Introduction to Machine Learning, Spring 2025
Homework 5

(Due May 5, 2025 at 11:59pm (CST))

April 21, 2025

1. Please write your solutions in English.
2. Submit your solutions to the course Gradescope.
3. If you want to submit a handwritten version, scan it clearly.
4. Late homeworks submitted within 3 days of the due date will be marked down 25% each day cumulatively Homeworks submitted more than 3 days after the due date will not be accepted unless there is a valid reason, such as a medical or family emergency.
5. You are required to follow ShanghaiTech's academic honesty policies. You are allowed to discuss problems with other students, but you must write up your solutions by yourselves. You are not allowed to copy materials from other students or from online or published resources. Violating academic honesty can result in serious penalties.

1. [15 points] [Convolutional Neural Networks]

   (a) Consider a sequential 2D convolution block: the input dimension is $4 \times 64 \times 64$ (channel, width, height) and we use **two** continously Conv2D layer. The first layer is with 4 channels input and 8 channels output, where the kernel size is 3×3 (width, height). And the second layer is with 8 channels input and 16 channels output, where the kernel size is 5×5 (width, height). Set both the layers with stride $= 1$ and pad $= 1$. What is the output dimension after each layer? How many parameters do we have in these two layers in total? [10 points] [Hint: Do not forget the bias term.]

   (b) The convolution block is followed by a max pooling layer with 2×2 (width, height) filter and stride $= 2$. What is the output dimension of the pooling layer? How many parameters do we have in the pooling layer? [5 points]

   [Hint]: You can check the pytorch document for details.

   Solution

2. [15 points] [Convolution and Cross-Correlation]

(a) Compute cross-correlation and convolution of A and B (Use B as the filter), respectively. Both the operations are with zero-padding, $P = 1$, and stride is set to be 1. [12 points]

$$A = \begin{bmatrix} 2 & 4 & -3 \\ 3 & -2 & 2 \\ -3 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & -2 \\ -2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$$

(b) Under what circumstances does convolution have the same result as correlation? [3 points]

[Hint: Please notice the mathmatical definition of convolution and cross-correlation.]

Solution

3. [5 points] [Convolution layer implementation details]

   Check the pytorch document, point out whether the convolution layer in PyTorch is using cross-correlation or convolution? And why it is defined like this? [5 points]

   Solution