

CS/CS 316/365 Deep Learning

Activity 2 Solution

September 3, 2024

Shallow Neural Networks

Activity needs to be handwritten.

- 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 $a[z] = z$?

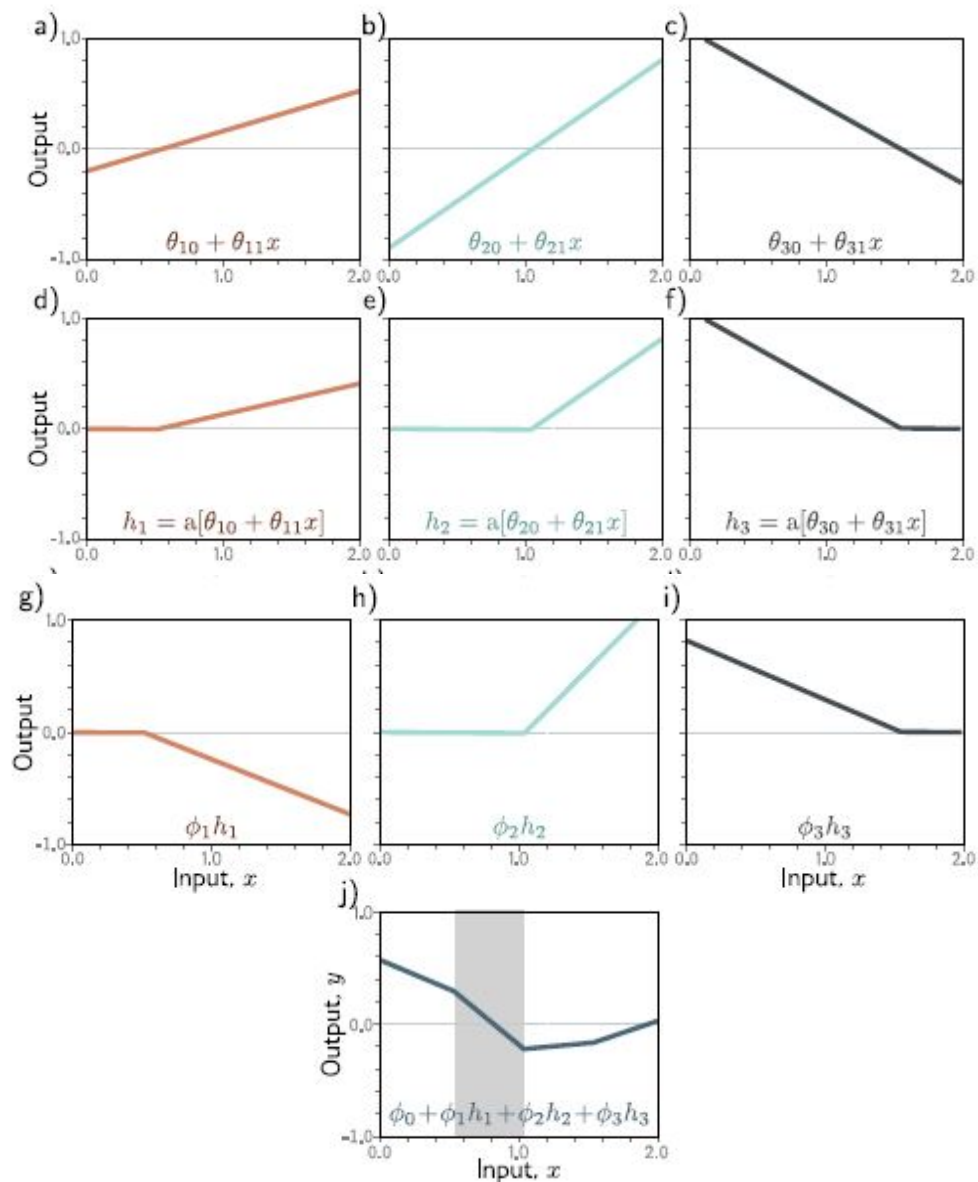
$$\begin{aligned} y &= f[x, \phi] \\ &= \phi_0 + \phi_1 a[\theta_{10} + \theta_{11}x] + \phi_2 a[\theta_{20} + \theta_{21}x] + \phi_3 a[\theta_{30} + \theta_{31}x] \end{aligned}$$

Solution: If this mapping is linear, then the overall function is a composition of three linear functions. A linear function applied to another linear function just creates a new linear function with a new slope and offset, so the model will only be able to describe lines. This is still true when the activation function is removed as this is just a special case where $\psi_0 = 0$ and $\psi_1 = 1$.

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

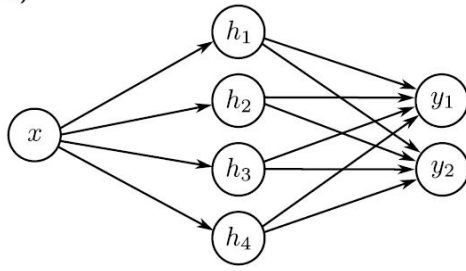


Provide an informal description of the family of functions that can be created by neural networks with one input, three hidden units, and one output for each activation function.

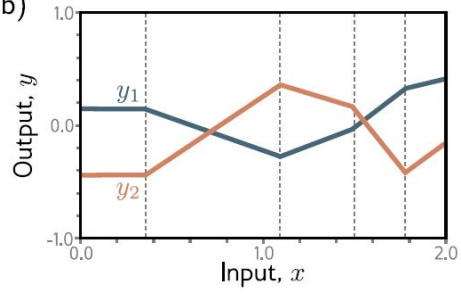
Solution: The Heaviside and rectangular activation function both make piecewise constant functions. The hyperbolic tangent and sinusoidal functions make smooth functions. The hyperbolic tangent functions are constant for large and small values of x .

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

a)



b)



Solution: It has $1 \times 4 + 4 \times 2 = 12$ slopes and $4 + 2 = 6$ intercepts giving a total of 18 parameters