

AI Hardware

Namratha Peddisetty

What is AI?

AI is an acronym for Artificial intelligence. As the name suggests, it is teaching intelligence to artificial (Non-living) things to perform actions, which humans usually perform using their cognitive ability i.e Process, Reason, Think, Learn, Attention, remember etc.

As soon as we hear the term AI, we associate it mostly with software or technologies like Chat GPT. AI in hardware is still not a highly explored topic because from an end user perspective, most of the user look at the capabilities of the product, but not exactly into what is powering into it. For example, I personally never bothered about every component of my laptop as far as it was fast. But, as I started exploring more in AI field, I realized that AI is also highly integrated into hardware. Infact, this has opened a lot of possibilities in hardware space like less power consumption and higher speed and is also the backbone to run lot of tasks for AI like training, Fine tuning etc.for AI to become reality in software world. In this article let's explore, what is AI in hardware and its importance.

What is AI Hardware?

In simple terms, specialized computational components like CPUs, GPUs, TPUs, ASICs to tackle complex AI algorithms, is AI Hardware.

The convergence of AI and hardware represents a shift in computing, enabling devices to perform complex tasks with unprecedented efficiency and intelligence. Unlike conventional hardware, AI-enabled hardware systems possess the ability to learn, adapt, and optimize their performance based on data and feedback. This integration of AI algorithms into hardware unlocks capabilities such as real-time data processing, autonomous decision-making, and advanced sensory perception, paving the way for innovative solutions across diverse domains.

Break throughs in AI Hardware

Some examples of key players researching and creating break throughs in AI hardware are below:

NVIDIA, a leading player in the field of AI hardware, has been at the forefront of developing GPUs (Graphics Processing Units) optimized for AI workloads.

Qualcomm has developed AI-enabled processors tailored for edge computing applications. These processors incorporate AI accelerators and specialized hardware components optimized for running AI workloads efficiently on edge devices such as smart phones, IoT devices, and autonomous systems.

Google's Tensor Processing Units (TPUs) are custom-designed ASICs (Application-Specific Integrated Circuits) optimized for running Tensor Flow, Google's open-source machine learning framework. TPUs excel at accelerating inference tasks, making them well-suited for applications like voice recognition, language translation, and image classification.

Similarly, Apple's Neural Engine is an AI accelerator integrated into its A-series and M-series

chips, powering features like Face ID, Animoji, and Siri on iPhones and iPads. By embedding AI capabilities directly into their silicon designs, companies can achieve greater efficiency, performance, and power optimization for AI-driven applications.

Intel, for instance, has introduced specialized processors like the Intel Xeon Scalable CPUs with built-in AI acceleration capabilities. These CPUs incorporate technologies such as Intel Deep Learning Boost, enhancing performance for AI workloads while maintaining compatibility with existing infrastructure.

Meanwhile, startups like Graphcore are also developing novel AI accelerators, such as the IPU (Intelligence Processing Unit), optimized for both training and inferencing tasks. These accelerators boast massive parallelism and efficiency, enabling organizations to achieve breakthroughs in AI research and deployment.

There are other players as well, from startups to large companies, driving the development of AI-centric hardware solutions, unlocking new possibilities and pushing the boundaries of computational capabilities.

Importance of AI hardware

As such AI itself does not require specialized hardware, it can be run even on our regular computing devices, but if we need faster and multi-processing with efficient results, choosing the right AI hardware based on the use case and data will become very important.

Especially in the current era where AI is the buzz word and everyone is trying to deploy AI for some or the other function, it brings up a need for processing large data sets and this need will only keep growing with more advancement in AI.

Many companies like Dell have already emerged at the forefront and have quickly pivoted to use AI hardware in almost every product of its portfolio, for better battery and faster performance.

Conclusion

As AI continues to permeate various industries, the demand for AI-optimized hardware will only grow, fueling further innovation and advancements. By leveraging AI hardware solutions, organizations can unlock new opportunities, accelerate innovation, and stay ahead in an increasingly competitive landscape. The convergence of AI and hardware holds immense potential to reshape industries, improve productivity, and drive transformative change in the years to come.

In conclusion, with the constant ongoing research and innovation, I strongly believe AI hardware will undergo significant advancements in the coming years. The future of AI hardware holds promise for unlocking new capabilities, enabling transformative applications, and shaping the next generation of intelligent systems.

References:

<https://www.nvidia.com/en-sg/deep-learning-ai/products/solutions/>

<https://www.dell.com/en-us/blog/ai-at-dell-technologies/>

<https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>

<https://www.qualcomm.com/products/technology/artificial-intelligence>

<https://www.mckinsey.com/~media/McKinsey/Industries/Semiconductors/Our%20Insights/Artificial%20intelligence%20hardware%20New%20opportunities%20for%20semiconductor%20companies/Artificial-intelligence-hardware.pdf>

<https://www.graphcore.ai/about>

<https://venturebeat.com/ai/what-is-ai-hardware-how-gpus-and-tpus-give-artificial-intelligence-algorithms-a-boost/#h-what-are-some-examples-of-ai-hardware>

<https://cloud.google.com/tpu>

<https://machinelearning.apple.com/research/neural-engine-transformers>

<https://www.intel.com/content/www/us/en/newsroom/news/5th-gen-xeon-data-center-news.html#gs.7wb1kj>