Types of GRADKNIT DESCENIT Butch Stochastic Mini Batch
Cnp Stochastic GD GD GD Batch GD: The first GD was Ratch Gr.D. It is generally slow and some perolelom with computation of slope is found out based mall In = bo - nx (Crope) with

To Power this, Use Stochastic Or D as it find barror updates are based I fore & then change in & b.

I fast but prone to error

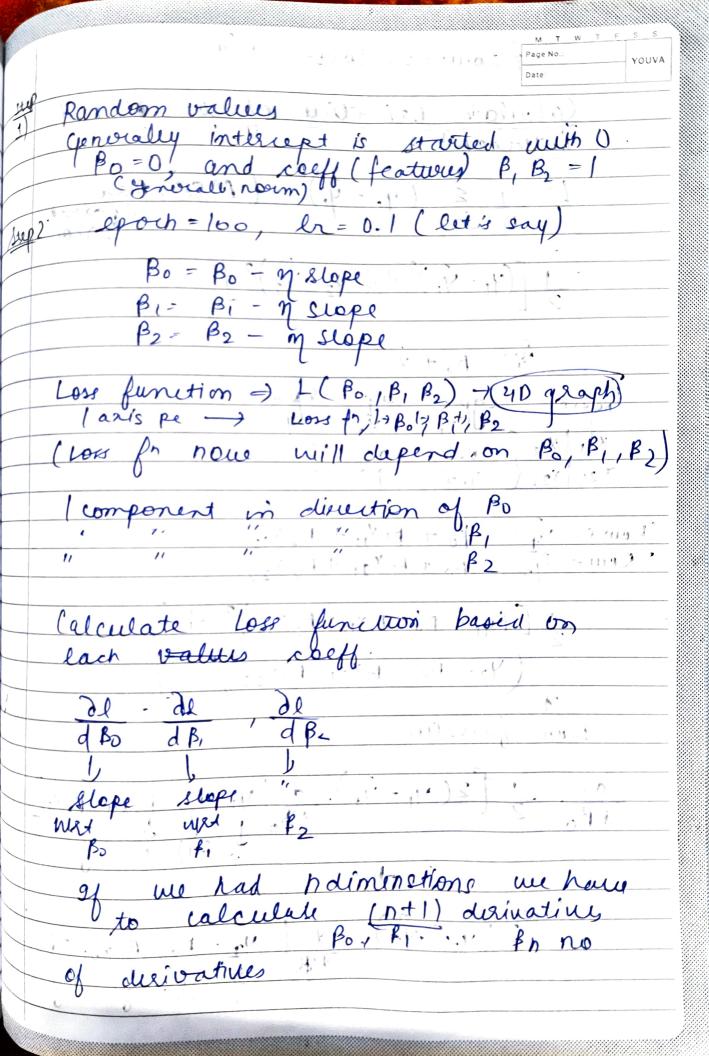
Courtalely for large datasets.)

Mini Buth Gib: In this we define batch size of 30 (originally 300) neans the Jupitate will in m & b pased on 30 1'ows, which mean Value of m & b will be updated 10 times (30×10 = 300) In stomastic, 1360 louis means 300 updates in m & b.

Note: Mostly use stochastic (and sentimes internally princ Batch) en have a conver function (data should not be too large) Batch CaD (move than 2 dimenstion Coa ig gender la Y= β0 + β1 x1 + β2 x2 + β3 x3
{ β0, β1, β2, β3 ceef to fine. If a dimension.

No of coef & Bo BI, B, Br MATHEMATICAL FORMULATION Lette say rather than in use hairs 3 dimension, and 2 rows (for understanding) 1960 - ig lea 8.1 93 3.2 7.5 95 35 y= β0 + β1 X1 + B2 X2

(cgpa) (12) (alulation -) (Po, P, , B2)



{ 20015 = 2, lals = 2+19 Calculate derivation of loss for mest intercept

L= 1 \(\geq (\frac{1}{2}, -\frac{1}{2}) \) (MSE) mean

Cyravet

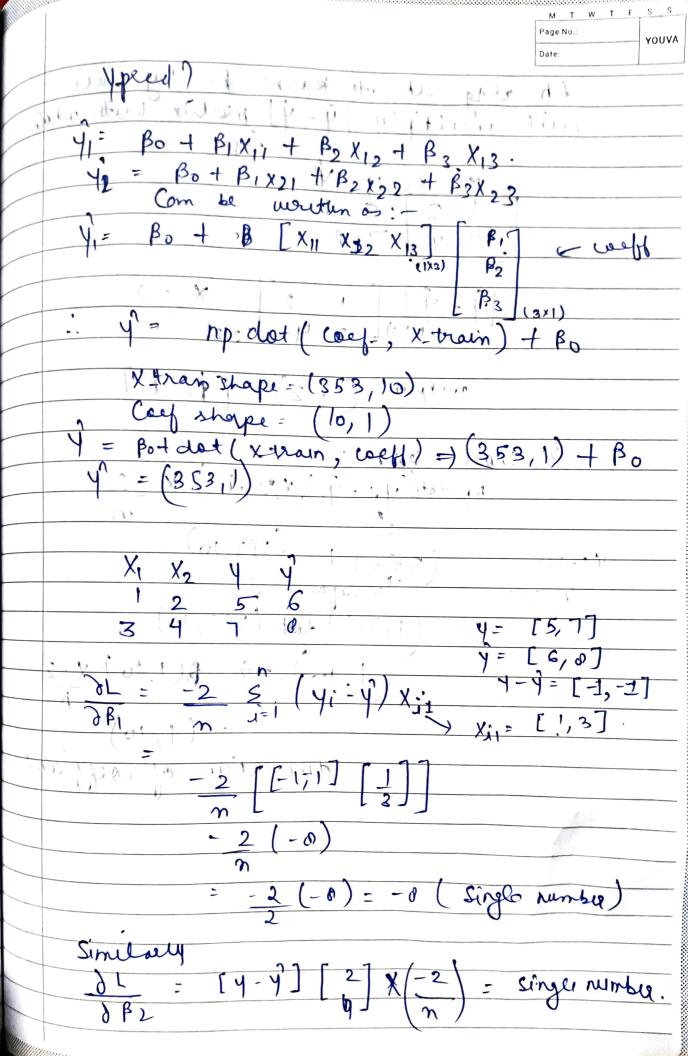
order 1 [(41 + Bo - B, X11 - B, X12) 2 + 17 Y2- β0 -β1 X21-β2 X22)2 Loss function 3l = 1 [2(41-71), d'(41-80-81X11-82X12)

3 (41) 0 $\begin{bmatrix} 2(4_2-4_2) & 10 & 4_2-\beta_0-\beta_1 x_{21}-\beta_2 x_{21} \\ d^2\beta_0 & 1 & 1 & 0 & 0 \end{bmatrix}$

30 = 2 [2(4-191)(-1) + 2(42-1/2)(-1)] JBO = -2 [(4-41) + (42-42)] Po = Po - M Slope Juli calculated calculate slope intert B. 8.1 93 3.2 7.6 9535 L= 15 (4: -7:)2 1 Long [(4) + 9,) 2 + 10 ky2 - 42 12] L-1 / POR - B2X11 - B2X12 + [42-B0-B1X21-B2X22]2 18, 2 (1,-1,) d (4,- Bo-BIX11-B2 x12) [2 (42-42) d (42-B) - B1X21 - B2X12)]

30 = 1 [2 (4= 41) (-XII) + 2 (42-42) (x Concerting for a numbers. $= -2 \left[\frac{(y_1 - y_1)(x_{11})}{(y_n - y_1)(x_{21})} \right]$ dl = -2 & (y; -4;) x;

dB1 m i=1 Similarly for B2 1002 1 dP2 n 1=1 (41-41) X12 general egn (B, 1B2... Bm) $\frac{dl}{d\beta m} = -2 \pm \left(\frac{y_1 - y_1}{y_1 - y_1} \right) \lambda_{im}$ This is the mathematical formul Slage y train yeard,



karne ke hije hug y-y) maleix mith u

1886 8 1 2 X -2 (1,2) 3 4 (2,2) h output (12) · In short trick (853,10) (353,10) 1, 10 / X-2 of welfiand duvatius 19 Lucky + 199.2 1000 gidalor comic