CS5202 Practice Questions (Set-I)
Spring 2026

1. Generative Models comparison

(a) Suppose we care about having a model with a tractable inverse mapping from data space to latent space. Which models provide this property by design?

(b) Suppose we want to minimize a divergence that is not likelihood-based, but instead depends on distinguishing real and fake samples. Which class of models does this correspond to?

(c) Suppose we want to compute the exact log-likelihood gradient with respect to model parameters for a given datapoint without using variational approximations. Which of these models allow this? Briefly explain.
a) Autoregressive models b) Latent variable models c) Flow-based models d) all of the above

(d) VAEs optimize:
$$\log p_\theta(x) = \text{ELBO} + D_{KL}(q_\phi(z \mid x) \,\|\, p_\theta(z \mid x)).$$

If the encoder has insufficient capacity, what effect does this have on generative quality?

2. Latent Variable Model and ELBO

Consider the joint distribution of a latent variable model denoted by $p(x, z)$. The model is capable of sampling only two images $\{x^{(1)}, x^{(2)}\}$, where $x^{(i)} \in \{0, 1\}^{784}$. The latent variable is scalar $z \in \mathbb{R}$. The model is defined as:

$$p(z) = \mathcal{N}(z; 0, 1)$$

$$p(x \mid z) = \begin{cases} 1 & \text{if } z \geq 0 \ \wedge \ x = x^{(1)} \\ 0 & \text{if } z \geq 0 \ \wedge \ x \neq x^{(1)} \\ 1 & \text{if } z < 0 \ \wedge \ x = x^{(2)} \\ 0 & \text{if } z < 0 \ \wedge \ x \neq x^{(2)} \end{cases}$$

(a) Compute the log-likelihood:
$$\text{like}(x) := \log p(x).$$

What is $\text{like}(x^{(1)})$?

(b) Compute the posterior distribution $p(z \mid x^{(1)})$ explicitly.

(c) Let the variational distribution be:

$$q(z) = \mathcal{N}(z; 0, 1).$$

Compute the KL divergence:
$$D_{KL}(q(z)\|p(z \mid x^{(1)})).$$

(d) Now consider instead:
$$q(z) = \mathcal{N}(z; 1, 1).$$

Compute numerically:
$$P_q(z \geq 0).$$

(e) Using your answer above, compute the ELBO:

$$\text{ELBO}(x^{(1)}; q) = \log p(x^{(1)}) - D_{KL}(q(z)\|p(z \mid x^{(1)})).$$

Is it finite?

3. Theory Questions

(a) What is the goal of Generative AI. How does it differ from discriminative modeling?

(b) Give two historical milestones in the development of modern Generative AI and explain why they were significant.

(c) Explain why latent variable models are useful for high-dimensional data such as images or speech.

(d) Suppose that you are modeling medical images where the interpretability of latent factors is important. What advantages do latent variable models provide compared to autoregressive models?

(e) Why is variational inference often preferred over exact Bayesian inference in deep generative models?

(f) What are the main limitations of variational inference? Discuss the role of the variational gap.

(g) Explain the key differences between a standard autoencoder and a VAE.

(h) In a real-world application such as anomaly detection in industrial equipment, which model would you choose and why?

(i) Discuss two advantages and two disadvantages of GANs compared to likelihood-based models.

(j) You are building a system to generate realistic sports player avatars for a video game. Would you choose a GAN, VAE, or diffusion model? Justify your answer.

(k) What are the main advantages of normalizing flows over VAEs?

(l) Why are normalizing flows typically harder to scale to very high-resolution images?

(m) Explain the main idea behind diffusion models.

(n) Why do diffusion models often produce higher-quality samples than GANs?

(o) What is the primary computational disadvantage of diffusion models compared to GANs?

(p) Explain the forward and reverse processes in Denoising Diffusion Probabilistic Models (DDPM).

(q) In a real-time content generation application (e.g., live video enhancement), what challenges would arise when deploying DDPM models?

(r) You are designing a generative model for the following tasks:

    (a) Generating realistic synthetic MRI scans where likelihood evaluation is required for anomaly scoring.

    (b) Generating creative artwork for marketing campaigns where visual quality is the primary objective. For each case, select an appropriate model class and justify your choice.

4. Numericals

1. Consider a DDPM defined with a forward diffusion process:

$$q(x_t \mid x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I)$$

where $\{\beta_t\}_{t=1}^T$ is a fixed variance schedule.

    (a) Show that the forward process admits the closed-form:

$$q(x_t \mid x_0) = \mathcal{N}\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)I\right)$$

    where

$$\alpha_t = 1 - \beta_t, \quad \bar{\alpha}_t = \prod_{s=1}^t \alpha_s.$$

    (b) Write the full training objective used in DDPM when the model predicts the noise $\epsilon_\theta(x_t, t)$.

Consider a 1-dimensional diffusion process with:

$$\beta_1 = 0.1, \quad \beta_2 = 0.2.$$

Let $x_0 = 2$.

    (a) Compute α_1, α_2 and $\bar{\alpha}_2$.

    (b) Compute the mean and variance of $q(x_2 \mid x_0)$.

(c) Explain why the reverse distribution

$$q(x_{t-1} \mid x_t)$$

is Gaussian.

(d) If the model perfectly predicts the true noise, what distribution does the generated sample converge to as $T \to \infty$?