Product Brief

ASRock Industrial AI LLM Developer Kits for Enterprise AI application

Advanced Edge AI System for Large Language Models

iEPF-10000S Series

iEPF-9030S Series

NUC Ultra 100 BOX Series

NUCS Ultra 100 BOX Series

About ASRock Industrial

ASRock Industrial is a leading provider of edge computers and industrial motherboards. Our vision is to co-create an intelligent world with vertical applications in CARES (Commerce, Automation Robot, Entertainment, Security) industries. ASRock Industrial works with systems integrators and independent software vendors to provide high-performance and flexible total solutions.

Develop and Train AI LLM Models with Ease

Developing and training large language models (LLMs) can be a daunting task due to their immense size, computational demands, and efficient memory management requirements. The majority of AI LLMs depend on cloud computing, rendering them incapable of executing offline or real-time processing tasks.

ASRock Industrial created AI LLM Developer Kits to ease the development process of private LLMs. Powered by 13th and 14th Gen Intel® Core™ processors and Intel® Core™ Ultra processors along with advanced AI technologies, these kits enable seamless execution of LLM tasks like multimodal language translation, text summarization, and question-answering. With applications spanning smart chatbots, interactive signage/ kiosks, smart factories, retail environments, transportation systems, and medical facilities, these kits provide a solid foundation for tackling diverse edge AI challenges with ease and efficiency.

Options to Meet the Needs of Different LLM Tasks

ASRock Industrial's AI LLM Fine-Tuning Kits feature the iEPF-10000S and iEPF-9030S Series edge training servers, powered by 13th and 14th Gen Intel® Core™ processors, along with advanced AI fine-tuning LLM technologies. These kits enable users to develop and train their own AI LLMs via the fine-tuning method, addressing specific task-oriented problems often missed by global performance models. The fine-tuning kits offer significant benefits, including reduced computation costs and access to state-of-the-art models without starting from scratch.

Meanwhile, ASRock Industrial's AI LLM Inference Kits provide a streamlined approach to improving model accuracy using the retrieval-augmented generation (RAG) method. The RAG method kits feature the NUC Ultra 100 BOX, NUCS Ultra 100 BOX, and NUC Ultra 100 MB Series, powered by Intel® Core™ Ultra processors, and leverage OpenVINO™ toolkit. These kits empower users to develop AI LLMs with enhanced accuracy without the need for re-training models, thereby increasing cost-effectiveness.

Benefits of Fine-Tuning Approach

- **Restricting the scope and context.** By inputting a defined data set and context, the model will be more specific for specific use cases.
- **Improved efficiency.** The model can run faster as it does not need to be reminded again on the context.
- **Running LLM on the edge.** Eliminating cloud upload helps maintain data privacy and avoid cloud service provider costs.

Benefits of RAG Approach

- **Up-to-date responses**. RAG ensures that responses are based on up-to-date external sources instead of static, stale training data.
- **More accurate responses.** The output is grounded on relevant, external knowledge that can include citation of original sources, allowing human verification.
- **Relevant, domain-specific responses.** The LLM can provide contextually relevant responses tailored to an organization's proprietary or domain-specific data.
- **Efficient and cost-effective.** Organizations can deploy RAG without customizing the model.

Intel Ingredients

14th and 13th Gen Intel® Core™ Processor

Optimized computing power with a performance hybrid architecture, delivering computing power to enable productivity and creativity for users.

Intel® Core™ Ultra Processor

Features a specialized AI processing unit alongside supercharged performance to support advanced, innovative use cases and demanding applications.

OpenVINO™ Toolkit

An open source toolkit that accelerates AI inference with lower latency and higher throughput while maintaining accuracy, reducing model footprint, and optimizing hardware use.

Learn More

- Intel® Foundational Developer Kits
- iEPF-10000S Series
- iEPF-9030S-EW4
- NUC Ultra 100 BOX Series and NUCS Ultra 100 BOX Series
- NUC Ultra 100 MB Series

Intel® Foundational Developer Kits allow you to get started on your targeted application development with a superior out-of-the-box experience. Deploy your application at scale by building customized systems via Intel ecosystem partners.

Performance varies by use, configuration, and other factors. Learn more on the Performance Index site.

No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software, or service activation.