

NVIDIA AI Foundation Models - PCs and Workstations

The following NVIDIA AI Foundation Models can be used on PCs and Workstations with NVIDIA GPUs without requiring an NVIDIA NIM in production.

Model	Variants
ChatGLM3	6B
CLIP	400M
Code Llama	7B, 13B
Gemma	3B, 7B
Llama 2	7B, 13B
Llama 3	8B
Llama 3.1	8B
Llama 3.2	3B
Mistral	7B
Phi 2	2.7B
Phi 3	Mini, Small, Medium, Vision
Qwen 2	7B
Whisper	Small, Medium

(v. September 25, 2024)