Q1 Instructions
1 Point

Open notes: The quiz is open notes. You are free to use any content from the course website or from your own personal notes.

No communication: ANY communication with other students about the quiz content is strictly forbidden and will result in a failing grade for the whole class (not just this quiz).

No partial credit: Every question is all or nothing credit. Thus, you must get the answer exactly right to get credit for the question (including SELECT ALL questions). No partial credit will be given on quizzes.

Number Format: When giving numbers as short answers, please give in standard decimal notation with preceding "0." for decimals if needed but no trailing 0s (e.g., "0.15", "2.9", "0.001", "100" but NOT "0.15000" NOR ".15" NOR ".001" NOR "6.0").

Honor Pledge: I assert that I have not received any information about this quiz and will not share any quiz content with anyone else. I understand that any violation of this will result in a failing grade for the whole class (not just this quiz).

Yes, I understand the policies above and assert the honor pledge.

No

Q2
1 Point

Which of the following is the loss function for denoising autoencoders where $\epsilon \sim \mathcal{N}(0, I)$ is a noise random variable?

$L(x, g(f(x + \epsilon)))$

$L(x, g(f(x) + \epsilon))$

$L(x, g(f(x)) + \epsilon)$

Q3
1 Point

The ELBO objective is always a ____ bound on the marginal log likelihood $\log p_g(x)$ regardless of the choice of q_f.

Lower

Tight

Upper

Q4
1 Point

Variational autoencoders (VAE) can generate new data after training (i.e., VAEs are generative models).

True

False

Q5
1 Point

The reparametrization trick changes the objective function value (i.e., the objectives before and after the reparametrization trick are different).

True

False

Q6
1 Point

Undercomplete autoencoders assume that the latent dimension is the same as the input but some of them are zero or near-zero.

True

False

Q7
1 Point

Suppose the input is an CIFAR image (32x32 size with only 1 channel) and we perform a 3x3 convolution with a padding of 0 on all sides and a stride of 1. What will be the width dimension of the output?

30

Q8
1 Point

Suppose $g(x) = \mathbb{E}_{z\sim q(z|x)}[f(x,z)\log f(x,z)] - \mathrm{KL}(p(z|x), q(z|x))$, where f is a concave function and p and q are conditional distributions of z given x. What can we claim without any more information?

$g(x) \leq \mathbb{E}_{z\sim q(z|x)}[f(x,z)\log f(x,z)]$

$g(x) \geq \mathbb{E}_{z\sim q(z|x)}[f(x,z)\log f(x,z)]$

$g(x) = \mathbb{E}_{z\sim q(z|x)}[f(x,z)\log f(x,z)]$

None of the above can be claimed without more information

Q9
1 Point

For any two distributions p and q, $KL(p(\cdot), q(\cdot)) = KL(q(\cdot), p(\cdot))$.

True

False

Q10
1 Point

In the Variational Autoencoders (VAEs), optimizing only over q_f while keeping p_g fixed will:

Improve the generative model.

✓ Tighten the lower bound.

Q11
1 Point

Let $\{x_i\}_{i=1}^n$ be samples from the true distribution p. If $\theta^*_{\text{MLE}} = \arg\max_\theta \sum_{i=1}^n \log \hat{q}(x_i;\theta)$ and $\theta^*_{\text{KL}} = \arg\min_\theta \widehat{KL}(p(\cdot), \hat{q}(\cdot;\theta))$ (where the KL term is approximated using samples as in lecture), then $\theta^*_{\text{MLE}} = \theta^*_{\text{KL}}$.

True

False

Q12
1 Point

Let X and Y be independent continuous random variables. If we know that $p(X=2|Y=1)=2$, $p(Y=2)=3$, and $p(Y=1|X=3)=4$, what is $p(X=2,Y=1)$? (Insufficient information [-1] is a valid answer.)

6

Answer
8

Quiz 7

● Graded

Student
Paloma Arellano

Total Points
10 / 12 pts

Question 1
Instructions 1 / 1 pt

Question 2
(no title) 1 / 1 pt

Question 3
(no title) 1 / 1 pt

Question 4
(no title) 1 / 1 pt

Question 5
(no title) 1 / 1 pt

Question 6
(no title) 1 / 1 pt

Question 7
(no title) 1 / 1 pt

Question 8
(no title) 0 / 1 pt

Question 9
(no title) 1 / 1 pt

Question 10
(no title) 1 / 1 pt

Question 11
(no title) 1 / 1 pt

Question 12
(no title) 0 / 1 pt

+ 1 pt Correct

✔ **+ 0 pts** Incorrect