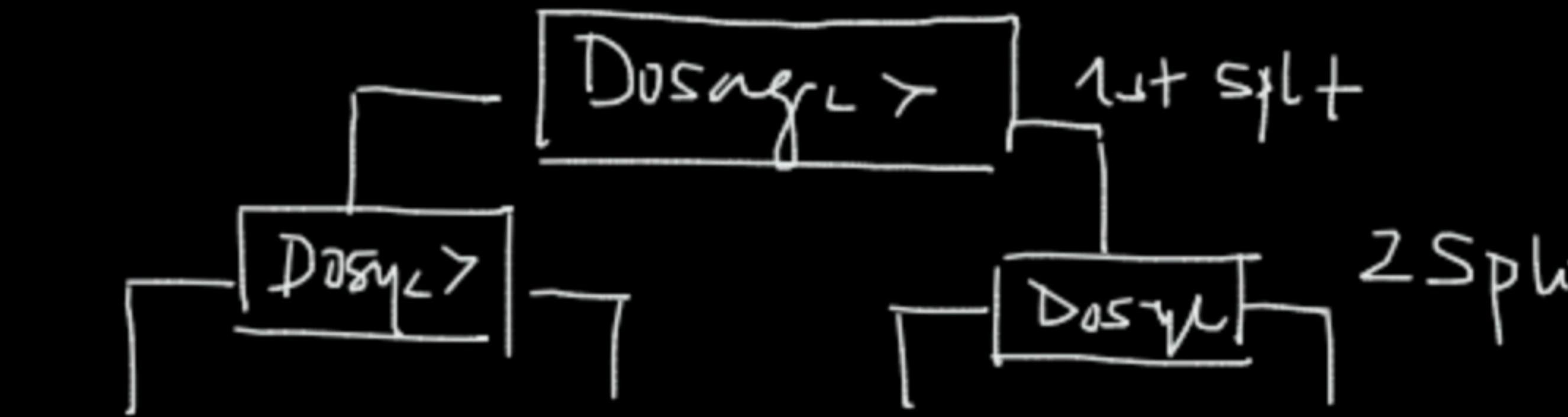


$$\frac{x}{\text{Drug Dosage}} \rightarrow \frac{y}{\text{Efficiency}}$$

1	-	-
2	-	-
3	-	-
4	-	-
5	-	-

XGBoost



Pip install lightgbm

$$\text{Gain} = \text{Left}_{\text{SVM}} + \frac{\text{Right}_{\text{SVM}} - \text{parent}_{\text{SVM}}}{\text{SVM}}$$

$$\text{SVM} = \frac{\sum (\text{Residuals})^2}{m \circ \text{rest} + \lambda \Rightarrow 0}$$

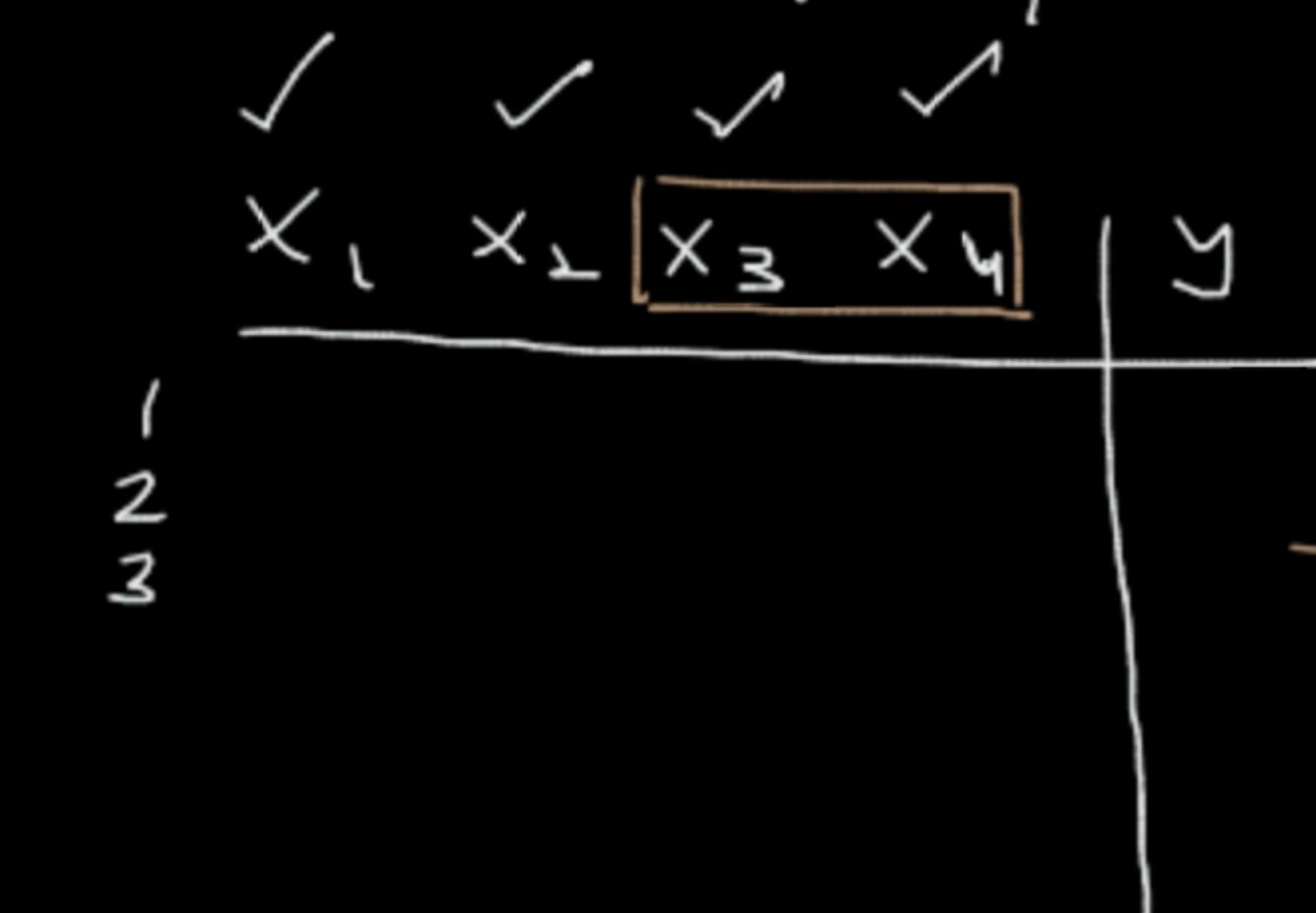
↳ Regularization

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

lightGBM → Some approximations

1 Binning / histogram → Reducing ^{rows} rows

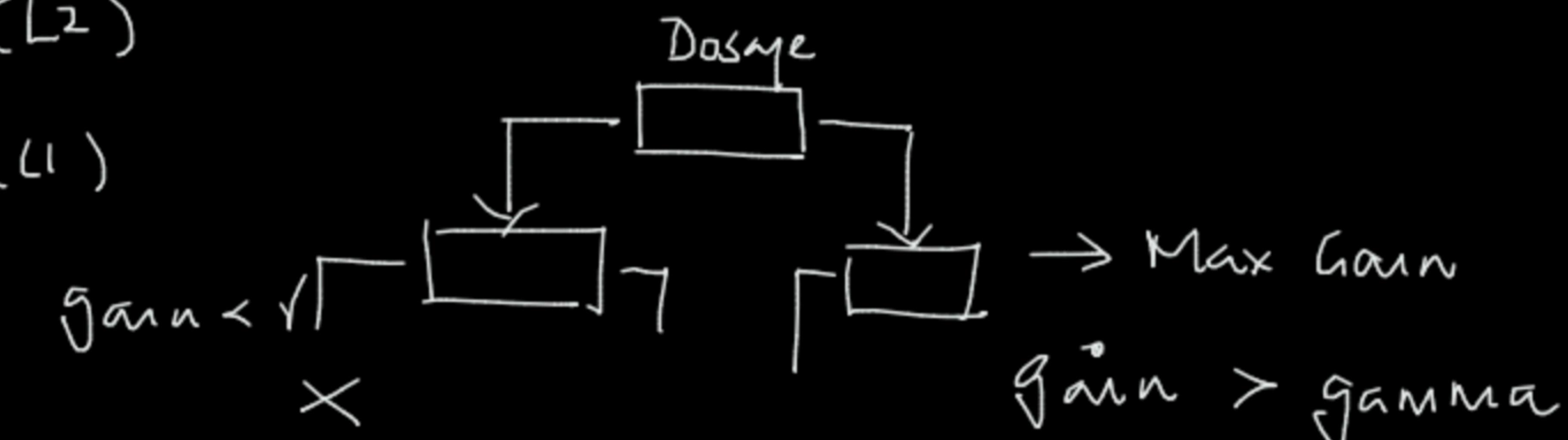
2 FFB → Reduce features

3 Goss $\frac{4000 \text{ trials}}{\rightarrow \text{Max Gain}}$ → Best

Hyperparameters

- 1 max_depth \rightarrow
 - 2 $\lambda \rightarrow$ Regularisation Parameter (L^2)
 - 3 $\alpha \rightarrow$ Regularisation Parameter (L^1)
 - 4 Gamma \rightarrow threshold for gamma
 - 5

≤ learning rate → How fast your model converges

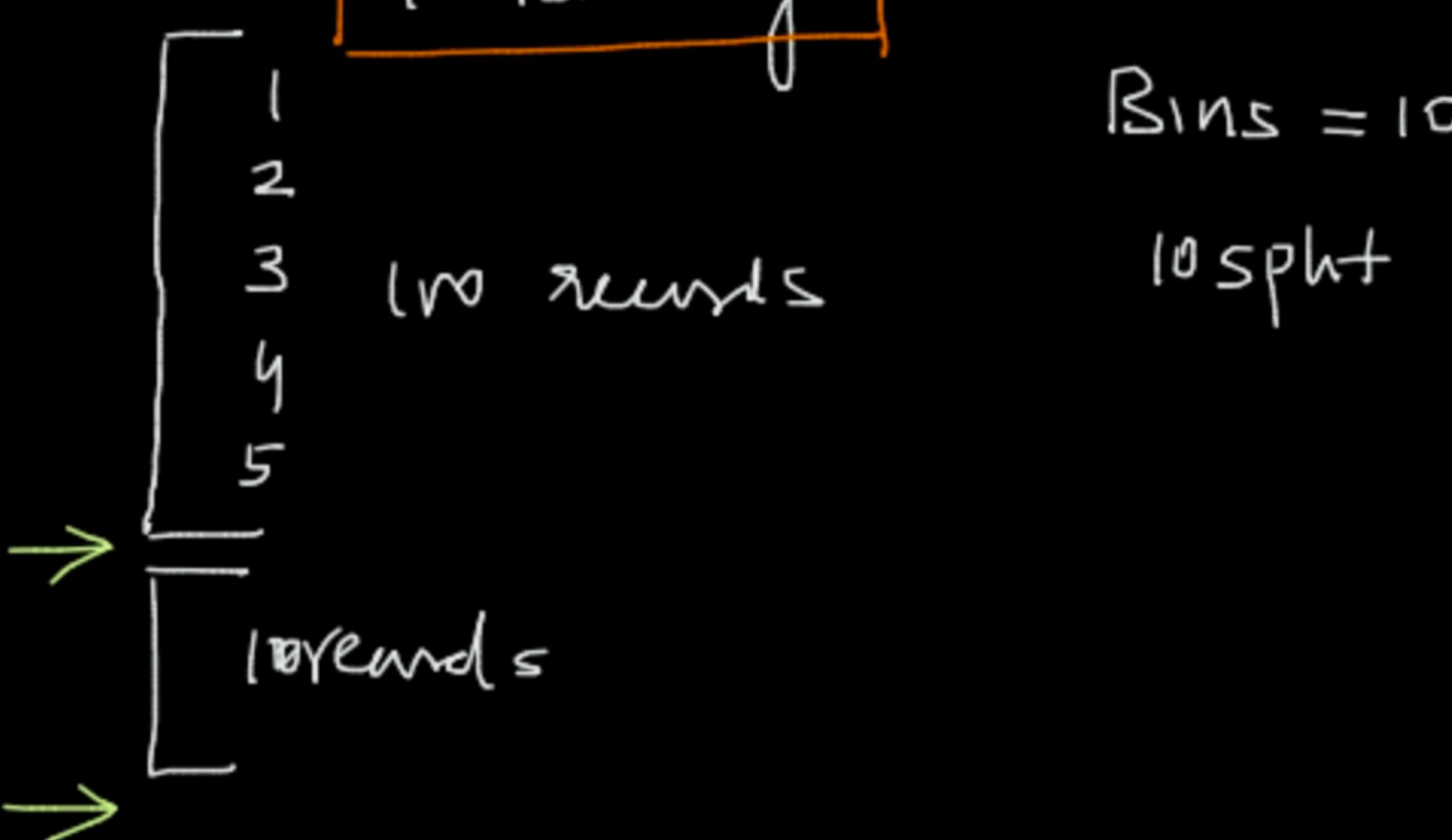


$$y = f_0(x) + \alpha_1 f_1(x) + \alpha_2 f_2(x)$$

- Regulam Ze ✓
- Hyperparam ✓
- Model Validation ✓

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

1. Binning



2 Effective Feature Bundling

- Multiple categorical columns are clubbed together
- reduces the no of features

100

(Error)

3 Gross → Gradient Based One Sided Sampling

$$M_1 \\ 100 \rightarrow 700 \checkmark \rightarrow 140 \checkmark \\ 300 \times \rightarrow 300$$

$$\frac{440}{440} \xrightarrow{M_2}$$

$$240 \checkmark \rightarrow 48 \checkmark \\ 200 \times \rightarrow \frac{200}{248} \xrightarrow{M_3}$$

$$x_1 \times_2 \frac{x_3 \times_4}{\text{cut} \quad \text{cut}}$$

4 features → 3 features

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

SVM - Support Vector Machines

— classification

— margin - maximization algorithm

$$2d \rightarrow \text{line} \quad y = \beta_0 + \beta_1 x$$

$$3d \rightarrow \text{plane} \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$\geq 3d \rightarrow \text{hyperplane}$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n$$

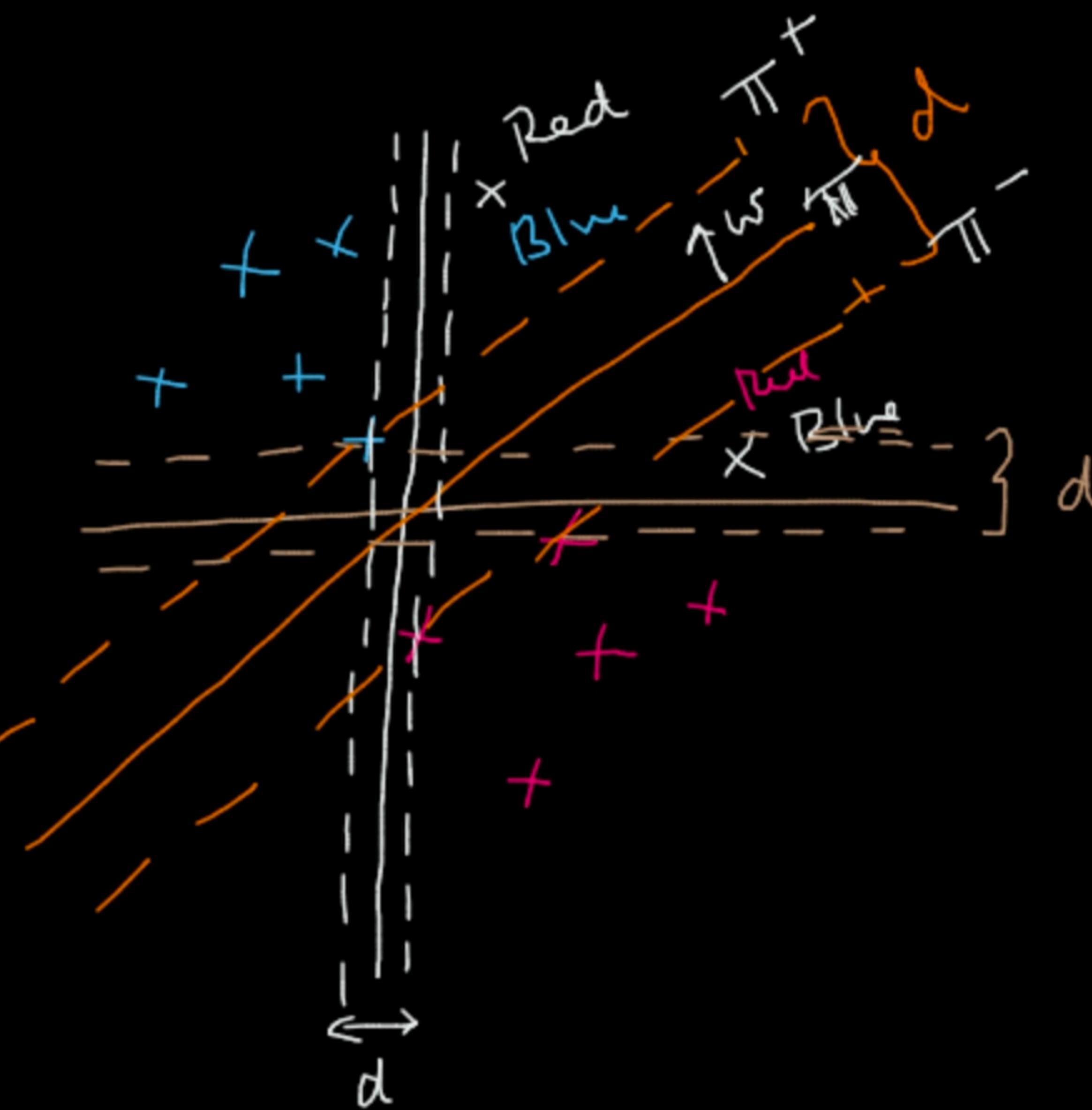
Task: Find the hyperplane \vec{w} for the margin between π^+ & π^-

$d \rightarrow \text{maximum}$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n$$

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

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



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

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

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

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

$$\pi^+ - \pi^- \Rightarrow d \Rightarrow \frac{2}{\|\vec{w}\|}$$

↳ Magnitude
- norm

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

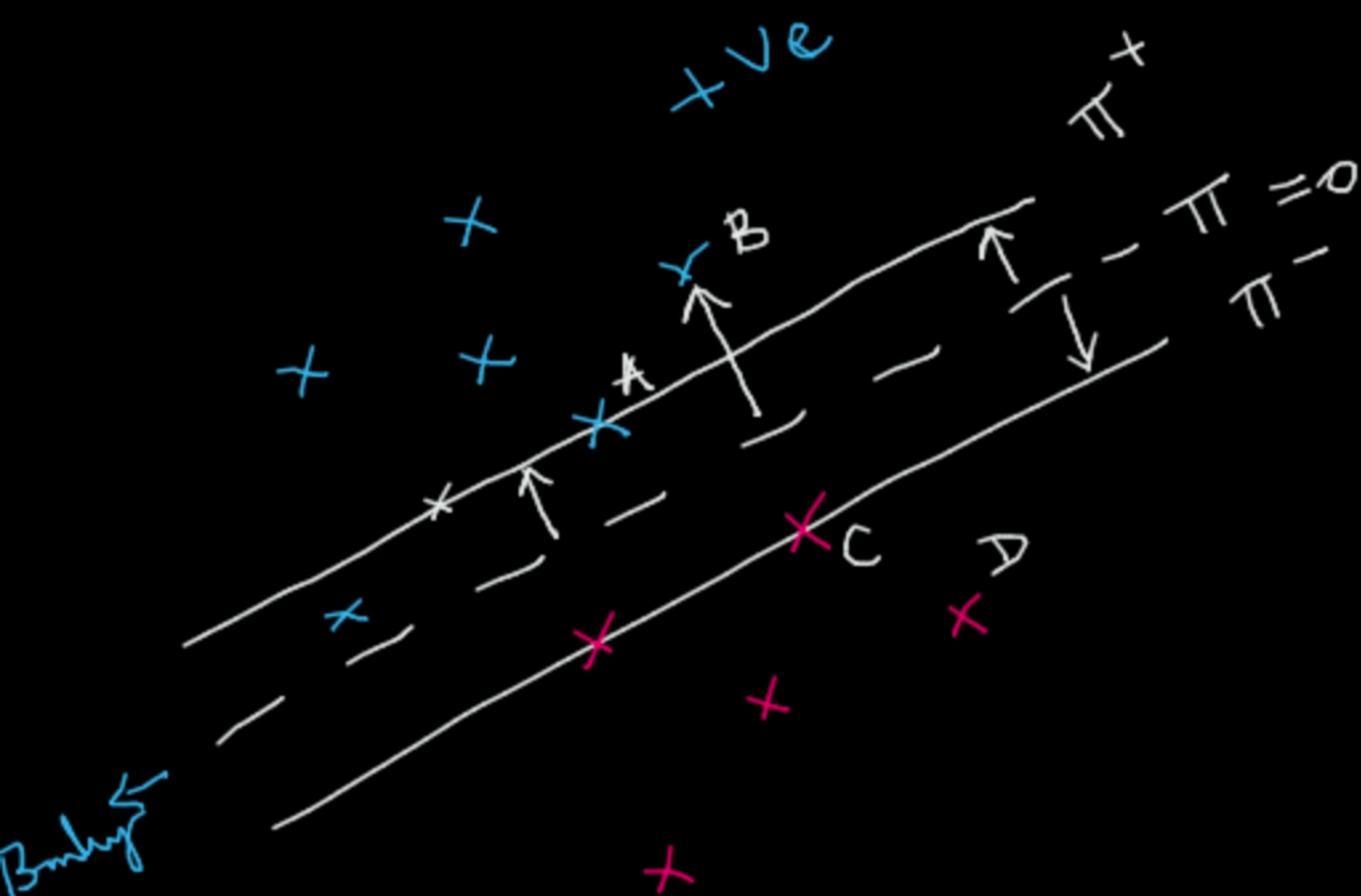
Hyperspace

Task Find the \vec{w} $\vec{\pi} +$

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

Constraint

$$\forall x_i = y_i (w^\top x_i + b) \geq 1.5$$



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

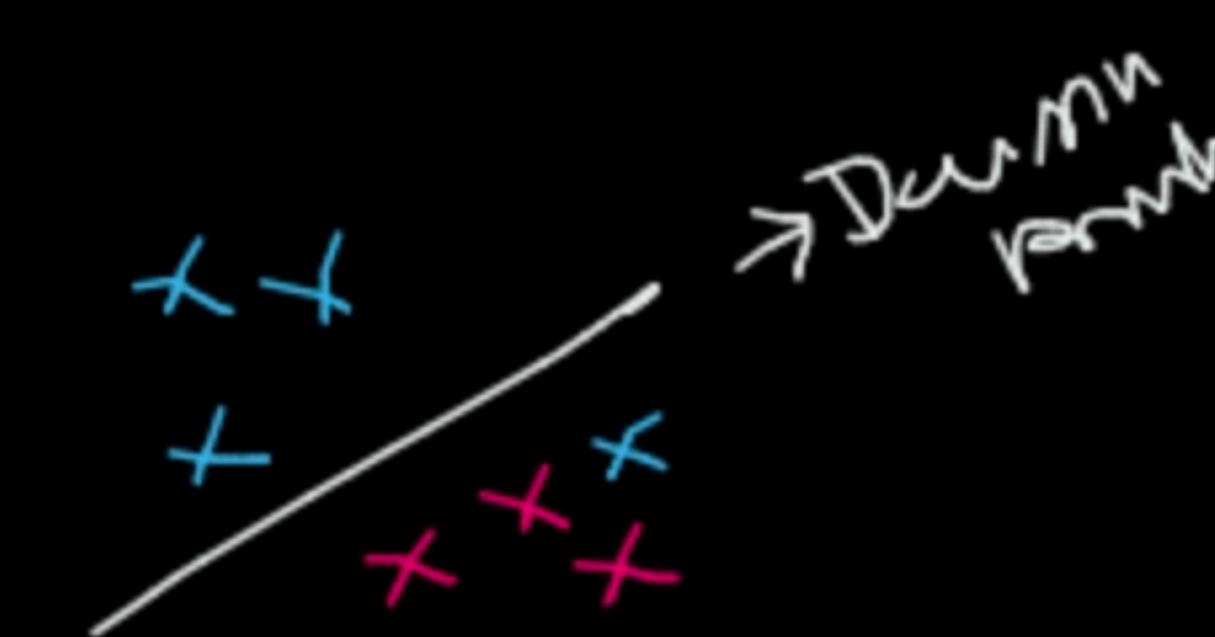
$$= y_1 (w^T x_1 + b) = +1$$

(+1) (+1)

$$B = w^T x_2 + b = +1.5$$

$$= y_2 (w^T x_2 + b) = +1.5$$

(+1) (+1.5)



$$C = w^T x_3 + b = -1$$

$$= y_3 (w^T x_3 + b) = -1$$

(-1) (-1)

$$D = w^T x_4 + b = -1.5$$

$$y_4 (w^T x_4 + b) = -1.5 \rightarrow "Hard Margin"$$

$$\Pi \quad w^T x + b = 0$$

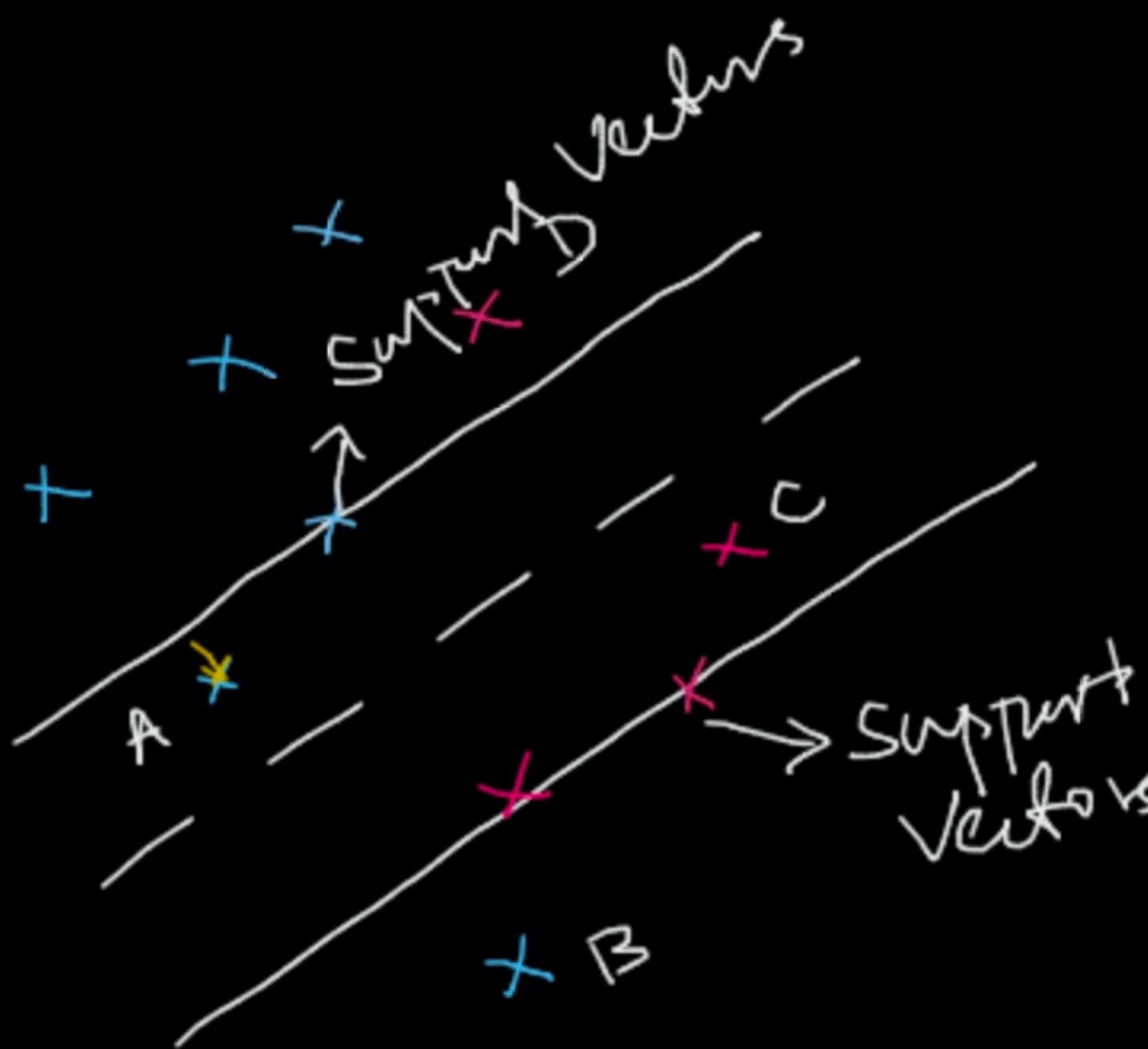
$$\Pi^+ \quad w^T x + b = 1$$

$$\Pi^- \quad w^T x + b = -1$$

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

$$y_1 (w^T x_1 + b) \geq 1 - \xi_1 \forall x_1$$

↳ 0



$$A: \mathbf{w}^\top \mathbf{x}_1 + b = +0.5$$

$$y_1 (\mathbf{w}^\top \mathbf{x}_1 + b) = +0.5$$

$$= (1 - 0.5) \quad \text{if } \rightarrow \text{+ve}$$

$$B: \mathbf{w}^\top \mathbf{x}_n + b = -0.5$$

$$y_1 (\mathbf{w}^\top \mathbf{x}_n + b) = -0.5$$

$$+1 (-1.5) = 1 - 2.5$$

$$y_1 (\mathbf{w}^\top \mathbf{x}_n + b) \geq 1 - \xi$$

$$C, \quad \mathbf{w}^\top \mathbf{x}_1 + b = -0.5$$

$$y_1 (\mathbf{w}^\top \mathbf{x}_1 + b) = 0.5$$

$$(-1) (-0.5) = (1 - 0.5)$$

$$D \quad \mathbf{w}^\top \mathbf{x}_1 + b = 1.5$$

$$y_1 (\mathbf{w}^\top \mathbf{x}_1 + b) = -1.5$$

$$(-1) (1.5) = 1 - 2.5$$

Task Find the \vec{w} (hyperplane) so that,

$$d \rightarrow \max \quad \Rightarrow \quad \frac{1}{\|\vec{w}\|} \rightarrow \max$$

$$\text{and } \forall n \Rightarrow y_n(\vec{w}^T x_n + b) \geq 1 - \xi_n$$

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

Task of the algorithm:

Find the \vec{w} (hyperplane) so that

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

ξ_i - Error in the model

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

$$\forall n, \quad y_n(\vec{w}^T x_n + b) \geq 1 - \xi_n \quad \xi_n > 0 \text{ wrong}$$

Linear SVC

$= 0$ correctly