ECDNN 2025Summer Assignment 2

Due: 23:59, July. 20, 2025

Q1 Below is the weight approximation numerical example of ABC-Net. We have $\boldsymbol{W} = \begin{bmatrix} -0.135 & 0.125 \\ -0.065 & 0.075 \end{bmatrix}$ as the weight matrix to approximate. There are three binary bases with $\mu_1 = -1$, $\mu_2 = 0$ and $\mu_3 = 1$. Assume $\mathrm{mean}(\boldsymbol{W}) \approx \mathbf{0}$ (2×2 matrix) and $\mathrm{std}(\boldsymbol{W}) \approx \mathbf{0.12}$ (2×2 matrix).

(a) (6%) Calculate three bases.

(b) (6%) Calculate the approximated $\boldsymbol{W}$ with $\boldsymbol{\alpha} = [0.0275, 0.07, 0.0325]$.

Q2 Suppose we have a weight matrix $\boldsymbol{W} = [-2, 1.5, 0.5, 2] \in \mathbb{R}^{1\times 4}$ and an input $\boldsymbol{x} \in \mathbb{R}^{4\times 1}$, the output of the neural network can be represented as $y = \boldsymbol{W}\boldsymbol{x}$. We now want to quantize this neural network to accelerate the inference speed.

(a) Suppose we use two-bit integers for quantization (we use -1, 0, 1 for the quantized values). Calculate the quantized weight $\boldsymbol{W}_q$.

(b) Supposing we use quantization-aware training and straight-through estimator (STE), the gradient with respect to $\boldsymbol{W}_{\boldsymbol{q}}$ is $[0.2, 0.3, 0.4, 0.5]$, what will be the gradient with respect to $\boldsymbol{W}$?

Q3 $\boldsymbol{A} = [-2.2, -1.1, 1.1, 2.2], \boldsymbol{B} = [0.5, 0.3, 0.3, 0.5]^\top, \boldsymbol{AB} = 0.$

(a) If we use 4-bit scale quantization, set the range of [-8, 7], please provide the quantization, calculation, and dequantization procedure.

(b) What if we set the range of [-7, 7], please calculate the above procedure again.

Q4 (a) Suppose we have two discrete distributions $p = [0.2, 0.8]$ and $q = [0.6, 0.4]$. Calculate the KL divergences $D_{KL}(p||q)$ and $D_{KL}(q||p)$. What can you tell from the results.

(b) Suppose we have a fixed distribution p, and we want to learn a distribution q_θ parameterized by θ. We choose KL divergence as the loss function. As we have seen in question (1), we can have either $D_{KL}(p||q_\theta)$ or $D_{KL}(q_\theta||p)$. Can you tell the differences on the learned distribution q_θ between the two choices?