ML

Regression Previous Year Question

- 1.The fuel efficiency of different cars in miles per gallon (mpg) with respect to its weight is given in the following table.

Weight	Mpg
3504	18
3693	15
3436	18
3433	16
3449	17
4341	15
4354	14
4312	14
4425	14
3850	15

- Find the least square estimation of the line y = β0 + β1X , such that β0 and β1 are the parameters of the line.
- **Ans:**
- Here n =10

$$\bar{x} = 38797/10 = 3879.7$$

$$\bar{y} = 156/10 = 15.6$$

$$Cov(x, y) = \frac{1}{9}(-5218.2) = -579.8$$

$$Var(x) = \frac{1}{9}(1680796.1) = 186755.12$$

- β1 = -579.8/186755.12 = -0.003104, β0 = 15.6 - - (-0.003104) * 3879.7 = 15.6 + 12.04 = 27.64

Regression Previous Year Question

- 2. Fit a straight line Y = a + bX to the data by the method of least square.

X	1	3	4	2	5
Y	3	4	5	2	1

- 3. Fit a straight line Y = a + bX to the data by the method of least square.

X	5	10	15	20	25
Y	16	19	23	20	30

- Q2 & Q3 are homework.

Regression Previous Year Question

❖ 4.Explain different evaluation metrics/errors used in measuring the performance of a regression problem?

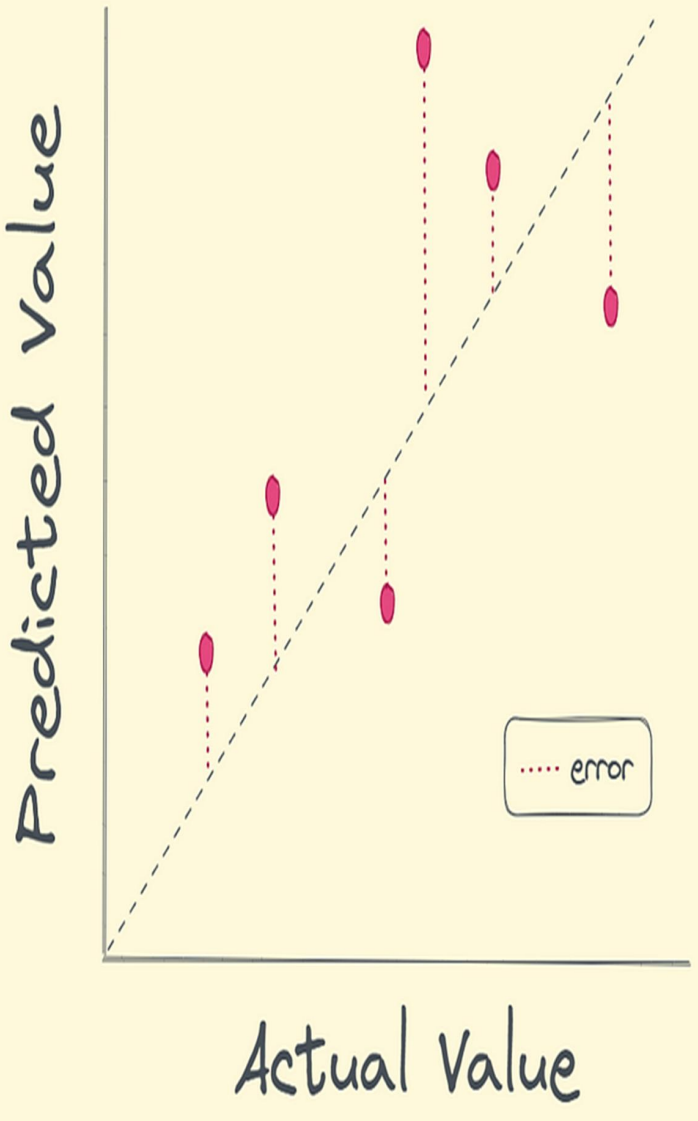
$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

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

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

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

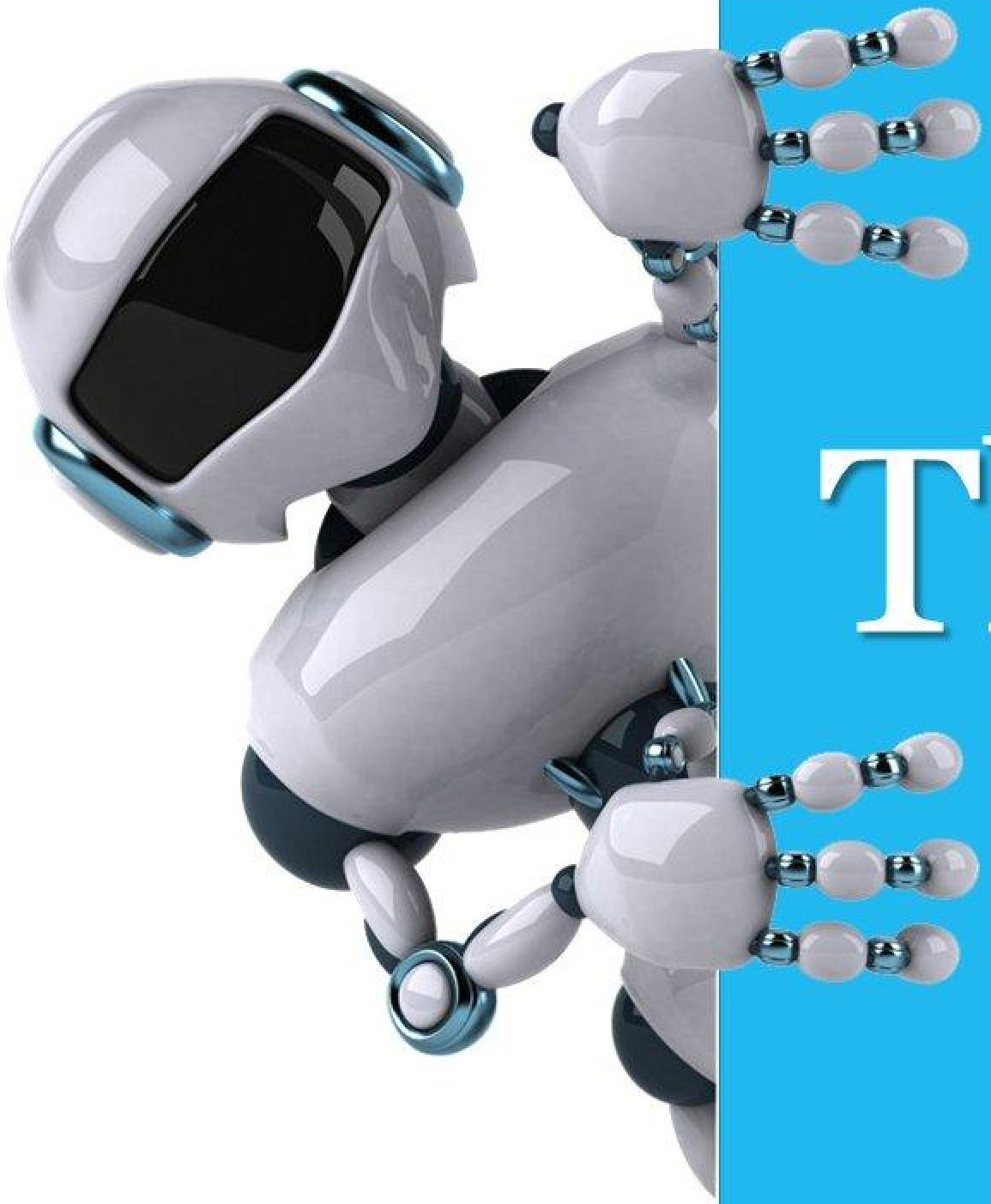
$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$



Metric	When to use
MAE	We want a metric that gives equal weight to all errors and is less sensitive to outliers. (Low value is good)
MSE	We want a metric that penalizes larger errors more than smaller ones.(Low value is good)
RMSE	We want a metric that penalizes larger errors more than smaller ones.(Low value is good)
R2	We want to understand the proportion of the variance in the dependent variable that is predictable from the independent variables. (High value is good)
MAPE	We want to express errors as a percentage of the actual values and want a metric that is easy to interpret.(Low value is good)

Regression Previous Year Question

- 5.Write the effect of learning rate on the performance of a Gradient Descent algorithm?
- **Ans:- High Learning Rate:** A high learning rate can lead to **faster convergence,** as the algorithm takes larger steps towards the minimum. However, it **may also cause overshooting**, where the algorithm **may oscillate around the minimum or even fail to converge.**
- **Low Learning Rate:** A low learning rate **slows down the convergence process**, as the algorithm takes smaller steps. While it may **increase the likelihood of convergence**, it can be **computationally expensive and time-consuming.**
- **Optimal Learning Rate:** An optimal learning rate allows the algorithm to **reach a sufficiently accurate solution within a reasonable number of iterations.**
- **Adaptive Learning Rates:** Some advanced optimization algorithms incorporate adaptive learning rates, **adjusting the learning rate during training based on the observed behavior of the optimization process.** This adaptability can **enhance robustness.**



Thank you