What to Expect in Machine Learning For 2024 and Beyond

Luís Fernando Torres · Follow
Published in Artificial Intelligence in Plain English
5 min read · 16 hours ago

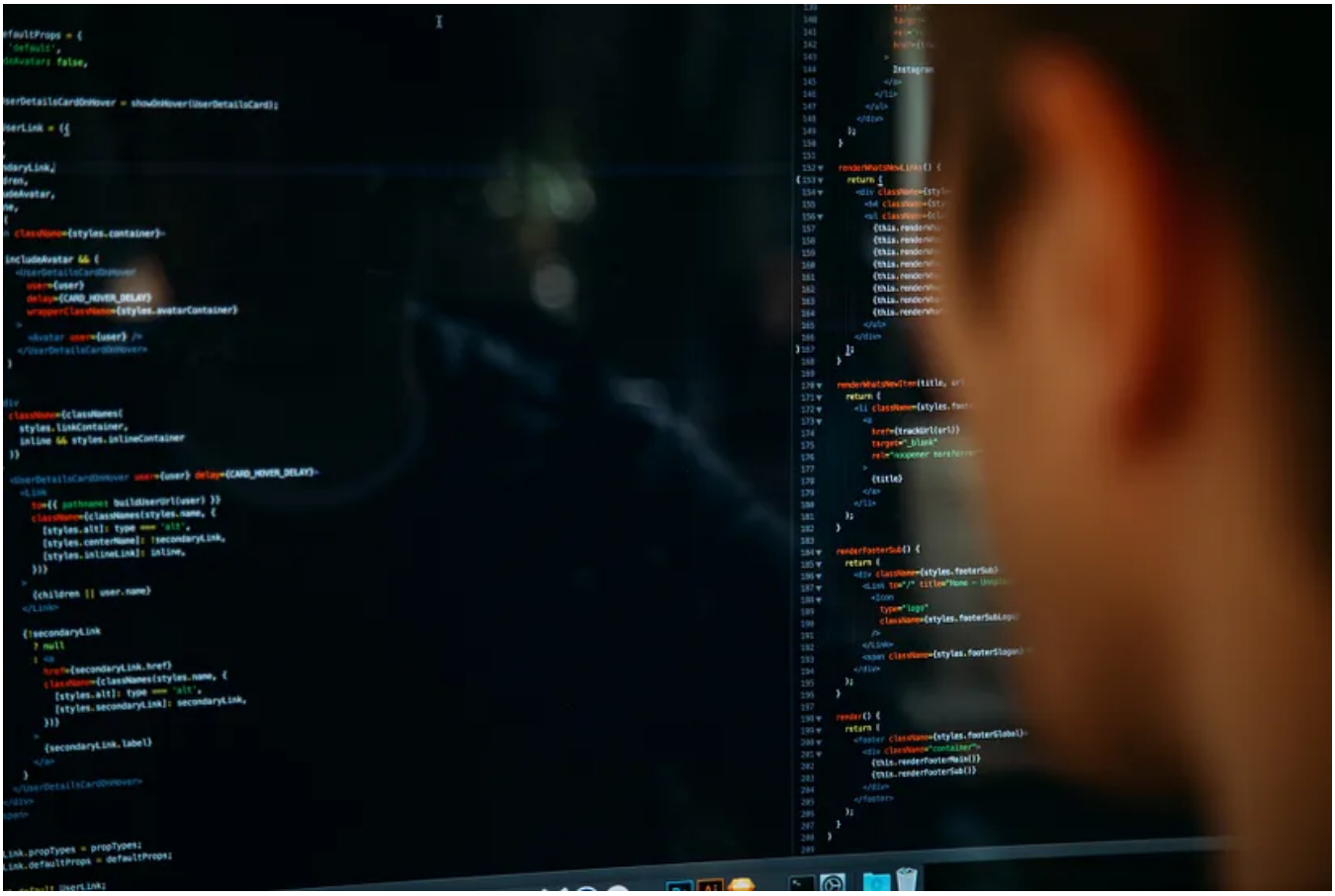


Photo by charlesdeluvio on Unsplash

No one can deny that 2023 was an exciting year for Machine Learning. Thanks to the release of ChatGPT — which offered Large Language Models as a product for the general public — AI has been in the spotlight like never before.

The term *"Artificial Intelligence"* is not new. Frank Rosenblatt presented the first neural network in 1958 — the perceptron — a significant milestone in using a machine to replicate human intelligence.

The Mark I Perceptron was a machine that implemented the perceptron algorithm for image recognition. Although revolutionary, it was clear that the project had several limitations, and with the limited computing power of the times and the failure to achieve high expectations, progress in AI slowed down significantly.

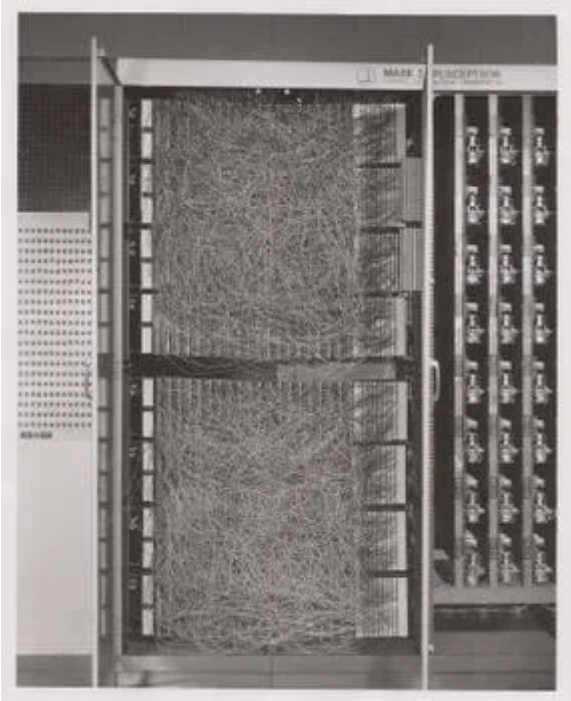


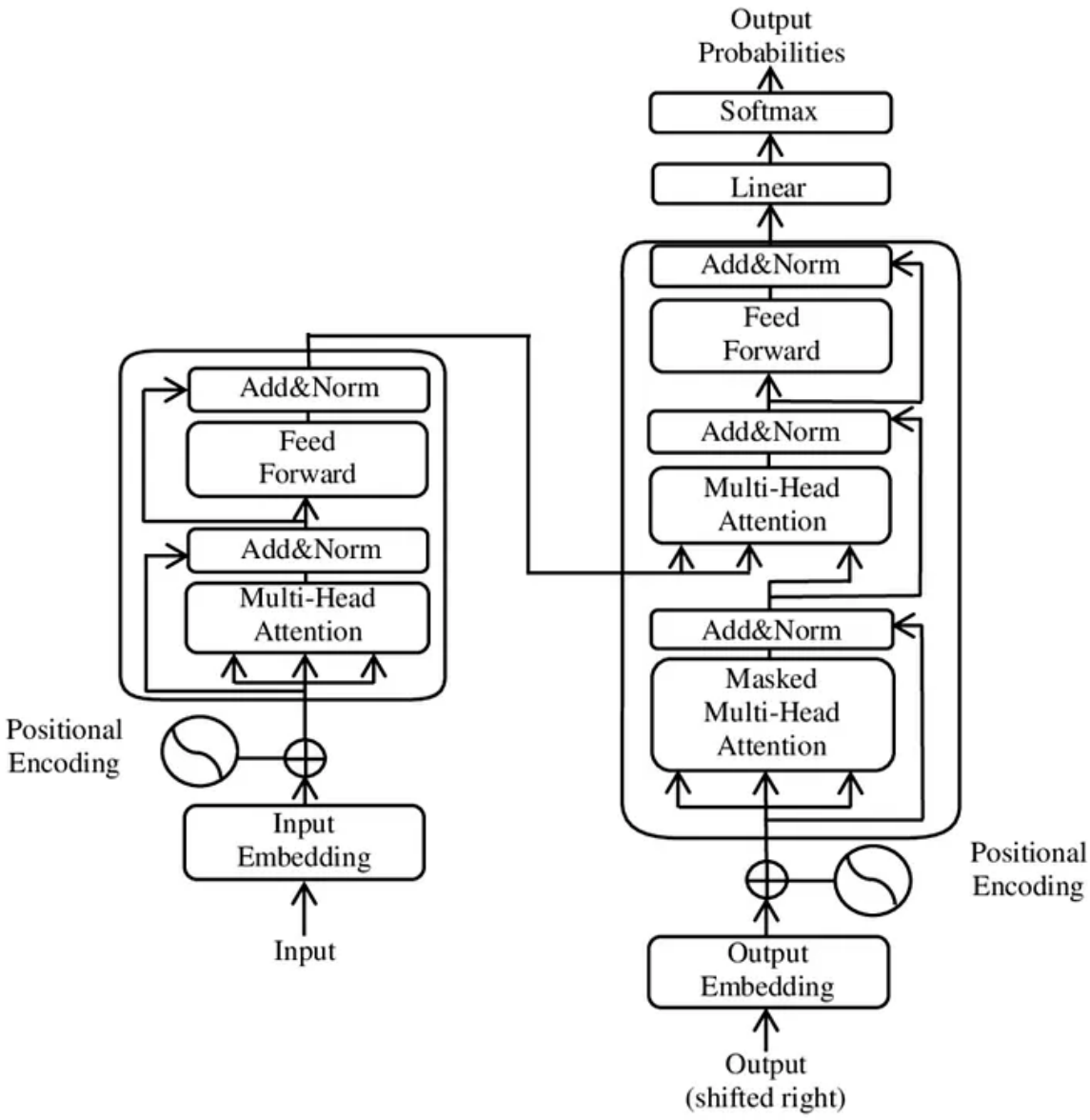
Photo of the Mark I Perceptron by an unknown photographer on Perceptron — Wikipedia

Slowly but surely, we saw the resurgence of interest in *intelligent* systems. The 1990s and 2000s brought several advancements in computing power and machine learning algorithms, along with a vast production of data never seen before in History, paving the way to more sophisticated models.

In 2012, a convolutional neural network called AlexNet achieved surprising results in the ImageNet Large Scale Visual Recognition Challenge, opening new doors for research in deep learning. Graphic processing units — *GPUs* — also allowed us to reach the necessary computing power for training deep and computationally expensive neural networks.

In 2017, the research paper *"Attention Is All You Need"* introduced the Transformer architecture, which proved highly effective in several domains, such as natural language processing, computer vision, and audio processing. The Transformer

architecture allowed for more parallelization, the attention mechanism improved performance, and the architecture served as a foundation for more advanced models like GPT — the model behind ChatGPT — and BERT.



Transformer (machine learning model). (2023, December 22). In Wikipedia. https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)

We now have several “AI” products available to the public. ChatGPT can chat with you via text and audio and comprehend images. MidJourney and Dall•E can create different types of images according to the text description given by the user. Google announced Gemini in an attempt to compete with OpenAI’s ChatGPT.

With all this in mind, we will probably see higher advancements in AI in the following years, such as more powerful models, integration with other technologies, and ethical concerns.

What should we look for in 2024 and the upcoming years?

• **Data, Data, Data**

Data is the backbone of AI. GPT4, the most potent LLM to date, was trained on vast amounts of data scrapped from the Internet.

As we move into 2024, a crucial story to follow is the current lawsuits filed against OpenAI and other AI companies.

Artists and writers say their work was used to train generative models without their permission and compensations (see: Artists take new shot at Stability, Midjourney in updated copyright lawsuit | Reuters and Some authors are suing OpenAI. Will it backfire? : NPR).

As we move forward with AI, more data will be required, and people with a deep understanding of how to extract, treat, analyze, and interpret data will probably be in high demand in the upcoming years.

We will also see higher concerns over copyright infringements and privacy matters.

• **Generative AI**

We probably will not see any slowing down in Generative AI. Suno AI just released a mind-blowing platform for generating music, while Pika promises ground-breaking results in text-to-video generation.

Midjourney also released its 6th version on Discord, with amazing results in image generation.

Expect higher-quality art, deepfakes, and more inovations in videos and music with GenAI by the end of 2024.

Black and White portrait generated with Midjourney 6 by @mberndtgen.

- **Robotics and Mechatronics**

In her recent book, *Essential Math for AI: Next-Level Mathematics for Efficient and Successful AI Systems,* Hala Nelson mentions *robotics* as a less-hyped but very important area within Data Science and Machine Learning. I definitely agree!

Computer Vision is highly compatible with robotics, allowing for the development of autonomous cars, drones, and several others.

Since we are witnessing larger adoption of LLM chatbots, it is also possible to see this technology used to improve real-world interactions between humans and machines.

Tesla Optimus Gen 2.0 humanoid robot handling delicate objects like eggs with ease using human-like tactile sensing. Credit: Courtesy of Tesla, Inc.

• **Augmented Reality**

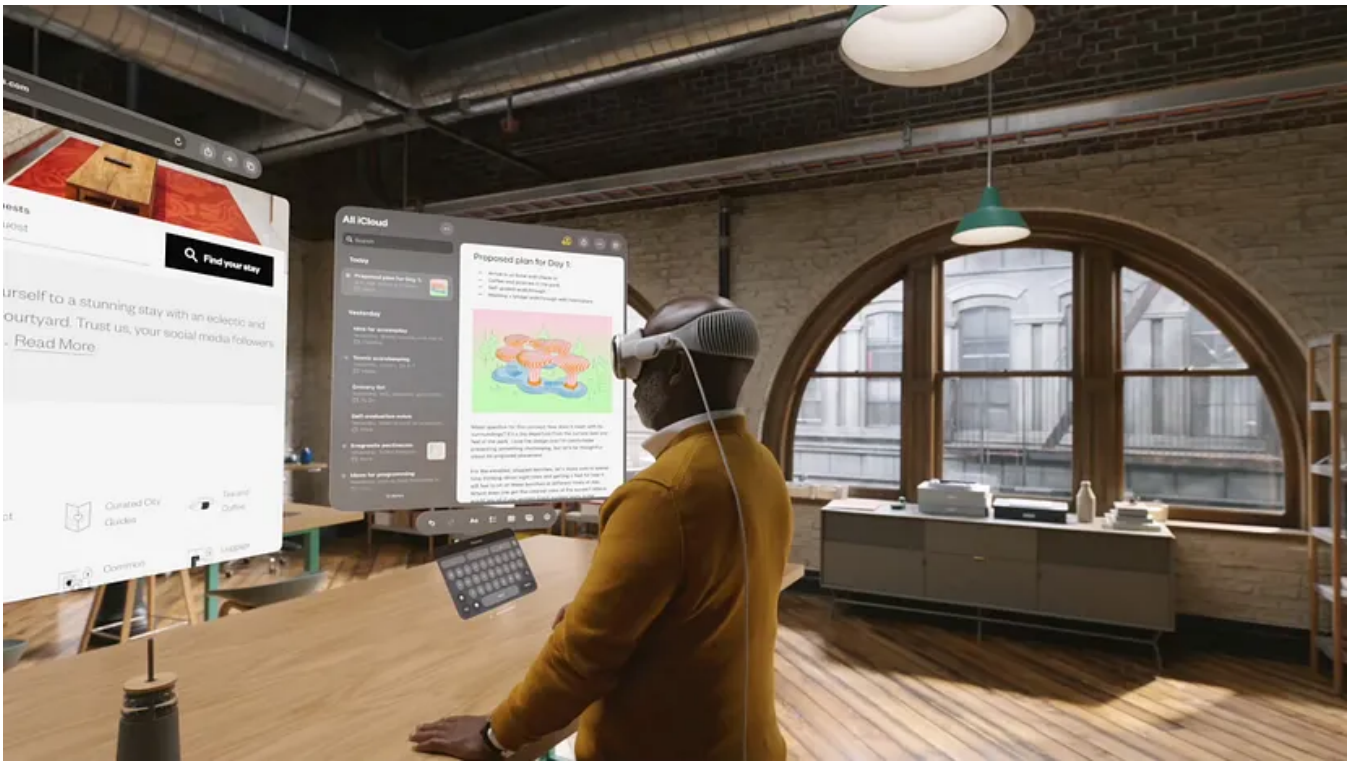
At the WWDC23, Apple announced the Apple Vision Pro, *a revolutionary spatial computer that seamlessly blends digital content with the physical world.*

This product is Apple's attempt at augmented reality, a field where numerous companies have repeatedly tried but struggled to achieve widespread adoption.

Machine learning models can generate 3D objects for apps and other uses in augmented reality, improve the interactions between people through facial

recognition in a virtual space, and enhance interaction between human users and LLM-powered chatbots.

Besides Apple Vision Pro, we may also see new attempts at wearable devices, similarly to Humane's AI Pin.



Introducing Apple Vision Pro: Apple's first spatial computer

• Quantum Computing

Giants like IBM are actively behind research in quantum computing. This technology harnesses the laws of quantum mechanics to solve problems too complex for classical computers.

As AI and its different use cases advance, models are getting bigger and need more computing power to break their current limits. Quantum computing can help provide more processing power to train more powerful models in a shorter period of time and at smaller costs.

However, it is relevant to say that the idea that larger models mean higher quality has been challenged in the past months, and there are several claims that smaller models may be a better option.

IBM Q System One, a quantum computer 20 superconducting qubits. Photo by IBM Research.

Without a doubt, 2023 was a pivotal year. As we move into 2024, we might expect even further enhancements in the currently existing tools and the emergence of new, groundbreaking technologies.

In this article, I highlighted the key areas I have been monitoring closely. Now, I am curious about your perspective. Which technologies are **you** aiming to master in the next years to come? Share your insights in the comments.

Thank you.

Stay curious!

Luis Fernando Torres

Let's connect! 🔗

LinkedIn • Kaggle • HuggingFace • X/Twitter

Like my content? Feel free to Buy Me a Coffee ☕ !

PlainEnglish.io 🚀

Thank you for being a part of the In Plain English community! Before you go:

- *Be sure to **clap** and **follow** the writer*
- *Learn how you can also **write for In Plain English***
- *Follow us: **X** | **LinkedIn** | **YouTube** | **Discord** | **Newsletter***
- *Visit our other platforms: **Stackademic** | **CoFeed** | **Venture***

Data Science | Machine Learning | AI | Artificial Intelligence | Technology

Follow

Written by Luís Fernando Torres

1.5K Followers · Writer for Artificial Intelligence in Plain English

Data Scientist | Machine Learning Engineer | Commodities Trader & Investor | https://luuisotorres.github.io/

More from Luís Fernando Torres and Artificial Intelligence in Plain English