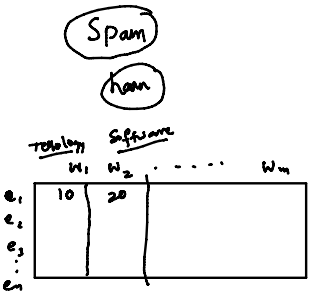
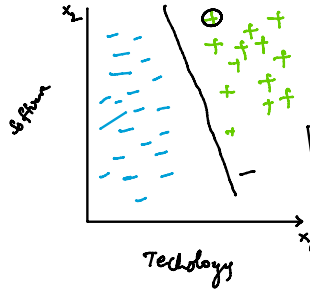
SVM-1

20 December 2022 20:42

Support Vector Machines



Logistic Regression

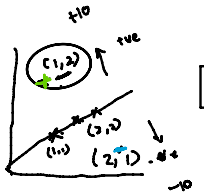


$w_0 + w_1 x_1 + w_2 x_2 = 0$ (eq 1)

$ax + by + c = 0$ ← line

$by = -c - ax$

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



$y = x$ $c = 0$ $m = 1$

$m = \frac{(y_2 - y_1)}{x_2 - x_1} = \frac{2-1}{2-1} = 1$

$y - x = 0$ ✓

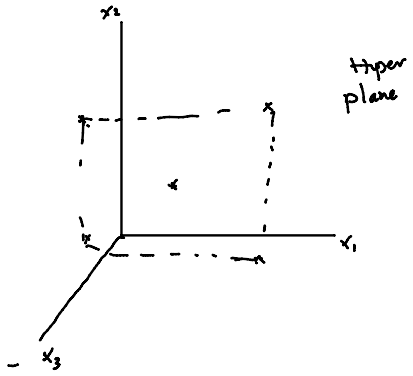
$\left(\frac{1}{1+e^{-x}}\right) \sim (0-1)$

$y - x = 0$ $(2,2)$ $(1,1)$ ✓

$(1,2)$ $y - x > 0$ $2 - 1 = 1$ ✓

$(2,1)$ $y - x < 0$ $1 - 2 = -1$

3-dimensional



hyper plane

Hyperplane

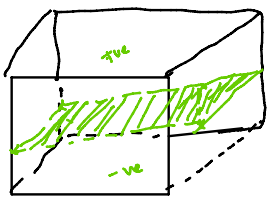
$ax + by + cz + d = 0$

$w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 = 0$ ✓ Sigmoid $p = 0.5$

$w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 > 0$ +ve $p > 0.5$

$w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 < 0$ -ve < 0.5

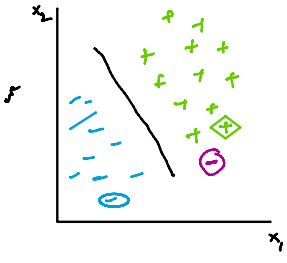
(10 input)



if $\sum_{i=1}^{p} w_i \cdot x_i + w_0 = 0$ ← hyperplane

$y \in \{0, 1\}$

$y \in \{+1, -1\}$ - pos, -1 - neg



$w_0 + w_1 x_1 = 0$

$w_0 + w_1 x_1 > 0$

$w_0 + w_1 x_1 < 0$

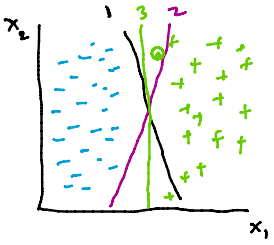
$-1 \times -ve \; (+ve)$

$y \in \{+1, -1\}$

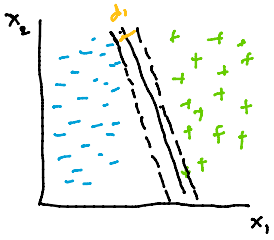
$y_i (w_0 + w_1 x_1) > 0$ (Correctly classified)

< 0 (Wrongly classified)

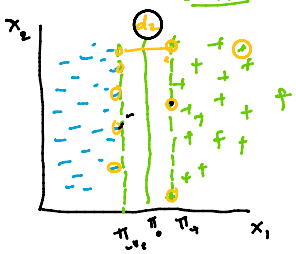
Margin classifier

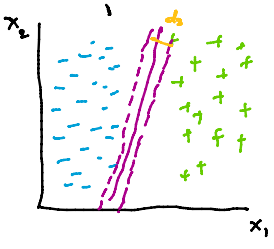


"Maximum Margin classifier"



'Support Vectors'

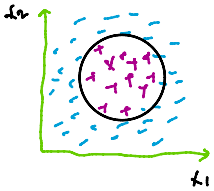




$ax + by + c_2 + d = 0$

$ax + by + c_2 + d = 0$

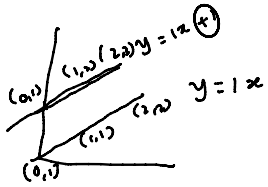
$ax + by + c_2 + d = 0$



Objective

To find a line that can have Maximum margin?

Max (d) → or Minimize $\left(\frac{1}{d}\right)$

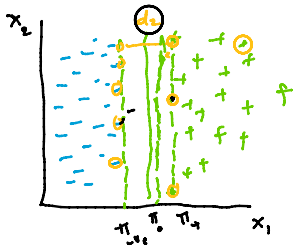


(Maximize the margin)

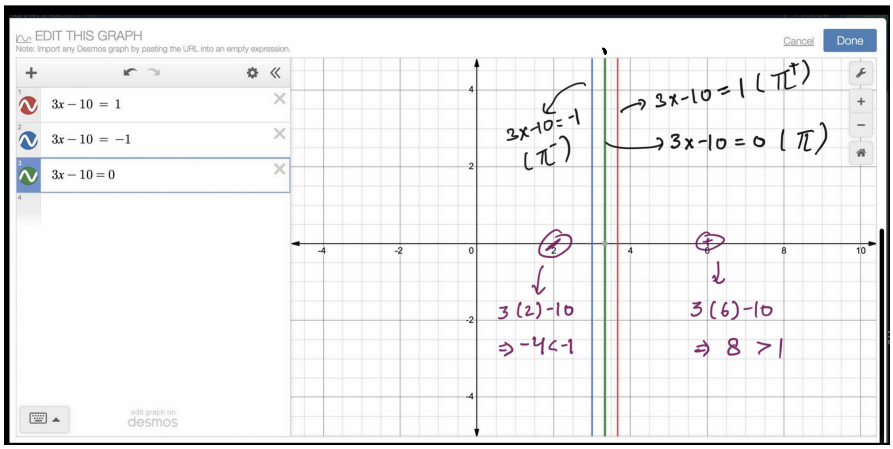
Hard Margin

Min $\left(\frac{1}{d}\right)$

Max (d)



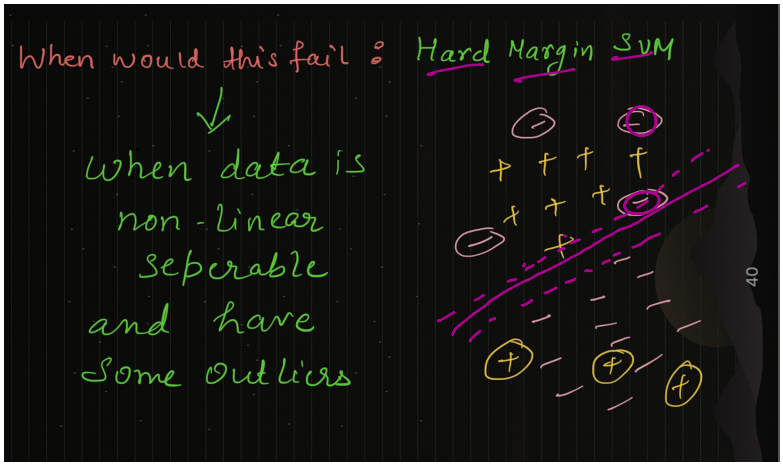
Driven by the Support vectors that are nearest and no Misclassification



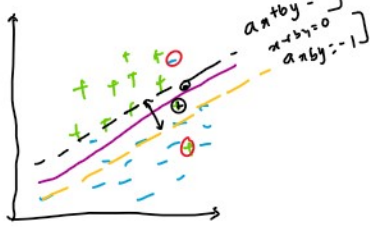
$ax+by=0$

$ax+by=1$ (K)

$ax+by=-1$ (-K)



Soft Margin classifier



$0-1$

$[0-1]$

(ϵ_i)

hard $\rightarrow$ soft margin

Max d + min (ϵ_i)

Cost = $\min\left(\frac{1}{d}\right) + \min(\epsilon_i)$ [Regularization Penalization]

how much error is acceptable?

Cost = $\min\left(\frac{1}{d}\right) + C * \min(\epsilon_i)$

C. hyper meter

$C = 10$

Cost = $\min\left(\frac{1}{d}\right) + 10 \cdot (\epsilon_i)$

$+ (20)$

$+ 10(1) = 10 \downarrow \uparrow$

ϵ_i (1, 2)

ϵ_i 1, 1

ϵ_i (1)

High value 'C' $\leftarrow$ overfitting

ϵ_i (10) looks wrong

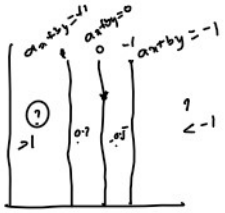
$C = 1$

$+ 10(10) \downarrow$

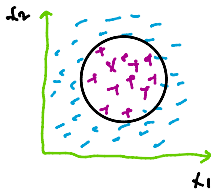
10 $\leftarrow$ (underfitting)

(1 - 10)

Break 10:29 pm

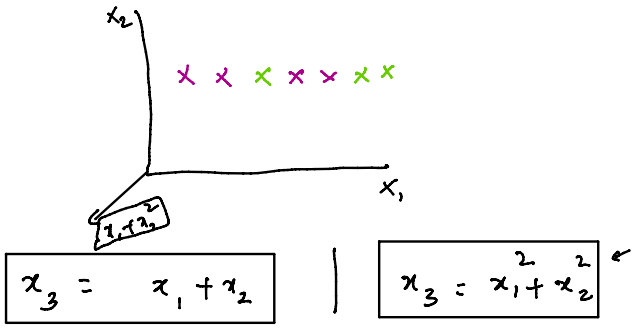


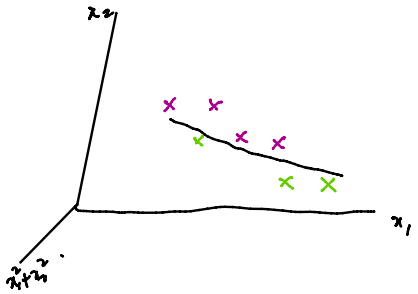
Non-linear

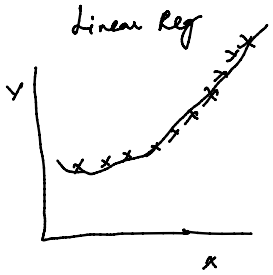


Kernels - (Mathematical functions)

Map $d \longrightarrow d'$ where $\boxed{d' >> d}$

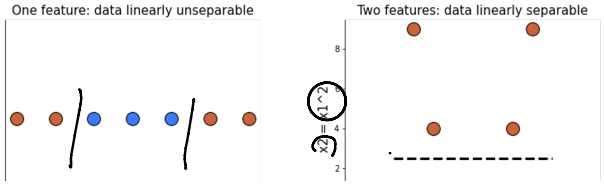


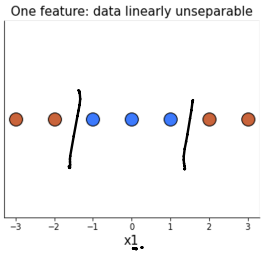


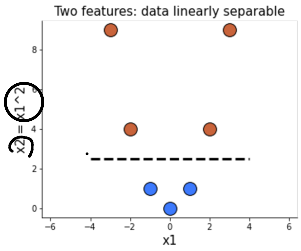


$y = w_0 + w_1 x_1 + \boxed{w_2 x_1^2}$

y	x_1	x_1^2	x_1^3





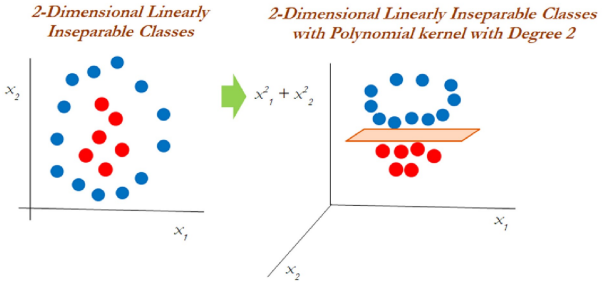


Linear →

polynomial

RBF (Radial Basis function)

Map (d) → (d) ← find a linear Separation in (higher dimensional)



"kernel trick"

Spam

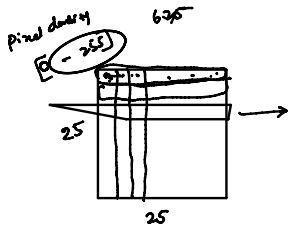
Text data

1000

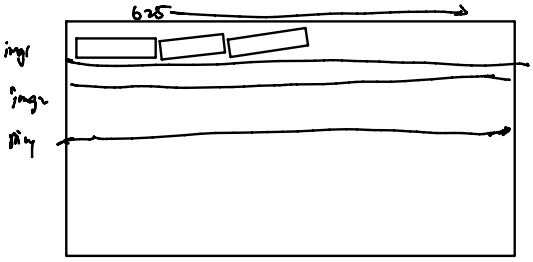
w_1 w_2 - - - - -

d_1
d_2
:
d_m

↗ high dimensional space

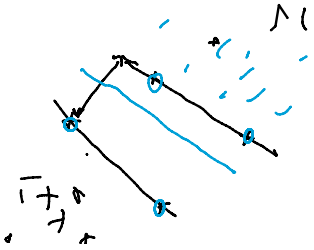


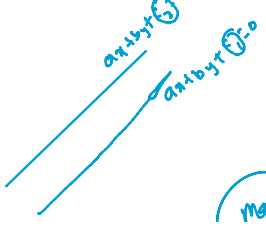
1000



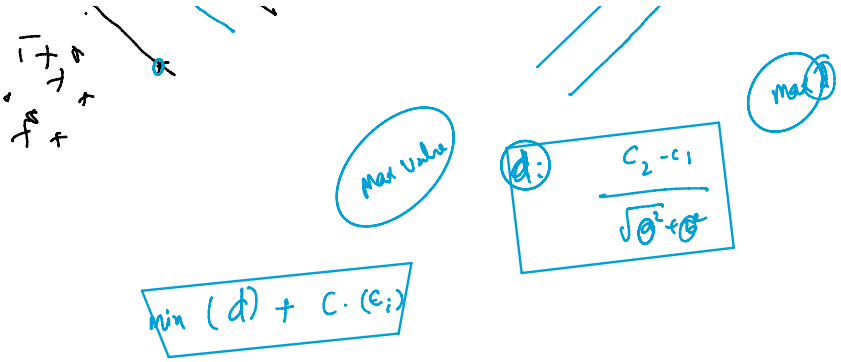
SVM

$\wedge(x_1, x_2)$ or $\wedge(x_1, x_2 - - x_{1000})$





max


max
max value
d_i: $\frac{c_2 - c_1}{\sqrt{\theta^2 + \theta}}$
min $(d) + C \cdot (\epsilon_i)$