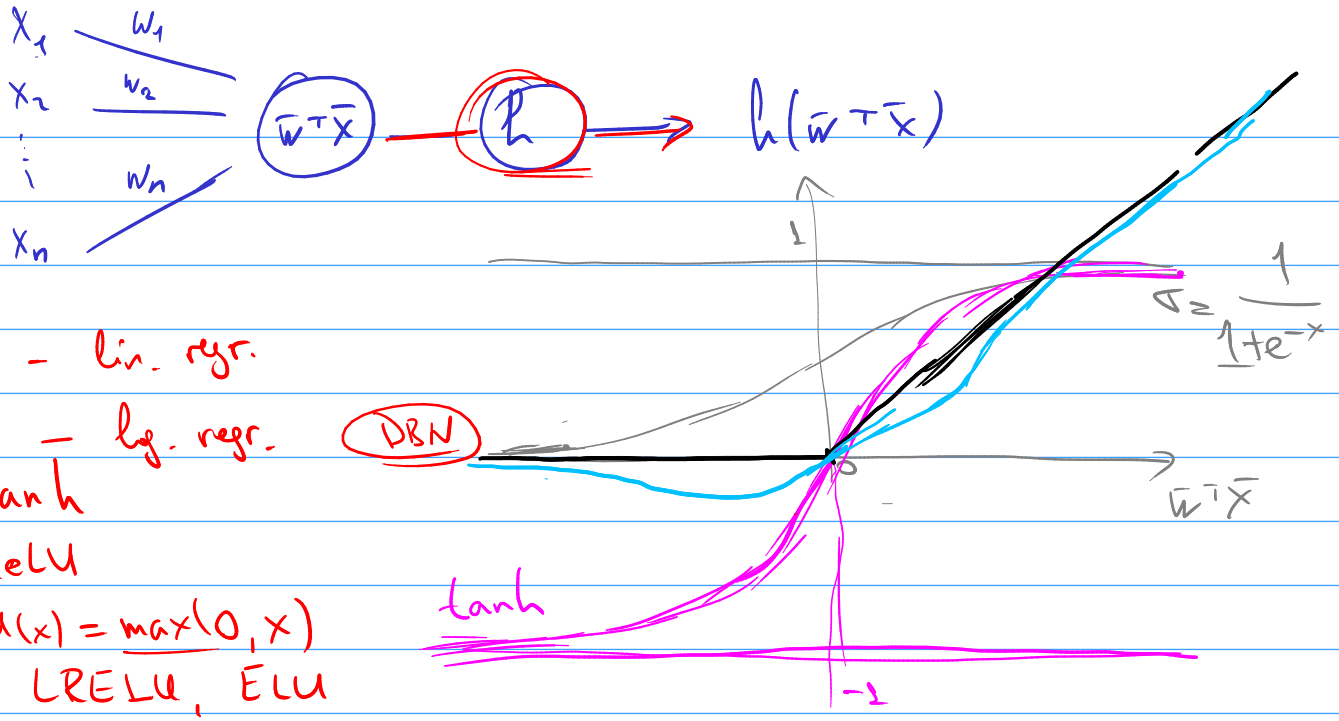

1) $h = id$ – lin. regr.

2) $h = \sigma$ – lg. regr.

3) $h = \tanh$

4) $h = ReLU$

$ReLU(x) = \max(0, x)$

5) $h = LRELU$, ELU

6) NAS – neural arch. search

$h = Swish(x) = x \cdot \sigma(\beta x) = \frac{x}{1+e^{-\beta x}}$

$h = Mish(x)$

7) $S_\beta(x_1 \dots x_n) = \frac{\sum x_i e^{\beta x_i}}{\sum e^{\beta x_i}}$ (September 2020)

$S_\beta(g(x), h(x)) = g(x) \frac{e^{\beta g(x)}}{e^{\beta g(x)} + e^{\beta h(x)}} + h(x) \frac{e^{\beta h(x)}}{- - -} = \dots$

$\dots = (g(x) - h(x)) \sigma(\beta(g(x) - h(x))) + h(x)$

$g(x) = x, \quad h(x) = 0$:

hard max

$ReLU(x) = \max(x, 0)$

$S_{\beta 1}$

$Swish(x) = x \cdot \sigma(\beta x)$

$g(x) = x \quad h(x) = ax$

LReLU

$(1-a)x \cdot \sigma(\beta(1-a)x) + ax$