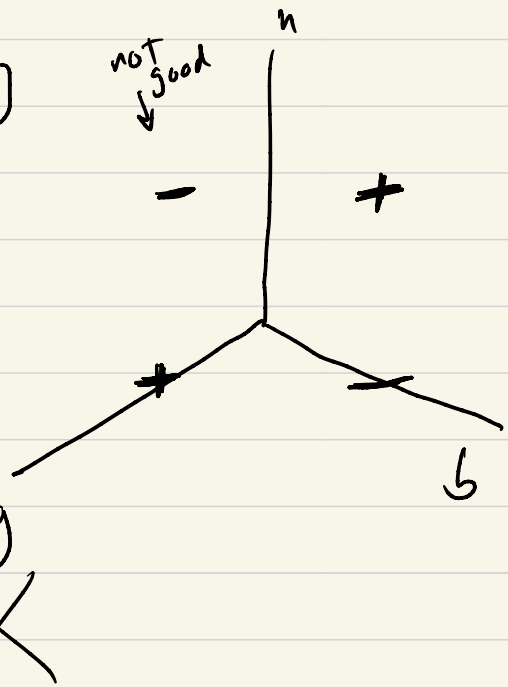
Neural Nets

$f(\bar{x}) = [\, g \;\; b \;\; n \,]$

good	1	0	0
bad	0	1	0
not good	1	0	1
not bad	0	1	1



One solution: bigrams

Our solution: nonlinear classifier (NN)

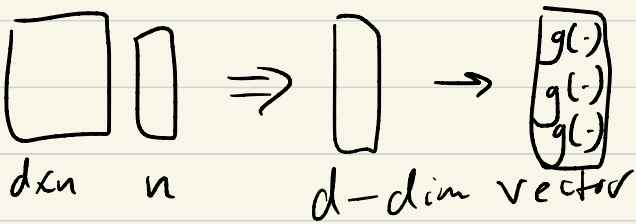
$\bar{w}^T \hat{y} f(\bar{x})$ for MC Log. Reg.

replace $f(\bar{x})$ w/a nonlinear function

intermediate features

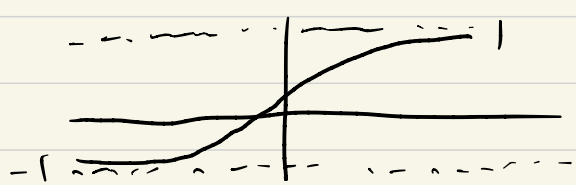
① $\bar{z} = g(V f(\bar{x}))$

nonlinearity, $d \times n$ matrix, BoW feats, n-dim vector

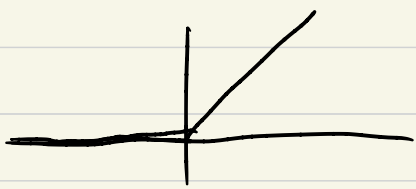


Final step: $\hat{y} = \text{argmax}_{\hat{y}} \bar{w}_{\hat{y}}^T \bar{z}$

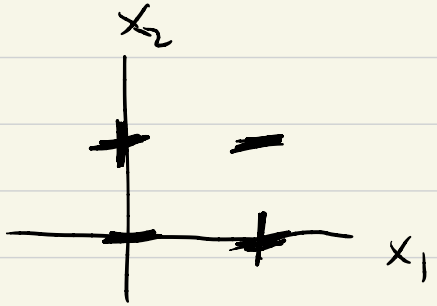
g: tanh



ReLU



Ex



x_1 XOR x_2

I made this up ↓

$$V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

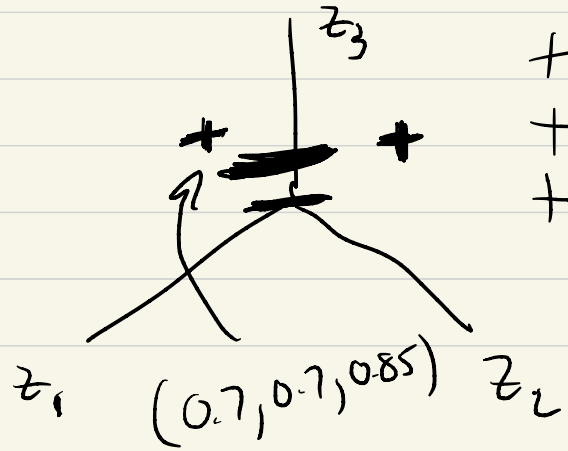
3×2 mat.

$$f(\bar{x}) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$g = \tanh$

$$Vf(\bar{x}) = \begin{bmatrix} x_1 & x_2 & x_1 + x_2 \end{bmatrix}$$

$$\bar{z} = \begin{bmatrix} \tanh(x_1), & \tanh(x_2), & \tanh(x_1+x_2) \end{bmatrix}$$



$\tanh(0) = 0$

$\tanh(1) \approx 0.7$

$\tanh(2) \approx 0.85$

Weights: V

$\overline{w}_{\hat{y}}$