

	x Drug Dosage	y Efficiency
1	—	—
2 →	—	—
3 →	—	—
4 →	—	—
5 →	—	—

XGBOOST



Pip install lightgbm

$$Gain = \text{Left Sim} + \text{Right Sim} - \text{parent Sim}$$

$$Sim = \frac{\sum (\text{residuals})^2}{n}$$

no } best $\lambda \Rightarrow 0$
 \rightarrow Regularization

$$x_1 > - \quad x_2 < \quad x_3 <$$

LightGBM \rightarrow Some Approximations

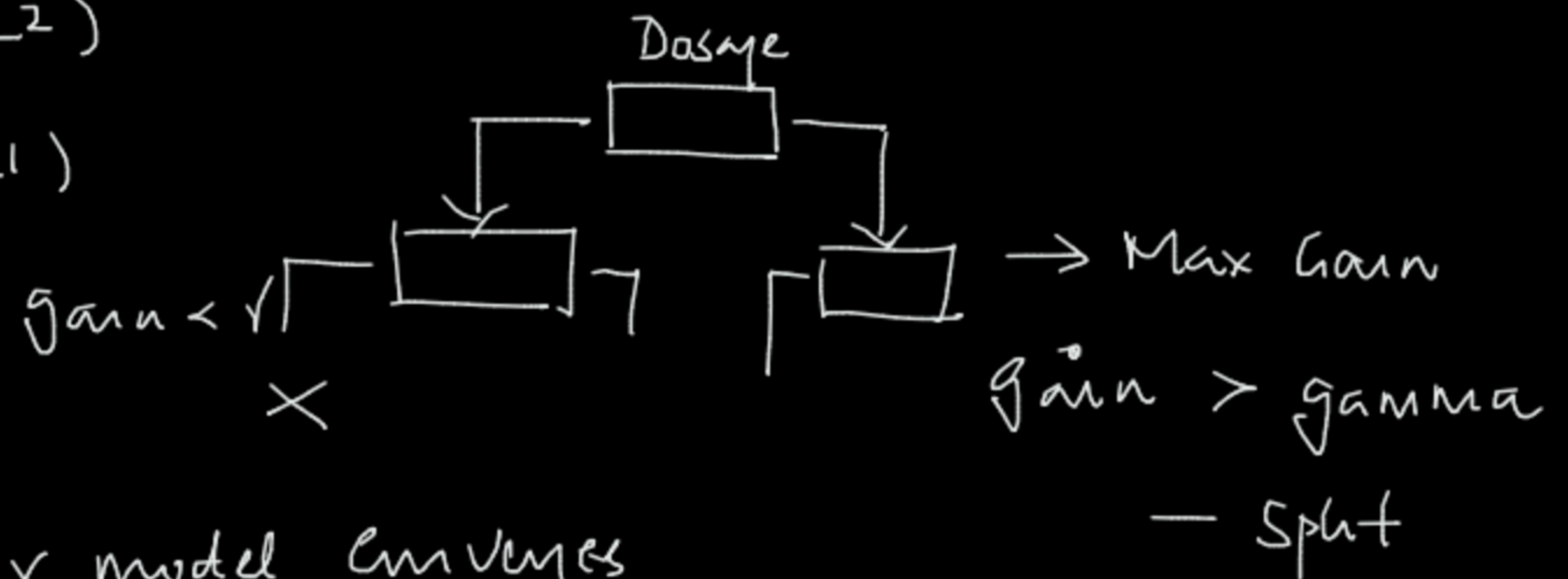
- 1 Binning / histogram \rightarrow reducing rows
- 2 FFR \rightarrow reduce features
- 3 Goss 4000 trials \rightarrow Best \rightarrow Max Gain

	✓	✓	✓	✓	
	x_1	x_2	x_3	x_4	y
1					
2					
3					

\rightarrow

Hyperparameters

- 1 max_depth \rightarrow
- 2 $\lambda \rightarrow$ Regularisation Param (L2)
- 3 $\alpha \rightarrow$ Regularisation Param (L1)
- 4 gamma \rightarrow threshold for gain



- 5 learning rate \rightarrow how fast your model converges

$$y = f_0(x) + \underset{\downarrow}{\alpha_1} \underset{\downarrow}{f_1(x)} + \underset{\downarrow}{\alpha_2} \underset{\downarrow}{f_2(x)}$$

0.3 0.1

$\left\{ \begin{array}{l} \text{Regularize} \checkmark \\ \text{Hyperparameter} \checkmark \\ \text{Model Validation} \checkmark \end{array} \right.$

α_i 's small \rightarrow long time to converge, more accurate
 large \rightarrow converge faster, less accurate

1. Binning

Bins = 10

10 split

10 records

1
2
3
4
5

10 records

2. Effective Feature Bundling

Multiple categorical columns are clubbed together

reduces the no of features

100

(Error)

3 Goss → Gradient Based One Sided Sampling

x_1 x_2 $\overset{x_3}{\boxed{x_5 \times x_7}}$
cat cat

4 features

→ 3 features

Pseudo Residuals → $\frac{\partial L}{\partial \hat{y}}$ ✓

M_1
1000 → 700 ✓ → 140 ✓
300 ✗ → 300

M_2
440

240 ✓ → 48 ✓
200 ✗ → 200
 M_3
248

SVM - Support Vector Machines

- classification
- margin - maximization algorithm

2d \rightarrow line $y = \beta_0 + \beta_1 x$

3d \rightarrow plane $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

$> 3d \rightarrow$ hyperplane

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_m x_m$$

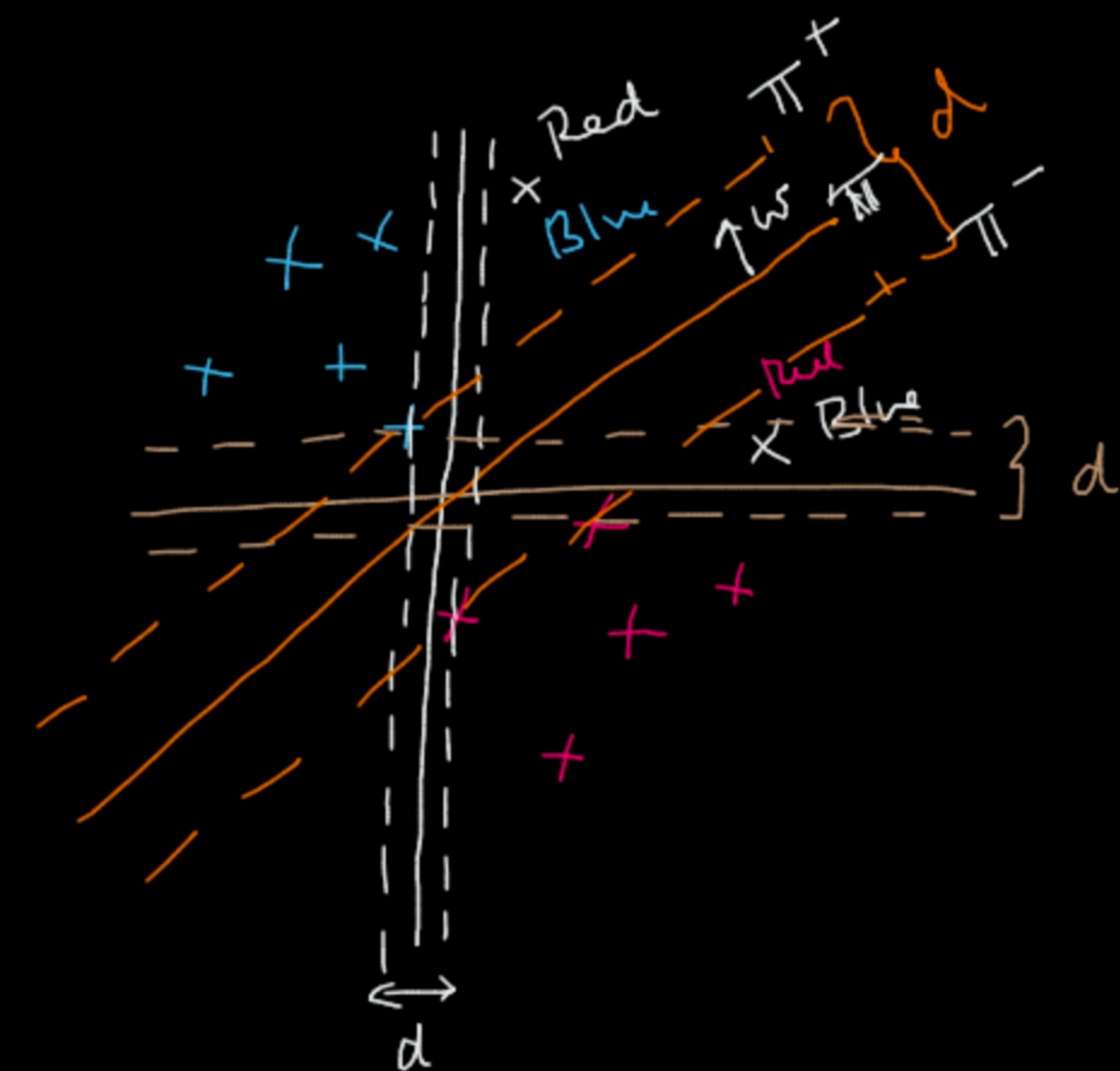
Task: Find the hyperplane s.t. the margin betw π^+ & π^-

$d \rightarrow$ maximum

β 's \rightarrow importance given to each feature
 \rightarrow weightage

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_m x_m$$

$$= \underbrace{w_1 x_1 + w_2 x_2 + w_3 x_3 + \dots + w_m x_m}_{w^T x} + b$$



$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}$$

$$w^T = [w_1 \ w_2 \ w_3 \ \dots \ w_m] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}$$

$$w^T x = w_1 x_1 + w_2 x_2 + \dots + w_m x_m$$

$$\pi^+ - \pi^- \Rightarrow d$$

$$\Rightarrow \frac{2}{\|\vec{w}\|}$$

↳ Magnitude
- norm

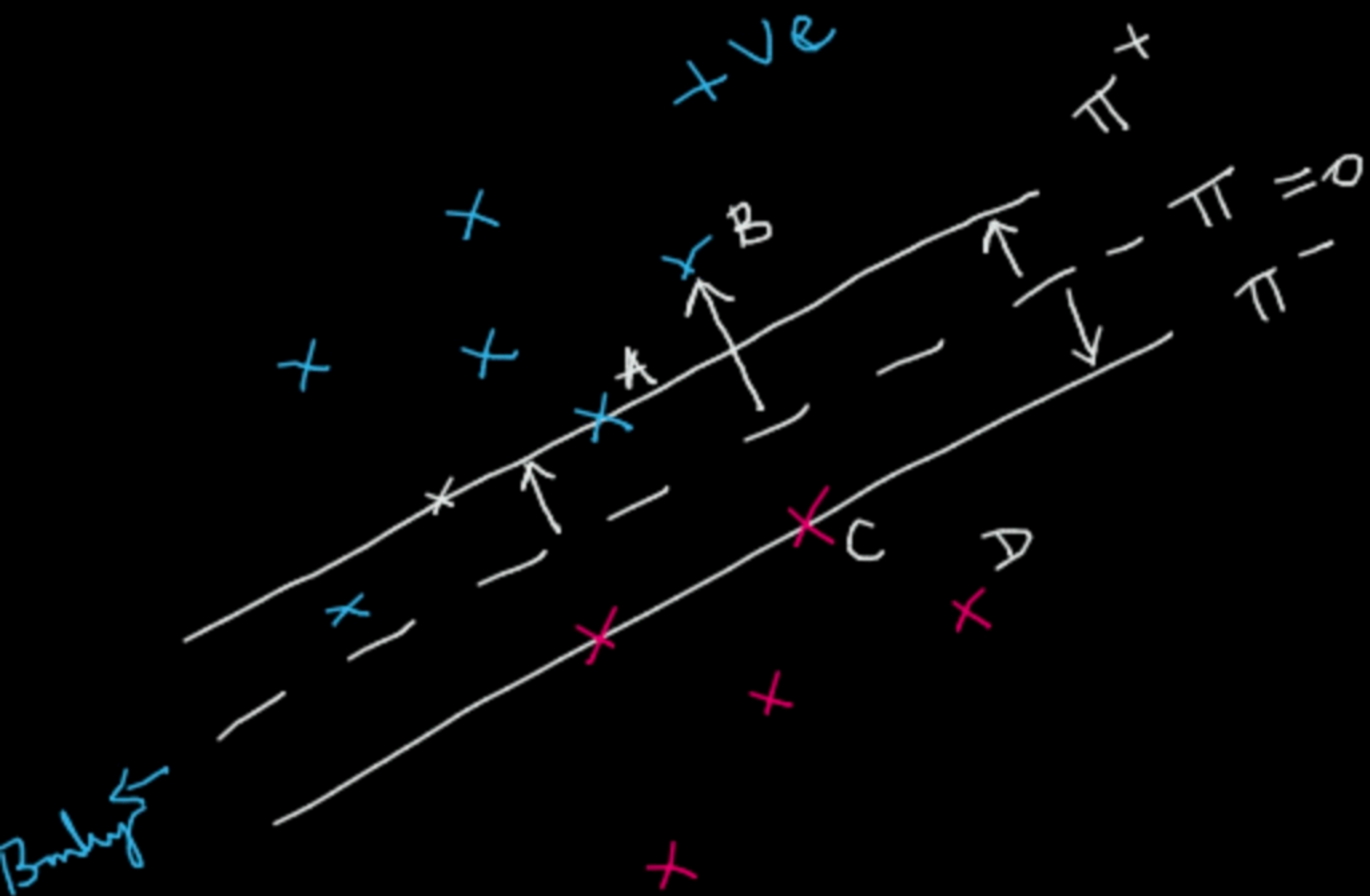
$$\pi = w^T x + b$$

Task Find the \vec{w} \rightarrow Hypoplane

$$\frac{2}{\|\vec{w}\|} \text{ is maximum}$$

Constraint

$$\forall x_i = y_i (w^T x_i + b) \geq 1.5 \quad \checkmark$$



$$A = W^T x_1 + b = +1$$

$$= y_i (W^T x_i + b) = \underline{+1}$$

(+1) (+1)

$$B = W^T x_1 + b = +1.5$$

$$= y_i (W^T x_i + b) = \underline{+1.5}$$

(+1) (+1.5)

→ Dummy points

$$\pi \quad W^T x + b = 0$$

$$\pi^+ \quad W^T x + b = 1$$

$$\pi^- \quad W^T x + b = -1$$

$$C = W^T x_1 + b = -1$$

$$= y_i (W^T x_i + b) = \underline{-1}$$

(-1) (-1)

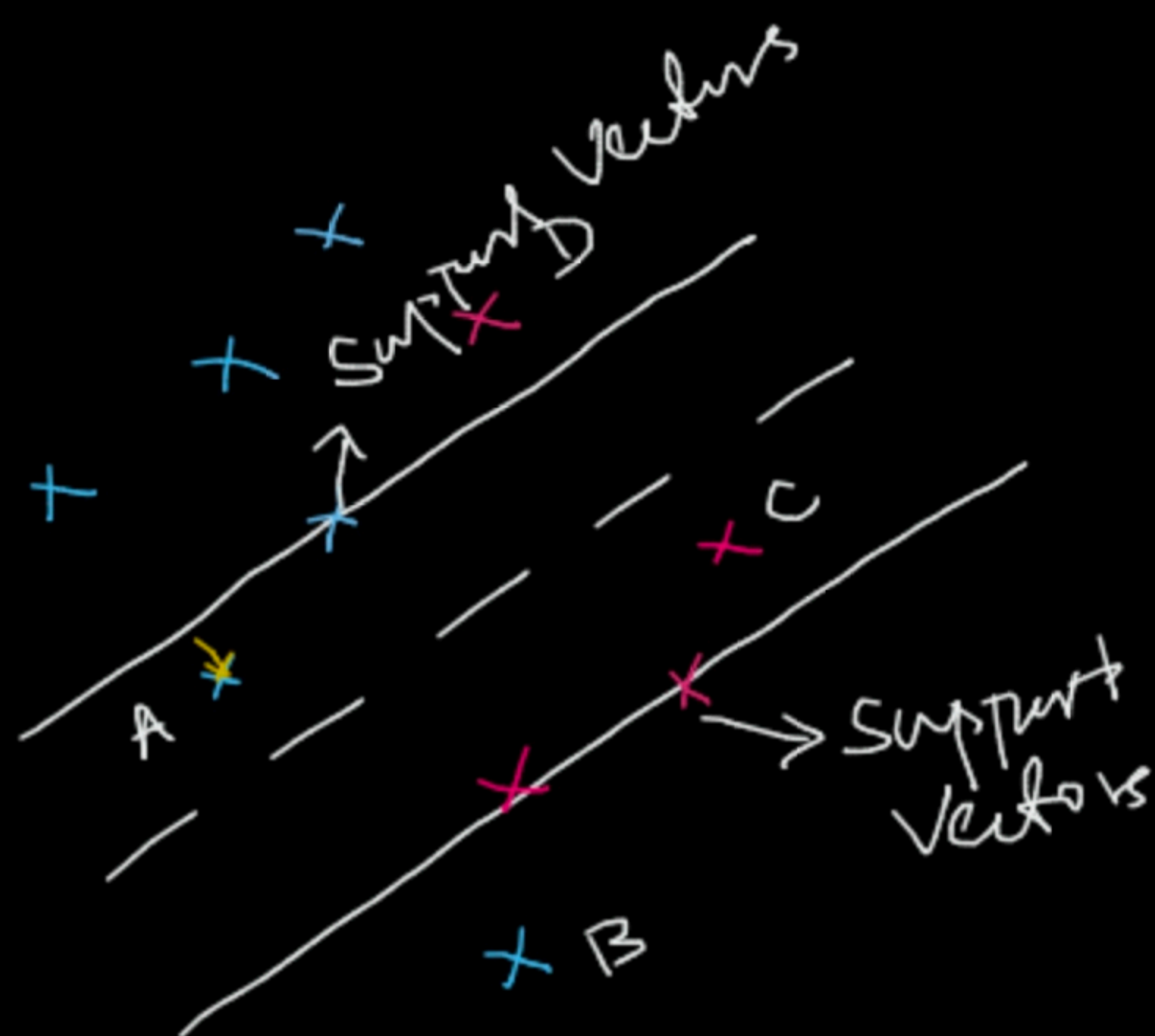
$$D = W^T x_1 + b = -1.5$$

$$y_i (W^T x_i + b) = \underline{+1.5} \rightarrow \text{"Hard Margin"}$$

$$y_i (W^T x_i + b) \geq \underline{1 - \epsilon} \quad \forall \quad x_i$$

↳ 0

x_1	x_2	x_3	y
-	-	-	+1 ✓
-	-	-	-1
-	-	-	-1
-	-	-	+1



$$A: (W^T x_1 + b) = +0.5$$

$$y_1 (W^T x_1 + b) = +0.5 \\ = (1 - 0.5) \\ \text{+ve}$$

$\xi \rightarrow +ve$

All Wrongly Cls Pts

$$y_1 (W^T x_1 + b) \geq 1 - \xi$$

$$B: W^T x_1 + b = -1.5$$

$$y_1 (W^T x_1 + b) = -1.5 \\ +1 (-1.5) = 1 - 2.5 \\ \text{+ve}$$

$$C: W^T x_1 + b = -0.5$$

$$y_1 (W^T x_1 + b) = 0.5 \\ (-1) (-0.5) = (1 - 0.5) \\ \text{+ve}$$

$$D: W^T x_1 + b = 1.5$$

$$y_1 (W^T x_1 + b) = -1.5 \\ (-1) (1.5) = 1 - 2.5 \\ \text{+ve}$$

Task Find the w^T (hyperplane) so $\frac{1}{d}$,

$$d \rightarrow \max \quad \Rightarrow \quad \frac{2}{\|\vec{w}\|} \quad \text{is max}$$

$$1/d \rightarrow \min$$

and $\forall x_i \Rightarrow y_i(w^T x_i + b) \geq 1 - \xi_i$

$\xi_i > 0$ - For wrong
 $= 0$

Task of the algorithm:

Find The \vec{w} (hyperplane) s.t

$$\Rightarrow \frac{1}{d} + \frac{1}{n} \sum_{i=1}^n \xi_i \quad \text{is min}$$

ξ_i - Error in the model

$$\rightarrow \frac{1}{n} \sum_{i=1}^n \xi_i \quad \text{Hinge Error}$$

min

$$\forall x_i, \quad y_i(w^T x_i + b) \geq 1 - \xi_i \quad \begin{matrix} \xi_i > 0 & \text{wrong} \\ = 0 & \text{correctly} \end{matrix}$$

Linear SVC