Testing model performance

Mean absolute error (MAE)

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

y_i = actual value (true mood score)

$\hat{y}_i$ = predicted value (from the model)

n = n test samples

Actual y	Predicted $\hat{y}$	Absolute error
7	6.5	0.5
4	5.2	1.2
9	8.7	0.3

$$MAE = \frac{0.5 + 1.2 + 0.3}{3} = \frac{2.0}{3} = 0.67$$

Lower MAE = better model performance

Mean squared error (MSE)

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

squares the error, so it penalize larger errors more heavily than MAE

Actual y	Predicted $\hat{y}$	Error $(y - \hat{y})$	Squared error
7	6.5	0.5	0.25
4	5.2	-1.2	1.44
9	8.7	0.3	0.09

$$MSE = \frac{0.25 + 1.44 + 0.09}{3} = \frac{1.78}{3} = 0.593$$

Lower MSE = Better model

R^2 score (coefficient of determination)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

y_i = actual val

$\hat{y}_i$ = predicted value

$\bar{y}$ = Mean of actual values

Measures how much of the variation in mood scores the model explains
Compares model error to baseline errors

$R^2 = 1$ perfect model (100% acc)

$R^2 = 0$ model is no better than predicting the mean

$R^2 < 0$ model is worse than guessing the mean

Actual y	Predicted $\hat{y}$	Error $(y-\hat{y})$	Squared error
7	6.5	0.5	0.25
4	5.2	-1.2	1.44
9	8.7	0.3	0.09

sum of squared errors
① $SSE = 0.25 + 1.44 + 0.09 = 1.78$

② SST (total variation)

$$SST = \sum (y_i - \bar{y})^2$$

if $\bar{y} = 6.67$ (avg mood score):

$$SST = (7-6.67)^2 + (4-6.67) + (9-6.67)^2 = 4.67$$

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{1.78}{4.67} = 0.62$$ → 62% of mood score variation by sleep, stress, and exercise