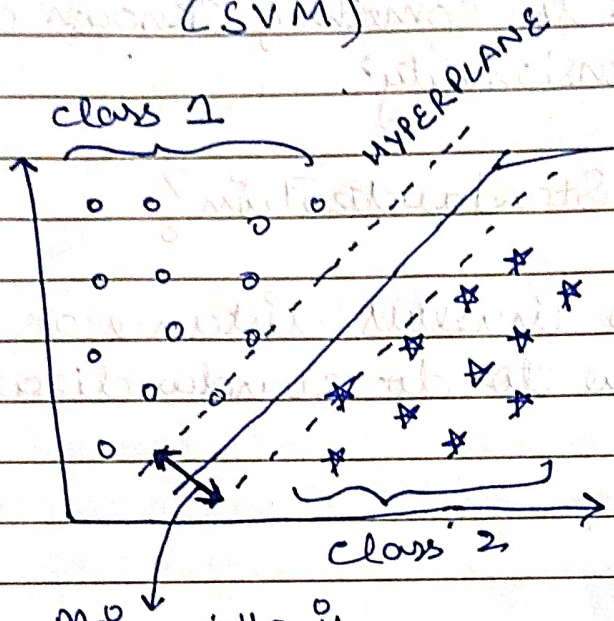
Date ___/___/_____

SUPPORT VECTOR MACHINE

Classification (SVM)

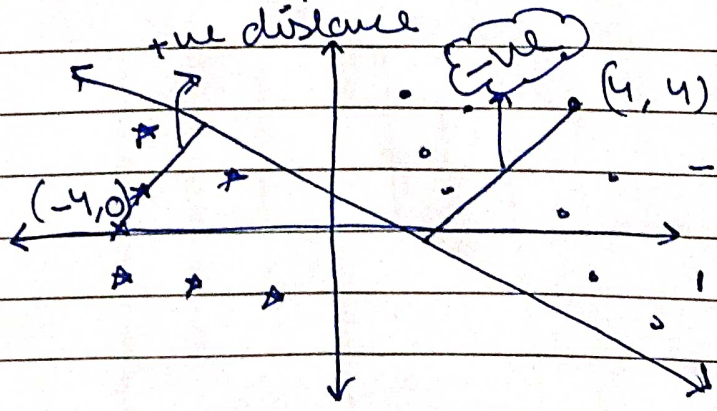
Regression (SVR)

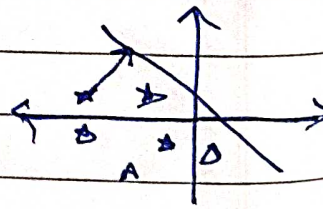


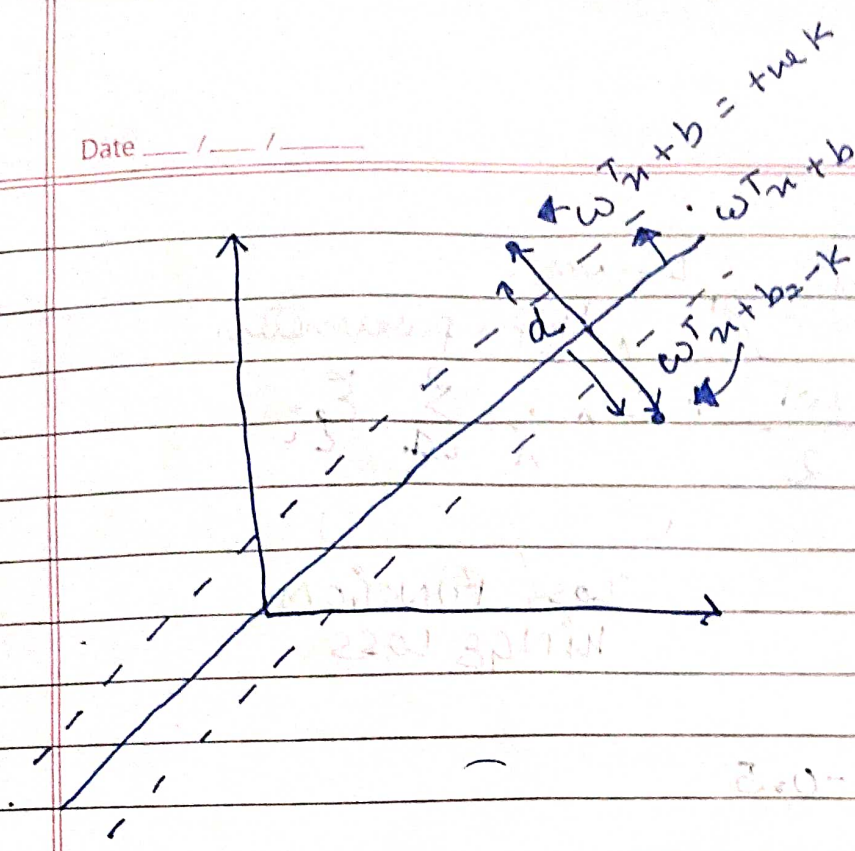
① Hard Margin SVM
② Soft Margin SVM

⇒ If your data is linearly Separable [If the data can be classified into two classes using a st. line] SVM works perfectly fine

Distance from a pt to plane = $w^T x$







Aim is to calculate the distance 'd' and maximize.

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

$$w^T(x_1 - x_2) = 2K \quad \text{—(i)}$$

Dividing (i) by $|w|$,

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

Such that

$y_i \times w^T x_i \geq 0$.

⇓

Can be applicable in Hard Margin or when data is linearly separable.

Let $K = 1$.

$$\max_{w,b} = \frac{2}{|w|} \Longrightarrow \text{OPTIMIZATION FUNCTION}$$

(⇒) Change w or b till the value is maximized.

Saathi

Date ___/___/___

Soft Margin

L1 norm.

Hyperparameter.

- $\underset{w,b}{\text{argmin}} \quad \frac{|w|}{2} + C \times \frac{1}{n} \sum_{i=1}^{n} \xi_i$

ξ_i → LOSS FUNCTION. HINGE LOSS

Q What is ξ

Pt one

Eg: $y_i \times w^T x_i = -0.5$

or

$y_i \times w^T x_i = 1 - (1.5)$

Pt two

$y_i \times w^T x = 1 - (2.5)$

$$\boxed{y_i \times w^T x_i = 1 - \xi_i}$$

ξ is the distance of the misclassified point from its actual plane.

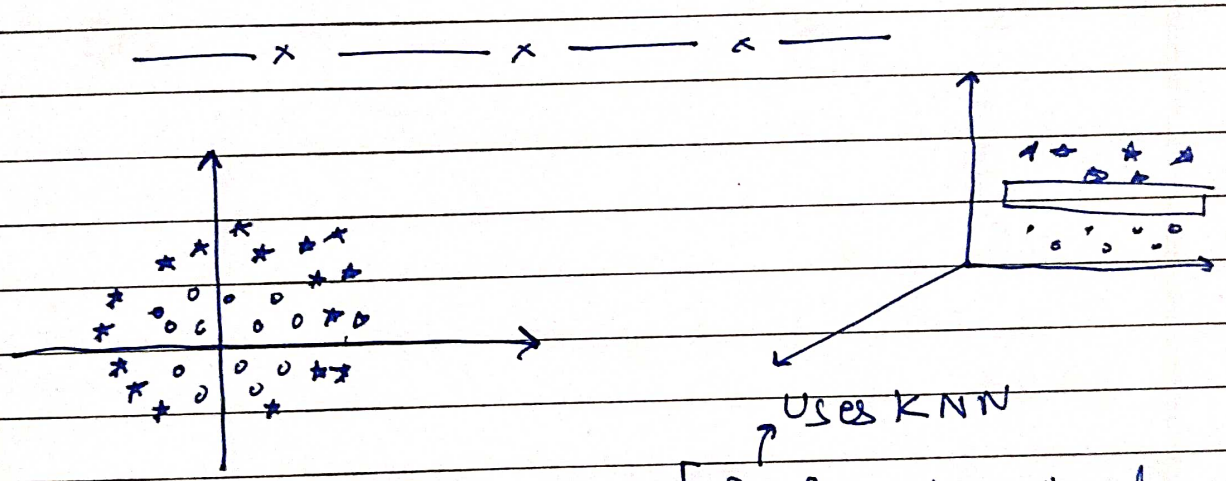
SVR (Support vector Regression)

Minimize.

$$\text{Argmin}\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}|\epsilon_i|$$

Constraints:

$$|y_i - w_i x_i| \leq \varepsilon + |\epsilon_i|$$



We use kernals here [Rbf, polynomial,
↓
Default