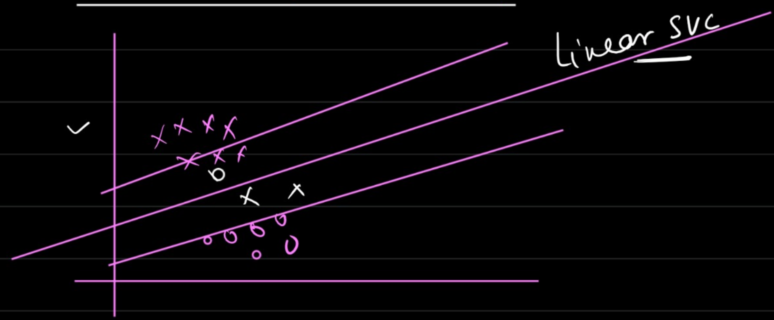
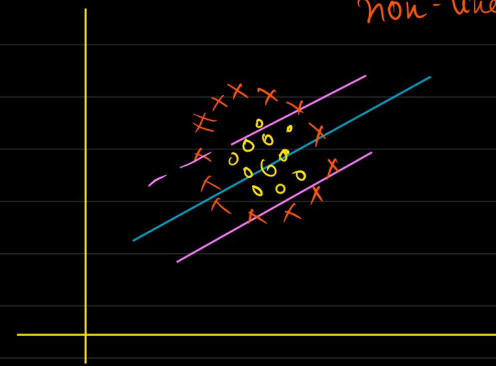
SVM - Kernel trick

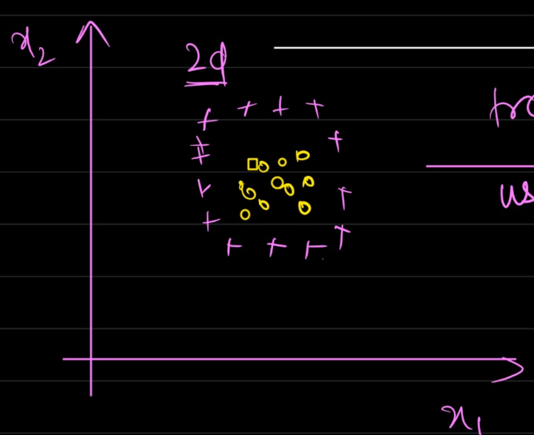


non-linear data

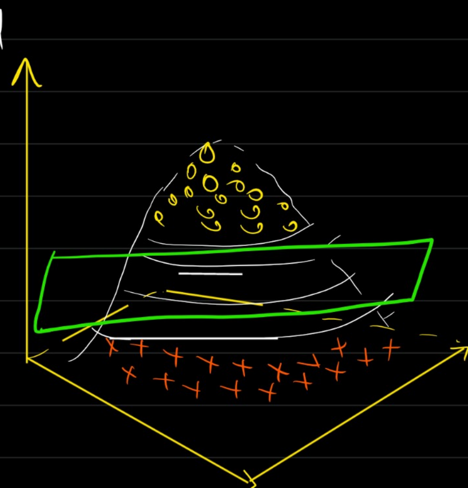


⇓ We can not solve using linear SVC.

* SVM kernel trick



transformation → using mathematical forunmulae



* 2d - 3d (Lower dimension to higher dimension using mathematical transformation, you can easily segregate between the classes.

$f(x) \Rightarrow$ Kernel $\Rightarrow$ kernel transformation.

(mathematical transf)

→ There can be some overlapping after kernel transformation.

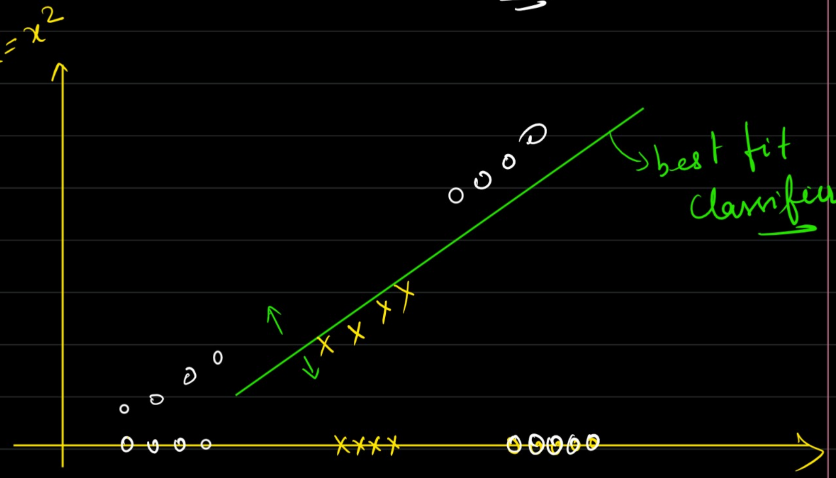
X	y
2	Yes
3	No
4	Yes
5	No
⋮	Yes

$\Rightarrow$ 

↳ You cannot segregate both the class using a linear SVC.

29

$\Rightarrow z = x^2$



x	x^2
2	4
3	9
4	16
5	25
⋮	⋮

idea :- to change the data from 1d to 2d, 2d - 3d, or increase the dimension by mathematical transformation to distinguish between the classes.

Why? → you are not sending the data in higher dimension, you are using mathematical transformation to achieve it

→ Using SVC kernel trick, you can also classify non-linear data.

* Kernel function

1. Polynomial
2. Rbf (Radial basis function)
3. Sigmoid

$ax^2 + bx + c = 0$

$ax^3 + bx^2 + cx + d = 0$

1. Polynomial Kernel

$$f(X_1, X_2) = (X_1^T \cdot X_2 + c)^d$$

→ d is degree of Polynomial

constant

$(X_1^T \cdot X_2 + 1)^q$

Suppose we have two features X_1 & X_2

Y is output variable

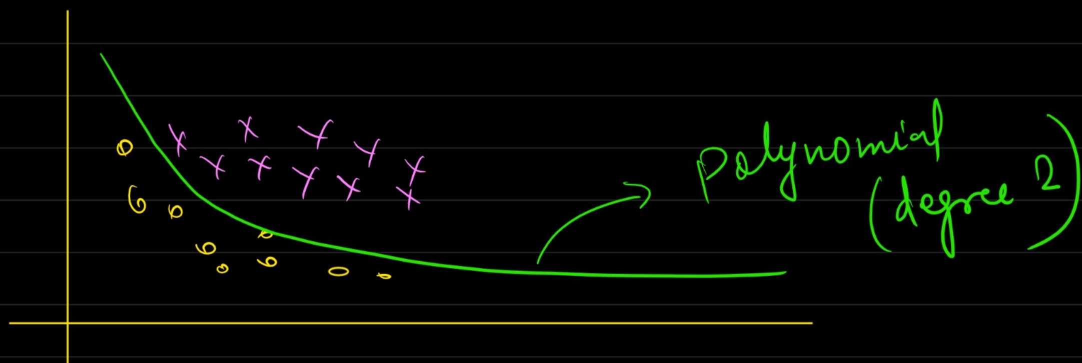
$$x_1^T \cdot x_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

dot product

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

$$x_1 x_2 \Longrightarrow x_1, x_2, x_1^2, x_1 x_2, x_2^2$$

2d $\Longrightarrow$ 5 dimension.



② Radial basis feature (Rbf kernel).

→ Creates non-linear combination of features to bring your feature in higher dimension.

$$f(x_1, x_2) = e^{\frac{-||(x_1 - x_2)||^2}{2\sigma^2}}$$

→ $||x_1 - x_2||$ – Euclidean distance b/w two point x_1 and x_2

→ σ – the variance / hyperparameter

To remember

Radial → radius.

③ Sigmoid Kernel

$$f(x,y) = \frac{1}{1+e^{-x}}$$

* Bessel Kernel ⟶ $f(x,y) = \dfrac{J_{v+1}(\sigma\|x-y\|)}{\|x-y\|^{-n(v+1)}}$

* ANOVA Kernel ⟶ $f(x,y) = \displaystyle\sum_{k=1}^{n} \exp\left(-\sigma\left(x^k-y^k\right)^2\right)^d$

⇓

multidimensional regression

* How to choose right kernel.

⇓

hyperparameter tuning.