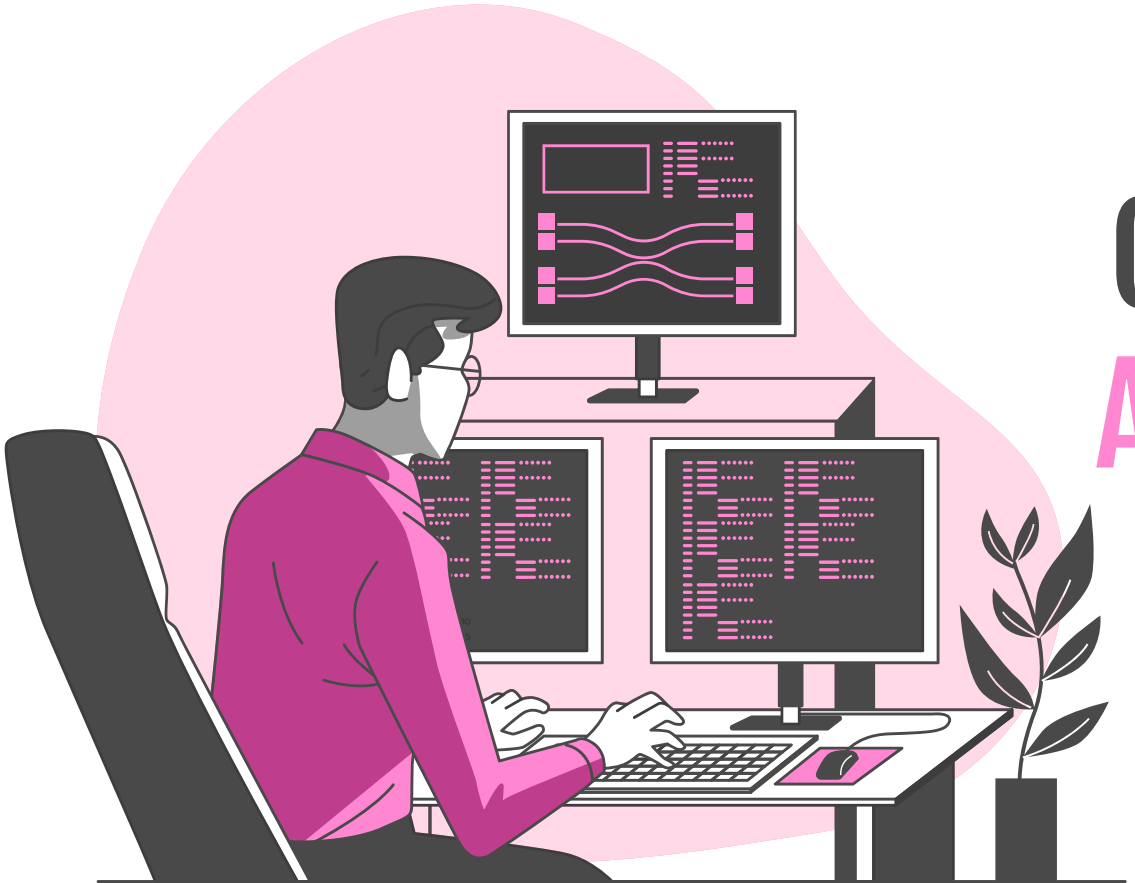
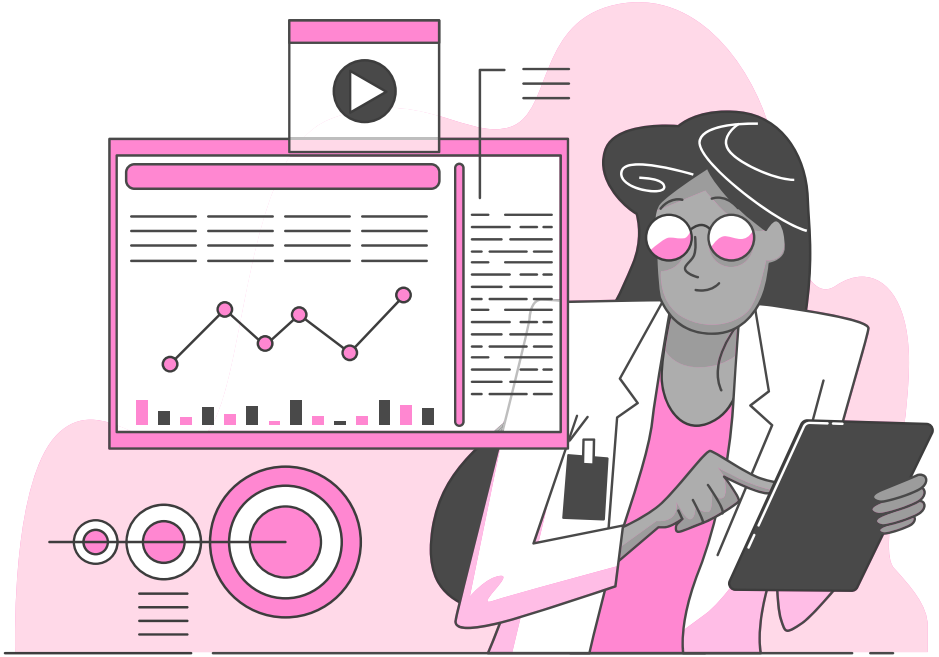


Generative Adversarial Networks

Table of Contents



Referencias:

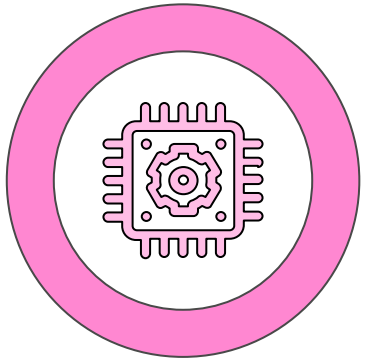
- Generative Adversarial Network - Ian J. Goodfellow, 2014
- GAN — What is Generative Adversarial Networks GAN? - Jonathan Hui, Medium, 2018

⭐ GAN — GAN Series - Jonathan Hui, Medium, 2018

- Modeling Tabular data using Conditional GAN - Lei Xu, 2019

01
GAN

Generative Adversarial Networks



It generates synthetic data

Two Neural Networks:
Generator & Discriminator

Generative Adversarial Networks

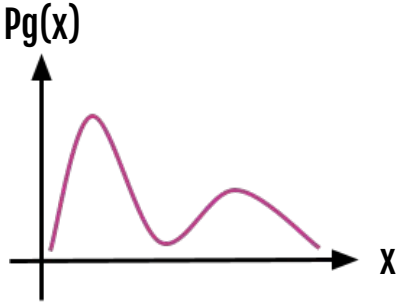
The two networks are trained against each other

02

Training Scheme

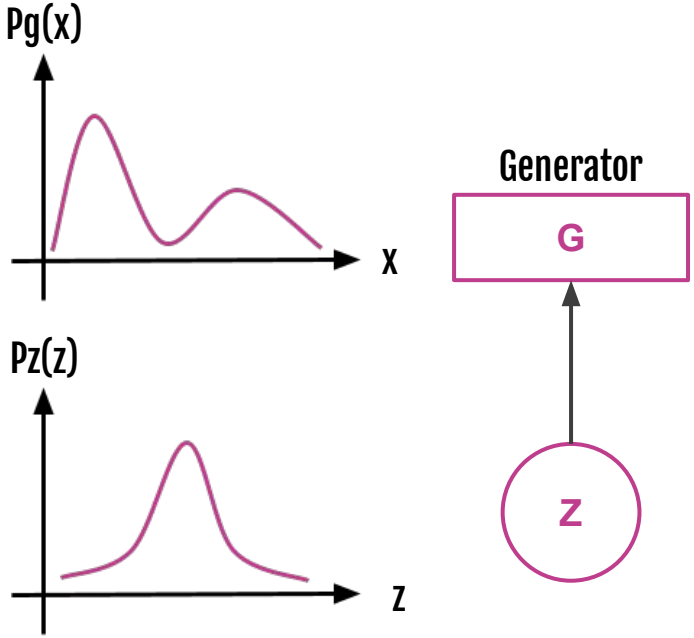
Training Scheme

- Pg(x): Generated data distribution.



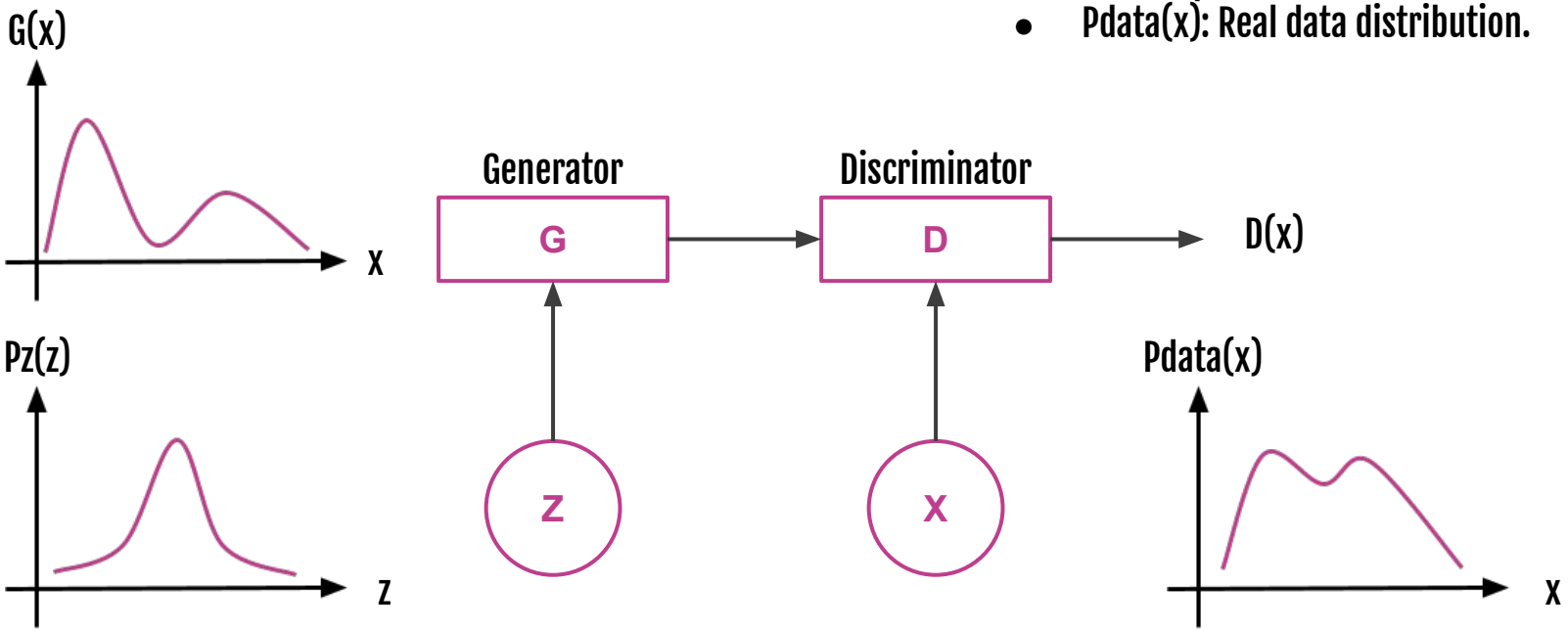
Training Scheme

- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.



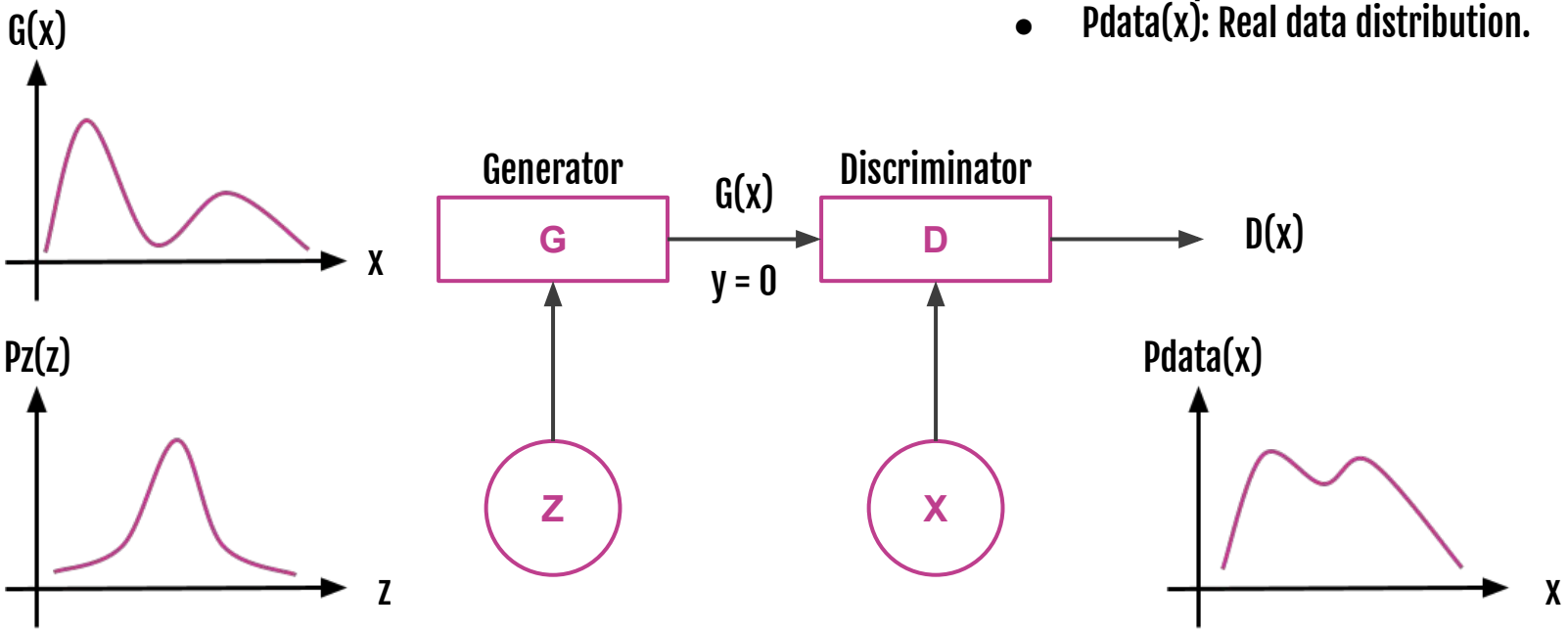
Training Scheme

- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.
- Pdata(x): Real data distribution.



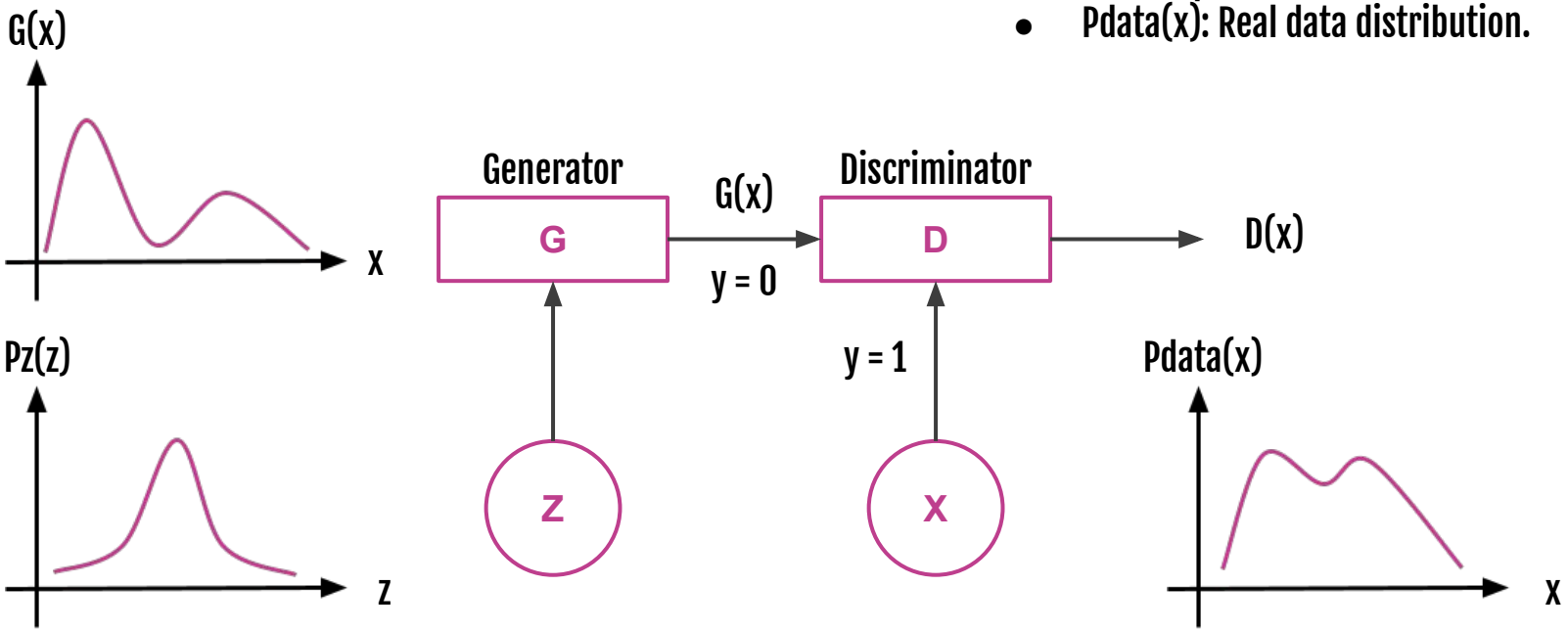
Training Scheme

- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.
- Pdata(x): Real data distribution.



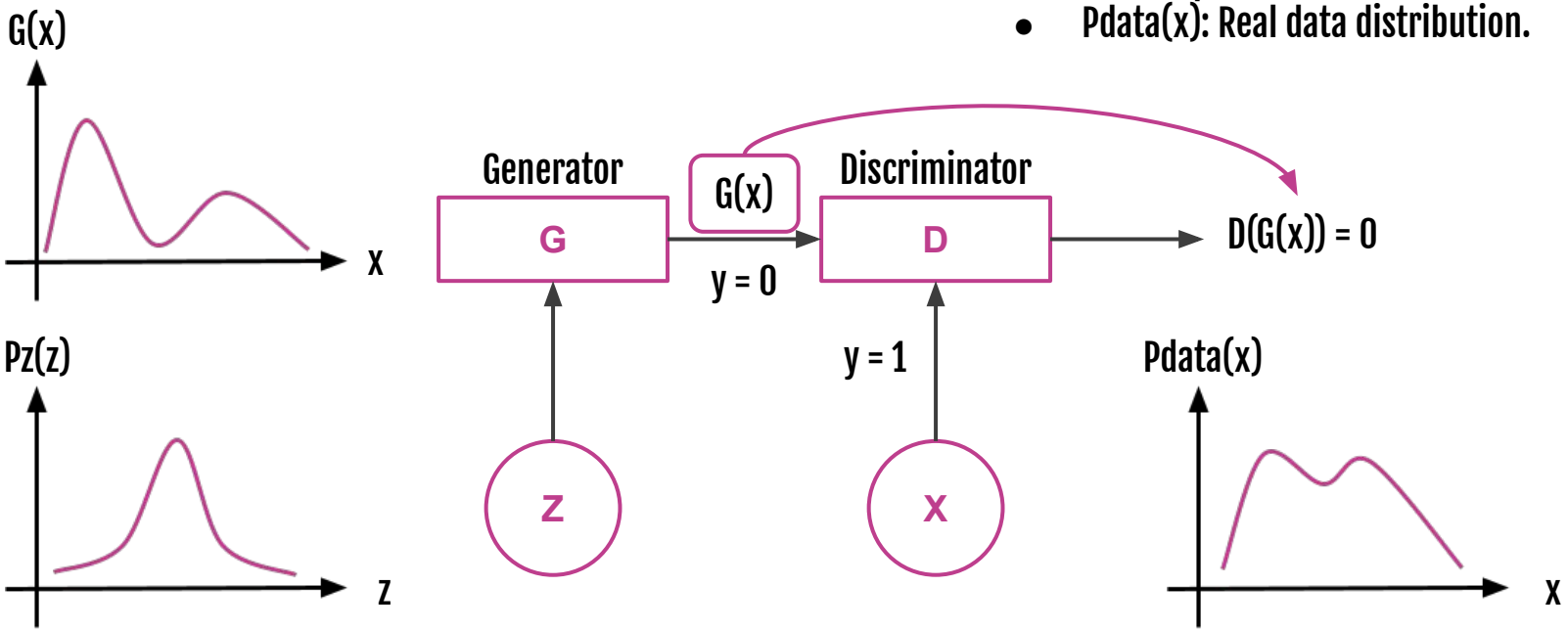
Training Scheme

- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.
- Pdata(x): Real data distribution.



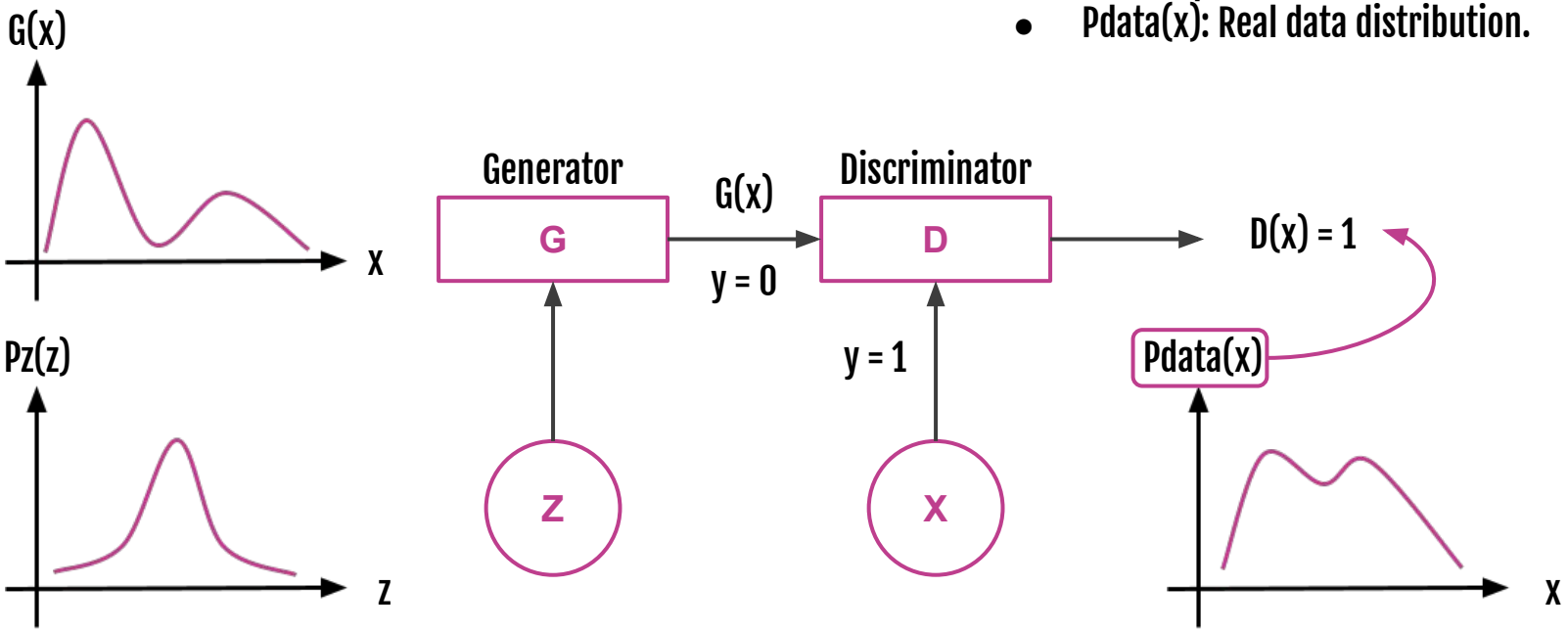
Training Scheme

- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.
- Pdata(x): Real data distribution.



Training Scheme

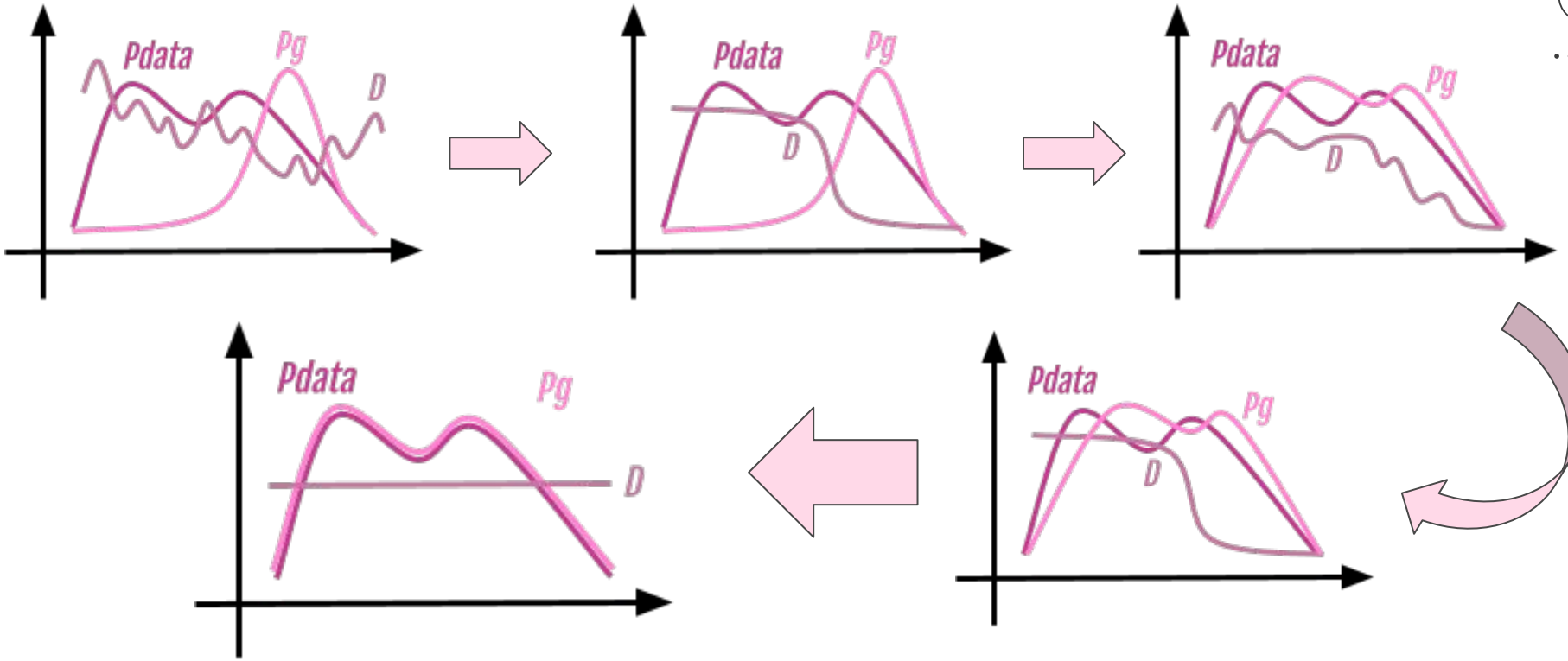
- Pg(x): Generated data distribution.
- Pz(z): Input noise distribution.
- Pdata(x): Real data distribution.



03

Training Sequence

Training sequence



MinMax Game

Value function

Entropía cruzada $= \sum_i p_i \cdot \log q_i$, Entropía Relativa: $D_{KL}(P||Q) = \sum_i P_i \cdot \ln \frac{P_i}{Q_i}$

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

Generator

$$\max_D V(D,G) = \mathbb{E}_{x \sim p_{data}(x)}[ln(D(x))] + \mathbb{E}_{z \sim p_z(z)}[ln(1 - D(G(z)))]$$

Discriminator

$$\min_G V(D,G) = \mathbb{E}_{z \sim p_z(z)}[ln(1 - D(G(z)))]$$

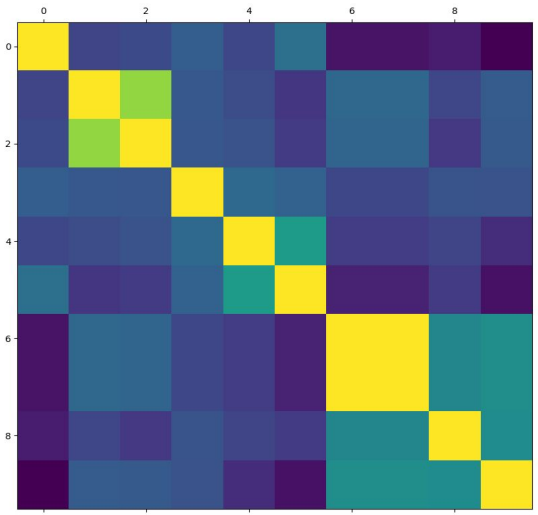
04

Evaluate Synthetic Data

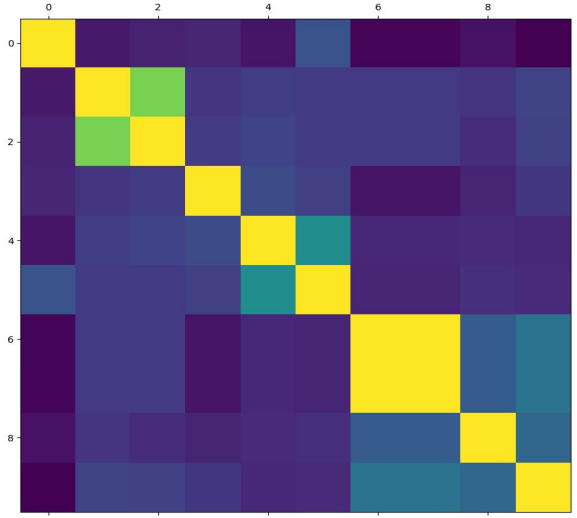
Evaluate Synthetic Data

Correlation Matrix (CM) to show the relations between the variables of the dataset.

Real CM

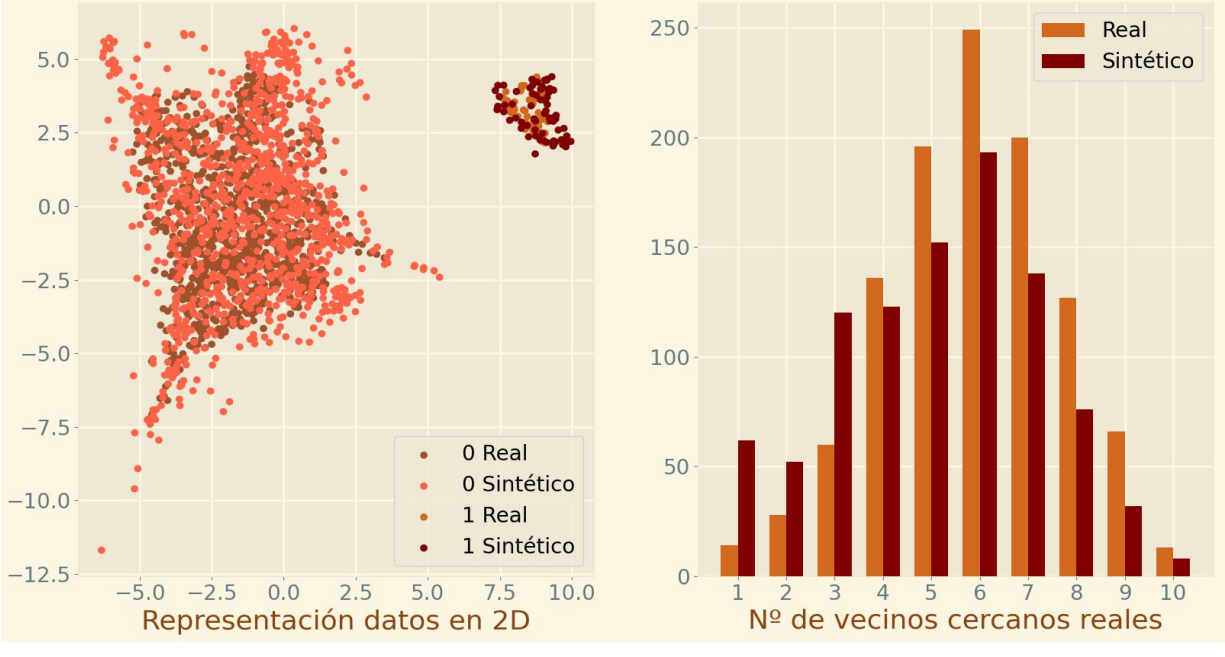


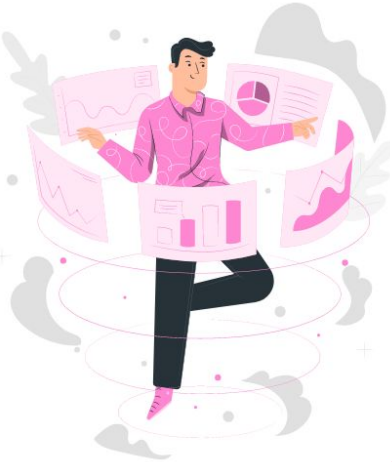
Synthetic CM



Evaluate Synthetic Data

Compress the dataset of high dimension to a 2D representation (UMAP) and show their 10 closest neighbours.





NOW, HANDS ON!