CSE 493 G1/ 599 G1
Deep Learning
Autumn 2024 Quiz 5

December 03, 2024

Full Name: _____

UW Net ID: _____

Question		Score
True/False	(4 pts)	
Multiple Choice	(8 pts)	
Short Answer	(8 pts)	
Total	(20 pts)	

Welcome to the CSE 493 G1 Quiz 5!

- The exam is 30 min and is **double-sided**.
- No electronic devices are allowed.

I agree to uphold the University of Washington Student Conduct Code during this exam.

Signature: _____ Date: _____

Good luck!

This page is left blank for scratch work only. DO NOT write your answers here.

1 True / False (4 points) - Recommended 4 Minutes

Fill in the circle next to True or False, or fill in neither. Fill it in completely like this: ●. No explanations are required.

Scoring: Correct answer is worth 1 points.

1.1 GPT gives an explicit probability associated with the likelihood of a generated sentence.

- ◯ True
- ◯ False

1.2 GANs work by having two networks which are both trying to minimize the same loss function.

- ◯ True
- ◯ False

1.3 Encoders are learned in VAEs, but not in Diffusion models

- ◯ True
- ◯ False

1.4 Diffusion models are made more stable by predicting the noise added to the image instead of the image itself.

- ◯ True
- ◯ False

2 Multiple Choices (8 points) - Recommended 8 Minutes

Fill in the circle next to the letter(s) of your choice (like this: ●). No explanations are required. **Choose ALL options that apply.**

Each question is worth 4 points and the answer may contain one or more options. Selecting all of the correct options and none of the incorrect options will get full credits. For questions with multiple correct options, each incorrect or missing selection gets a 2-point deduction (up to 4 points).

2.1 Which of the following can be used to generate new images?

- ◯ A: VAEs
- ◯ B: Diffusion models
- ◯ C: SimCLR
- ◯ D: Autoregressive models
- ◯ E: GANs
- ◯ F: CLIP
- ◯ G: Discriminator models

2.2 Consider a classification model which maps images, x, to labels, y. Which of the following is a valid interpretation of $p(y)$ of the model?

- ◯ A: The model's predicted probability that an input image belongs to a given class
- ◯ B: The (normalized) bias term in the final layer
- ◯ C: An estimate that approximates the class frequencies from the training data
- ◯ D: The model probabilities for class labels for a given image, without taking into account the actual image content
- ◯ E: The probability of finding this particular image in the training data

3 Short Answers (8 points) - Recommended 8 Minutes

Please make sure to write your answer only in the provided space.

3.1 Sampling from the uniform distribution in a differentiable way

1. Recall the reparameretization trick which we used to make sampling differentiable for VAEs. We modeled the distribution of z values as a Normal Distribution with mean μ and standard deviation σ. Then, instead of directly sampling from z, we sample ϵ from $\mathcal{N}(0, 1)$ and calculate $z = \mu + \epsilon\sigma$.

   Now instead of modeling the distribution of z values as a normal distribution, you want to model it as a uniform distribution between a and b. Can you perform a similar reparameretization trick in this setting? If yes, give the specific formula of how to do this. If not, explain why not.