LECTURE NOTES

Campus: PCE Course: BTECH Class/Section: CS-III - C Section Date: 22-02-22

Name of Faculty: Praveen Kumar Yadav Name of Subject: Machine Learning Code: 6CS4-02

Date (Prep.): 22-02-22 Date (Del.): 16-03-22 Unit No./Topic: 4 Lect. No: 05

OBJECTIVE: To be written before taking the lecture (Pl. write in bullet points the main topics/concepts that will be taught in this lecture)

Support Vector Machine

IMPORTANT & RELEVANT QUESTIONS:

① what is Support Vector Machine? Explain in detail.

FEED BACK QUESTIONS (AFTER 20 MINUTES):

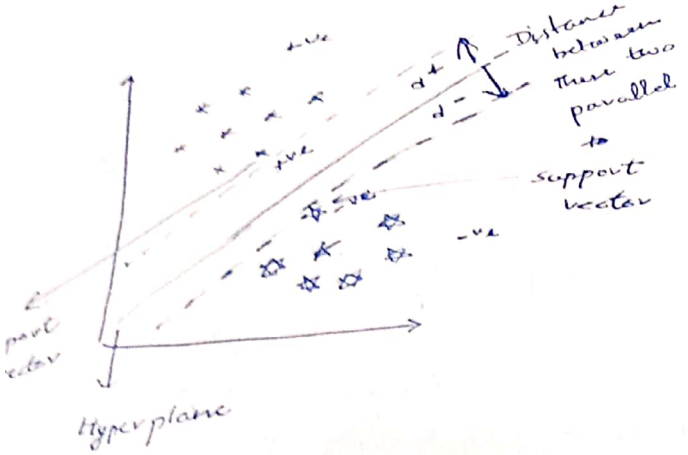
① what is Support Vector.

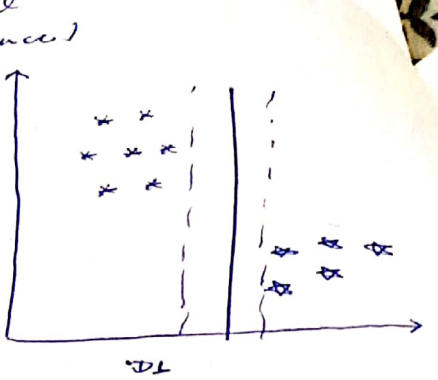
② what is Margin in SVM?

OUTCOME OF THE DELIVERED LECTURE: To be written after taking the lecture (Pl. write in bullet points about students' feedback on this lecture, level of understanding of this lecture by students etc.)

REFERENCES: Text/Ref. Book with Page No. and relevant Internet Websites:

SVM, Simpli Learn.





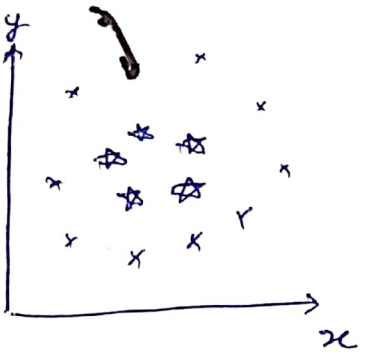
Support Vectors

Hyperplane

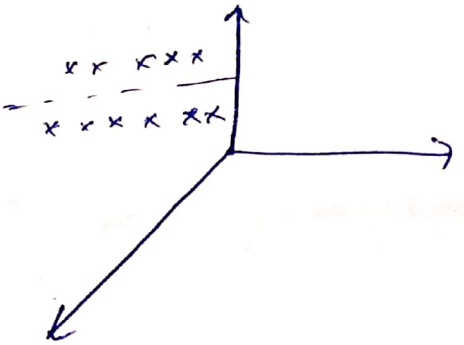
Marginal Distance

Linear Seperable

Non Linear Seperable



2D

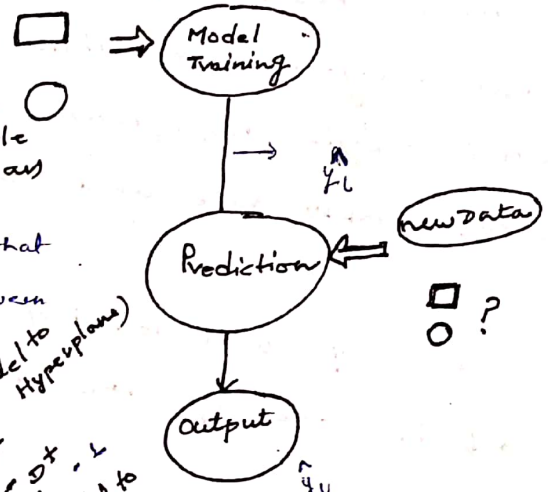


Kernal - Try to convert low Dimension into high Dimensional space so that we easily can classify.

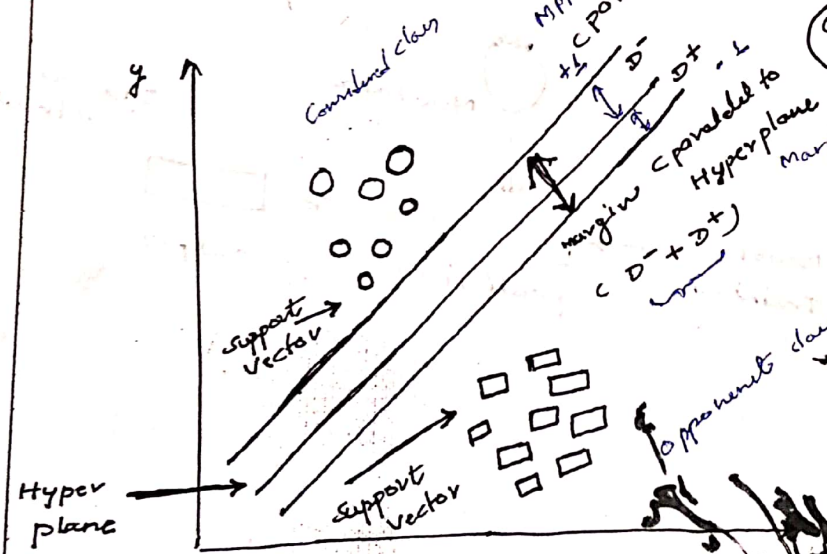
Support Vector Machine:-

- Its a supervised Learning Algo^m.
- SVM used for classification.
- objective of SVM Algo^m is to find a hyper plane in N-dimensional space that distinctly classified the data points.
- Our objective to find a plane that has maximum margin - distance between the data points of the classes.

$D = \{x_i, y_i\}$



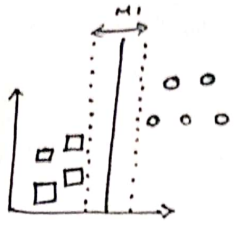
circle class



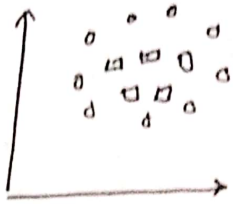
- Maximum margin distance provides some confidence so that future data points can be classified with more confidence.
- nearest data point which passes to MP-1 and MP-2 are called as support vector.

opponents class

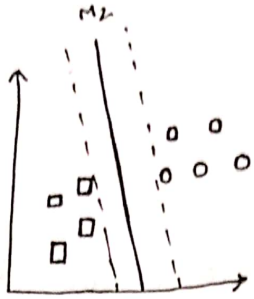
since it decide the Test data is belongs to which class either circle or square.



Linearly seperable plane

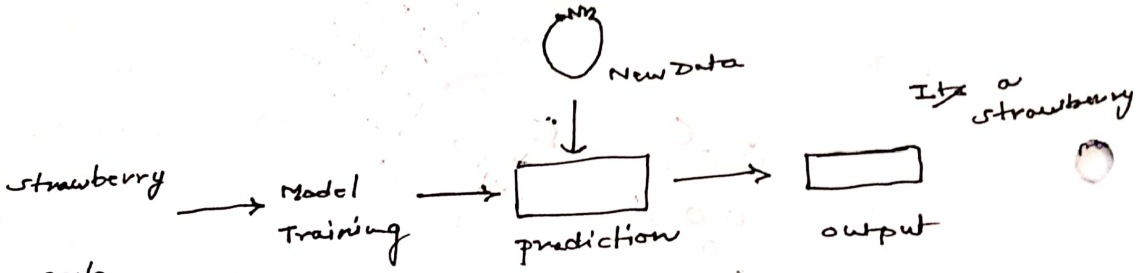


Non-Linearly seperable plane.



→ we Always choose those hyper plane which is having largest margin. (max - width margin).
(Maximal margin Hyperplane should be selected which decreases ~~error~~ error rate and ~~decr~~ increases accuracy.

SVM is a supervised learning method that looks at data and sort it one of the two category.



PAGE NO. .....

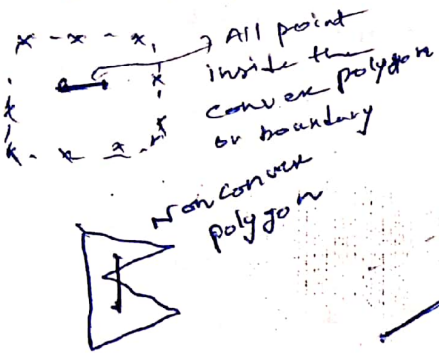
Application of Support Vector Machine :-

- Face detection
- classification of images.
- Bioinformatics.

AND			OR			EXOR		
	0	1		0	1		0	1
0	0	0	0	0	1	0	0	1
1	0	1	1	1	1	1	1	0

Geometric Intuition - II -

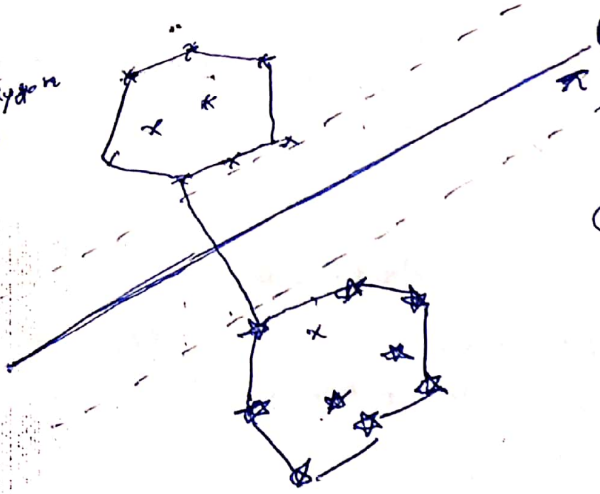
Convex polygon



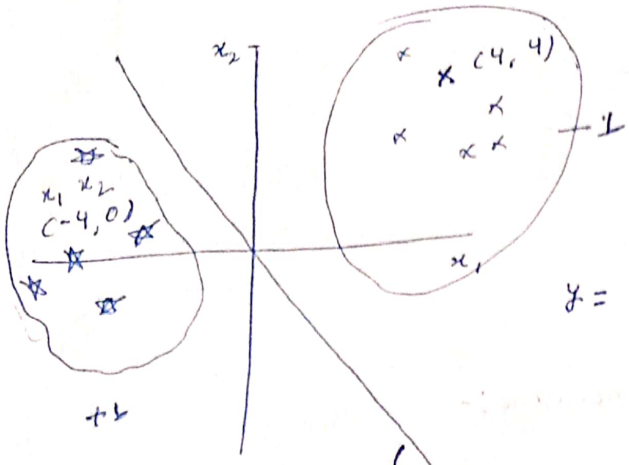
1. construct a convex hull for all +ve points and -ve points

2. find the shortest line connecting these hulls.

3. bisect the line (Margin max plane).



SVM MATHS INTITUTION :-



(-4, 0)

$y = w^T x + b$

$m = -1$

$b = 0$ (∵ since origin)

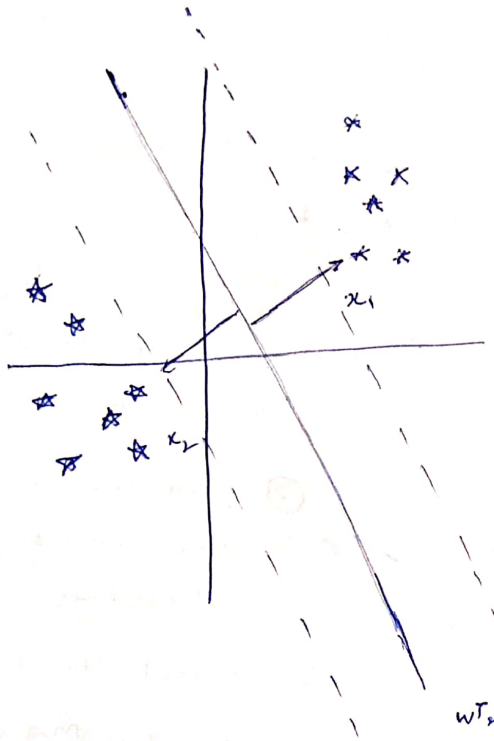
slope = -1 $\quad y = \begin{bmatrix} -1 \\ 0 \end{bmatrix} [-4 \;\; 0]$

$= 4 >$ +ve value

$y = w^T x$

$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} [4, 4]$

$= -4 >$ -ve value



Negative plane $\quad w^T x_1 + b = -1$

$w^T x + b = 0$

+ve plane $\quad w^T x_2 + b = 1$

$w^T x + b = 1$

$w^T x_1 + b = -1$

$w^T x_2 + b = +1$

$w^T (x_2 - x_1) = 2$

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

optimization Fun^n

$(w^*, b^*) \max \frac{2}{\|w\|}$

such that

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

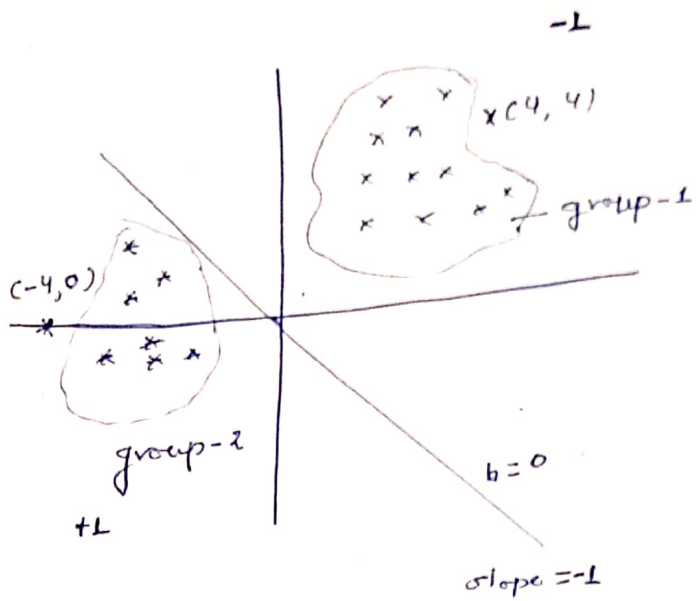
PAGE NO.

SVM MATHS INTITUTION :-

eq^{n}. of Hyperplane

$$y = w^T x + b = 0$$

Here $m = -1$, $b = 0$ ∴ (passes through origin)



$$y = w^T x + 0$$

$$= \begin{bmatrix} -1 \\ 0 \end{bmatrix}^{x_1} \begin{bmatrix} x_2 \\ -4, 0 \end{bmatrix}$$

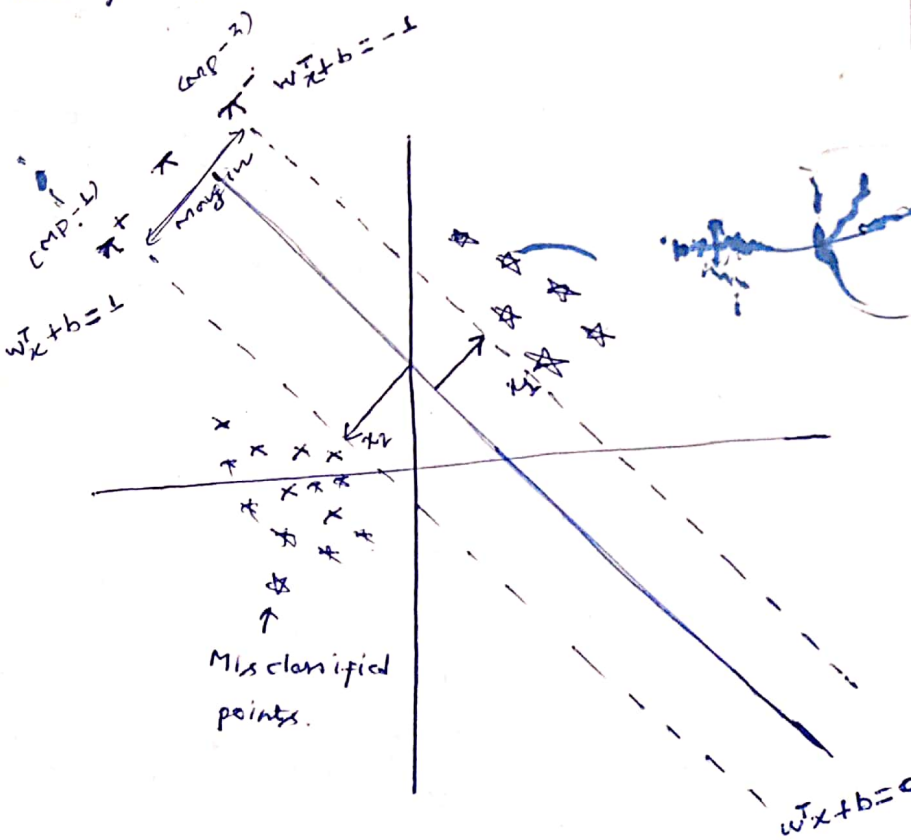
$= 4 \Rightarrow$ +ve value

likewise for any other point is always going to be an +ve value.

$$y = w^T x$$

$$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} [4, 4]$$

$= -4 \Rightarrow$ -ve value



For MP-2 -

$$w^T x_1 + b = -1$$

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

$$w^T (x_2 - x_1) = 2$$

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

Norm of w → $\|w\|$

This is optimization Function we need to maximize it

optimization Function -

$$(w^*, b^*) \quad \underset{(w,b)}{\text{argmax}} \frac{2}{\|w\|} \quad \text{such that}$$

$\forall_i$ for all

$$y_i \begin{cases} 1 & w_0^T x + b \geq 1 \\ -1 & w_0^T x + b \leq -1 \end{cases}$$

Hard Margin SVM

in other way

$$(w^*, b^*) \quad \underset{(w,b)}{\text{argmax}} \left\{ \text{margin}, \frac{2}{\|w\|} \right\} \quad \text{such that} \quad y_i * w^T x_i + b_i \geq 1 \quad \forall_i$$

[For correctly classified point]

otherwise its an mis classification

$$\max_x f(x) \Rightarrow \min_x \frac{1}{f(x)}$$

Take a reciprocal of the function

$$(w^*, b^*) \quad \underset{(w,b)}{\text{argmin}} \left(\frac{\|w\|}{2} + C \sum_{i=1}^{n} \xi_i \right)$$

Regularization

zeta i

zeta of i

denoted as 1/margin

How many errors we considered in the model?

value of error

avg dist of misclassified point for correct $\xi_i \uparrow$; point is further away from the correct & incorrect dist

$y_i * (w^T x_i + b) = 0.5$

$= 1 - (0.5)$

$\uparrow \xi_i$

if $\xi_i > 0$ (0.5 units away from the correct hyper plane)

$y_i * (w^T x_i + b) = 0.5$

$= 1 - (1.5)$

$\hookrightarrow \xi_i$

$y_i (w^T x_i + b) = 1.5$

$= 1 - (-0.5)$

$\uparrow \xi_i$

if $\xi_i = 0$ means correctly classified