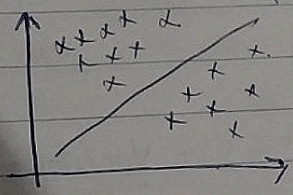
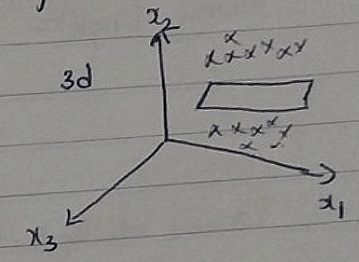
Support vector machine ML Algorithm.

① SVC (Support Vector Classifier) → Classification

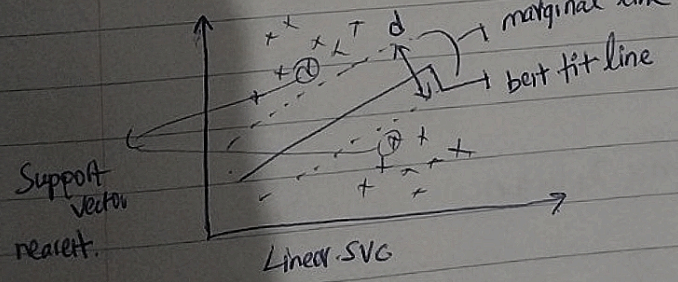
② SVR (Support Vector Regressor) → Regression

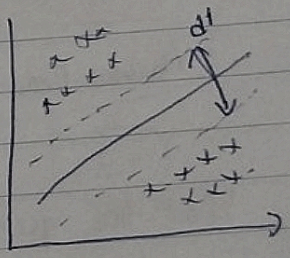




① Support Vector Classifier (SVC)

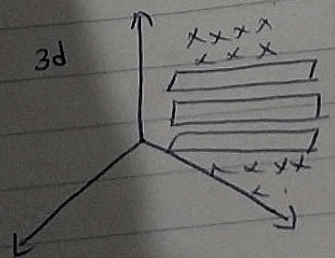
d = marginal plane distance



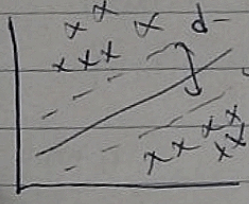


if $d' > d$ → go with d'

distance is maximum



Soft Margin And Hard Margin In SVC



Hard margin = None of the datapoints are misclassified

⇓

Impossible

Soft Margine → Some data points are misclassified (overlapping) we get some error compulsory

③ Support Vector Machines (SVC) Math Intuition

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

$$y = b + [w_1 x_1 + w_2 x_2 + w_3 x_3]$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$w^T = [w_1\ w_2\ w_3] \cdot \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$w^T x = [w_1 x_1 + w_2 x_2 + w_3 x_3]$$

$$\boxed{y = w^T x + b} \rightarrow y = mx + c$$

$$\rightarrow ax + by + c = 0$$

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

equation of straight line

$$y = mx + c$$

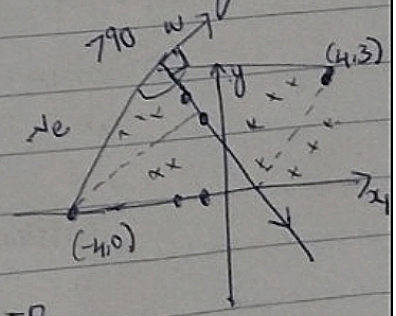
$$h_\theta(x) = \theta_0 + \theta_1 x_1$$

$$ax + by + c = 0$$

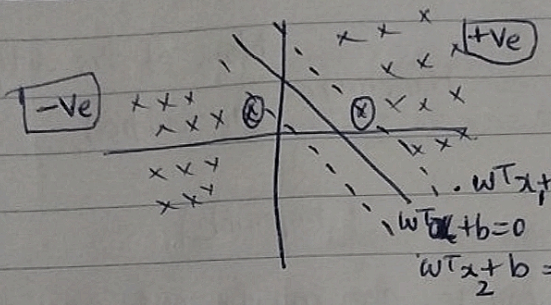
$$by = -ax - c$$

$$y = \boxed{\frac{-a}{b}} x + \frac{c}{b}$$

$$y = mx + c$$



Marginal plane in SVC



$P_1 \in x_1 \Rightarrow (x_1, x_2)$

$P_2 \in x_2 \Rightarrow (x_1, x_2)$

$w^T x_1 + b = +1$

$w^T x_2 + b = -1$

$(-) \quad (-1 \quad (+)$

$\vec{w} = \dfrac{w^T(x_1 - x_2) = 2}{\|w\|} \Rightarrow \dfrac{2}{\|w\|}$ distance between marginal plane

Cost function

$$\underset{w,b}{\text{Maximize}} = \boxed{\frac{2}{\|w\|}} \Rightarrow \text{Distance between marginal plane}$$

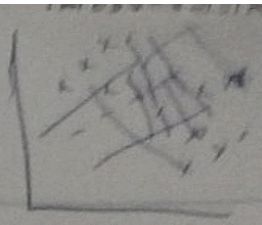
Constraint such that $y_i = \begin{cases} +1 & \text{if } w^T x + b \geq 1 \\ -1 & \text{if } w^T x + b \leq -1 \end{cases}$

For all correctly classified data points

Modified cost function of SVC

$$\underset{w,b}{\text{maximize}} \ \frac{2}{\|w\|} \Rightarrow \boxed{\underset{w,b}{\text{minize}} \ \frac{\|w\|}{2}}$$

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



Cost function of soft margin (lot of overlaps)

$$\text{Cost fn} = \min_{w,b} \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{n} \xi_i} \rightarrow \text{Hinge loss}$$

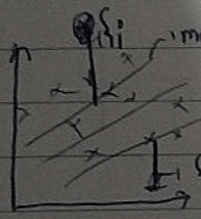
ξ_i → summation of the distance of incorrect data points to margin plane

C → Hyperparameter {how many points we can consider for misclassified}

$$C = \frac{1}{\lambda}$$

Support Vector Regressor (SVR)

ϵ = marginal error



$w^T x + b + \epsilon$

$w^T x + b - \epsilon$

ξ_i = Error above the margin

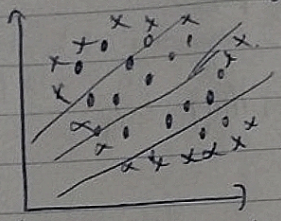
Cost fn

$$\min_{w,b} \frac{\|w\|}{2} + \boxed{C \sum_{i=1}^{n} \xi_i} \rightarrow \text{Hingeloss}$$

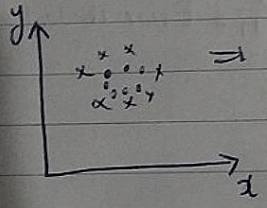
Constant

$$\boxed{|y_i - w^T x_i| \leq \epsilon + \xi_i}$$

↳ Why Kernel ↳ The dataset was more overlapping above the margin line and bellow margin. with linear svm we cannot get best fit line. so we use kernel in there it convert 2d to 3d dimension then we get best fit line.



↳ see the two features are half are above and half are bellow. The accuracy will not be more than 50% so we have to use svm-kernels



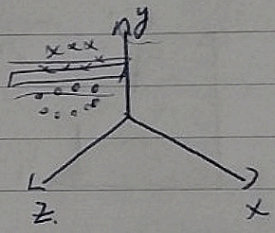
⇒ Transformation

‖

Mathematic formula

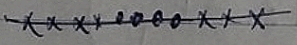
$z = f(x, y)$

↳ mathematical formula



Dataset		$z = x^2$
X	Y	
2	Yes	4
3	No	9
4	Yes	16

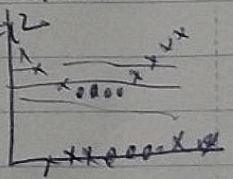
identify ⇒



svm kernel

↳ Data transformation method

$z = x^2$



1) Polynomial Kernel
2) RBF Kernel
3) Sigmoid kernel

} Transformation ⇒ mathematical formula

① Polynomial kernel

$$f(x,y) = (x^T y + c)^d$$

$c = 1$

$x = [x_1, x_2]$

$y = [x_1, x_2]$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} [x_1, x_2] = \begin{bmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{bmatrix}$$

2d
x_1 x_2 y

↓

3d
x_1^2 $x_1 x_2$ x_2^2

= 1 Polynomial kernel.

② RBF kernel ③ Sigmoid kernel