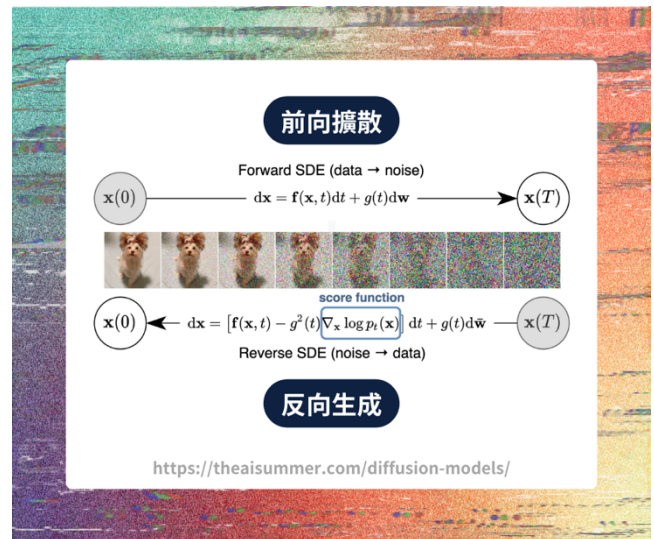


影像生成式 AI 的現任霸主：擴散模型 Diffusion Model

關注生成式 AI 發展的讀者們，一定還記得最近轟動一時的文字轉影片 AI Sora，這個厲害的模型就是 diffusion transformer。之前我們已經簡單聊過 diffusion model，今天就讓我們來深入挖掘，來探討這個目前稱霸影像生成 AI 領域的擴散模型吧！



擴散模型究竟是？

擴散模型是一種基於深度學習的生成模型，通過模擬數據的擴散過程，來生成新的數據樣本。想象將一滴墨水滴入水中，墨水會逐漸擴散開來，直到整杯水顏色變為一致。

原理基於兩個主要步驟：前向擴散和反向生成。前向擴散就是逐步給原始數據加噪聲，直到變成純噪聲；反向生成則是學習如何從這堆隨機噪聲中，一步步恢復原始的數據樣本。

擴散模型的研究

擴散模型的概念最早可以追溯到 2015 年，由 Sohl-Dickstein 等人發表的論文《Deep Unsupervised Learning using Nonequilibrium Thermodynamics》。該論文首次提出了用擴散過程來逆向生成數據的想法，為擴散模型奠定了基礎。

2020 年，Jonathan Ho 和 Stefano Ermon 推出了《Denoising Diffusion Probabilistic Models》，首次把去噪引入擴散模型，提出了一套包含一連串條件概率分佈的模型設計，讓生成過程既精確又靈活。

接著，《Improved Denoising Diffusion Probabilistic Models》對這套模型進行了優化，包括模型架構調整、訓練過程中的梯度計算和優化方法，以及更精細的噪聲控制和損失函數設計等，讓模型不只在圖像生成上大放異彩，在文本、音頻等其他領域也展現了潛力。

擴散模型的應用

擴散模型已經在多個領域展現了它強大的生成能力。在圖像生成上，像是著名的 Stable Diffusion、OpenAI 的 DALL·E 都是靠擴散模型技術來創造高品質圖像，在藝術創作、遊戲開發和媒體產業等領域大展身手。而在自然語言處理上，它也能用來生成文本，甚至在音頻和影片合成上也有一席之地。

另外，擴散模型在藥物設計領域也有應用潛力，可用來預測分子結構，幫忙加速新藥開發。

擴散模型憑其獨特而強大的生成能力，在機器學習領域引起了廣泛的關注。從最初的理論探索到現在實際應用不斷完善，我們引頸企盼更多創新和應用的出現。