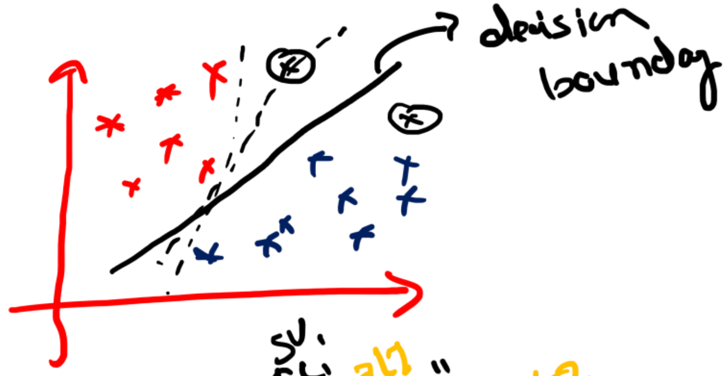
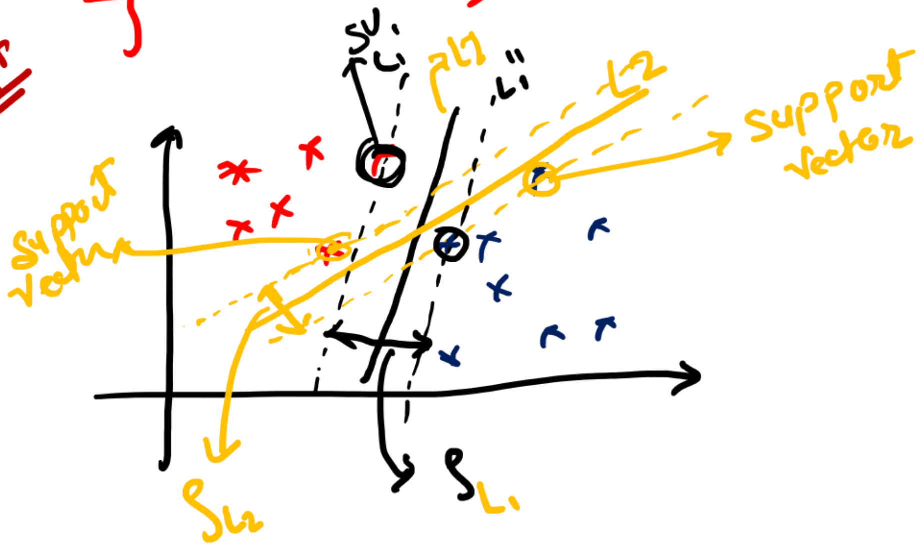
Support Vector Machine

→ Classification
→ Regression

Piyush Wairale



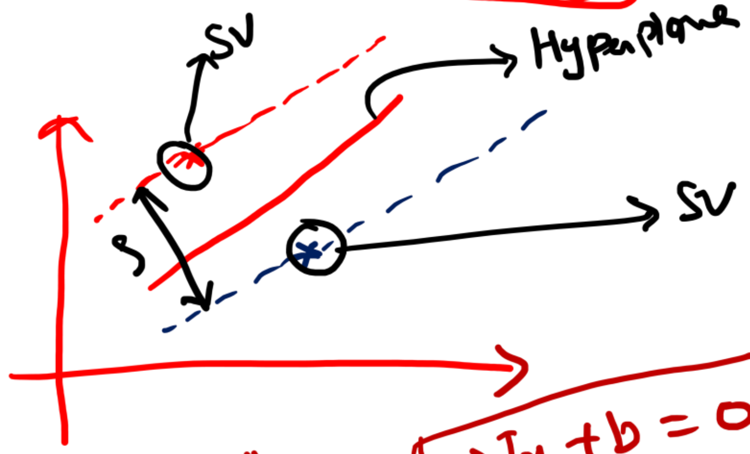
concept



* margin

distance betn L'_1 & L''_2 is margin

$$\boxed{\max(S_{L_1}, S_{L_2})}$$



$y = mx + c$

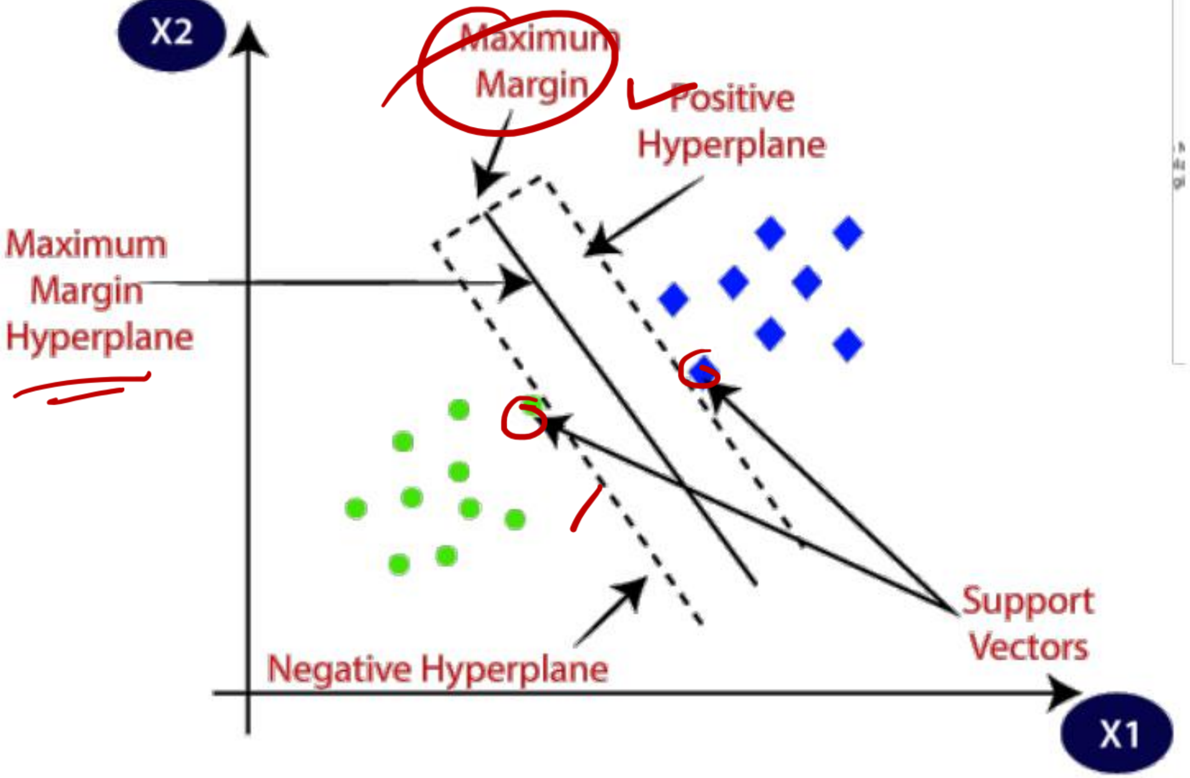
$\Downarrow$

$$\boxed{y = w^T x + b}$$

① for hyperplane, $\boxed{w^T x + b = 0}$

② SV line$_{(red)}$ = $w^T x + b = -1$

③ SV line$_{(blue)}$ = $\boxed{w^T x + b = 1}$


X2
Maximum
Margin
Positive
Hyperplane
Maximum
Margin
Hyperplane
Support
Vectors
Negative Hyperplane
X1

Note

- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.
- SVMs pick best separating hyperplane according to some criterion e.g. maximum margin
- Training process is an optimisation
- Training set is effectively reduced to a relatively small number of support vectors
- SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.

  Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

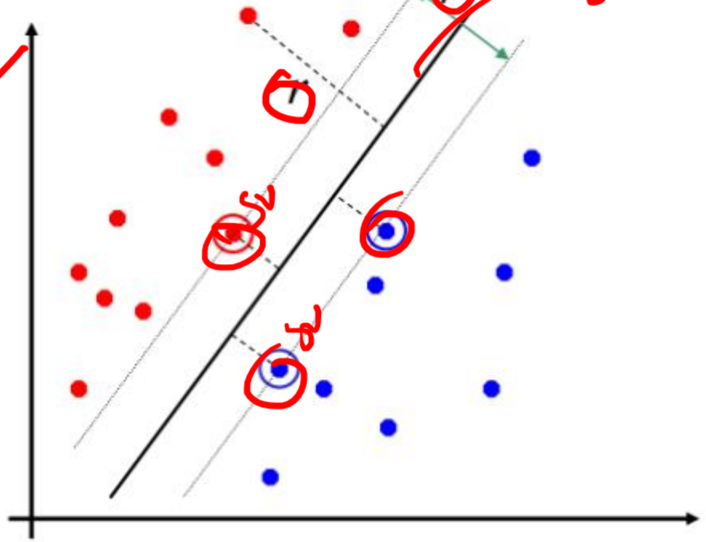
SVM Notes

- Distance from example data to the separator is $r = \frac{\mathbf{w}^T\mathbf{x} + b}{\|\mathbf{w}\|}$
- Data closest to the hyperplane are ***support vectors***.
- ***Margin*** ρ of the separator is the width of separation between classes.

v.v.imp

margin (ρ)

$\|w\| = \sqrt{w_1^2 + w_2^2 + \dots w_n^2}$



v.imp

max margin

$$\boxed{\rho = \frac{2}{\|w\|}}$$

Q Consider a binary classification problem involving two classes, $Class\_A$ and $Class\_B$. We have a dataset with the following feature vectors and class labels:

Feature Vector	Class Label
$\mathbf{x}_1 = [3, 4]$	Class_A
$\mathbf{x}_2 = [1, 2]$	Class_A
$\mathbf{x}_3 = [4, 2]$	Class_A
$\mathbf{x}_4 = [5, 6]$	Class_B
$\mathbf{x}_5 = [2, 1]$	Class_B
$\mathbf{x}_6 = [6, 5]$	Class_B

$$\boxed{\delta = \frac{2}{\|w\|}}$$

$$\sqrt{w_1^2 + w_2^2}$$

We want to train a Support Vector Machine (SVM) to classify these samples. Assume that the SVM finds the optimal hyperplane $\boxed{w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0}$ for the separation.

Solving the SVM optimization we get the value of $w_1=1$, $w_2=1$, and $b=-5$. Then find out the margin of the hyperplane

For SVM, $\mathbf{w} = [w_1, w_2]$, where w_1 and w_2 are the weights associated with the features x_1 and x_2 respectively.

So, for the optimal hyperplane $w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0$, $||\mathbf{w}||$ in this case would be:

$$\boxed{||\mathbf{w}|| = \sqrt{w_1^2 + w_2^2}}$$

In the numerical example, if we found $w_1 = 1$ and $w_2 = 1$, then:

$$||\mathbf{w}|| = \sqrt{1^2 + 1^2} = \sqrt{2}$$

The margin is given by $\frac{2}{||\mathbf{w}||} = \frac{2}{\sqrt{2}}$. $= 1.414$