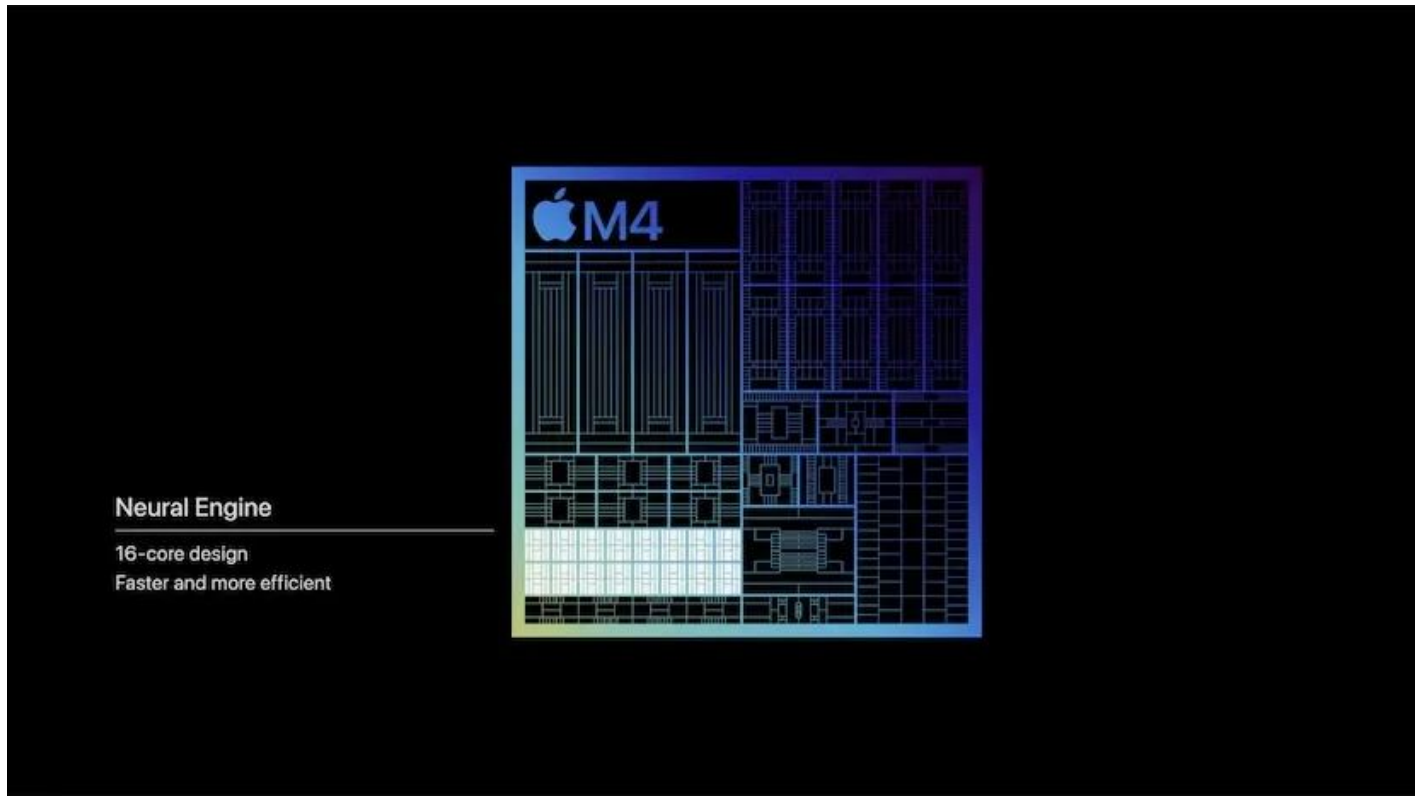


# Neural Processing Unit

May 9, 2024 • vajiramandravi.com



## About Neural Processing Unit (NPU):

- It is a **dedicated processor designed** specifically for accelerating **neural network processes**.
- A neural network is essentially a type of **machine learning algorithm** that mimics the human brain for processing data.
- It is highly capable for handling machine learning operations that form the basis for AI-related tasks, such as speech recognition, natural language processing, photo or video editing processes like object detection, and more.
- In most consumer-facing gadgets such as smartphones, laptops and tablets, the NPU is integrated within the main processor, adopting a **System-on-Chip (SoC)** configuration.

## How is NPU different from the Central Processing Unit (CPU)?

- CPUs employ a **sequential computing method**, issuing one instruction at a time, with subsequent instructions awaiting the completion of their predecessors.
- In contrast, NPU **harnesses parallel computing** to simultaneously execute numerous calculations. This parallel computing approach results in swifter and more efficient processing.
- NPU is also able to learn based on accumulated data in order to deduce optimal solutions.

---

## Q1: What is Deep learning?

Deep learning is a subset of machine learning, has revolutionized the way machines process and interpret data. It teaches computers to do what comes naturally to humans; for example, in the case of Self-Driving cars, recognizing the signals to stop or go.

**Source:** NPU: What is it and why is it gaining importance in semiconductor industry

---

Source: <https://vajiramandravi.com/current-affairs/neural-processing-unit/>

© Vajiram & Ravi. For personal use only.