

## ① LR with single Variable

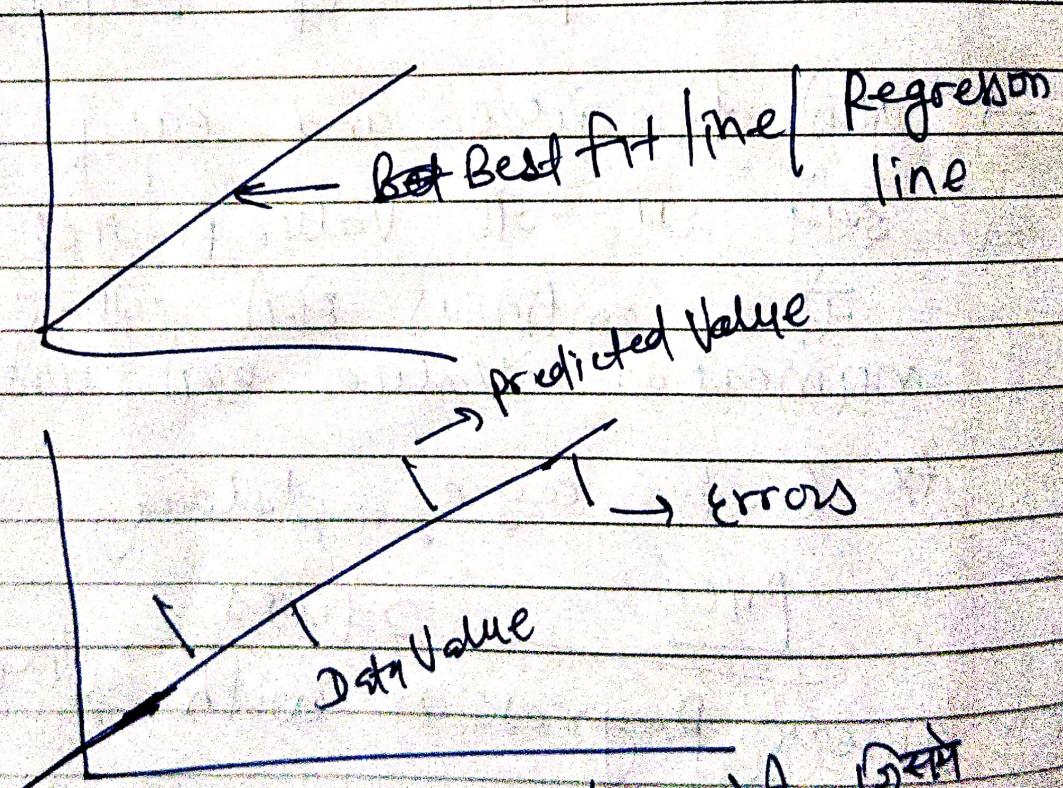
Use

case  $f(x) = \text{पानी की आवश्यकता} = \text{water requirement}$

② Home की price  $\rightarrow$   $\text{Area}$   
की Base  $\rightarrow$   $\text{PT}$

③ Student  $\rightarrow$  Hr '46 हो दी  
result कैसा था?

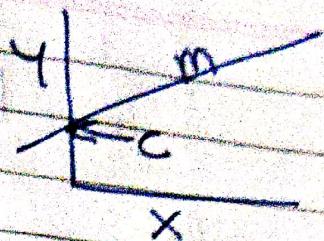
इन सब numerical Value



Best fit line से line की दिसने  
errors कम से कम हो।

$$y = mx + c \rightarrow \text{Intercept}$$

premium      Age



$y$  = dependent Variable

$x$  = independent Variable

$m$  = slope / gradient / coefficient

$c$  = intercept

$$\text{premium} = m \times \text{age} + c$$

Linear regression on single variable (P41b)

(ii) Linear Regression with Multiple  
- Variable

Continuous or Numerical  
Value

Cost, price, sell, firm  
Marks, amount,

23



## LR with Multiple Variable

Ex Student's score 45% एवं 37%

or 2 hours 45% & PC 19 level

2 hours का तो 54% एवं 37%

Multipal Variable एवं गणित

formula

$$y = m_1 * n_1 + m_2 * n_2 + m_3 * n_3$$

$$\text{premium} = m_1 * \text{age} + m_2 * \text{height} + m_3 * \text{weight}$$

$y = \text{premium}$  (dependent)

$\text{age}$  independent

$+ 1$  weight width fm

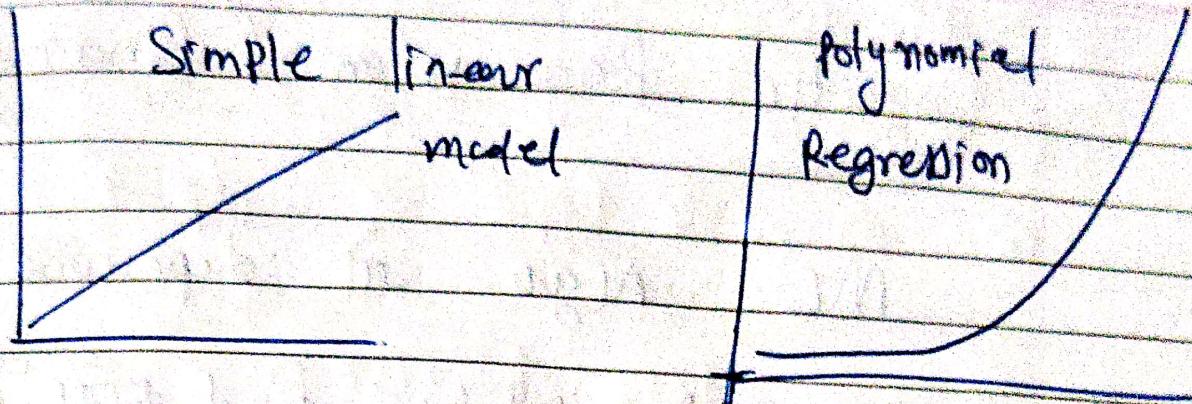
$+ 2$  weight

$+ 3$  slope

$m_1, m_2, m_3$  intercept

$c =$

# Poly nomial Regression



Polynomial  $\rightarrow$

o Degree Polynomials

$$y = \text{Constant}$$

1 Degree polynomials

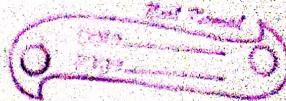
$$y = mx + c$$

Generalized  
Polynomial Equation

$$y = q_0 + q_1x + q_2x^2 + \dots + q_nx^n$$

2 Degree polynomial

$$y = ax^2 + bx + c$$



# Logistic Regression (Binary classification)

ML Algo  $\rightarrow$  supervised learning

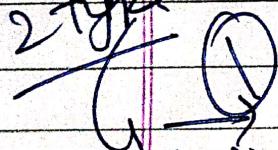
$\rightarrow$  Probability find out  $\hat{P}(Y)$

Classification problem of probability  
predict  $\hat{P}(Y)$

$\rightarrow$  non continuous form of  $\hat{P}(Y)$

(x)- English or Hindi, True or False  
1 or 0, High or Low, Good or

~~2 types~~



Binary classification → win or loss  
dead or alive

II Multiclass Classification

$\rightarrow$  I Onion or Potato or  
Sweet potato

II Gilly or Sunflower or Rose

only one class

Mutipal  
class

96

Ex. ① Heart attack risk or probability

② Tumor prediction



③

④ Credit - Card fraud

⑤ Spam Detection  $\rightarrow$  free gift, offer, WM

class A

Sigmoid function

Cut off line

class B

probability  $\geq 0$ 

$$y = 1$$

$$y = 1 / (1 + e^{-x})$$

Sigmoid function

$y$  = dependent variable

$x$  = independent variable

e = Euler's constant

2.71828

\* Multiclass Classification  $\Rightarrow$

Target Variable of prediction  
का लिए

Example  $\Rightarrow$  True or False

True  
False

सत्य व वैज्ञानिक

English or Hindi

अंग्रेजी के भाषा

Wrong or Right

चance

Cat or dog or goat

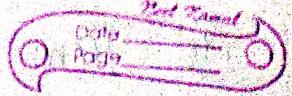
Male III Female

Continuous form

अंग्रेजी के भाषा



22) Classification - Target Variable को Yes या No देना होता है।



e.g. identify Types of animal

(i) Vehicle को से type को Air Road

(ii) News के types को Sport,

(iii) Iris dataset use

Decision

Tree Classification

classifier model

Supervised Learning

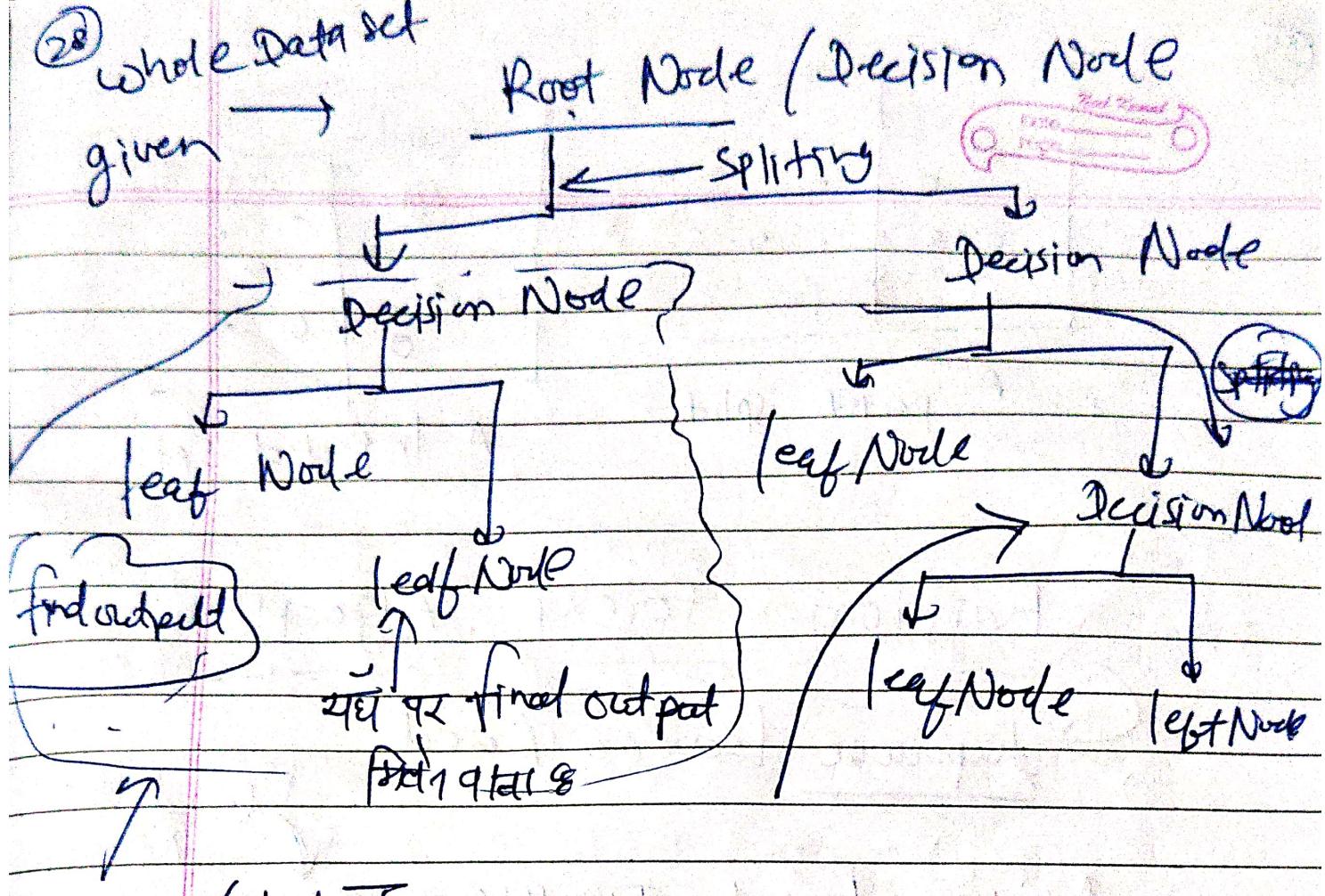
Regression or Classification एप्पों के लिए

Variable Continuous  
use की तरह एक होती है  
Target Variable को एक वर्गीकरण करती है  
Nature की तरह एक होती है  
Income, Age, etc.

(iv) Classifier Model को बनाना होता है

class का Value को predict करता है

Target Variable के लिए



Branch/ Sub Tree

Ex- ① Medical field → patient को Covid है या नहीं  
 patient की fever है, तुम्हारे

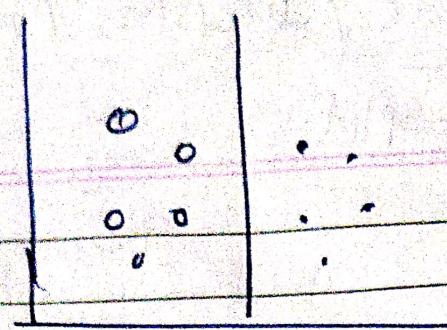
Age (वयस्सी),  
 weight

② Student mark Based पर Decision of hire  
 60% से कम fail तभी तक 60% pass

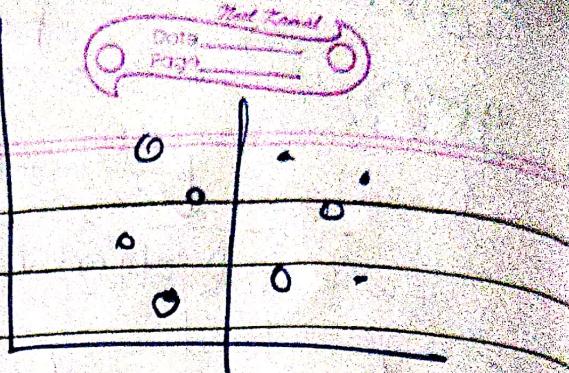
→ Teacher promotion

→ Banking → Loan Approve (Y/N)

29

 Date \_\_\_\_\_  
 Page \_\_\_\_\_


A perfect split



A Imperfect split

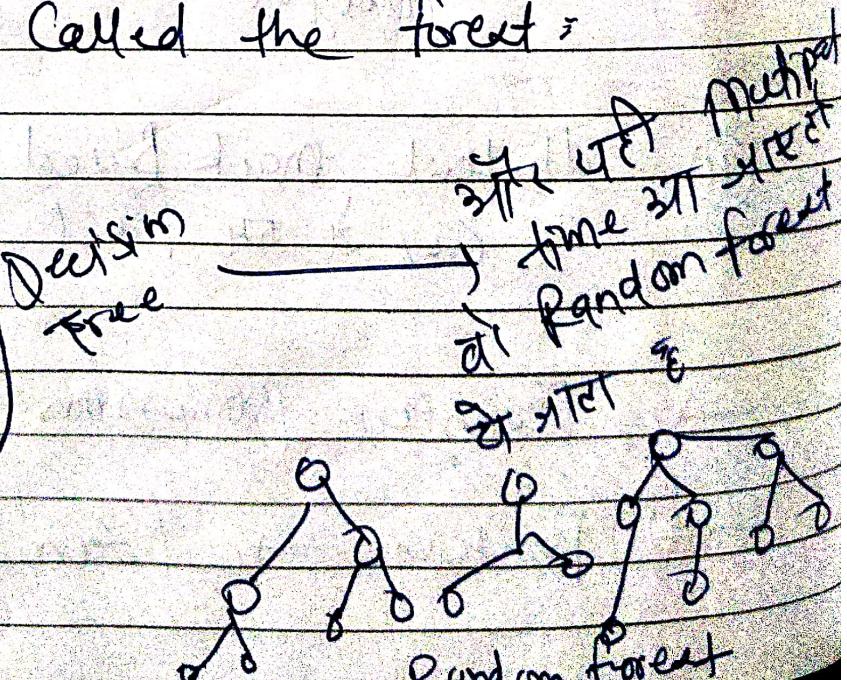
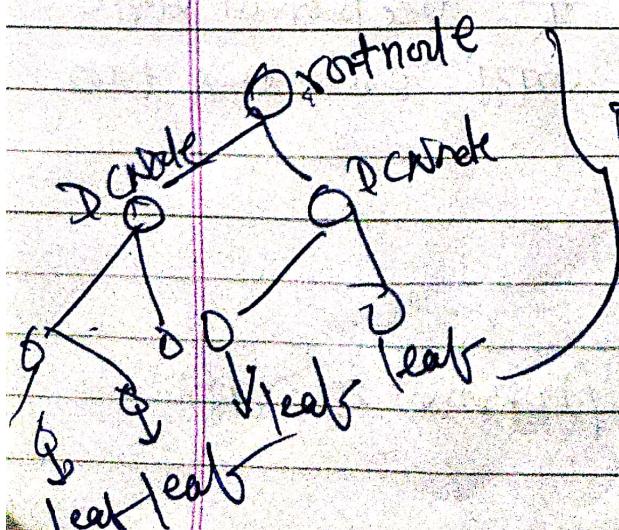
## Random Forest Algorithm

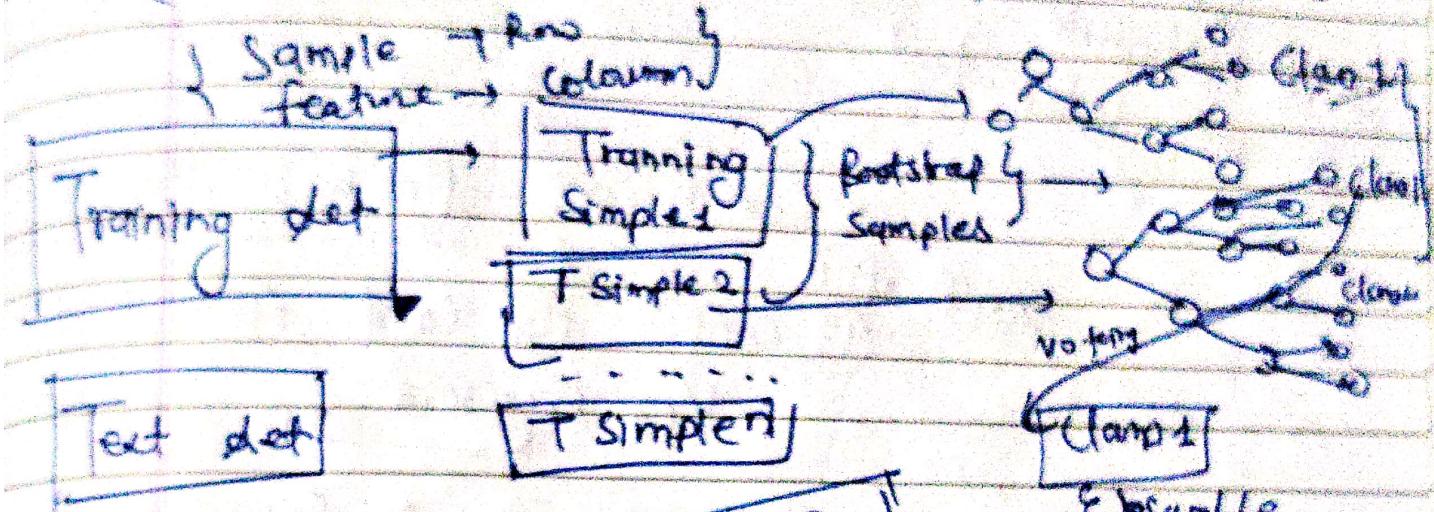
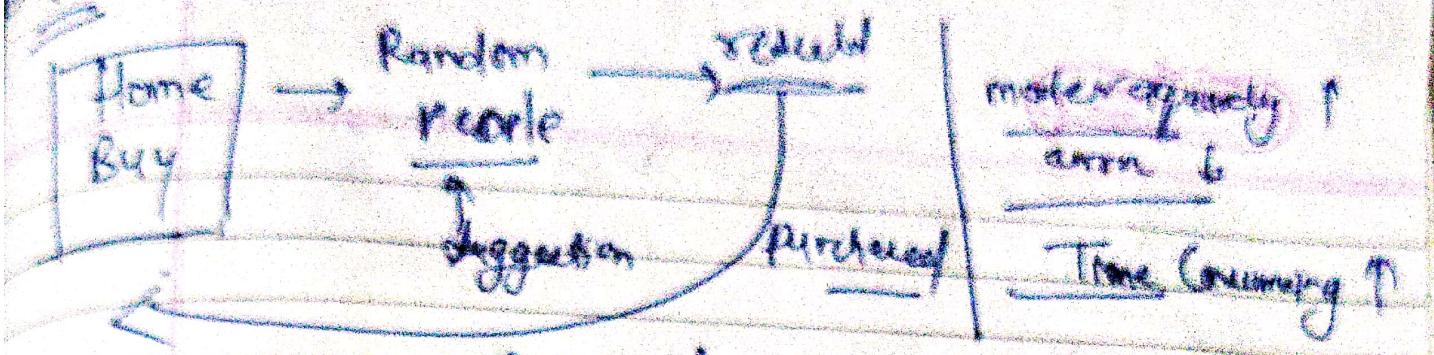
Supervised learning Based

→ Random forest का उपयोग regression or classification के लिए किया जाता है।

Target Variable  
Yes or No  
Male or Female  
Etc.

⇒ It is a collection of multiple random decision trees. Which is called the forest.





Naive Bayes Classifier  
 Algorithm

ML Algo

→ supervised learning  
 → classification problem lots of probability

(e.g. Corn in唐玉  
 not eat) → Tail of Tail  
 50% Corn damage + half

② Corn chance 50%

→ 2 digit & 6 digit  
 → object at probability HC prediction

31

→ Bayes theorem का इसका नाम

## Naïve Bayes :-

जो भी attribute  
हैं उन्होंने independent  
भारत में

$$\begin{array}{c} * \text{Condition} \quad \text{Probability} = \\ \text{प्रश्न की प्राप्ति में } \\ (1) \text{ card } \quad (2) \text{ Card } \quad (3) \text{ Card } \quad (4) \\ \sqrt{1/13 \times 1/13} = 0.076 \text{ Card} \end{array}$$

$P(A|B)$  = probability of A occurring  
given that B has already  
occurred.

Bayes theorem

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \rightarrow A, B \rightarrow \text{Events}$$

$P(A|B)$  → probability  
A is given  
that B is the

$$\begin{array}{c} \sqrt{1/13} * \sqrt{1/13} = \sqrt{1/13} \\ \sqrt{1/13} \end{array} = 0.076$$

→ Face recognition software

- Used →
- ① Face recognition
  - ② Weather forecast
  - ③ News Categorization
  - ④ Article
  - ⑤ Sport news
  - ⑥ World

# SVM

## Support Vector Machine

SVM → ML Algo

↳ Supervision learning

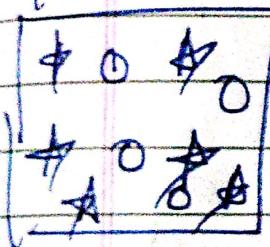
Regression or Classification दोनों  
problem solve करता है

mostly एक Classification

↳ इसमें Classify

label पात्र  
प्रदान किया जाता है

label पात्र



Model  
Training

→ prediction

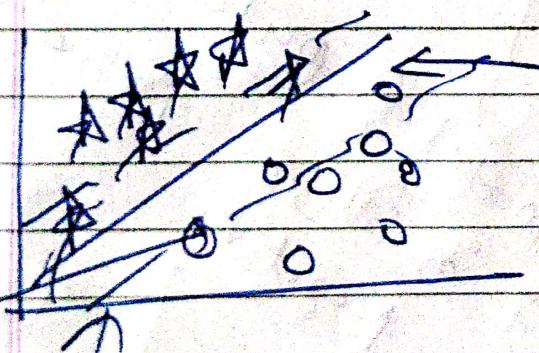
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\rightarrow$  Grade  
1 Star

Marginal Distance

Precision  
Boundary

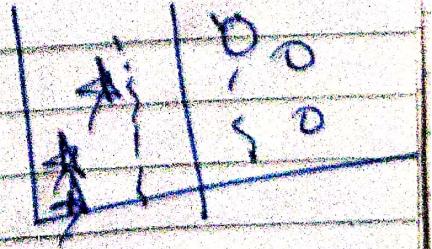
SD



Marginal line

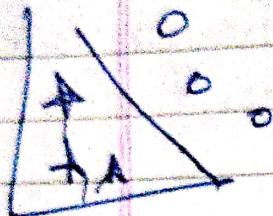
Marginal (between)  
after nearest sample

hyperplane

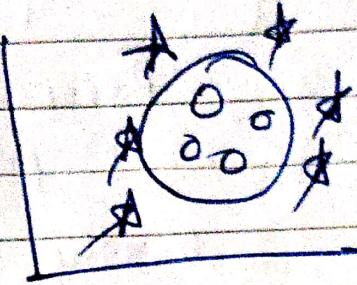


## Q) Types of SVM

Linear SVM

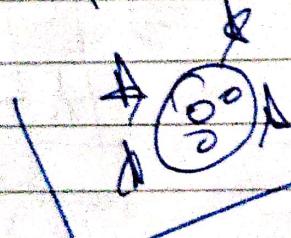


Non Linear SVM



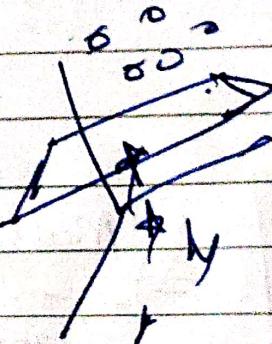
soft state line

Kernel function



Kernel

Matrix



Confusion

Classification model off Performance

evaluate  $\text{TP}$ ,  $\text{TN}$ ,  $\text{FP}$ ,  $\text{FN}$

Negative

Actual values

Targets



(2)

Accuracy  $\rightarrow$  Total no of prediction right %

Error Rate  $\rightarrow$  Total % wrong predict rate %

Precision  $\rightarrow$  Total Right predictions %

Recall  $\rightarrow$  actually Yes how often it predict Yes

Actual: No	200	TN = 100	FP = 15	115
Actual: Yes		FN = 5	TP	85
		105	95	

## Oversampling and Undersampling

Performance check करने के लिए उन-

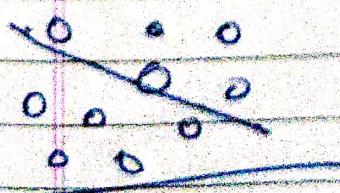
Accuracy  $\rightarrow$  किसी X का predict किया है

Generalization  $\rightarrow$  New test data predict

करने के / New Data के Test  
करने के

## Underfitting

$\Rightarrow$  Training or test हो जाएंगे।  
accuracy सेवे।

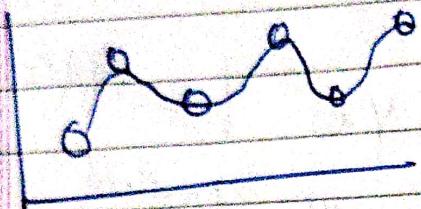


Underfitting  $\rightarrow$  Underfit Jaisa हो जाएगा।

35



## Overfitting =



दो दाता को fit कर  
की Try करती है

Tranning में की High  
Accuracy score करती है

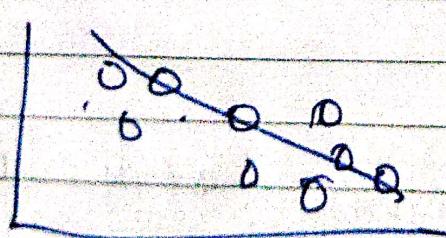
पर Test में score घटता है

## Ex Over size Jense

Machine को इस से ज्यादा प्रधान करते हैं  
हो machine error से ज्यादा लगता है

Right fit जब training or test परीक्षा  
में same data मिल तो यह model

best होता है



नहीं की overfit हो  
हो under fit