ECDNN 2025Summer Assignment 1

Due: 23:59, July. 13, 2025

Q1 We provide a sinlge-layer neural network (NN), as shown in Figure 1. There are three learnable weights w_1, w_2, and w_3 in this NN and the activation function is sigmoid function.

(a) Given an input $\boldsymbol{x} = (x_1, x_2, x_3)$, give the explicit expression of the output y.

(b) Suppose the ground truth is t and we use an L_2 loss ($loss = (y - t)^2$) to train this network. Derive the gradient of w_1, w_2 and w_3.

(c) Based on previous results, can you tell what problem would happen if we train a deep neural network with sigmoid activation? Can you give any suggestions to avoid this problem?

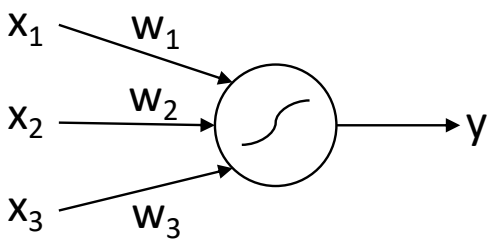

Figure 1: A simple single-layer neural network

Q2 **Regularization**. In this question, you are going to solve a toy problem. Consider a function $J(x) = (x - 2)^2, x \in \mathbb{R}$.

(a) Find the global minimum of function $J(x) + 4x^2$. Justify your answer.

(b) Find the global minimum of function $J(x) + 4|x|$. Justify your answer.

(c) Consider the following optimization problem.

$$\min_{x \in \mathbb{R}} J(\alpha; x) = (x - 2)^2 + \alpha|x|.$$

How should we determine α so that the minimizer is at $x = 0$?

(d) How do you get inspired from the above questions about ℓ_1, ℓ_2 and sparsity? Please explain *briefly*.

Q3 Convolution is the most important operation in CNN. As shown in Figure 2, the input activation tensor is $\mathcal{X} \in \mathbb{R}^{H \times W \times C}$. Weight tensor is $\mathcal{W} \in \mathbb{R}^{R \times S \times C \times K}$. The output activation tensor is $\mathcal{Y} \in \mathbb{R}^{P \times Q \times K}$. Here we set $H = W = 5, C = 8, R = S = 3, K = 4$. Besides, the stride number is 1 and the padding number is 0.

(a) What are the height and width of the output feature map (P and Q)?

(b) Write down direct convolution by C++ language style.

(c) The loop unrolling is one of loop optimization techniques to make full use of the hardware on-chip storage resources. Write down the loop unrolling at input channel level and output channel level, respectively, by C++ language style.

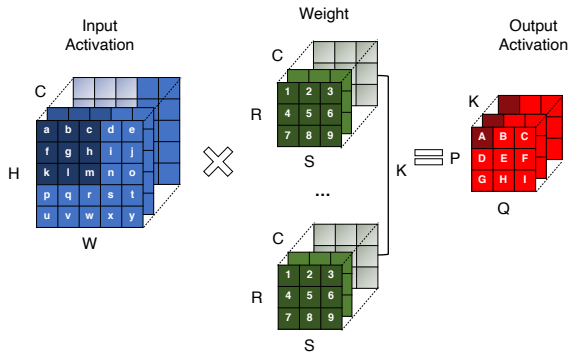

Figure 2: Convolution.

Q4 Suppose we have a feature map (Figure 3) of size $5 \times 5 \times 3$, where P_1, P_2, and P_3 have pixel value of 1 in 3 channels, i.e. $P_1 = P_2 = P_3 = [1, 1, 1]$. Consider the case: Conv2D($kernel\_size = 3$, $out\_channels = 2$, stride=1, padding=0). Please make use of what you learnt in the Lecture to solve the following questions:

(a) Build the input hash tables for P_1, P_2, and P_3 to store the input feature map.

(b) Build the output hash table.

(c) Build the rulebook for sparse convolution.

(d) What will happen during the calculation of the sparse convolution result when two rows of the rulebook have the same out index? Give an example to show this.

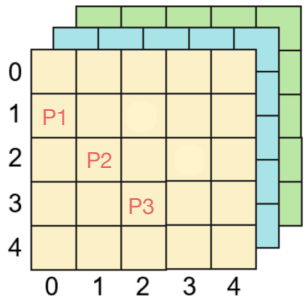

Figure 3: Feature map.