Date.....................

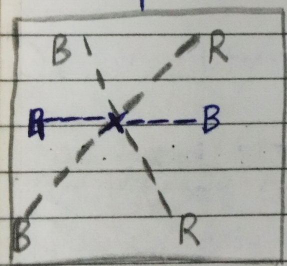
# Types of Kernels:-

- rbf Kernel
- Polynomial Kernel
- Sigmoid Kernel
- Custom Kernel

## 1. Radial Basis Function

Red points ⇒ (1,2) (2,3) (3,4)
Blue points ⇒ (6,5) (5,6) (7,7)

In 2D plane, the data might look like:-




- As, we can see that, in this plane, the straight line cannot seperate the 'Red' and 'blue' points effectively.
- Here, comes the use of RBF Kernal.

rbf ⇒ $\boxed{K(x,y) = e^{-\gamma \|x-y\|^2}}$

"Where:-
$\gamma$ [gamma] ⇒ It is a hyper parameter which is used to increase the smoothness of efficiency."

Let's compute the Kernel values as per the data points given:-

{Assuming that, $\gamma = 1$}

1. $K((1,2),(6,5))$
$\Rightarrow e^{-1\|(1,2)-(6,5)\|^2}$
$\Rightarrow e^{-1\|(6-10)\|^2} \Rightarrow e^{-1\times 16} \Rightarrow e^{-16} \Rightarrow \boxed{0.000335}$ — (1)

2. $K((2,3),(5,6))$
$\Rightarrow e^{-1\|(2,3)-(5,6)\|^2} \Rightarrow e^{-1\|(10-18)\|^2} \Rightarrow e^{-1.64} \Rightarrow \boxed{0.000045}$ — (2)

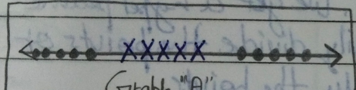
Date.........

3. $K((3,4),(6,5)) \Rightarrow e^{-1\|(3,4)-(6,5)\|^2}$
$\Rightarrow e^{-1\|(18-20)\|^2} \Rightarrow e^{-1\times 4} \Rightarrow \boxed{0.006738} \Rightarrow$ (3)

From (1), (2) and (3), We can notice that as Kernel values are <u>decreasing</u> as the distance between the data points <u>Increases.</u>

# Polynomial Kernel

Example 1:- $\boxed{\begin{array}{l} x \rightarrow \text{Positive} \\ \bullet \rightarrow \text{Negative} \end{array}}$

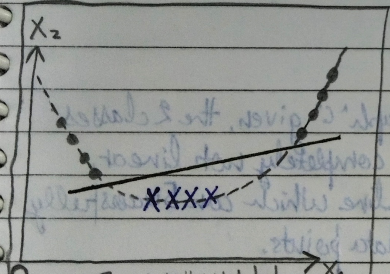


Graph "A"

- As per the Graph 'A', 1D, we can see that the data is not linear. So, we cannot draw a single line which can classify the data points correctly.

$$\text{poly} \Rightarrow \boxed{K(x,y) = (x \cdot y + c)^d}$$

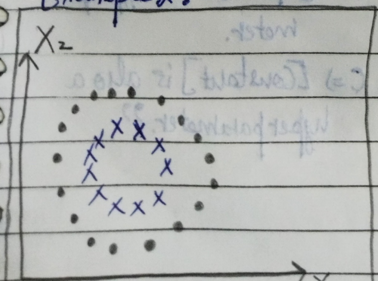
where :-
"c ⇒ constant term.
d ⇒ dimension / degree of the polynomial."



Graph "A" Updated

- After applying the polynomial Kernal on the data points, now we can see that, a plane or line has successfully classified the points. So, this line is the best distribution line.

Example 2:-



Graph "B"

- As per the Graph "B", we can observe that the 2D data is not linearly seperable at all. So, in that case drawing a single line to perfectly classify the data points..

Spiral

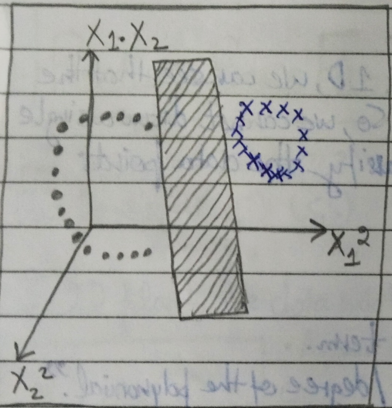
Date..........

$$X_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \qquad X_2 \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \qquad X_1^T \cdot X_2 \Rightarrow \begin{bmatrix} x_1^2 & x_1 \cdot x_2 & x_1 \cdot x_3 \\ x_1 \cdot x_2 & x_2^2 & x_2 \cdot x_3 \\ x_3 \cdot x_1 & x_3 \cdot x_2 & x_3^2 \end{bmatrix}$$

$$X_1^T \Rightarrow [x_1 \; x_2 \; x_3 \cdots x_n]$$

So, the 3 main Components are:-

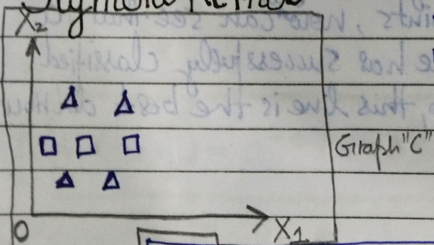
$$\boxed{x_1^2, x_2^2, x_1 \cdot x_2}$$




Graph "B" Updated

• We can see that, after applying the "poly" Kernal function to the datapoints of '$X_1$' and '$X_2$'. We get a hyperplane which can easily divide the points or bestly classify the points.

# Sigmoid Kernel




Graph "C"

• As per the Graph "C" given, the 2 classes "Δ" and "□" are completely not linear. So there is no line which can successfully classified the data points.

$$\text{sigmoid} \Rightarrow \boxed{K(x,y) = \tanh * (\alpha * (x-y) + C)}$$

"Where :-
$\alpha$ = [alpha] is a hyperparameter.
C => [Constant] is also a hyperparameter."

Let's say the Δ data points = [2, 3]
□ data points = [1, 4]

Date........................

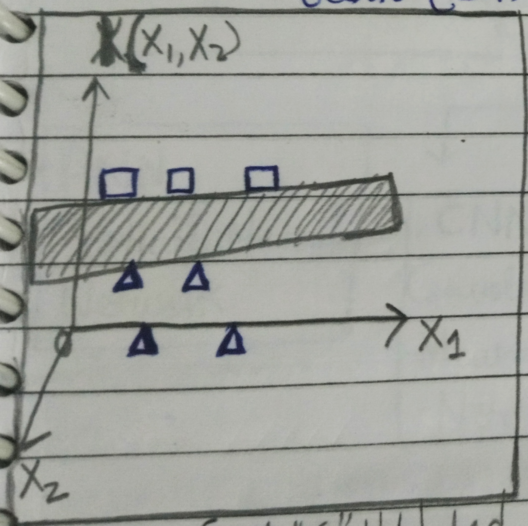
Computing "K" value as per sigmoid Kernel.

$$K(x,y) = \tanh(1 * ([2,3] \cdot [1,4]) + 0)$$

[ Assuming that:- $\alpha = 1$, $C = 0$ ]

$$= \tanh(1 * (2 + 12))$$

$$= \tanh(14) \Rightarrow \boxed{0.09999969}$$




Graph "C" Updated