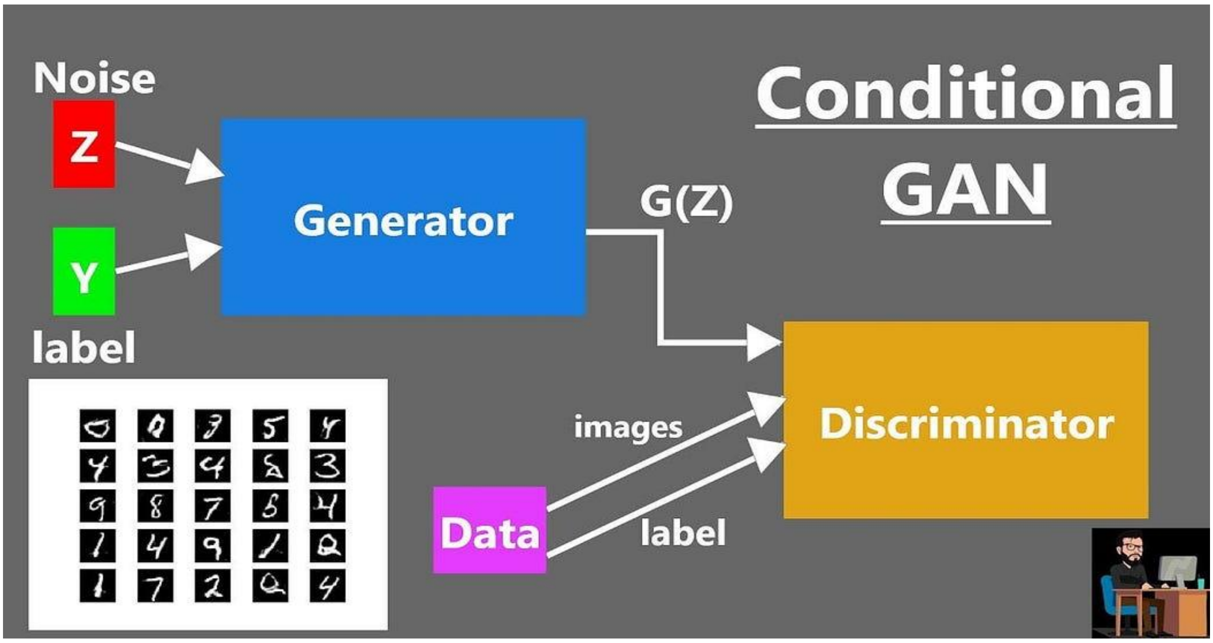
Semester: VIII Subject: Advanced AI Academic Year:2024-2025

Module 2

Conditional Generative Adversarial Network(C-GAN):

- A conditional generative adversarial network (cGAN) is a type of neural network that uses labels—or conditions—to generate novel text or images that have characteristics similar to its training data set.
- The difference between a typical generative adversarial network (GAN) and a conditional GAN is that you use labeled data to provide context to a conditional GAN, allowing you to get better, more targeted results from the generator.
- Ex: Imagine the need to generate images that are of only Mercedes cars when you have trained your model on a collection of cars. To do that, you need to provide the GAN model with a specific "condition," which can be done by providing the car's name (or label).
- Conditional generative adversarial networks work in the same way as GANs. The generation of data in a CGAN is conditional on specific input information, which could be labels, class information, or any other relevant features. This conditioning enables more precise and targeted data generation.

Architecture and working of CGANs:



Semester: VIII Subject: Advanced AI Academic Year:2024-2025

Conditioning in GANs:

- GANs can be extended to a conditional model by providing additional information (denoted as y) to both the generator and discriminator.
- This additional information (y) can be any kind of auxiliary information, such as class labels or data from other modalities.
- In the generator, the prior input noise (z) and y are combined in a joint hidden representation.

Generator Architecture:

- The generator takes both the prior input noise (z) and the additional information (y) as inputs.
- These **inputs are combined in a joint hidden representation**, and the generator produces synthetic samples.
- The adversarial training framework allows flexibility in how this hidden representation is composed.

Discriminator Architecture:

- The discriminator **takes both real data (x) and the additional information (y)** as inputs.
- The discriminator's task is to distinguish between real data and synthetic data generated by the generator conditioned on y.

Loss Function:

The objective function for the conditional GAN is formulated as a two-player minimax game:

$$minG\ maxD\ \ V(D,G) = \mathrm{E}x{\sim}pdata(x)\ [logD(x|y)] + \mathrm{E}z{\sim}pz(z)\ [log(1-D(G(z|y)))]$$

- E represents the expected operator. It is used to denote the expected value of a random variable. In this context, Ex∼pdata represents the expected value with respect to the real data distribution pdata(x), and E*z*∼*pz*(*z*) represents the expected value with respect to the prior noise distribution *pz*(*z*).
- The objective is to simultaneously minimize the generator's ability to fool the discriminator and maximize the discriminator's ability to correctly classify real and generated samples.
- The first term (logD(x|y)) encourages the discriminator to correctly classify real samples.
- The second term (log(1−D(G(z|y)))) encourages the generator to produce samples that are classified as real by the discriminator.

- This formulation creates a balance in which the generator improves its ability to generate realistic samples, and the discriminator becomes more adept at distinguishing between real and generated samples conditioned on y.

The cGAN and its multiple uses:

1. Image-to-image translation:

cGANs, in particular, allow images to evolve by considering additional information, such as labels. cGANs have enabled the development of the Pix2Pix method, some applications of which include object reconstruction from edges, photo synthesis from label maps, and image colorization.

2. Creating images from text:

Thanks to cGANs, it's possible to create high-quality photos based on text. Using text and the richness of its vocabulary enables the generation of much more precise synthetic images.

3. Video generation:

In video, cGANs can also predict future frames of a video based on a selection of previous images.

4. Face generation:

cGANs can be used to generate images of faces with specific attributes, such as hair or eye color.