

# Guide to Choosing a Generative AI Model Type

## Types of generative AI models

Model	Key features	Applications
Generative adversarial networks (GANs)	<ol style="list-style-type: none"><li>1. Two competing neural networks: generator and discriminator.</li><li>2. The generator learns to create realistic data, while the discriminator learns to distinguish real from fake.</li><li>3. The adversarial training process continuously improves both networks.</li><li>4. Can be challenging to train and achieve stable results.</li></ol>	<ol style="list-style-type: none"><li>1. Image generation: faces, landscapes, objects</li><li>2. Text generation: poems, code, scripts</li><li>3. Video generation: realistic videos, animation</li><li>4. Drug discovery: generate molecules with intended properties</li><li>5. Music generation: composing new songs</li></ol>
Variational autoencoders (VAEs)	<ol style="list-style-type: none"><li>1. Encode input data into a lower-dimensional latent space</li><li>2. Learn a probability distribution over the latent space</li><li>3. Decode samples from the latent space to generate new data points</li><li>4. Focuses on learning a meaningful representation of the data</li></ol>	<ol style="list-style-type: none"><li>1. Image compression: efficiently stores and transmits images</li><li>2. Anomaly detection: identify unusual data points</li><li>3. Dimensionality reduction: compress high-dimensional data</li><li>4. Text summarization: generate concise summaries of text documents</li></ol>
Autoregressive models	<ol style="list-style-type: none"><li>1. Generate data point by point, conditioned on previously generated points</li><li>2. Use recurrent neural networks (RNNs) or transformers to capture long-term dependencies</li><li>3. Can be computationally expensive for long sequences</li></ol>	<ol style="list-style-type: none"><li>1. Text generation: realistic and coherent text sequences</li><li>2. Music generation: generating music that follows genre and style</li><li>3. Time series forecasting: predicting future values of a time series</li><li>4. Image inpainting: filling in missing parts of an image</li></ol>
Diffusion models	<ol style="list-style-type: none"><li>1. Start with a simple noise and gradually "de-noise" it into realistic data</li><li>2. Use a U-Net architecture with skip connections to preserve information</li><li>3. Can be more stable and easier to train than GANs, but often slower</li></ol>	<ol style="list-style-type: none"><li>1. Image generation: high-quality and diverse images</li><li>2. Text generation: coherent and grammatically correct text</li><li>3. Audio generation: realistic and musical audio</li><li>4. Inpainting and denoising: improving the quality of images or audio</li></ol>
Flow-based models	<ol style="list-style-type: none"><li>1. Transform a simple distribution (Gaussian) into a complex one using invertible transformations</li><li>2. Learn the parameters of these transformations from the data</li><li>3. Can be efficient and accurate for high-dimensional data, but training can be challenging</li></ol>	<ol style="list-style-type: none"><li>1. Image generation: realistic and diverse images</li><li>2. Density estimation: modeling the probability distribution of data</li><li>3. Dimensionality reduction: compress high-dimensional data</li><li>4. 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



# Skills Network