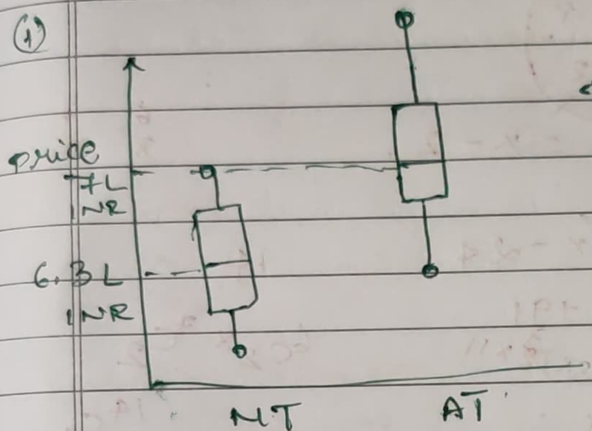
DATE: / /
PAGE:

Questions.

(1)



$\hat{price} = \beta_0 + \beta_1 (MT|AT)$

what is β_0 & β_1
(No usage of computers)

$= 0.7 +$

(30 - 35%)
rate

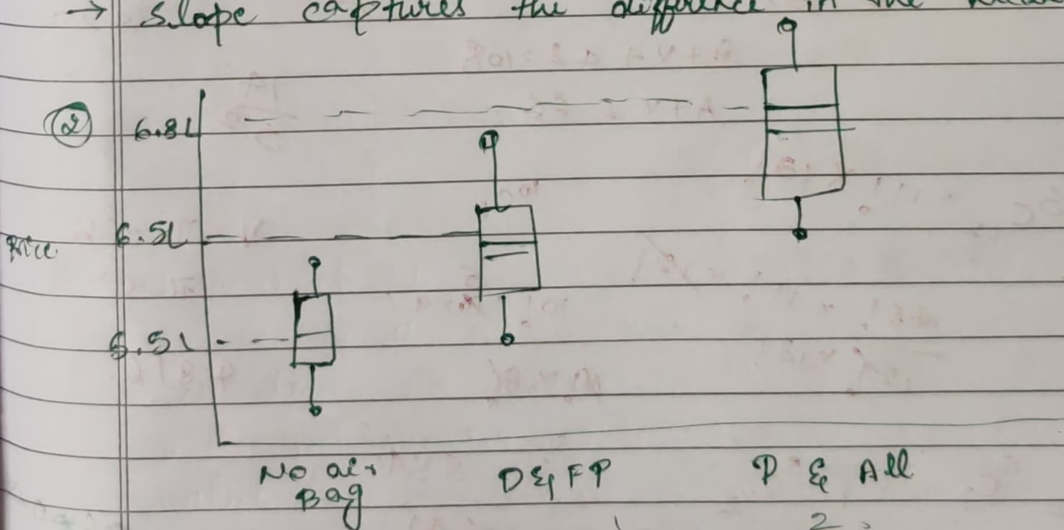
soln :- $6.3 + 0.7\ X$

→ Encode MT = 0 & AT = 1
when I need to predict price of MT, (MT = 0) I need to get 6.3. i.e $\beta_0 = 6.3$.

→ If Encoded as MT = 1 & AT = 0
$7 - 0.7X$

→ Slope captures the difference in the mean.

(2)



$\beta_0 + \beta_1 \binom{Airbags}{DFP} + \cancel{\beta_2\ DFP}$.

5.5 + (1 DFP) + 1.3 DA

- For ordinal scale category, there is no consistancy in change / difference
- Hence, we dont do ordinal encoding for linear regression

- So we do dummy encoding

Nominal } Categorical
Ordinal

Interval } Numeric
Ratio

DFP	D All
1	0
0	1
0	0

$\hat{P} = B_0 + B_1 d_1 + B_2 d_2$.

$= 5.5 \quad 0 \quad 0 \rightarrow$ No air bags.

$\quad 1 \quad 0 \rightarrow$ DFP.

$= 5.5 + 1 d_1 + 1.3 d_2$.

$5.5 \quad 0 \quad 0$

$= \cancel{6.5} \quad 1 \quad 0 = 6.5$ DFP.

$0 \quad 1$

5.5

Now Drop DFP.

	NA	D All
NA	1	0
DFP	0	0
D All	0	1

$\hat{P} = B_0 + B_1 d_1 + B_2 d_2$

$= 6.5 - 1 d_1 + 0.3 d_2$.

Now Drop D All

	NA	DFP
NA	1	0
DFP	0	1
D All	0	0

$\hat{P} = 6.8 - 1.3 d_1 - 0.3 d_2$

(3) Covariance & Correlation coefficient Understanding.

Soln:- Unit gets disrupted. In Covariance.

Correlation of Coefficient $r_{x,y} = \frac{Cov(x,y)}{\sigma_x, \sigma_y}$; Takes cov(x, y) in z score (x,y)

$r_{xy} = cov(\text{standardized } (x \,\&\, y))$

Calculating using formula - Analytical / Parametric Approach

DATE: / /
PAGE:

4) Strength & Weakness of Linear Regressor Model

- Strength → It is very simple - decoded easily
- Weakness → Accuracy is low - slightly high bias error

for benefit of Inference

can reduce variance error at the cost of high bias error - Ridge, Lasso, EN.

overfit ← High variance Error
↳ No consistency but low error (RMSE)

P_1	P_2	P_3
89	10	100
85	18	60
86	12	12
83	14	20
82	13	62
μ = 84	μ = ~~20~~ 16	μ 64

High bias Error
↳ Consistency but ~~low~~ high error (RMSE)
↓
underfit

variance ~~error~~ ↑ ↓ bias error ↓

(5) Drawback of Train - Test - split validation.

Soln: It gives a number & point estimation is done. We cannot tell if new data will be greater or lesser than this estimated value.

Hence, we go with K fold cross validation.

(6) What is the R^2 of Null model regressor?

Null model → Model having no feature.

~~R^2 of null model~~ $\hat{y} = \bar{y}$ - mean of target actual

R^2 of null model is 0

For Null Model :-

$\cancel{\bar{y} = \beta_0 +}$ for Null, $x = 0$.

$\beta_0 = \bar{y} - b_1 \bar{x}$

$\beta_0 = \bar{y}$

$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.

$\hat{y} = \beta_0 = \bar{y}$

7) What is the AUC score of Null model classifier

Soln:- 0.5.

8) Use of Base line score / why do we talk about null model (Base line ref)

9) What do you mean by deviance = 0.

→ Indicates perfectly fit data

10) classical probability - Sample space is fixed

Empirical probability - Having historical probability

Subjective probability - No historical probability
Need domain expert

11) If something is revealed in conditional probability, the sample space reduced.

12) Training, Validation & Testing records.

1000 → No of sample records

Test / pre production data ← 100

900 Train

Train

Test → Validation data

DATE: / /
PAGE:

12) Impact when n neighbour is increased.
Soln:- Bias Increases. - Model behaves like a Null model
Eg:- 1's → 23
0's → 18.
If I give n=41, it will consider all the points 41 as its neighbours and picks the one which has maximum or more ~~&~~ no of records.

- If n neighbour is decreased - It leads to overfit (high variance Error)

13) KNN cannot be boosted but can be bagged
Boost to reduce bias error
No luxury of reducing bias error later.
Try to design with least bias error.

Cannot boost as there are no coeff kind of thing.

14) Difference b/w hyperparameter & model parameter

All models have model parameter except KNN
model understanding data in its own way
using these parameters, model learns the ~~pa~~ data

Hyperparameter defines the structure / design spec of the model

15) Neural n/w without hidden layer acts as (logistic) Linear Model.