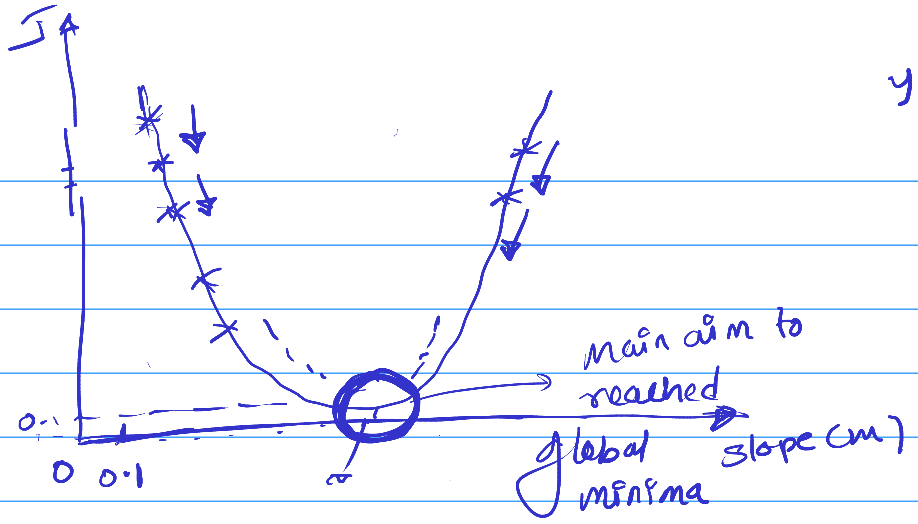
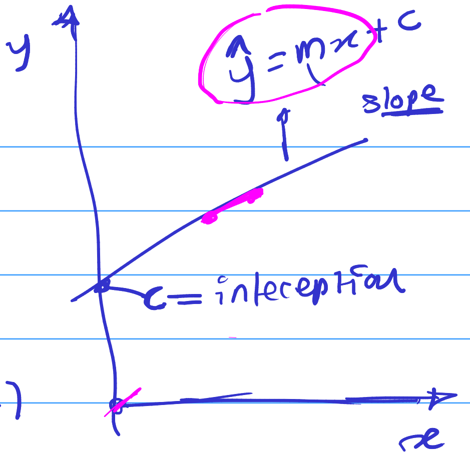
ML
Day 2





	x	y_i	$\hat{y}$
m →	1	1 →	
	2	2	
	3	3	

Performance Matrix ⟹ R^2 or Adj R^2

R^2 ↓ single i/p

Adj R^2 ↓ multiple i/p.

x weight kg →

Cost Function ⟹

MSE → ~~X~~ $\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{n}$ → unit^2 kg^2 → outliled

MAE → $\sum \frac{|y_i-\hat{y}|}{n}$ → it take time

⟹ RMSE → "Neg. Mean. Sq. error"

→ $\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{n}}$ → ~~kg~~^{2} → o dum → time

✗ Assumptions of Linear Regassion

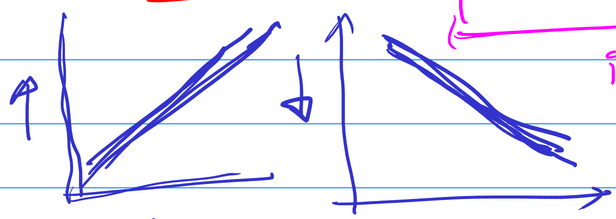
① Linearity → Linear relationship betn i/p & o/p

↓ i/p ↑ = o/p ↑

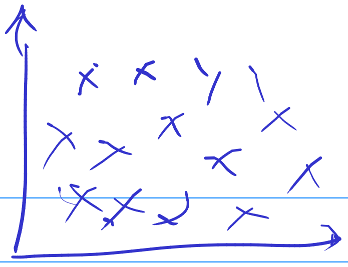
↓ i/p ↑ = o/p ↓

0.1 = 0.1

① Scatter plot



Positive linear → negative →



— No relationship

No - Linearity.

2 — rainbow plot

3 — reg. plot — sns.

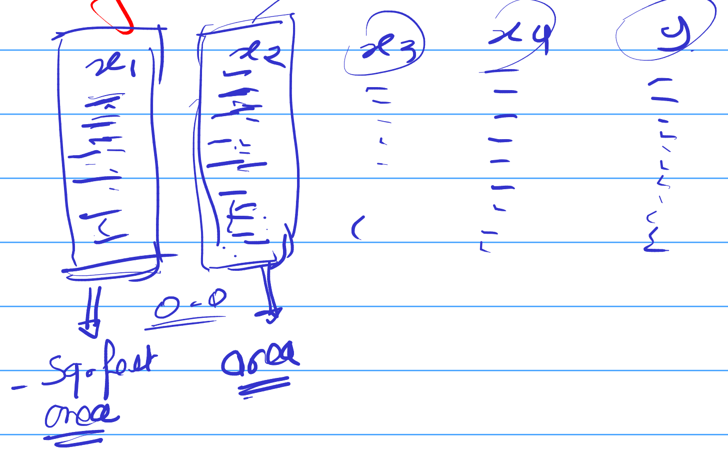
2 Multicolinearity → Relation betn all i/p

x_1 x_2 x_3 x_4 ...

1 Correlation Coeff (x_1, f_x)

2 Heatmap -

3 VIF - varient inflation factor.



3 Multicovenient normality —

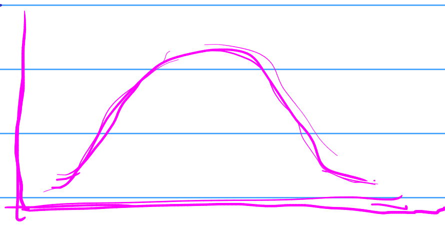
(Normality by Ressidual)

Residual = y test - y predict

1 kde plot =

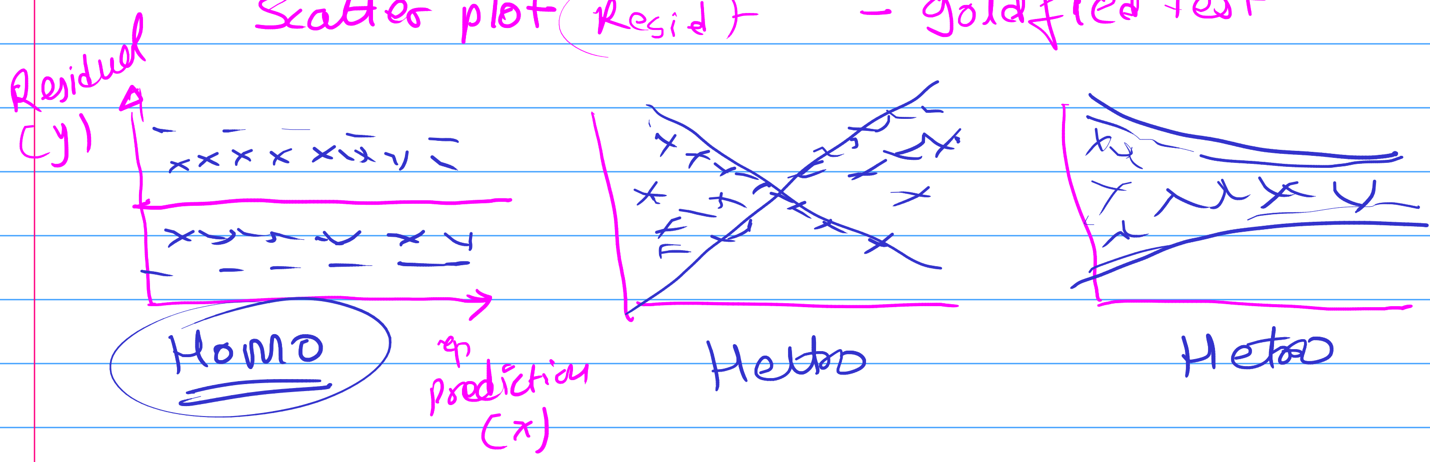
2 Q-Q plot

3 Skewness



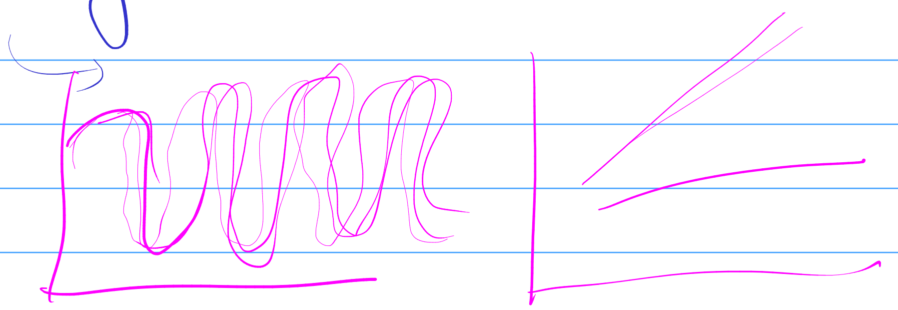
④ Homoscedasticity (having the same Scatter)

Scatter plot (Resid) - goldfied test



⑤ Autocorrelation of error (Residual)

Plt.plot(residual)



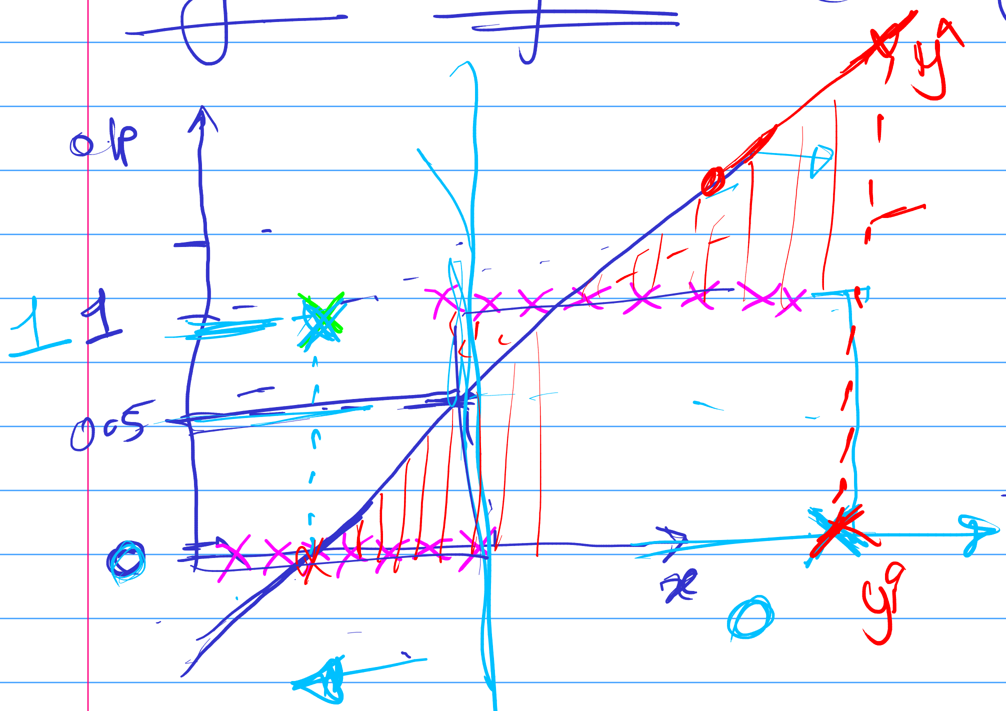
Regression (Linear Reg) ⇒ Prediction use Case

⇓

O/P Continious

multiple i/p - multi ~~ple~~ Linear Regression.

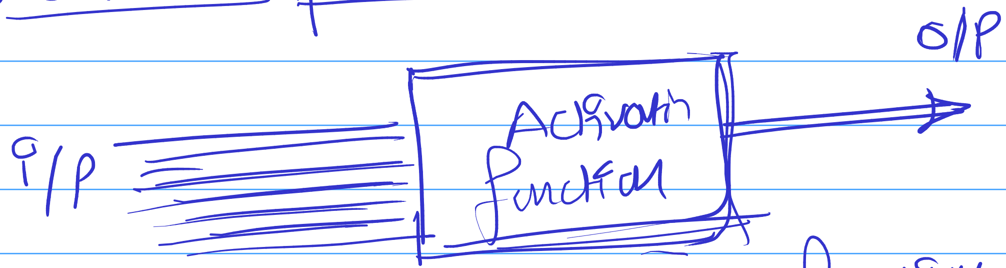
② Logistic Regression (Classification prob)



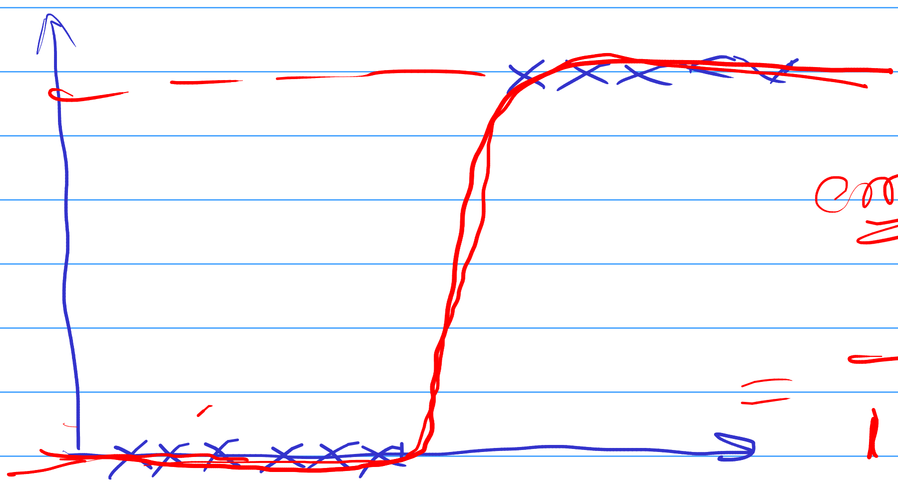
x	y / yes / No
—	0
—	1
—	0
—	1
—	0

binary form

Activation function



Sigmoid activation function



error

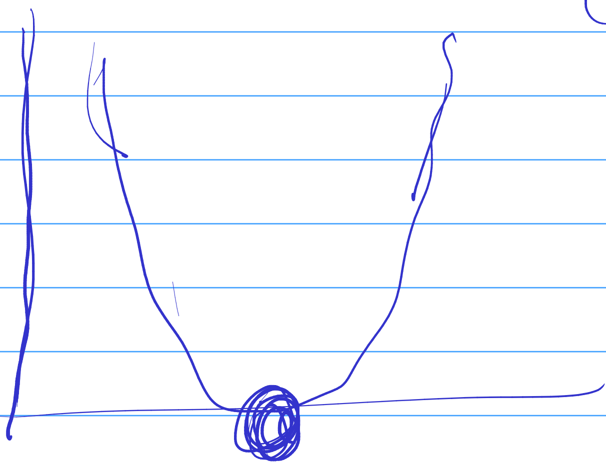
$$= \frac{1}{1+e^{-\hat{y}}}$$

$$\hat{y} = mx + c$$

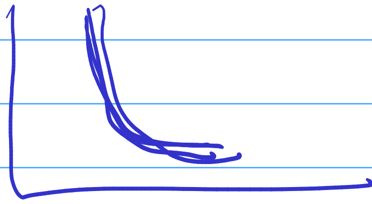
0 to 1

Cost function = $\begin{cases} -\log(J_{new}) & \text{if } y=1 \\ -\log(1-J) & \text{if } y=0 \end{cases}$

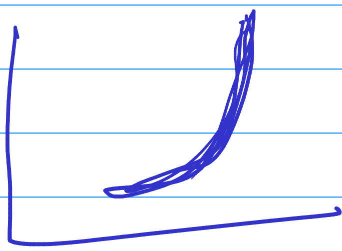
Convex function

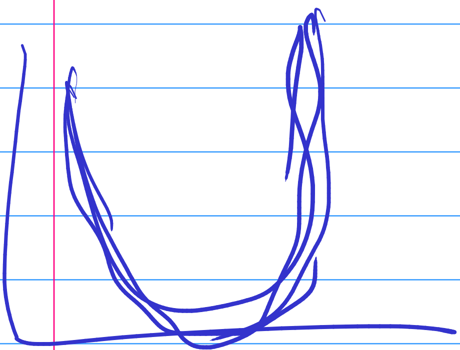


$-\log(J) = y=1$

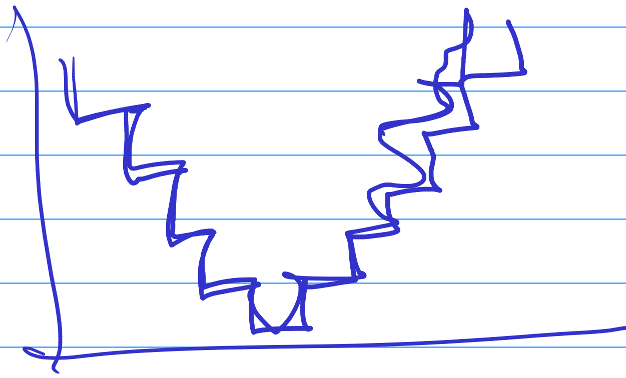


$-\log(1-J) \quad y=0$



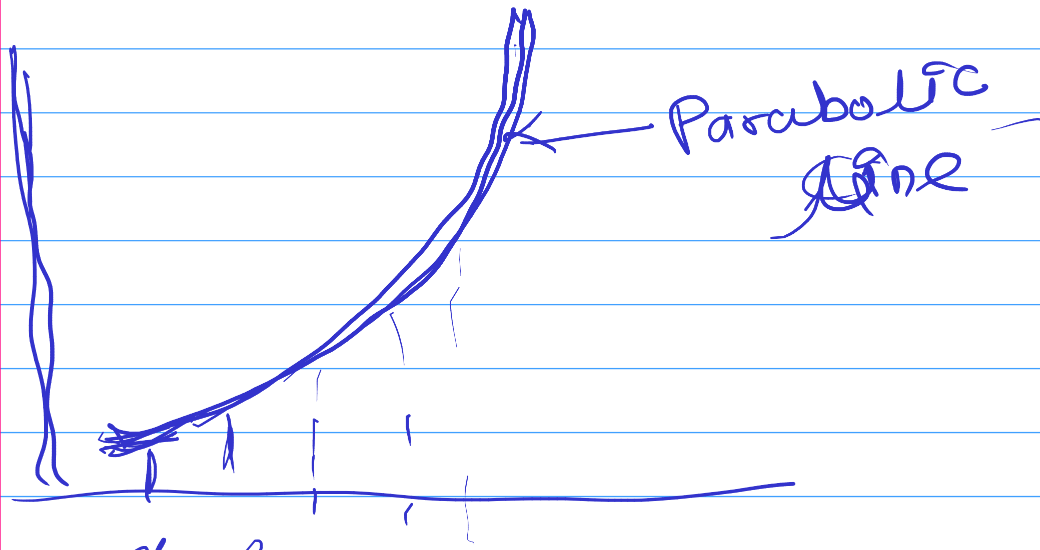


Convex



non-convex

③ Polynomial Regression

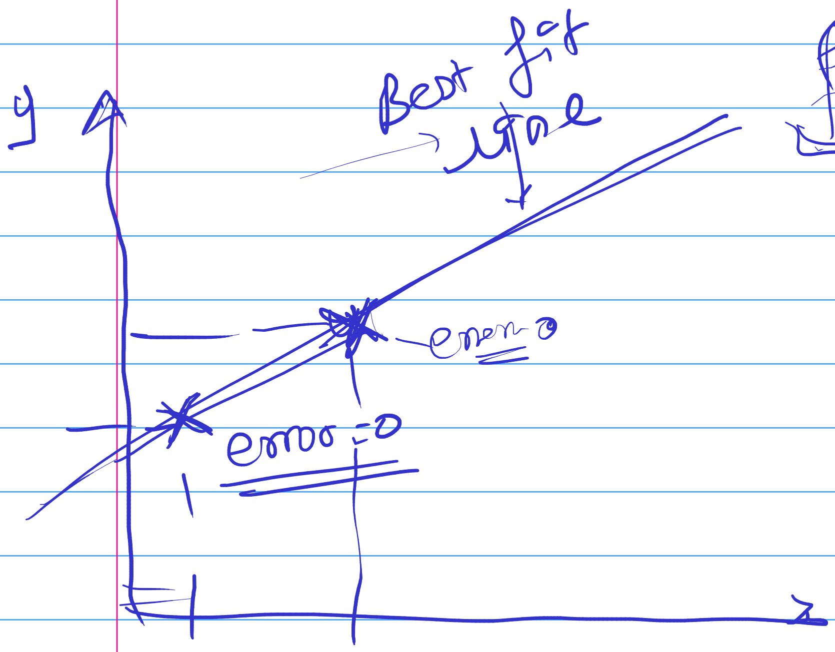


x, x^2, x^3, x^4 → polynomial degrees

④ Ridge (L2) & Lasso (L1) (Regularization)

⇓

Feature Selection



$(y_i - \hat{y})^2$

residual error

L2-regulation = Ridge

L_1 - regularization = Lasso

error = 0

overfitting

Residual = $(y_i - \hat{y})^2$ + error

0

error add

$y_i = \hat{y}$

1 = 1

① all data (

② test data same as train data

$\lambda (\text{slope})^2$ → if slope 0

small value

$L_2 = (y_i - \hat{y})^2 + \lambda (\text{slope})^2$ → overfitting prevent

$= 0 + 4 = 4$ → error

4 16

— error ↑

L_1 reg (Lasso)

lambda

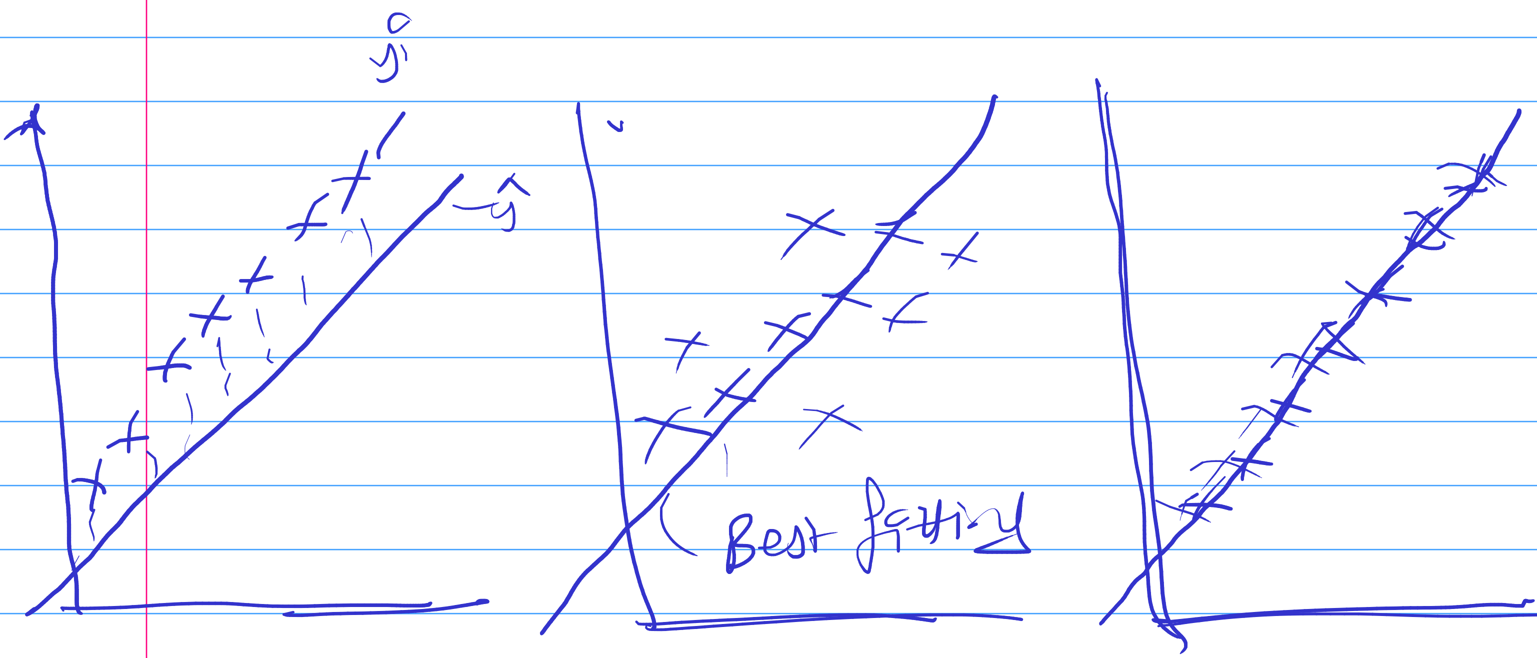
$= (y_i - \hat{y})^2 + \lambda \|\text{slope}\|$

⇒ Feature selection

Overfitting & Feature Selection

ElasticNet (Regulation)

$$= (y_i - \hat{y})^2 + \lambda (\text{slope})^2 + \lambda |\text{slope}|$$



Underfitting | Good fit / Best fit | Overfitting

Bias & Variance

Bias = error of training data

Variance = error of testing data

Underfitting	Best fit	Overfitting
Train Acc. ↓ error ↑	Train & Test Acc. ↑	Train Acc ↑
Test Acc ↓ error ↑	error ↓	Test Acc ↓
High bias &	Low Bias & ↓	Low Bias
High Variance	low Variance ↓	High Variance

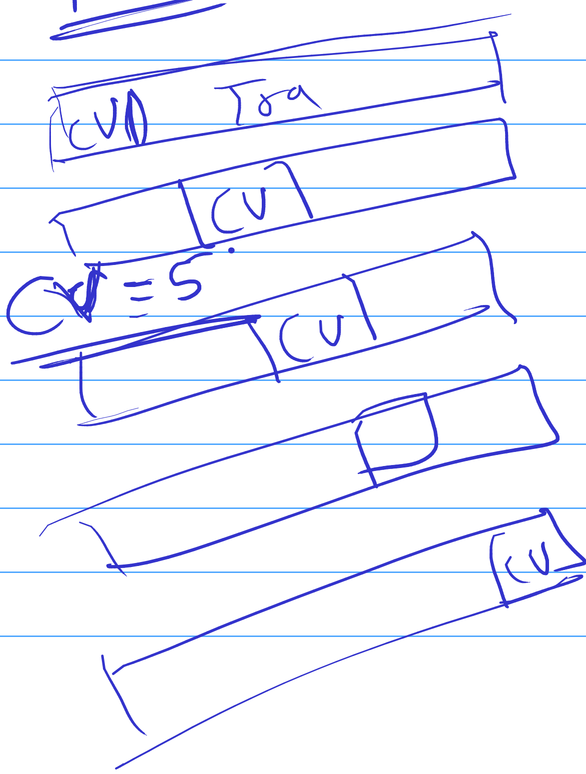
Data

→ train → 1000

→ test data

train → Train data

train → Validation data



$(y_i - \hat{y})$ — Polynomial Reg

$(y_i - \hat{y}) + L2$

$(y_i - \hat{y}) + L1$

$(y_i - \hat{y}) + L_1 + L_2$ — elasticnet

reg =

binary (Regression)

0 — 1

i/p → 54%

error

class

Classification

Classes

0 → 1 → 2 → 0 1 2

Sq. feet area → Room size → Bed Room size

1000 sqft — 1000 sqft

2000 sq feet — 2000 sq feet

x_1 — x_2 — x_n

yes / No

True / false

0 1 0 1 0 1

log → i/p → log reg

diff — diff

~~All i/p features are independent~~

A → / B →

B → / A ↔

o/p

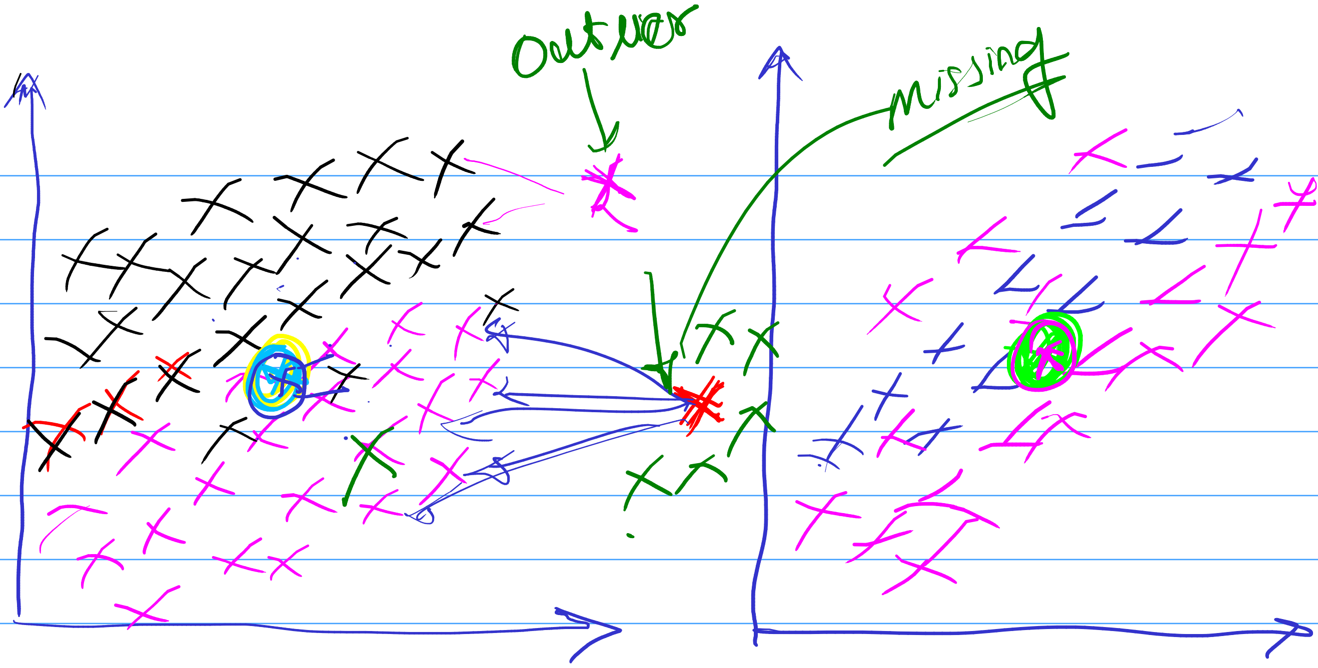
Bays: ⇒ Naive Bais classification

$$P(A/B) = \frac{P(A) \cdot P(B/A)}{P(B)}$$

1. Gaussian NB ⇒ o/p is in numerical form.

2. Bernolli Naive Bias ⇒ o/p is in True/False
   yes/no
   → Binary

3. multinomial Naive Bias ⇒ o/p is is classes



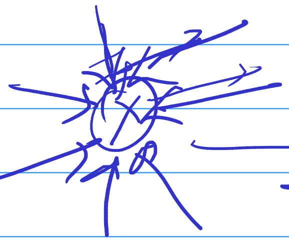
KNN — K-nearest Neighbor

(Reg — Class)

K-value — Hyperparameter K=2

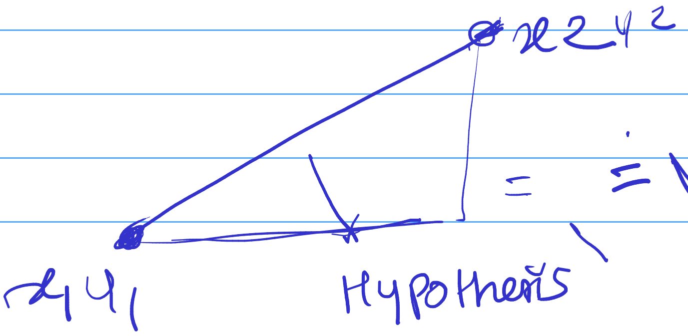
K=3 K=4 ⇒ 3=0, 1=1 } 0

Nearest ?



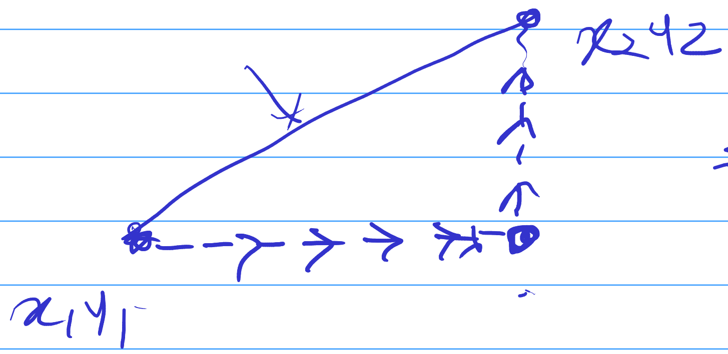
— distance

① Eucledian Distance



$$= \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$$

② Manhatten Distance



$|x_2 - x_1| \; |y_2 - y_1|$

Limitation :

- Do not work with Huge data because of distance.
- Sensetive to outlier
- sensetive to missing value.