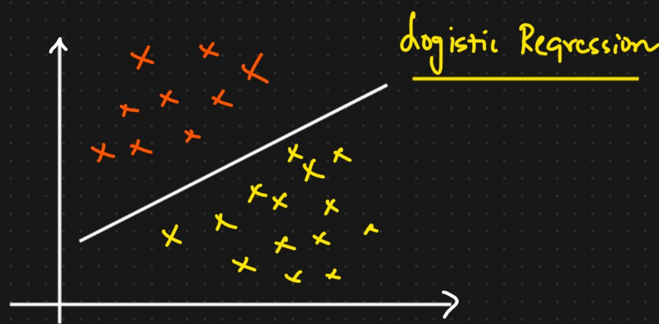
Support Vector Machines ML Algorithm

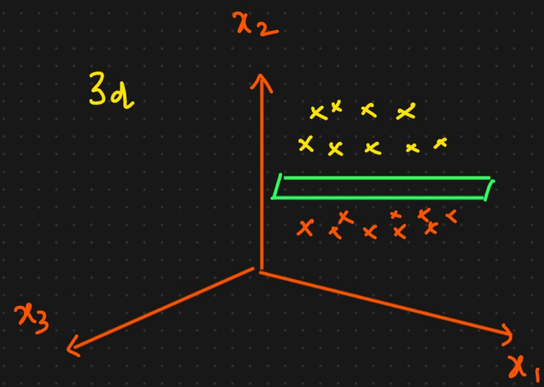
① SVC (Support Vector Classifier) ⟶ Classification

② SVR (Support Vector Regressor) ⟶ Regression

2d

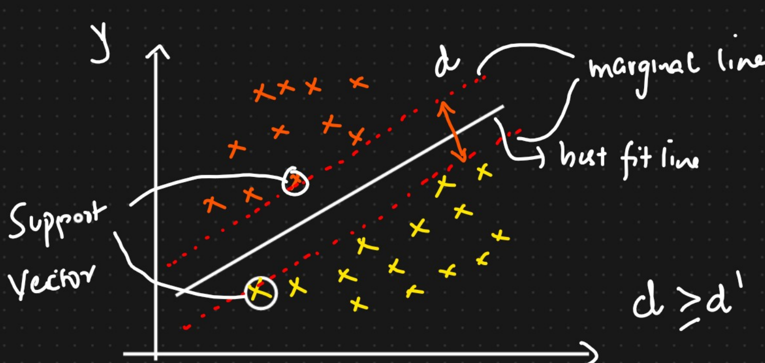


3d

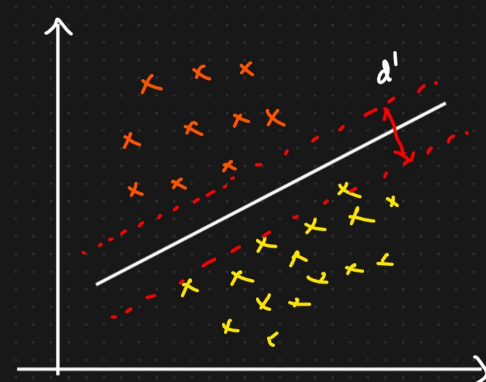


① Support Vector Classifier (SVC)

d = marginal plane distance



$d \geq d'$

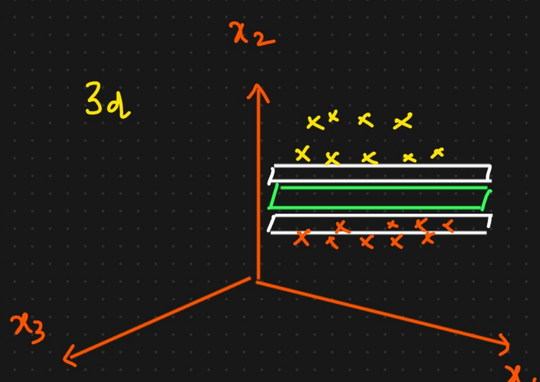


Linear SVC

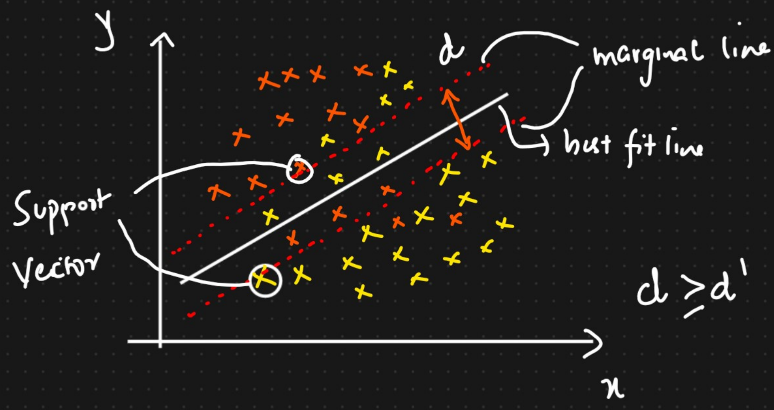
distance is maximum

if $d' \geq d$

3d



Soft Margin And Hard Margin In SVC

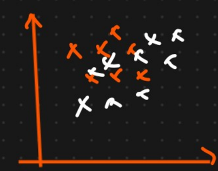


$d \geq d'$

Hard Margin = None of the datapoints are missclassified

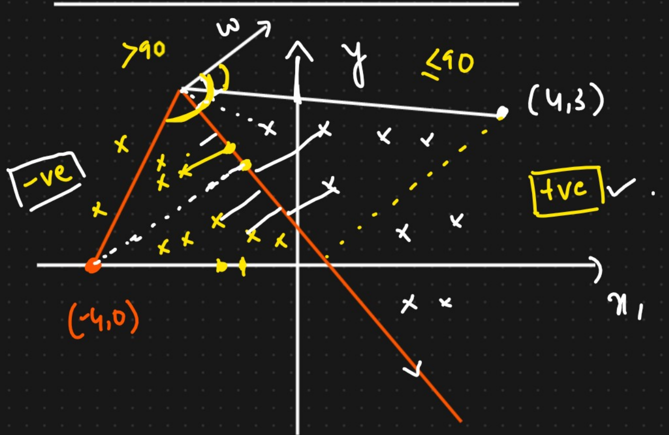
⇓

$\boxed{\text{Impossible}}$



② **Soft Margin** ÷ Some data point are missclassified [Error]

③ Support Vector Machines (SVC) Maths Intuition



$\boxed{w^T x + b = 0}$

Equation of a straight line

$\hat{y} = mx + c \iff ax + by + c = 0$

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

$by = -ax - c$

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

$y = mx + c$

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

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

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

$W^T = [w_1 \ w_2 \ w_3] \bullet X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

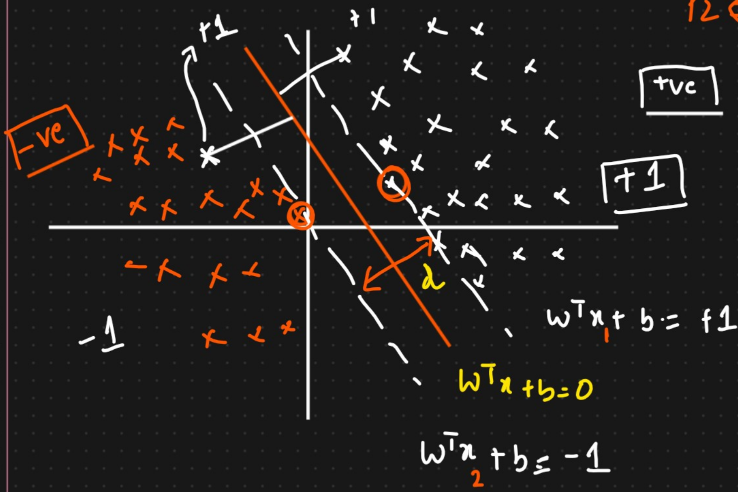
$\boxed{w^T x} = [w_1 x_1 + w_2 x_2 + w_3 x_3]$

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

$ax + by + c = 0$

$$\boxed{w^T x + b = 0}$$

Marginal plane in SVC



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

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

$w^T x_1 + \cancel{b} = +1$

$w^T x_2 + \cancel{b} = -1$

(-) (-) (+)

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

$\Downarrow$

Maximize

Cost function

Maximize $\boxed{\dfrac{2}{\|w\|}} \Rightarrow$ Distance between Marginal plane
w, b

Constraint such that

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

$\Downarrow$

For all correctly classified data points

$$\boxed{y_i * w^T x + b} \geq 1.$$

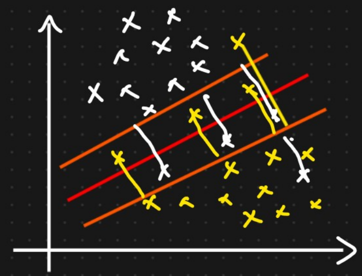
Modified Cost function of SVC

Maximize $\boxed{w,b}$ $\frac{2}{||w||}$ $\Rightarrow$ $\boxed{\text{Minimize}_{w,b} \ \frac{||w||}{2}}$

Constraint such that

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

$\boxed{C_i = 5}$



Cost function of Soft Margin SVC

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

C_i → hyperparameter

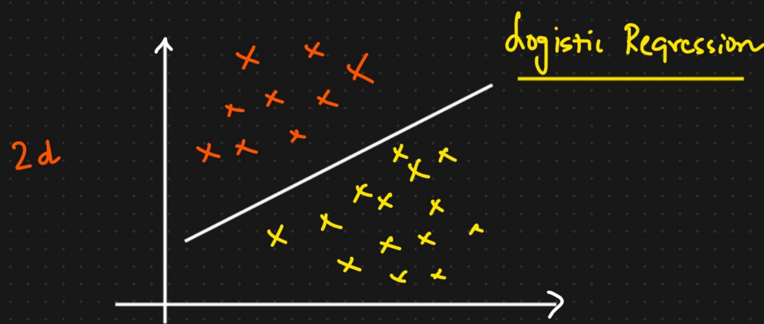
{How many points we can consider for misclassification}

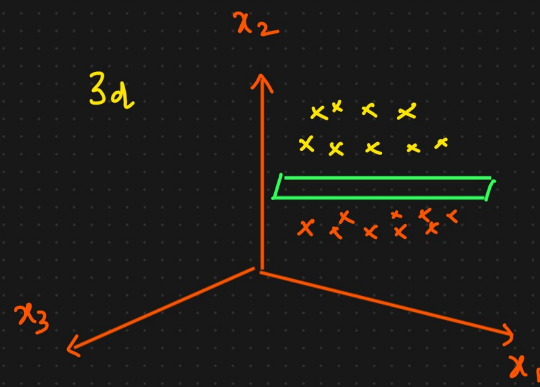
ξ_i → Summation of the distance of incorrect data points to the marginal plane

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

Support Vector Machines ML Algorithm

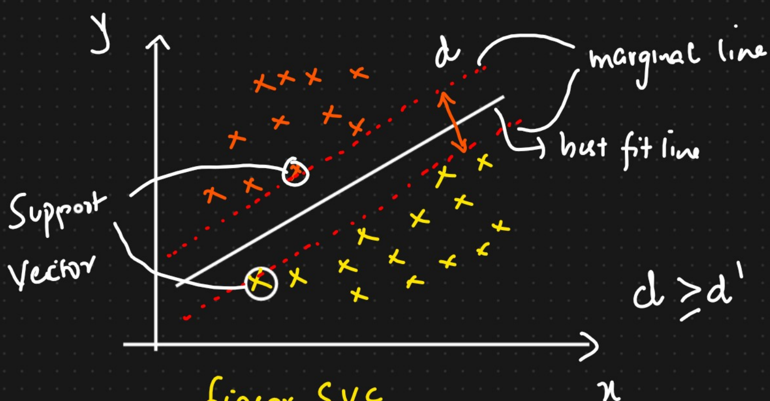
1. SVC (Support Vector Classifier) ⟶ Classification
2. SVR (Support Vector Regressor) ⟶ Regression

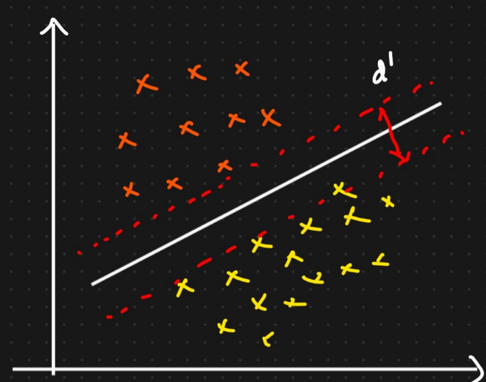




1. Support Vector Classifier (SVC)

d = marginal plane distance

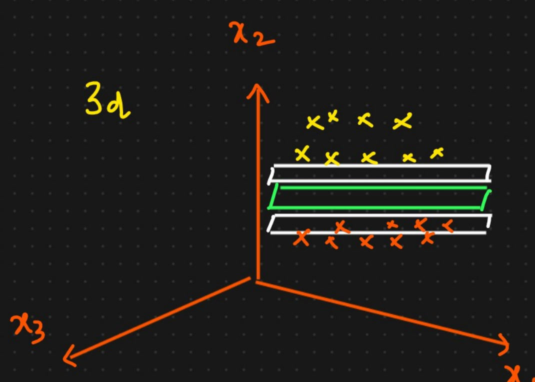




Linear SVC

distance is maximum

if $d' \geq d$

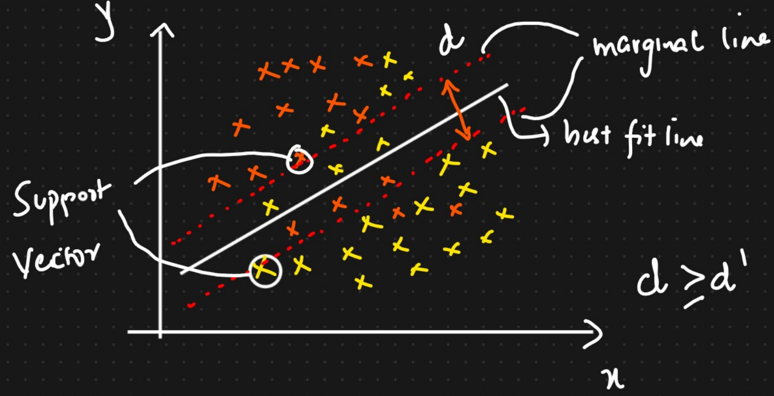


Soft Margin And Hard Margin In SVC

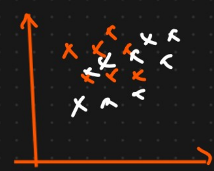
Hard Margin = None of the datapoints are missclassified

⇓

$\boxed{\text{Impossible}}$

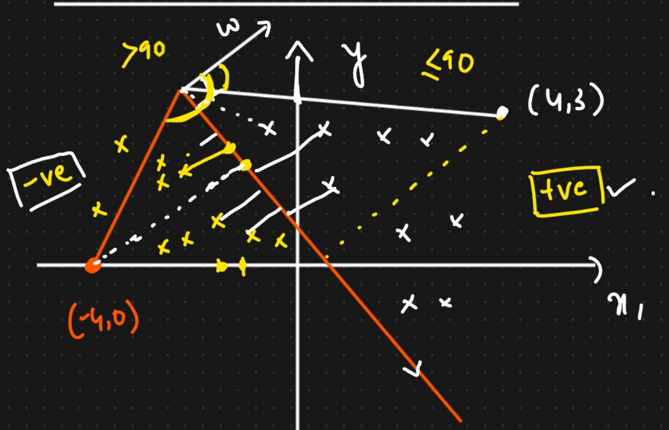


$d \geq d'$



② Soft Margin ÷ Some data point are missclassified [Error]

③ Support Vector Machines (SVC) Maths Intuition



$\boxed{w^T x + b = 0}$

Equation of a straight line

$\hat{y} = mx + c \quad \Leftrightarrow \quad ax + by + c = 0$

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

$by = -ax - c$

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

$y = mx + c$

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

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

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

$W^T = [w_1 \; w_2 \; w_3] \cdot X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

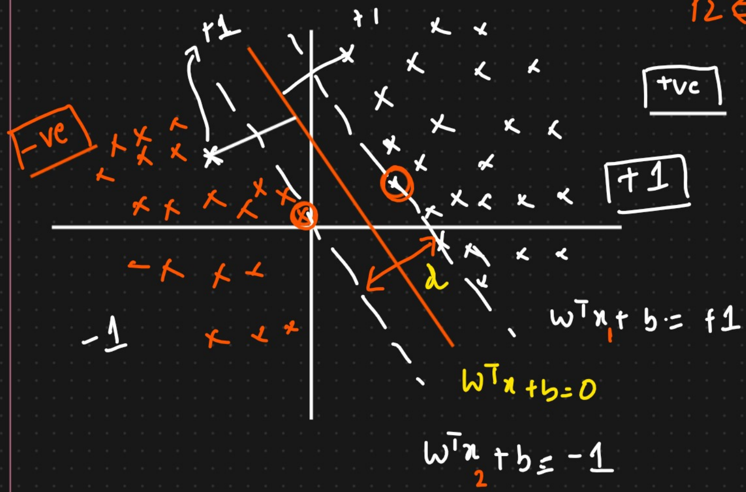
$\boxed{W^T x} = [w_1 x_1 + w_2 x_2 + w_3 x_3]$

$\boxed{y = W^T x + b} \Rightarrow y = mx + c$

$ax + by + c = 0$

$$\boxed{w^T x + b = 0}$$

Marginal plane in SVC



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

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

$w^T x_1 + \cancel{b} = +1$

$w^T x_2 + \cancel{b} = -1$

(-) (-) (+)

$$\vec{w} \Leftarrow \frac{w^T(x_1 - x_2)}{\|w\|} = \boxed{\frac{2}{\|w\|}} \Rightarrow \text{distance between Marginal planes.}$$

$\Downarrow$

Maximize

Cost function

Maximize $\boxed{\frac{2}{\|w\|}}$ $\Rightarrow$ Distance between Marginal plane
w, b

Constraint such that

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

$\Downarrow$

For all correctly classified data points

$$\boxed{y_i * w^T x + b} \geq 1.$$

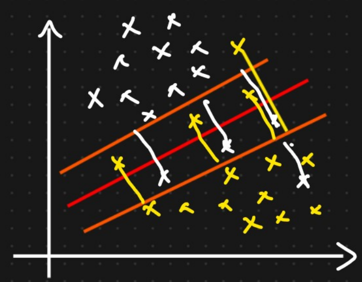
Modified Cost function of SVC

Maximize $\frac{2}{||w||}$ over $\boxed{w,b}$ $\Rightarrow$ $\boxed{\text{Minimize}_{w,b} \ \frac{||w||}{2}}$

Constraint such that

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

$\boxed{C_i = 5}$



Cost function of Soft Margin SVC

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

C_i: hyperparameter

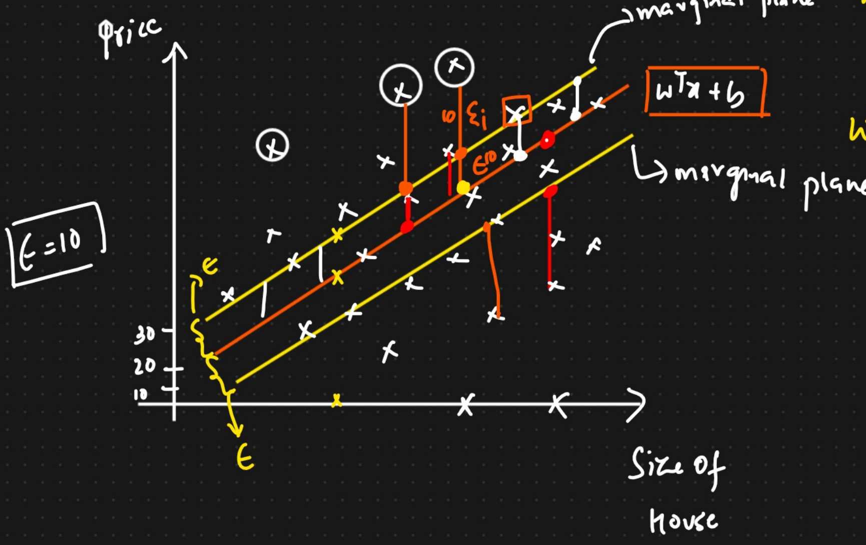
{How many points we can consider for misclassification}

ξ_i: Summation of the distance of incorrect data points to the marginal plane

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

Support Vector Regressor (SVR)

ϵ : Marginal Error



$\xi_i = 10$, $\epsilon = 10$

$\boxed{\leq 20}$

Cost fn

$$\min_{w,b} \frac{\|W\|}{2} + \boxed{C \sum_{i=1}^{n} \xi_i} \Rightarrow \text{Hinge loss}$$

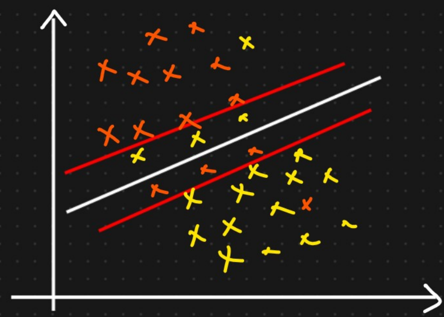
Constraint

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

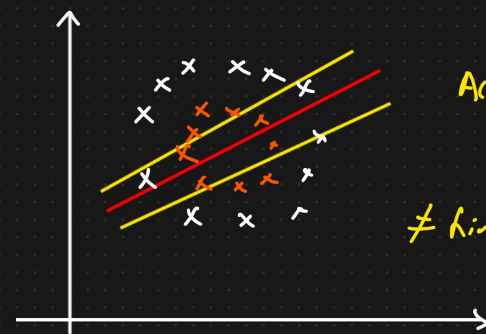
ϵ Marginal Error

ξ_i = Error above the Margin

SVM Kernels



Linear SVC

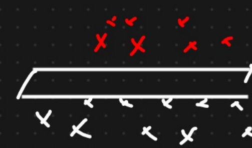


Accuracy = 50%

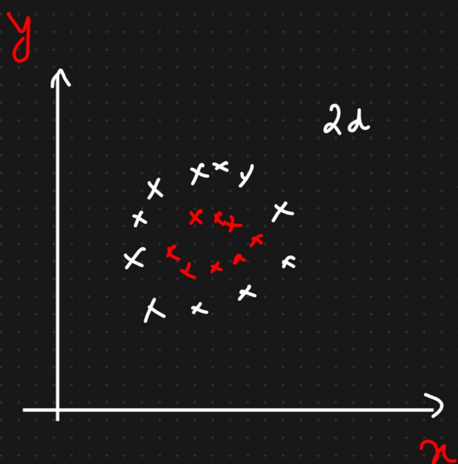
≠ Linear SVC

⇓

SVM Kernels



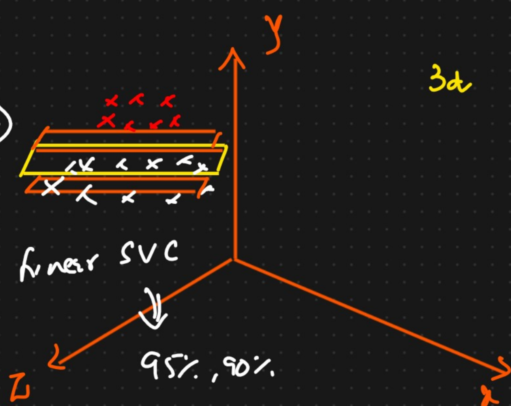
⇓



⇒ TRANSFORMATION ⇒

⇓

Mathematical Formula



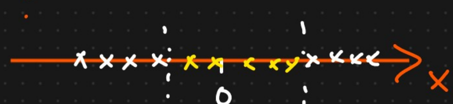
$z = f(x, y)$

↳ Mathematical formula.

Dataset

X	y	$z = x^2$
2	Yes	4
3	No	9
4	Yes	16
–	Yes	–
–		

1 dimension



⇒

$\boxed{1d - 2d}$

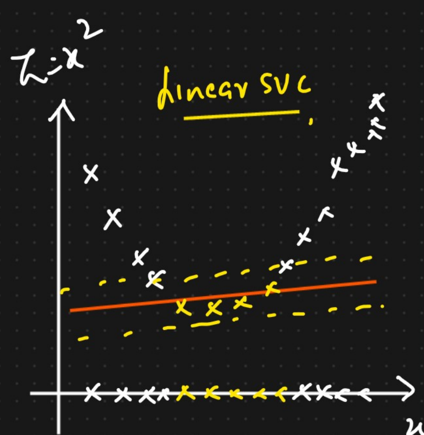
SVM Kernel

⇓

Data Transformation

$z = x^2$

Linear SVC

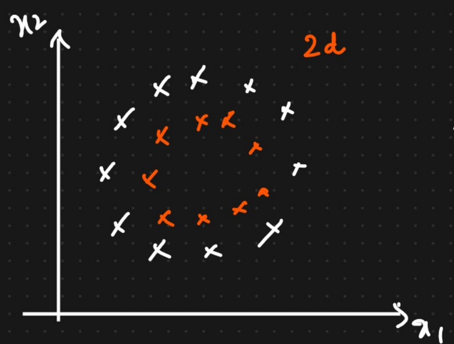


① Polynomial Kernel
② RBF Kernel
③ Sigmoid Kernel

} ⇒ Transformation ⇒ Mathematical Formula

SVM Kernels

1. Polynomial Kernel
2. RBF Kernel
3. Sigmoid Kernel



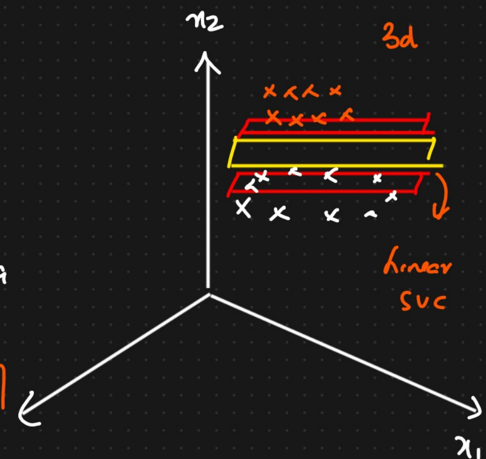
$\Rightarrow$ TRANSFORMATION

$\Downarrow$

Mathematical Formula

1. Polynomial Kernel
2. RBF Kernel
3. Sigmoid Kernel

$\boxed{Z}$



$x = [x_1, x_2]$

$y = [x_1, x_2]$

1. Polynomial Kernel

$$f(x,y) = (x^T y + c)^d \qquad \boxed{c=1}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} x_1, x_2 \end{bmatrix}$$

$\boxed{2d}$

x_1 x_2 y

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

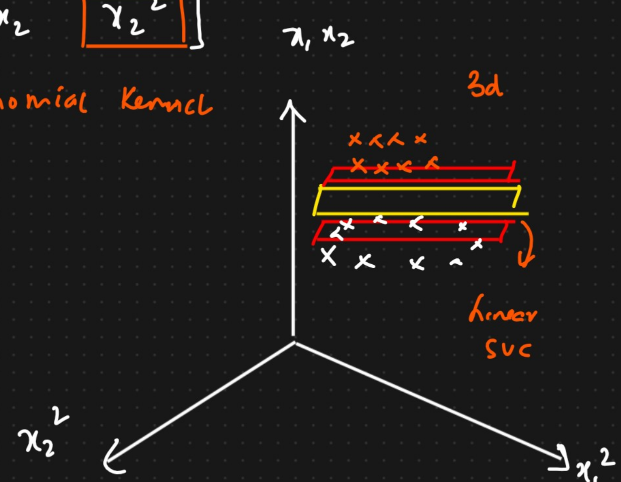
$\Downarrow$ Transformation

$\boxed{3d}$

x_1^2 $x_1 x_2$ x_2^2 y

x_1^2	$x_1 x_2$	x_2^2
–	–	–
–	–	

$\Rightarrow$ Polynomial Kernel



② RBF Kernel

⇓

Formula

③ Sigmoid Kernel

⇓

Formula