Perception :

$[2, 2, 1]$

+ve class : $A = \{[2,1,1], [3,2,1], \cancel{\quad}\}$ $\quad w^T x > 0$

$\overset{0}{}$

-ve class : $B = \{[-1,-1,1], [-1,-2,1], [-2, \cancel{\ }, 1]\}$ $\quad w^T x < 0$

Let us Initialize W as $[1,-1,1]$

for class A : $[1,-1,1] \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 2 > 0 \Rightarrow$ correct prediction

$\boxed{y_i = +1}$

$[1,-1,1] \cdot \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = 2 > 0 \Rightarrow$ correct prediction

$[1,-1,1] \cdot \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = 1 > 0 \Rightarrow$ correct prediction

for class B : $[1,-1,1] \cdot \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = 1 > 0 \Rightarrow$ wrong prediction since for this class $w^T x$ has to be < 0

$\boxed{y_i = -1}$

$[1,-1,1] \cdot \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} = 2 > 0 \Rightarrow$ wrong prediction

$[1,-1,1] \cdot \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = -1 < 0 \Rightarrow$ correct prediction because $w^T x \geq 0$

Therefore error set will have samples $\Rightarrow \{[-1,-1,1], [-1,-2,1]\}$

Perception update : $\boxed{W^{k+1} = W^k + \sum_{x_i \in \text{error set}} x_i y_i}$ $\quad \boxed{\xi \to \text{error set}}$

$\therefore \quad W^2 = W^1 + \sum_{x_i \in \xi} x_i y_i \quad , \quad W^1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ (Initialized W)

$$\therefore w^2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

$$\therefore w^2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$

Now, Repeat and calculate error set for w^2 and update using the new error set.

For w^2 you will note that w^2 will be empty and hence all samples are correctly classified, output w of Perceptron will be $\Rightarrow [3, 2, -1]^T$

For perceptron with L_2 Regularization

$$\text{Loss} = -\sum_{x_i \in \text{error set}} y_i w^T x_i + \left(\sum_{j=1}^{d} w_j^2\right)\lambda$$

$\boxed{W \Rightarrow d \text{ dimensional vector}}$

$$\frac{\partial(\text{loss})}{\partial w} = -\sum_{x_i \in E} y_i x_i + \lambda\left(2\sum_{j=1}^{d} w_j\right)$$

$\therefore$ update Rule : $W^{k+1} = w^k - \left(-\sum_{x_i \in E} y_i x_i + 2\lambda \sum_{j=1}^{d} w_j\right)$

$$\boxed{W^{k+1} = w^k + \sum_{x_i \in E} y_i x_i - 2\lambda \sum_{j=1}^{d} w_j}$$

$\sum_{j=1}^{d} w_j$ is nothing but ~~some~~ sum of values of w.

λ = hyperparameter