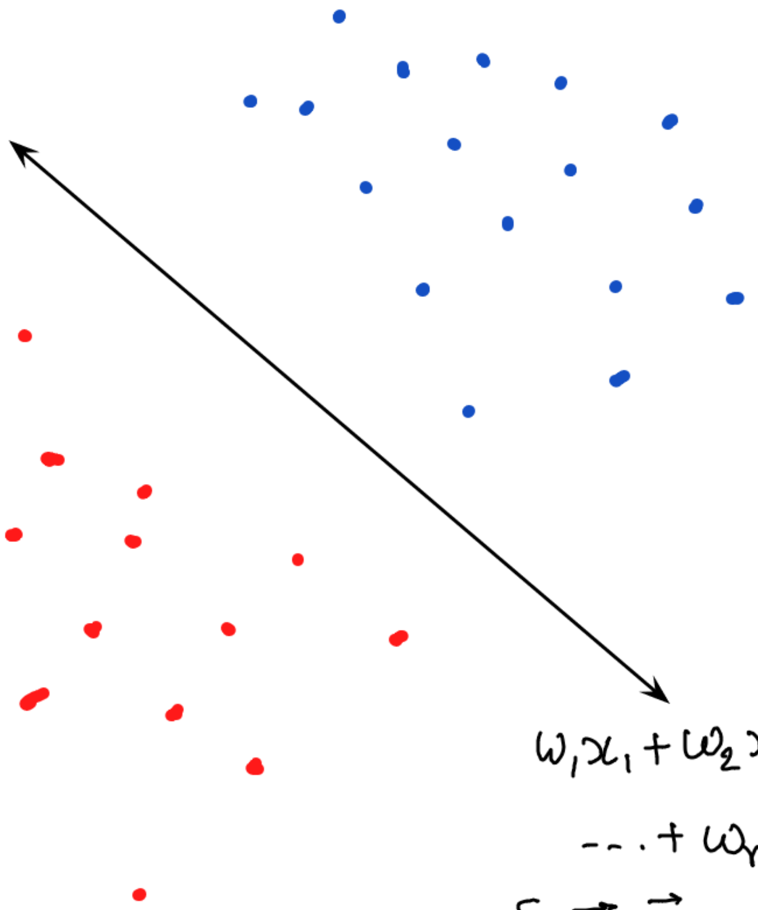
☆ Loss Function / Loss Minimization



$w_1 x_1 + w_2 x_2 + \dots$

$\dots + w_n x_n + w_0 = 0$

$[\ \vec{w} \cdot \vec{x} + w_0 = 0\ ]$

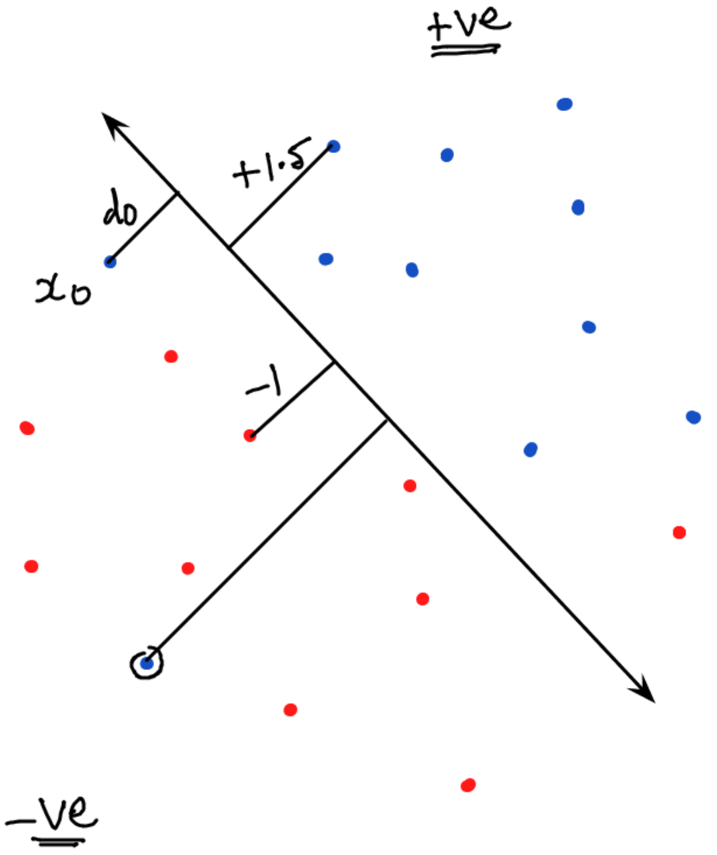
The goal is to maximize the distance of each point from the line.

∴ This distance (that I want to maximize) is our gain function.

∴ We want to maximize:

$$\boxed{\sum_{i=0}^{n} \frac{w^T \cdot x_i + w_0}{\| \vec{w} \|}}$$

This logic has a problem. What if a point is misclassified?



$d_0 \rightarrow$ calculated $= -1$
actual $= +1$ $\Rightarrow x_0$ is misclassified.

Problems:

1. +ve & -ve distances will cancel out each other
2. Misclassified points are not penalized.

A nice 'hack' was introduced to solve both these issues.

★ Some terminology :- y_i = Actual label of point x_i

$\hat{y}_i$ = Predicted label of point x_i

→ Suppose we are solving that fish-sorting problem. The data is as below:

Sr. No.	length	breadth	weight	type ↳ y_i	distance	Prediction ↳ $\hat{y}_i$
1	10	2	20	$\boxed{1}$	+2.5	$\boxed{1}$
2	70	30	500	$\boxed{-1}$	+3	$\boxed{1}$
3	8	1	12	$\boxed{1}$	+1.5	$\boxed{1}$
4	100	40	700	$\boxed{-1}$	−1.9	$\boxed{-1}$

The hack: Multiply distance of each point by its actual label.

$\therefore d_1 = 2.5 * 1 = 2.5$ (appreciation)

$d_2 = 3 * (-1) = -3 \Rightarrow$ penalty

$d_3 = 1.5 * 1 = 1.5 \Rightarrow$ appreciation

$d_4 = -1.9 * (-1) = 1.9 \Rightarrow$ "

→ Putting this idea into formula:

$$\boxed{\sum_{i=0}^{n} \frac{w^T . x_i + w_0}{\|\vec{w}\|} \cdot y_i} \longrightarrow \text{Gain Function}$$

→ This is called Gain Function because we want to maximize the output of this function. Then what if we want to find loss function?

$$\boxed{L(x, y, w, w_0) = - \sum_{i=0}^{n} \frac{w^T . x_i + w_0}{\|\vec{w}\|} \cdot y_i} \longrightarrow \text{Loss Function}$$