Data Science & AI

Machine Learning

Unsupervised Learning

Lecture No.- 01

By- Krish Naik Sir

Recap of Previous Lecture

Topic

Topic

Topic

Topic

Topic

Randon Forest Classifier
KNN
K-means

Topics to be Covered

Topic SVM

Topic PCA

Topic

Topic

Topic

Machine Learning

① SVM (SVC And SVR) ✓.

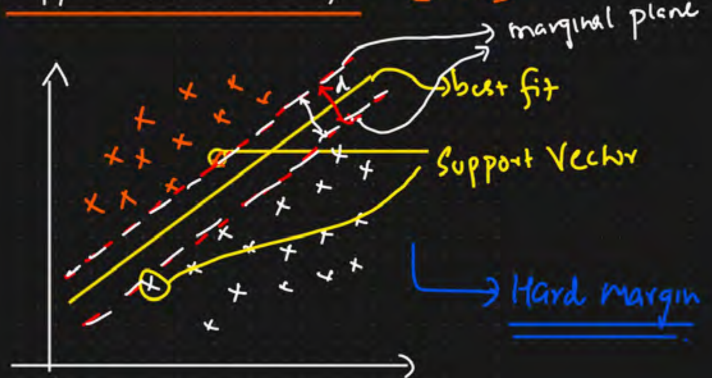
② PCA [Curse of dimensionality, Dimensionality Reduction] ✓.

① Support Vector Machines

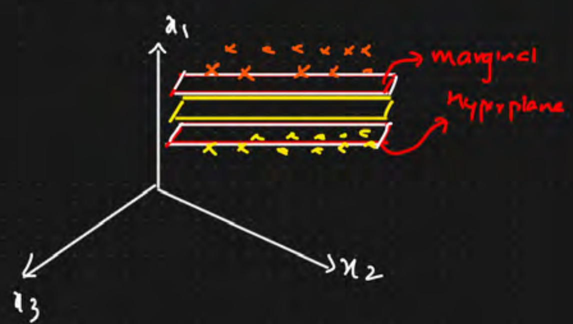
① Support Vector Classifier [SVC] ⇒ Classification.

② Support Vector Regressor [SVR] ⇒ Regression.

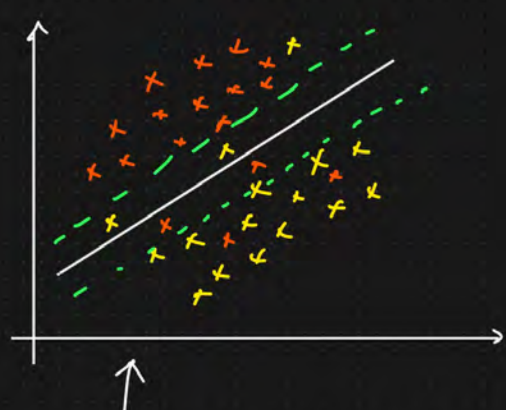
① Support Vector Classifier [SVC]



d = marginal plane distance

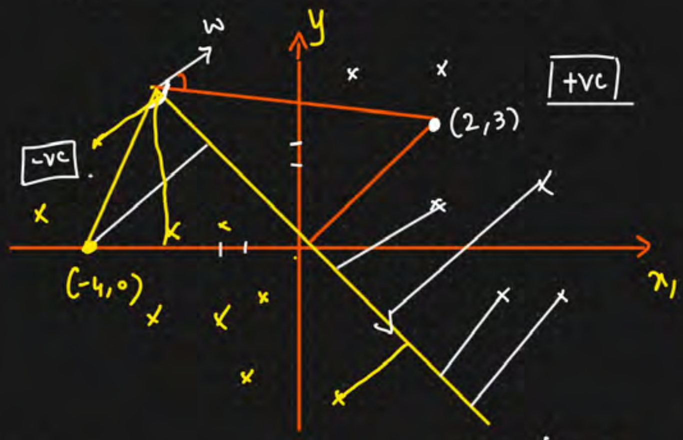


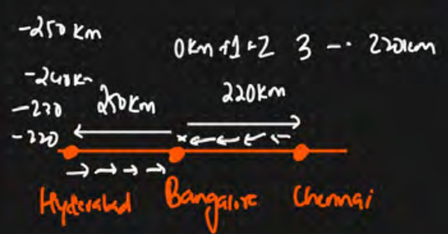
Soft Margin And Hard Margin In SVC



Soft Margins : Some data points are Misclassified.

⊛ Support Vector Machine Intuition





Equation of a straight line

$\boxed{y = mx + c} \iff \boxed{ax + by + c = 0}$

$h_\theta(x) = \theta_0 + \theta_1 x_1$

$by = -ax - c$

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

$\hookrightarrow m$ $\quad \downarrow c$

$y = \theta_0 + \theta_1 x_1 + \theta_2 x_3 + \theta_3 x_3$

$y = b + [w_1 x_1 + w_2 x_2 + w_3 x_3]$

$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$w^T \cdot x = [w_1 \; w_2 \; w_3] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

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

$x_1 [x, y]$

$x_2 [x, y]$.

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

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

$w^T x_1 + \cancel{b} = +1$

$\underline{w^T x_2 \cancel{+} b = -1}$
(-) (-) (+)

$\frac{w^T (x_1 - x_2)}{\|w\|} = \boxed{\frac{2}{\|w\|}}$ ⇒ Distance between marginal plane.

Cost fn

$$\underset{w,b}{\text{Maximize}} \quad \boxed{\frac{2}{\|w\|}} \Rightarrow \text{Distance between Marginal planes.}$$

Constraint such that

$$y_i \begin{cases} +1 & \text{if } w^T x + b \geq 1 \\ -1 & \text{if } w^T x + b \leq -1. \end{cases}$$

For all correct classified data points

$$\boxed{y_i * [w^T x + b] \geq 1.}$$

Modified cost fn

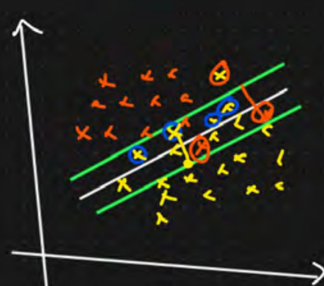
$$\underset{w,b}{\text{maximize}} \quad \frac{2}{\|w\|} \Rightarrow \boxed{\underset{w,b}{\text{Minimize}} \quad \frac{\|w\|}{2}}$$

Constraint such that

$$y_i \begin{cases} +1 & \text{if } w^T x + b \geq 1 \\ -1 & \text{if } w^T x + b \leq -1. \end{cases}$$

$\Rightarrow$ Hard Margin

Cost fn of SVC [Soft Margin]



$$\text{Cost fn} = \underset{w,b}{\text{Min}} \quad \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{n} \xi_i} \Rightarrow \text{Hinge Loss}$$

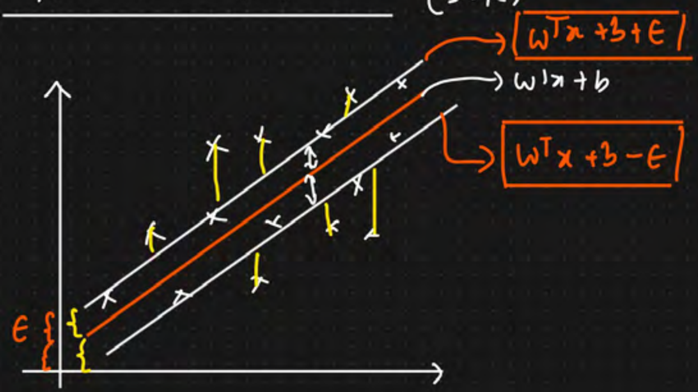
Hyperparameter = $\{ C_i \rightarrow \}$

{Summation of the distance of incorrect data points from marginal plane}

$\downarrow$

{How many datapoints we can consider for misclassification}.

Support Vector Regressor (SVR)



Cost fn

$$\min_{w,b} \frac{\|w\|}{2} + \boxed{C \sum_{i=1}^{n} \xi_i} \Rightarrow \text{Hinge Loss}$$

Constraint

$$\boxed{|y_i - w^T x_i| \le \epsilon + \xi_i} \Rightarrow \text{Minimize Error}$$

$\Uparrow$

$\epsilon \Rightarrow$ Marginal Error

$\xi_i \Rightarrow$ Error above the Margin.

Principal Component Analysis [PCA] [Dimensionality Reduction]

(i) Curse of Dimensionality

3 feature	6 features	10 features	100 features
M_1 →	M_2 →	M_3 →	M_n
Acc	Acc ↑↑	Acc ↑↑	

Reduce the Dimensions

(1) Feature Selection

⇓

Imp features

(2) Dimensionality Reduction (PCA)

⇓

Feature Extraction

⇓

PCA

↓

Eigen Values And Vectors

[500 features]

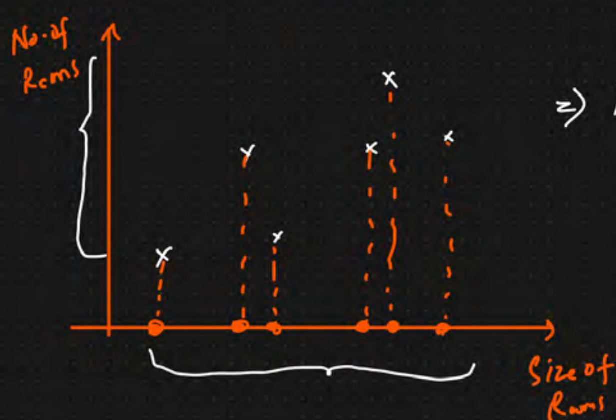
⇓

[20 features]

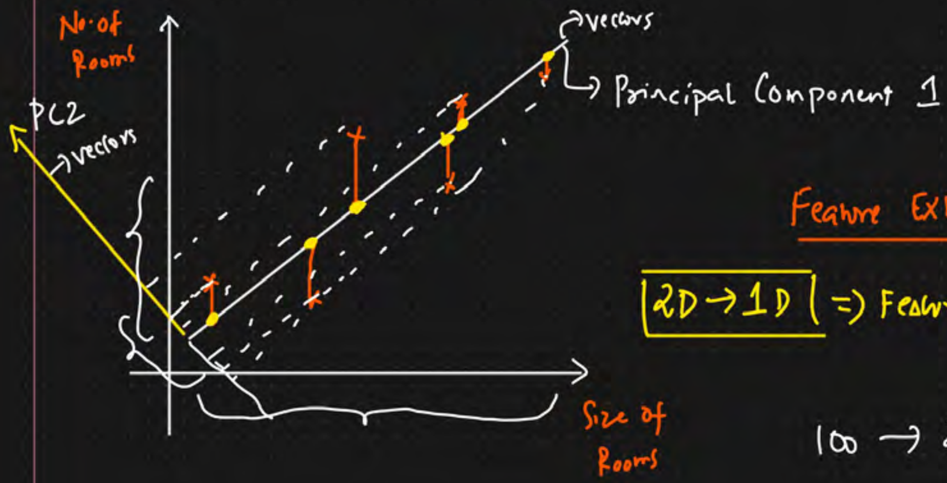
⇒ Variance Information Captured

PCA Geometric Intuition

x_1	x_2	Price
Size of Rooms	No. of Rooms	



⇒ Feature Selection



$3D \to 2D$

Feature Extraction

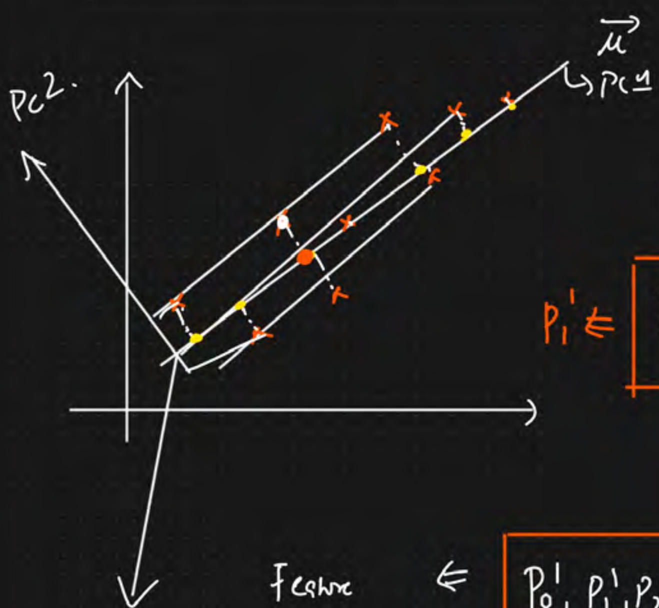
$\boxed{2D \to 1D}$ ⇒ Feature Extraction

$100 \to 2$ feature ⇒ PC1, PC2

$100 \to 3$ feature ⇒ PC1, PC2, PC3.

⇒ PC1.

Mathematical Intuition behind PCA Algorithms



① Projection

② Optimization → Max Variance

$P_1' \Leftarrow \boxed{\text{Proj}_{P_1} u = P_1 \cdot u} = \begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$

$\Leftarrow x_1 x_2 + y_1 y_2 \Rightarrow$ Scaler Value.

Feature Extracted. $\Leftarrow \boxed{P_0', P_1', P_2', P_3', P_4' \ldots P_n'}$

⇓

Scaler value

⇓

Max Variance

Q) → How to find the vectors?

Eigen Value Decomposition $\Rightarrow$ Eigen values And Eigen vector

① Covariance Matrix between Features. $Cov(f_1, f_2)$

$100 \rightarrow 5d$

$\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5$

② Eigen values and Eigen vector will be computed using the Covariance Matrix = $\boxed{Av = \lambda v}$

$A \rightarrow$ Covariance Matrix

$\lambda \rightarrow$ Eigen value

③ Eigen Vector $\longrightarrow$ Eigen Values $\longrightarrow$ Capture the Maximum Variance

$Cov(x, y)$

$$A = \begin{array}{c|cc} & x & y \\ \hline x & Var(x) & Cov(x,y) \\ y & Cov(x,y) & Var(y) \end{array}$$

$$\boxed{A \cdot v = \lambda \cdot v}$$

$\lambda_1 \rightarrow PC_1$ $\quad$ $\lambda_2 \rightarrow PC_2$.

$3D \rightarrow 2D$

$\boxed{\lambda_1, \lambda_2} \Rightarrow$ Eigen values

THANK - YOU