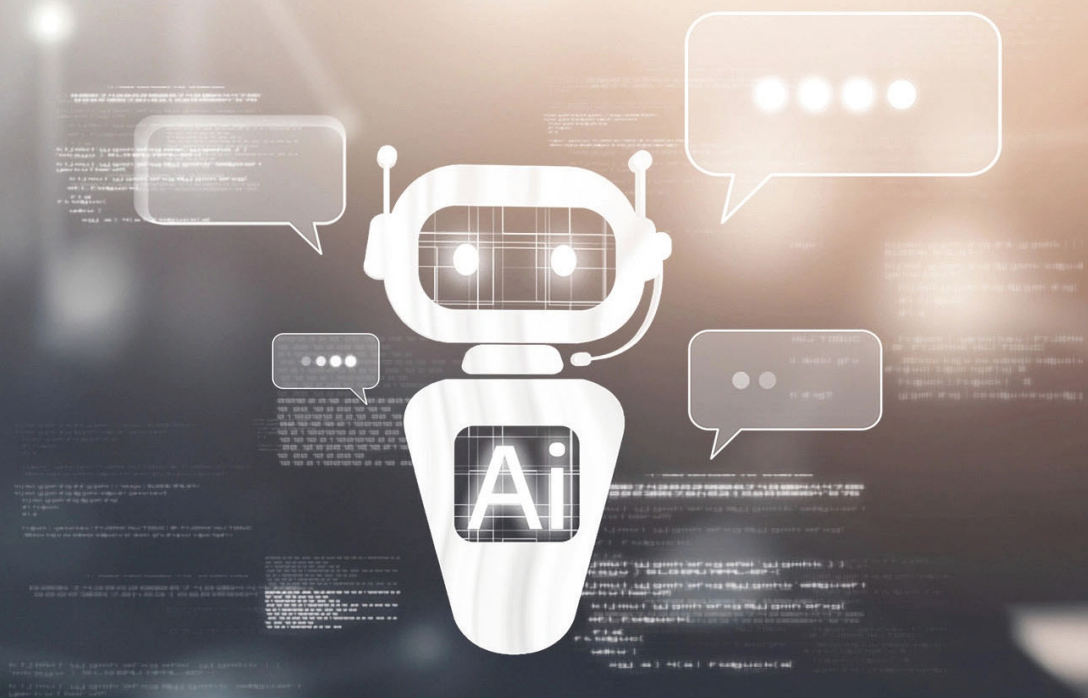




淺談大型語言模型在醫療應用 發展與挑戰

蔡曉婷 / 高雄醫學大學 人工智慧生醫研究院 博士後研究員

吳崇豪 / 高雄醫學大學 人工智慧生醫研究院 博士後研究員

李沁璟 / 高雄醫學大學附設中和紀念醫院 體系智能醫療推動辦公室執行秘書

盧鴻興 / 高雄醫學大學附設中和紀念醫院 副院長

一、前言

人工智慧 (Artificial Intelligence, AI) 在近年間取得了顯著進展，其中自然語言處理 (Natural Language Processing, NLP) 作為 AI 的核心應用之一，在人機互動領域發展尤為迅速。而大型語言模型 (Large Language Models, LLMs) 作為 AI 在 NLP 領域的突破性成果，通過大規模預訓練，使機器在文本生成、語言翻譯及互動式對話等方面展現卓越能力。在現代醫療體系中，不同的大型語言模型各有其特色，其中原生型大語言模型 (Non-distilled LLMs) 擁有廣泛的知識範圍與強大的生成能力，而蒸餾的大語言模型 (Distilled LLMs) 則具備了高效率、低成本特點；另外，通用型語言模型 (General-purpose LLMs) 提供全方位的人工智慧能力，可應用於多種場景，包括對話、內容生成、翻譯、情感分析、代碼輔助等；特定用途語言模型 (Domain-specific LLMs) 專注於特定領域，例如醫學領域，具備更精確的專業知識與術語處理能力。然而，如何在醫療學術研究與臨床應用中合理選擇與運用這些 AI 工具，並在提升醫療效率的同時，兼顧數據隱私、倫理風險及模型可靠性，是當前醫療界關注的核心議題。

二、大型語言模型在醫學領域的應用與發展

大型語言模型 (LLMs) 在醫學領域的應用日益廣泛，涵蓋醫學知識問答、醫療文書輔

助處理、臨床診斷輔助、病人衛教、醫學寫作、醫學教育與訓練、醫務管理及護理實務等多個層面，主要可分為學術研究與臨床實務兩個面向。

在學術面向上，包含有但不限於（1）**醫學知識問答與決策支持**：LLMs 透過訓練大量醫學文獻與臨床資料，可為醫療人員及病人提供實證醫學（Evidence-based Medicine）的資訊。例如，Google 的 Med-PaLM 2 在醫學問題回答上的表現已接近專業醫師水準。

（2）**醫學寫作與教育**：提升醫學論文撰寫的效率，並在醫學教育與培訓中提供個人化學習資源，幫助醫學生及住院醫師掌握最新的臨床實踐與研究進展。（3）**醫學專案與研究管理**：LLMs 可加速醫學文獻分析，提高研究人員的資訊獲取效率，並結合生物資訊學（Bioinformatics）應用於新藥開發、基因組研究與個人化治療。

在臨床實務面向上，包含有但不限於（1）**病歷與醫療文書處理**：LLMs 可輔助自動生成病歷草稿，提高醫師書寫效率，減少行政負擔，並能優化醫療紀錄檢索及結構化資料整理，提升臨床資訊管理的效率。（2）**醫病溝通與護理實務應用**：能夠協助醫護人員進行病人教育，提供**個人化健康諮詢**，如慢性病管理、用藥指導及術後照護，並透過醫療聊天機器人（Chatbots）解答常見健康問題，減少門診負擔，提高病人滿意度。（3）**臨床診斷與輔助治療**：LLMs 可結合電子病歷（Electronic Health Records, EHRs）等數據，進行疾病篩檢，如**糖尿病與高血壓的早期預測**，並在罕見疾病診斷中提供鑑別診斷建議，以提高診斷準確率。特別是在癌症精準醫療領域，LLMs 可結合個人化醫療與癌症基因檢測，透過大數據分析與科學數據支援，制定個人化治療計劃。這樣的癌症治療模式可應用於多個層面，包括早期診斷與預防、治療方法及用藥決策、提升用藥安全與效能，以及癌症的長期追蹤與照護。

三、比較不同的大型語言模型

大型語言模型（LLMs）按照訓練方式可分為原生型大語言模型（Non-distilled LLMs）及蒸餾型語言模型（Distilled LLMs）；依據應用範疇區分，可分為通用型大語言模型（General-purpose LLMs）及特定用途的語言模型（Domain-specific LLMs），其相關的定義、模型範例與優勢及劣勢可見下表一。

表一 比較不同的大型語言模型

1. 訓練方式				
類型	定義	模型範例	優勢	劣勢
原生型大語言模型 (Non-distilled LLMs)	直接基於大規模語言模型架構訓練，較大且性能較強	Claude 3, Gemini 1.5, GPT-4, Llama 3	生成能力強，理解能力佳，能處理更複雜的任務	計算資源需求大，推理速度較慢，部署成本高
蒸餾型語言模型 (Distilled LLMs)	從現有的 LLMs 提取精華並壓縮模型參數，通常較輕量化	Claude 3.5 Haiku, DeepSeek-R1-Distill-Qwen-1.5B	計算成本低，運行速度快，適合邊緣設備和小型應用場景	可能會損失部分推理能力，知識範圍可能較窄
2. 應用範疇				
通用型語言模型 (General-purpose LLMs)	廣泛使用在各種不同的領域的大語言模型	Claude 3, DeepSeek-V3, Gemini 1.5, GPT-4, Llama 3	生成能力強，理解能力佳，能處理廣泛的任務	可適用範圍寬廣，用於各類設定的基本用途，專業術語精準度較低
特定用途語言模型 (Domain-specific LLMs)	透過專業領域數據進行從零訓練 (training from scratch) 或微調 (fine-tuning)，以更適用於特定領域 (如醫學、金融等)	BioGPT (生物醫學), BloombergGPT (金融), Med-PaLM 2 (醫學)	專業領域表現優越，能提供更專業、準確的知識，能處理行業特定的術語與專業數據	適用範圍狹窄，不適用於非專門設定的用途

為了更形象化地說明原生型與蒸餾型、通用型與特殊用途語言模型之差異，我們可以借用烹飪、製藥等比喻來描述。

以烹飪來說，原生型 LLMs 與蒸餾型 LLMs，類似於「米其林**頂級料理**」與「米其林聯名款**平價料理**」之差異。米其林頂級料理製作需耗費相當多的時間心力，製作成本高昂，因此費用並非所有人都負擔得起。而米其林聯名款平價料理則以萃取精華來降低成本、提高效率以滿足大眾需求。通用型 LLMs 與特定用途 LLMs，就像是「**通用料理**」與「**主題料理**」。通用料理偏向適合各種文化、飲食習慣和需求的菜餚，例如三明治（內餡可自由搭配），主題料理（Themed Cuisine）則是以特定概念、文化、節日或故事為核心設計的菜餚或餐飲體驗，例如情人節浪漫料理。

以製藥來說，原生型 LLMs 與蒸餾型 LLMs 就像是「**原廠藥**」與「**學名藥**」的差異。原廠藥經過大量研究與測試，確保在多種臨床場合中都有良好的表現。然而，它的開發成本高昂。學名藥則透過降低開發成本，仍能達到與原廠藥相似的療效。而通用型 LLMs 與特殊用途 LLMs 就像是「**傳統化療藥物**」與「**標靶藥物**」的差異。「傳統化療藥物」是針對所有活躍分裂的細胞，因此在攻擊癌細胞的同時，也可能影響正常細胞，導致較大的副作用。而「標靶藥物」則透過精準識別特定類型的癌細胞，針對其訊息傳導、增殖機制或生長環境進行干擾，以抑制癌細胞的生長或促使其凋亡。

四、結論

LLMs 與多模態 LLMs 在智慧醫療領域的發展潛力巨大，但在應用過程中，仍須審慎考量**隱私保護、數據合規、資安與法律限制**，以確保應用安全性與可靠性。高醫醫療體系作為南臺灣智慧醫療的重要推動者，透過醫療數據資料庫平台 (KMURD) 結合 LLMs 技術，將臨床數據、醫學影像、基因分析等結構化應用於智慧醫療，如智慧加護病房、手術麻醉管理、遠距醫療與智能病歷協作，並促進醫療與產業的跨域合作。為提升護理效率與職場友善度，高醫「護理小護手」LLM 系統，導入整合自建的 KMUGPT 生成式 AI，輔助護理人員快速生成護理與交班紀錄。該系統可自動擷取醫病數據，將**書寫時間從約 1 小時縮短至 3-10 分鐘**，提升紀錄準確性並減輕護理壓力，讓護理專業回歸病人照護。未來，「護理小護手」將優化**生成紀錄準確性**，並整合**護理決策支持系統 (CDSS)**，透過 AI 識別**病人潛在風險**並提供**護理建議**，提升護理決策的精準度，並可進一步將結合**大數據分析**，協助**趨勢預測、工作負荷評估與人力配置優化**，打造更高效的護理環境。在享受 AI 帶來便利的同時，高醫亦**落實資料保護與法規合規性**，確保 AI 在智慧醫療的負責任應用，並攜手醫療夥伴共同推動**AI 臨床落地應用**，提升醫護生產力與醫療服務品質，造福更多病人。



蔡曉婷 博士後研究員

高雄醫學大學人工智慧生醫研究院 博士後研究員
樹人醫護管理專科學校 兼任助理教授
輔英科技大學 兼任助理教授
高雄醫學大學護理學系 博士



吳崇豪 博士後研究員

高雄醫學大學 人工智慧生醫研究院 博士後研究員
國立陽明交通大學 統計學研究所 博士後研究員
國立交通大學 電子研究所 博士



李沁璟 執行秘書

高雄醫學大學附設中和紀念醫院資訊室 副主任
高雄醫學大學附設中和紀念醫院體系智能醫療推動辦公室 執行秘書
社團法人臺灣醫學資訊學會 理事
國立中山大學公共事務管理所 博士



盧鴻興 副院長

高雄醫學大學附設中和紀念醫院 副院長
高雄醫學大學 人工智慧生醫研究院 院長
高雄醫學大學學士後醫學系 特聘教授
國立陽明交通大學統計學研究所 特聘教授