Generative Adversarial Networks

Raamkumar Thiyagarajan
Michael Wang
Andres Silva

GANs: Generative Models vs. Supervised Learning

Supervised Learning

- Labeled data
- Learns a function of **x**.

$$\hat{y} = \hat{f}(\vec{x})$$

- Given a set of features, predict the response.
- Eg. Age+Height → Weight.

↓

Prediction

GANs (Generative Adversarial Networks)

- Unsupervised Learning
  - No labels
  - Underlying structure
- Generative Models
  - Learn the intrinsic distribution function of the data. Ie:

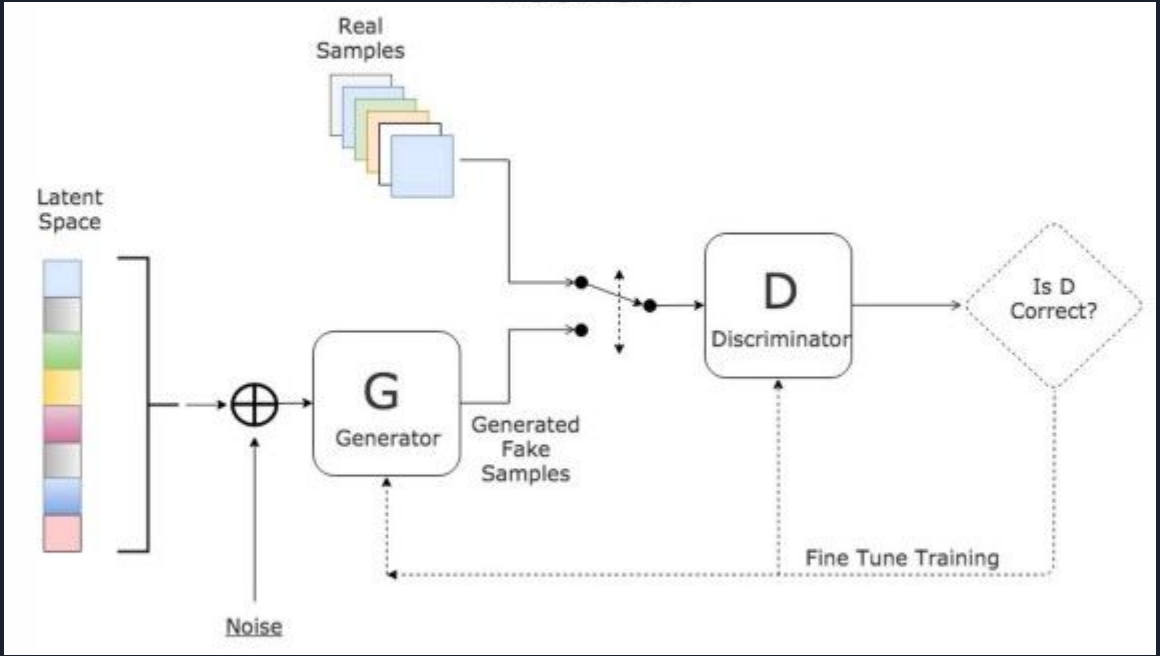
$$\hat{p}(\vec{x})$$

  - **X** is a vector: $\vec{x} = [x_1, x_2, \ldots, x_p]$
- If we now the probability distribution, we can **generate** new samples.

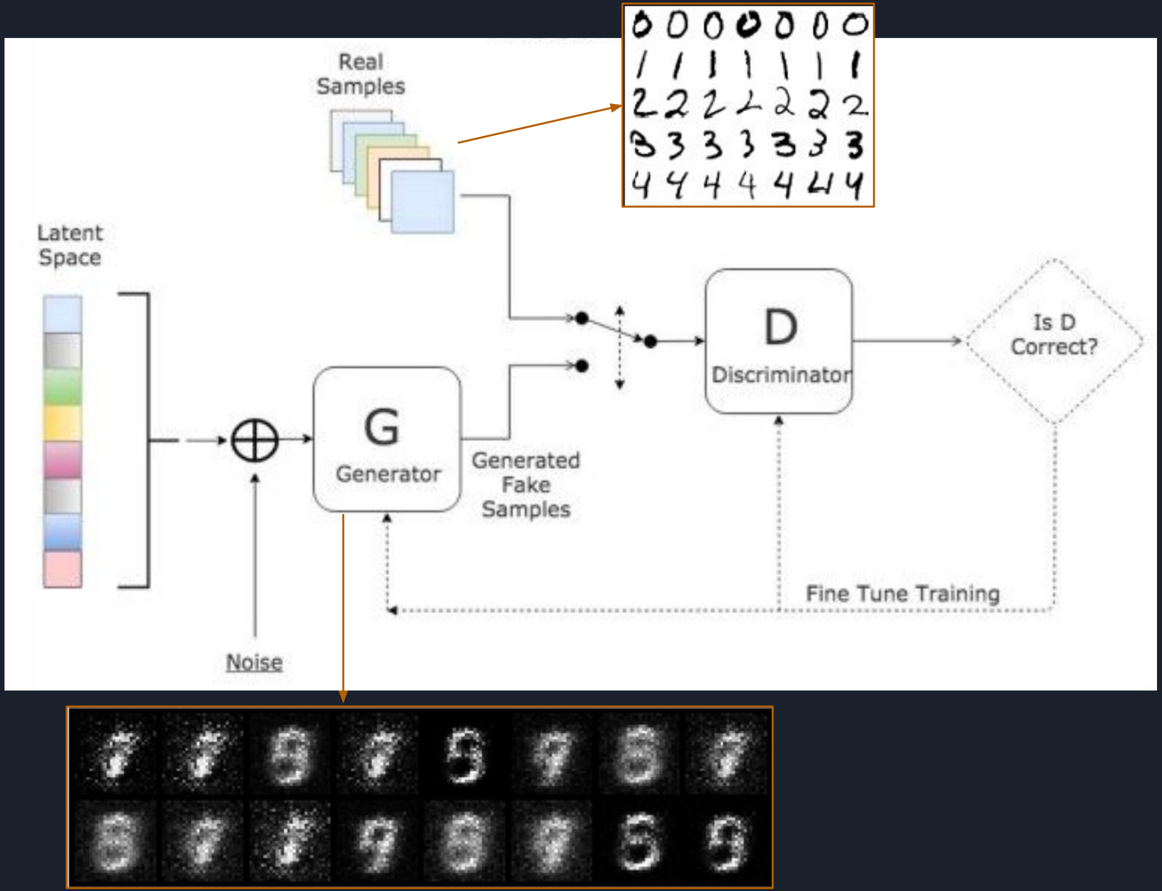
↓

Creating new Samples

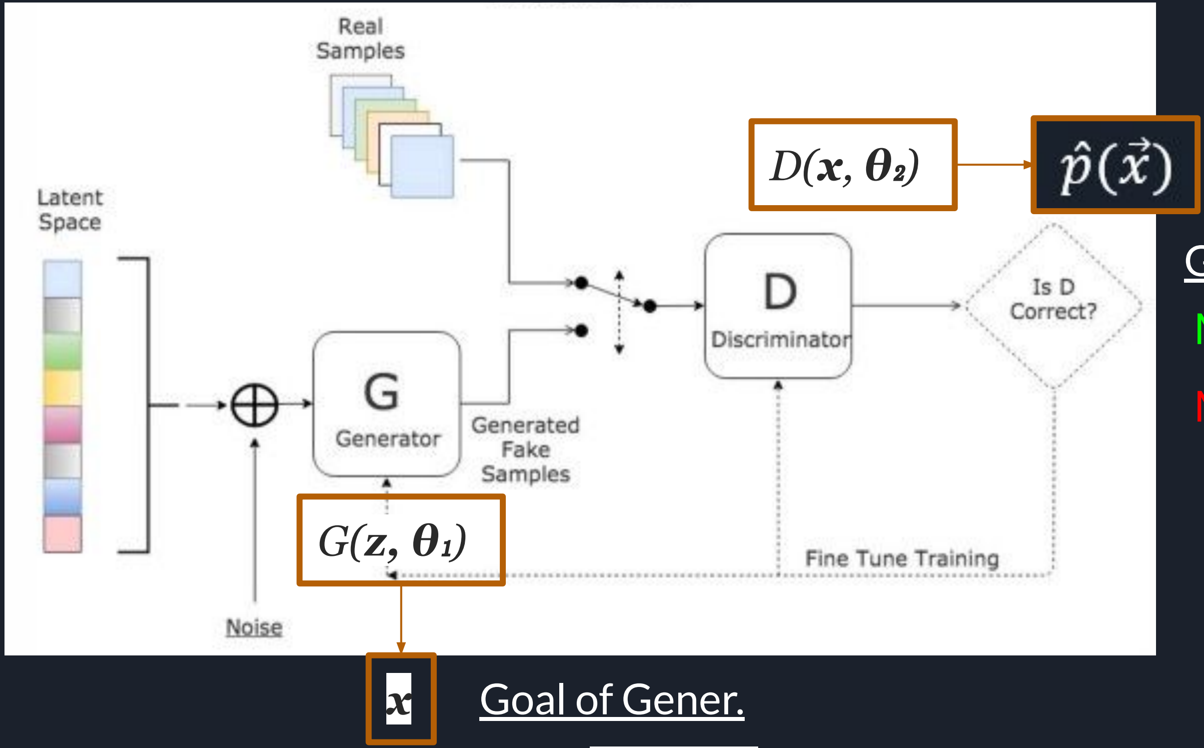
GANs: How do they work?



GANs: How do they work?



GANs: How do they work?



Goal of Discrim.

Max $D(x)$

Min $D(G(z))$

Goal of Gener.

Max $D(G(z))$

GANs: How do they work?

$G(z, \theta_1)$ → x

Goal of Gener.

Max $D(G(z))$

$D(x, \theta_2)$ → $\hat{p}(\vec{x})$

Goal of Discrim.

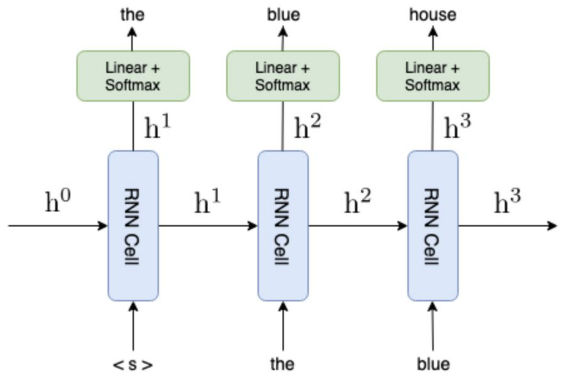
Max $D(x)$

Min $D(G(z))$

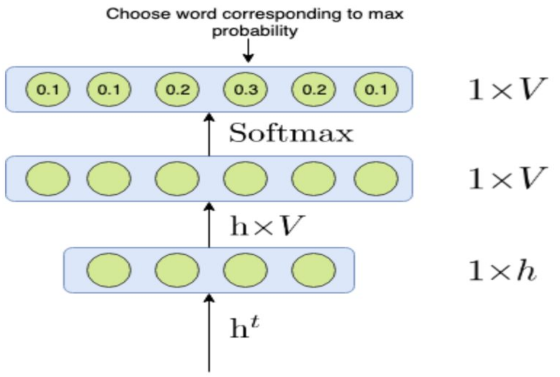
Minimax Game: GANs Loss Function

$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Issues with GANs on Text Generation



RNN Text Generator



Decoding RNN hidden state

At every time step t, the RNN takes the previously generated tokens and the previous hidden state as input and generates the new hidden state ht.

The hidden state is then passed through a linear layer and softmax layer followed by argmax to yield the next word

Why is text generation an Issue for GAN?

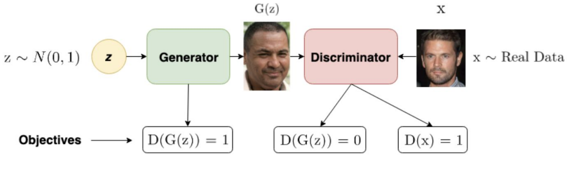


Image GAN

At every step of RNN, we make the choice of next word by picking the word corresponding to the max probability from the output of softmax function. This picking operation is non-differentiable.

It is an issue because for GANS, in order to train the generator to minimize 1 - D(G(z)), we need to feed the output of the generator to the discriminator and back-propagate the corresponding loss of the discriminator.

Because back-propagation relies on the differentiability of all the layers in the network and the picking operation in text generation is non-differentiable, thus it becomes problematic

This is not a problem for image generation because the generated data is continuous.

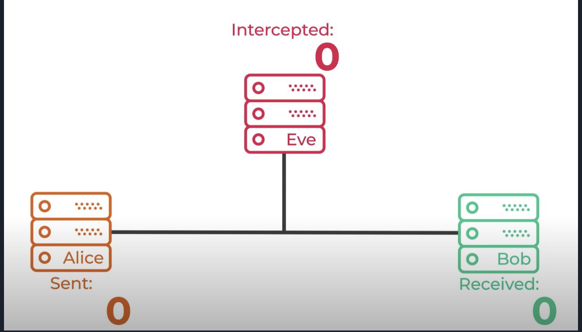
Solutions for generating text

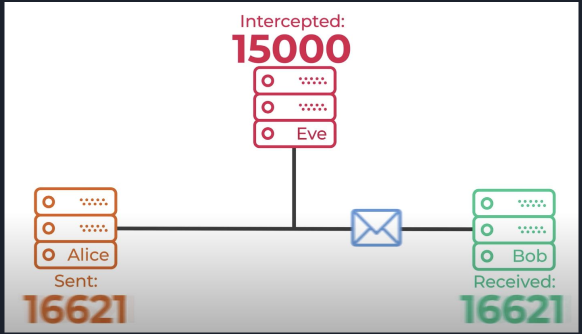
1. The REINFORCE algorithm and policy gradients (Reinforcement Learning-based solutions)
2. The Gumbel-Softmax approximation (A continuous approximation of the softmax function)
3. Avoiding discrete spaces altogether by working with the continuous output of the generator

Use Cases of GAN - Google Brain (2016)

Researcher used GANs to develop a method of encryption

3 neural networks are used in this project, Alice, Bob and Eve





- Continued (Google Brain project)

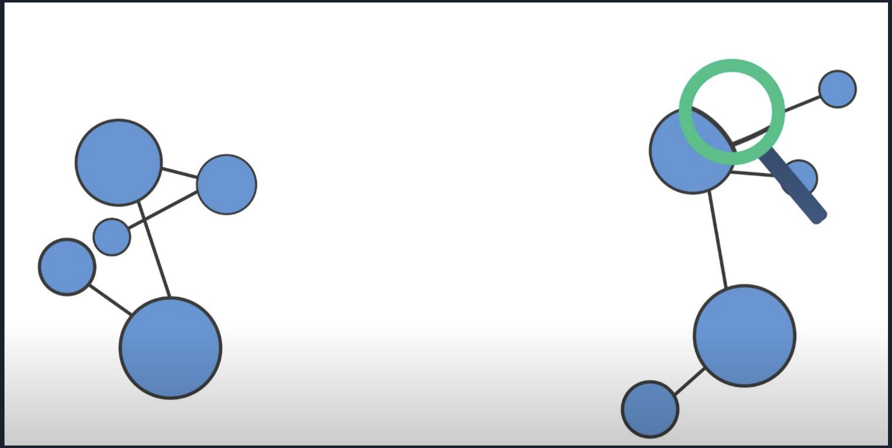
To make sure the message remained secret, Alice had to convert her original plain-text message into complete gobbledygook, so that anyone who intercepted it (like Eve) wouldn't be able to understand it. The gobbledygook – or "cipher text" – had to be decipherable by Bob, but nobody else. Both Alice and Bob started with a pre-agreed set of numbers called a key, which Eve didn't have access to, to help encrypt and decrypt the message.

Initially, the neural nets were fairly poor at sending secret messages. But as they got more practice, Alice slowly developed her own encryption strategy, and Bob worked out how to decrypt it.

After the scenario had been played out 15,000 times, Bob was able to convert Alice's cipher text message back into plain text, while Eve could guess just 8 of the 16 bits forming the message. As each bit was just a 1 or a 0, that is the same success rate you would expect from pure chance.

More Use Cases of GAN

Synthetic Drugs development: The neural net can be trained on existing drug structures and suggest new chemical structures to improve the already existing drug

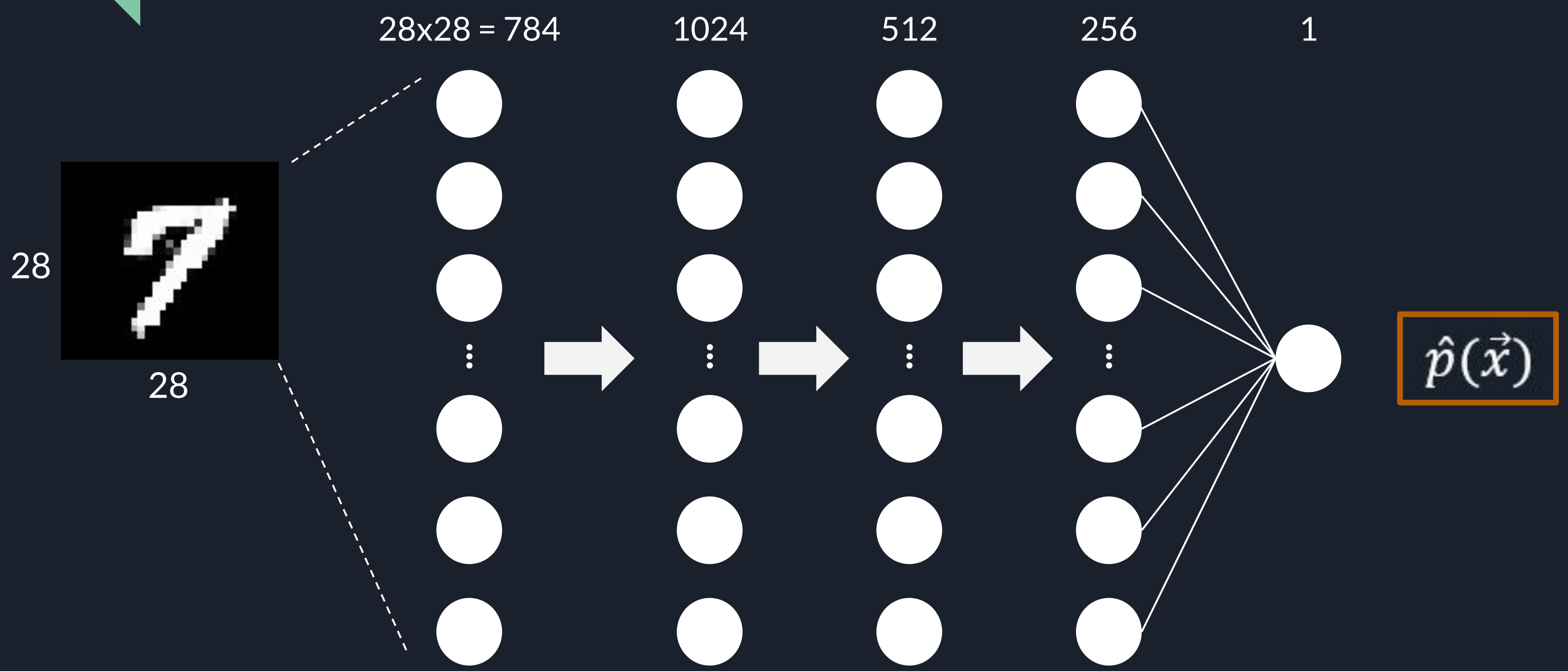


VISIT thispersondoesnotexist.com for more pictures generated by GAN

Fake video generated by AI

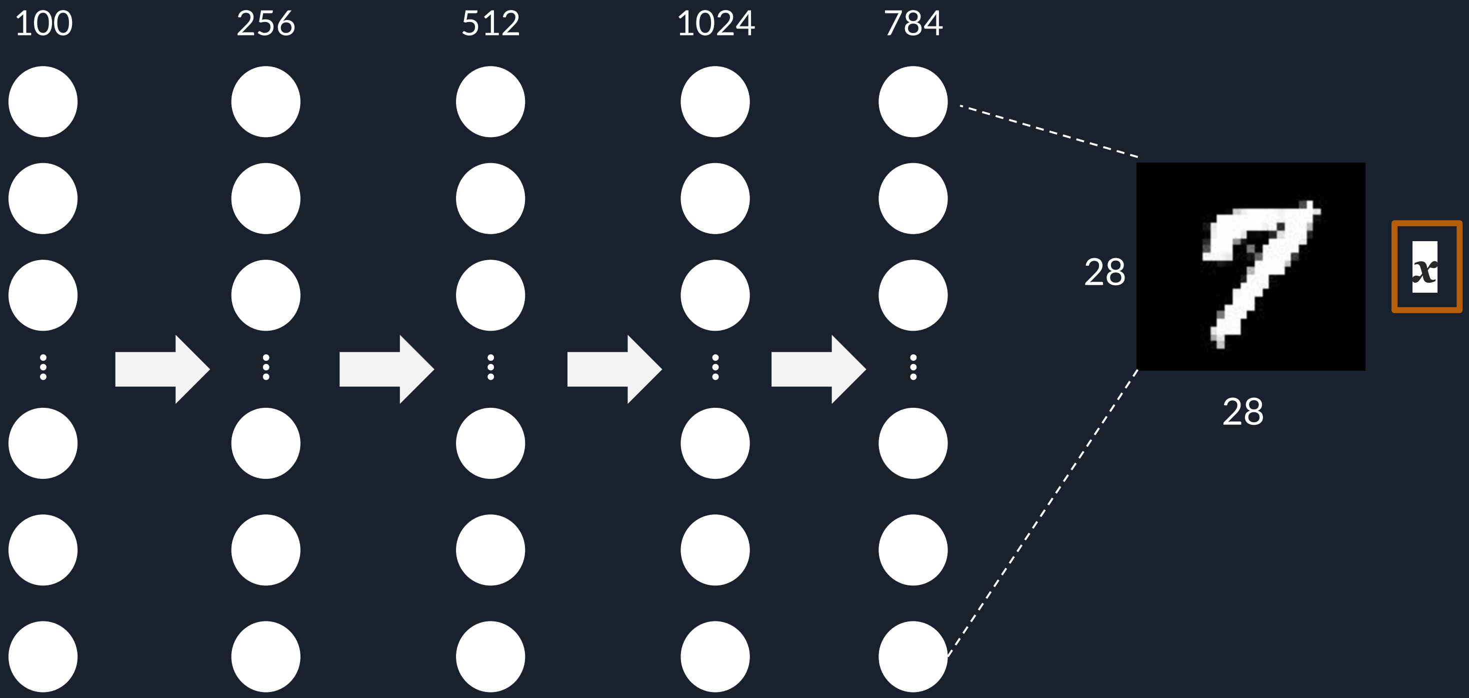
GANs: Our implementation

Discriminator



GANs: Our implementation

Generator



Thanks

Now, we will go to the Coding session.