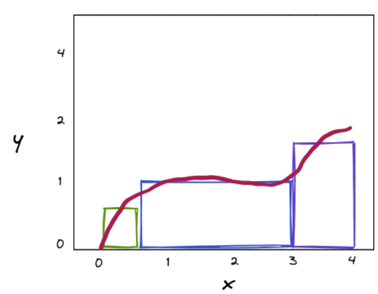
Introduction to Machine Learning
Problem Set: Deep and Convolutional Neural Networks

Summer 2021

1. **Neural network as a universal approximator.** In our lesson this week, we said that a fully connected neural network with one hidden layer can approximate a continuous function to a desired degree of precision, by using the hidden units to create "bumps" at the desired positions and heights.

   Can you use this approach to approximate the function drawn in red here, in the range $0 \leq x \leq 4$ as shown, using three "bumps" as in the illustration?

   Figure 1: Approximating a continuous function with three "bumps".

   Draw a network that is capable of this approximation. Your network should have an input unit, a hidden layer with a sigmoid activation function at each hidden unit, and an output unit (with a linear activation). Indicate *actual, numeric values* of the weights on each edge and the bias input to each node.

   Hint: if you are not sure what to do, try working through this interactive tutorial which will show you exactly how to make "bumps" using sigmoid units.

2. This animation shows how the output volume of a convolutional layer of a neural network is computed for a specific input. Every time you refresh the page, it will show you a new, randomly generated example.

   Compute the output of the center pixel at each depth in the output volume, i.e. the missing squares in the image below, for the following example:

   Input Volume (+pad 1) (7x7x3)

   `x[:,:,0]`

0	0	0	0	0	0	0
0	1	2	0	2	1	0
0	2	2	1	2	1	0
0	1	1	2	1	1	0
0	0	1	1	1	1	0
0	1	2	2	2	1	0
0	0	0	0	0	0	0

   `x[:,:,1]`

0	0	0	0	0	0	0
0	0	1	2	1	0	0
0	0	0	2	1	2	0
0	1	2	1	2	2	0
0	0	2	1	1	1	0
0	1	0	2	2	2	0
0	0	0	0	0	0	0

   `x[:,:,2]`

0	0	0	0	0	0	0
0	2	1	0	1	1	0
0	2	2	1	1	1	0
0	2	0	0	0	2	0
0	0	0	2	0	1	0
0	0	2	0	2	0	0
0	0	0	0	0	0	0

   Filter W0 (3x3x3)

   `w0[:,:,0]`

1	0	1
-1	0	1
-1	1	-1

   `w0[:,:,1]`

1	0	0
0	0	1
1	-1	1

   `w0[:,:,2]`

-1	0	1
1	1	-1
-1	-1	0

   Bias b0 (1x1x1)

   `b0[:,:,0]`

1

   Filter W1 (3x3x3)

   `w1[:,:,0]`

-1	0	-1
-1	-1	1
1	0	0

   `w1[:,:,1]`

-1	1	1
-1	-1	0
-1	1	0

   `w1[:,:,2]`

0	1	0
1	-1	-1
-1	-1	-1

   Bias b1 (1x1x1)

   `b1[:,:,0]`

0

   Output Volume (3x3x2)

   `o[:,:,0]`

3	-5	-3
11		3
2	7	3

   `o[:,:,1]`

-6	-3	-3
-3		-8
-1	-4	-5

3. Suppose the input to a 2D max pooling layer with 2x2 filters is as follows:

$$X = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \\ 2 & 1 & 0 & 0 \end{bmatrix}$$

   (a) Show the output of the pooling layer with stride 1.

   (b) Show the output of the pooling layer with stride 2.

4. Transfer learning notebook

   Please refer to the Colab notebook on the course site.