

# Homework Assignment 5

CSE 251A: Introduction to Machine Learning

**Due: June 8nd, 2022, 9:30am (Pacific Time)**

**Instructions:** Please answer the questions below, attach your code in the document, and insert figures to create **a single PDF file**.

Grade: \_\_\_\_ out of 120 points

## 1 Activation functions (40 Points)

1. Which one of these is a valid layer? For this question,  $w.shape=(input, output)$  with  $x.shape=(batch\ size, input)$  and  $b.shape=(1, output)$ . Note, below that anywhere we use dot, we could have instead used matmul.
  - (a)  $z = \text{activationFunction}(\text{np.dot}(x, b) + w)$
  - (b)  $z = \text{activationFunction}(\text{np.dot}(x, w)) + b$
  - (c)  $z = \text{activationFunction}(\text{np.dot}(x, w) + b)$
2. Name at least two possible activation functions and explain the reason why they are used as activation functions.
3. What will happen if the activation function is a linear function in Multi-layer Perceptron?

## 2 Overfitting and Regularization (30 Points)

1. What are the common techniques to alleviate overfitting in the neural network training?
2. Can we still apply the L1/L2 regularization in NN? If we can, how; if we cannot, why?

### 3 Compute output for a Convolutional Neural Network (30 Points)

Consider the image X and filter F given below. Let X be convolved with F using no padding and a stride of 1 to produce an output Y . What is the output Y ?

$$X = \begin{bmatrix} 1 & 0 & -2 & 3 & 4 & 1 \\ 2 & 9 & 5 & 6 & 0 & -1 \\ 0 & -3 & 1 & 3 & 4 & 4 \\ 6 & 5 & 2 & 0 & 6 & 8 \\ -5 & 4 & -3 & 1 & 3 & -2 \\ 4 & 1 & 2 & 8 & 9 & 7 \end{bmatrix}$$
$$F = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

## 4 (Bonus, 20 points) Experiment with CNN using Keras

In this question, you will experiment with Convolutional Neural Networks using the deep learning framework Keras (<https://keras.io/api/>). Please download the Jupyter notebook HW5\_CNN.ipynb and fill in the blanks and answer the questions. Please attach your **code** and **answers** in Gradescope submission.

**Note:** Make sure this notebook is launched in an environment with Numpy, Tensorflow, matplotlib and Keras installed. You can refer to: [https://www.tutorialspoint.com/keras/keras\\_installation.htm](https://www.tutorialspoint.com/keras/keras_installation.htm) if you need help with creating a virtual environment with all required dependencies.

## 5 (20 points) Convolution with padding and dilation.

In Lecture 15, slide 12 (and Q3 of this assignment) we have gone through the process of calculating the output size of a convolutional layer. Besides the height and width, dilation, stride, and padding are also important parameters of a CNN layer.

- dilation controls the spacing between the kernel points; also known as the atrous algorithm.
- stride controls the stride for the cross-correlation, a single number or a tuple.
- padding controls the amount of padding applied to the input's outer edges.

This [link](#) provides a great visualization of what each parameter controls.

1. Given an input of size (16, 16), a filter of size (2, 2) with **stride**=2, **padding**=1, **dilation**=4, what's the output size of it after a **one** layer 2d-convolution?
2. Given an input of size (32, 32), a filter of size (3, 3) with **stride**=2, **padding**=1, **dilation**=1, what's the output size of it after a **two** layer 2d-convolution?
3. (**Receptive Field**) The Receptive Field (RF) is defined as the size of the region in the input or previous layers that produce the current feature. For example, Figure 1 shows a 3-layer network of filter size (3, 3), with **stride**=1, **padding**=2, **dilation**=1. The receptive field in layer 2 for position (2, 2) of layer 3 is  $3 \times 3 = 9$ , namely, position (1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3) of layer 2.

When we want to figure out the RF in layer 1, we need to take the **union** of the receptive field, i.e. the 9 positions, in layer 2. **Question:** assumed the input is large enough, say, (1000, 1000), what's the size of the receptive field in layer 1 for the position (2, 2) of layer 3?

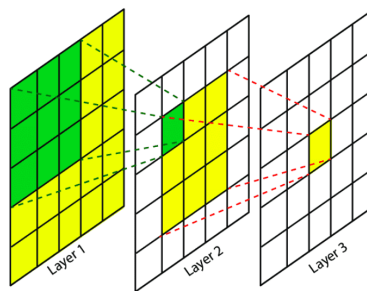


Figure 1: Example of RF for a 3-layer CNNs.