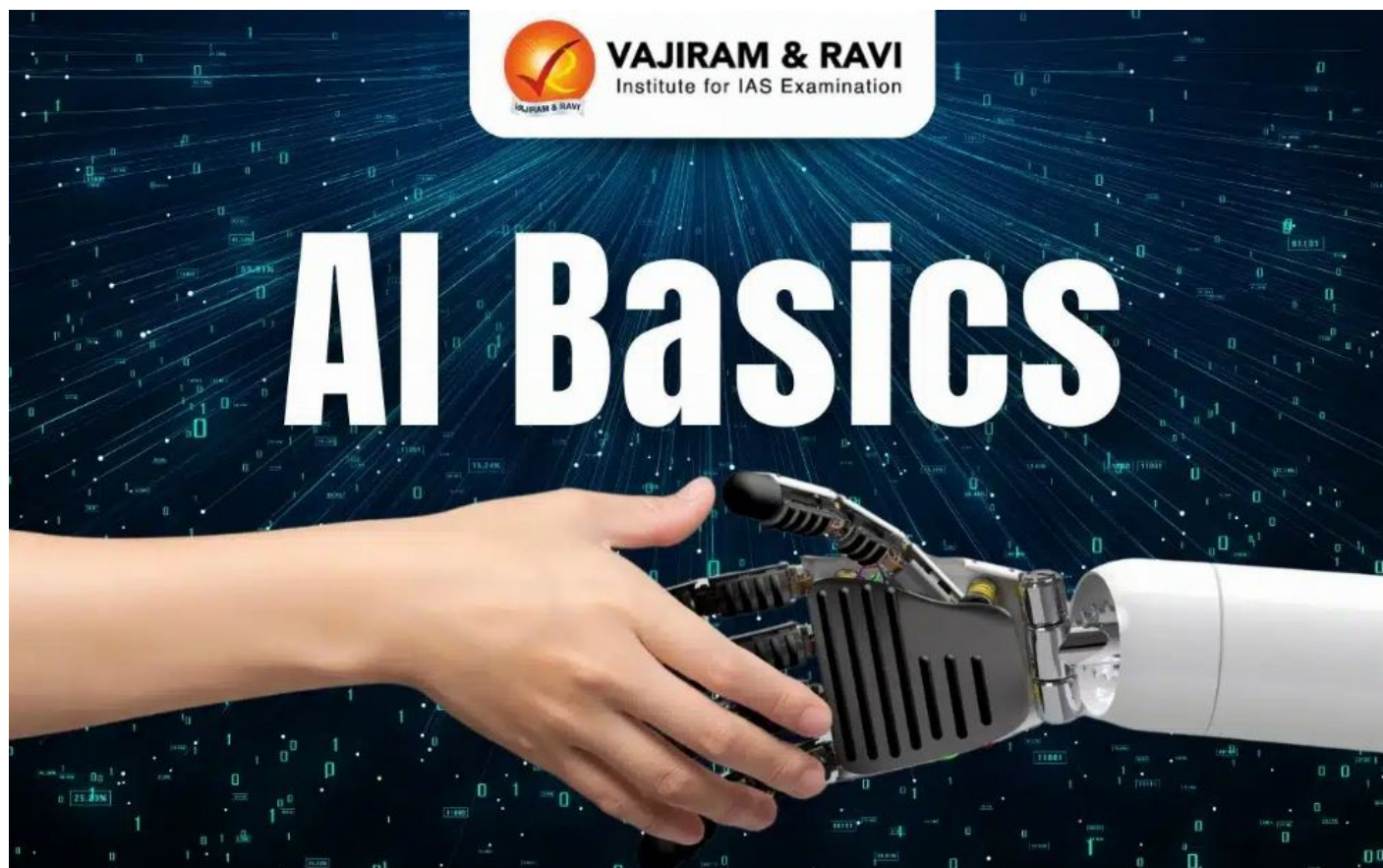


AI Basics: CPU, GPU and TPU

April 16, 2025 • vajiramandravi.com



AI Basics: CPU, GPU and TPU Latest News

How TPU is Different from CPU and GPU

About Processing Units

- **Processing units** are the **core hardware components** that act as the **brain of a computer**. They perform tasks such as **calculations, processing images, and transmitting data**, similar to how the **human brain executes complex tasks**.

What is a CPU (Central Processing Unit)?

- The **CPU** is a **general-purpose processor** that was developed in the **1950s** and can handle **a wide variety of tasks**.
- It functions like a **conductor in an orchestra**, coordinating the operations of all other computer parts like **GPUs, disk drives, and memory units**.

- A CPU contains **cores** — individual units that execute instructions. Early CPUs had **only one core**, but modern CPUs may contain **2 to 16 cores**.
- **Each core can handle one task at a time**, so a CPU's **multitasking capacity** depends on the **number of cores**.
- For everyday users, **2 to 8 cores** are usually sufficient, and CPUs are so efficient that users rarely notice that tasks are completed **sequentially, not simultaneously**.

What is a GPU (Graphics Processing Unit)?

- A **GPU** is a **specialised processor** designed to **perform many tasks simultaneously**, using a technique called **parallel processing**.
- Unlike CPUs, which process tasks **sequentially**, GPUs break down complex tasks into **thousands or millions of smaller problems**, solving them in **parallel**.
- Modern GPUs contain **thousands of cores**, making them far more suitable for **intensive computational tasks**.
- Initially developed for **rendering graphics** in gaming and animation, GPUs are now widely used in **machine learning and artificial intelligence**.
- GPUs have evolved into **general-purpose parallel processors**, making them a **key tool** in running **AI models** and handling large data operations.
- However, GPUs have **not replaced CPUs**, because certain operations are **better handled sequentially**, which is the **strength of CPUs**.

What is a TPU (Tensor Processing Unit)?

- A **TPU** is also a **type of ASIC (Application-Specific Integrated Circuit)**, meaning it is **built for a specific function** — in this case, **AI tasks**.
- First introduced by **Google in 2015**, TPUs are **specially designed hardware units** built **from the ground up** to handle **machine learning operations**.
- TPUs focus on processing **tensors** — the **multidimensional data arrays** used in AI model computations.
- They are optimised to **run neural networks efficiently**, enabling **faster training** and **execution of AI models** than GPUs or CPUs.
- For example, **training an AI model** that may take **weeks on a GPU** can often be completed in **hours using a TPU**.
- TPUs are used at the core of **Google's major AI services**, such as **Search, YouTube**, and **DeepMind's large language models**, illustrating their real-world application in high-scale AI infrastructure.

Source: [IE](#)

Source: <https://vajiramandravi.com/current-affairs/ai-basics-cpu-gpu-and-tpu/>

© Vajiram & Ravi. For personal use only.