE = E[(\hat{Y} - Y)^2] = (Bias(\hat{Y}))^2 + Var(\hat{Y})$$

Expand the squared error :

$$(\hat{Y} - Y) = (\hat{Y} - E[\hat{Y}]) + (E[\hat{Y}] - Y)$$

Substitute this into the squared terms

$$(\hat{Y} - Y)^2 = [(\hat{Y} - E[\hat{Y}]) + (E[\hat{Y}] - Y)]^2$$

Expand the square

$$(\hat{Y} - Y)^2 = (\hat{Y} - E[\hat{Y}])^2 + 2(\hat{Y} - E[\hat{Y}])(E[\hat{Y}] - Y) + (E[\hat{Y}] - Y)^2$$

Applying expectation E[.] to both sides:

$$E[(\hat{Y} - Y)^2] = E[(\hat{Y} - E[\hat{Y}])]^2 + 2E[(\hat{Y} - E[\hat{Y}])(E[\hat{Y}] - Y)] + E[(E[\hat{Y}] - Y)^2]$$

Simplify terms

first term $E[(\hat{Y} - E[\hat{Y}])^2]$ variance of $\hat{Y}$, $Var(\hat{Y})$

last term $E[(E[\hat{Y}] - Y)^2]$ is square of the $(Bias(\hat{Y}))^2$

middle term $2E[(\hat{Y} - E[\hat{Y}])(E[\hat{Y}] - Y)]$ since $E[\hat{Y} - E[\hat{Y}]] = 0$

we are left with :

$$MSE = E[(\hat{Y} - Y)^2] = Var(\hat{Y}) + (Bias(\hat{Y}))^2$$

MSE equals variance plus squared bias