

# CS/CE 316/365 Deep Learning

## Activity 02

September 2, 2025

### Shallow Neural Networks

Activity needs to be handwritten.

1. What kind of mapping from input to output would be created if the activation function in equation given below was linear so that  $a[z] = \psi_0 + \psi_1 z$ ? What kind of mapping would be created if the activation function was removed, so  $a[z] = z$ ?

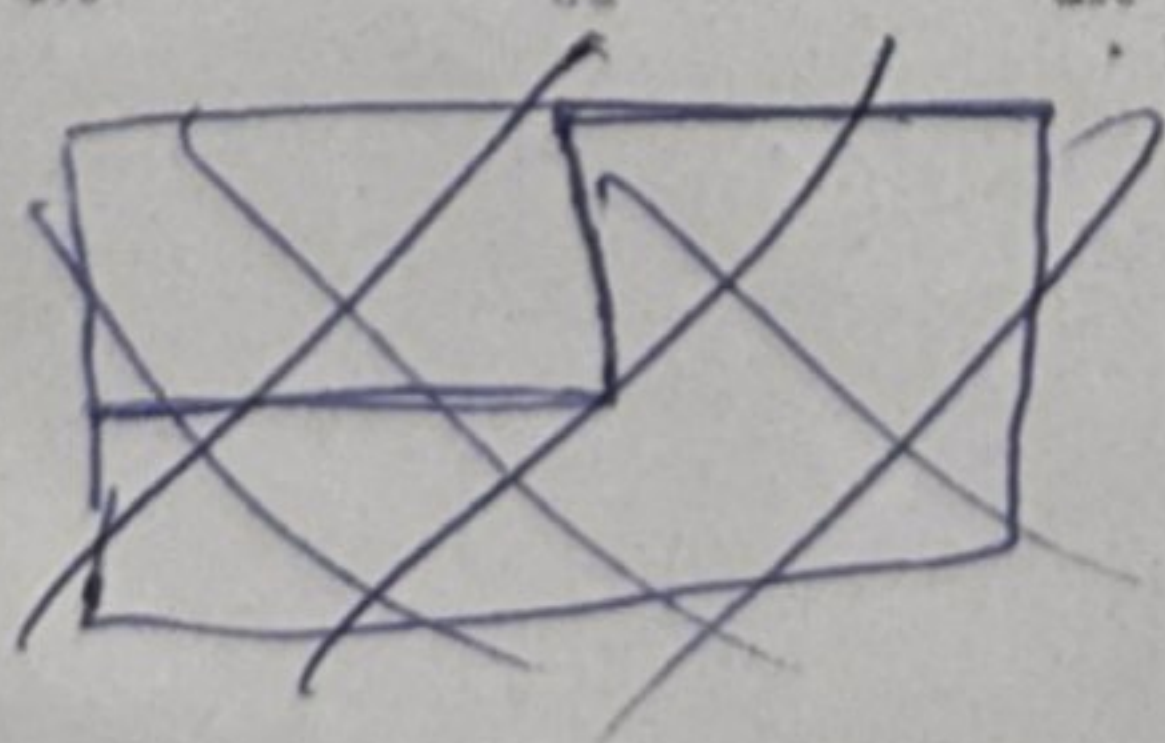
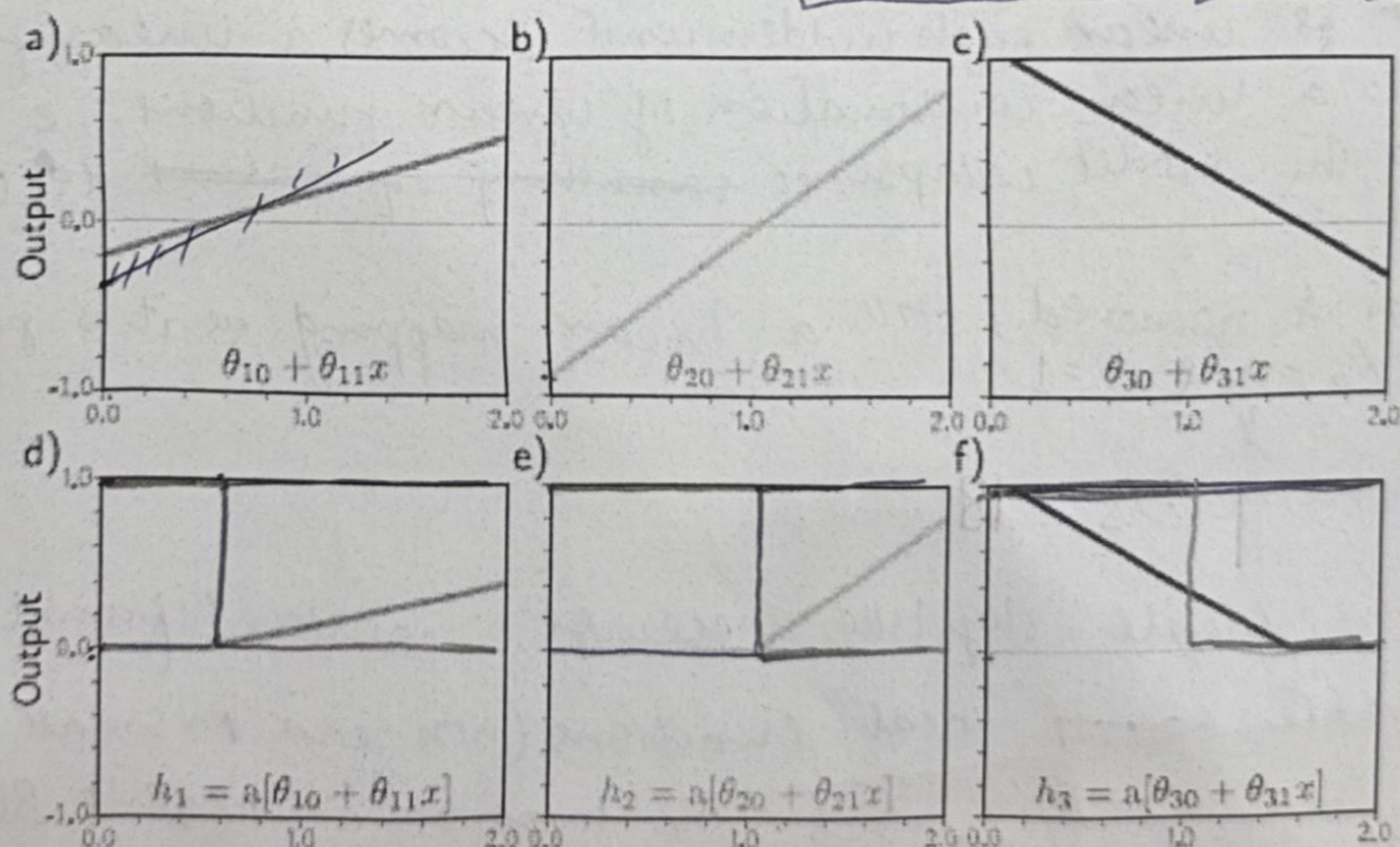
$$y = f[x, \phi] \\ = \phi_0 + \phi_1 a[\theta_{10} + \theta_{11}x] + \phi_2 a[\theta_{20} + \theta_{21}x] + \phi_3 a[\theta_{30} + \theta_{31}x]$$


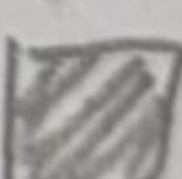
2. Consider replacing the ReLU activation function with (i) the Heaviside step function heaviside  $[z]$ , (ii) the hyperbolic tangent function  $\tanh[z]$ , and (iii) the rectangular function rect  $[z]$ , where:

$$\text{heaviside}[z] = \begin{cases} 0 & z < 0 \\ 1 & z \geq 0 \end{cases} \quad \text{rect}[z] = \begin{cases} 0 & z < 0 \\ 1 & 0 \leq z \leq 1 \\ 0 & z > 1 \end{cases}$$

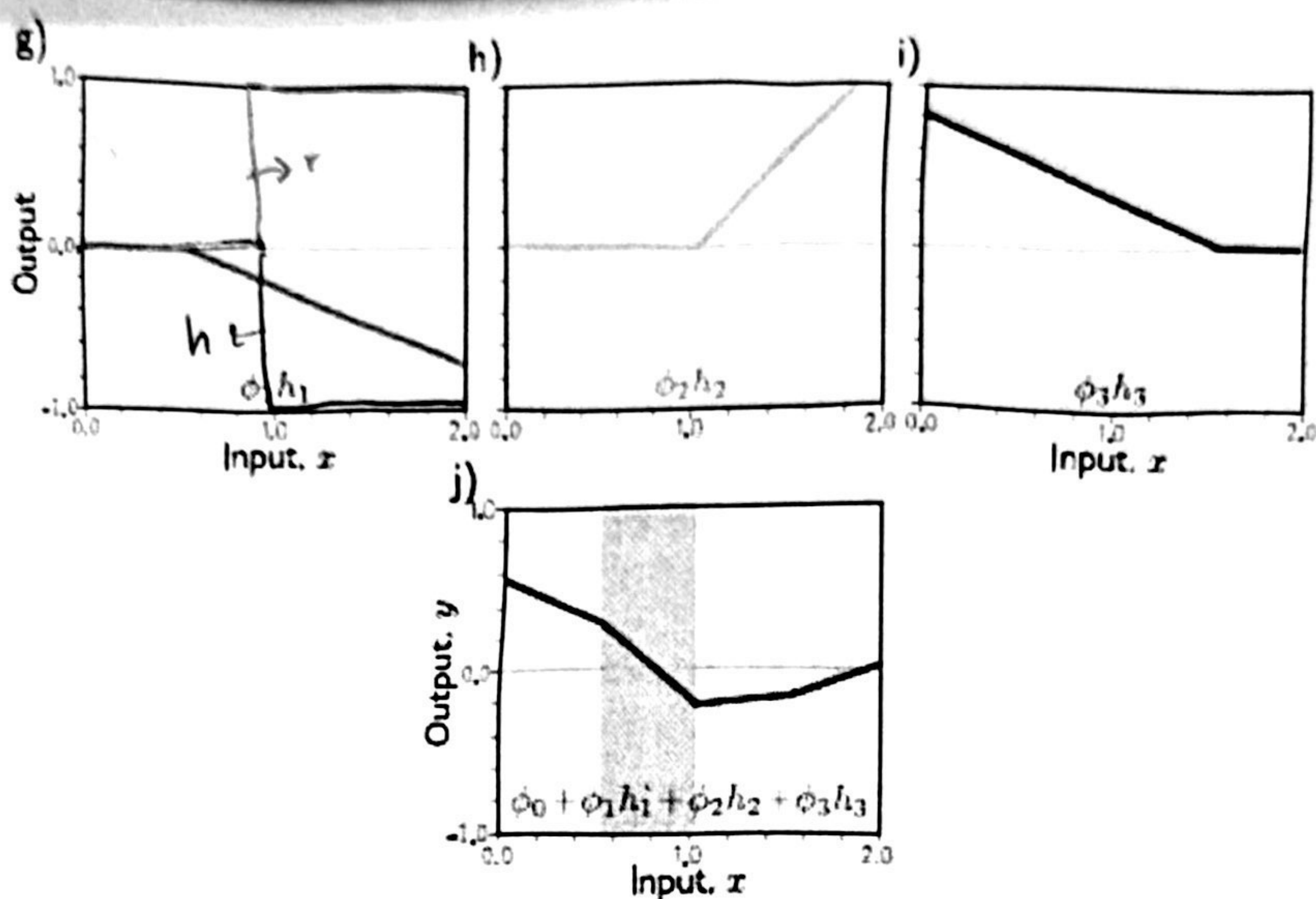
Look at figure given below (divided in 2 parts) which is a version of what was shown in class. Redraw a version of this figure for each of these functions. The original parameters were:

$$\phi = \{\phi_0, \phi_1, \phi_2, \phi_3, \theta_{10}, \theta_{11}, \theta_{20}, \theta_{21}, \theta_{30}, \theta_{31}\} = \{-0.23, -1.3, 1.3, 0.66, -0.2, 0.4, -0.9, 0.9, 1.1, -0.7\}.$$



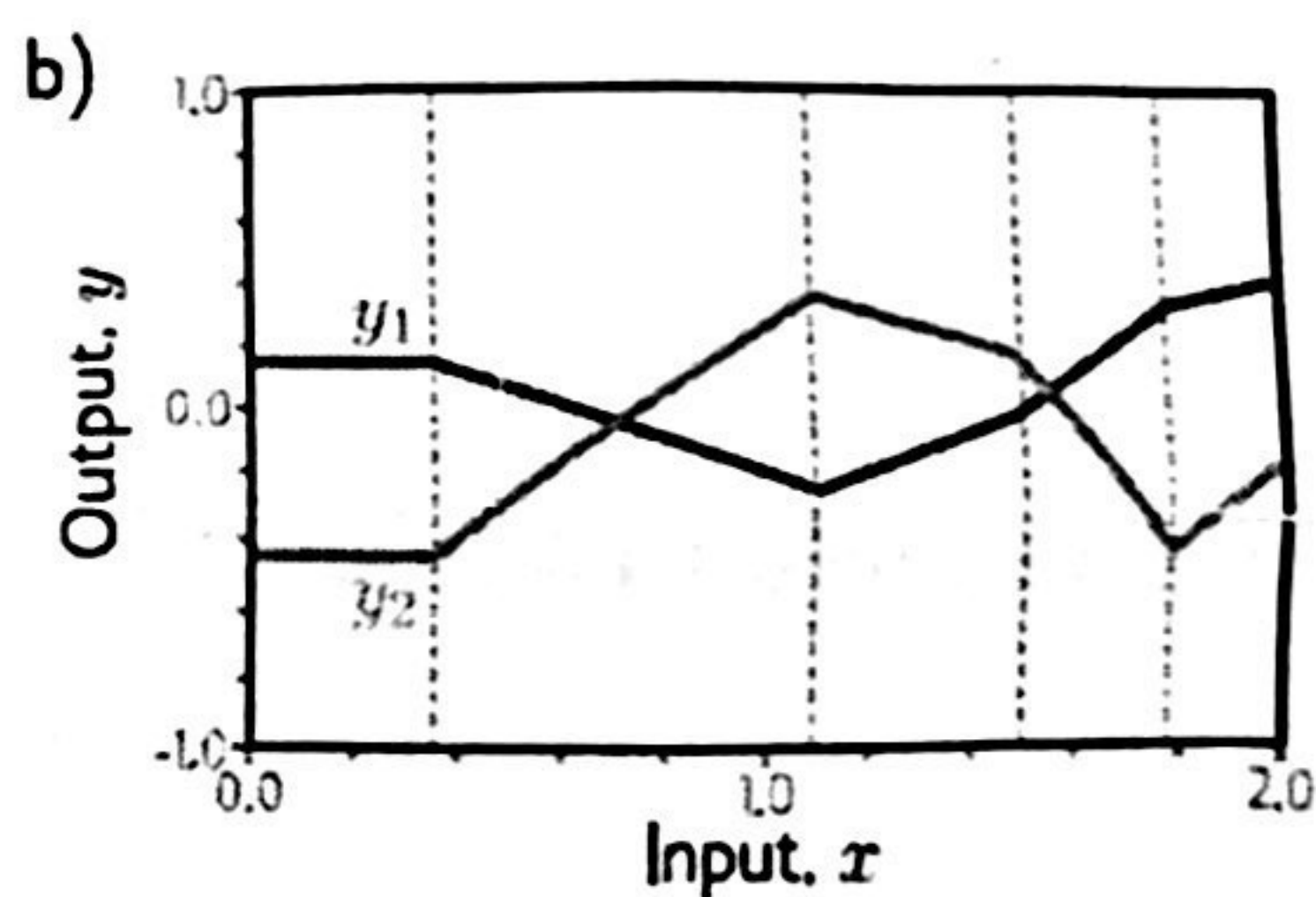
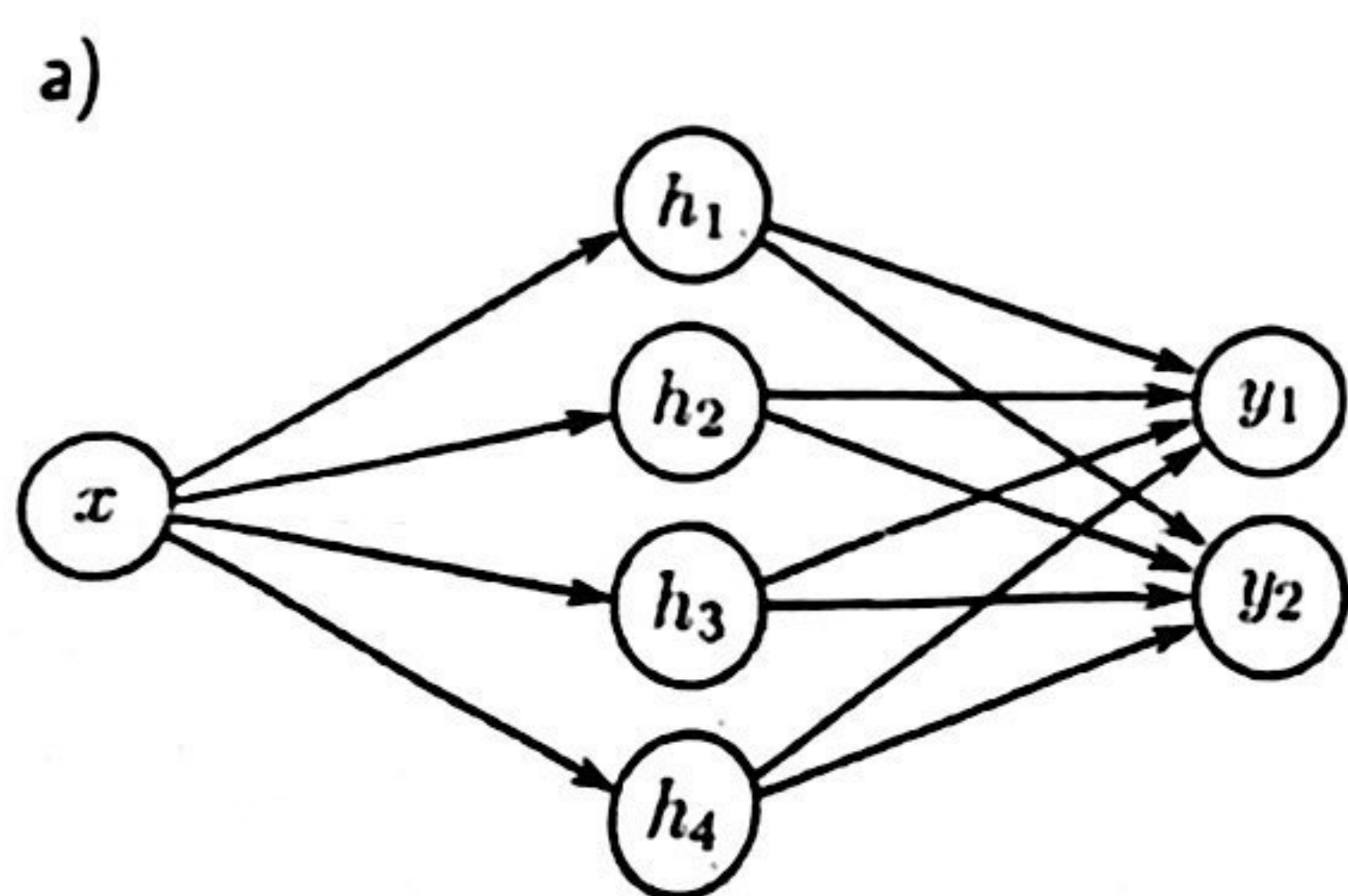
  $\rightarrow \text{heaviside}[z]$   
  $\rightarrow \text{rect}[z]$





Provide an informal description of the family of functions that can be created by neural networks with one input, three hidden units, and one output for each activation function.

3. How many parameters does the model in below given figure have?



$$\begin{aligned}
 1 \times 4 &= 4 \\
 4 \times 2 &= 8 \\
 4 + 0 &= 12 \\
 4 + 2 &= 6 \\
 12 + 6 &= 18
 \end{aligned}$$

A1 : If activation is linear each hidden unit becomes a linear function of  $x$ . The output is a linear combination of linear functions, i.e. still a linear mapping. so the model collapses to ~~something equivalent~~ to linear regression.

If activation is removed, still a linear mapping as it is just a special case where  $\phi_0 = 0$  &  $\psi = 1$ .

A3 : total parameters : 18.

A2 : Heaviside : create step like piecewise constant function.

rect : create ~~small~~ local functions (non zero in small ranges)

tanh : smooth, wavy curves. The hyperbolic tangent functions are constant for large and small values of  $x$