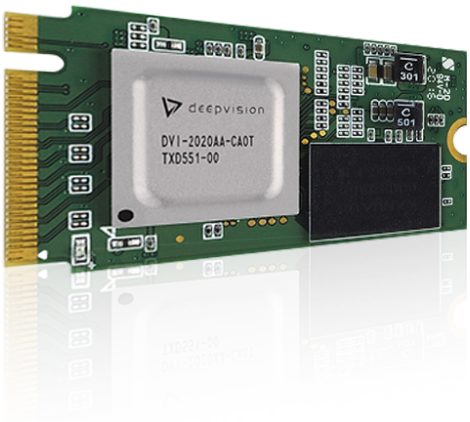
kinara

ARA-1 M.2 MODULE

BREAKTHROUGH AI PERFORMANCE



Kinara Ara-1 M.2 modules deliver plug-and-play compatibility to PCIe-enabled host systems. Start running inferences in minutes with a fully integrated module that includes the Kinara Ara-1 Edge AI processor, local memory and support circuitry.

kinara

EASILY DEPLOY AI AT THE EDGE

The Kinara Ara-1 M.2 module enables high-performance, low-power devices.

Its advanced technologies mean you can quickly build and run AI models. When combined with the Kinara SDK, these powerful M.2 modules are ideal for full-scale commercial deployment or rapid prototyping for new AI models.

Versatile M.2 Form Factor and Host System Integration
Design into everything from notebooks to traditional PCs, to Intel® NUCs or embedded boards with Arm® processors. A Kinara-provided Linux driver supports runtime communication between most Linux-based host systems and the M.2 module.

Increase Performance by Offloading Inferencing
Like all Kinara AI acceleration products, the Ara-1 M.2 modules can be designed into a wide variety of applications including smart retail, smart city, industrial automation, robotics, and automotive.

While the host system to performs all pre- and post-processing functions, the Ara-1 M.2 module handles the application's inferencing requirements. When applications require high-performance, low-power AI acceleration, let the Ara-1 M.2 module offload AI inferencing from the host system.

Integrate with PCIe Plug and Play

Exceptional performance/watt

Common frameworks and the Kinara SDK

No Cloud compute needed

SPECIFICATIONS	
AI Model Frameworks Supported	TensorFlow, PyTorch, MxNet, ONNX, Caffe
Form Factor	M.2-2280 (M-Key) compliant; (22mm x 80mm x 10mm)
Weight	3.018g (with heat sink)[1]
Interface	PCIe Gen3 x4
Integrated Memory	1 GByte local memory stores all user models
Performance (batch=1)[2]	Resnet50-v1: 100 inferences/sec. Mobilenet-v1: 554 inferences/sec
Latency[2]	Resnet50-v1: 10 msec. Mobilenet-v1: 1.8 msec
Module Power Consumption (Typical)	3.9W @ 600MHz
Operating System Support	CentOS 8, Ubuntu 20.04
Temperature	0oC to +70oC (Commercial) -40oC to +85oC (Industrial)
Part Numbers	DVI-P311M4-CM2TB (600MHz core frequency) DVI-P312M4-CM2TB (800MHz core frequency) DVI-P311M4-IM2TB (600MHz core frequency) DVI-P312M4-IM2TB (800MHz core frequency)
Warranty	Ninety (90) days[3]

[1] Kinara M.2 modules can be ordered without heat sink. Contact Kinara sales.

[2] Maximum performance based on peak computational throughput of Ara-1 (800MHz) and Host System. Specification subject to change without notice. Performance may vary depending on system configuration.

[3] Discuss details with Kinara on production orders.

KINARA | LEADING EDGE AI

Kinara delivers unrivaled edge AI solutions to accelerate and optimize real-time decision making. Our AI accelerators power smart edge devices and gateways that demand responsive AI computing at high energy efficiency. The Kinara AI team, based in Silicon Valley as well as Hyderabad, India, includes Silicon Valley innovators, technology experts from Stanford University, and a world-class hardware and software development group. The company derives its name from the Hindi word for 'edge' and reflects the commitment we make to our customers to build extremely innovative edge devices for retail, smart cities, industry 4.0, and automotive.

www.kinara.ai

KINARA | LEADING EDGE AI

www.kinara.ai