

Generative AI (AI4009)

Course Instructor(s):

Dr. Akhtar Jamil

Sections: A,B,C,D

Sessional-I Exam

Total Time (Hrs): 1

Total Marks: 50

Total Questions: 5

Date: Sep 22, 2025

Roll No

Course Section

Student Signature

Do not write below this line.

Attempt all the questions.

[CLO 1: Explain and interpret the foundational principles of generative AI, including Bayesian learning, neural architectures (CNNs, RNNs, Transformers), and their role in generative modeling.]

Question No. 1. MCQ [1 x 25 = 25]

Answer the MCQs on the given answer sheet attached at the end of the question paper. Answers marked on the question paper will not be evaluated.

1. You are using semantic hashing for image de-duplication as given in DCGANs paper. Three images are represented by their binary hash codes:

1	1
0	0

I₁

1	0
0	0

I₂

0	0
1	1

I₃

If Image **I₁** must be kept, which one of the other two images should be dropped based on semantic hash codes?

- A. Image **I₁**
 - B. Image **I₂**
 - C. Image **I₃**
 - D. None
2. Why do Variational Autoencoders (VAEs) often generate images that appear blurrier than those in the original dataset?
- A. Adversarial loss smooths out fine details
 - B. Latent space regularization removes sharp edges
 - C. Encoders output nearly identical latent codes
 - D. **Reconstruction loss averages variations, reducing sharpness**
3. Why do Mini-Batch GANs reduce mode collapse compared to standard GAN training?
- A. They force the generator to memorize training data.
 - B. **They allow the discriminator to compare relationships among multiple samples.**
 - C. They eliminate all gradient noise during optimization.
 - D. They increase the generator's parameter space.

National University of Computer and Emerging Sciences

Islamabad Campus

4. The discriminator loss in GANs is defined as:

$$L_D = -\frac{1}{N} \sum_{i=1}^N [y_i \log(D(x_i)) + (1 - y_i) \log(1 - D(G(z_i)))]$$

For a mini-batch of $N = 2$ with just two classes:

- Sample 1: real data x_1 , $y_1 = 1$, $D(x_1) = 0.8$
- Sample 2: generated data $G(z_2)$, $y_2 = 0$, $D(G(z_2)) = 0.3$

What is the discriminator loss L_D (use natural log, round to 3 decimal places)?

- A. 0.123
- B. 0.290**
- C. 0.456
- D. 0.678

5. In the Diagonal BiLSTM of PixelRNN, why is processing done along diagonals instead of rows?
- A. Because diagonals contain independent pixels that can be computed in parallel
 - B. Because row-wise processing ignores vertical dependencies
 - C. Because diagonal traversal reduces vanishing gradient issues
 - D. Because diagonals allow each pixel to access a larger context (receptive field) than row-wise processing**
6. In GANs, we can generally use a pretrained Discriminator to improve the quality of the Generator and reduce training time.
- A. True
 - B. False**
7. Which one of the following is NOT a solution to mode collapse?
- A. Mini-batch discrimination
 - B. Multi-label Supervised training
 - C. Unsupervised Training without Labels**
 - D. None of the above
8. What is the effect of setting the regularization parameter λ too large?
- A. The model becomes more likely to underfit the data.**
 - B. The model becomes more likely to overfit the data.
 - C. The model's capacity to generalize improves.
 - D. The optimization process becomes faster.
9. What is the primary purpose of transpose convolutions (also known as deconvolutions) in neural networks?
- A. To increase the depth of the feature maps.
 - B. To reduce the spatial dimensions of the input.
 - C. To perform element-wise multiplication between feature maps.
 - D. To increase the spatial dimensions of the input.**

National University of Computer and Emerging Sciences

Islamabad Campus

10. Since the learning rate changes dynamically across epochs in adaptive learning rate strategies, can it be considered a model parameter?
- A. True
 - B. **False**
11. For which of the following problems, RNN is not suitable:
- A. Time series forecasting
 - B. Sentiment analysis on text data
 - C. **Image classification**
 - D. Speech recognition
12. What is the purpose of the entropy term in a VAE?
- A. **To encourage diversity in the latent space**
 - B. To sharpen the reconstructed outputs
 - C. To minimize reconstruction error
 - D. To enforce a strict mapping between inputs and latent vectors
13. What does the receptive field of a neuron in a convolutional neural network (CNN) refer to?
- A. The number of filters applied to an input.
 - B. The number of parameters in a convolutional layer.
 - C. The stride and padding used in convolution.
 - D. **The area of the input image that influences the activation of a neuron.**
14. In autoregressive models such as PixelCNN, how can we increase the context (receptive field) seen by each pixel without significantly increasing computation?
- A. Stack many more convolutional layers
 - B. **Use dilated (atrous) convolutions**
 - C. Replace convolutions with fully connected layers
 - D. Increase filter size for every convolutional layer
15. During GAN training, the system may suffer from local convergence. Which of the following best explains why this happens?
- A. Generator reaches global optimum too early
 - B. Discriminator keeps improving indefinitely
 - C. Generator overfits to training samples
 - D. **Discriminator feedback becomes uninformative, leading to collapse**
16. You are given a very large dataset where most of the data is unlabeled and only a small portion is labeled. Which of the following approaches is most suitable for this scenario?
- A. Train a standard feedforward ANN directly on the small, labeled portion
 - B. Use a recurrent neural network trained only on the labeled portion
 - C. **First train an autoencoder on the unlabeled data, then use the encoder to train a classifier on the small, labeled set**
 - D. Train a GAN, since GANs do not require large amounts of data

National University of Computer and Emerging Sciences
Islamabad Campus

17. Long-range dependencies can be introduced in autoregressive models using
- a). CNN
 - b). ANN
 - c). **RNN**
 - d). None of the above
18. Consider a model that learns the following joint probability distribution for three sequential inputs X_1, X_2, X_3 and the label Y :

$$P(X_1, X_2, X_3, Y) = P(Y) \cdot P(X_1|Y) \cdot P(X_2|X_1, Y) \cdot P(X_3|X_2, X_1, Y)$$

Is this model generative or discriminative?

- A. **Generative**
 - B. Discriminative
 - C. None
19. Which of the following is not a hyperparameter
- A. Learning rate
 - B. Batch size
 - C. Epochs
 - D. **Bias**
20. Consider an input of size $64 \times 64 \times 10$. To reduce the number of channels to 5 without changing the spatial dimensions, which of the following strategies would be the most suitable?
- A. Use a convolution with a kernel size of 3×3 , and 5 filters.
 - B. **Use a convolution with a kernel size of 1×1 , and 5 filters.**
 - C. Apply max pooling with a pool size of 2×2 .
 - D. None of the above.
21. For a convolutional layer with an input size of $5 \times 5 \times 3$ and 3 filters of size 3×3 , how many trainable parameters (including bias) are required?
- A. 27
 - B. 54
 - C. **84**
 - D. 243
22. Given the following probabilities, what is the joint probability $P(X = \text{Generative}, Y = \text{AI})$?
- $P(X = \text{Discriminative}) = 0.4$, $P(X = \text{Generative}) = 0.6$,
 $P(Y = \text{AI} | X = \text{Discriminative}) = 0.3$, $P(Y = \text{AI} | X = \text{Generative}) = 0.5$
- A. 0.20
 - B. 0.24
 - C. **0.30**
 - D. 0.50

National University of Computer and Emerging Sciences
Islamabad Campus

23. In a Many-to-Many RNN architecture for machine translation, how is the total loss computed during training?
- A. By computing the loss only at the final time step
 - B. By comparing output with the given label at the output layer
 - C. By using only the first time step for loss computation
 - D. **By summing the loss over all time steps**
24. In the context of training neural networks, what is the role of the momentum term in gradient descent?
- A. It increases the learning rate dynamically when gradients are small
 - B. **It helps escape local minima by adding a fraction of the previous update to the current update**
 - C. It reduces overfitting by penalizing large weights
 - D. It normalizes gradients to keep them within a fixed range

25. Which of the following kernels is most suitable for detecting diagonal lines in an image?

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

A

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

B

$$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

C

None of these

D

[CLO 1: Explain and interpret the foundational principles of generative AI, including Bayesian learning, neural architectures (CNNs, RNNs, Transformers), and their role in generative modeling.]

National University of Computer and Emerging Sciences

Islamabad Campus

Question No 2. Write short answers to the following questions. [3 x 5=15]

- 1) Suppose the input z has dimension 10. We want to project it through a fully connected (dense) layer so that it can be reshaped into a $2 \times 2 \times 3$ feature map. How many parameters does this fully connected layer need to have (including biases)?

For $2 \times 2 \times 3 = 12$ outputs

- Input dimension = 10
- Output dimension = 12 (for $2 \times 2 \times 3$)
- Parameters in weight matrix = $10 \times 12 = 120$
- Bias terms = 12
- **Total = $120 + 12 = 132$ parameters**

- 2) Why do we train the generator and discriminator of a GAN alternately using the min-max loss function?

We train the generator and discriminator alternately using the min-max loss because both networks have opposing goals: the discriminator learns to distinguish real from fake samples, while the generator learns to fool it. Alternating updates maintain balance and prevent one from overpowering the other, ensuring stable convergence.

- 3) What is local minima issue that we see in neural networks? Propose a solution to overcome it?

The local minima issue occurs when the training process gets stuck in a point where the loss is low but not optimal. This prevents the model from reaching the best solution. Using techniques like momentum-based optimizers (e.g., Adam), proper weight initialization, and batch normalization can help overcome it.

- 4) Why are we forcing the encoder output distribution to match with Gaussian distribution with zero mean and unit variance? Explain with reasons.

In VAEs, forcing the encoder's output distribution to match a Gaussian (zero mean, unit variance) ensures a continuous and smooth latent space. This allows meaningful sampling and interpolation, making the model capable of generating coherent and realistic data.

- 5) In a Variational Autoencoder (VAE), what happens if the KL divergence weight (β) is set too high or too low?

If β (KL divergence weight) is too high, the latent space becomes overly constrained, and the model may ignore input details (underfitting). If β is too low, the latent space loses structure, leading to poor generation quality and overfitting to training data.

[CLO 1: Explain and interpret the foundational principles of generative AI, including Bayesian learning, neural architectures (CNNs, RNNs, Transformers), and their role in generative modeling.]

National University of Computer and Emerging Sciences
Islamabad Campus

Q4: Long questions

[2x5= 10 marks]

1) Consider two discrete probability distributions P and Q defined over the same events:

$$P = [0.5, 0.3, 0.2], Q = [0.4, 0.4, 0.2]$$

Compute the Kullback–Leibler (KL) Divergence:

$$D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

(use natural logarithm, base e).

$$\begin{aligned} D_{KL}(P \parallel Q) &= 0.5 \log \frac{0.5}{0.4} + 0.3 \log \frac{0.3}{0.4} + 0.2 \log \frac{0.2}{0.2} \\ &= 0.5 \log (1.25) + 0.3 \log (0.75) + 0.2 \log (1) \\ &= 0.5(0.2231) + 0.3(-0.2877) + 0 \\ &= 0.11155 - 0.08631 = 0.0252 \end{aligned}$$

2) Consider two images 2x2. Using L_2 distance (Euclidean) find the diversity score between the two images.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

Consider the following formula to calculate diversity score.

$$T_b(x_i) = \sum_{j=1}^N e^{-||o(x_i) - o(x_j)||}$$

1. Flatten each image:

- $x_1 = [1, 2, 3, 4]$
- $x_2 = [2, 3, 4, 5]$

2. Compute L_2 norms:

$$\begin{aligned} o(x_1) &= \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30} \approx 5.477 \\ o(x_2) &= \sqrt{2^2 + 3^2 + 4^2 + 5^2} = \sqrt{54} \approx 7.348 \end{aligned}$$

3. Plug into formula for $T_b(x_1)$:

$$\begin{aligned} T_b(x_1) &= e^{-|o(x_1) - o(x_1)|} + e^{-|o(x_1) - o(x_2)|} \\ &= e^{-(5.477 - 5.477)} + e^{-(5.477 - 7.348)} \\ &= e^0 + e^{-1.871} \\ &= 1 + 0.154 = 1.154 \end{aligned}$$

National University of Computer and Emerging Sciences

Islamabad Campus

Answer Sheet MCQs

Fill the correct option. Only one option must be selected. Selection of multiple options or overwriting will result in ZERO marks.

CORRECT METHOD				WRONG METHOD			
<input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>				<input checked="" type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input checked="" type="radio"/>			
NAME:				ROLL NO:			
<div style="display: flex; justify-content: space-around;"> <div> <p>Roll No</p> <div style="border: 1px solid black; width: 30px; height: 20px; margin: 0 auto;"></div> <div style="border: 1px solid black; width: 30px; height: 20px; margin: 0 auto;"></div> <div style="border: 1px solid black; width: 30px; height: 20px; margin: 0 auto;"></div> <div style="border: 1px solid black; width: 30px; height: 20px; margin: 0 auto;"></div> </div> <div> <p>A B C D</p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>0 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>1 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>2 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"> <p>11 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>12 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>13 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>3 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>4 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>5 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"> <p>14 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>15 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p style="text-align: center;">A B C D</p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>6 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>7 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>8 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"> <p>16 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>17 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>18 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>9 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p style="text-align: center;">MCQs</p> <p style="text-align: center;">Section1</p> <p style="text-align: center;">A B C D</p> </div> <div style="width: 45%;"> <p>19 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>20 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p style="text-align: center;">A B C D</p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>1 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>2 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>3 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"> <p>21 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>22 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>23 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>4 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>5 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p style="text-align: center;">A B C D</p> </div> <div style="width: 45%;"> <p>24 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>25 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>6 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>7 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>8 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"></div> </div>							
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>9 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>10 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> </div> <div style="width: 45%;"></div> </div>							