

Guide to Choosing a Generative AI Model Type

Types of generative AI models

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

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

