Quiz - Solution
B657 Discussion Section

Abhay Singh

1. As the name suggests generative models for image generation, these are computer programs that learn how to generate new images.

For better analogy we can take example of **DALL-E used by GPT for image generation, the core idea behind the image generation is the same i.e. the model (GAN) learns from the data during training, and it can generate new images from scratch.**

How does GAN works?
Generative Adversarial Nets contains two components **Generator and Discriminator**. The generator develop images from scratch and discriminator's work is to detect fake images (created by generator) or real (based on training data). With iterations the generator gets betters at creating images that look real and on the other hand discriminator gets better at detecting fake images. This process ensures that the image creation by generator comes out so realistic that even discriminator can't spot the difference.

2.
(i) True
(ii) False
(iii) True
(iv) False

3. This step in a VAE is a problem because we can't directly take gradients through random sampling and without gradients the model can't learn using backpropagation. To fix this VAEs use reparameterization trick. Instead of sampling z directly from a distribution, we write it as **z = mu + sigma times epsilon** where epsilon is just random noise from a normal distribution (like N(0,1)) This makes the whole thing work like a normal function, so we can still take gradients and train the model. It's like saying **"Let's keep the randomness but move it outside the tricky part so learning still works.**"

4. The two player minimax game is the core idea behind how GANs work.

The Generator creates images that look real starting from a vector of random numbers. Its goal is to make the Discriminator believe that the generated images are real. On the other hand, the Discriminator looks at both real images (from the training data) and fake images (from the Generator) and tries to correctly tell them apart.

This creates a minimax game a kind of competition where one player tries to minimize the loss (Generator) while the other tries to maximize it (Discriminator). The Discriminator tries to max its ability to guess real from fake while Generator tries to min this ability by getting better at generating realistic images.

Over time both models improve. The Generator produces better and more realistic images and the Discriminator becomes sharper at detecting fakes until ideally the fake images become so good that the Discriminator

can't tell the difference anymore.

For example, if trained on handwritten digits the Generator will start producing fake digits that look just like real ones and the Discriminator will have a hard time guessing which is real. This back and forth improves the quality of the generated data.

5. VAEs
Intuition:

To have intuition about how does VAEs work, let's go back to an infamous concept in machine learning which is Dimensionality Reduction PCA. In PCA we reduce the dimensionality of the dataset to have maximum variance.

Now, suppose we have a dataset of rock images where the original dimensionality of the dataset is around 980 fields and using PCA we bring it down to 150 fields preserving maximum variance. Now, using the PCA reduced dataset if we want to reconstruct the original images then we can do it using PCA inverse transform function, that will not give us exact images, but some noise will be there as we have lost some of the details while reducing the dimensionality.

So, in short, we go from Original → Compressed → Reconstructed

Encoder:
VAEs, the core idea is to compress the input image into smaller and simpler code which is known as latent space (basically summarizing an image, like about its style, color or shape), but instead of giving it a single summary we give it a range of summaries.

Decoder:
Then we learn how to decode the image into original image. This takes the summary generated by encoder and creates image based on it. The model basically learns how to make the encoded version follow a smooth continuous distribution.
The image generated might not be perfect, but it is very similar. VAEs are trained using backpropagation which is similar to calculating the loss function and reducing the loss function just in regular neural networks and they include a term called KL divergence which makes sure the encoded summary stays close to normal distribution.

GANs:
Generative Adversarial Nets have two parts: a Generator and a Discriminator. The Generator develops images from noise while Discriminator checks if an image is real (from the training data) or fake (made by the Generator). Over time the Generator improves at creating close to real or actual images, and the Discriminator gets better at checking differences. This process continues until the Generator creates images that are so real that the Discriminator can't tell the difference.

To summarize VAEs focus on learning structured smooth representations, while GANs focus on generating the most realistic images possible through competition between two networks (generator and discriminator).

6. (A) Encoder
An encoder is the part of a model (like a Variational Autoencoder) that takes input data and compresses it into a smaller meaningful representation called latent space. For example, given an image the encoder learns to represent it as a short vector that captures the most important features of that image.

(B) Decoder

A decoder receives the compact representation given by encoder and tries to recreate the original input. In VAEs it transforms points from the latent space into images. A well trained decoder generates outputs that closely resemble the original data.

(C) Intractability
Intractability refers to situations where a mathematical operation like computing a probability or integral is too complex or impossible to solve exactly. In many generative models, computing the exact likelihood or posterior is intractable so we use approximations instead.

(D) GAN
It consists of a generator and a discriminator. The generator develop images and the discriminator checks the difference in generated images (comparing it with training data). They compete in a minimax game like setup and improving each other until the generated data becomes very realistic.

(E) Discriminator
The discriminator is the part of a GAN that receives both real and fake data and tries to correctly classify each one. Its goal is to tell apart real examples from the training data and fake examples made by the generator.

(F) Generator
The generator is the part of a GAN that creates new, fake data starting from random noise. It aims to produce data that is real and the discriminator cannot tell the difference. As training progresses, the generator gets better at creating convincing data.

References

1. Auto-Encoding Variational Bayes
2. Generative Adversarial Network
3. Reddit CV Forums
4. Medium
5. 3Blue1Brown Youtube Channel