

Shown on the vertical axis is the training computation that was used to train the AI systems.

10 billion petaFLOP

Computation is measured in floating point operations (FLOP). One FLOP is equivalent to one addition, subtraction, multiplication, or division of two decimal numbers.

100 million petaFLOP

The data is shown on a logarithmic scale, so that from each grid-line to the next it shows a 100-fold increase in training computation.

1 million petaFLOP

10,000 petaFLOP

100 petaFLOP

1 petaFLOP = 1 quadrillion FLOP

10 trillion FLOP

100 billion FLOP

1 billion FLOP

10 million FLOP

100,000 FLOP

1,000 FLOP

10 FLOP

$$\frac{a^2 + b^2}{2} > \left(\frac{a + b}{2}\right)^2$$

Minerva: built in 2020  
Minerva can solve complex mathematical problems.

PaLM: built in 2022  
PaLM can generate high-quality text, explain some jokes, cause & effect, and more.

GPT-3: 2020; 314 million petaFLOP  
GPT-3 can produce high-quality text that is often indistinguishable from human writing.

DALL-E: 2021; 47 million petaFLOP  
DALL-E can generate high-quality images from written descriptions.

NEO: 2021; 1.1 million petaFLOP  
Recommendation systems like Facebook's NEO determine what you see on your social media feed, online shopping, streaming services, and more.

AlphaGo: 2016; 1.9 million petaFLOP  
AlphaGo defeated 18-time champion Lee Sedol at the ancient and highly complex board game Go. The best Go players are no longer human.

AlphaFold: 2020; 100,000 petaFLOP  
AlphaFold was a major advance toward solving the protein-folding problem in biology.

MuZero: 2019; 48,000 petaFLOP  
MuZero is a single system that achieved superhuman performance at Go, chess, and shogi (Japanese chess) – all without ever being told the rules.

AlexNet: 2012; 470 petaFLOP  
A pivotal early "deep learning" system, or neural network with many layers, that could recognize images of objects such as dogs and cars at near-human level.

NLP

Decision tree

TD-Gammon: 1992; 18 trillion FLOP  
TD-Gammon learned to play backgammon at a high level, just below the top human players of the time.

LSTM

LeNet-5

NetTalk: 1987; 81 billion FLOP  
NetTalk was able to learn to pronounce some English text by being given text as input and matching it to phonetic transcriptions. Among its many limitations, it did not perform the visual recognition of the text itself.

ALVINN  
Zip CNN

RNN for speech

Pandemonium (Morse)  
Samuel Neural Checkers

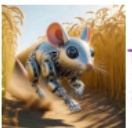
System 11

Back-propagation  
Neocognitron: 1980; 228 million FLOP  
A precursor of modern vision systems. It could recognize handwritten Japanese characters and a few other patterns.

Fuzzy NN

Perceptron Mark I: built in 1957/58; 695,000 FLOP  
Regarded as the first artificial neural network, it could visually distinguish cards marked on the left side from those marked on the right, but it could not learn to recognize many other types of patterns.

ADALINE: built in 1960 and trained on around 9,900 FLOP  
An early single-layer artificial neural network.



Theseus: built in 1950 and trained on around 40 floating point operations (FLOP)  
Theseus was a small robotic mouse, developed by Claude Shannon, that could navigate a simple maze and remember its course.

The first electronic computers were developed in the 1940s

Pre Deep Learning Era  
Training computation grew in line with Moore's law, doubling roughly every 20 months.

Deep Learning Era  
Increases in training computation accelerated, doubling roughly every 6 months.

1940 1950 1960 1970 1980 1990 2000 2010 2020

1956: The Dartmouth workshop on AI, often seen as the beginning of the field of AI research

1997: Deep Blue beats world chess champion Garry Kasparov



ChatGPT

