

What is generative AI?

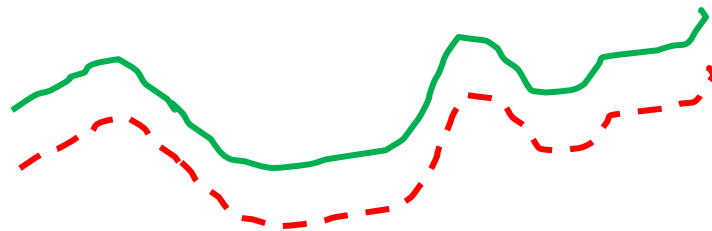
AI

make use of computing to mimics how humans think to work like a human being to complete a task (e.g. conversation, decision making, translation.....)

Machine learning

{ Deep learning)
Supervised/unsupervised
.....

Pattern recognition



e.g. Stock market prediction, DNA expression profile

What is generative AI?

AI

make use of computing to mimics how humans think to work like a human being to complete a task (e.g. conversation, decision making, translation.....)

Machine learning

{ Deep learning)
Supervised/unsupervised
.....

Pattern recognition

Rule-based

E.g. Rule-based expert system for medical diagnosis
Computer virus detection system
If sub-pattern: 01xx0001 => virus

Most of these techniques:
NO new contents will be generated

What is generative AI?

AI

make use of computing to mimics how humans think to work like a human being to complete a task (e.g. conversation, decision making, translation.....)

Machine learning

Pattern recognition

Rule-based

Generative AI

{ Deep learning
Supervised/unsupervised
.....

Text =>
drawing

Here's DALL-E: An algorithm learned to draw anything you tell it

The drawings may look simple (some are better than others), but it's the kind of progress that highlights how artificial intelligence is continuing to gain humanlike capabilities.



Focus on models & algorithms that can be used to "create new contents" such as images and art (DALL-E2), text (chatGPT), music, videos

Generative AI (different models)

```
graph TD; A[Generative AI (different models)] --- B[1. Generative adversarial networks (GAN)]; A --- C[2. Transformer-based model <= chatGPTs are using this]; A --- D[3. Variational autoencoders (VAEs)]; A --- E[4. Autoregressive models]; A --- F[5. Recurrent neural networks (RNN)]; A --- G[6. Reinforcement learning]; A --- H[.....];
```

1. Generative adversarial networks (GAN)

2. Transformer-based model <= chatGPTs are using this

3. Variational autoencoders (VAEs)

4. Autoregressive models

5. Recurrent neural networks (RNN)

6. Reinforcement learning

.....

Transformer:

- Neural network architecture that **can automatically transform one type of input** into **another type of output**

In the context of "text"
to solve the problem of "**sequence to sequence learning**"

English:

I come from China

Translation

=>

Chinese:

我来自中国

I pour coffee from a pot to a cup to fill **it** up.

I pour all coffee from a pot to a cup to empty **it**.

我把咖啡从壶里倒到杯子里，把**它**装满。

我将所有咖啡从壶中倒入杯子中以将**它**倒空。

More example (更多例子):

猫喝了牛奶, 因为**它**饿了。

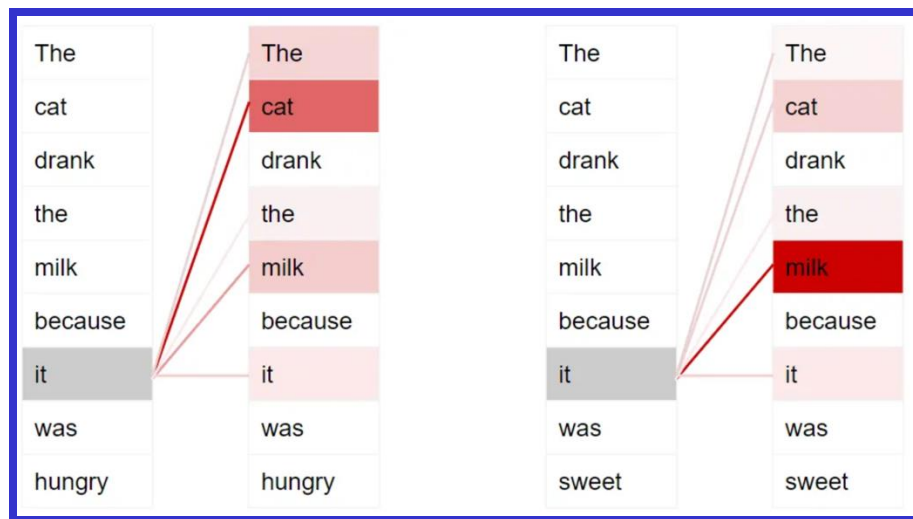
猫喝了牛奶, 因为**它**很甜。

The cat drank the milk because **it** was hungry.

The cat drank the milk because **it** was sweet.

Technical details (omitted)

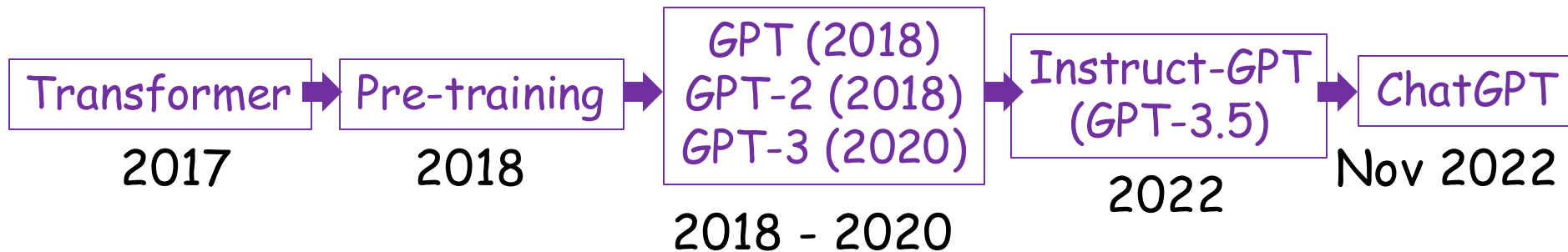
Technical details:
embedding,
encoder, decoder,
attention
mechanism...



=> can also model relationship between different modes of data
(e.g. text to picture) [**multimodal AI**]

How about ChatGPT?

- **Chat** Generative **P**re-trained **T**ransformer:
- Developed by OpenAI, launched in Nov 2022
- Accumulated quite a lot of research



What ChatGPT can do?

e.g. Answer simple queries => writing proposals, codes.....

What is bitcoin?

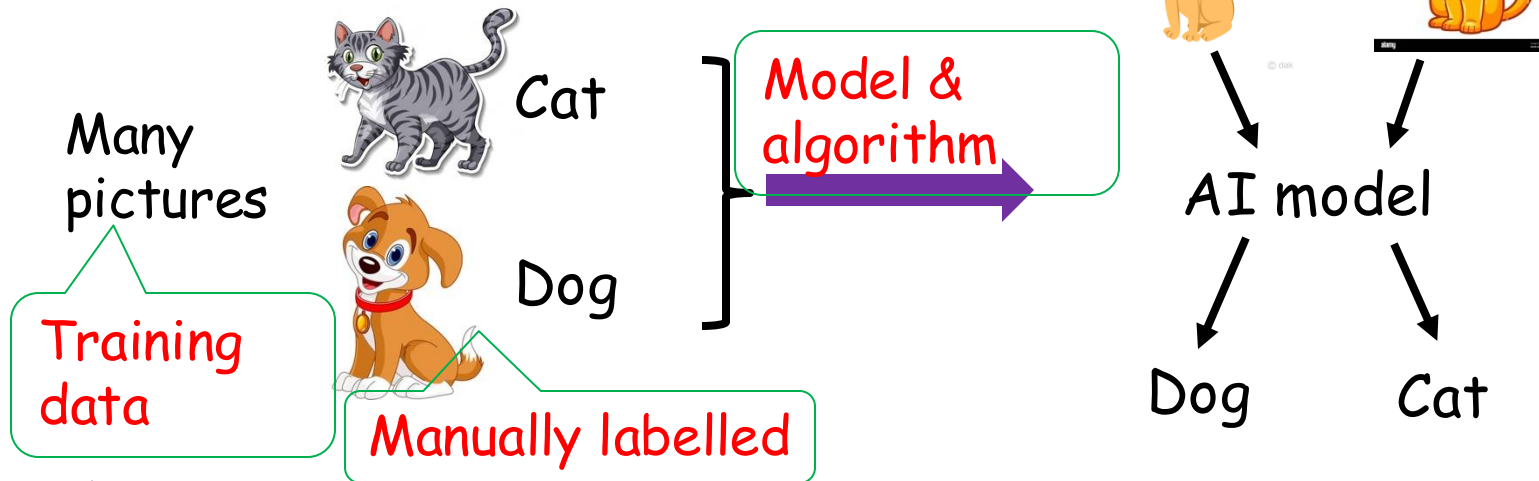
Bitcoin is a decentralized digital currency that operates without the need for a central bank or administrator. It allows for peer-to-peer transactions that can be completed without the need for

Write a c code for generating a prime number

```
Sure, here's a simple C code for generating prime numbers:
#include <stdio.h> int is_prime(int n) { /* Returns 1 if n is prime, 0
otherwise */ if (n < 2) { return 0; } for (int i = 2; i*i <= n; i++) { if (n
% i == 0) { return 0; } } return 1; } int main() { int max_num;
printf("Enter the maximum number to generate primes up to: ");
scanf("%d", &max_num); for (int num = 2; num <= max_num;
num++) { if (is_prime(num)) { printf("%d ", num); } } printf("\n");
return 0; } This program prompts the user for a maximum
number and then generates all prime numbers up to that number
using a basic primality test algorithm. The is_prime() function
returns 1 if a given number is prime, and 0 otherwise. The main()
function calls this function for all numbers up to the maximum
specified by the user, and prints out each prime number found.
```

How ChatGPT learns?

(a) How to train an AI system ?



=> computing resources

OpenAI uses Common Crawl db (billions of webpages) + private data

(b) For ChatGPT:

- How to capture the meaning of words/sentences?
- To manually label so many texts, too time-consuming?

Transformer-based model - to build up a language model for related words within certain context

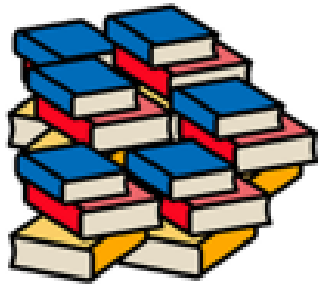
Pre-training - using training data without labels (e.g. techniques such as masked language modeling)

Specific Knowledge
Base

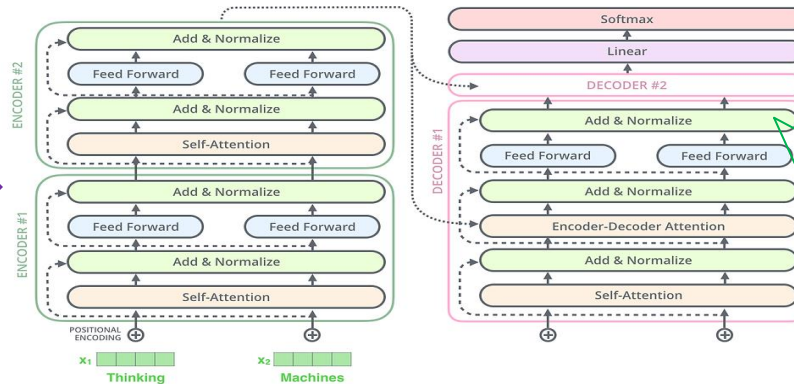


Fine-tuning

Pre-training



Gigantic
LLM Training Set



Read documents,
analyze them,
build up a large
language model

Transformer-based
Model (Large language
model, LLM)

LLM model with trillions
parameters

Requirement of computing
=> resources is huge