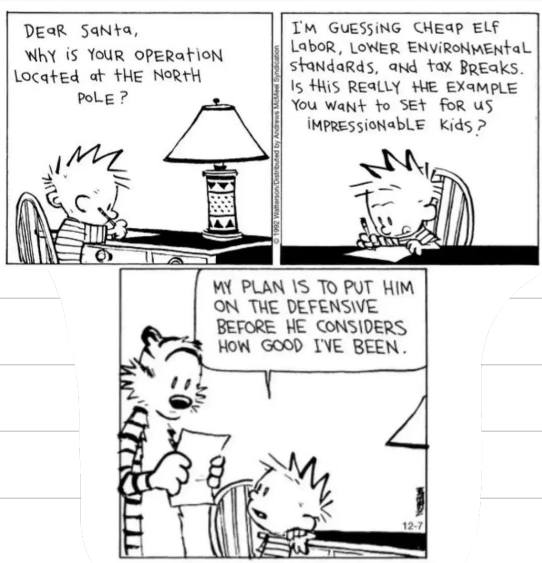
Session 12

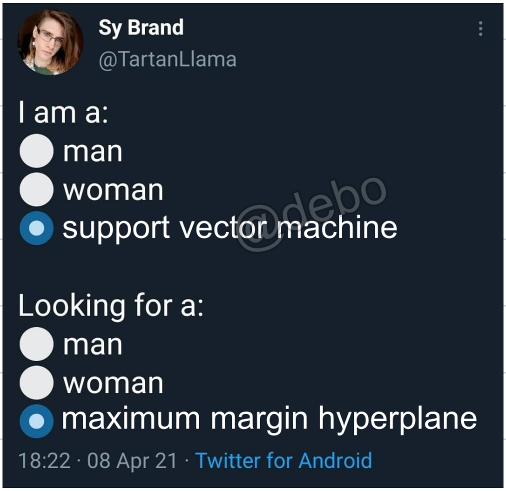
SVM - 2

Sep 19, 2025



Agenda

1. Recap SVM - 1
2. Hinge Loss visualization
3. Primal Dual Equivalence form
4. Kernel Function

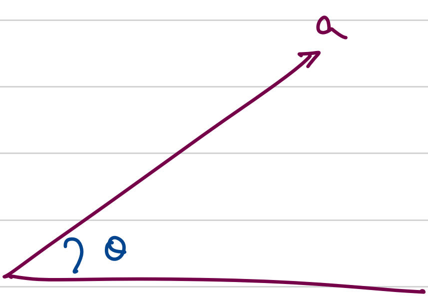


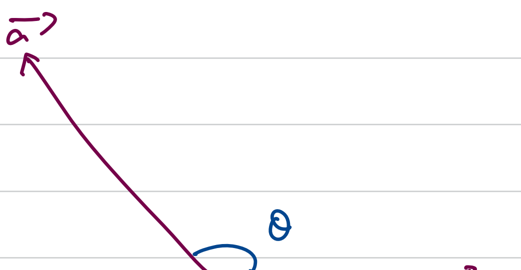
Hinge Loss

Revision

$\vec{a}$ & $\vec{b}$; $\vec{a} \cdot \vec{b}$ = Similarity

Cosine Sim: $\dfrac{\vec{a} \cdot \vec{b}}{||a|| \cdot ||b||}$





dot Product = +ve $\vec{b}$

dot pr = -ve

SVM - Primal Dual Form

1 Primal form

$$\underset{}{\text{argmin}} \frac{||w||}{2} + \frac{C}{N}\sum_{i=1}^{N} \xi_i \quad \text{S.T. } (w^T x + w_0) * y_i \geq 1 - \xi_i$$

Dual Form is a simpler form of primal form, which if solved, gives you same results, as that of primal form.

2 Dual Form

Kernel

$$\underset{\alpha}{\text{arg. max}} \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j \boxed{x_i^T x_j} y_i y_j$$

$$\text{S.T. } 0 \leq \alpha \leq C$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

Sequential minimal Optimization

Train: $O(n^2)$

D → 10K rows

R.F: 30-40 second.

SVM: 40 min

D → 1 mill rows

R.F → 3 min

S.V.M → 10 hours

1. $\alpha \rightarrow$ weightage for each row of dataset

		w_1	w_2	w_3	w_4	w_5	w_6	
		f_1	f_2	f_3	f_4	f_5	f_6	y
α_1	x_1							
α_2	x_2							
α_3	x_3							
α_4	x_4							
α_5	x_5							
α_6	x_6							
α_7	x_7							
α_8	x_8							

$\forall \rightarrow$ for all

2. $\alpha = 0 \ \forall$ rows which are not support vectors

   a. What is a support vector

   b. How does support vector help us?

(9) What is support vector?

Conditions for a vec to be support vector.

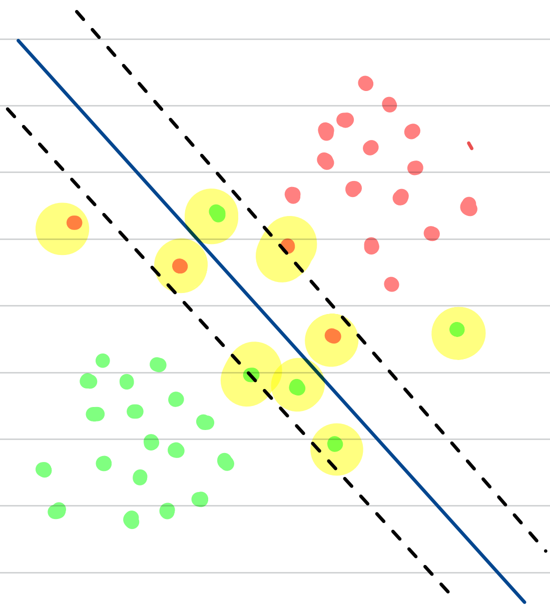
(a) It lies either on π^+ or π^-

OR

(b) It lies within (+)ve or (-)ve margin

OR

(c) If they are misclassified.



(b.) How does support vector help us

x_q (Query point)

→ dot product

(1) with all the (+)ve S.V.

(2) " " " (-)ve S.V.

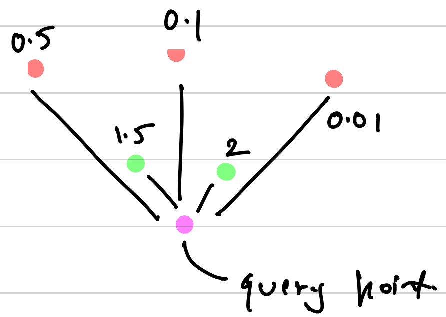
R G.

5 + (-10)

= -5 = G

Test : $O(k)$

$$f(x_q) = \sum_{i=1}^{K} \alpha_i \times y_i \times \underbrace{x_i^T \cdot x_q}_{\text{dot Product}}$$



$0.5 \times 1 + 0.1 \times 1 + 0.01 \times 1$

$+ 1.5 \times (-1) + 2 \times (-1)$

0.5 * 1 + 0.1 * 1 + 0.01 * 1 + 1.5 *(-1) + 2 * (-1) = -2.89

$f(x_q) \geq 0$ → (+)ve class

< 0 → (-)ve class

Which of the following are support vectors?

18 users have participated

A	Points which are within the margin	6%
B	Points which lie on +ve/-ve hyperplane (π^+/π^-)	17%
C	Points which are misclassified	11%
D ✓	All of the above	67%

End Quiz Now

Based on all quizzes from the session

1	Sumana Pedapudi	1/1	94.59
2	Purushottam Ku...	1/1	94.10
3	Nachiket Pawar	1/1	92.90
4	SHASHANK JHA	1/1	91.13
5	tejas sinha	1/1	90.60
6	Hanumanthgouda Patil	1/1	89.67
7	RAHUL	1/1	87.27
8	Snehal Adhikary	1/1	84.57
9	Ankita Parhi	1/1	81.33
10	Sri Harsha Nanduri	1/1	76.03

We have 100 datapoints out of which 5 are Support Vectors, then which is True:

0 users have participated

A	α >0 for 95 datapoints	0%
B	α <0 for 95 datapoints	0%
C	α = 0 for 5 datapoints	0%
D ✓	α >0 for 5 datapoints	0%

End Quiz Now

Based on all quizzes from the session

1	Purushottam Ku...	2/2	185.70
2	SHASHANK JHA	2/2	184.05
3	Nachiket Pawar	2/2	177.07
4	tejas sinha	2/2	169.60
5	Ankita Parhi	2/2	166.40
6	Sri Harsha Nanduri	2/2	163.30
7	Samyucktha Ramesh	2/2	151.56
8	Deeksha Sharma	1/2	98.06
9	Sumana Pedapudi	1/2	94.59
10	Hanumanthgouda Patil	1/2	89.67

Which of the following SVM can handle non linear data?

0 users have participated

A	Hard Margin SVM	0%
B	Soft Margin SVM	0%
C	Dual form of SVM	0%
D ✓	None of the above	0%

End Quiz Now

1	Purushottam Ku...	3/3	273.63
2	SHASHANK JHA	2/3	184.05
3	Nachiket Pawar	2/3	177.07
4	tejas sinha	2/3	169.60
5	Ankita Parhi	2/3	166.40
6	Sri Harsha Nanduri	2/3	163.30
7	Samyucktha Ramesh	2/3	151.56
8	Deeksha Sharma	1/3	98.06
9	Sumana Pedapudi	1/3	94.59
10	Hanumanthgouda Patil	1/3	89.67

Non-Linear SVM

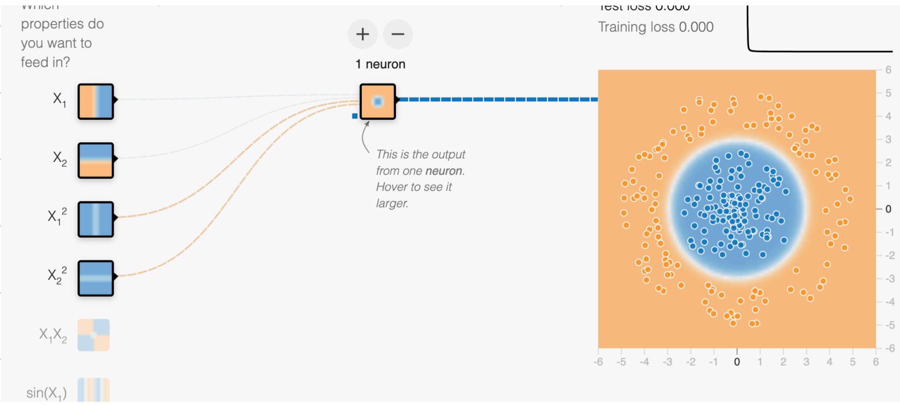
Linear dot Product

$$\arg\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \boxed{x_i^T x_j} y_i y_j$$

S.T. $0 \leq \alpha \leq C$

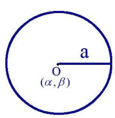
$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

Sequential minimal Optimization



x_1, x_1^2, x_2, x^2

The equation $x^2 + y^2 + 2gx + 2fy + c = 0$
always represents a circle.

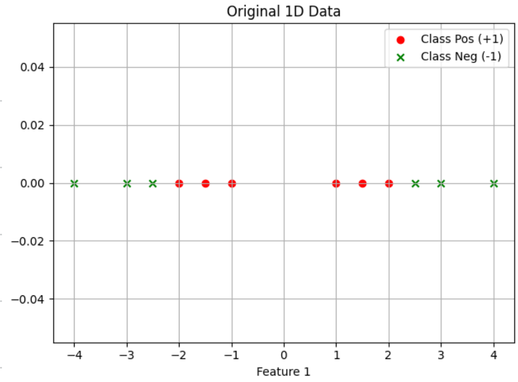


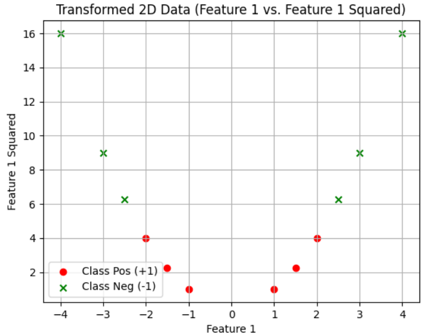
Its centre is $(-g, -f)$ and radius is $\sqrt{g^2 + f^2 - c}$

If $g^2 + f^2 > c$, the radius is real

If $g^2 + f^2 = c$, the radius vanishes and circle becomes a point circle.

If $g^2 + f^2 < c$, the radius is imaginary.





Polynomial Kernel

$$K(\vec{x_1}, \vec{x_2}) = \vec{x_1}^T \cdot \vec{x_2} \quad (\text{Linear})$$

$$= x_{11} \cdot x_{21} + x_{12} \cdot x_{22}$$

$$X_1 = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} \qquad X_2 = \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}$$

$$\text{Polynomial Kernel} = (c + x_1^T x_2)^n ; \; C = 1$$

$$n = 2 \quad = (1 + x_1^T x_2)^2$$

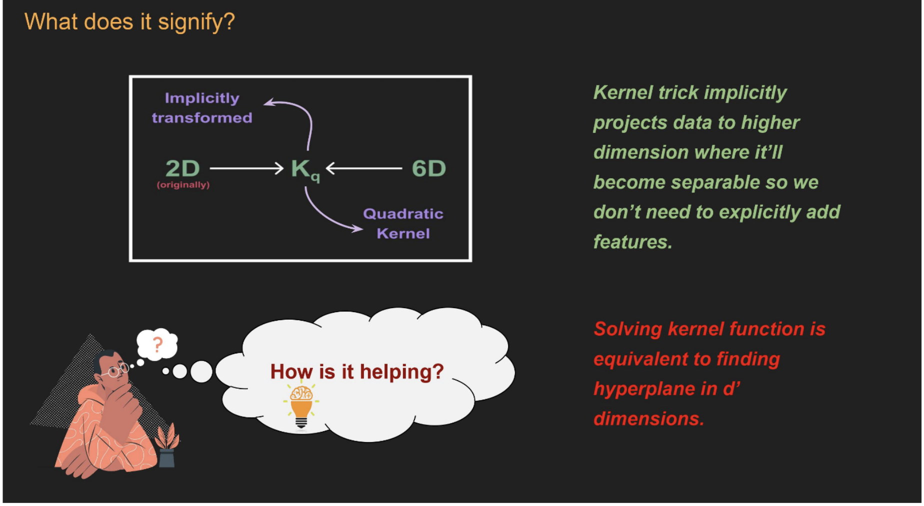
$$= (1 + x_{11} \cdot x_{21} + x_{12} \cdot x_{22})^2$$

$$(a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2bc + 2ac$$

$$\Rightarrow 1^2 + x_{11}^2 \cdot x_{21}^2 + x_{12}^2 \cdot x_{22}^2 + 2 \times x_{11} \cdot x_{21} \cdot x_{12} \cdot x_{22}$$

$$+ 2 \times 1 \times x_{11} \cdot x_{21} \quad \text{— } 2ac$$

$$+ 2 \times 1 \times x_{12} \cdot x_{22}$$



What of the following statement(s) is/are true about Kernel in SVM? Statement 1: Kernel function map low dimensional data to high dimensional space
Statement 2: It's a similarity function

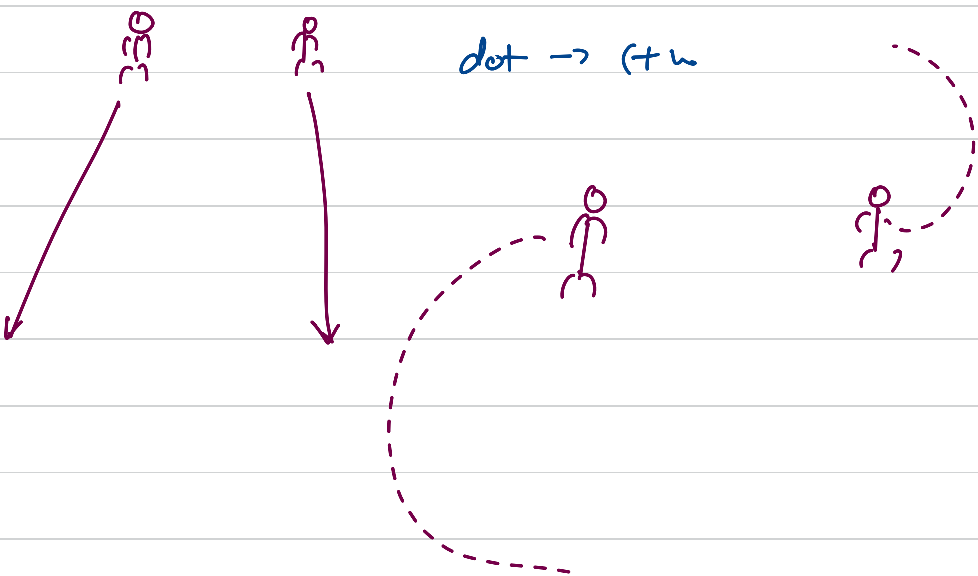
0 users have participated

A	Statement 1	0%
B	Statement 2	0%
C ✓	Statement 1 and 2	0%
D	None of the above	0%

End Quiz Now

Rank	Name	Score	Points
1	Purushottam Ku...	4/4	356.73
2	SHASHANK JHA	3/4	275.18
3	Nachiket Pawar	3/4	263.03
4	Ankita Parhi	3/4	250.70
5	Samyucktha Ramesh	3/4	228.50
6	Sri Harsha Nanduri	3/4	226.60
7	Deeksha Sharma	2/4	190.49
8	Sumana Pedapudi	2/4	185.52
9	Snehal Adhikary	2/4	178.77
10	RAHUL	2/4	172.20

dot → (+ w



RBF KERNEL

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

γ x_q S.V.

$$RBF = e^{-\frac{||x_1 - x_2||}{2\sigma^2}}$$

$$= e^{\gamma ||x_1 - x_2||} \quad x$$

$$= e^{\gamma \times x}$$

Taylor's expansion.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$||x_1 - x_2||^3$

$$e^{\gamma x} = e^{\gamma}\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \ldots\right.$$

$\gamma \uparrow \rightarrow$ Over

$\gamma \downarrow \rightarrow$ Under

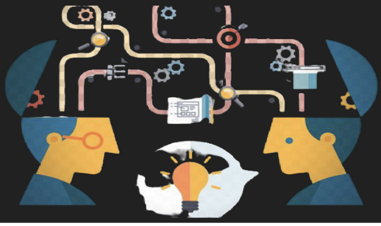
Will outliers affect SVM?

If using SVM without any kernels hyperplane will only depend on support vectors.

∴ So, outliers will not impact the hyperplane.

However, since kernel SVMs are based on distance/ proximity and similarity.

∴ The outliers will affect them.



9790723608