

Сервер SNR-SR4410RG с LLM Deepseek-671B, 8U, 1024GB RAM, 768GB vRAM, резервируемый БП

SNR-SR4410RG-DPSK-671B

Описание

Готово к использованию

- Модель Deepseek предустановлена и поставляется в интегрированном виде — программное обеспечение и оборудование в одном решении.

Быстрый запуск

- Предоставляем подробное руководство по дообучению и развертыванию модели, что снижает порог внедрения и позволяет оперативно перейти к использованию приватной модели.

Надёжность и безопасность

- Локальное развертывание на стороне заказчика. Все данные остаются внутри компании, что полностью исключает риск утечки.

Технические характеристики SNR-SR4410RG-DPSK-671B:

Модель LLM	Deepseek R1-671B, Deepseek V3-671B
Скорость вывода	До 1400 токенов в секунду для 100 одновременных подключений
Размер платформы	8U
Видеопамять	768GB GDDR6X
GPU	16x NVIDIA RTX4090
Вычислительная мощность AI	Производительность FP8 – 10560 TFLOPS
CPU	4 процессор по 32 ядра
RAM	1024GB
Сетевой интерфейс	2 порта 25Gbps Base-X SFP28

Идеальное решение для различных задач:

Корпоративные сценарии

- База знаний: Создание бизнес-агентов, устранение инфобарьеров между отделами
- Документы: Проверка тендеров, контрактов и юридических файлов
- Производство: Анализ и оптимизация процессов, поддержка итераций продукта

Сфера общественной безопасности

- Анализ дел: Прогноз криминальных очагов по времени, месту, типу
- Антифрод: Выявление мошеннических паттернов в звонках и сообщениях
- Мониторинг соцсетей: Отслеживание ключевых слов, связанных с рисками

Высшее образование

- Наука: Автосоздание графов знаний и поиск связей в экспериментах
- Публикации: Подбор тем и улучшение логики научных работ
- Междисциплинарность: AI в фин. стратегиях и биоинформатике (редактирование генов)

Госсектор

- Соцзащита: Анализ потребностей уязвимых групп
- Бизнес: Льготные меры и сопровождение заявок.
- Строительство: Градплан и контроль безопасности
- Контроль качества: Общественное мнение о пищевой и фарм. безопасности.

Медицина

- Больницы: Диагностика по КТ/МРТ + тексту
- Биология: Анализ связей между геномом и метаболомом
- Фармакология: Виртуальный скрининг и оптимизация молекул
- Эпидемиология: Прогнозы распространения заболеваний во времени и пространстве

Школьное образование

- AI-ассистент: Автогенерация уроков и заданий по программе
- Проверка работ: Распознавание формул и логики решений
- Психология: Оценка состояния по данным поведения

Общие

Назначение сервера

GPU сервер