KNERON-KN330P-001

KNEO330 Plus Specification

KNEO330 Plus

info @ kneron.us

Copyright Notice

©2025 Kneron Inc. All rights reserved. This material and documentation contain confidential and proprietary information that is the property of Kneron Inc. The material and documentation are furnished under a license agreement and may be used or copied only in accordance with the terms of the license agreement. No part of the material and documentation may be reproduced, transmitted, or translated, in any form or by any means, electronic, mechanical, manual, optical, or otherwise, without prior written permission of Kneron Inc., or as expressly provided by the license agreement.

Export Control Statement

All technical data contained in this publication is subject to the Taiwan export control laws, and may be subject to export or import regulations in other countries ("Export Law"). Disclosure to nationals of other countries contrary to Export Law is prohibited. It is the reader's responsibility to determine the applicable regulations and to comply with them.

Trademarks (™)

KNERON is a trademark of Kneron Inc.

ARM® and AMBA® are registered trademarks of ARM® Limited.

All other product or company names may be trademarks of their respective owners.

info @ kneron.us

San Diego HQ
10052 Mesa Ridge Court, Suite 202,
San Diego, CA 92121, U.S.A.
https://www.kneron.com/en/

Taipei Office
12F-1., No. 386, Sec. 6, Nanjing E. Rd.,
Neihu Dist. (114), Taipei City, Taiwan, R.O.C.
https://www.kneron.com/tw/

Overview

The KNEO330 Plus is an edge computing-based AI server developed by Kneron™ Inc., specifically designed for efficiently processing large-scale language model tasks such as GPT. Below are its key features and advantages:

Features

1. Edge Computing Capabilities:
    - Supports running large-scale AI models at the edge, reducing reliance on cloud computing resources.
    - Delivers low-latency, high-efficiency inference capabilities, ideal for real-time applications.
    - Equipped with Kneron's proprietary NPU (Neural Network Processor) to optimize AI model execution efficiency.
    - Supports multiple deep learning frameworks and models, including the GPT series.
2. Energy Efficiency and Cost Optimization:
    - Reduces cloud computing costs through edge processing.
    - Hardware design prioritizes power optimization, suitable for deployment in energy-constrained environments.
3. Security:
    - Edge computing eliminates data uploads, safeguarding data privacy and model security.

Advantages

- Low Latency: Process data at the edge to reduce transmission time.
- Efficient Deployment: Supports localized deployment without reliance on remote cloud services.
- Dynamic External Data Updates: To maintain data freshness, Kneron's solution supports asynchronous file updates and refreshes embedded representations, ensuring users receive the latest and most accurate information.
- Preload knowledge bases relevant to the application domain to reduce hallucinations and lower user requirements for engineering background.
- Implement keyword blocking to prevent erroneous responses

info @ kneron.us

Hardware Specifications

Items	Description
Product Name	KNEO330 Plus Edge GPT Server Edge Computing AI Inference Server, supporting GPT model operation
CPU	Intel® i5 10 cores 16 threads 4.6GHz or higher
NPU	Kneron™ embedded KNEO series NPU
AI Computing Power	400 TOPS equivalent
LLM Support	Kneron™ LLM model inside. Token size: 16K
Memory	32GB DDR4
Storage	2TB NVMe SSD
Power Supply	100-240V 50/60Hz, Max 850W
Operating System	Ubuntu Linux
Dimensions	428 x 350 x 89mm (rack mount 2U)
Cooling System	• Air cooling (fan) • Heat dissipation design: ◦ High-efficiency fan module ◦ Aluminum heat sink
Operating Temperature	0°C to 40°C
Indicator lights	• Power Input • Hard Drive Access
External Interfaces	Front Panel: • Power Switch*1 • System Reset Button*1 • USB2 TYPE A *2 Rear Panel: • Power Cable Jack*1 • USB TYPE C *1 • USB3.2 TYPE A *4 • Gigabit Ethernet Port*1 (RJ45) • HDMI *1 • Audio Input/Output Ports ◦ MIC *1 ◦ LINE OUT* ◦ SPDIF *1

info@kneron.us

info @ kneron.us