4/11/25, 10:30 AM about:blank

## **Guide to Choosing a Generative AI Model Type**

## Types of generative AI models

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

## 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

about:blank 1/2

4/11/25, 10:30 AM about:blank



about:blank 2/2