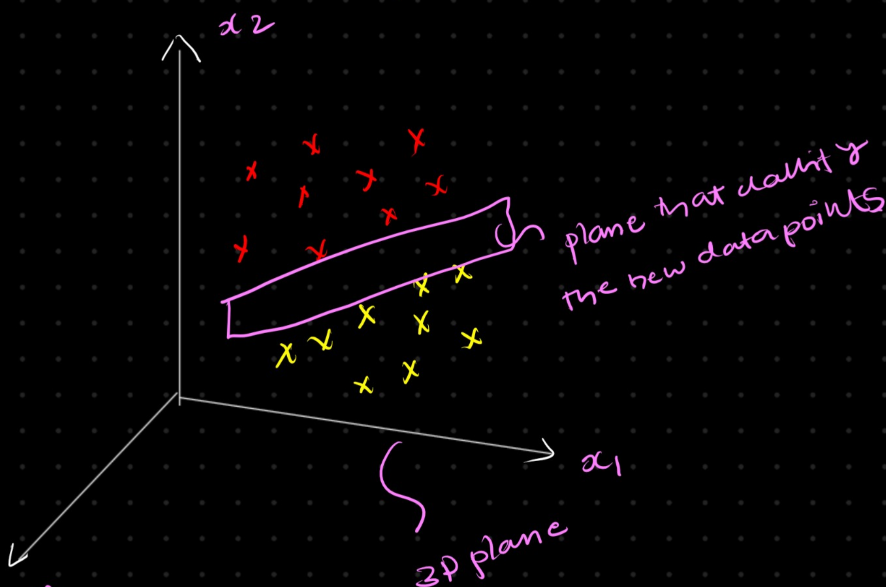
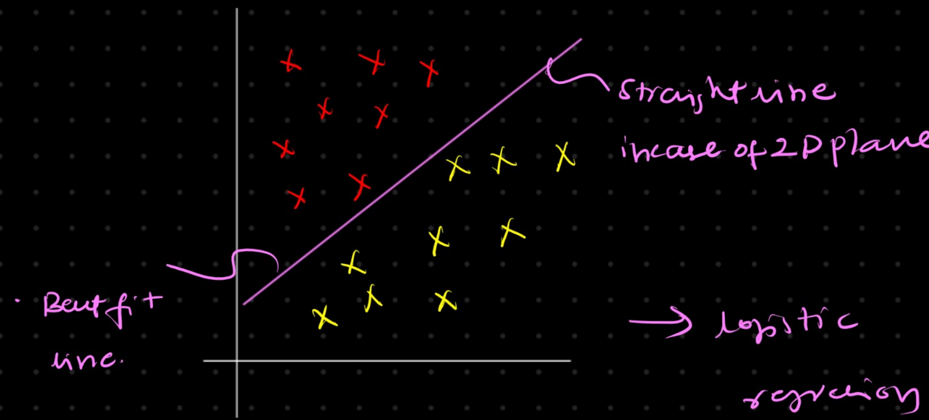
* Support vector machine

1. Support vector classifier.
2. Support vector regression.

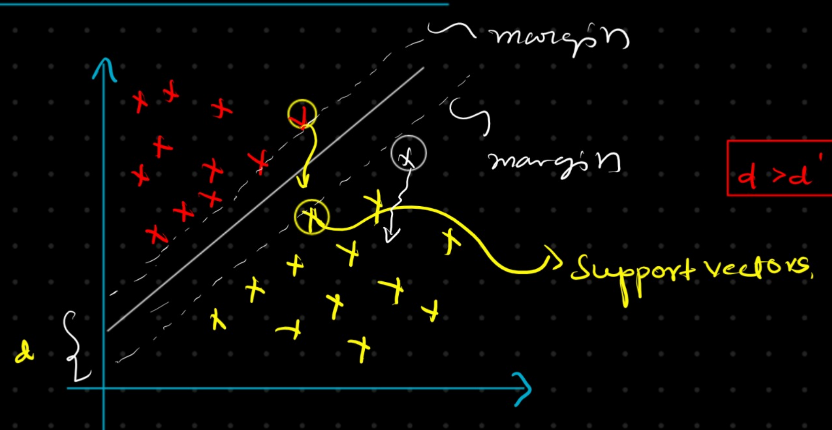


logistic regression

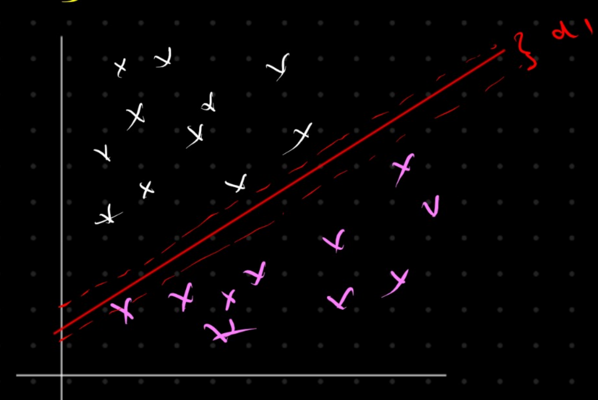


→ logistic regression

1. Support vector machine (geometrical intuitions)

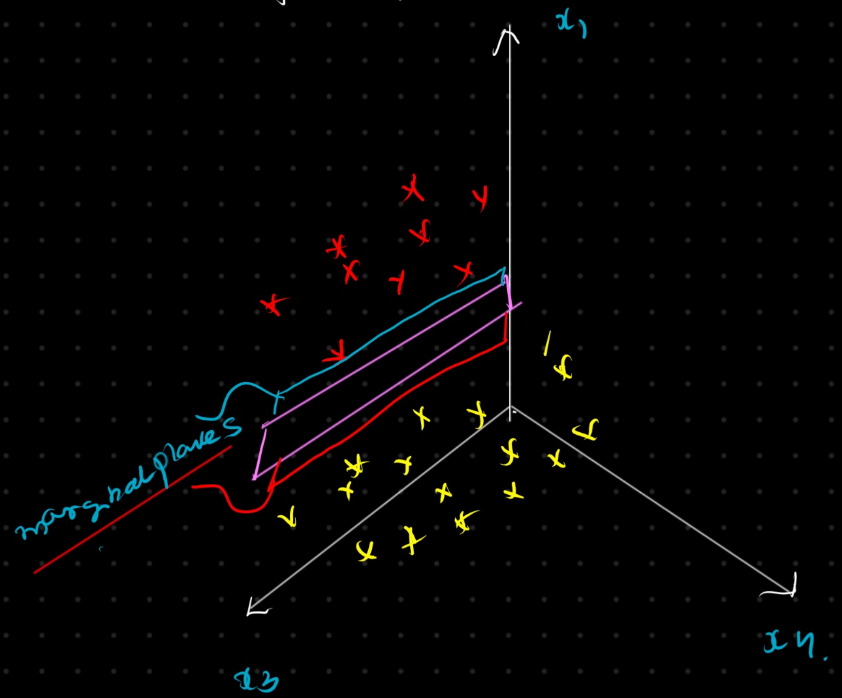


$\boxed{d > d'}$



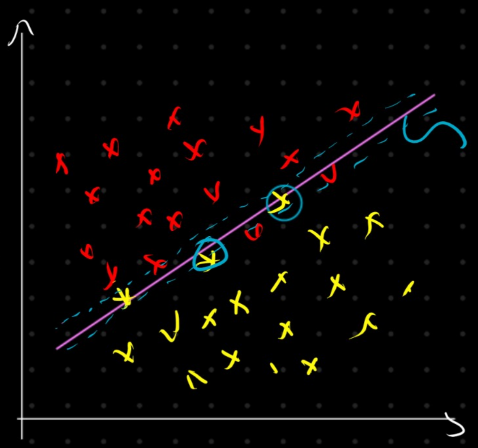
* The main aim is to maximize the marginals.

in the case of 3D-plane

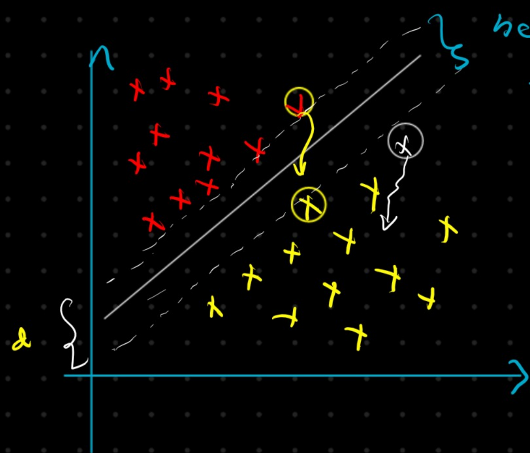


Soft margin & hard margin

* In real time the datapoints will be overlaped and we cant actually classify the datapoints

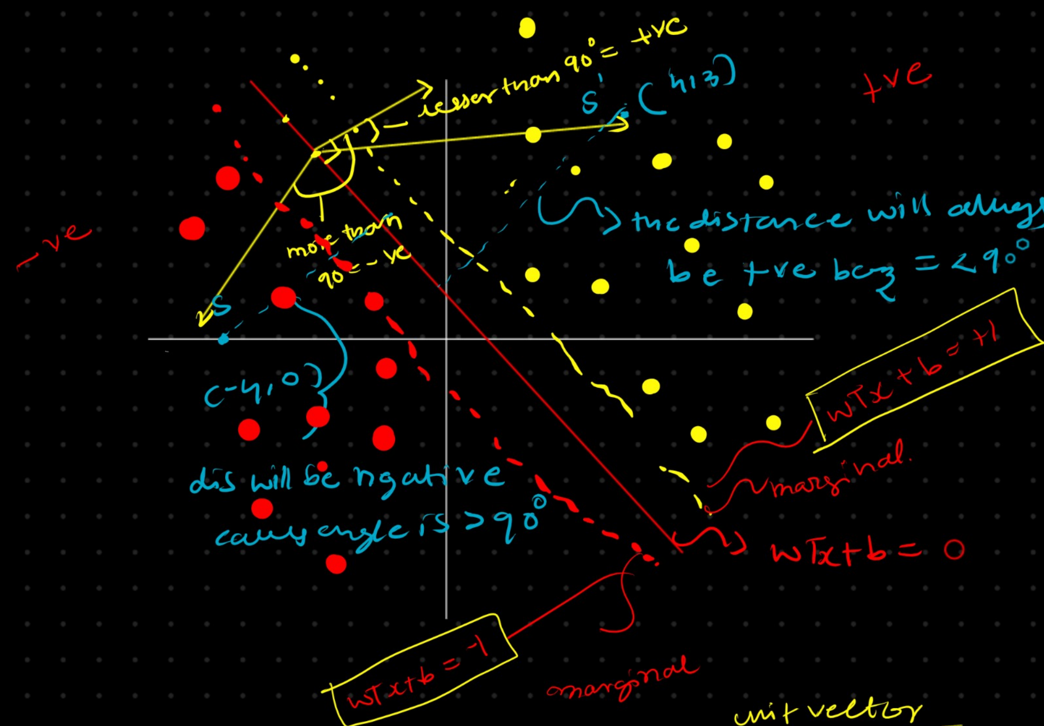


there are Errors and we can't actually classify the datapoints this is called Soft margin



here we can clearly classify the datapoints this called hard margin.

Support vector machine math intuition



$ax+by+c=0$

$w_1x_1+w_2x_2+b=0$

$\boxed{w^Tx+b=0}$

$\Downarrow$ — when our best fit line passes through origin $b=0$

$\boxed{w^Tx=0}$

unit vector

unit vector { where the magnitude of the vector is '1' }

$w^Tx_1+\cancel{b}=+1$

$w^Tx_2+\cancel{b}=-1$

$(-)\quad(-)\quad(+)$

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

cost function = $\frac{2}{||w||}$ = Distance b/w marginal plane

Constraint such that $y_i \begin{cases} +1 & w^Tx+b \ge 1 \\ -1 & w^Tx+b \le -1 \end{cases}$

how we gonna get it → by dividing W^T with magnitude of the w

it is always a good idea to come up with unit vector Because all the points will be normalized b/w 0 and 1

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

RHS

this will be our cost function

For all correct points

constraint → $\boxed{y_i * (w^Tx+b) \ge 1}$

+ve => +1 * +1 = +ve

-ve => -1 * -1 = +ve

Const_function = we have to change w, b to maximize the distance => $\frac{2}{||w||}$ => distance b/w marginal plane.

$$= \boxed{\min_{(w,b)} \frac{||w||}{2}}$$

i can write like this.

cost_function of svm (SVC)

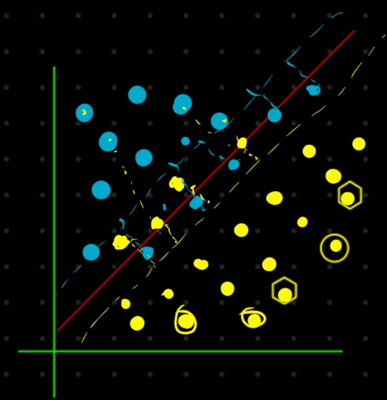
$$\min_{w,b} \frac{||w||}{2} + \boxed{C_i \sum_{i=0}^{n} \xi_i}$$

=> hinge loss.

hyper parameter

this suitable for the situation where overlapping less

↳ But these type of situation will be less in real time.



$$\text{Cost function} = \min_{(w,b)} \frac{\|w\|}{2} + c_i \sum_{i=1}^{n} \xi_i$$

hinge loss.

summation of the distance of the incorrect distance from the marginal plane

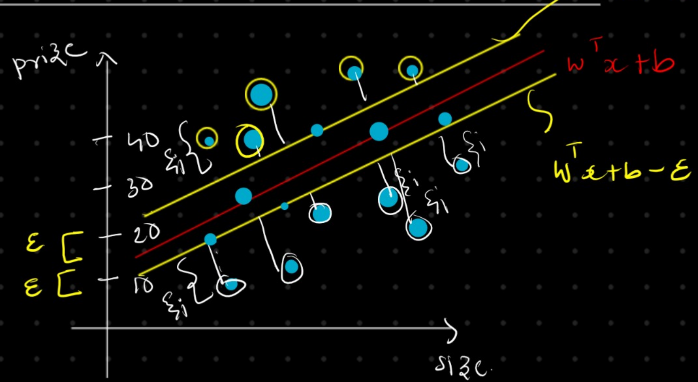
how many points or misclassified points we should allow

soft margin //

hyper parameter => c_i => tells us How many misclassified data points are allowed

ξ_i => tells us how much distance is allowed from the marginal plane of misclassified datapoint.

Support vector Regression



ϵ : marginal error //

* in regression the data points should between the marginal Error, then we can tell that our model is good.
* the Error between the predicted & actual value should be so minimal.

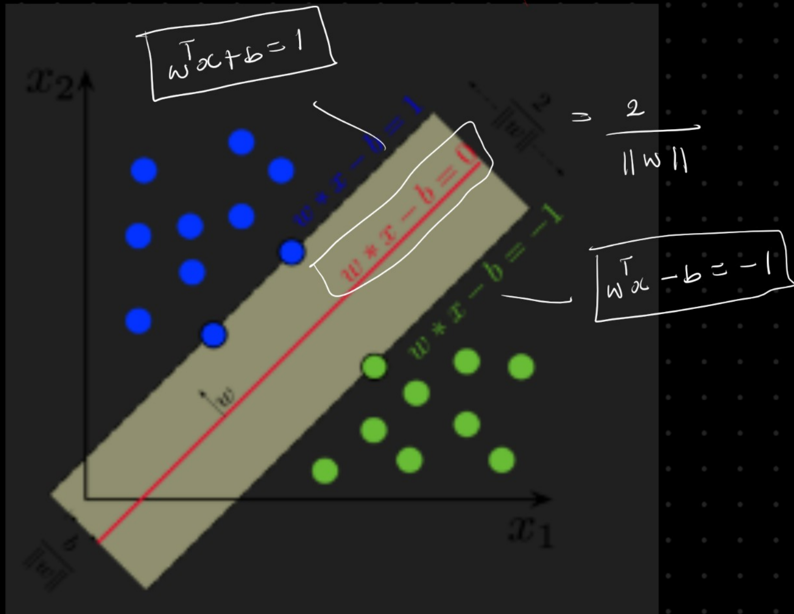
Constraints :

$$\boxed{|y_i - w_i x_i| \le \epsilon + \xi_i}$$

* cost function

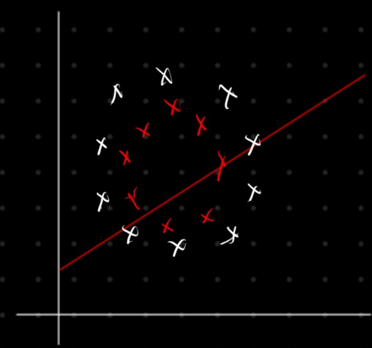
$$\boxed{\min_{w,b} \frac{\|w\|}{2} + C \sum_{i=0}^{n} \xi_i}$$

Sumation of the Error b/w the margin line and the datapoints.



* whenever our data is not linearly separable

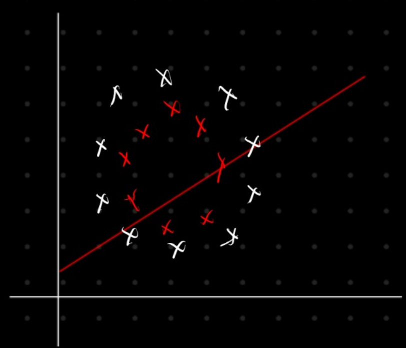
0



if my data is distributed like this my accuracy will be low.

we use SVM kernals in this case.

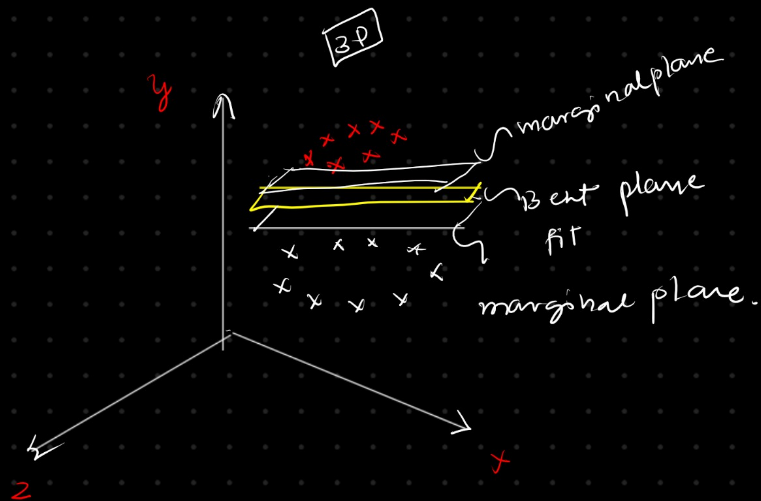
Svm kenals



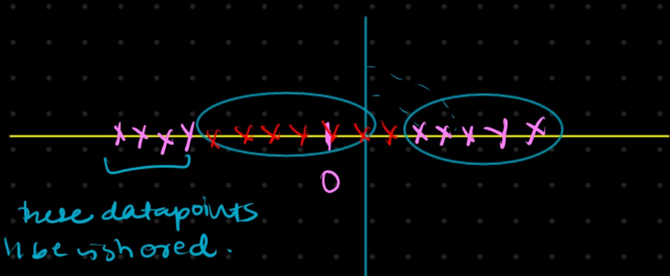
=> transformations =>

$\|$

$\boxed{\text{mathematical formula}}$



Dataset : 1d



• these datapoints will be ignored.

=>

transformation