Guide to Choosing a Generative AI Model Type

Types of generative AI models

Model	Key features	Applications
Generative adversarial networks (GANs)	1. Two competing neural networks: generator and discriminator. 2. The generator learns to create realistic data, while the discriminator learns to distinguish real from fake. 3. The adversarial training process continuously improves both networks. 4. Can be challenging to train and achieve stable results.	1. Image generation: faces, landscapes, objects 2. Text generation: poems, code, scripts 3. Video generation: realistic videos, animation 4. Drug discovery: generate molecules with intended properties 5. Music generation: composing new songs
Variational autoencoders (VAEs)	1. Encode input data into a lower-dimensional latent space 2. Learn a probability distribution over the latent space 3. Decode samples from the latent space to generate new data points 4. Focuses on learning a meaningful representation of the data	1. Image compression: efficiently stores and transmits images 2. Anomaly detection: identify unusual data points 3. Dimensionality reduction: compress high-dimensional data 4. Text summarization: generate concise summaries of text documents
Autoregressive models	1. Generate data point by point, conditioned on previously generated points 2. Use recurrent neural networks (RNNs) or transformers to capture long-term dependencies 3. Can be computationally expensive for long sequences	1. Text generation: realistic and coherent text sequences 2. Music generation: generating music that follows genre and style 3. Time series forecasting: predicting future values of a time series 4. Image inpainting: filling in missing parts of an image
Diffusion models	1. Start with a simple noise and gradually "de-noise" it into realistic data 2. Use a U-Net architecture with skip connections to preserve information 3. Can be more stable and easier to train than GANs, but often slower	1. Image generation: high-quality and diverse images 2. Text generation: coherent and grammatically correct text 3. Audio generation: realistic and musical audio 4. Inpainting and denoising: improving the quality of images or audio
Flow-based models	1. Transform a simple distribution (Gaussian) into a complex one using invertible transformations 2. Learn the parameters of these transformations from the data 3. Can be efficient and accurate for high-dimensional data, but training can be challenging	1. Image generation: realistic and diverse images 2. Density estimation: modeling the probability distribution of data 3. Dimensionality reduction: compress high-dimensional data 4. Anomaly detection: identify unusual data points

Comparison of models on different considerations

Feature	GANs	VAEs	Autoregressive models	Diffusion models	Flow-based models
Data type	Images, text, audio	Images, text, continuous data	Images, text, sequences	Images, text	Images, continuous data
Task objective	High-fidelity generation, data augmentation	Encoding/decoding, representation learning	Sequence generation, text-to-image translation	Image generation, editing, inpainting	Image generation, conditional generation
Quality of samples	High-fidelity, diverse	Often blurry, less realistic	Sharp, high-resolution	High-fidelity, diverse	High-fidelity, controllable
Control over generation	Limited	Moderate	High	Moderate	High
Training complexity	High	Moderate	High	Moderate	High
Interpretability	Low	Moderate	High	Moderate	Low

Author(s)

Abhishek Gagneja