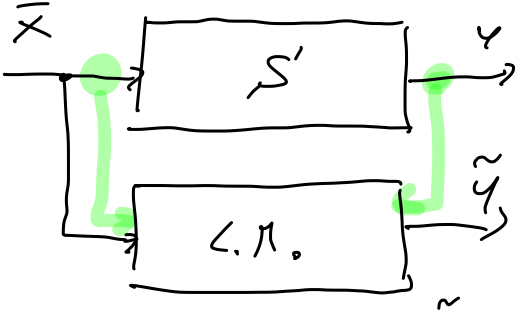
$D_m = \{(\bar{x}_1, y_1), \dots, (\bar{x}_m, y_m)\}$ $\quad y = \begin{cases} \{0,1\} \\ \mathbb{R} \\ \{1,\dots,c\} \end{cases}$

↳ i.i.d.

$f(\bar{x}) = \underline{w}\,\phi(\bar{x})$ — ϕ: LEARNED

$l(f(\bar{x}), y) = \begin{cases} \text{MSE} \\ \varepsilon\text{-INSENSITIVE} \\ \text{LINEAR} \\ \text{HINGE LOSS} \\ \vdots \end{cases}$

$C(f) = \begin{cases} \|\underline{w}\|_2 \\ \|\underline{w}\|_p \\ \int \left|\frac{d^1 f}{dx^1}\right| dx \\ \vdots \end{cases}$

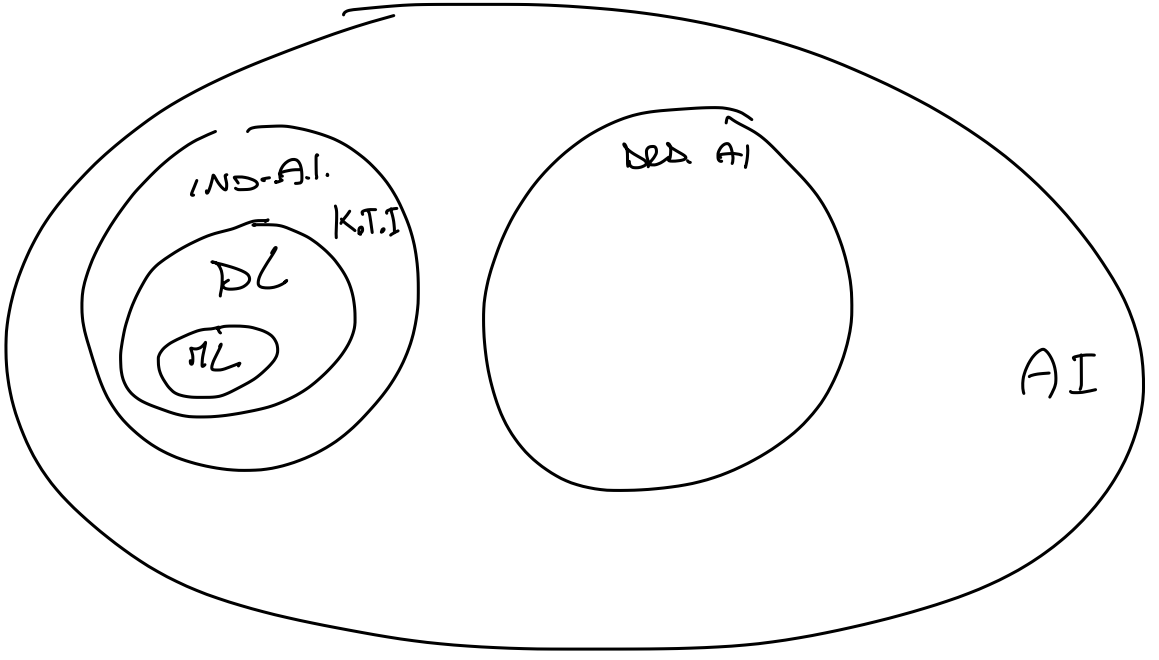


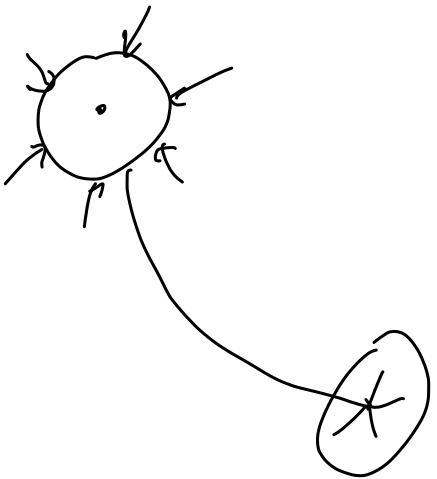
$\tilde{y} = y$

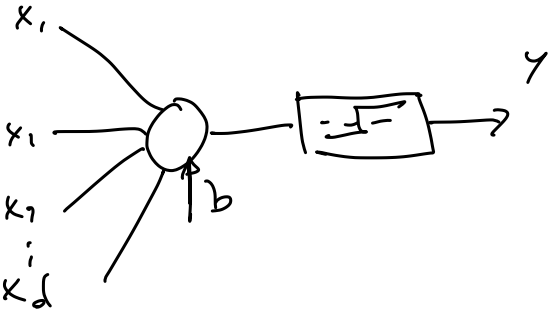
M.L. → D.L.

P. → N.P.

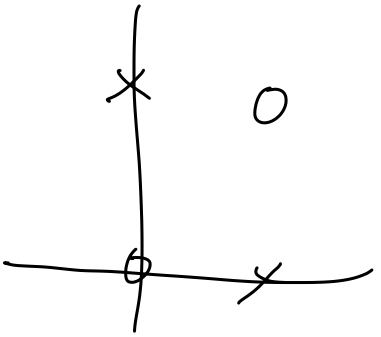
$\min_f \; \underbrace{\hat{R}(f) + \lambda C(f)}_{\leq R(f)}$

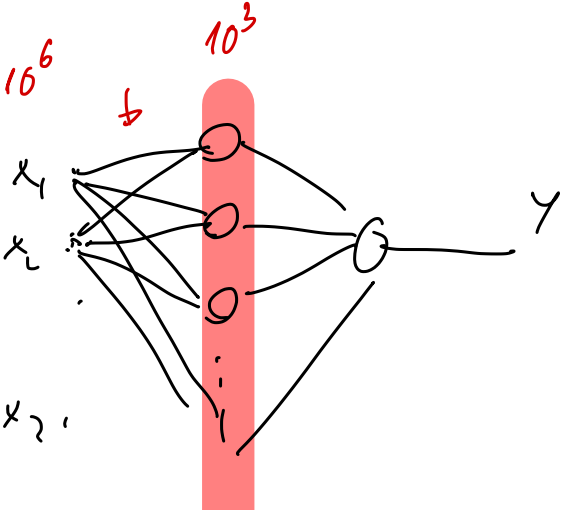

IND-A.I.
K.T.I
DL
ML
AI





$$y = \sigma(\underline{w}\,\underline{x} + b)$$

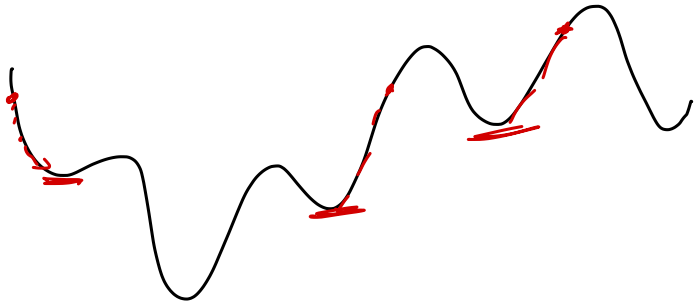




$10^6 \cdot 10^3$

10^9

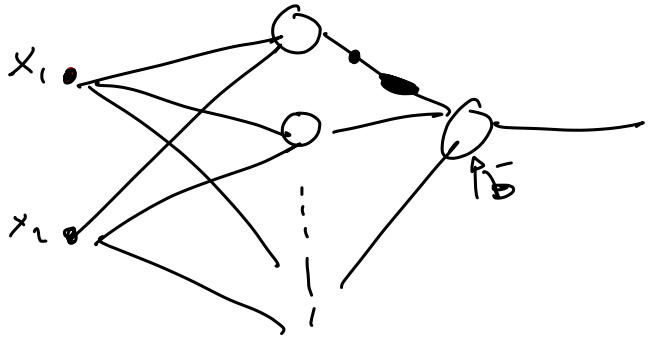
$$y = \underline{w} \cdot \sigma(W x + b) + b$$



$$= \underline{w}\ \phi(x)$$

$$\min_{\underline{w}, W, b, \ldots} \sum_{i=1}^{n} \left(f_{\underline{w}, W, b, \ldots}(x_i) - y_i \right)^2 + \lambda \left(\|\underline{w}\| + \|W\| \right)$$

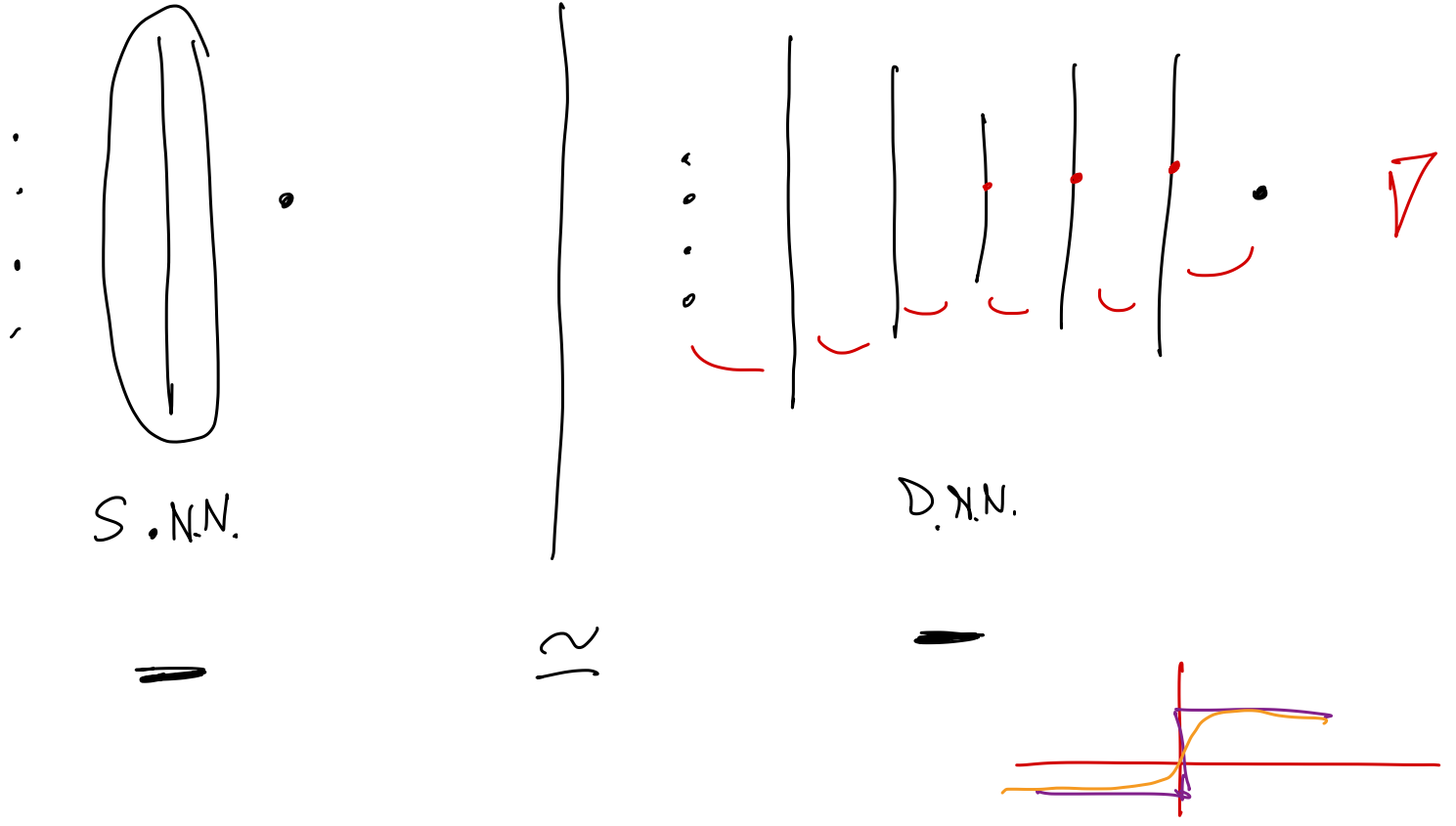
$$\sum_{i=1}^{n} \nabla f(x)$$

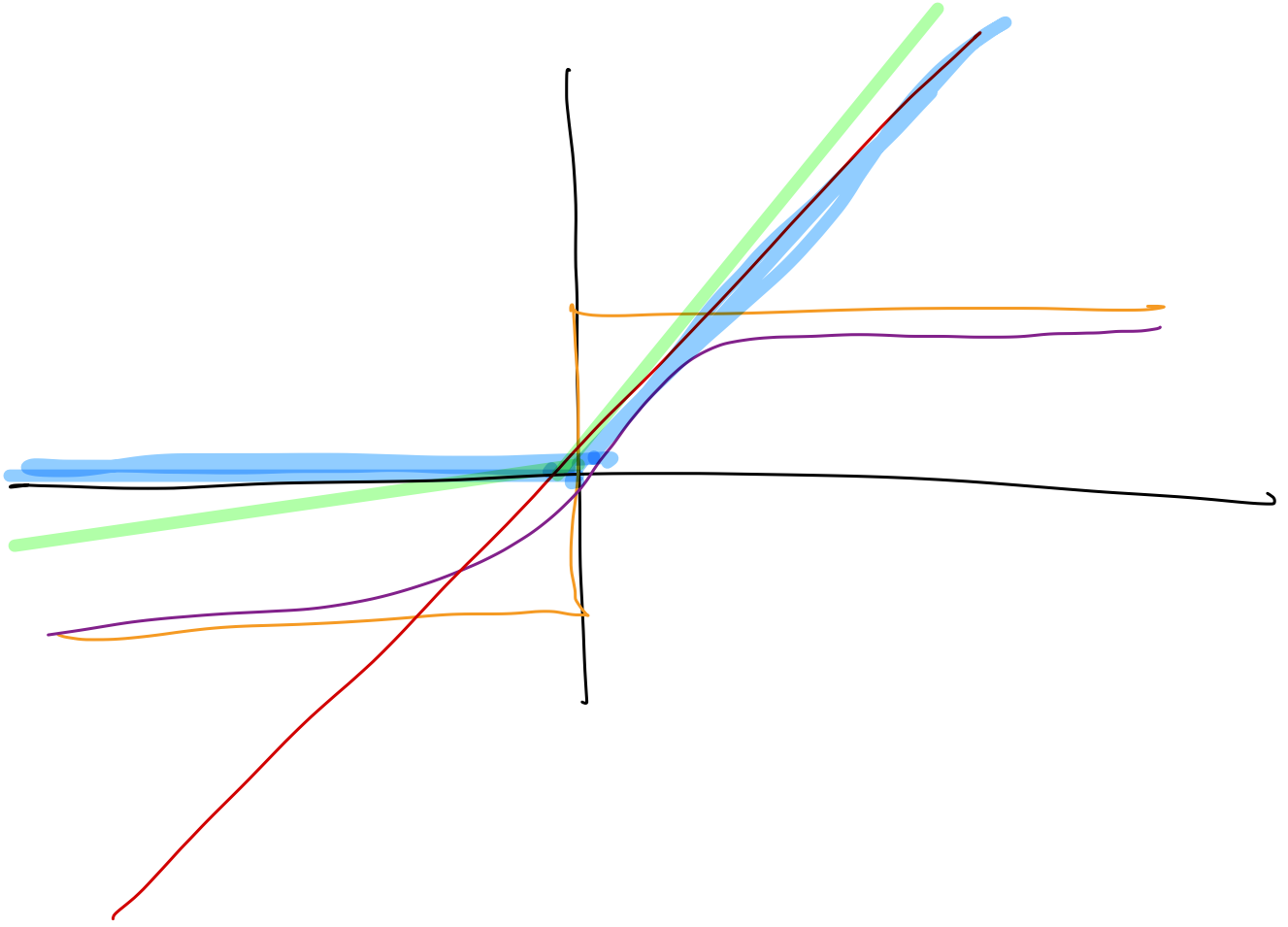


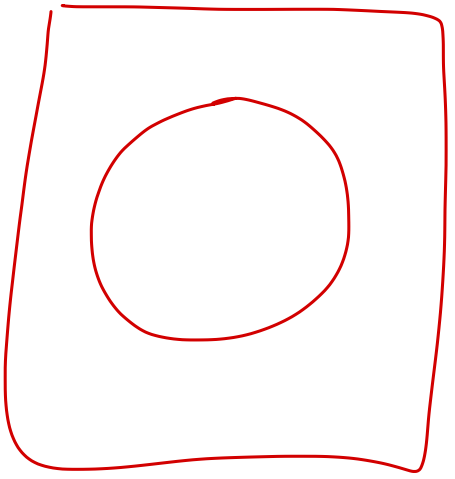
$$\bar{w}_1(w_{11}x_1 + w_{12}x_2 + b_1) +$$
$$\bar{w}_2(w_{21}x_1 + w_{11}x_1 + b_2) +$$
$$\vdots$$
$$+ \bar{b}$$

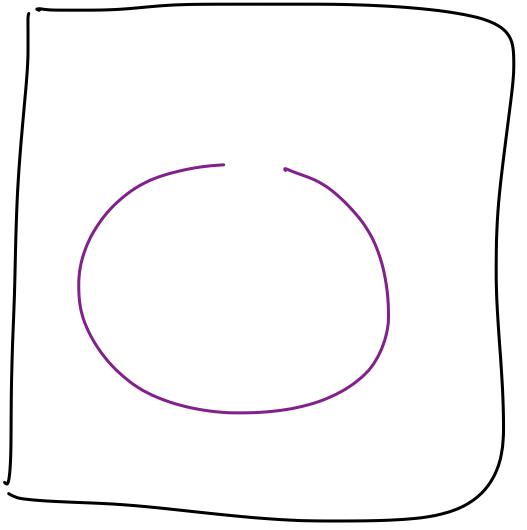
$$= \left(\bar{w}_1 w_{11} + \bar{w}_2 w_{21} + \dots\right)x_1 + \left(\bar{w}_2 w_{12} + \bar{w}_2 w_{12} + \dots\right)x_2 + \left(b_1\bar{w}_1 + b_2\bar{w}_{11}\dots\right) + \bar{b}$$

$$= w_1 x_1 + w_2 x_2 + b = \bar{w} * + b$$

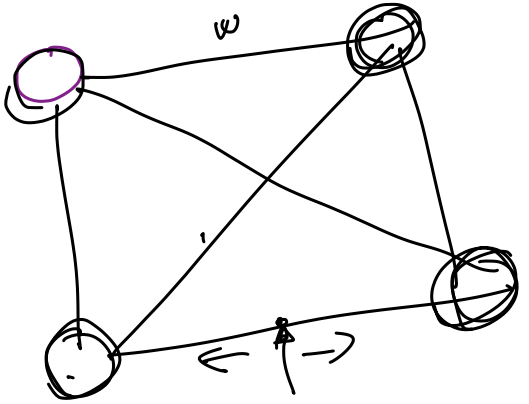

S.N.N.
D.N.N.
GRADIENT
NANI STYING



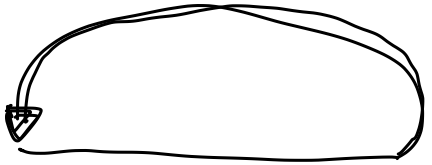




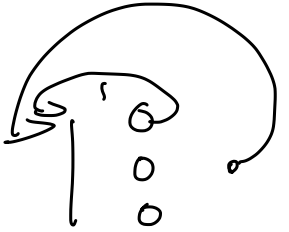
$$\int (y) x$$

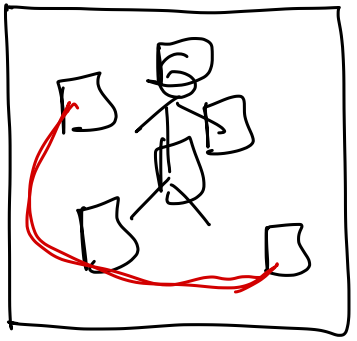


ABDUCTION



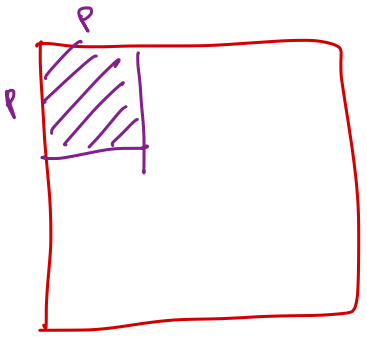
R.N.N.

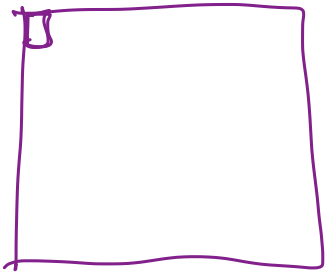


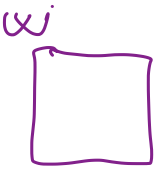


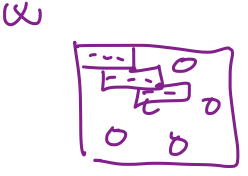
CONVOLUTION

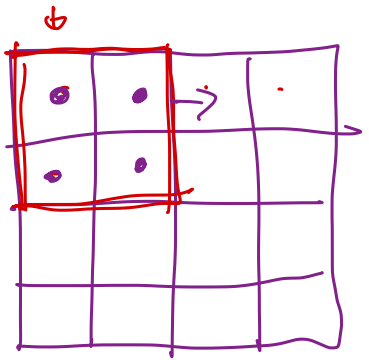
ATTENTION

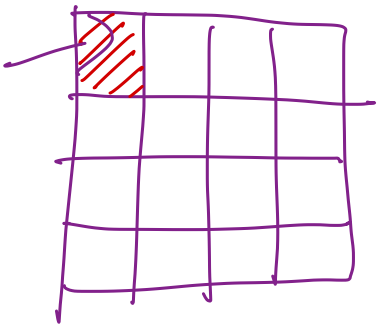




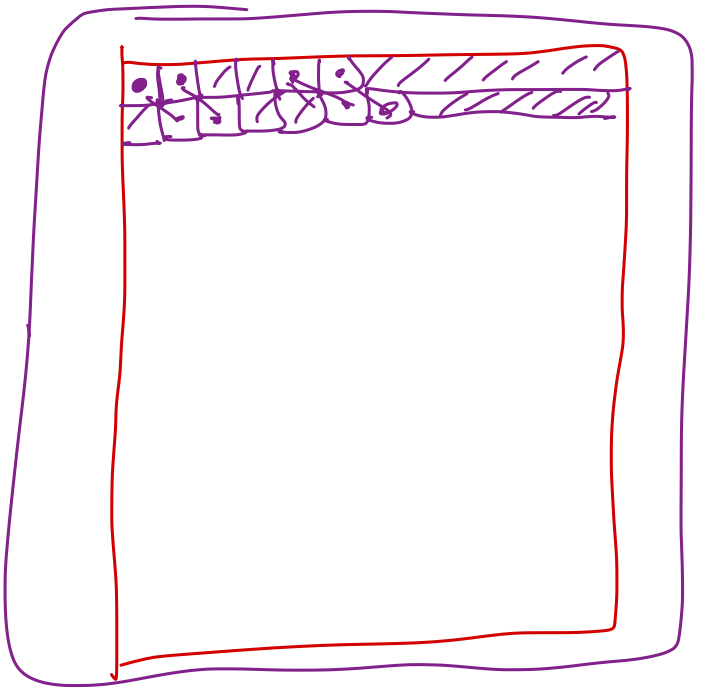


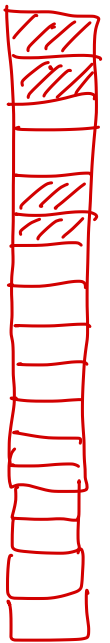


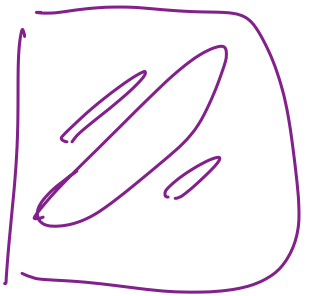




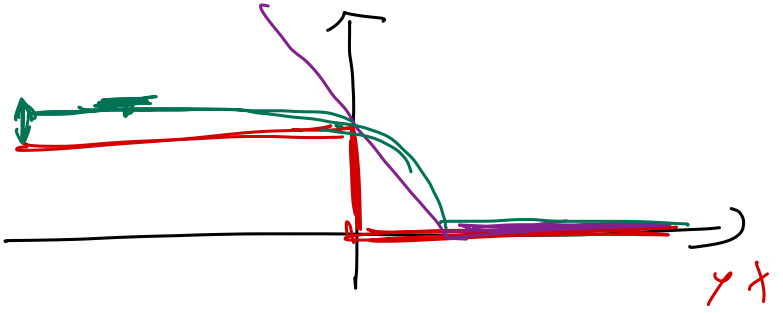
$$\sigma(W \times + b)$$



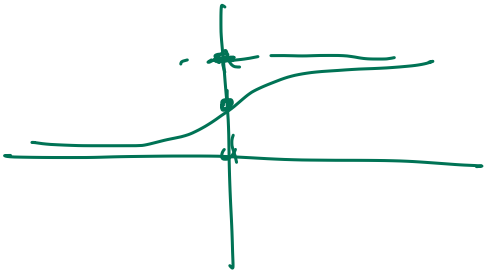


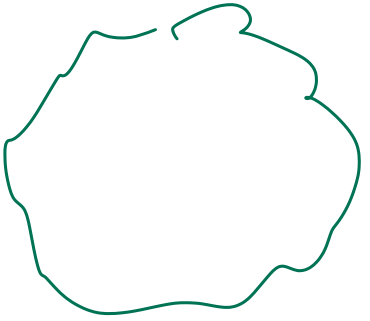


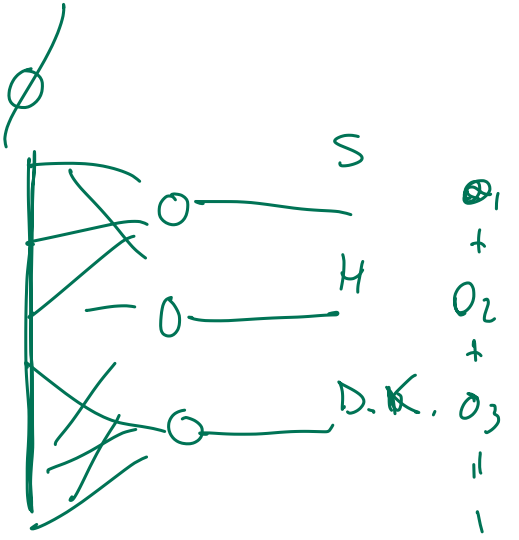
losses -> DIFF.

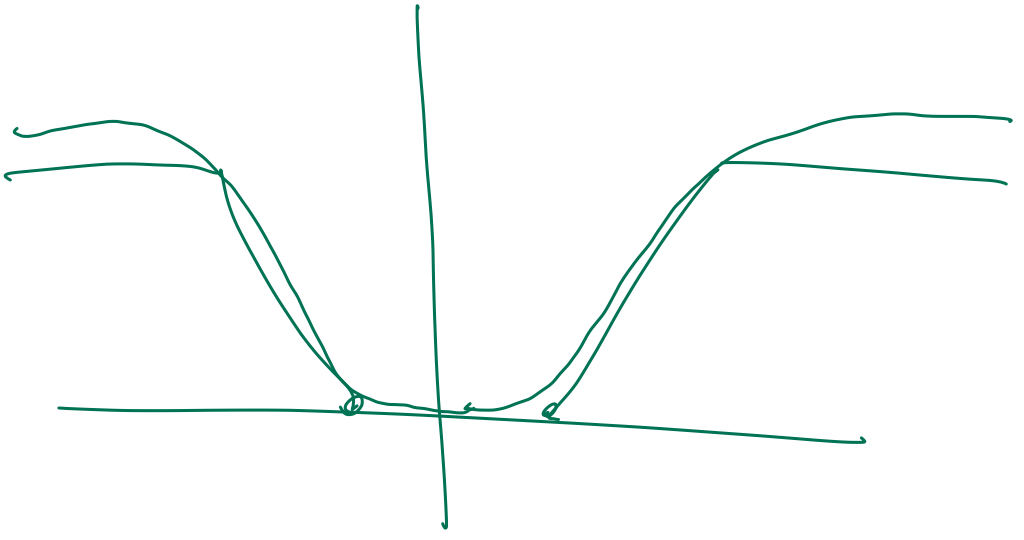


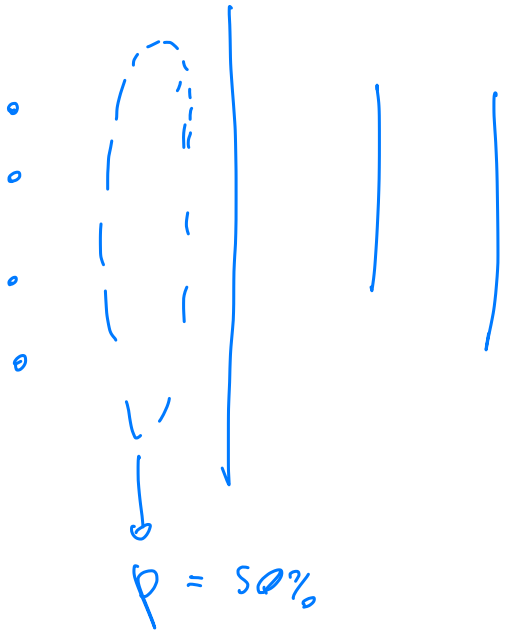
SOFT MAX



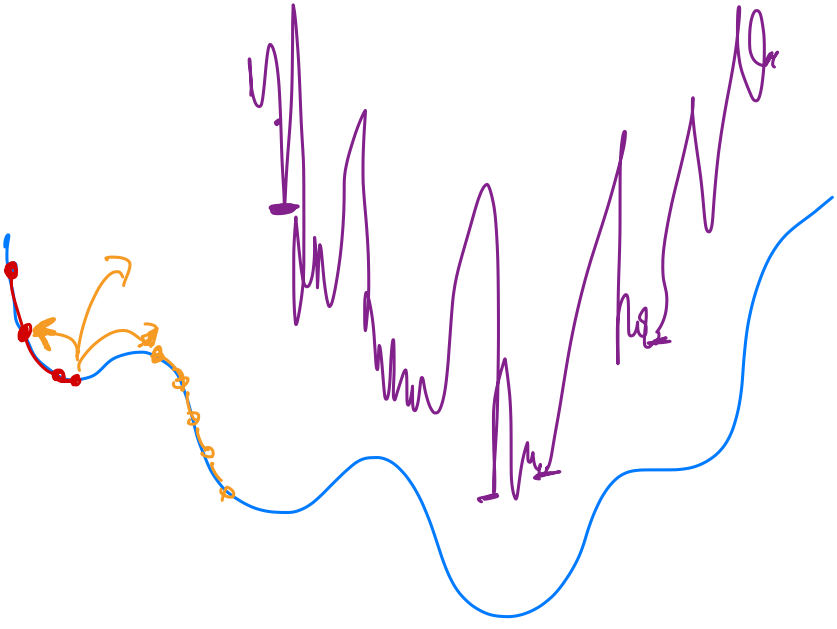








DROP OUT



$\lambda \; \underline{w}' \; B \; \underline{w}$

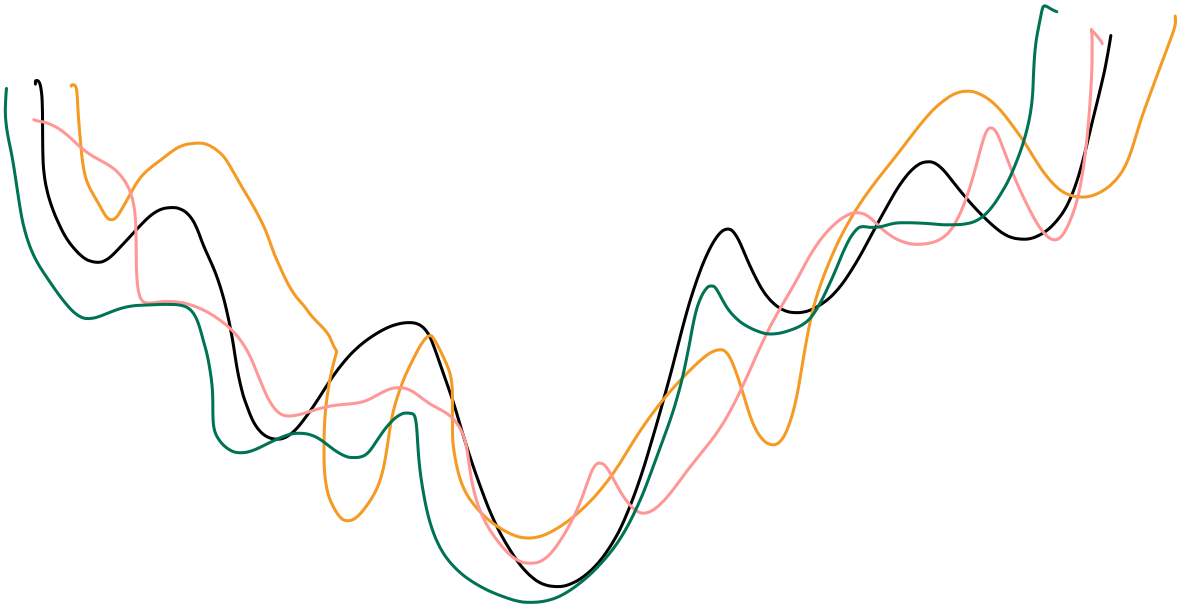
$B = I \rightarrow$ Ridge

$$w_{i+1} = w_i - \gamma \nabla_w^{obs} \Big|_{w_i}$$

L.R.

A.D.M.

S. G. D.



$$\min \quad \hat{R}(f) + \lambda\, c(f)$$

$$\hat{R}(f) = \sum_{n=1}^{P_m}$$

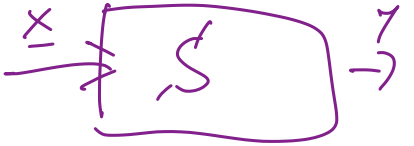
EFFICIENCY

$$f(\underline{x}) = \underline{\alpha}\ \underline{\sigma}\left(\underline{\omega}\,\underline{x} + b\right) + \bar{b}$$

ELR.

E.S.N.

INFORMING



$$\min_{y} \quad \hat{R}(y) + \lambda' C(y) + \lambda^2 (y - \hat{y})^2$$

$$\rightarrow \left\{ \frac{dy}{dt} = v_s \right\}$$