

Modelos de difusión

Elizabeth Toro Chalarca

Daniel Alejandro Garcia Zuluica

Edward Alejandro Rayo Cortez

Universidad EAFIT

Aprendizaje Automático – Semestre 2025-2.

¿Qué son los modelos de difusión?

Definición básica

Son modelos generativos que aprenden a crear datos a partir de ruido aleatorio.

Premisa básica: Convertir los datos en ruido para luego convertir el ruido en datos





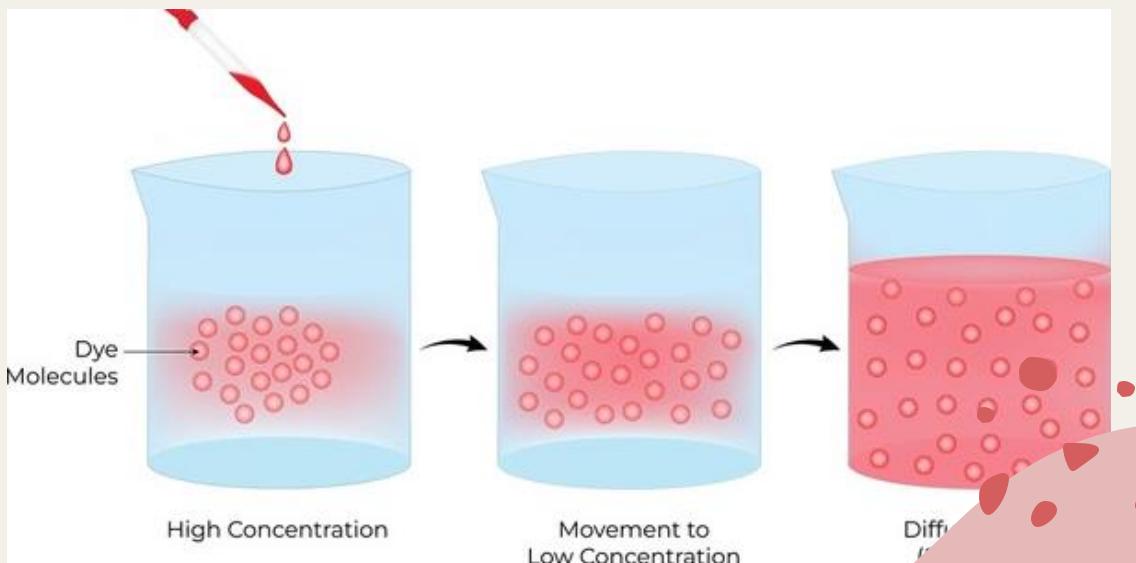
¿Por qué son relevantes?

- Son **modelos generativos avanzados** que producen resultados de alta calidad.
- Superan limitaciones de otros enfoques (GANs, VAEs) en:
Estabilidad del entrenamiento (menos problemas de colapso).
Calidad y diversidad de las muestras.
- Permiten generar contenido **realista y creativo** en múltiples dominios.

Inspiración científica

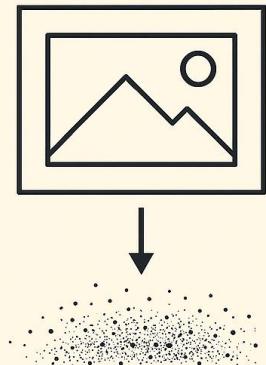
Inspiración en la física

- Basados en procesos como la **difusión molecular**.
- El ruido se comporta como partículas que se dispersan.
- El modelo aprende a revertir esa dispersión.
- Simulan un proceso físico de difusión, donde el ruido se añade progresivamente hasta convertirse en puro ruido y luego se aprende a revertir ese ruido.



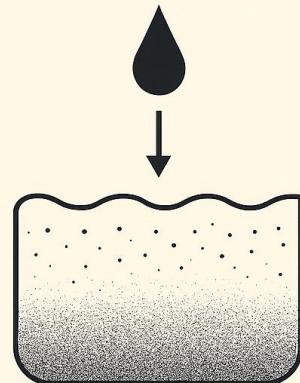
Conceptualmente simples, matemáticamente complejos

PROBABILISTIC (MARKOV CHAINS)



ADDING DUST

SCORE-BASED



ADDING PAINT

Tiene como base dos enfoques matemáticos:

- **Probabilístico (Markov Chains)**

- El proceso de añadir ruido y quitarlo se ve como una **serie de pasos discretos**.
- Modela la densidad completa $p(x)$.
- Necesita normalización (constante Z).

Cómo se calcula la probabilidad inversa.

- **Basado en ecuaciones diferenciales (Score-based).**

- En lugar de pasos discretos, el proceso se describe como **continuo en el tiempo**
- Modela el gradiente $\nabla_x \log p(x)$.
- No necesita normalización.

Cómo entrenar el modelo para estimar la "dirección" correcta en cada paso.

¿Dónde se usan los modelos de difusión?

- En tareas relacionadas con la generación y calidad de imágenes (ej. Stable Diffusion, DALL·E, , Midjourney).
- Generación de audio y video.
- Diseños de fármacos

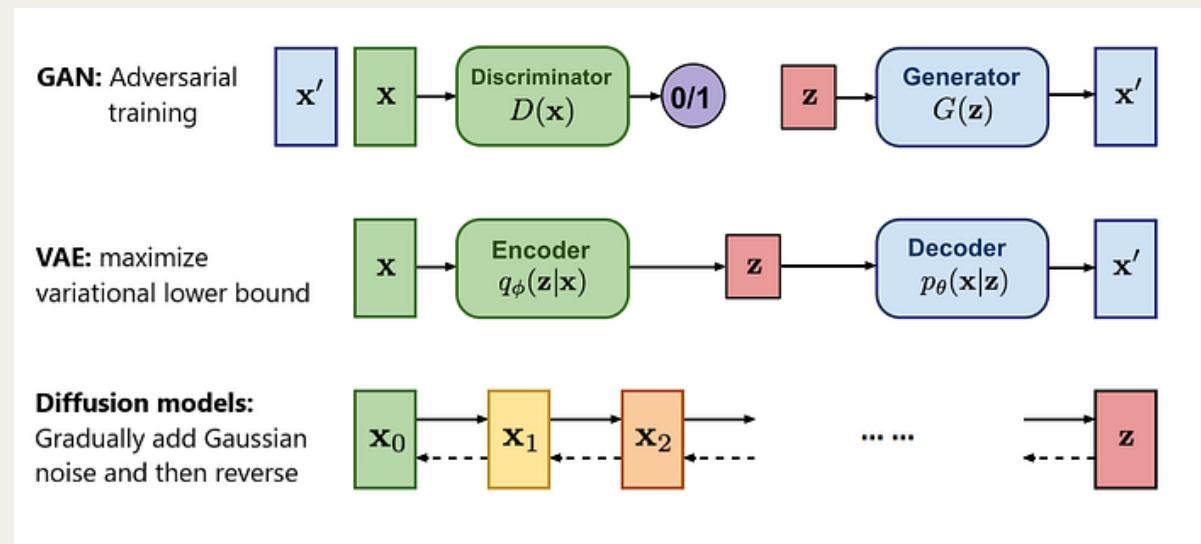


comparación con otros modelos generativos

- **GANs (Redes Generativas Antagónicas):**
Buenas imágenes, pero difíciles de entrenar.
- **VAEs (Autoencoder Variacional):**
Algoritmos más simples, pero menor calidad.
- **Difusión:** Alta calidad, control y estabilidad.

Desventajas:

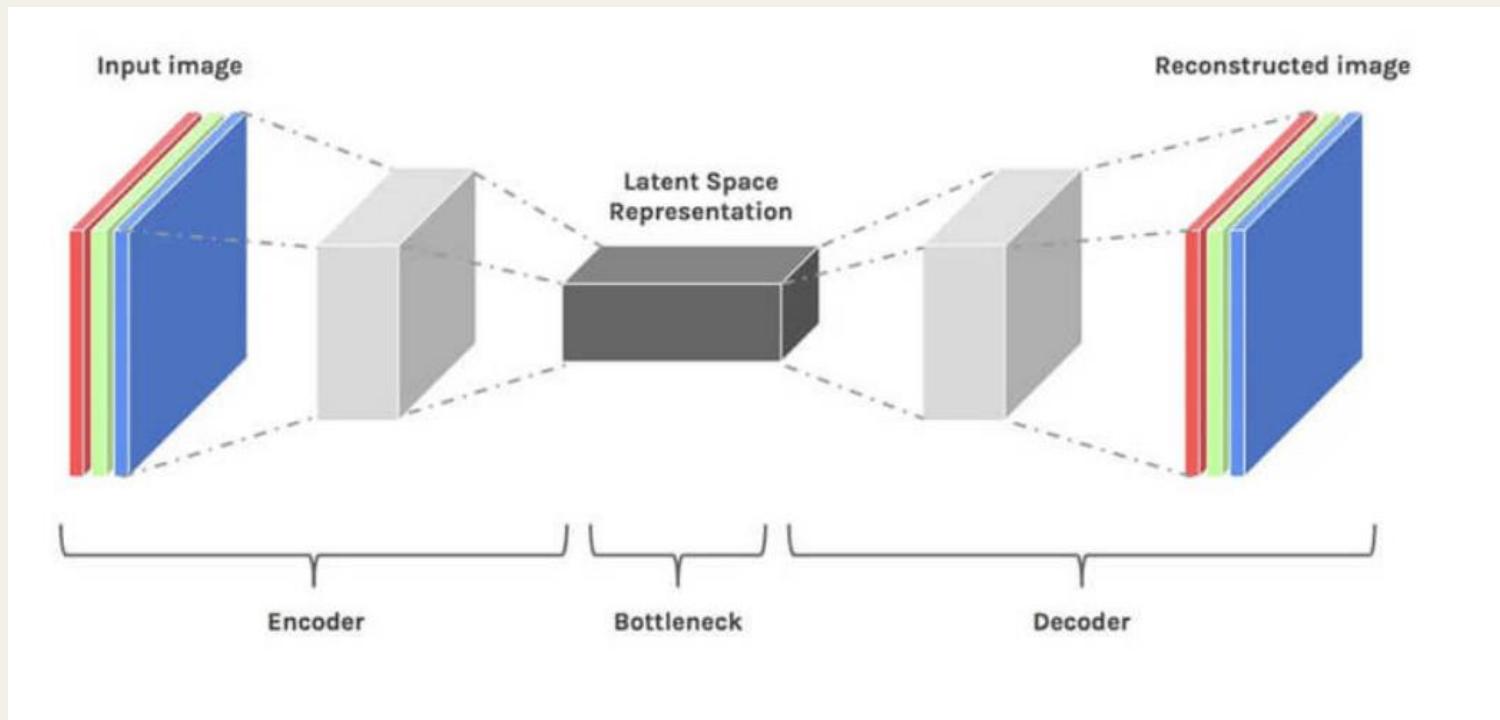
- Más lentos que GANs
- Requieren más recursos computacionales



¿Cómo funcionan?

Proceso de difusión

- **Difusión directa:** se añade ruido paso a paso a una imagen.
- **Difusión inversa:** el modelo aprende a revertir ese ruido para reconstruir la imagen original o generar una nueva.

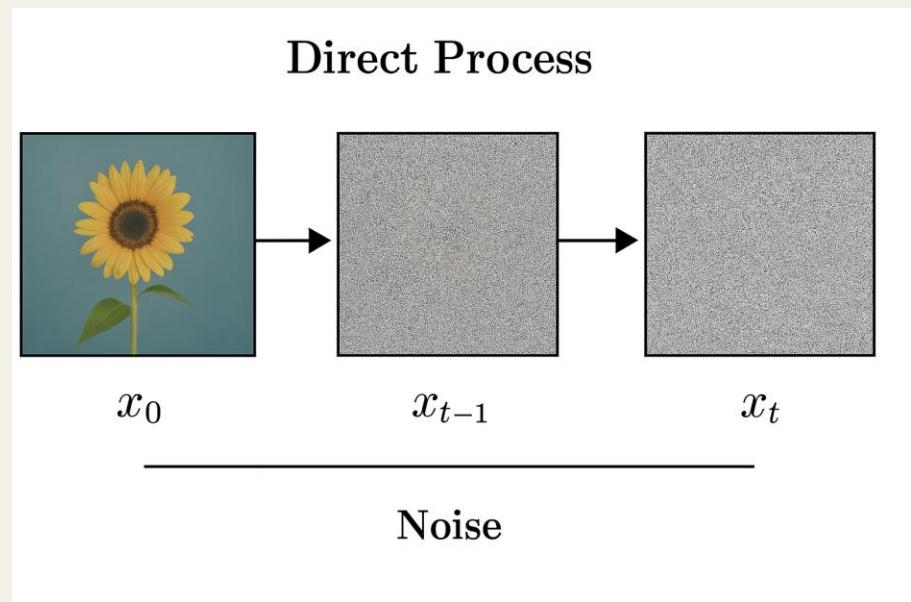


Difusión directa

Este es el proceso de añadir gradualmente ruido a los datos. Comenzando desde los datos originales x_0 , añadimos ruido en T pasos para llegar a ruido puro X_t

cómo se genera el siguiente paso con ruido:

$$q(x_t|x_{t-1}) = N(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I)$$



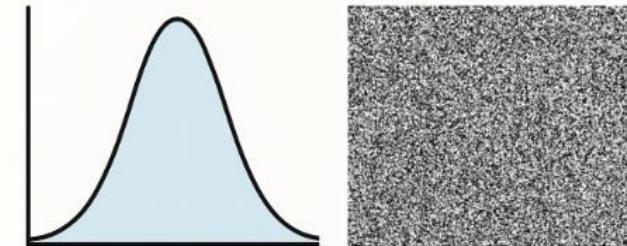
Que tipo de Ruido añadimos?

Ruido gaussiano:

Eso significa:

- Su media es **0** (el ruido no tiende a aumentar ni disminuir la imagen).
- Su varianza es **1** (controlada con el parámetro β_t para cada paso).
- Los valores se distribuyen como una **campana de Gauss** → la mayoría pequeños, pocos muy grandes.

Gaussian Noise



Difusión inversa

Este es el proceso aprendido que remueve el ruido paso a paso, comenzando desde x_t y trabajando hacia atrás hasta x_0 .

reconstruir los datos originales:

$$p(x_{t-1}|x_t) = N(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I)$$

