Activity: MLP Structure
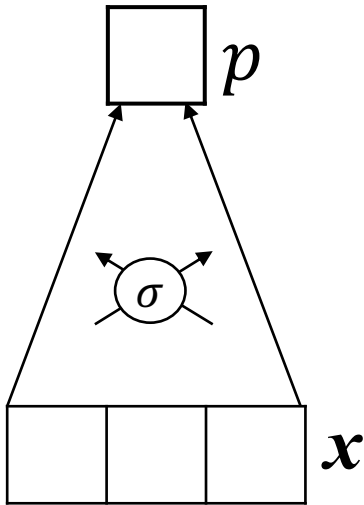
ML for Health, Week 3

ANSWER KEY

Logistic Regression
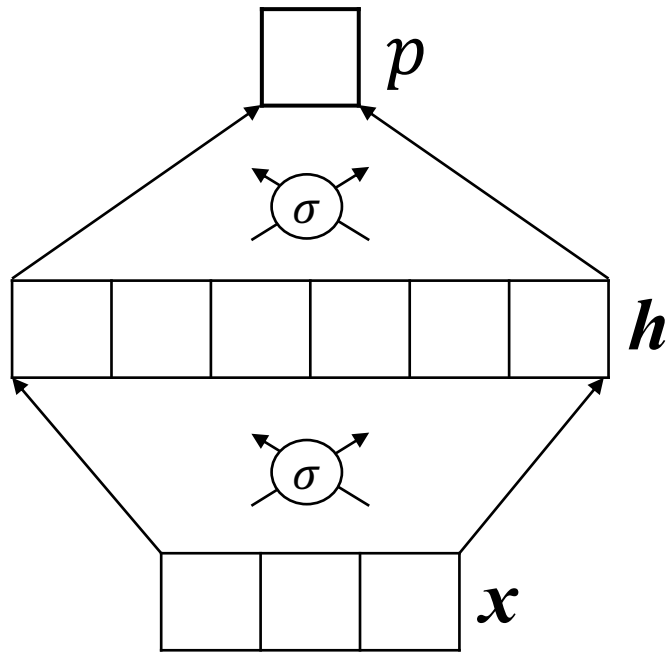
- 0 hidden layers

- 3 + 1 = 4 parameters

- $p = \sigma(Wx + b)$

p

σ

x

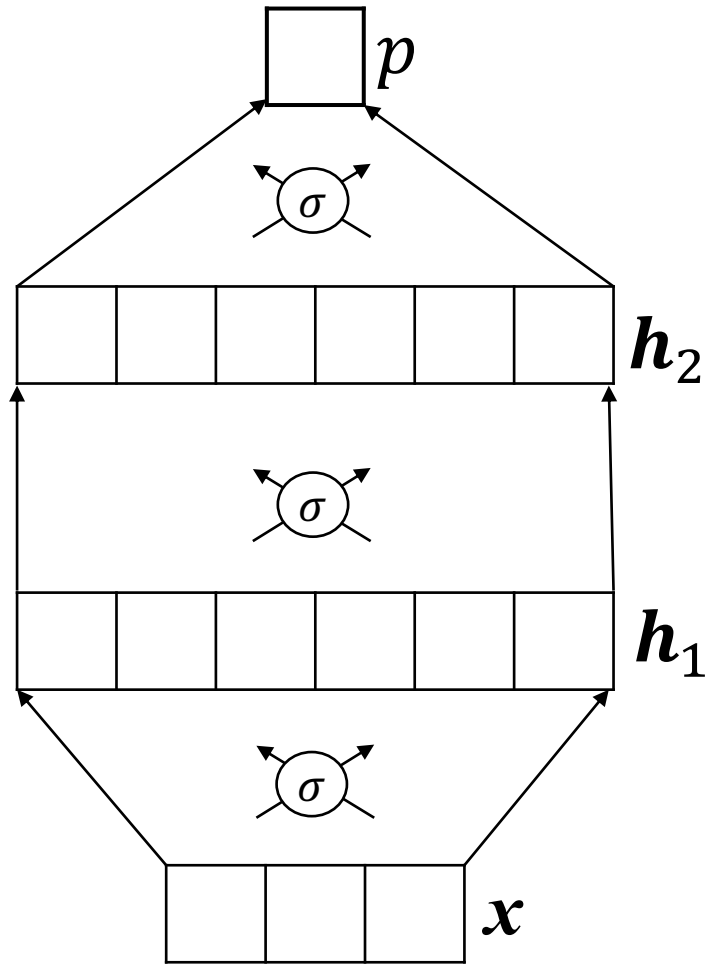# Shallow MLP
(with σ activations)

- 1 hidden layer of size 6

- (3 x 6) + 6 + (6 x 1) + 1 = 31 parameters

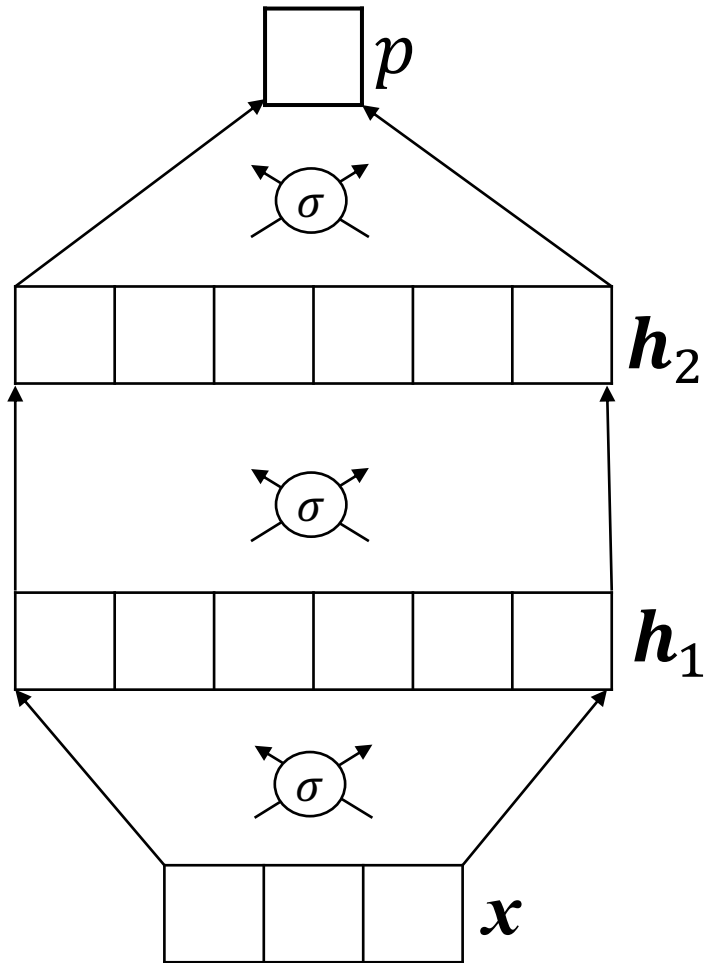- $p = \sigma(W_2 * \sigma(W_1\boldsymbol{x} + \boldsymbol{b}_1) + \boldsymbol{b}_2)$

Slightly Deeper MLP
(with σ activations)

- 2 hidden layers, each of size 6

- (3 x 6) + 6 + (6 x 6) + 6 + (6 x 1) + 1 = 73 parameters

- $p = \sigma(W_3 * \sigma(W_2 * \sigma(W_1 x + b_1) + b_2) + b_3)$

Which model contains more parameters?

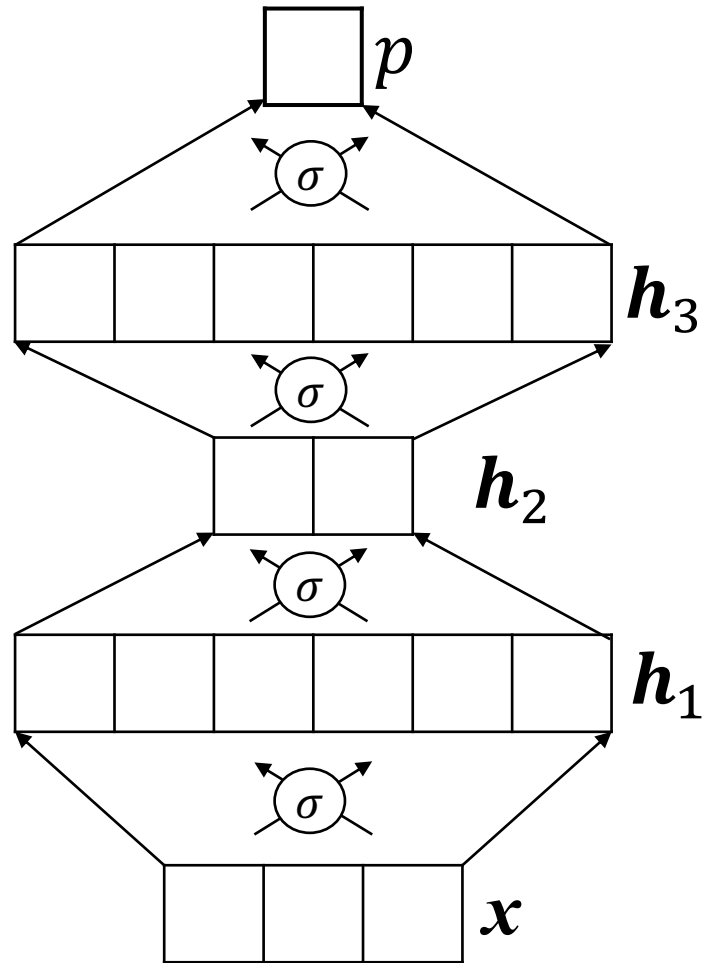$(6 \times 1) + 1$

$+$
$(6 \times 6) + 6$

$+$
$(3 \times 6) + 6$

$= 73$ params

$(6 \times 1) + 1$

$+$
$(2 \times 6) + 6$

$+$
$(6 \times 2) + 2$

$+$
$(3 \times 6) + 6$

$= 63$ params

- Effect of x_1 on p (LogReg): you need to know the value of b_1, the parameter associated with x_1; and since σ is non-linear, you also need to know the value of z

- Effect of $h_{2,1}$ on p (MLP): similar to the previous answer, you need to know the parameter connecting $h_{2,1}$ to z and $p = \sigma(z)$; as well as the value of z

- Effect of x_1 on p (MLP): here you need to know (a) the values of all parameters connecting x_1 to $\boldsymbol{h}_1$; (b) the current value of $\boldsymbol{h}_1$; (c) the values of all parameters connecting $\boldsymbol{h}_1$ to $\boldsymbol{h}_2$; (d) the current value of $\boldsymbol{h}_2$; (e) the values of all parameters connecting $\boldsymbol{h}_2$ to z; and the current value of z. In short, you need to know all current parameters and all current values in all hidden layers, because x_1 exerts an effect on p through all possible paths connecting x_1 to p in the graph at right. In contrast, in the graph at left, there is only one path connecting x_1 to p.