



什麼是生成式人工智慧？

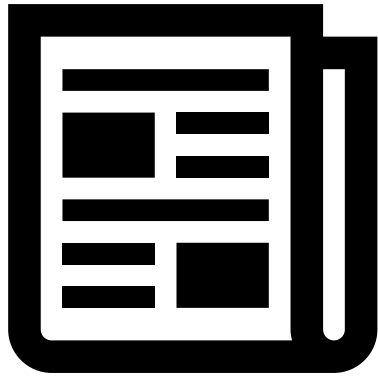
李宏毅 Hung-yi Lee

人工智慧 (Artificial Intelligence, AI)

人工智慧 (目標)

讓機器展現「智慧」

生成式人工智慧 (Generative AI) : 機器產生複雜有結構的物件



文章

由文字所構成



影像

由像素所組成



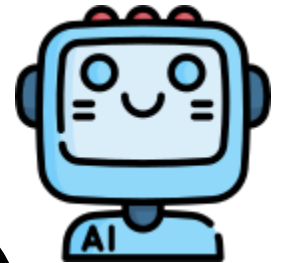
語音

由取樣點構成

生成式人工智慧 (Generative AI) : 機器產生複雜有結構的物件 **盡乎無法窮舉**



一篇 100 字的中文文章，標題為《縫隙的聯想》



用中文字組成 100 字的文章有多少的可能性

假設中文常用字為 1000

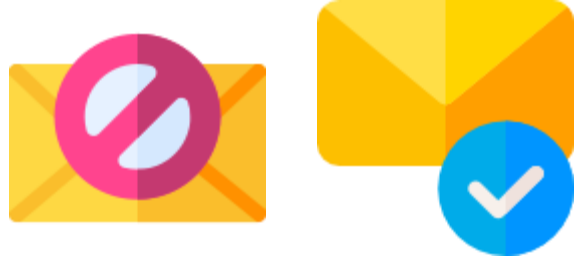
$$1000 \times 1000 \times 1000 \times 1000 \dots = 1000^{100} = 10^{300}$$

宇宙中的原子總數估計 10^{80}

從近乎無窮的可能中找出適當的組合

生成式人工智慧 (Generative AI) : 機器產生複雜有結構的物件

- 分類(Classification)：從有限的選項中做選擇 (不是生成式人工智慧)



垃圾郵件偵測

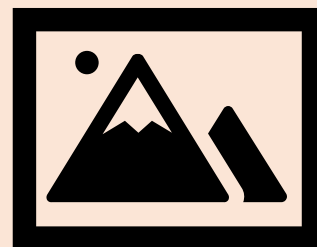


貓狗分類器

生成式人工智慧

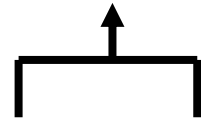
人工智慧 (目標)

生成式人工智慧 (目標之一)



機器學習 \approx 機器自動從資料找一個函式 (Machine Learning)

參數 (Parameter)



有一個函式 $y = f(x) = ax + b$ 假設輸入 $x = 1$, 輸出 $y = ?$

假設輸入 $x = 4$, 輸出 $y = 5$

假設輸入 $x = 2$, 輸出 $y = -1$

請找出 a, b $a = 3, b = -7$

國中的時候你用人力算出參數，機器學習是自動找出參數的方法

假設我要做一個機器可以分辨貓和狗

模型

測試, testing (推論, inference)

上萬個參數

有一個函式 貓 or 狗 = $f(\text{)}$ = ...a...b...c...d...e...f...g.....
 $a = 1.3, b = -7.2, c = 0.4, \dots$



輸出： 貓

輸出： 狗

輸出： 貓

輸出： 狗

機器學習可以把上萬個參數找出來

訓練, training (學習, learning)

$a = 1.3, b = -7.2, c = 0.4, \dots$

假設我要做一個機器可以分辨貓和狗

模型

有一個函式

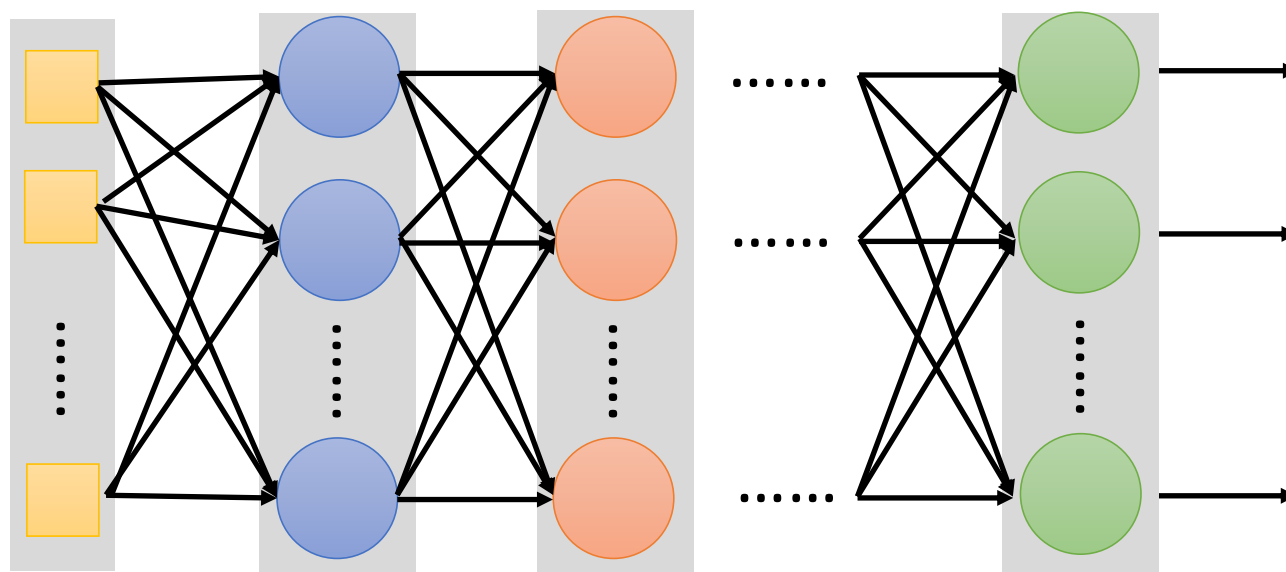
貓 or 狗 = f (



) = ...a...b...c...d...e...f...g.....

上萬個參數

深度學習(Deep Learning)
是一種機器學習技術



類神經網路 (Neural Network)

生成式人工智慧

人工智慧 (目標)

生成式人工智慧 (目標之一)

機器學習 (手段)

深度學習 (更厲害的手段)

生成式人工智慧

人工智慧 (目標)

機器學習 (手段)

深度學習 (更厲害的手段)

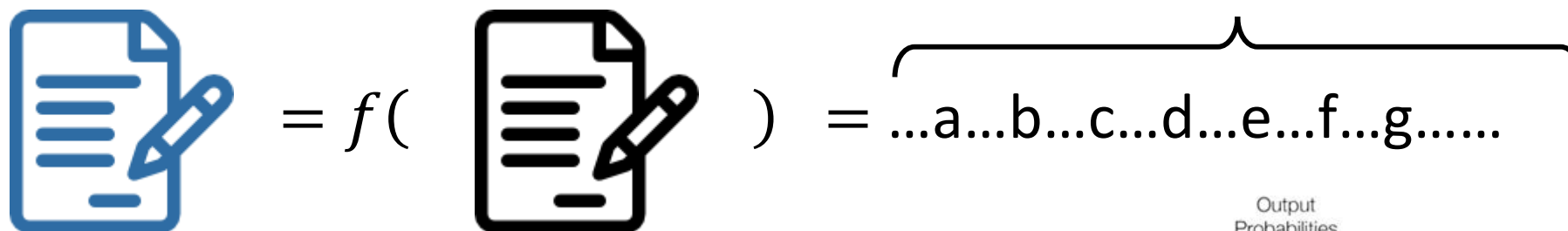
生成式人工智慧

今日的生成式人工智慧多以深度學習達成

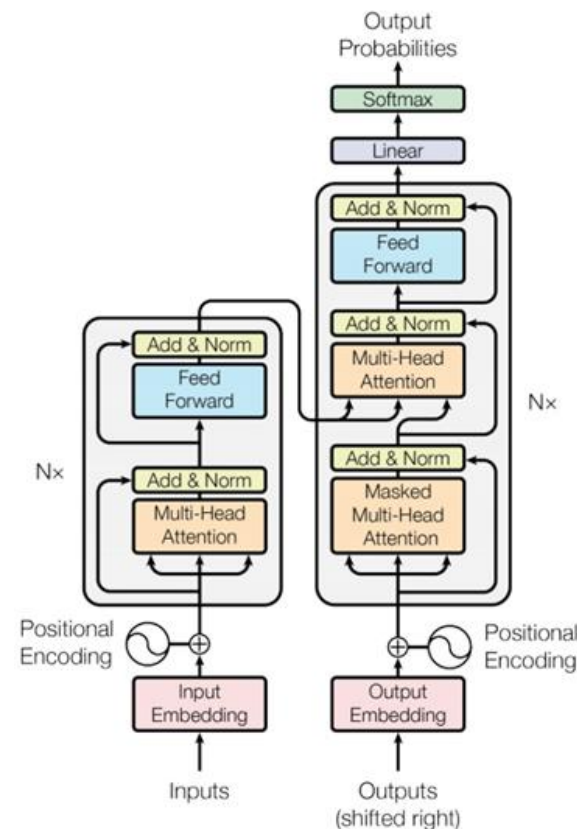
ChatGPT 也就是個函式

模型

上億個參數

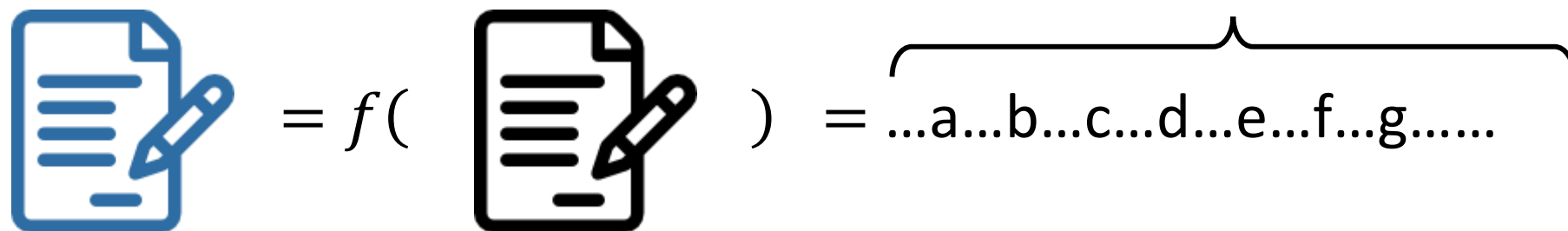


Transformer
(類神經網路的一種)



ChatGPT 也就是個函式

模型



輸入: 何謂人工智慧?

輸出: 人工智慧就是

輸入: 說個跟人工智慧有關的故事

輸出: 很久很久以前

輸入: 寫一首詩

輸出: 床前明月光

輸入: 人工智慧的英文翻譯

輸出: Artificial Intelligence (AI)

機器學習把上億個參數找出來

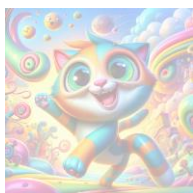
AI 畫圖也就是個函式 (Stable Diffusion, Midjourney, DALL·E)

模型

上億個參數

$$\text{Image Icon} = f(\text{Text Icon}) = \overbrace{\dots a \dots b \dots c \dots d \dots e \dots f \dots g \dots}^{\text{上億個參數}}$$

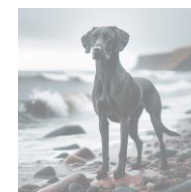
彩色卡通貓



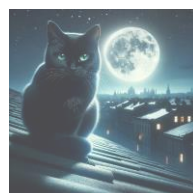
奔跑的貓



海濱的黑狗



月夜下的黑貓



在室內的波絲貓



花叢中的狗



機器學習把上億個參數找出來

2019 年的機器學習投影片

2024 年：

雖然庫拉皮卡抵達暗黑大陸
遙遙無期，但是人類可說已
經見到暗黑大陸的守門人了

Generation (生成)

產生有結構的複雜東西
(例如：文句、圖片)

擬人化的講法—創造

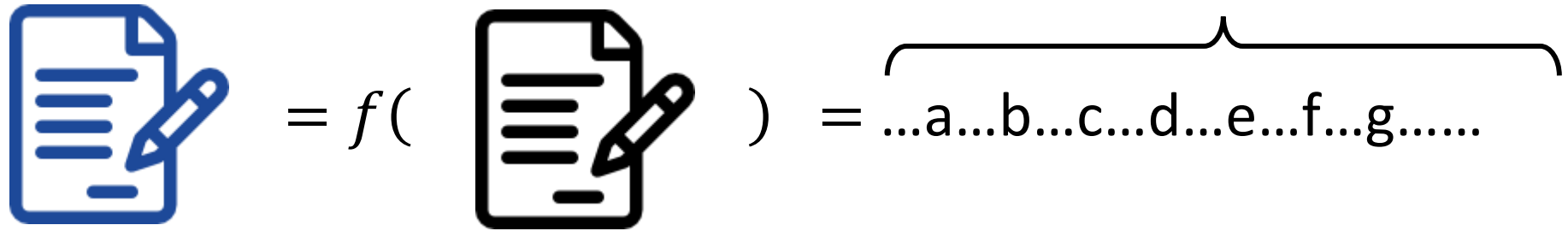
Regression,
Classification



機器需要能夠產生在訓練時從來沒有看過的東西

也就是需要有「創造力」

上億個參數



訓練資料

輸入: 何謂人工智慧?

輸出: 人工智慧就是

輸入: 說個跟人工智慧有關的故事

輸出: 很久很久以前

輸入: 寫一首詩

輸出: 床前明月光

輸入: 人工智慧的英文翻譯

輸出: Artificial Intelligence (AI)

測試時 ...

需要創造
全新的文句



= f (



) 文字: 寫一篇題為
「縫隙的聯想」的文章

文字接龍

原本的目標 臺灣最高的山是哪座？



玉山

可能性
窮盡無盡

拆解成一連串文字接龍 分類問題

臺灣最高的山是哪座？



玉

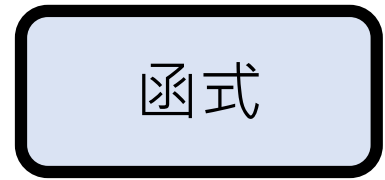
答案有限!

臺灣最高的山是哪座？玉



山

臺灣最高的山是哪座？玉山



[END]

語言模型

生成式人工智慧

人工智慧 (目標)

機器學習 (手段)

深度學習 (更厲害的手段)

生成式人工智慧

語言
模型

生成策略

複雜的物件

文章

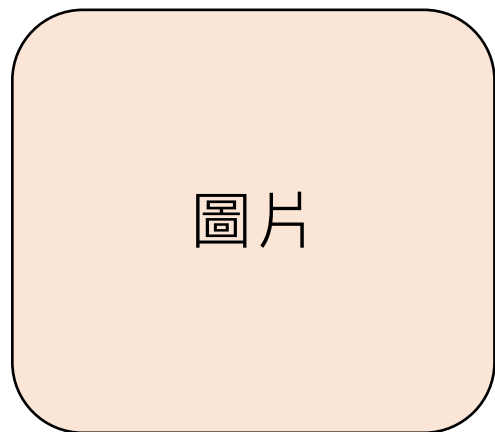
由文字構成



較小的單位

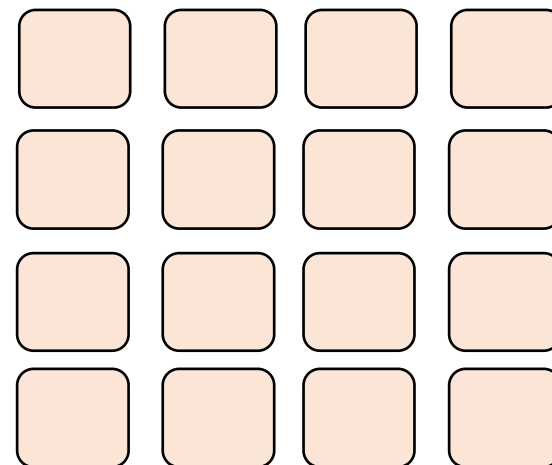
字 字 字 字 字

依照某種固定的順序依序生成



圖片

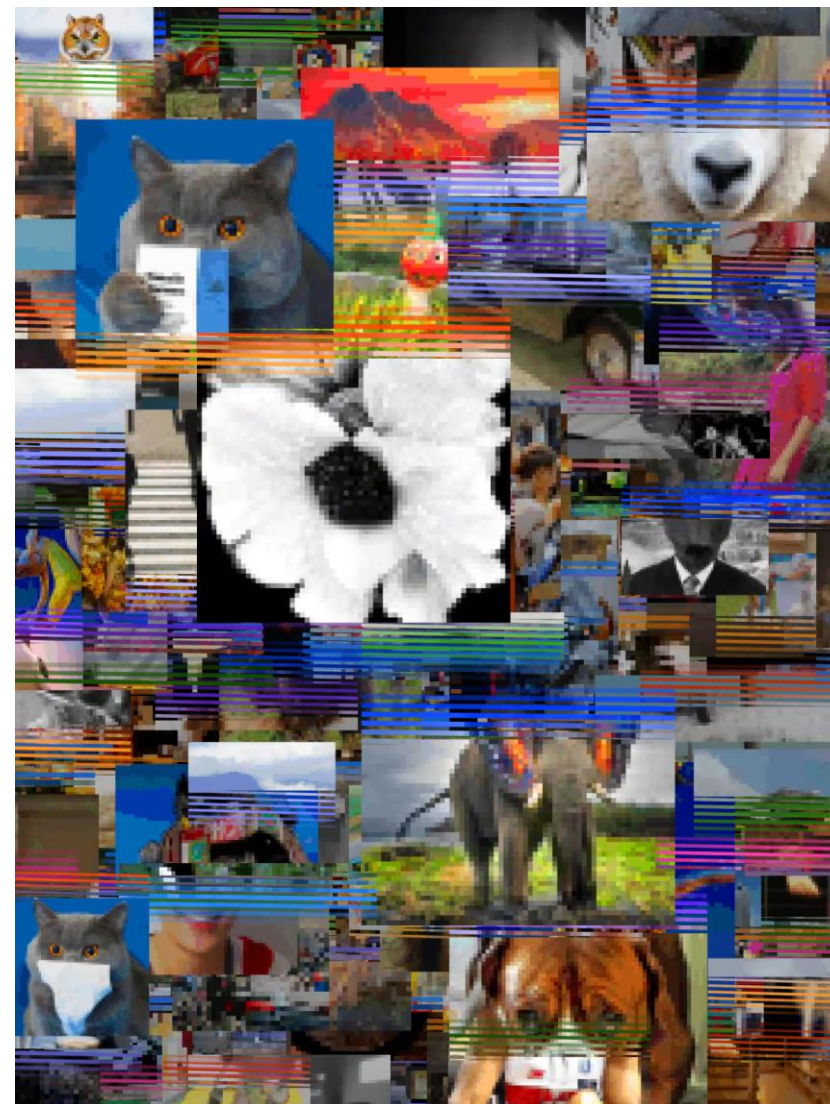
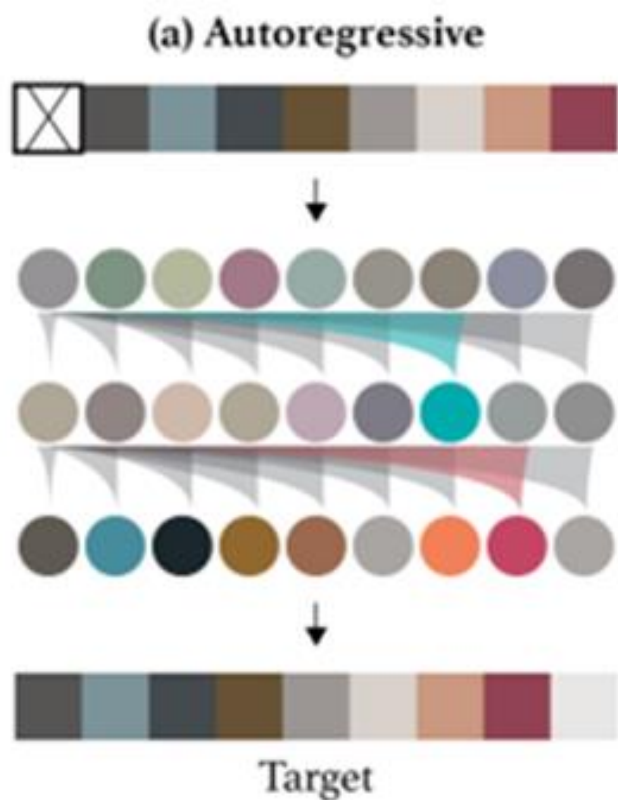
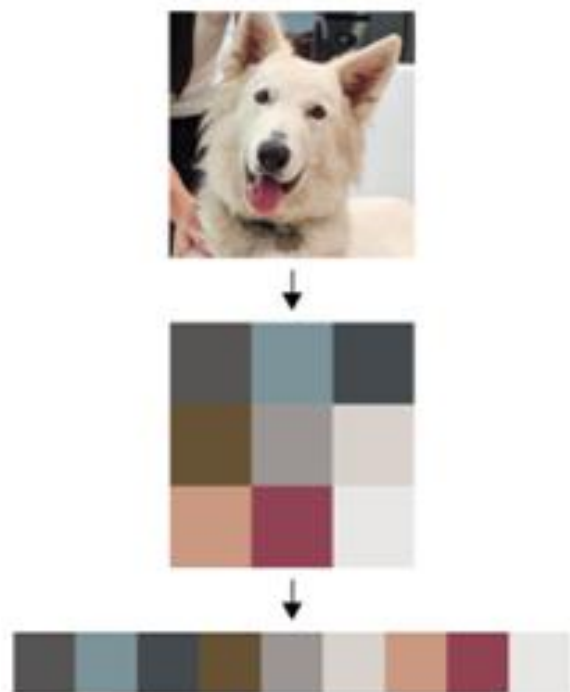
由像素構成



Autoregressive Generation

生成策略

影像版 GPT



<https://openai.com/blog/image-gpt/>

生成式人工智慧不是今天才有 ...

- 機器學習及其深層與結構化 (2015)

https://speech.ee.ntu.edu.tw/~tlkagk/courses_MLSD15_2.html

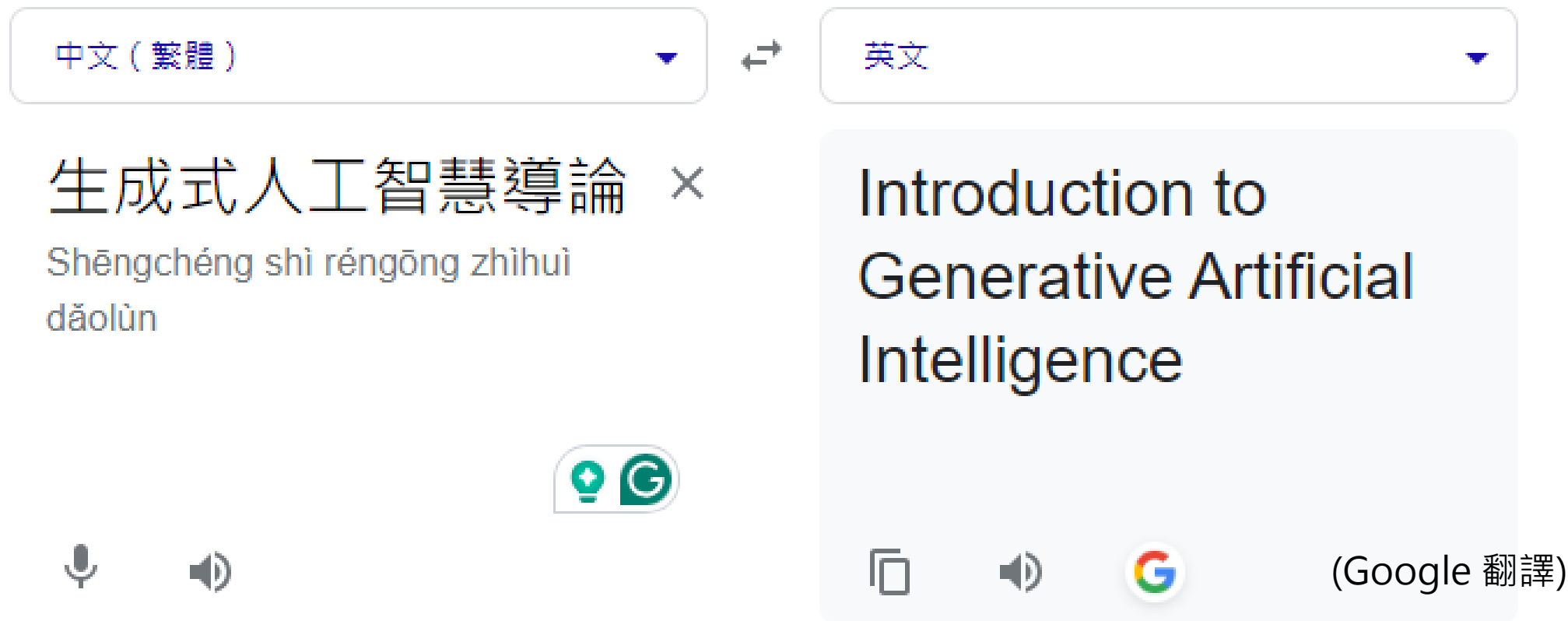
Deep Learning

- Single Neuron - Logistic Regression [ppt, pdf](#) (2015/03/13)
- Neural Network - Introduction [ppt, pdf, EverCam, mp4](#) (2015/03/13)
- Backpropagation - An efficient way to compute the gradient [ppt, pdf, EverCam, mp4](#) (2015/03/20)
- Tips for Training Neural Network - scratch the surface [ppt, pdf, EverCam, mp4](#) (2015/03/20)
- More Tips for Training Neural Network [ppt, pdf, EverCam\(part1\), EverCam\(part2\), mp4\(part1\), mp4\(part2\)](#) (2015/04/24)
 - Motivation of Adagrad and RMSprop [ppt, pdf, EverCam](#) (2015/05/08)
- Do machines know the meaning of a word? [ppt, pdf, EverCam](#) (2015/05/01)
- Neural Network with Memory (1) [ppt, pdf, EverCam, mp4](#) (2015/05/01)
- Neural Network with Memory (2) [ppt, pdf, EverCam, mp4](#) (2015/05/08)
 - Techniques related to RNN based Language Model [ppt, pdf, mp4](#) (2015/05/29)
- Introduction of cuDNN [pdf](#) (2015/05/15)
- Deep Learning Methods for Image Classification [pdf](#) (2015/05/22)
- Exploiting Unlabeled Data [ppt, pdf, mp4](#) (2015/05/29)
- Deep Learning for Speech Recognition [ppt, pdf, mp4](#) (2015/06/05)

Structured Learning

- Three Basic Problems [ppt, pdf, EverCam, mp4](#) (2015/03/20)
- Sequence (1) [ppt, pdf, EverCam, mp4](#) (2015/04/10)
- Sequence (2) [ppt, pdf, EverCam](#) (2015/04/17)
- Structured SVM (Objet Detection) [ppt, pdf, EverCam, mp4](#) (2015/04/17)
- Structured SVM + Hidden Information (Objet Detection) [ppt, pdf, EverCam](#) (2015/05/08)
- Structured SVM + Hidden Information (Summarization) [ppt, pdf, EverCam, mp4](#) (2015/05/08)
- Graphical Model and Gibbs Sampling [ppt, pdf, mp4](#) (2015/05/15)
 - A little bit of theory for Gibbs Sampling [ppt, pdf, mp4](#) (2015/05/29)
- Markov Logic Network [ppt, pdf, mp4](#) (2015/06/12)

生成式人工智慧不是今天才有 ...



The screenshot displays the Google Translate interface. On the left, the source language is set to '中文 (繁體)' and the text '生成式人工智慧導論' is entered. Below the text is its pinyin: 'Shēngchéng shì réngōng zhìhuì dǎolùn'. On the right, the target language is '英文', and the translated text is 'Introduction to Generative Artificial Intelligence'. The interface includes a bidirectional arrow between the language boxes, a close button for the source text, and icons for voice input, output, and a Google logo. The text '(Google 翻譯)' is visible at the bottom right of the translation area.

今日的生成式人工智慧有甚麼特別厲害的地方呢？