=> Linear Regression Algorithm  $h\theta(x) = \theta_0 + \theta_1 \propto - \text{Simple linear Reglession}.$ 

NB (91) = 00+0131, + 02 €2+03 x3 + .... + On xn

=> Convergence Algorithm:-
$$J(\theta_0, \theta_1) = \frac{1}{m} \left( h \theta(x)^{(i)} - y^{(i)} \right)^2$$

Lost Function

Loss Function = 
$$\left(h_{\theta}(x)^{(i)} - y^{(i)}\right)^2 \Rightarrow \left(\hat{y}_i - y^{(i)}\right)^2$$

where  $\hat{y_i}$  - Predicted value.

as slow us hapothesis testing wind in linear sequestion? Derivatives:

the multipariote resignation of the south

$$\frac{\partial}{\partial x} (x)^m = m_2 x^{m-1} (x) \text{ and } 2x^m = 0.12$$

$$\frac{\partial}{\partial x} (x+1)^2 = 2 \times (x+1) \times (1+0)$$

$$= 2(x+1)$$

J=0 
$$\frac{1}{3}$$
  $\frac{1}{3}$   $\frac{1}{3}$ 



100	Disadvantage:-
	Disadvantage:- wistory gomon agrid
(3)	
	Penalizing the Essos { Changing the unit }
(2)	Penalizing the Essos & Changing the unit?
	In perhaps makes mater as as to the us
(IT)	MAE Smean Abdolute Flags?
0	MAE [Mean Absolute Elava]  MAE = $\frac{1}{n} \frac{1}{i=1} \frac{1}{y-y} \frac{15-3}{15-3} = 1-3 = 1.3$
	MDE - 1 6 1 1 - 2 1 - 3 = 1 - 3 = 1 - 3 = 1 - 3 = 1 - 3 = 1 - 3
	m != 1 9 11
	And the state of t
	Advantages 6000002 9 0
	To van faces
0	This is latent to puttions
_	This is hobrest to outliers  It will also be in the same unit
3	The way and the same will be a same
	Diadaha Biasa San Dan Maria
	Diadvantageer- song of more and the more and
0	Converge usually takes make time Optimization
	convergence usually takes more time. Optimization is a complex task.
.00	Asserted to the second
(2)	Time Compunition
	Time Comsuming
3	Huber Loss It is a combination of MSE and MAE
0	The town of the to
(TV)	RMSE = MSEOW LATER & John out IT
	- Unit will be same in RMSE
	- It is not robust to outliers
	- This is also differentiable

	Performance Matrix:	
	To check wheather the model is good of not,	in
	case of linear regression, we have to see	
	case of linear regression, we have to see performance matthix	0
	60 7 90	
	In performance Mathix, we see two things.	
	In Performance Mathix, we see two things.	(m)
	1 R Squared Error  2 Adjusted R Squared Error	
	(2) Adjusted R Squared Edrox.	
	Grandlake Faustian and Miller to so	
0	R Squared 1994 March 19	
	R squared = 1 = 85 Res of tankal in all	0
0	time and Scrotal will ad allo like IT	
	Where SSRes = Sum of Square Residuals.	
	SSTotal = Sum of Squale Average.	
Nother	R Squared = 1 - E (yi - yi)	0
	Resquared = 1- & (4) = 4) solgmon se	
	$\frac{1}{2}\left(y_1-\overline{y}\right)^2$ musumal smit	(9)
	j=1	
	(alass) y	
AM I	where y = Average of y	(2),
	0 00	
	If the model is fitted well &	A Just
		لهو
	R Squared = 1 - Small No. x	0
	Bigger No. of	
	- This is also at tecentable	1
		90

B Somared = 1- (Small Number)	
R Squared = 1- (Small Number -> Small Number -> Small Number	mber
TRATA TATA YELD GOOD ACCURATE AND FRANK FRANK	
Therefore, If the model is fitted well	
SINDH ? ( ( = Ke je 3'Sige of Pouse ? And Par Ke VIII Pet	en se
R Squaled & 1 mit and the state of the squaled	
K 88 / 2 Mo 8/ 8/20m	
Suppose R Squared = 0.85 1.85%. Accurate.	949,
11 R Squared = 0.75 : 75% Accurate.	用
Guestian What Adjusted R2 200	
Question: What R Squared Do?	the T
are when fined on the very important	
Answer: R Squared basically measures the perfor	hmance
of the model that you have created.	
Mole -	
Answer: If R Squared value tomes -ve, then model is very bad	
Answer: If R Squared value comes - ve, then	the
model is very bad	
Low Team Duta Acculacy on Wigh English	
Doloff & (not) 12 To Tage of	
6 Adjusted R Squared	
Topo destapointes	
Adjusted $R^2 = 1 - (1 - R^2)(N-1)$	
1-9-M (889) - M-P-1	0.
Man Good Martage 700/4 Long Train	
where N= No. of data points.	BIE
P = No. of Independent features	
(002) (003)	
MORTAGIJAV DWIMIAGT	
At The state of th	
Their the model the permitted Tennity the	

City Location No. of Gender Price. size of god, muches House Bed hooms Mals According & Islam est I shopers of R2 65% sige of House } Adj R2 = 63% R<sup>2</sup> 75% & City Location? Adj R<sup>2</sup> = 73%. R2 88% { No. of Bedroom} otolosof 190%. Esta a boson Adj R2 sang 5%. Question What Adjusted R2 Do? Answer 1. The Adjusted R2 is going to evalute the features. of the model that you have cheated. Overfitting And Underfitting 10 100 4 20 MG (Bias And Variance) Thaining (700) > Model
Thaining DATA SET 1000 datapoints Test (300) -> model Test TRAINING DATASET (700) VALIDATION TRAINING . Thain the model Hyperparameter Tunning the

	-1 in 6 in march 1
	Model
29	TRAIN DATA Very Good Accuracy (85%) [BIAS]  TEST DATA Very Good Accuracy (85%) [Variance]
Sel S	TRATE DATA Meen Good Acculage (gor) [ PTAC]
	TRAIN DATA CON SPECIAL ME WORLD [ 107.) [ BINS]
T	Tear Dans Good America Comit T Mariance
- 1	1557 DATA very Good Accuracy (85%) [ variance]
	- Whenever we talk about theiring data acculary, we are brasically going to use [BIAS]
	We are basically going to use BIAS!
	- whenever we talk about test data accuracy, we are basically going to use variance.
. 83	are basically going to use variance.
	The postagnists
	Mote:-
	High Thain Data Accuracy - Low Bias.
	High Test Data Accuracy - Low Valiance
	Low Their Data Acculacy - High Bias.
2 time	·
. 2	Low Test Data Acculacy - High Valiance.
	Franch
P	Let Say for Training Data, we have.  Very Good Accuracy [90%] [ Low Bias]
	Les Say for raining sala, we have.
	very good Haway 140% I Low Blas
	For Test Data, Bad Accuracy [50%] [ High Variance]
	This condition where there is Low Bias and High
	This condition where there is Low Bias and High Variance is called Overfitting

Example:-Model For Thain Data Model Accuracy is Low [High Bias] For Test Data | Model Accuracy is Low High sompilar (128) wassing book flow of High Accuracy model is Underfitting somewhat -- whenever we talk shout tot that xouring /+ × > Thaining Datapoints.

X > Test Datapoints. g w en I wo Example of overfitting. Apil and doil to warmen atold mind well and mind datapoints Test datapoints. => Example of Underfritting of Fairing Data, we have very Good Accusacy 7907. I low Rias For Test Data, Rad Accusary [50] High Variance This condition where there is low side and High variance is called everything 1