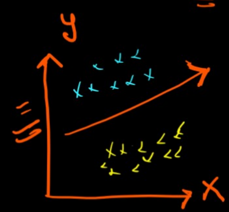
Log reg :- Prob → Line → Sigmoid fun → (Prob) → final out 0 | 1

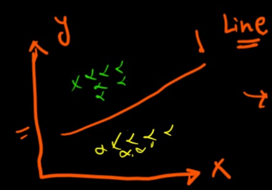
(Line → $y = mx + c$)

Log reg → (Linear seprable)

classification :- binary
multiclass (forest cover class)

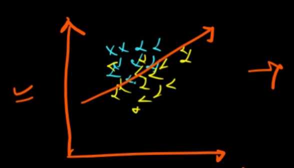


→ Linear seprable data.

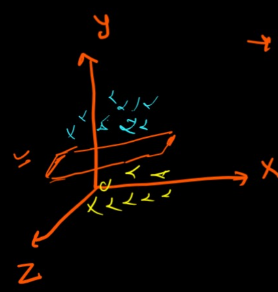


→ 2D

→ 11, 12

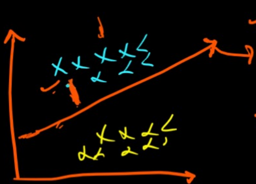


→ non linear seprable data



SVM → kernel trick

Linear Algebra



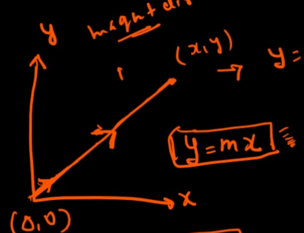
$y = mx + c$ → $y = m_1x_1 + m_2x_2 + m_3x_3 + c$

$y = W_0 x + C$

$(m^T x)$ → [ ] [ ] → Dot Product

ML mathematics

(Norms) → (distance b/w)



→ $y = mx + c$

$\boxed{y = mx}$

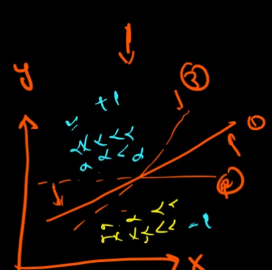
$\boxed{x\hat{i} + y\hat{j}}$

Point and Plane

Norms = $\dfrac{mx + c}{\|m\|}$ ← (vector) value + direction

Norms = $\dfrac{W^T x + C}{\|W\|}$

Norms = $\dfrac{W^T x}{\|W\|}$ → unit vector = 1

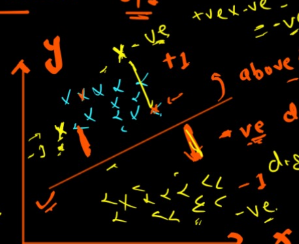


$\boxed{\text{Norms} = W^T x = mx}$ ⇒ $mx + c$

machine learning :- math → optimal target

1. cal.
2. loss.
3. optimization

Binary class : 0, 1



→ Correct class

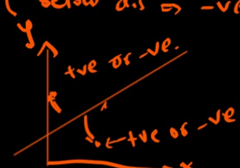
distance

(+1, -1)

(1, 0)

y actual	y Predict
0	0
0	0
1	1
1	0 x
0	1 x

{ above dis ⇒ +ve
below dis ⇒ -ve } → mathematically



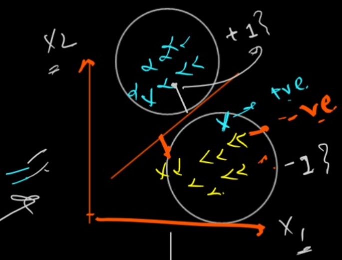
Case 1 :- +ve x y_i

→ +ve x +ve $\boxed{\text{blue}}$

⇒ +ve $\boxed{\text{Correct class}}$

Case :- 2 → +ve x -ve

dis   Point class



X_1 X_2 Y (0,1) or (-1,+1)

X_1	X_2		Y
55kg	160	→	1
50	170	→	0
60	180	→	0
70	185	→	0
80	190	→	1

+1, -1

→ -ve → wrong

Case:-3 $\boxed{-ve} \times \boxed{-ve}$

dist

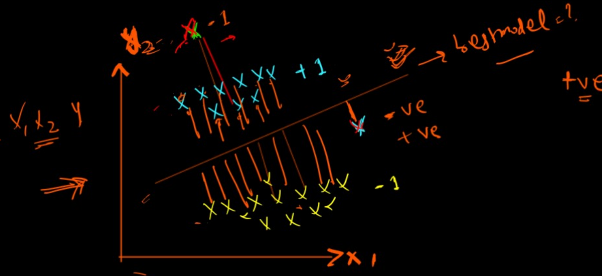
→ +ve → $\boxed{\text{Correct}}$

Case:4:- -ve × +ve

dis, ident point

→ -ve →

x_1, x_2, y



+ve

(Correct classification)

+ve =?

distance(+ve) × (+ve)

⇒ +ve

-ve × -ve = $\boxed{+ve}$

$\boxed{+ve \times -ve = -ve}$

wrong classification

Model

Line-1

5 Point

$\boxed{5\times+1} = 5$
$6\times+1 = 6$
$7\times+1 = 7$
$8\times+1 = 8$
$9\times+1 = 9$

= 35

+ ⇒ 70

5 Point

$-5\times-1 = 5$
$-6\times-1 = 6$
$-7\times-1 = 7$
$-8\times-1 = 8$
$-9\times-1 = 9$

= 35 = add(dis)

$\boxed{-ve \times +ve = -ve \text{ Wrong classification}}$

if my classifier is best classifier over here distance addition max

$$\boxed{\max \left(\sum_{i=1}^{n} y_i * w^T x \right)}$$

y_i: -1 or +1, $w^T x$: (dis)

+ve × +ve = +ve

-ve × -ve = +ve

= add

Line:-2

$\boxed{5x-1} = -5$
$6x+1 = 6$
$7x+1 = 7$
$8x+1 = 8$
$9x+1 = 9$

⇒ 25 add(dis)

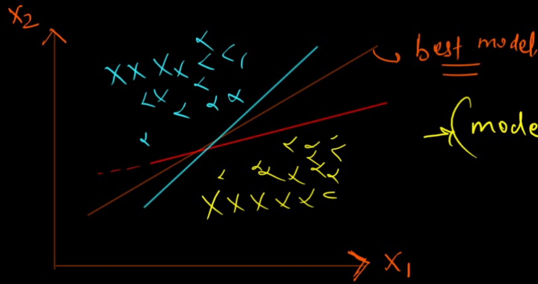
$\boxed{-5x+1} = -5$
$\boxed{-6x+1} = 6$
$-7x-1 = 7$
$-8x-1 = 8$
$-9x-1 = 9$

19

Line:-3

1. support vector → max margin (linear)
2. loss function →

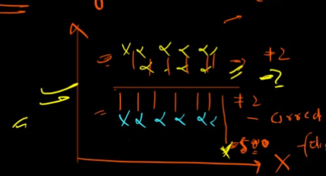
NLSD → 1. Polynomial, 2. RBF } → kernal trick



→ (model → sensitive)

$$\boxed{obj \Rightarrow \max\left(\sum_{i=1}^{n} y_i \times w^T x_i\right)}$$

Outlier :- y



model-1

$(abs) \times +ve \Rightarrow (-ve)$

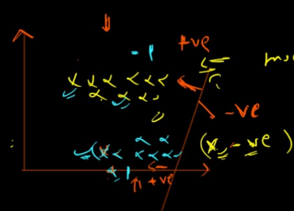
$\boxed{m_1}$

$\boxed{12}$

12

$(+2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2) + (-500)$

$24 - 500 \Rightarrow \boxed{-476} \Leftarrow$ max.



model-2

100 0 80

$\Rightarrow (+2 +3 +4 +5 +6 + 7) -2 -3 -4 -5 -6 -7$

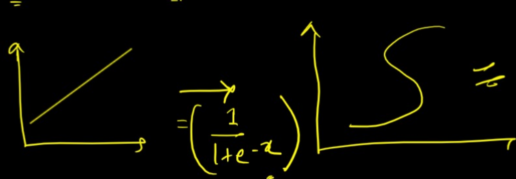
$(+500)$

$\boxed{500}$

$\boxed{m_2 \to 7}$

7

Sigmoid :- Squash



$mx \Rightarrow \frac{1}{1+e^{-x}} = \frac{1}{1+e^{-mx}}$

m

$\boxed{0-1}$

(logit)

(logloss)

so

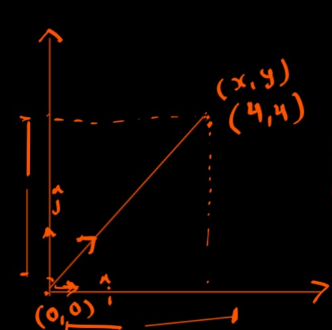
- Model 1 → $\boxed{\text{correct identify}}$ ⇒ (40)
  - $\boxed{\text{obj sum} \Rightarrow 200}$
- Model 2 → Correct identification ⇒ (23)
  - $\boxed{\text{objsum} \Rightarrow 600}$
  - sigmoid

$(2x + 2) = ✓$

$(2x + 3) = ✓$

$(2x\ 500) = ✓$

(outliers)



vector

$\boxed{\text{Vector} = \text{magnitude} + \text{direction}}$

$x\hat{i} + y\hat{j} = 0$

$ax + by + c = 0$

general eq

0

$\sqrt{(4-0)^2 + (4-0)^2}$

$\sqrt{16+16} = \boxed{\sqrt{32}} + 0 \cdot 1 x$

distance

$\boxed{ax + by = 0}$

$\boxed{a x\hat{i} + b y\hat{j} = 0}$

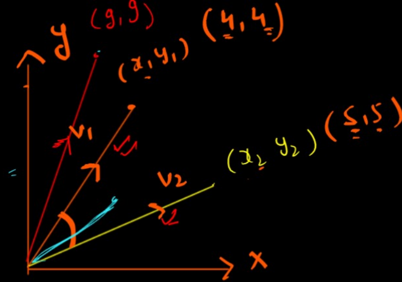
$= \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ or $[4, 4]$

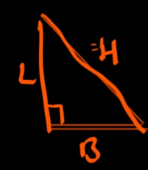
$[x_1 y_1]$

$= [x_2 y_2]$

$= (V_1 + V_2)$

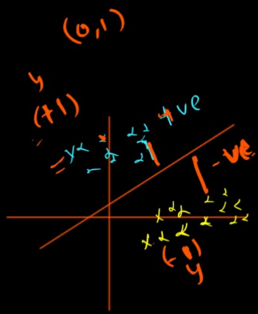
$(V_1 - V_2)$

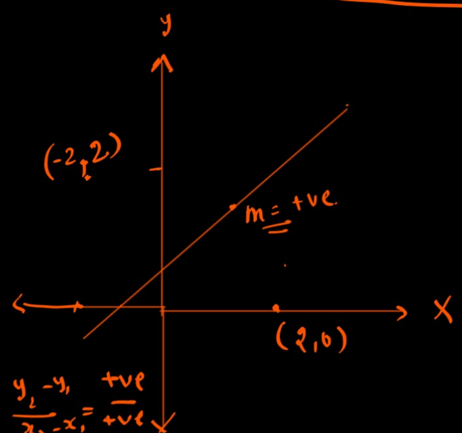


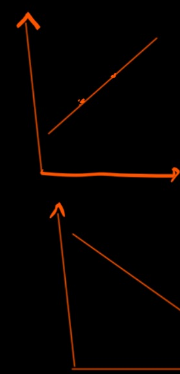


$\tan\theta = \frac{L}{B}$

$$\boxed{\theta = \tan^{-1}\frac{L}{B}}$$





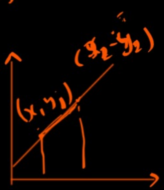


Slope = +ve

m =

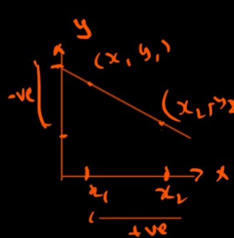
$m = \frac{y_2 - y_1}{x_2 - x_1}$

Slope = -ve.



$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{+ve}{+ve}$

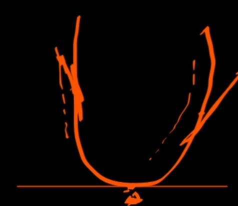
$= +ve$

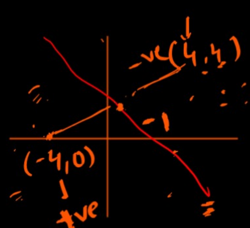


$\Rightarrow m = \frac{y_2 - y_1}{x_2 - x_1}$

$= \frac{-ve}{+ve}$

$= -ve$





$\rightarrow$ origin

$y = mx + c$

$y = w_0 x + c$

$y = w^T x + c$

$y = mx$ or $\boxed{y = w^T x}$

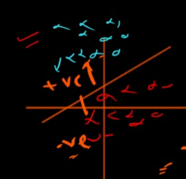
(+ve) (-ve)

$= -ve \times -ve$

$= +ve$ c

$= -ve \times +ve$

$= \boxed{-ve}$



$\rightarrow$ random

Prob

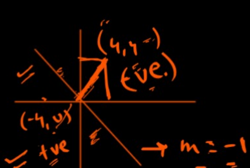
$\boxed{y = m^T x}$

$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + c$

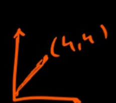
$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

$[\quad] \times 1$



$\rightarrow m = -1$



Dot

$(m_1 m_2 m_3) \cdot (x_1 x_2 x_3)$

$m_1 x_1 + m_2 x_2 + m_3 x_3$

$y = w^T x + c - 0$

$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} [4, 4]$

$= \boxed{-4}$

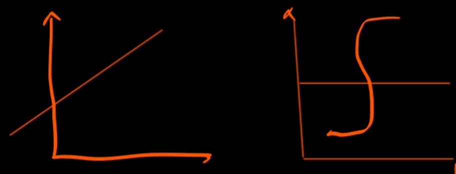
$y = mx + c$

$= -1 \times 4 + 0$

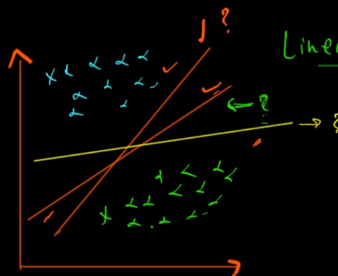
$= -4$

$\begin{bmatrix} m_1 & m_2 & m_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$m_1 x_1 + m_2 x_2 + m_3 x_3$



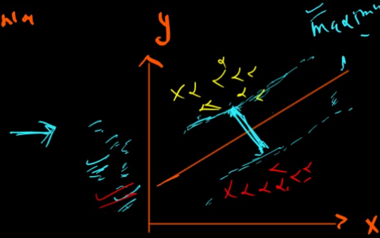
(100, 200, 300) → (sigmoid) → (logistic)
soft
high - low



Linear separable data

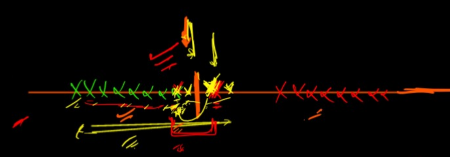
$\Rightarrow$ obj fn = $\arg\max\left[\sum_{i=1}^{n} y_i \times m x_i\right]$

↑
add of all d's
correct classification
+ve.
+ve, +ve, +ve, +ve



maximum
① cal
② cost
③ optimization
(best to best)

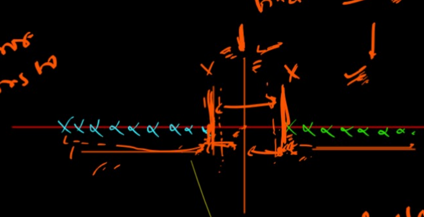
= squash
tanh softmax (sigmoid)
(logloss / obj function)



hard classifier

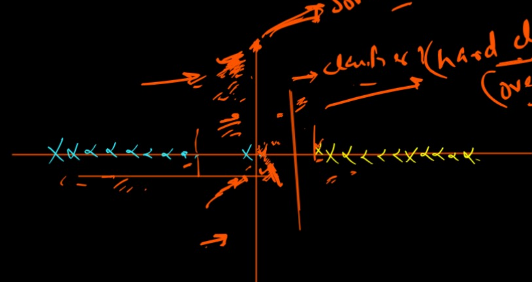
SVM → loss
≡ |
how many errors
you are going to
consider

find classifier



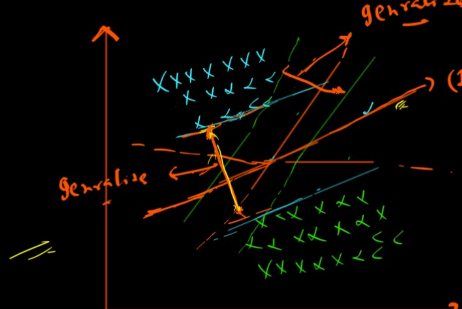
→ soft classifier
→ classifier (hard classifier)
(overfitting)

C = 0
$\boxed{C = 1}$ testing
C = error point
(cross validation)

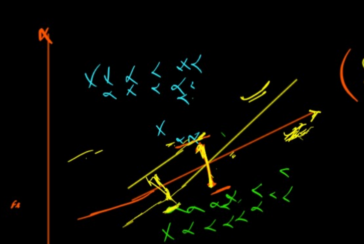


generalize
(1)
generalise

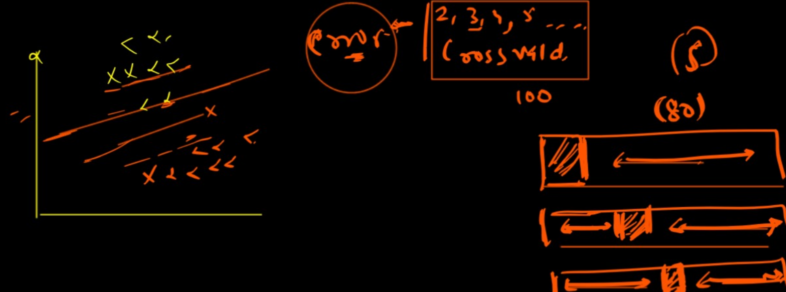
SVM → max dist should be there.
(maximum margin)
→ $\boxed{\text{Data}}$

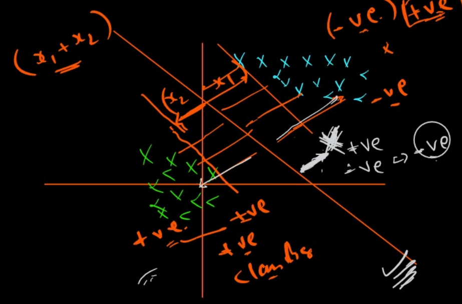


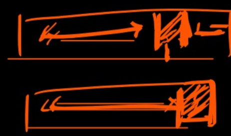
(SVM) → good accuracy
generalize model



error ← $\boxed{\text{2, 3, 4, 5 ... cross valid}}$
100
(5)
(80)







$y = mx + c$

$mx + c = +1$

$mx + c = -1$

$w^T x_1 + c = +1$

$w^T x_2 + c = -1$

$w^T x_1 - w^T x_2 = +2$

max $\boxed{\frac{w^T x + c}{\|w\|}}$ 0 1

$\boxed{w^T (x_1 - x_2) = +2}$

$\boxed{\frac{w^T (x_1 - x_2)}{\|w\|} = \frac{+2}{\|w\|}}$ max.

$\boxed{w^T x_1}$ Distance each and every

(obj.) $\boxed{y_1 \times w^T x}$

+v +ve ⇒ +ve

-ve × -ve ⇒ +ve } → correct.

error :- 2, 3, 4, 5 --- N

ϵ

$\epsilon \times$ distance

$(1 \times w^T x + c)$

$y \times$ distance

$\boxed{\text{error}}$ → generalization

+

$\frac{w^T (x_2 - x_1)}{\|w\|} = \frac{+2}{\|w\|}$

max $\left(w^T (x_2 - x_1) = +2 \right)$

distance $\frac{1}{w^T (x_2 - x_1)} = \frac{1}{2}$

(min)

$f(x) = x$

$f(1) = 1$

$f(2) = 2$

$f(3) = 3$

$f(x) = \frac{1}{x}$

$f(1) = \frac{1}{1}$

$f(2) = \frac{1}{2}$

$f(3) = \frac{1}{3}$