

23– 06 – 2024

News: GPT – 4o

- Recently, OpenAI introduced its latest large language model (LLM) called GPT-4o, billing it as their fastest and most powerful AI model so far.

GPT – 4o

- GPT – 4o ("o" stands for "Omni" here) is a revolutionary AI model developed by OpenAI to enhance human-computer interactions.
- It allows users to input any combination of text, audio, and image and receive responses in the same formats, making it a multimodal AI model.
- **Technology Applied:** LLMs are the backbone of GPT-4o. Large amounts of data are fed into these models to make them capable of learning things themselves.
- GPT-4o differs from its predecessors by using a single model to handle text, vision, and audio tasks, eliminating the need for multiple models.
- For example, previous models required separate models for transcription, intelligence, and text-to-speech in voice mode, but GPT-4o integrates all of these capabilities into a single model.

- It can process and understand inputs more holistically, including tone, background noises, and emotional context in audio inputs.
- GPT-4o excels in areas like speed and efficiency, responding to queries as fast as a human does in conversation, in around 232 to 320 milliseconds.

Key Features and Abilities

- Enhanced audio and vision understanding allow GPT-4o to process tone, background noises, and emotional context, and identify objects.
- GPT-4o demonstrates significant advancements in handling non-English text, catering to a global audience.

Safety Concerns

- Despite its advancements, GPT-4o is still in the early stages of exploring unified multimodal interaction, with ongoing development required.
- OpenAI emphasises built-in safety measures and continuous efforts to address risks like cybersecurity, misinformation, and bias.

Large Language Model (LLM)

- A Large Language Model (LLM) is an AI program capable of recognising and generating text. LLMs are trained on vast datasets using machine learning and

deep learning, particularly transformer models that mimic the human brain's neural structure.

- LLMs typically rely on transformer models, consisting of an encoder and a decoder. LLMs can be categorised based on architecture, training data, size, and availability.
- LLMs are used for generative AI tasks like producing text, assisting programmers with coding, and various applications like sentiment analysis and chatbots.
- They excel at understanding natural language and processing complex data, but can also provide unreliable information or "hallucinate" responses if given poor input data, and pose security risks if misused.