2025 年【科學探究競賽-這樣教我就懂】

大專/社會組　科學文章格式

文章題目： 人工智慧怎麼懂人話？我們真的知道它「在想什麼」嗎？

摘要：大型語言模型（LLM）是透過 Transformer 架構建立的人工智慧，能理解並產生自然語言。它的核心在於自我注意力機制、多層結構與大量語料訓練，透過預測下一個字來學習語言模式，模仿出「理解」人類語言的能力。

文章內容：（限 500 字~1,500 字）

人工智慧怎麼懂人話？

你有沒有想過現在在你手機裡、電腦裡、甚至家裡智慧音箱裡的 AI，它們是怎麼知道你在說什麼的？你傳訊息時，它能幫你自動補字；你跟它聊天時，它回你的話不僅通順，還有點像「真的在思考」。但問題來了：" 它真的「懂」你嗎？還是只是在裝懂？"

這背後的關鍵角色，是一種叫做**大型語言模型**（Large Language Model，簡稱 LLM）的技術。GPT-4、ChatGPT、Claude、Gemini，全都靠它在運作。

這篇報告不講艱深數學、不用看程式碼，而是用**你能聽懂的語言**，揭開這台「語言機器」的真面目——它是怎麼學會人類語言？怎麼組織知識？怎麼在你說話的那一瞬間，給你一個像是「真的有在思考」的回答？

什麼是大型語言模型？

大型語言模型（LLM）是一種專門設計來「理解和產生語言」的人工智慧。你可以把它想像成一個超級強大的自動補字機器，但它不只是會補單字，而是能理解句子的意思，甚至產生新的句子來回答問題、寫故事、解釋知識。

它是怎麼做到的？模型架構：Transformer

大部分 LLM（像是 GPT-4、Claude、Gemini）都是基於一種叫做 **Transformer** 的模型架構。這個架構是 2017 年由 Google 的研究團隊提出的，被認為是近代 AI 革命的基石。

- Transformer 的三個關鍵概念：

1. **Self-Attention（自我注意力）**

   想像你在看一句話：「她拿著雨傘因為下雨了。」要理解「她為什麼拿雨傘」，你得注意到「因為下雨了」。Transformer 就是用這種「注意力機制」來判斷句子裡哪些字與哪些字有關聯。

2. **Layer（層）**

Transformer 不是一層模型，而是由很多層組成（例如 GPT-3 有 96 層）。每一層都會做類似的事情：判斷單字之間的關係、萃取語意，但每層會看得更廣、更深。

3. **Embedding（嵌入表示）**

電腦不懂「字」，所以我們要把文字轉換成數字，這就叫做 Embedding。比如「貓」可能會被轉成 [0.21, -0.35, 0.98, ...] 的數字向量，這些數字可以讓模型捕捉到「語意」。

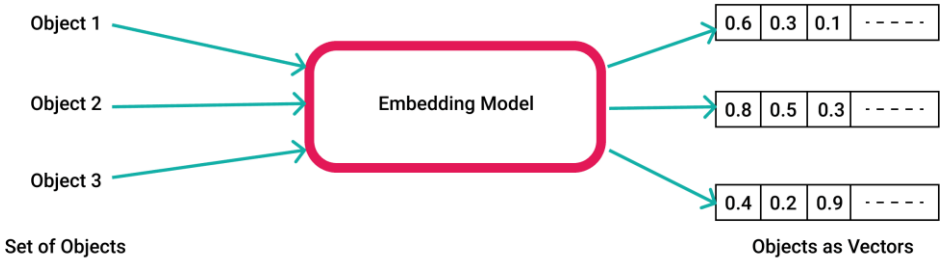

圖 1 Embedding

模型是怎麼「長出知識」的？

- **訓練的方式：預測下一個字（Next Token Prediction）**

訓練這些模型的方法很簡單，讓它們讀大量的文字資料（像是維基百科、書籍、網站），然後問它：「下一個字是什麼？」

例如：

輸入：我今天心情很好，因為天氣

預測：好（或晴、很棒……）

這種方式叫做 **自監督學習（self-supervised learning）**，不需要人類標註，只要大量資料。

- **Loss Function（損失函數）：告訴模型學得好不好**

每次預測錯字，模型就會被「扣分」，這個扣分叫做 Loss。模型會不斷調整裡面的參數，目的是「少犯錯、多預測對」。

模型的內部長什麼樣？Layer 架構淺談

每一層 Transformer Layer 通常包含以下元件，參考圖 2:

1. **Multi-Head Attention（多頭注意力）**
   把注意力拆成很多「頭」，同時注意句子不同部份。例如一頭關注主詞，一頭關注時間，一頭關注地點。
2. **Feed Forward Network（前饋神經網路）**
   這部分類似傳統的神經網路，幫助模型將剛剛注意到的資訊做進一步處理。
3. **Layer Normalization（層正規化）**
   保持資料穩定，不會因為某些數值太大太小而導致整體訓練不穩。
4. **Residual Connection（殘差連接）**
   幫助資訊在多層之間流通不會消失，讓深層網路也能有效學習。

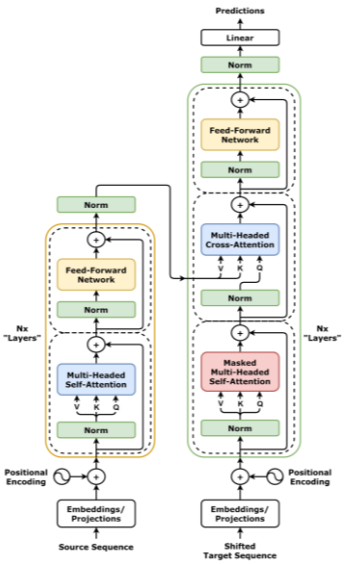

圖 2 模型架構

模型大小代表什麼？

模型大小通常指的是「參數的數量」。參數越多，模型越聰明，但也越需要資源。

- GPT-2：約 1.5 億個參數
- GPT-3：約 1750 億個參數
- GPT-4：官方沒說，但外界推測可能超過 1 兆參數

參數就像是模型的「腦細胞」，數量越多，學習能力越強，但也越貴、越難訓練。

模型怎麼知道「人話」的意思？

它其實「不懂人話」，它只是非常擅長找出「字與字的關係」。但因為訓練資料夠多，當你問它問題時，它會從類似問題中學到「什麼樣的回答看起來最合理」。這種「合理性的模仿」，讓它看起來像真的「懂」了，但其實它是在「模仿懂的樣子」。這就是 LLM 神奇但也讓人敬畏的地方。

參考資料

1. A Very Gentle Introduction to Large Language Models without the Hype (Mark Riedl, 2023)
   https://mark-riedl.medium.com/a-very-gentle-introduction-to-large-language-models-without-the-hype-5f67941fa59e
2. How Transformers Work: A Detailed Exploration of Transformer Architecture (Josep Ferrer, 2024)
   https://www.datacamp.com/tutorial/how-transformers-work
3. Introduction to Large Language Models | Machine Learning (Google, 2024)
   https://developers.google.com/machine-learning/resources/intro-llms

4. What Is a Transformer Model? | NVIDIA Blogs (Rick Merritt, 2022)
https://blogs.nvidia.com/blog/what-is-a-transformer-model/
5. Understanding the Transformer architecture for neural networks (Jeremy Jordan, 2023)
https://www.jeremyjordan.me/transformer-architecture/
6. What are Vector Embeddings (圖 1, Rajat Tripathi, 2023)
https://www.pinecone.io/learn/vector-embeddings/
7. Transformer 模型 (圖 2, Wikipedia, 2025)
https://zh.wikipedia.org/zh-tw/Transformer%E6%A8%A1%E5%9E%8B

註：

1. 未使用本競賽官網提供「科學文章表單」格式投稿，**將不予審查**。
2. 字數沒按照本競賽官網規定之限 500 字~1,500 字，**將不予審查。**
   PS.摘要、參考資料與圖表說明文字不計入。
3. 建議格式如下：
   - 中文字型：微軟正黑體；英文、阿拉伯數字字型：Times New Roman
   - 字體：12pt 為原則，若有需要，圖、表及附錄內的文字、數字得略小於 12pt，不得低於 10pt
   - 字體行距，以固定行高 20 點為原則
   - 表標題的排列方式為向表上方置中、對齊該表。圖標題的排列方式為向圖下方置中、對齊該圖