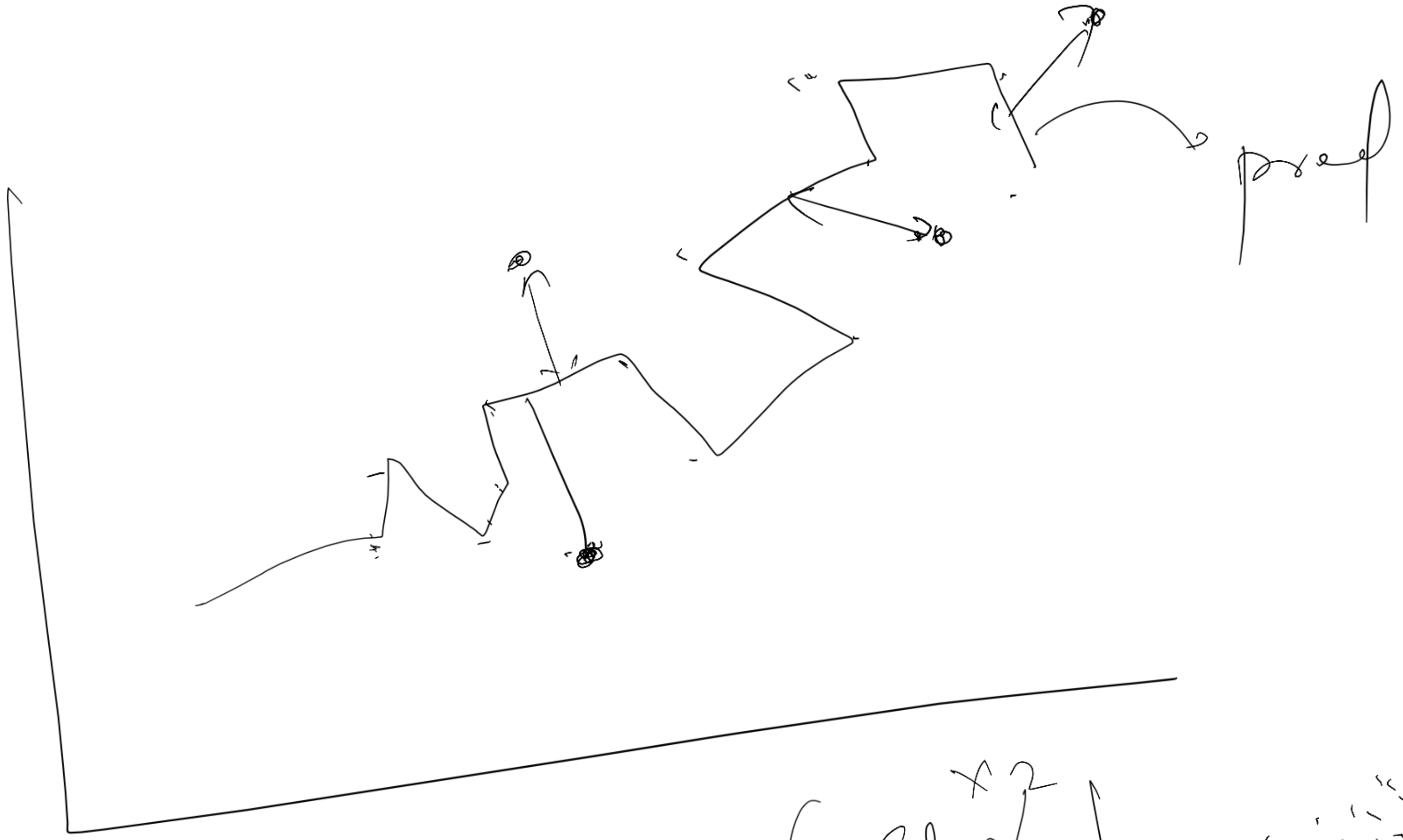
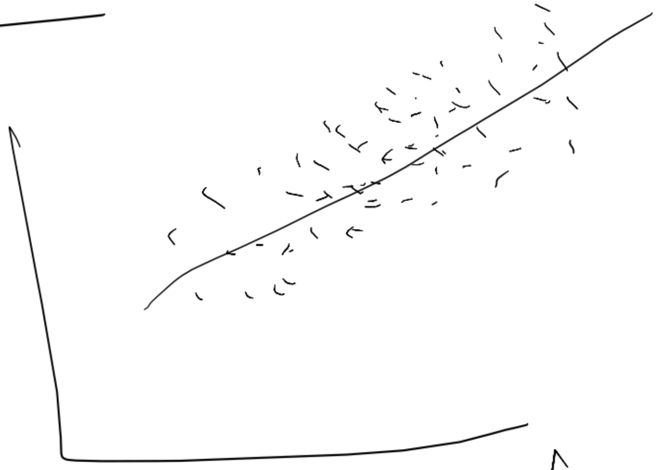
01/05/2025

Regularization :-

Overfitting :- Train error is less (eg 5% mape)
Test " " high (20% mape)

Undefitting : Train error is high
Test error is high


pref



L1 → LASSO $\boxed{\text{coeff} = 0}$ → 3 ×

$\boxed{\lambda = 1}$

L2 → RIDGE

Elastic Net $\boxed{\text{LASSO} + \text{RIDGE}}$

$$E = \sum (y - mx + b) + \lambda \sum \left| (m_1 + m_2 + m_3) \right|$$

Penalize the coeff that are high in magnitude

$$E = \sum y - mx + b + \lambda \sum (m_1^2 + m_2^2 + m_3^2)$$

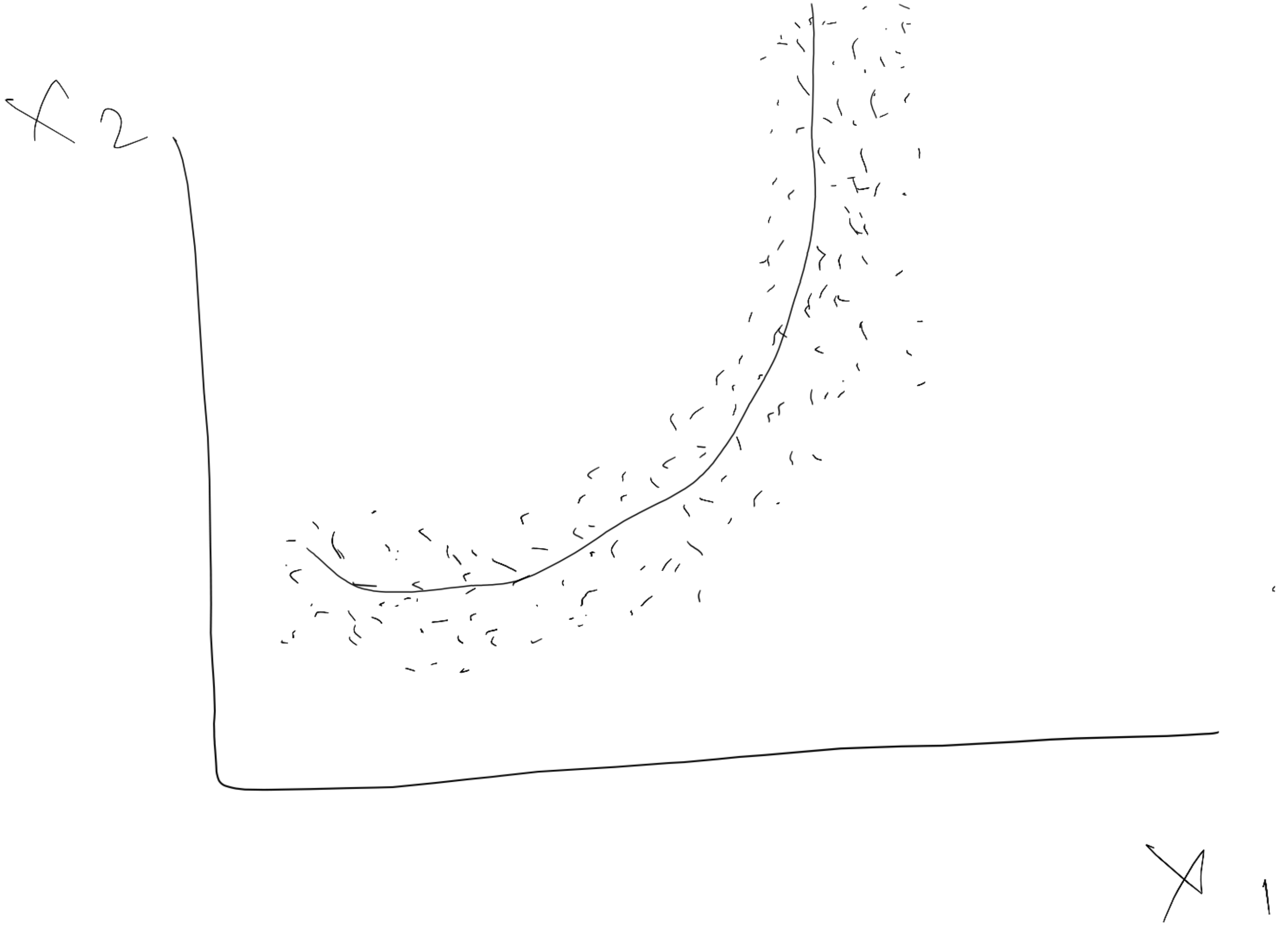
Elastic Net:

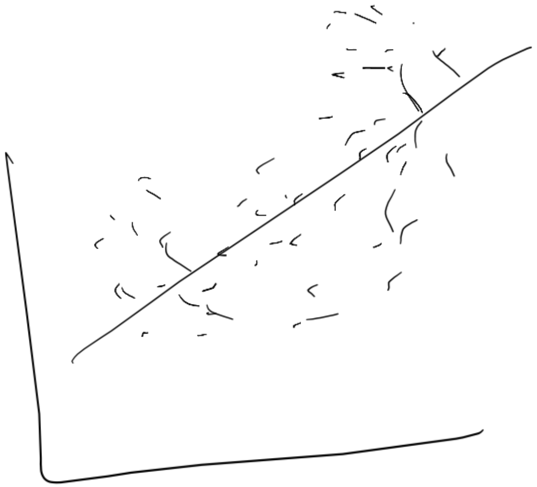
$\boxed{70\%}$ $\boxed{30\%}$

$$(y - mx + b) + \underbrace{\lambda \sum |m_1 + m_2|}_{\text{Lasso}} + \underbrace{\lambda \sum (m_1^2 + m_2^2)}_{\text{Ridge}}$$

2

coeff


X2
X1



$\boxed{y = mx + b}$ → simple line eqn

$x^2 + 2x + 4$ → Quadratic eqn

$x^3 + x^2 + 2x + 4$ → Cubic eqn

$x^4 + x^3 + x^2 + 2x + 4$ → Quatric eqn

Models → fit (X_train, y_train)

Scaling: → 1 Std scalar

70% 70 → Min Max Scalar

i/p ⇒

X_ train ⇒ fit_transform

1 → 5
2 → 6
3 → 4

$$\frac{X_i - \mu}{\sigma}$$

$\mu = 7 \quad \sigma = 1$

X_ test 30% 30 ~~transform~~

$y = mx + b \Longrightarrow$ Model

$m_2 x_2^2 + \boxed{m_1 x_1^1} + b$

$+$ inter

poly = 3

$m_3 x_1^3 + m_2 x_1^2 + m_3 x_1^1 + b$

$$\boxed{y = mx + b}$$

$(n = 100)$

$2 \rightarrow$ Variable

$(100, 2)$

$x_1 \rightarrow 1V$

$= m_1 x_1 + m_2 x_2 + b$

$x_2 \rightarrow 2V$

polynomial degree = 3

$$m_1 x_1^1 + m_2 x_1^2 + m x_1^3 + m_1 x_2^1 + m_1 x_2^2 + m_1 x_2^3 + b$$

$m_1 \Rightarrow 0.02 \quad m_2 \Rightarrow 0.1 \quad (100, 6)$

Regression: (Target → Continuous)

- Linear Regression
- L1
- L2
- ElasticNet

Logistic Regression:-

Target - Discrete

CLASSIFICATION

0

1

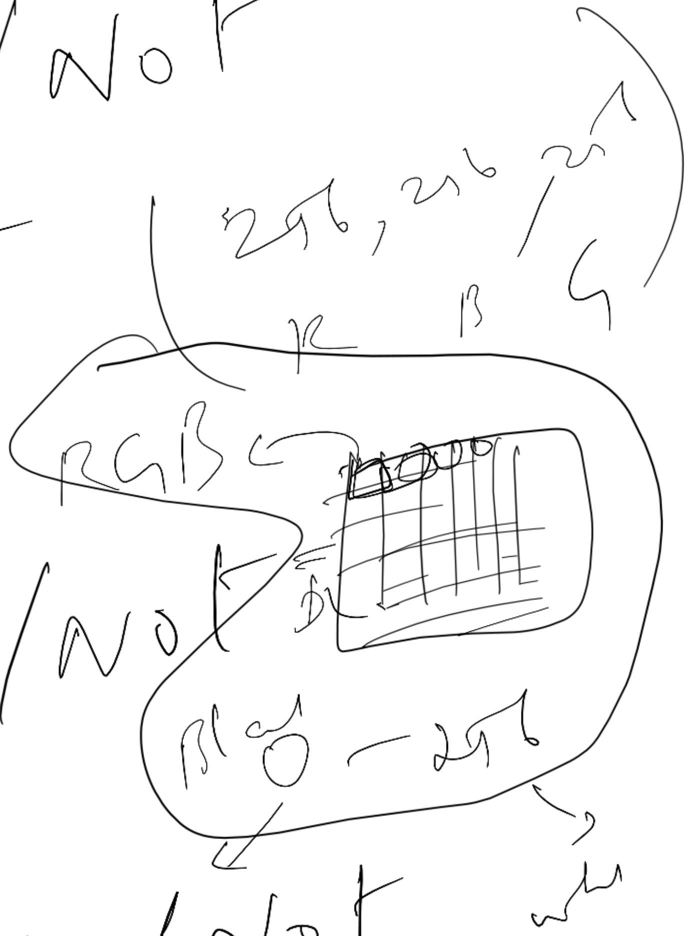
Binary (0, 1)

Multi-class (0,1,2,..)

On top of linear Regression → Continuous

Classification model Application:

1) Weather prediction → Rain / Not
2) Email → Spam / Not
3) Exam → pass / Fail
4) Online Fraud → fraudulent / Not
5) Election prediction → Win / Not
6) Credit / Loan → Give / Approved / Not



Output of Linear Regression:

$$\langle -\infty , +\infty \rangle$$

-5

$\boxed{\exp^{-5}}$

range $(0 \quad +\infty)$

-2	-1	0	1	2
0.13	0.36	1	2.71	7.38

Step :-> 1

exp

random = 7000

$(-\infty \text{ to } \infty)$

$x = 7000$

$(0 \text{ to } \infty)$

$(0 \longleftrightarrow 1)$

Step 2 ->

$$\frac{x}{x+1} = \frac{7000}{7000+1}$$

$(0, 1)$

$(0, 1)$

0

1

$$\boxed{y = mx + b} \longrightarrow \text{Linear Regression}$$

$$y = (-\infty, \infty) \Longrightarrow (0, 1) \Rightarrow \text{Logistic Reg}$$

$$\boxed{\frac{e^{(mx+b)}}{e^{(mx+b)} + 1}} = P \quad (0, 1)$$

y

$$\boxed{P = \frac{e^{y}}{e^{y} + 1}}$$

$$\frac{\cancel{e^y}}{\cancel{e^y}} P = \frac{e^y}{e^y+1} = \frac{\cancel{e^y}/\cancel{e^y}}{\frac{e^y+1}{e^y}}$$

Both numerator & Denominator, $\times$ by e^y

$$P = \frac{1}{\frac{\cancel{e^y}}{\cancel{e^y}} + \frac{1}{e^y}} = \frac{1}{1 + \frac{1}{e^y}} = \frac{1}{1+e^{-y}}$$

$$P(1 + e^{-y}) = 1 \implies P = \frac{1}{1+e^{-y}}$$

$$\boxed{P = \frac{1}{1+e^{-(mx+b)}}}$$

Num

Num

$$f \rightarrow (0, 1)$$

$$\boxed{\text{Sigmoid fn}} \rightarrow \text{D.L}$$

$$\boxed{(0, 1)}$$

$$P = e^y / e^y + 1$$

$$P(e^y + 1) = e^y$$

$$Pe^y + P = e^y \implies P = e^y - Pe^y$$

$$P = e^y(1 - P)$$

$$\boxed{e^y = \frac{P}{1-P}}$$

Applying log on both sides

$$\cancel{\log}\left(\cancel{e}^{y}\right) = \log\left(\frac{p}{1-p}\right)$$

$$y = \log\left(\frac{p}{1-p}\right)$$

$$mx + b = \log\left(\frac{p}{1-p}\right)$$

p = prob of success (0.7)

$1-p$ = " " " $1-0.7=0.3$

(0, 1)

$\frac{p}{1-p}$ => Odds

Continu

$\boxed{\log(\text{odds}) = mx + b}$

$\longleftrightarrow$ (0 . . . . 1)

$$\log\left(\frac{p}{1-p}\right) = \boxed{mx+b} \Rightarrow \text{Continua}$$

$$\log\left(\frac{p}{1-p}\right) = \boxed{-0.066 \times 0} + 1.8185$$

$$= 1.8185$$

$m \Rightarrow -0.066$

$x_1 \Rightarrow 0$

$b \Rightarrow 1.8185$

$$\frac{p}{1-p} = \exp(1.8185) = 6.16$$

$$\frac{p}{1-p} = 6.16$$

$$p = 6.16 - 6.16p$$

$$p + p\,6.16 = 6.16$$

$$p(1 + 6.16) = 6.16$$

$$p = \frac{6.16}{7.16} = 0.86$$