

AI 大型模型比一比！一探 LLM、LAM 與 LVM

你知道大型模型其實是個家族嗎？隨著人工智慧的進步，除了 LLM（Large language model, 大型語言模型），還有 LAM（Large action model, 大型行動模型）和 LVM（Large vision model, 大型視覺模型）各展所長，成為推動科技革新的重要力量。

本文將帶來一探這三大模型的特色與應用。



LLM (Large language model, 大型語言模型)

透過深度學習技術和大量的文本資料，來理解和生成自然語言。模型主要基於變壓器架構，特別是生成預訓練變壓器（如 GPT 系列），使其在處理文本輸入時表現出色。會透過學習文本中的統計關係來預測下一個詞，從而生成語境相關的文本，應用包括聊天機器人、內容創建、語言翻譯等。

LAM (Large action model, 大型行動模型)

結合了語言理解與執行具體任務的能力，例如在各種應用中預訂機票或安排會議。這類模型通常整合了神經符號編程，允許模型理解使用者的意圖，並與其他應用程序或系統互動來執行任務，例如基於 ChatGPT 的 Rabbit R1 就是一種整合了 LAM 的裝置，可以直接通過使用者介面來執行多項任務，大幅簡化操作。

LVM (Large vision model, 大型視覺模型)

專注於處理和分析視覺數據，如圖像和影片，通常會利用深度學習架構——尤其是卷積神經網絡（CNN）——訓練大量數據集，來進行物件識別、圖像分類和場景解析等任務。LVM 具有高精度度，能分析複雜的視覺數據，從而改善醫學成像、提升診斷效率，並通過優化物體分類和駕駛輔助技術來增進汽車系統的安全性。

LLM 專注於語言的理解和生成；LAM 則結合了語言理解與行動執行，能夠與應用程式直接互動；而 LVM 處理的是視覺數據，適用於圖像相關的任務。這三者的主要區別在於它們的應用焦點和技術實現方式。

隨技術不斷演進，小編相信三大模型將在更多領域內展現出潛力和影響力，而深入理解這些模型的運作機制及其未來發展，將是掌握趨勢的關鍵。你覺得大型模型會如何改變世界呢？