

History of Deep Learning

Chao-En Yu

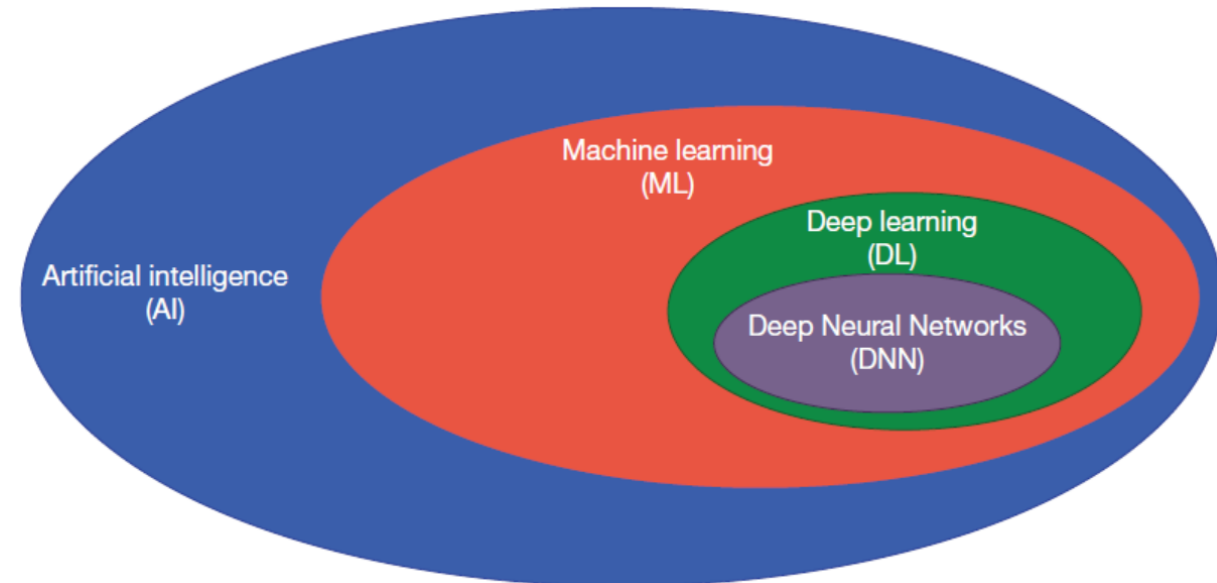
National Tsing Hua University

Deep Learning

- Deep learning (DL) is a quickly evolving field, which has demonstrated amazing results in performing tasks that traditionally have been performed well only by humans.
- DL is a class of machine learning algorithms that use multiple layers of computational units where each layer learns its own representation of the input data.
- These representations are combined by later layers in a hierarchical fashion.

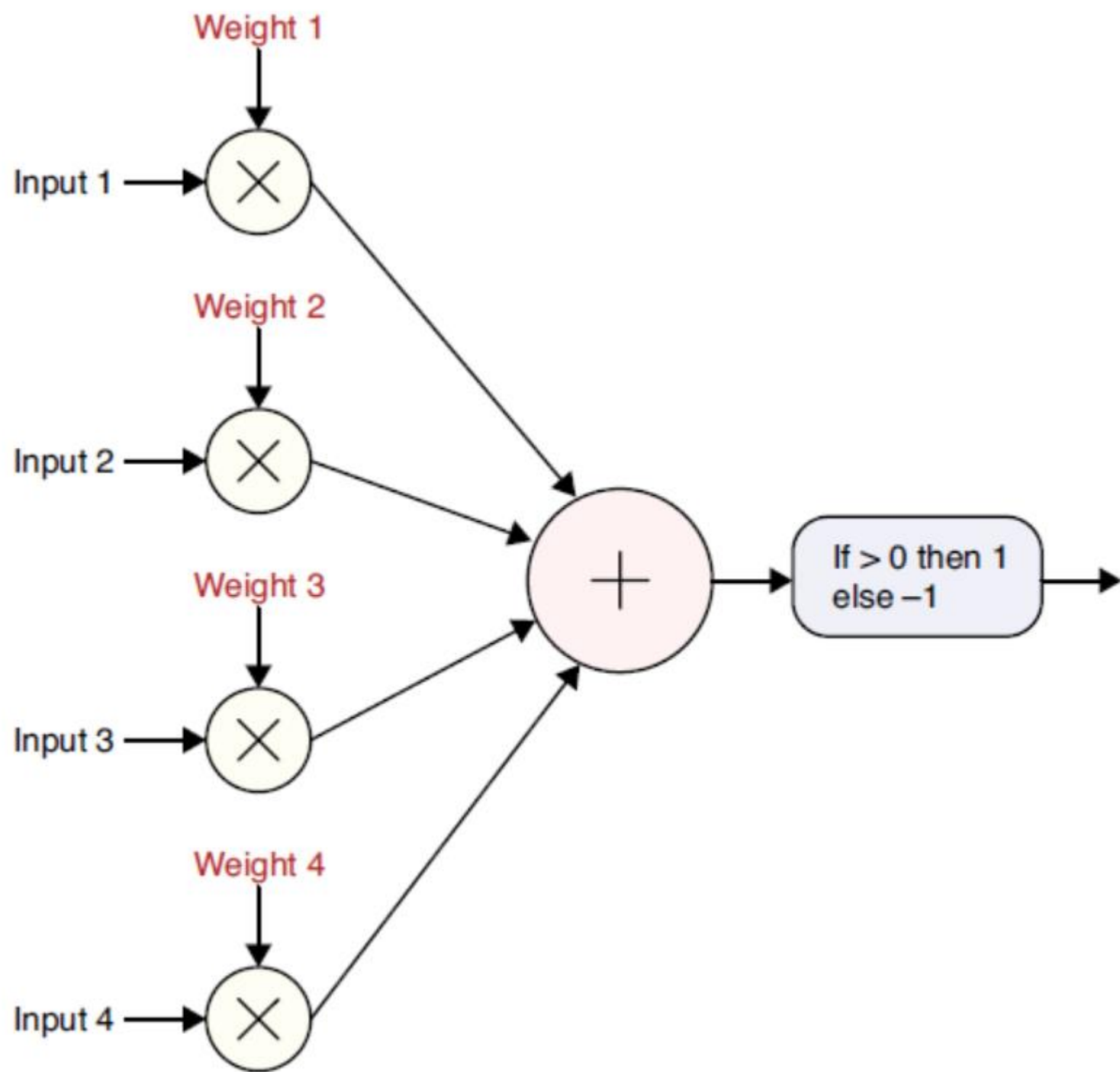
Deep Neural Network

- A fundamental part of DL is the deep neural network (or ANN) which is inspired by the biological neuron.
- The algorithms presented are more tied to machine capabilities than to how an actual human brain works.
- DL is a subset of machine learning (ML), which is a subset of AI.



Dartmouth Conference

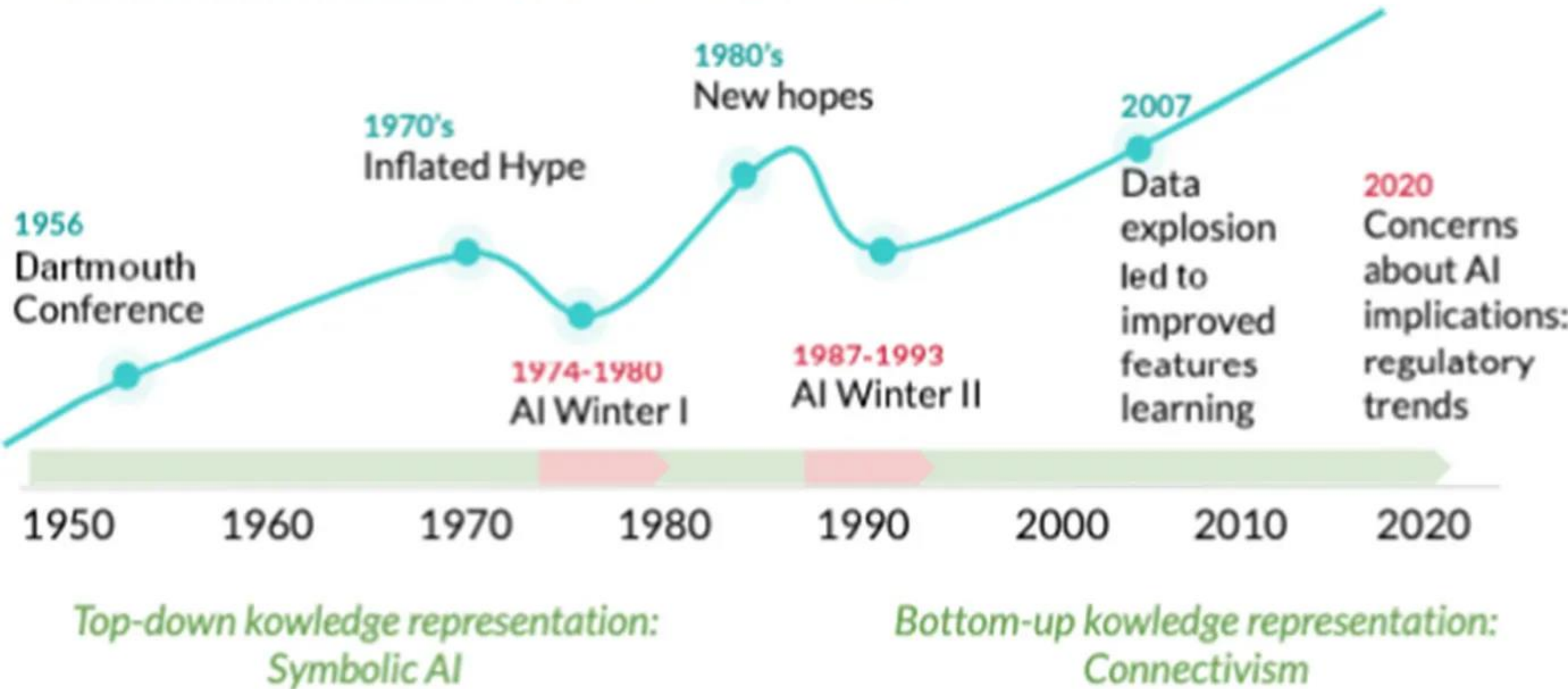
- The first model of an artificial neuron was introduced in 1943 (McCulloch and Pitts, 1943).
- Artificial intelligence was introduced as a name for the field of thinking machines in 1956 in a workshop at Dartmouth College. John McCarthy: “Since we’ve managed to model a neuron, why not create a network of neurons/a brain/artificial intelligence?”
- In 1957, the perceptron was proposed as a simplified mathematical model of a neuron (Rosenblatt, 1958).



The First AI Winter

- The perceptron has some fundamental limitations: The original learning algorithm did not extend to multilayer network.
- This resulted in neural network research falling out of fashion, which is referred as the *first AI winter* (1974-1980 by Minsky and Papert, 1969).
- Although there did exist a learning algorithm for multilayer networks (Ivakhnenko and Lapa, 1965).

— investment & research = $f(\text{expectations, results})$



Expert System Boom

- By the early 1980s, programmers handcrafted a large set of explicit rules for manipulating knowledge stored in explicit databases, involving *no* learning.
- Rule-based expert systems like MYCIN (medical diagnosis) and XCON (computer configuration) delivered real business value.
- Around 1985, companies were spending over \$1 billion each year on the technology.

The Second AI Winter

- By the early 1990s, these systems had proven expensive to maintain, difficult to scale, and limited in scope, and interest died down.
- Meanwhile, simpler approaches often worked better for many problems.
- Major projects failed to deliver on their promises, and the expert systems market collapsed in 1987, which began the second AI winter (1987-1993).

Neural Network Revived

- The second wave of neural network (NN) research was heavily influenced by the backpropagation algorithm for automatic training of multilayer networks (Werbos, 1982; Rumelhart et al., 1986).
- An important outcome was the development of LeNet in 1989, which was shown to be able to recognize handwritten zip codes (LeCun et al., 1990).
- LeNet was built on Fukushima's Neocognitron (1980) which is the first convolutional neural network (CNN).

Out of Fashion Again

- An enhanced version of LeNet was later used by major US banks to read handwritten checks.
- It thereby became one of the first big commercial applications of neural networks.
- Despite the progress, the limited computational capability at the time prevented the networks from scaling to larger problems, and other traditional ML approaches were perceived as better alternatives.

Third Wave of Neural Network

- The third wave of neural network research was enabled by a combination of algorithmic progress, availability of massive data sets, and the ability to use graphics processing units (GPU).
- The field started using the term deep networks (DL) in 2006 (University of Toronto) and was popularized due to AlexNet (Krizhevsky et al., 2012).  Though it is not the first one to use GPUs.
- AlexNet scored significantly high in a computer vision competition known as the ImageNet challenge.

AlexNet, AlphaGo, ChatGPT

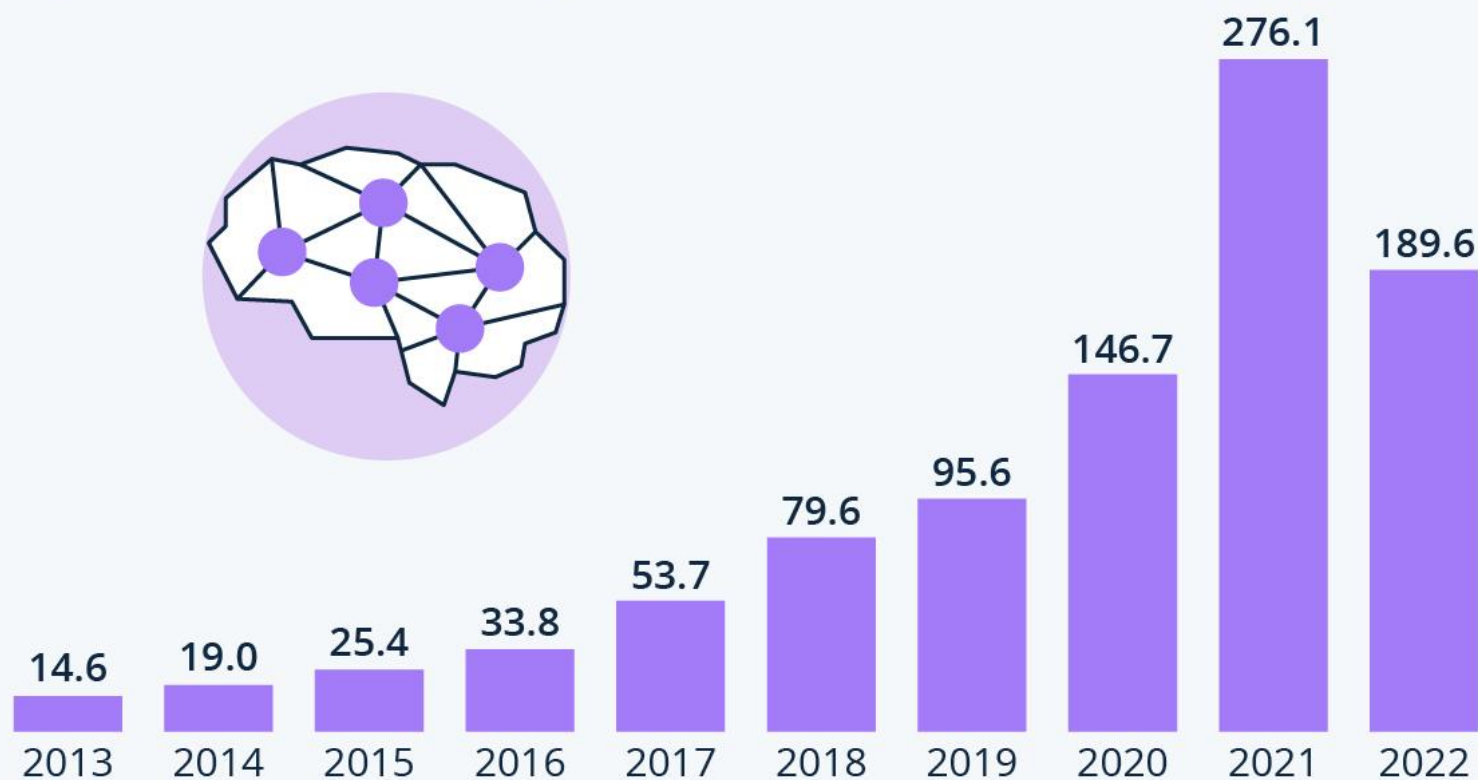
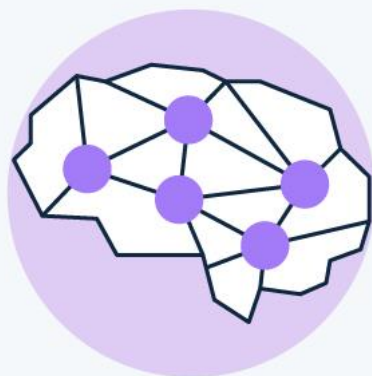
- AlexNet was a CNN consisting of 650,000 neurons with 60 million parameters that were trained using two early Nvidia GPUs.
- AlphaGo (2016) used convolutional neural networks with 12 layers, trained by reinforcement learning.
- Deep learning is now the rule in AI, supported by the Attention mechanism and Transformers that underpin the large language models, like ChatGPT (2022), Claude (2023), Gemini (2023) and others.

Hype vs. Hope

- Currently, AI investment surpasses \$200 billion annually while revenue is close to \$30 billion.
- AI is currently being judged by executives and investors not by what it has accomplished. What will happen precisely as the AI bubble deflates is still up in the air.
- Don't believe the short-run hype, but prepare for the long-term impact of AI.

How Much Are Companies Investing in AI?

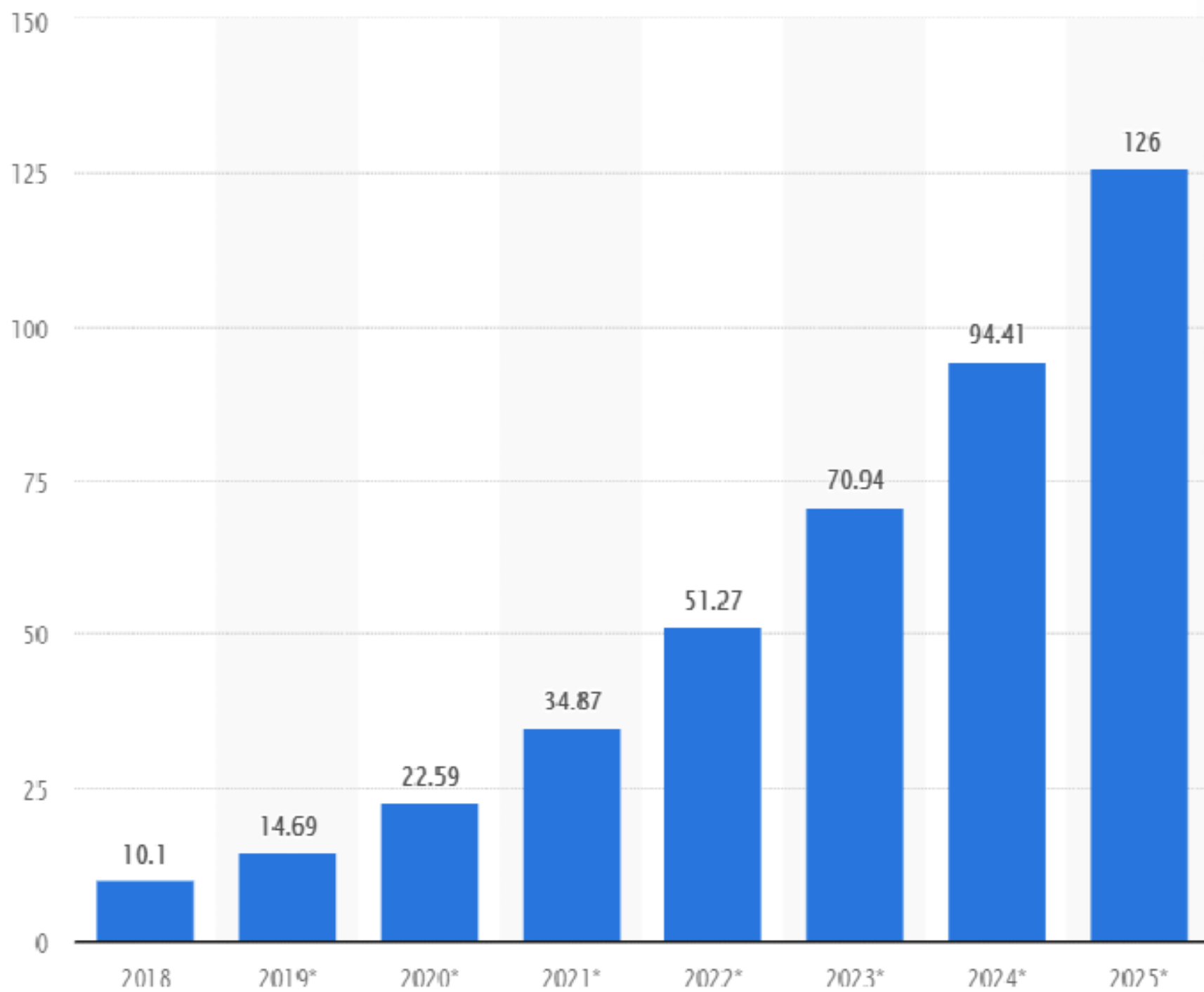
Total global corporate investment in artificial intelligence*
(in billion U.S. dollars)



* Mergers/acquisitions, minority stake investments,
private investments, public offerings

Source: Stanford University

Market revenue in billion U.S. dollars



Required Background

- Anybody who wants to get serious about DL needs to also master traditional ML.
- DL is an applied field, so coding with examples is necessary, and it is allowed to use AI tools to generate or complete source code.
- Many topics in DL are inherently mathematical, so it is necessary to include some of the mathematics to provide a good description of how things work.

Prerequisites

- Statistics and probability theory.
- Linear algebra: dot products, matrix-vector multiplications, and matrix-matrix multiplications.
- Calculus: partial derivatives, chain rule, and optimization.
- Python programming: integer vs. floating point datatypes, if-statements, for-loops, list and Numpy arrays, methods and functions, matplotlib, and pandas.

Economic Data Analysis and Predictions

Chao-En Yu

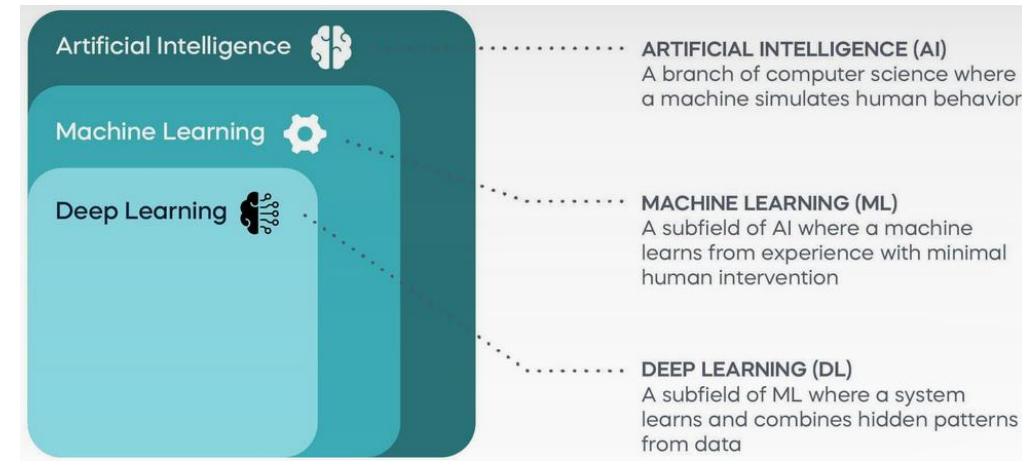
National Tsing Hua University

Why do We Need Predictions?

- Rationality: Using all available information to achieve your goal. Does rationality require people to make predictions? Guess 2/3 (Keynesian Beauty Contest)
- Rational expectations (Muth 1961): Expectations are essentially the same as the *predictions* of the relevant economic theory.
- John McCarthy (1997): “Intelligence is the *computational part of the ability* to achieve goals in the world.”

What is Machine Learning?

- Machine learning (ML) is a subfield of statistics (termed statistical learning) and artificial intelligence, also related to behavioral economics.
- Machine learning (ML) uses algorithms trained on data sets to *create or adjust models* capable of performing tasks that would otherwise only be possible for humans.



Why Machine Learning?

- Prediction vs explanation (Pre vs. Post)
- Model selection
- Bias-variance trade-off
- Bayesian: To blend data analysis with background professional knowledge
- Human heuristic vs. AI



Sebastian Raschka 

@rasbt



Maybe a hot take, but what about the following advice to the next gen:

Don't get an AI degree; the curriculum will be outdated before you graduate. Instead, study math, stats, or physics as your foundation, and stay current with AI through code-focused books, blogs, and papers.

Why Coding?

- Exact vs. approximate solutions
- Verification of theoretical models
- Making predictions
- To have practical works
- To have an advantage on your resume

What is Python?

- The implementation of programming language Python was started in December 1989 by Guido van Rossum at CWI in the Netherlands.
- It was developed at CWI as a successor to the ABC language.
- Its name was inspired by the BBC comedy series Monty Python's Flying Circus.

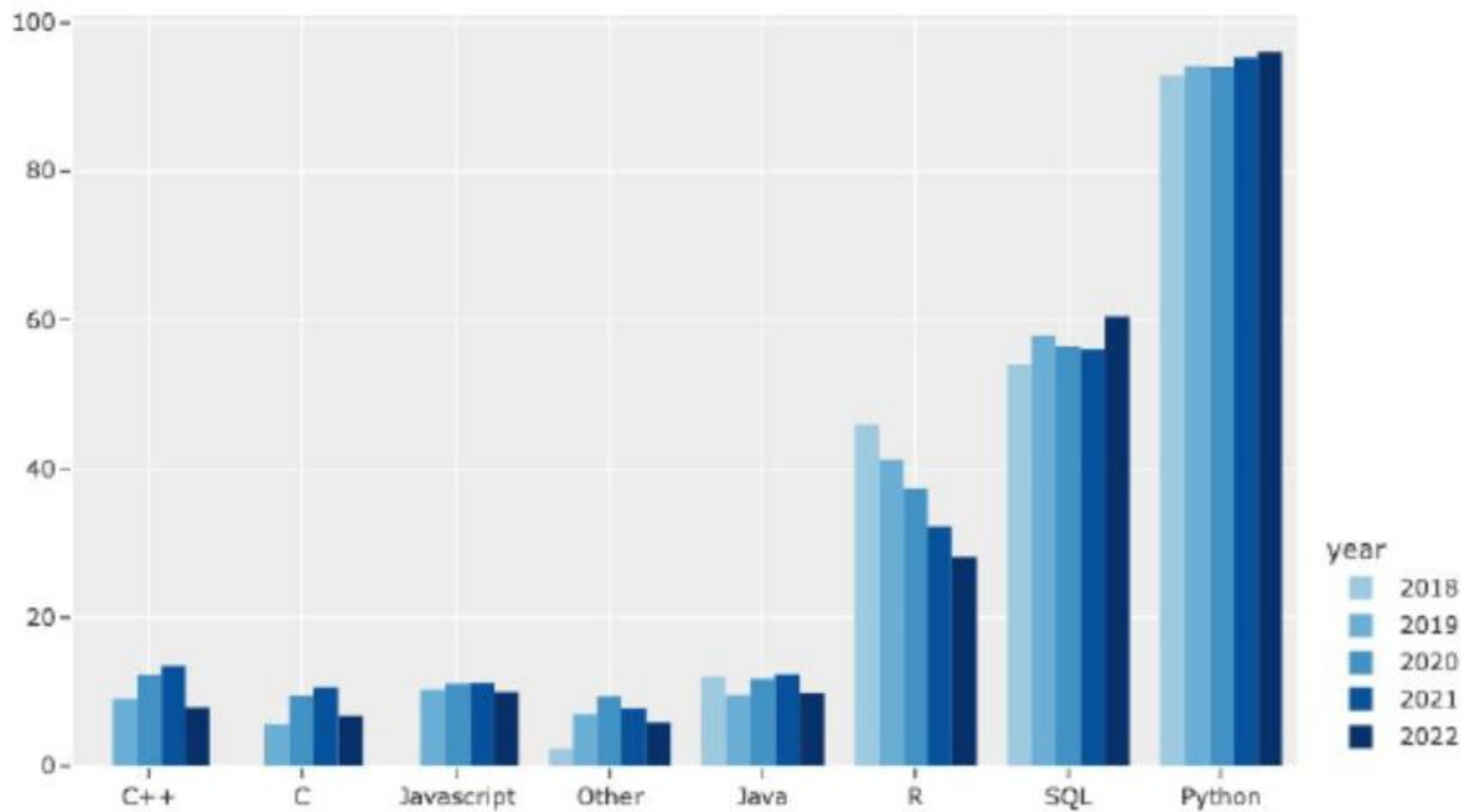


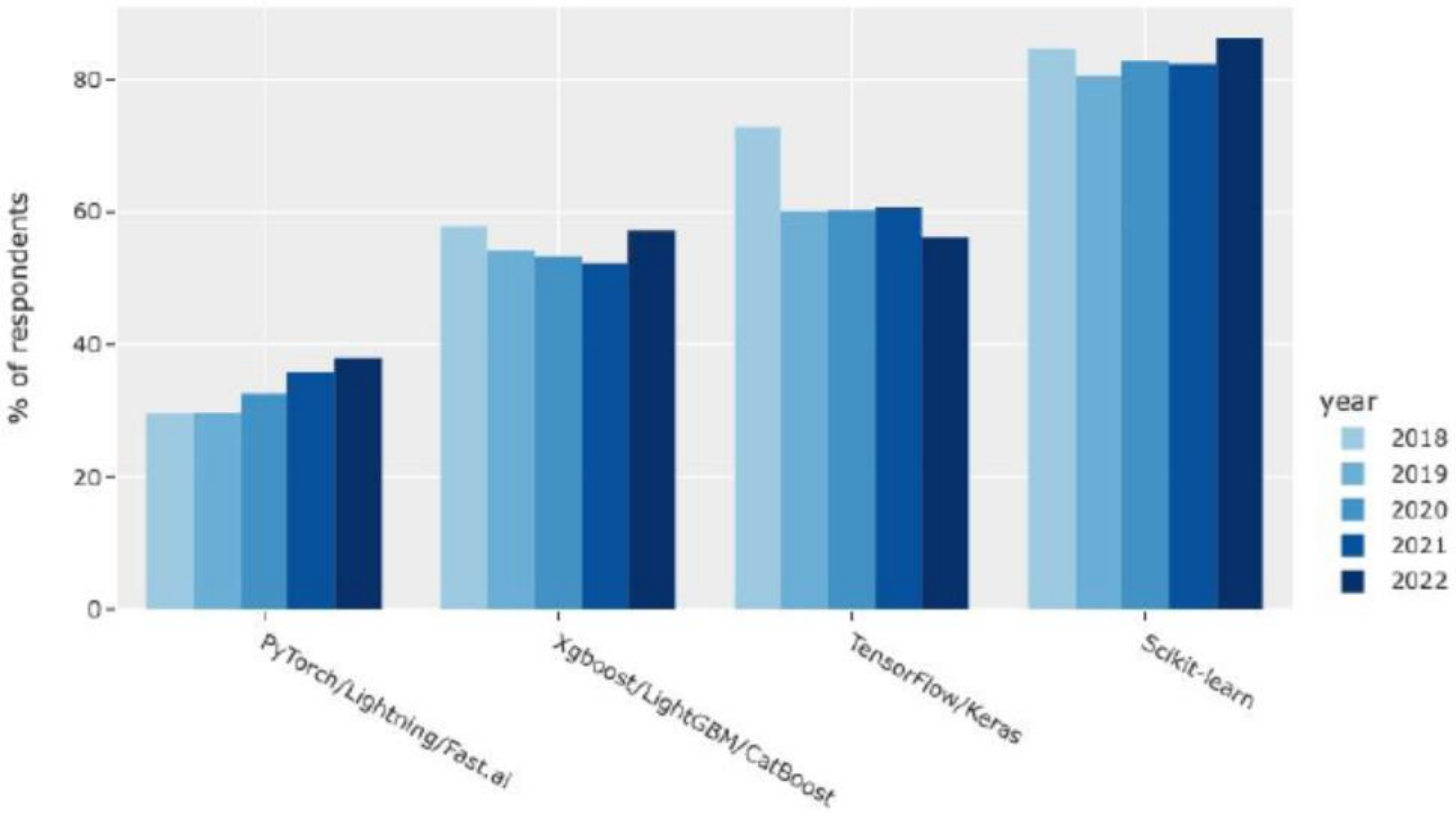
Why Python?

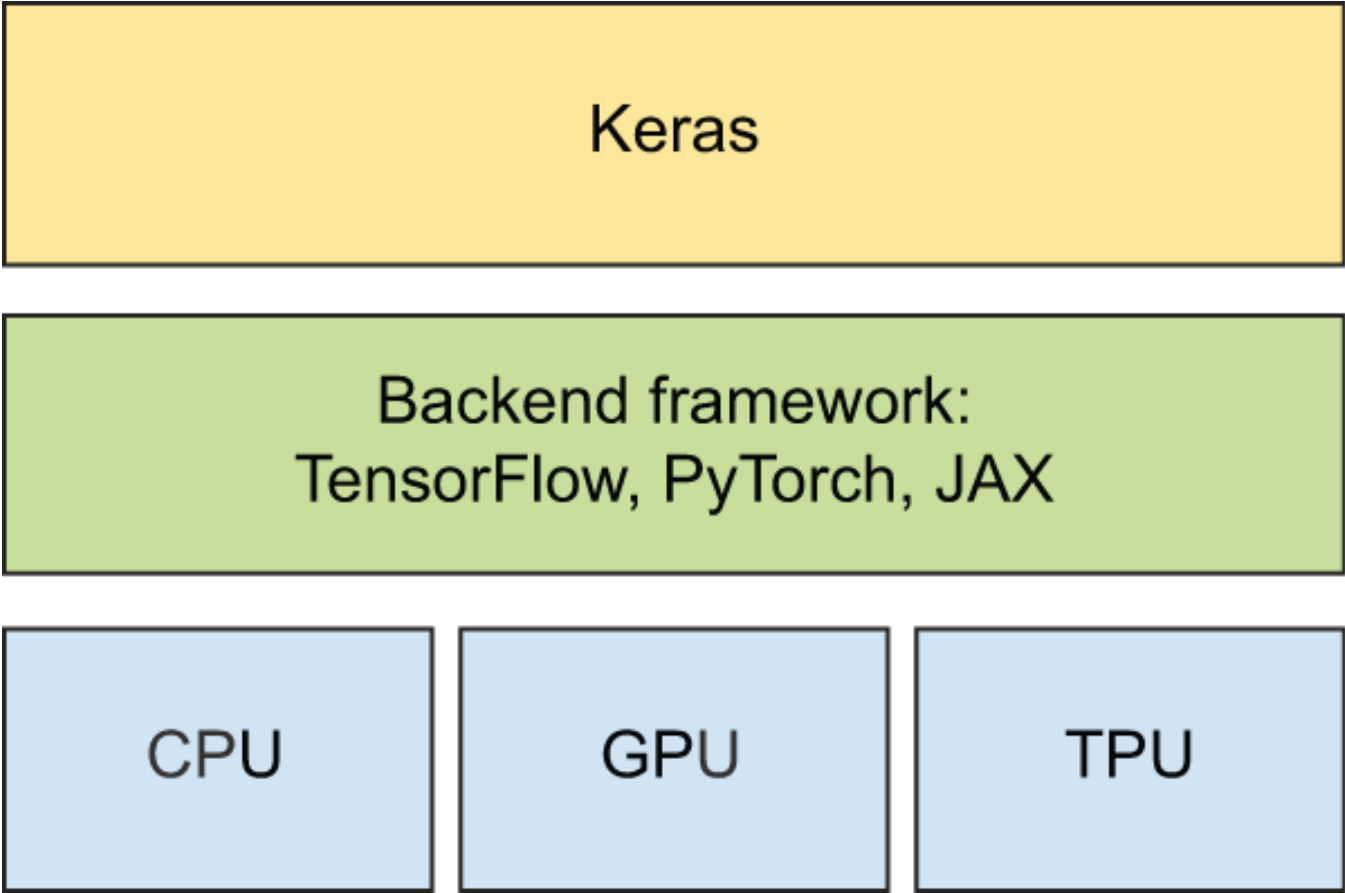
- Free
- Friendly for beginners
- Available resources and packages
- Powerful in web scraping
- Easy to extend (CUDA vs JAX)



% of respondents







Keras

Deep learning development:
layers, models, optimizers, losses,
metrics, training loops ...

Backend framework:
TensorFlow, PyTorch, JAX

Tensor manipulation infrastructure:
tensors, variables, automatic
differentiation, distribution ...

CPU

GPU

TPU

Hardware: execution

Anaconda

- Download Anaconda distribution:
https://repo.anaconda.com/archive/Anaconda3-2026.07-1-Windows-x86_64.exe

- The answers of quizzes and homework will be grading based on the outputs offered by Anaconda. You may use a different IDE but please ensure that it produces the same outcome as follows:

```
import numpy as np  
np.random.seed(234)  
np.random.rand()
```

Out[3]: 0.03..... (there should be 17 digits after the decimal point)