Generative AI Model Categories

1. Language Models (LLMs)

Generate or understand human language.

Examples: GPT-4, Claude 3, Gemini, LLaMA 3, Mistral

2. Embedding Models

Convert text or images into vector form for search or clustering.

Examples: text-embedding-3, bge, e5, MiniLM

3. Code Generation Models

Generate, explain, or complete code.

Examples: Codex, StarCoder2, Code LLaMA, DeepSeek-Coder

4. Multimodal Models

Understand or generate from multiple input types (text + image).

Examples: GPT-4o, Gemini 1.5, Claude 3 Opus

5. Vision Models

Understand visual content.

Examples: CLIP, DINOv2, SAM, BLIP

6. Text-to-Image Models

Generate images from text prompts.

Examples: DALL·E 3, Stable Diffusion, Midjourney

7. Image-to-Text Models

Generate descriptions of images.

Examples: BLIP, GPT-4o (vision), Gemini

8. Text-to-Video Models

Generate videos from text.

Examples: Sora (OpenAI), Runway Gen-2, Pika

Generative AI Model Categories

9. Speech Models

Speech-to-text or text-to-speech models.

Examples: Whisper, Bark, Deepgram, ElevenLabs

10. Audio Generation Models

Create music or audio from text or prompts.

Examples: MusicGen, AudioCraft

11. Conversational Agents

Built for chat and dialogue.

Examples: ChatGPT, Claude, Gemini Chat, Pi

12. RAG Models / Frameworks

Combine search and generation.

Examples: LangChain, LlamaIndex, Haystack

13. Agentic Models

Execute multi-step tasks autonomously.

Examples: AutoGPT, BabyAGI, OpenAI GPTs

14. Diffusion Models

Used in image/audio generation via denoising steps.

Examples: Stable Diffusion, Imagen

15. Fine-Tuning & Adapter Models

Tune models efficiently for specific tasks.

Examples: LoRA, QLoRA, PEFT