Xgboost Regressor Mh Algorithm :→

Dataset {Regression}

Exp	Gap	Salary	R_1	$\hat{y}$	R_2
2	Yes	40K	-11	49.9	-9.9
2.5	Yes	42K	-9	49.9	-7.9
3	No	52K	1	51.5	0.5
4	No	60K	9	51.5	8.5
4.5	Yes	62K	11	52.1	9.9

$[51 + (0.1)(-10)] = 51 - 0.1 = 49.9$

$[51 + (0.1)(5)] = 51.5$

$[51 + 0.1(11)] = 51 + 1.1 = 52.1$

Average ≈ 51K ($\hat{y}$)

$$\text{Similarity weight} = \frac{\sum(\text{Residual})^2}{\sum p_r(1-p_r)}$$

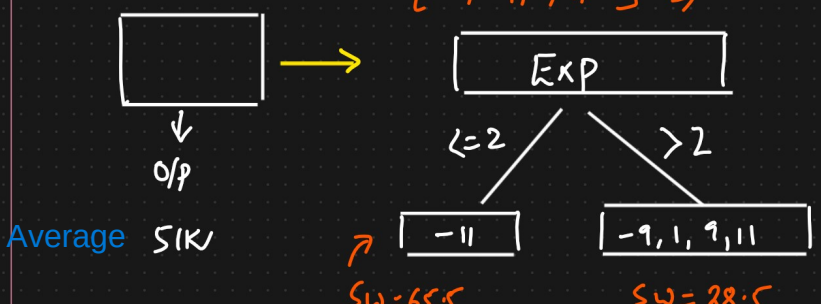
For regression, it will change slightly.

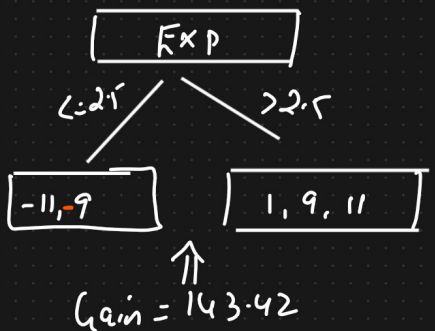
Gain

Steps

① Create a Base Model ② Residual Computation ③ Construct DT1 using {x_i, R_1} (As I/P, As O/P)

[-11, -9, 1, 9, 11] ⇒ SW = 0.16.





④ $$\text{Similarity weight} = \frac{\sum(\text{Residual})^2}{\text{No. of Residuals} + \lambda}$$ → Hyperparameter

$\lambda = 1$

$\lambda \uparrow$ SW ↓

$SW(\text{Right child}) = \frac{121}{1+1} = 121/2 = 65.5$ //

$$SW(\text{Right child}) = \frac{(-9+1+9+11)^2}{4+1}$$

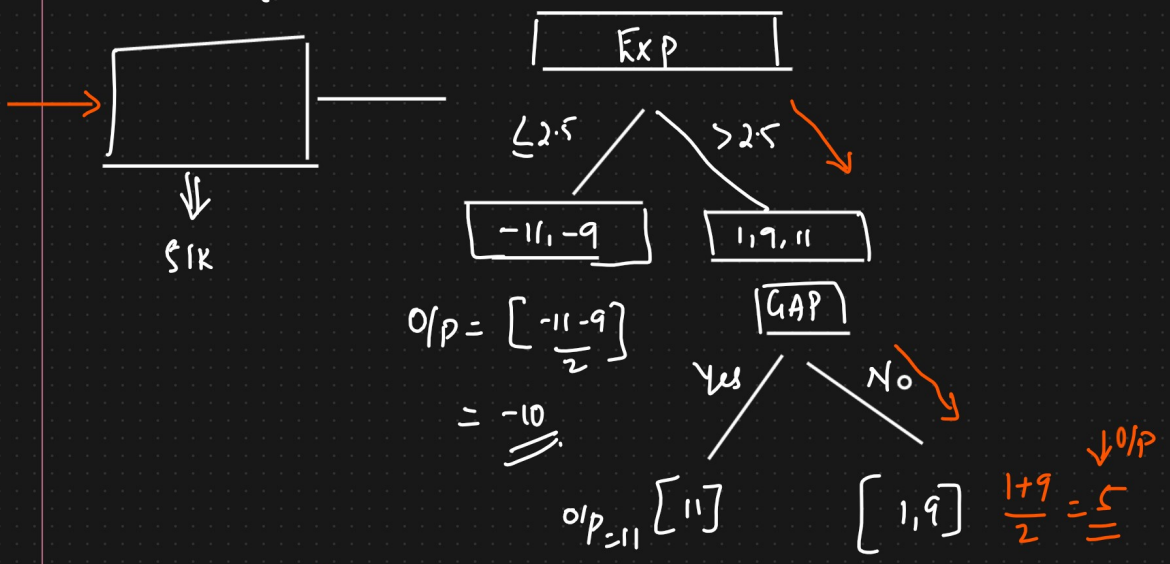
$$= \frac{144}{5} = 28.5$$

(5) Calculate Gain

$$\text{Gain} = 65.5 + 28.5 - 0.16$$

$$= 98.34$$

DT1



α = Learning Rate $\alpha = 0.1$ => Hyperparameter

$$\text{XGBoost Classifier} = \text{Base Learner} + \alpha_1(DT1) + \alpha_2(DT2) + \cdots + \alpha_n(DTn)$$

$$\text{XGBoost Regressor} = \overset{O/P}{51K} + 0.1(5)$$

$$= 51 + 0.5$$

$$= 51.5$$

For regression :-

$$\text{Similarity weight} = \frac{\Sigma(\text{Residual})^2}{\text{No. of Residuals} + \lambda}$$

[redacted]

For classification :-

$$\text{Similarity Weight} = \frac{\sum (\text{Residual})^2}{\sum Pr(1-Pr) + \boxed{\lambda}}$$