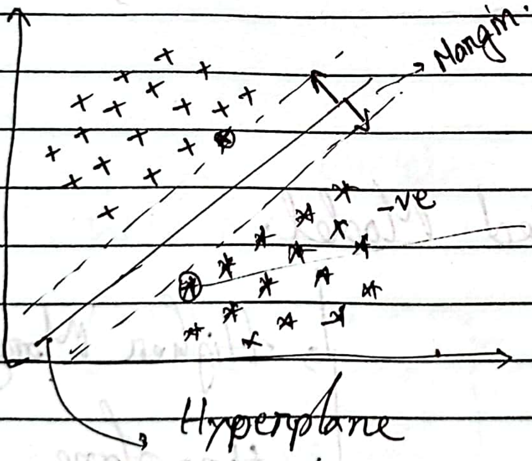
Support Vector Machine.

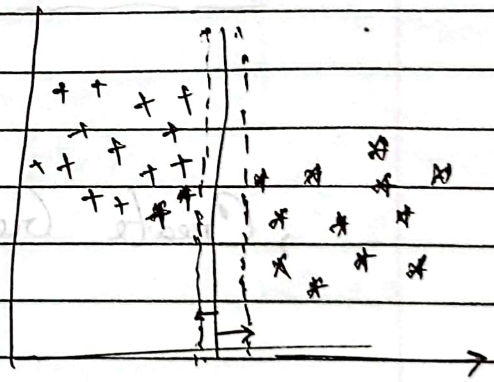
Solve → Classification or Regression.

Term should know.

1. Support vector.
2. Hyperplanes.
3. Marginal Distance
4. linear Seperable Point's.
5. Non-linear Seperable Point's.

let's Consider a Classification Problem:





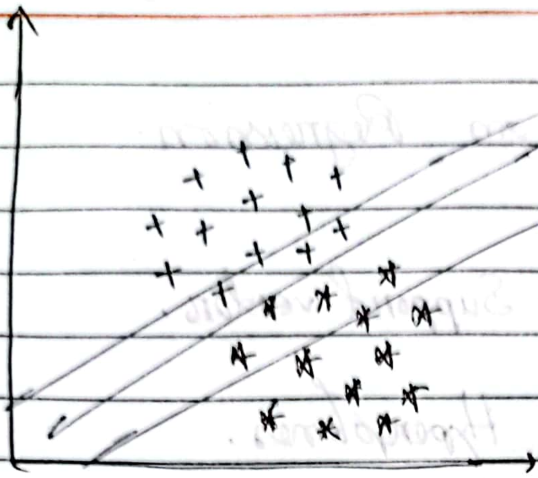
Aim → Maximize the "marginal Distance"

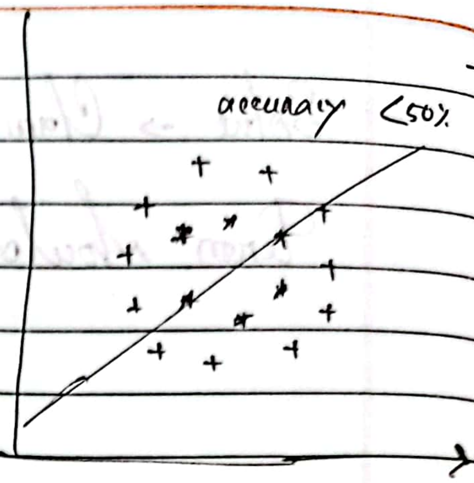
[Selected]

it's may be many way, many way's of hyperplane.

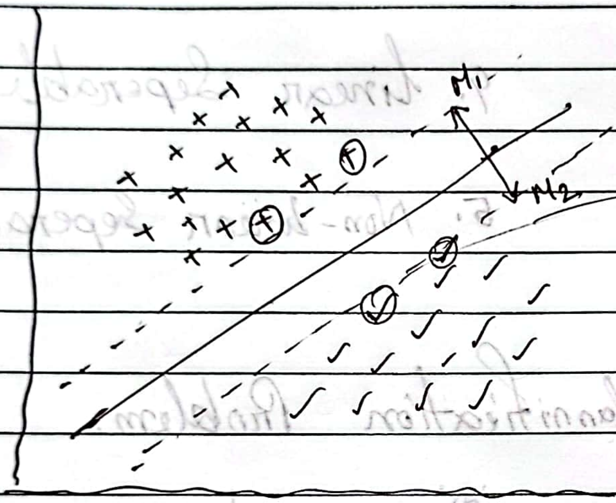
linerly Separable. / draw straight line.

2D





No linear



Nearest Point
↓
Support Vector
↓
Any Num of Points.

→ Create Generalized Model.

↳ Higher Marginal Distance

↳ Hyperplane

Support Vector Machine Vs Logistic Regression.

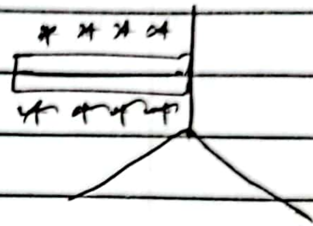
Marginal Distance

Kernel.

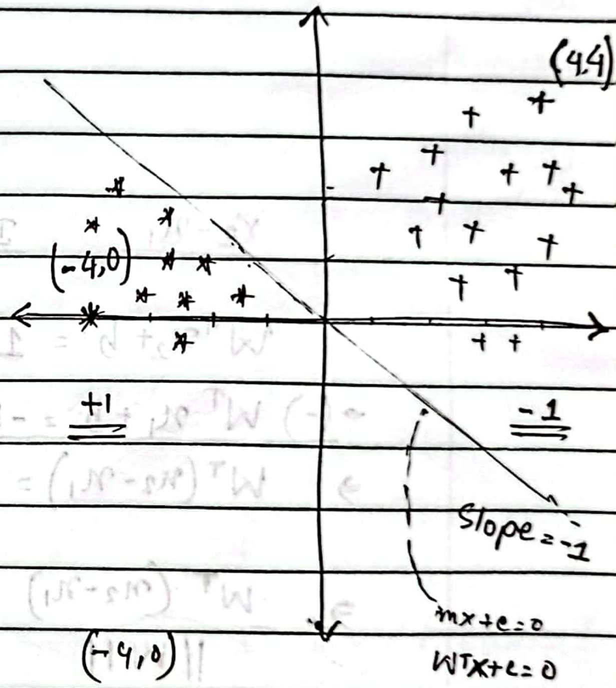
SVM → Converts 2 Dimension (low dimension) into a

↳ High Dimension.

if 3



Logistic Regression



(4, 4)

(-4,0)

$y = W^T x + c$

$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} 4 & 4 \end{bmatrix}$

$= -4$ -ve value

Going to be always -ve.

$y = mx + c$

$= W^T x + c$

$m = -1$

; $c = 0$

$y = W^T x + c$

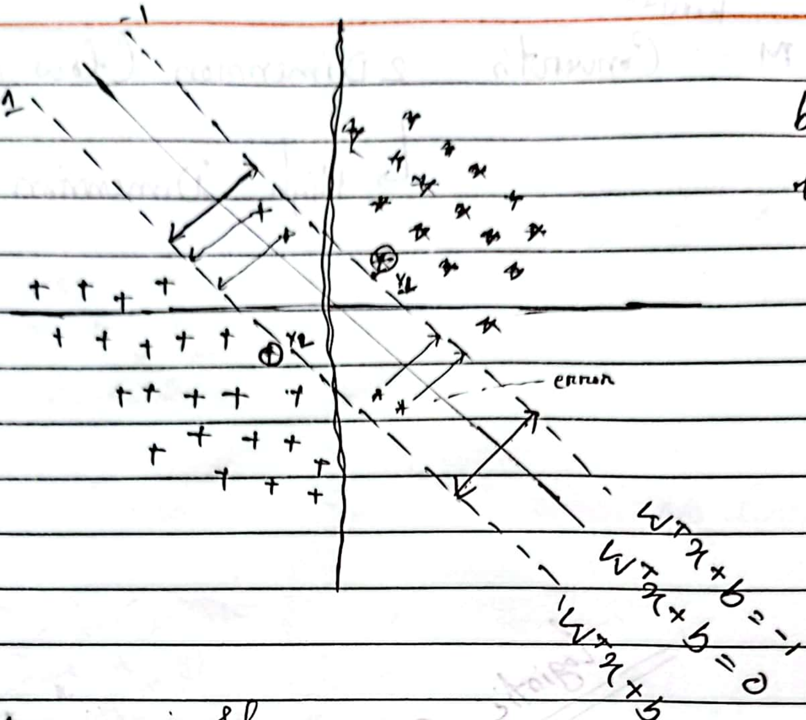
$= W^T x + 0$

$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} -4 & 0 \end{bmatrix}$

$= 4$ +ve value

→ Going to be always +ve.

based on logi-Reg Scenario.



$x_2 - x_1$ Difference Distance

$$W^T x_2 + b = 1$$

$$(-)\ W^T x_1 + b = -1$$

$$\Rightarrow W^T (x_2 - x_1) = 2$$

$$\Rightarrow \frac{W^T (x_2 - x_1)}{||W||} = \frac{2}{||W||}$$

Optimization function.
We need to maximize this !!

$$y_i \begin{cases} 1 & \text{if } W^T x + b \geq 1 \\ -1 & \text{if } W^T x + b \leq -1 \end{cases}$$

$$\underset{\text{update}}{(w^*, b^*)} \ \max\left(2/||w||\right)$$

$$y_i * W^T x_i + b_i \geq 1$$

+ve → +ve
-ve → +ve

Real world Scenario data is not distributed on a Proper way.

$$(W^*, b^*) \to \min \left\{ \frac{||W||}{2} + C_i \sum_{i=1}^{n} \xi_i \right.$$

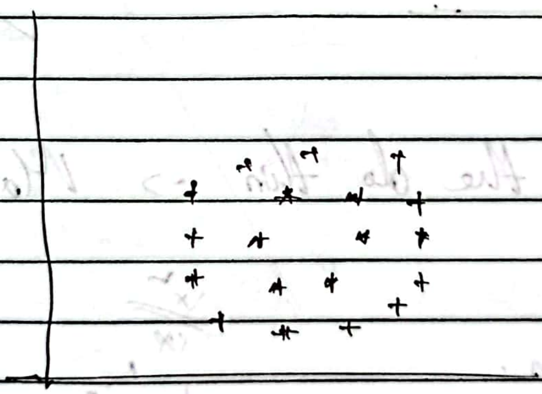
Regularization

how many errors my model Consider

value of of the error.

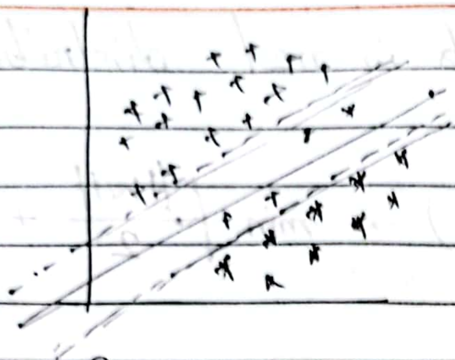
(4 * 4 Point distance sum) this scenario.

SVM



SVM Kernel trick.

SVM Kernels.



① Polynomial Kernel.

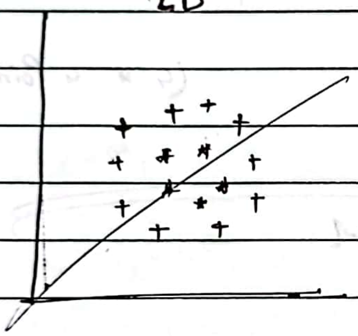
② RBF Kernel.

③ Sigmoid Kernel.

Soft margin -> small % of error

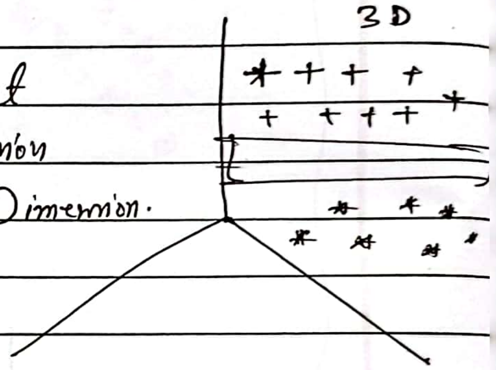
~~Hard margin -> Properly classify~~

2D



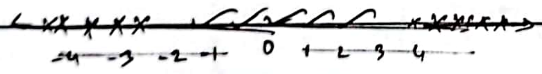
SVM Convert lower Dimension to higher Dimension.

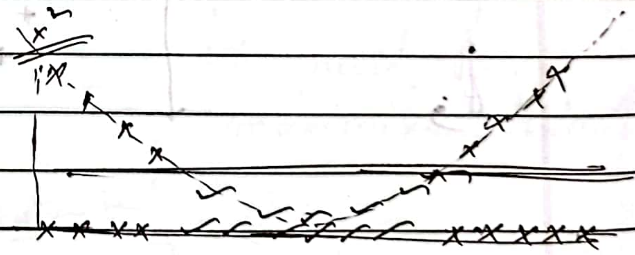
3D



How the do the do this --> Mathmatical formula.

1 Dimension.





2 Dimension.

$y = f(x) \cdot x^2$

that mean SVC = linear.

1st apply SVM NORMALLY → 50% <

Apply SVM Kernels.

2 Dimension

x_1	x_2	y

$y = f(x_1, x_2)$

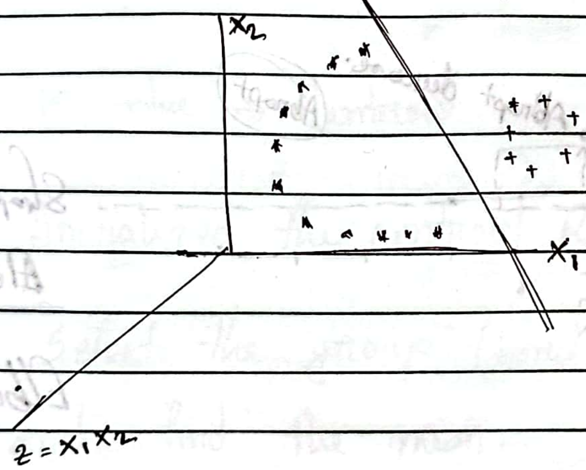
Polynomial Kernels → $(x_1^T \cdot x_2 + 1)^{\text{dimension}}$

$$= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{x_1^2} & \underline{x_1 x_2} \\ x_1 x_2 & \underline{x_2^2} \end{bmatrix}$$ unique

x_1	x_2	y_0	x_1^2	$x_1 x_2$	x_2^2

← 3 Dimension.



SVM Kernel = "Poly" or "rbf"

Consider like this.

How do we decide that weather we should use Polynomial or RBF or Sigmoid.

Determine by Hyperparameter tuning.

4

K= 2 means.

2 Centroid /

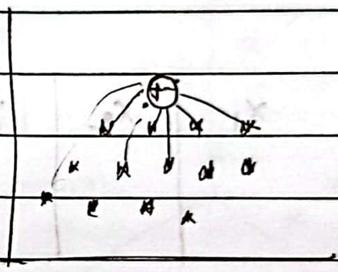
2 cluster /

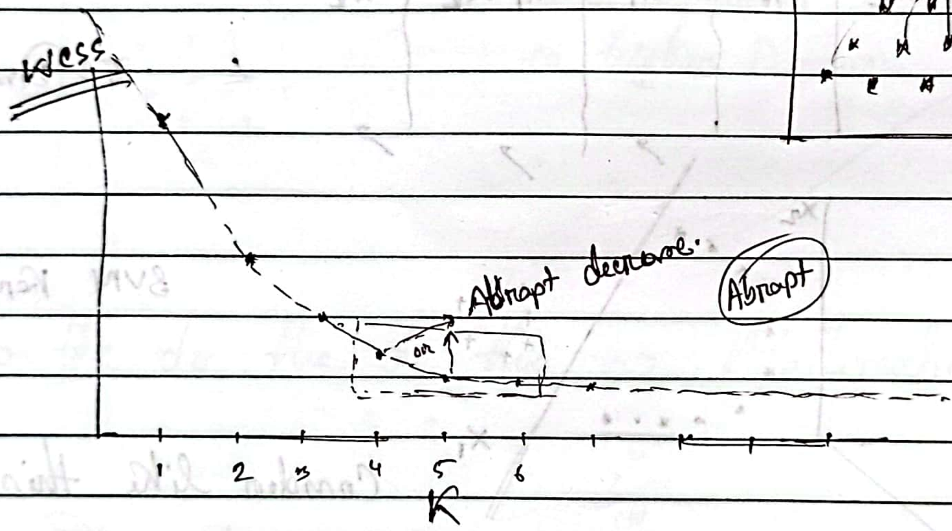
2 Group.

"Krish Naik."

WCSS = Within cluster Sum of Square.

$$= \sum_{i=1}^{n} (c_i + x_i)^2$$





Shape of Elbow

Elbow method.

K=4

K=5

[Analyticsvidhya detail explanation.]

K= highest possibility
= numb. of observation.