SVM

25 May 2023 22:07

SVM → Support Vector Machines

* It is one of the most popular Supervised learning algorithms.

* SVM is used to solve both Regression and classification.

→ SVR → Support vector regression used to solve data which has continuous target.

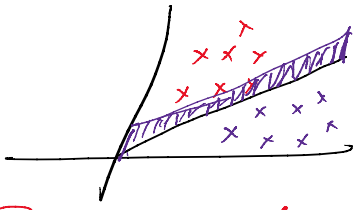
→ SVC → Support vector classifier used to solve data which has categorical target

Goal of SVM

The Goal of SVM is to create best fit line / hyperplane / decision boundary that segregates data into different classes.

In simple words "It has to find a line which seperates data into classes".

The line or decision boundary is called as "Hyperplane".



How SVM works?

→ Let us say that we have diabetes data which has two classes [diabetic, Nondiabetic]

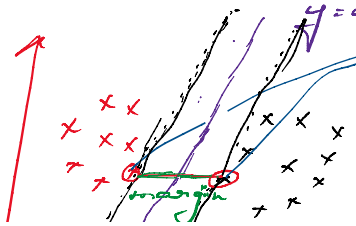
* Now the task is to fit a line / hyperplane which seperates data

* out of all the possible lines we have to select the more generalized one.

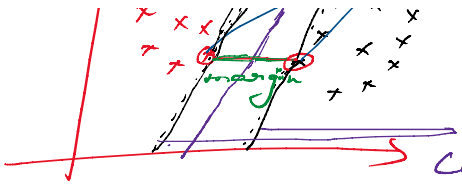
* Challenging task in SVM is selecting best line.

How to Select Best fit line using SVM.

$y = w^T x + b$

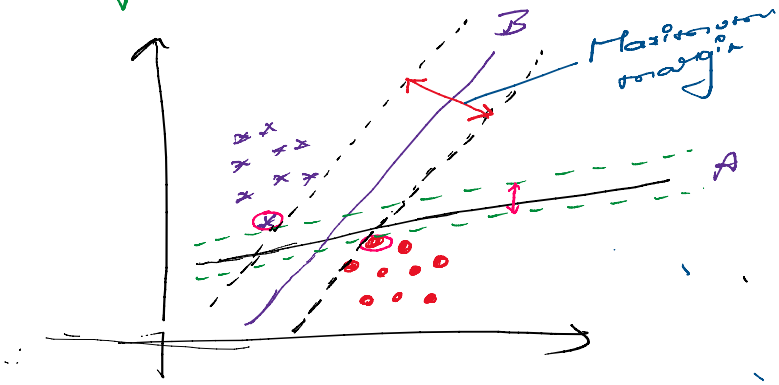


* Find the points near to line
* Closest points to the hyperplane are called as support vectors
* The line drawn at nearest points / support vectors is called as Marginal line



* The line drawn at nearest point to support vector is called as Marginal line
* Compute distance between hyperplane and support vectors. The distance is called as Margin / Marginal distance

* Goal of SVM is to maximize the margin
* Select the hyperplane which gives the maximum margin.
* Select the hyperplane what has the highest distance between hyperplane and support vector.



* Clearly B has highest margin compared to A.
So B is best hyperplane.

① Support vectors
Support vectors are nearest points to the hyperplane or points which lie on marginal line

② Marginal line
Marginal lines are parallel lines hyperplane which are drawn close to the support vectors. / parallel line to hyperplane which passes through support vector.

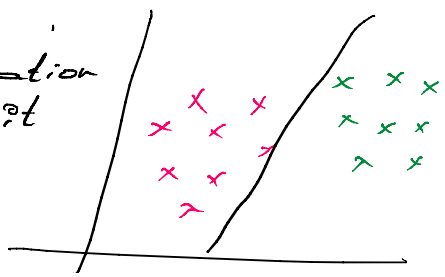
③ Margin
The distance between hyperplane and support vectors are called as Margin.

Goal of SVM
* Maximize the margin, $d = \frac{2}{\|w\|}$

Imp
SVM ignores outliers (*)

Linear Data
If there is a clear separation between two classes, we call it as Linear.

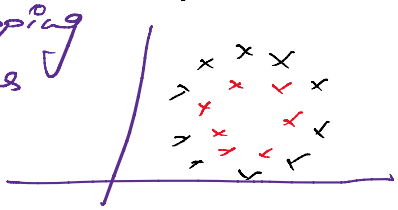
x for Linear Data



as Linear.

Non Linear Data

If there is a overlapping of classes, we call it as Non linear

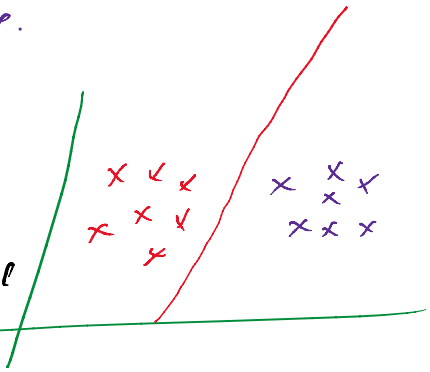


Types of SVM

1. Linear SVM
2. Non Linear SVM.

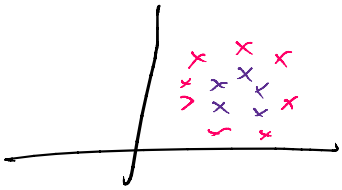
Linear SVM

If we can find a hyperplane which seperates two classes in a better way, we call that method as Linear SVM.



What if Data is Non Linear?

→ Non-Linear SVM is used to overcome non-linearity.



Kernels

Kernels are mathematical functions that transforms converting lower dimension data into higher dimensions.

Non Linear

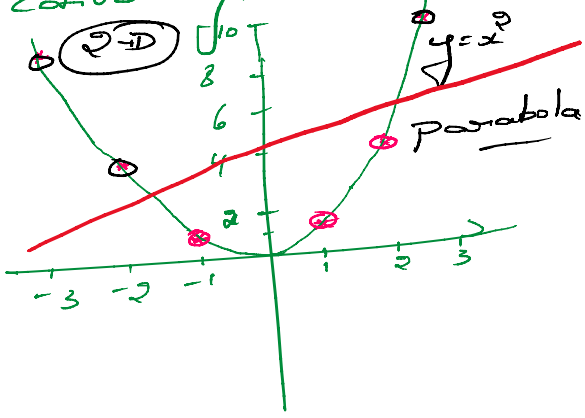
1-D

x	
-1	-ve
-2	+ve
-3	+ve
1	-ve
2	-ve
3	+ve

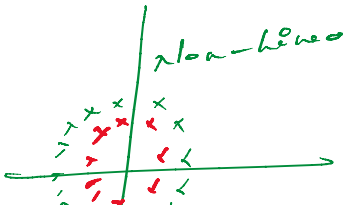
St-Line

2-D

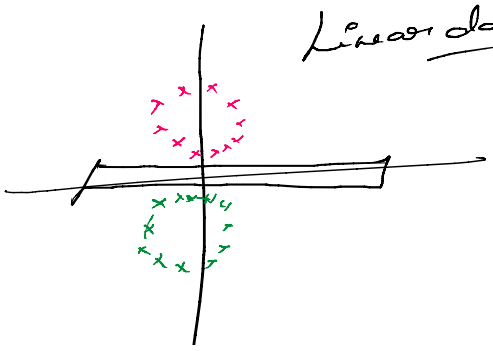
x	$y=x^2$
-1	1
-2	4
-3	9
1	1
2	4
3	9

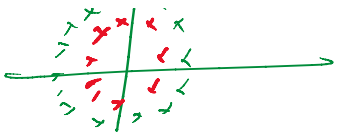


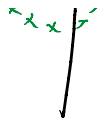
Case 2



$z = x^2 + y^2$ → circle





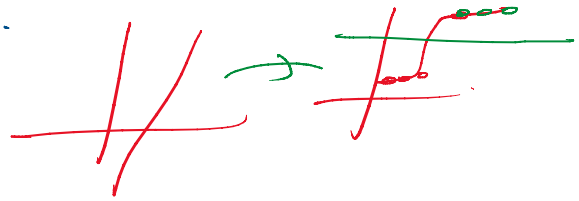


Types of Kernels

① Polynomial

Eg:- $y = x^2$, $y = x^3$, $z = x^2 + y^2$

Eqns which has powers of x greater than 1; such equations are called as polynomial Eqns.



② Sigmoid

$$h_\theta(x) = \frac{1}{1+e^{-z}}$$

③ Exponential

Eg: $y = e^x$, $y = e^{3x}$.

④ RBF → Radial Basis function.

$$F(x, x_j) = \exp(-\gamma |x - x_j|^2)$$

γ = Hyperparameter which tells us about how squared distance between two observations are varying.

$\gamma \rightarrow$ Gamma

RBF is the popular kernel often used in SVM to convert non Linear data into Linear data.