CMSC5743 2021F Homework 3

Due: Dec. 16, 2021

All solutions should be submitted to the blackboard in the format of **PDF/MS Word**.

Q1 (12%) Quantization terminology in neural network comes from digital signal processing. Following is an example. An audio compact disc (CD) holds up to 74 minutes and 33 seconds of sound, just enough for a complete mono recording of Beethoven's Ninth Symphony at probably the slowest pace. CDs use a sampling rate of 44.1 kHz (i.e., sample 44100 signals per second) with 16-bit scale quantization for each sample. When the CD was first introduced in 1983, every 8 bits of data were encoded as 17 bits of signal and error correction data together. Assume 8 bits are 1 byte and 2^20 bytes are 1 megabyte (MB).

(a) (4%) How many samples can a CD store?

(b) (4%) How many bytes can a CD record?

(c) (4%) Calculate the capacity of a CD.

Q1 (12%) 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 $\boldsymbol{\alpha} = [0.0275, 0.07, 0.0325]$.

Q3 (12%) IEEE 754 double-precision binary floating-point (FP64) format is shown in Figure 1.

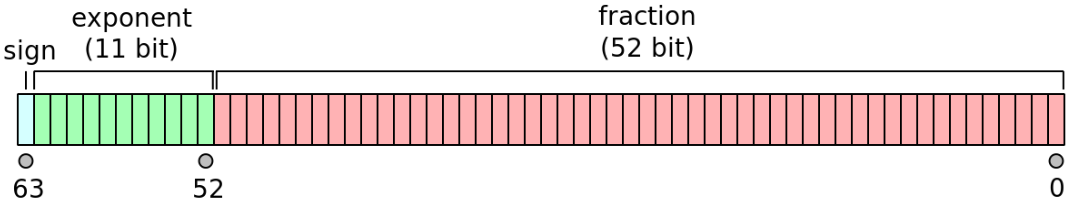


Figure 1: IEEE 754 double-precision binary floating-point format

(a) (6%) What is the IEEE 754 double-precision number 40C0 0000 40C0 0000$_{16}$ in decimal?

(b) (6%) What is -0.510_{10} in the IEEE 754 double-precision floating-point format?

Q4 (13%) We provide a very simple neural network as shown in Figure 2.

(a) (3%) Please calculate the result in the blank neuron.

(b) (7%) We use one data $(\boldsymbol{x}, y)$ to train the binary weight network. The structure of binary weight network is shown in Figure 2. $\boldsymbol{x} = [x_1, x_2, x_3] =][0.3, -0.2, 0.3]$ and $y = 1$. All weights are initialized as Figure 2. The loss function is defined as mean square error. The iteration number is set as 1. Please show all binary weights.

(c) (3%) We have another data $(\boldsymbol{x}', y')$, where $\boldsymbol{x}' = [x_1, x_2, x_3] = [0.5, 0.1, -0.4]$ and $y' = 3$. Please show estimation absolute error by the binary weight network in (b).

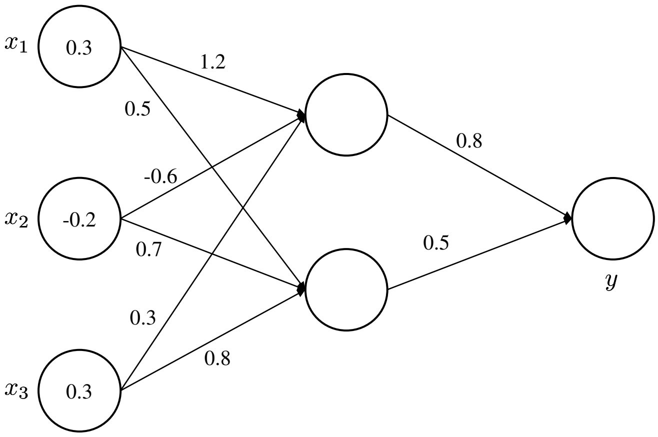


Figure 2: A simple 2-layer neural network

Q5 (10%) Quantization is a widely used technique in many areas besides network compression. For example, in electronics, an analog-to-digital converter (ADC, A/D, or A-to-D) is a system that converts an analog signal, such as a sound picked up by a microphone or light entering a digital camera, into a digital signal. Now we have a 12 bit ADC, with analog input voltage ranging from -2V to 2V. Answer the following questions.

(a) (5%) Calculate the quantization resolution.

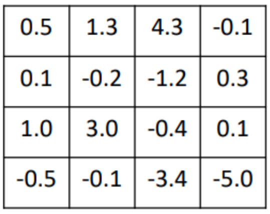
(b) (5%) Give the quantization result when the analog voltage is 1.33V.

(c) (3%) What's the quantization error in the above problem.

Q6 (10%) K-means clustering is a method of vector quantization, which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Based on the following parameters as shown in Figure 3, give the cluster results. To simplify your computation, we would recommend that you adopt the implemented K-means algorithm provided by the sklearn library in Python.

(a) (5%) If $k = 2$, what's the clustering result. Assume the initialized centers are 0.5 and -5.0.

(b) (5%) If $k = 4$, what's the clustering result. Assume the initialized centers are 0.5, -1.2, 4.3 and -5.0.



0.5	1.3	4.3	-0.1
0.1	-0.2	-1.2	0.3
1.0	3.0	-0.4	0.1
-0.5	-0.1	-3.4	-5.0

Figure 3: The weight parameters.