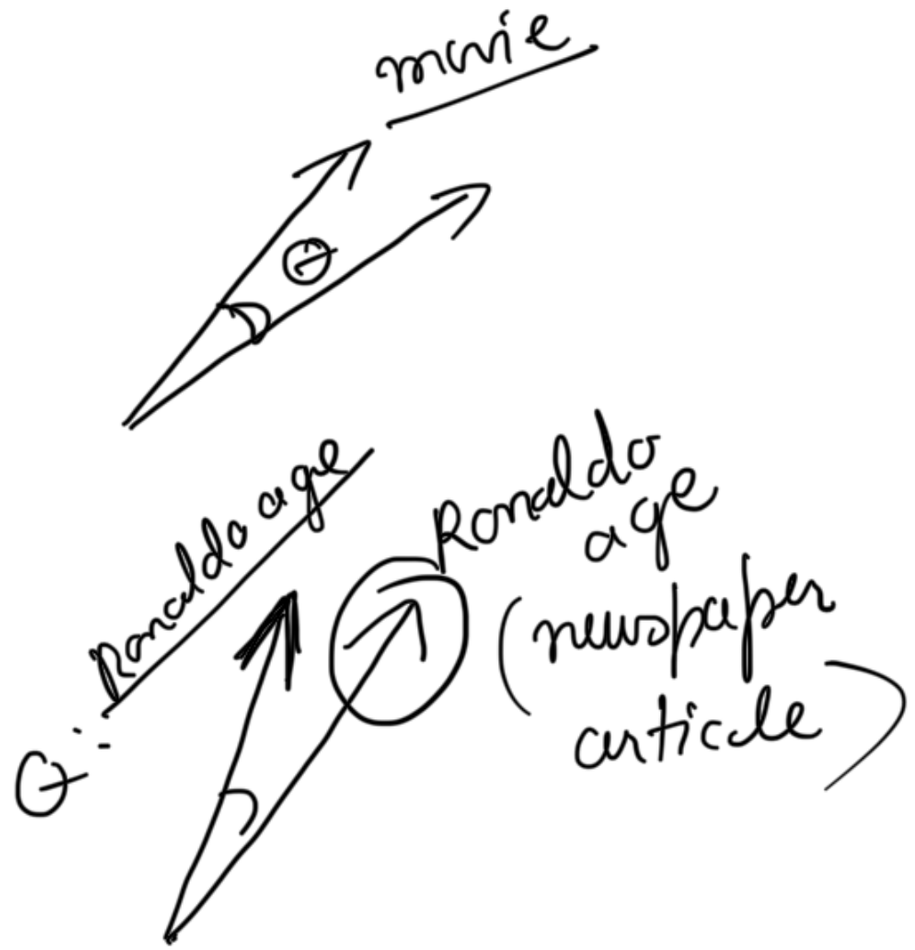
Optimization - 1

Why angle b/w 2 vectors is important in ML?



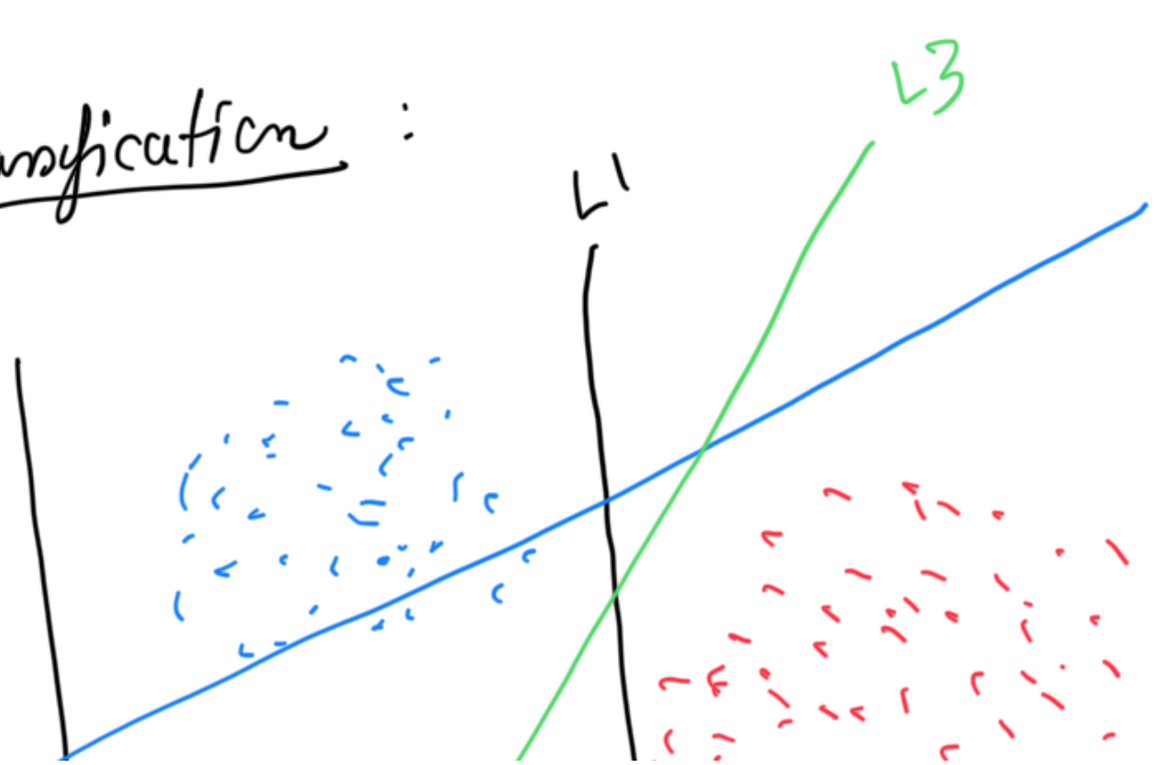
→ Classifier algo

⇒ optimization problems.

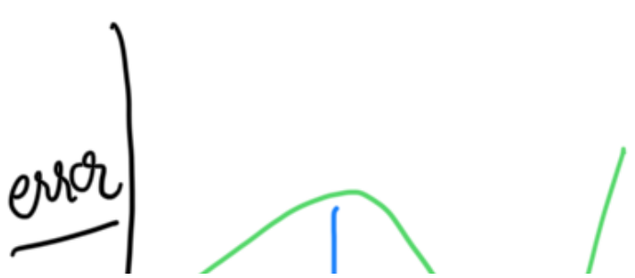
→ Limits → continuity.

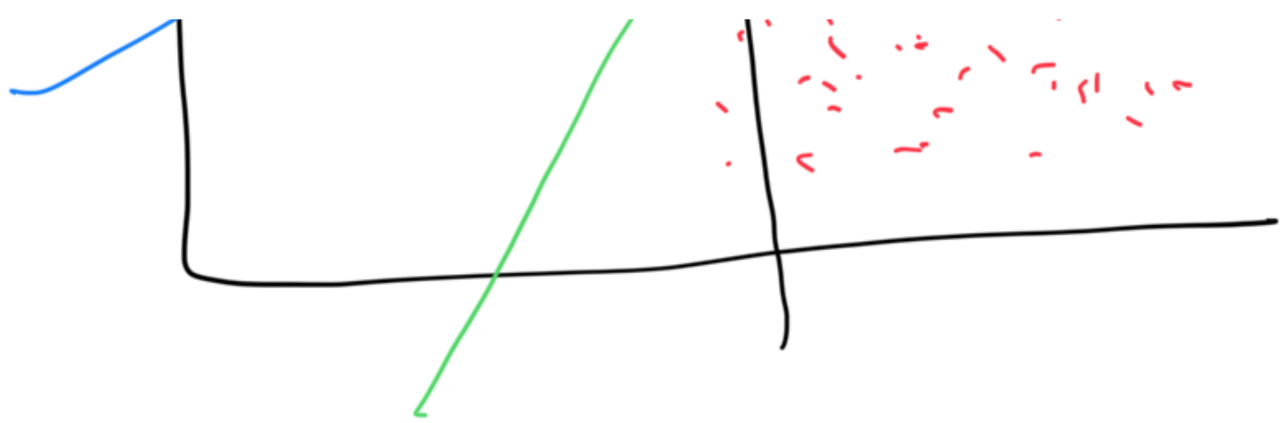
→ domain and range.

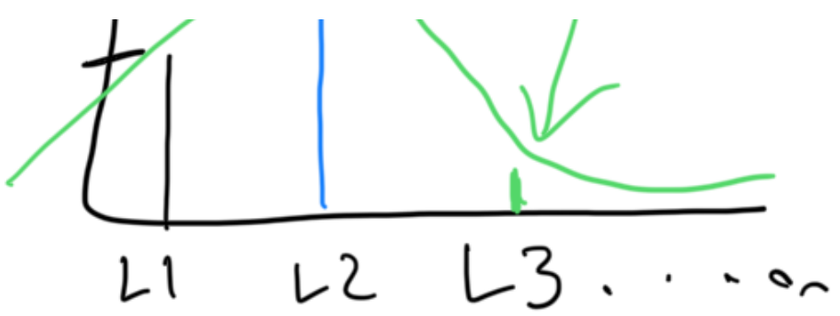
Classification :

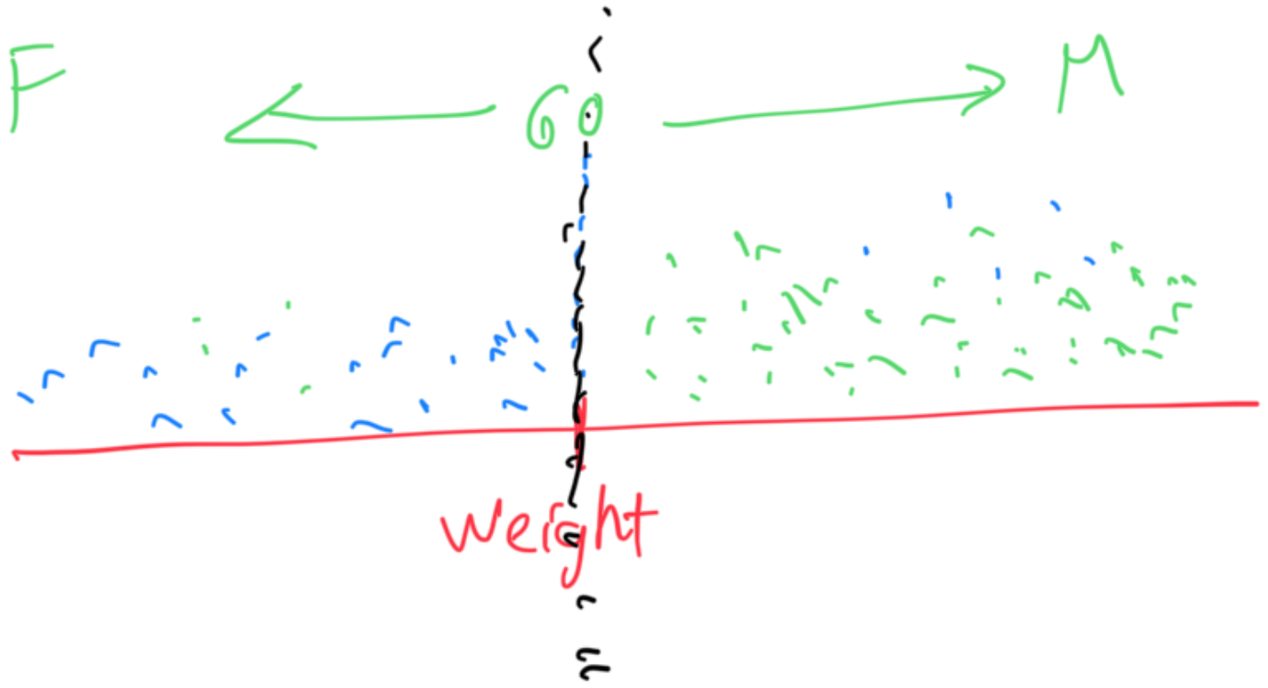


Linear boundary



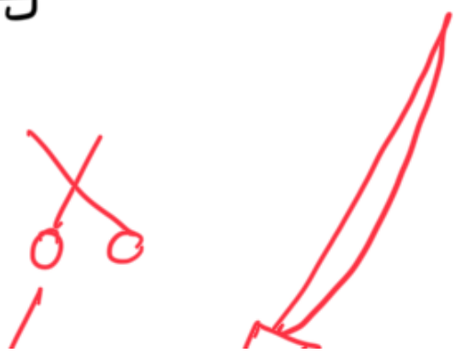






Ocam's

Razar



simple choice.

heavy op.

Join

Group · By

Selec

A Join B

ON

Group by

(Selec

Group by) the

join B

---

line : slope and intercept

$$w_n x_n = 0$$

$w_0 + w_1 x_1 + w_2 x_2$

$$W = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$\rightarrow \boxed{4}\ \boxed{\ }\ \boxed{\ }\ \boxed{\ }$

$\boxed{10}^4$

Linear Search

$w_0 \rightarrow [-10 \text{ to } 10]$ $\boxed{0.1}$

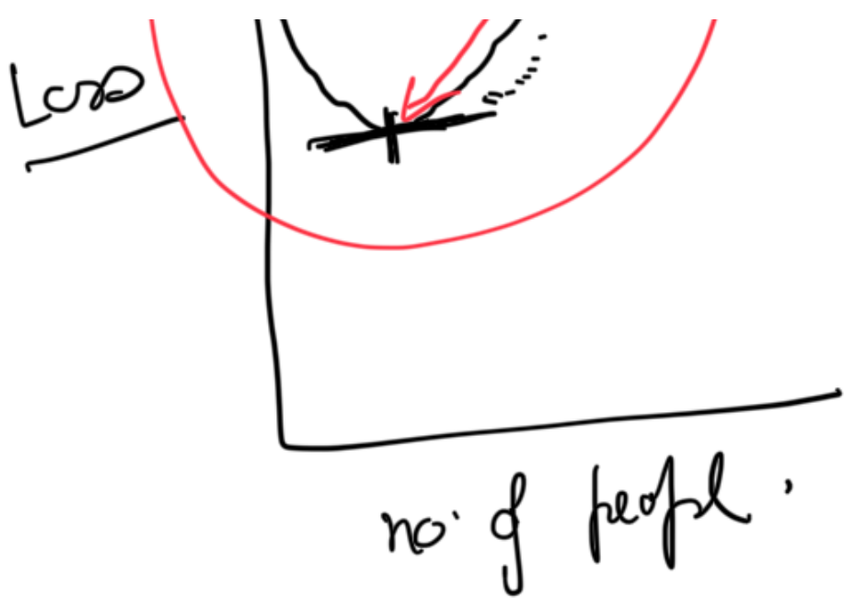
$w_1 \rightarrow [-10 \text{ to } 10]$

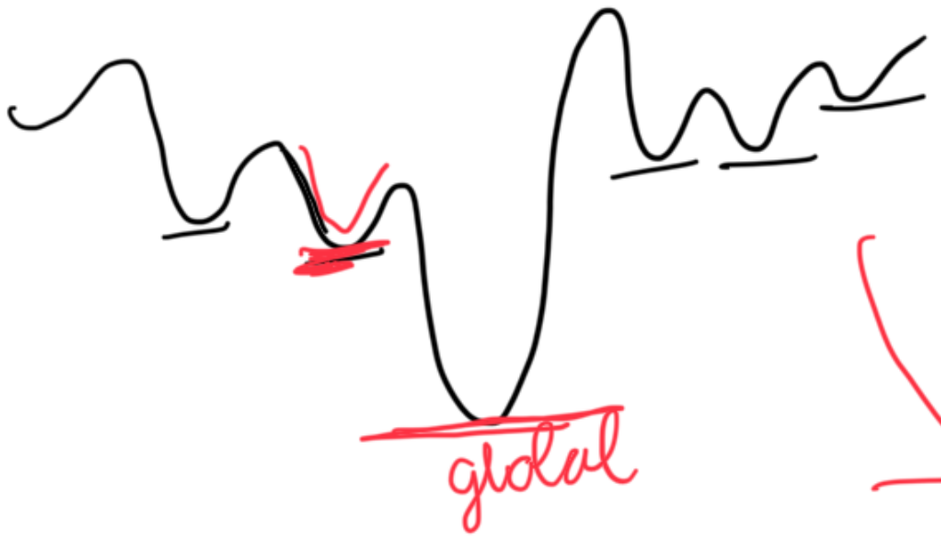
w_3

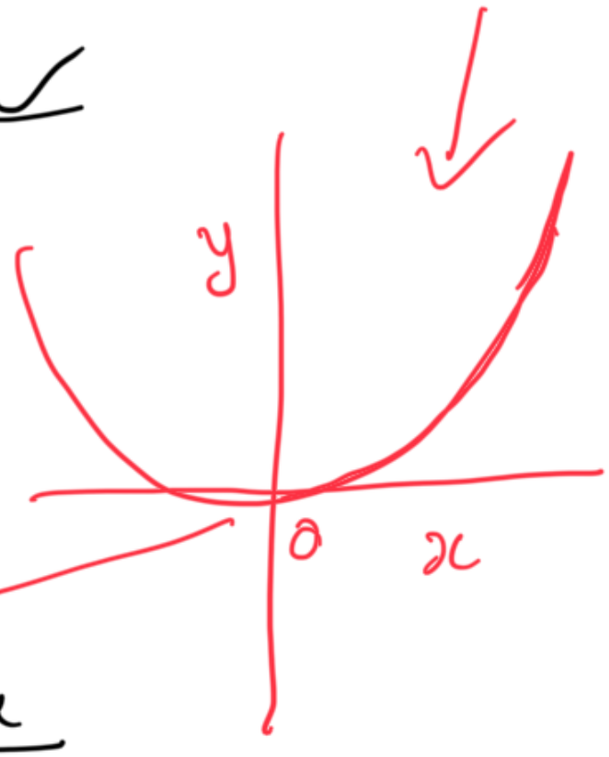
$\rightarrow \boxed{201}^3$

$\boxed{210}^3$

x^2







$y = x^2$

global-minima

Gradient Descent

Calculus → multi-variable calculus.

→ slope, derivative.

$$\frac{d}{dx}(x^2) = \boxed{2x}$$

Binary classification → label

$D: (x_i, y_i)$ $\quad y_i \in \{-1, 1\}$

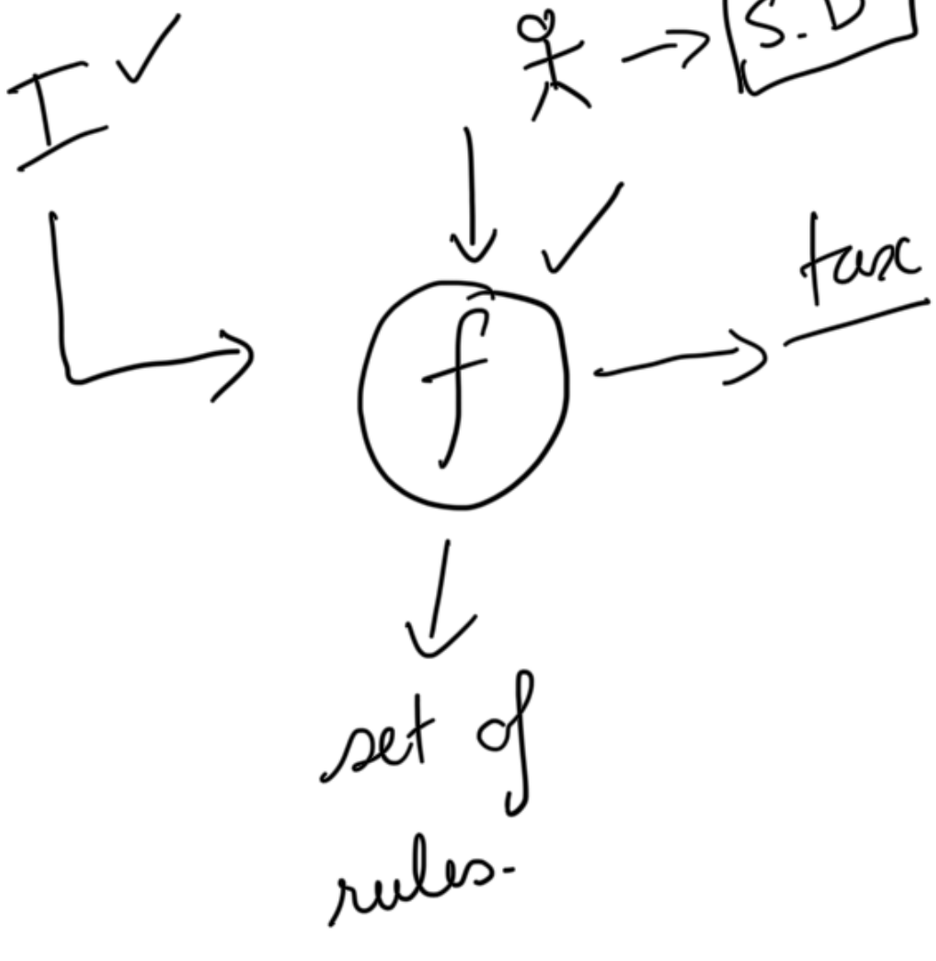
↓ actual target

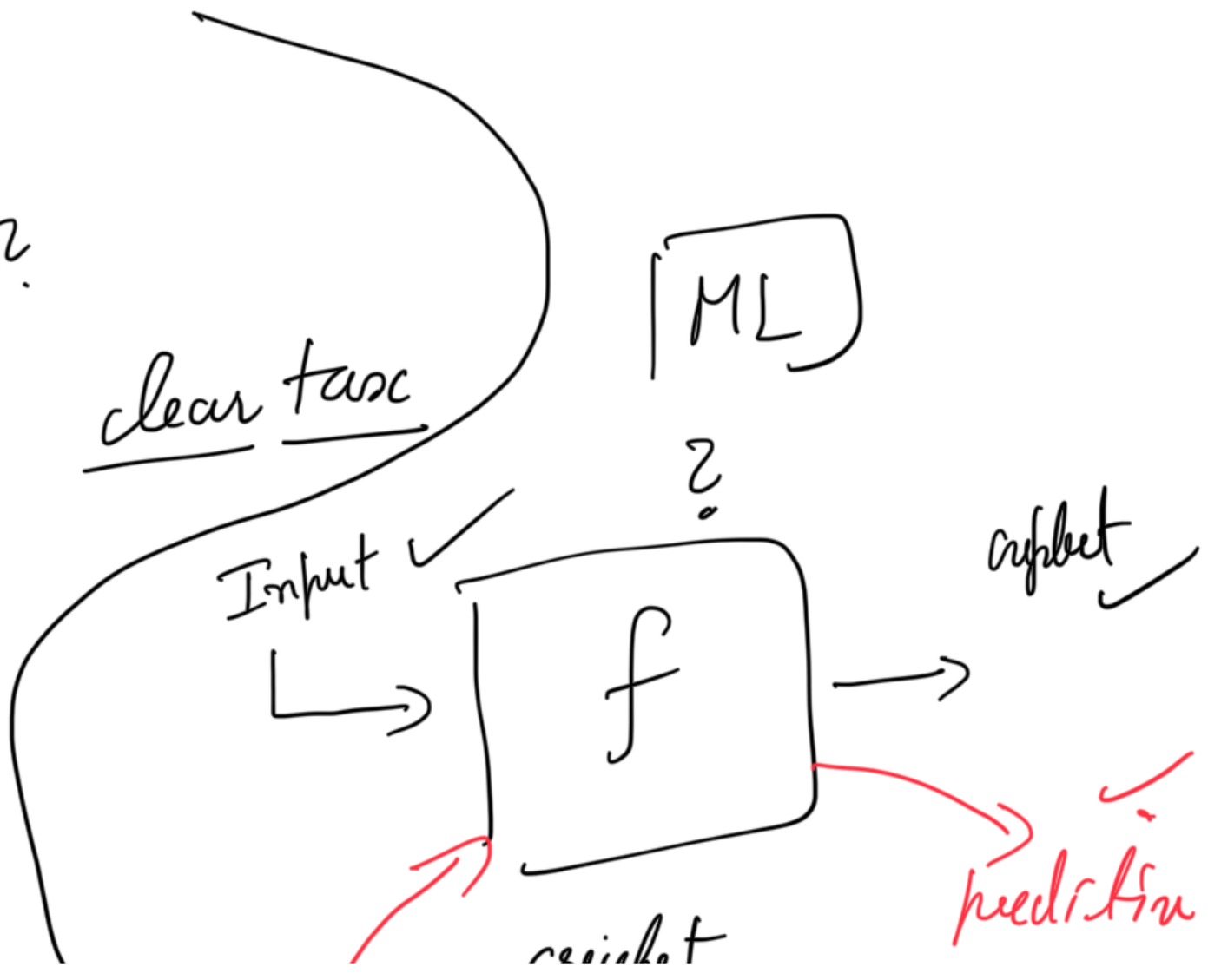
features

$f(\hat{x}_i) = \hat{y}_i$ (predicted value)

$\hat{y}_i = y_i$

new data





score prediction

new output

linear f

$f(\bar{x}) = y_i$

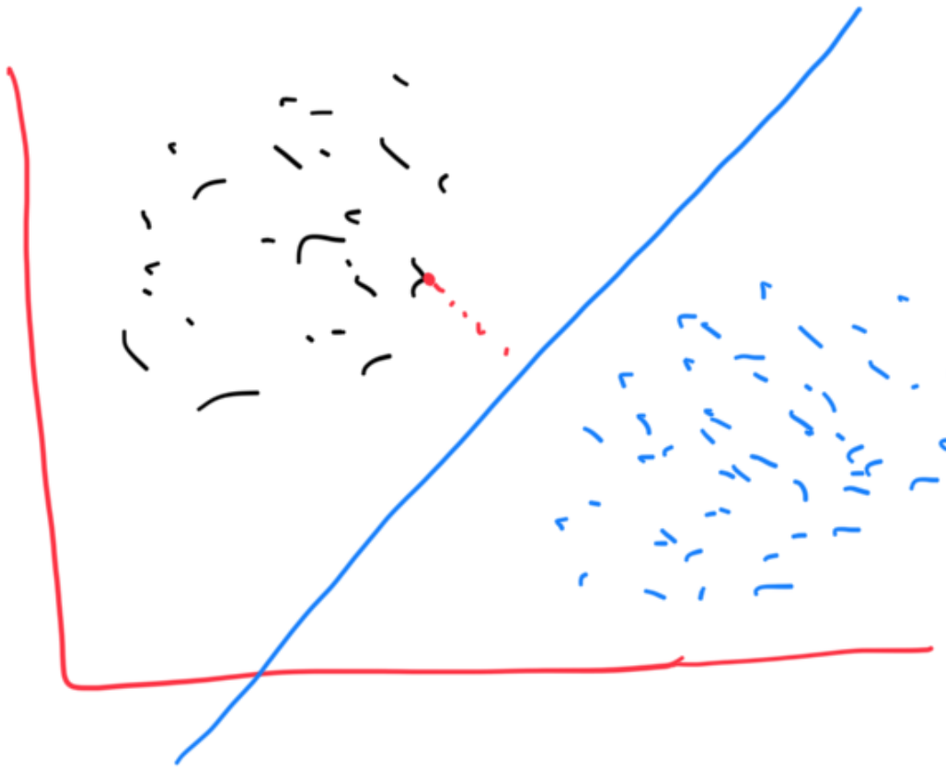
$\bar{w}^T \bar{x} + w_0$

$\rightarrow G(D, \bar{w}, w_0)$

optimize

↓

gain function (maximize dist of points from boundary)



$$g = \left(\frac{w^T x_i + w_0}{\|w\|} \right) y_i$$

$\bar{w}^*, w_0^*$ → optimal values.

$$\bar{w}^*, w_0^* = \underset{\bar{w}, w_0}{\arg\max} \left\{ G(D, \bar{w}, w_0) \right\}$$

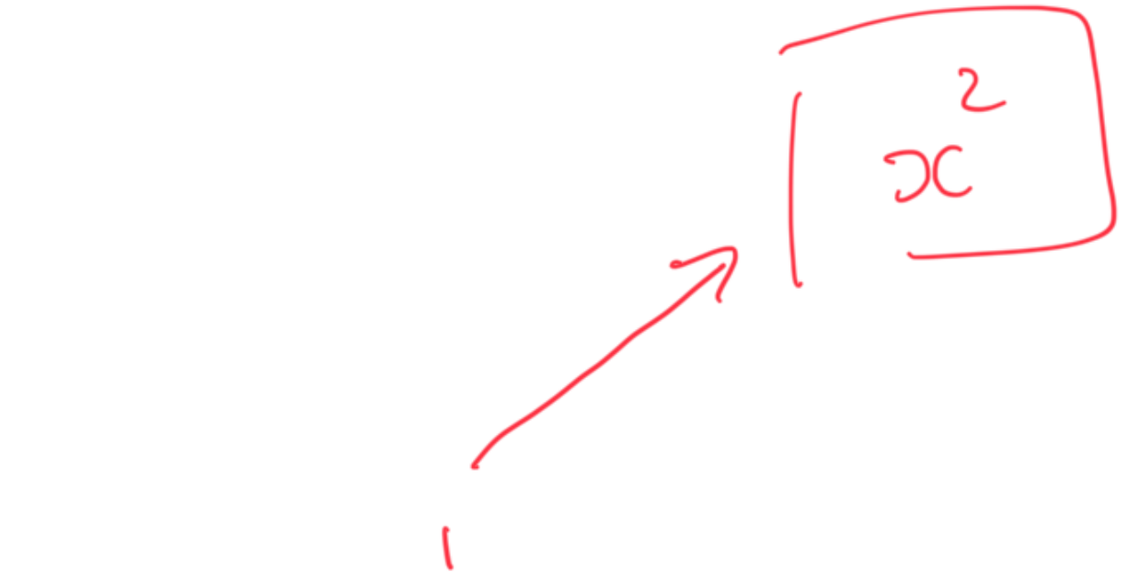
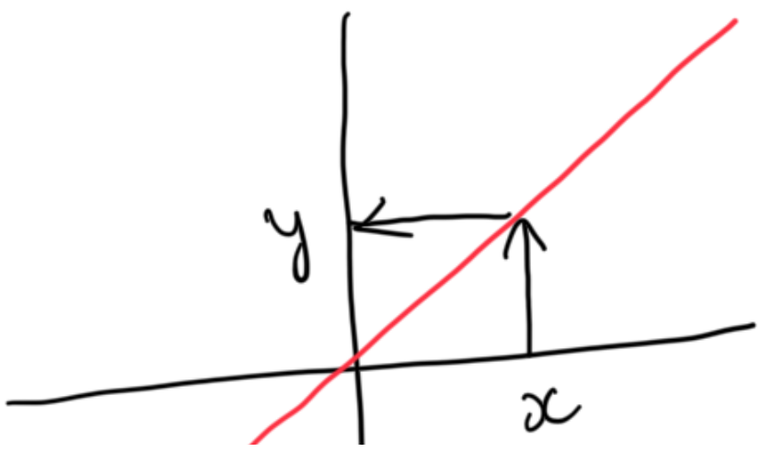
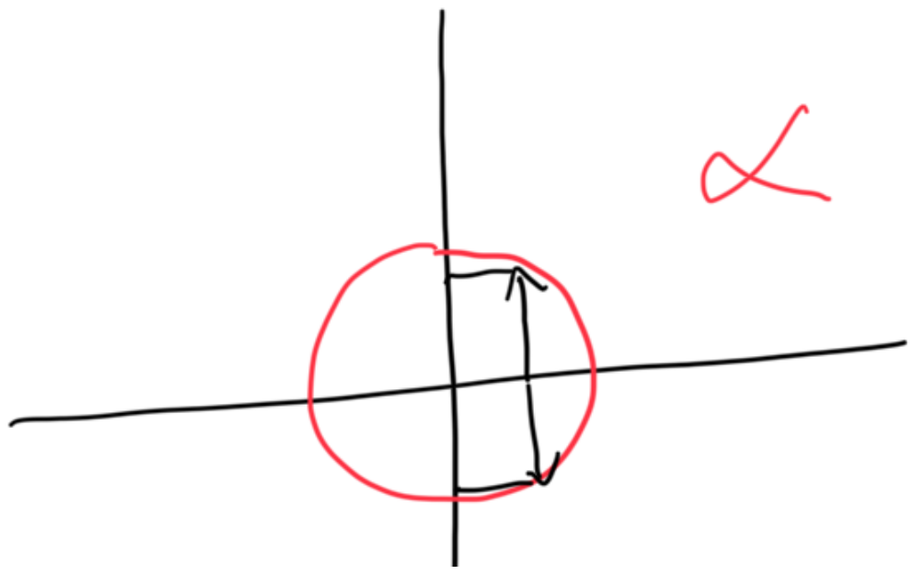
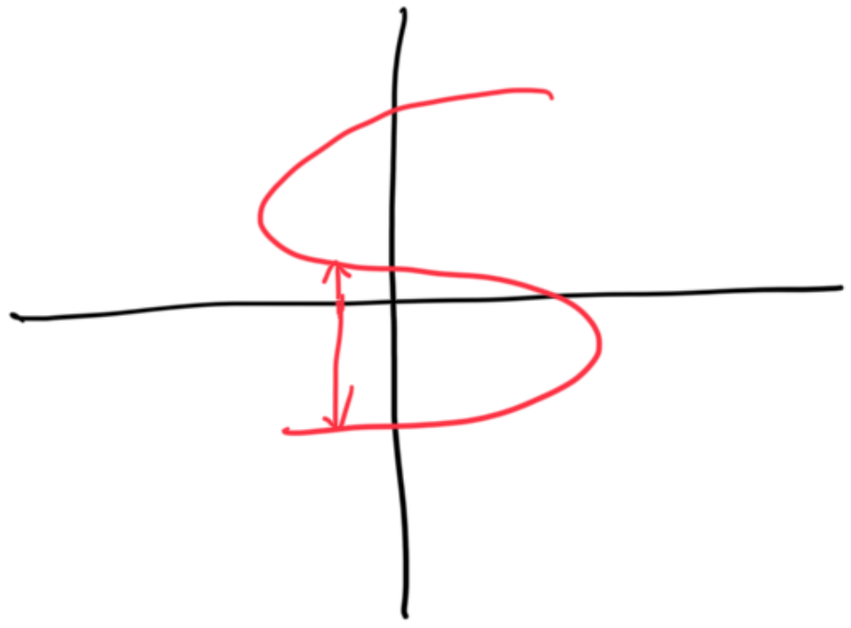
$$G(D, \bar{w}, w_0) = \sum_{i=1}^{n} g(x_i, y_i, \bar{w}_0, w_0)$$

$\downarrow$

$$\prod_{i=1}^{n} \mathcal{L}$$

---

inputs

output

an input has a unique output