Lecture 10 – Diffusion Models

Provide short answers (max. 25 words each) to the following questions.

- What is the main idea behind Diffusion Models?
- What does the encoder in a Diffusion Model do?
- What does the decoder in a Diffusion Model do?
- What type of probabilistic structure does the forward process follow?
- What is a noise schedule?
- What is the role of the diffusion kernel?
- Why is training diffusion models time-consuming?
- What distribution is typically used for noise in the forward process?
- What does the reverse process aim to do?
- How are reverse distributions approximated?
- What sampling method is used for data generation in Diffusion Models?
- What is the objective function for training the reverse process?
- Why is ELBO used in Diffusion Models?
- What does the diffusion loss function train?
- What is the reparameterization trick in Diffusion Models?
- What is the primary application of Diffusion Models?
- How is conditional generation achieved in Diffusion Models?
- What is classifier guidance?
- How can diffusion models generate text-conditioned images?
- Why are diffusion models considered robust?
- What is the main computational drawback of Diffusion Models?
- Why are Diffusion Models suitable for conditional generation?
- What is the role of Monte Carlo estimation in Diffusion Models?

Multiple-choice questions

What is the main idea behind Diffusion Models?

A) Encode data using adversarial loss
B) Approximate data likelihood via a forward noising process and reverse denoising process
C) Use GANs for image generation
D) Apply clustering to latent space

What does the encoder in a Diffusion Model do?

A) Removes noise from data
B) Generates samples from latent space
C) Predicts latent variables directly
D) Adds noise progressively to data

What does the decoder in a Diffusion Model do?

A) Learns to reverse the noising process
B) Adds noise to latent variables
C) Computes adjacency matrix
D) Performs clustering

What structure does the forward process follow?

A) Fully connected network
B) Convolutional layers
C) Markov chain
D) Attention mechanism

What is a noise schedule?

A) Sequence of latent variables
B) Rule for blending noise at each step
C) Loss function for diffusion
D) Sampling method

What distribution is typically used for noise in Diffusion Models?

A) Uniform distribution
B) Exponential distribution
C) Bernoulli distribution
D) Standard Normal distribution

Why is training Diffusion Models time-consuming?

A) Requires adversarial optimization
B) Uses small batch sizes
C) Needs multiple noisy samples per input across time steps
D) Requires GAN discriminator

What does the diffusion kernel provide?

A) Loss function for GANs
B) Sampling strategy for latent space
C) Closed-form expression for noisy variable distribution
D) Attention weights

What is the goal of the reverse process?

A) Add noise to data
B) Map noisy variables back to clean data
C) Compute latent embeddings

D) Perform clustering

How are reverse distributions approximated?

A) Using normal distributions parameterized by neural networks
B) Using uniform distributions
C) Using GAN discriminators
D) Using attention layers

What sampling method is used for generation in Diffusion Models?

A) Beam search
B) Random cropping
C) Ancestral sampling
D) Gibbs sampling

What is the objective function for training the reverse process?

A) Minimize KL divergence only
B) Maximize log-likelihood of training data
C) Minimize reconstruction error only
D) Use adversarial loss

Why is ELBO used in Diffusion Models?

A) To approximate intractable log-likelihood
B) To compute GAN loss
C) To normalize latent space
D) To enforce Lipschitz constraint

What does the diffusion loss function train?

A) A single network for all steps
B) A network for each diffusion time step
C) GAN discriminator
D) Attention weights

What is the reparameterization trick in Diffusion Models?

A) Predict clean data directly
B) Predict the noise added during forward process
C) Remove latent variables
D) Use GAN generator

What metric measures reconstruction accuracy?

A) Cosine similarity
B) Wasserstein distance

C) KL divergence only
D) Distance between normal distributions

How is conditional generation achieved?
A) By adding random noise
B) By removing latent space
C) By guiding denoising using class labels or text embeddings
D) By using GAN discriminator

What is classifier guidance?
A) Uses gradients from a classifier to influence generation
B) Uses GAN loss
C) Removes noise schedule
D) Applies dropout

How can Diffusion Models generate text-conditioned images?
A) By using attention only
B) By conditioning on text embeddings
C) By removing latent space
D) By adversarial training

What technique is used for generating higher-resolution images?
A) Random cropping
B) Conditional generation with resolution-specific conditioning
C) GAN-based upsampling
D) Removing noise schedule

Why are Diffusion Models considered robust?
A) They avoid adversarial training and mode collapse
B) They use GAN discriminators
C) They ignore latent space
D) They require fewer steps

What is the main computational drawback of Diffusion Models?
A) High memory usage only
B) Lack of latent space
C) Slow sampling due to many reverse steps
D) Poor image quality

Which property of latent space is desirable in Diffusion Models?
A) Random transitions

B) Smooth transitions between noise and data
C) Discrete jumps
D) Ignoring latent variables

Why are Diffusion Models suitable for conditional generation?

A) They allow flexible conditioning during reverse process
B) They ignore latent space
C) They use GAN discriminators
D) They require paired data only

What is the role of Monte Carlo estimation in Diffusion Models?

A) Compute GAN loss
B) Approximate ELBO during training
C) Normalize latent space
D) Remove noise schedule