

Smart Healthcare and Pharmacy: Building LLM and RAG Solutions with NVIDIA NeMo

Po - Fu Kuo

Accelerate Private Machine Intelligence Company AI Enterprise Team LLM Engineer

With the advancement of AI technology, Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), and Vision-Language Models (VLMs) are profoundly transforming the fields of medicine and pharmacy. These technologies empower healthcare professionals with rapid information retrieval, clinical decision support, medical image analysis, medical literature interpretation, and even the optimization of drug development and precision medicine.

This talk will introduce real-world applications of NVIDIA NeMo in medical AI development, including:

- Fine-tuning domain-specific LLMs: Optimizing models with medical data to enhance diagnostic recommendations and clinical decision-making accuracy.
- Integrating RAG with medical knowledge bases: Combining real-time retrieval with generative AI to improve the reliability and explainability of medical text-based information.
- VLM applications in medical imaging and text: Leveraging multimodal AI to enhance pathology image analysis, medical scan diagnostics, and automated report generation.

