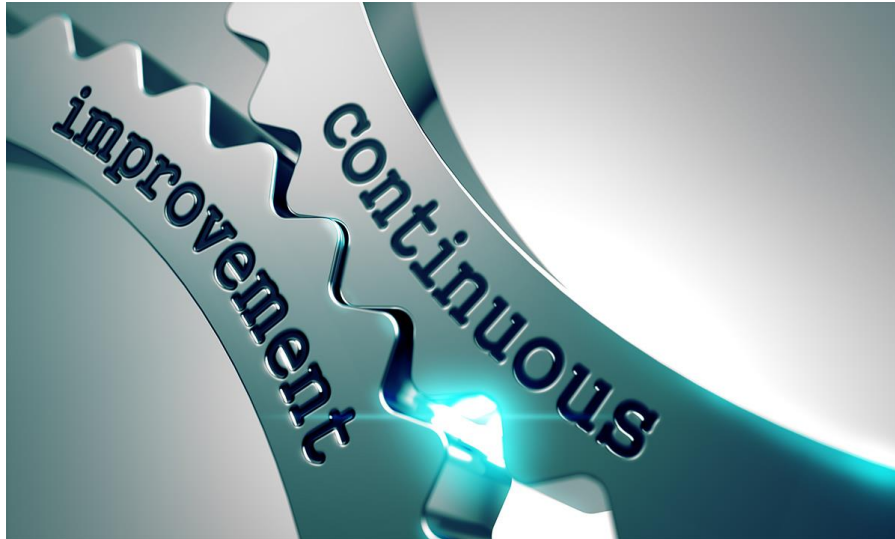


標題：AI 的昔與今

關鍵字：類神經系統、歐拉方法、微積分、ODE net

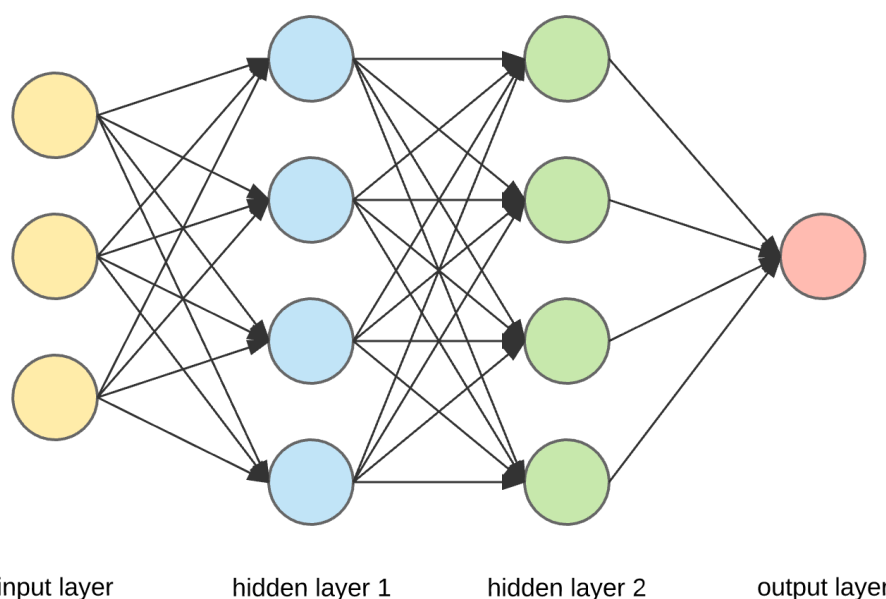


一生當中生病難免，而病歷可說是個人的健康資料庫。不過，生病的時間無法掌握，有時候我們一天要看兩次醫生，又有時候兩年看一次，使得病歷參差不齊，導致在選取不同時間段進行健康評估時，會產生截然不同的結果。這類的連續型問題，讓擅長處理離散型資料的傳統類神經系統難以招架，若病歷分布不夠平均，結果和預期可能相差千里，無法如實反應出個人的健康狀況，實用價值有待商榷。

類神經系統

如果說深度學習是 AI 軍團中的四星上將，那麼類神經網路就是上等兵，透過模擬人類腦神經系統的運作，自行析出輸入資料的特徵後進行階段式學習，再依照邏輯逐步解題，為 AI 軍團進攻人類社會相當重要的一項基礎。

通常來說，神經網路的訓練就是從解謎的過程中找出規則，利用經驗面對未來的種種未知。舉圖像辨識為例，我們可以饋送一堆狗的照片進入神經系統進行訓練，讓神經元從中抽出共同的特徵，進而成為「圖片中是否有狗」的判斷依據。然而，這世界並不如理想中的簡單，單一標準其實不夠嚴謹，極有可能導致過高的失誤率。例如，我們不能單憑「耷拉的大耳朵」分類出含有狗的照片，因為例外不在少數，像是吉娃娃的耳朵是直挺的，而大象的耳朵卻是下垂的等等。



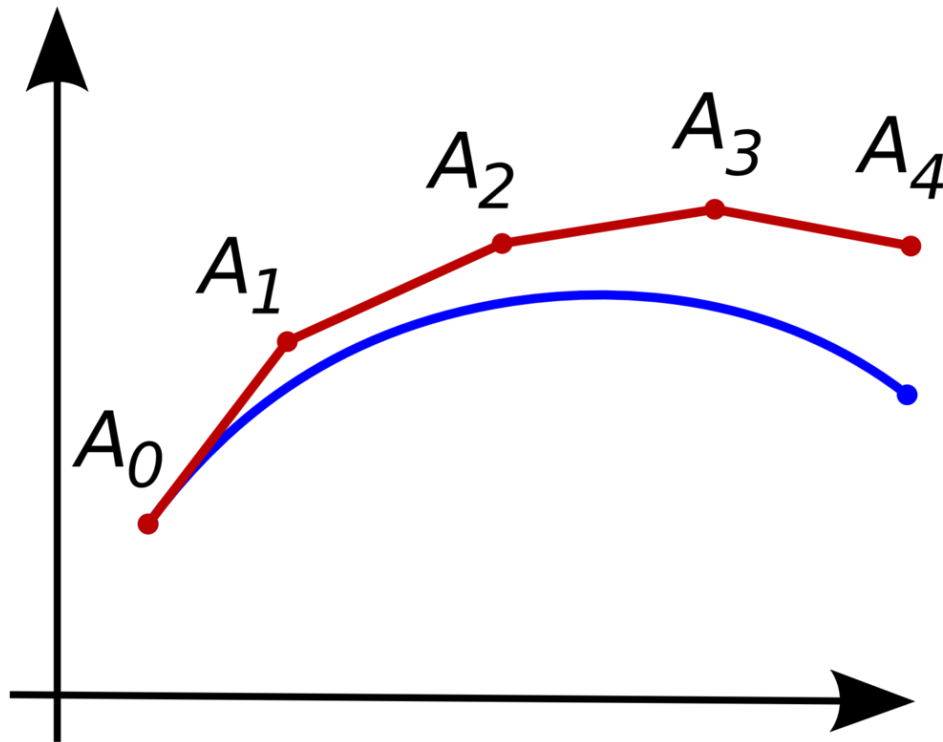
（圖一、傳統神經系統）

為此，「階層」的概念萌出新芽，用以修正太過粗糙的判斷策略。如圖一所示，整體而言有分成三類：負責接收圖片的輸入層(input layer)、進行一連串分析的隱藏層(hidden layer)、發表判斷結果的輸出層(output layer)。當我們在執行受訓過的網路時，第一階層(hidden layer 1)可能會接收照片中的所有畫素，並挑出哪些相對符合狗的模樣；接著第二階層(hidden layer 2)可能藉由別的像素建出一個更大的圖樣，同時判斷這張圖片是否含有鬍鬚或耳朵；依此類推，往後的每一階層會逐步推敲出圖上生物的各種特徵，計算也越來越精細複雜。而最後的結果取決於不同階層、不同面向的判斷，相比於先前的單一標準，處理過程明顯細膩不少。

當然，若各層神經元的連結方式有所差異，執行效率和應用方面也會不盡相同，造就了現今各式的神經網路，包括簡單的前饋神經系統(FNN)、深度神經系統(DNN)、卷積神經系統(CNN)和遞歸神經系統(RNN)等等，都受到了相當程度的重視和研究。

ODE net

然而，當我們回到病歷問題，「階層」的「離散性」卻成了一大阻礙。由於輸入資料和所求都是隨時間連續變動的，卻必須被切成諸多時間段方能訓練或分析，誤差因而產生了。舉例來說，假設我們以年為單位分類資料，就會顯得「每個月看一次病」和「一年中單在某個月看病十二次」毫無差別，因而導致健康評估的結果和不如預期。



（圖二、歐拉方法）

換個角度思考，連續型的判斷可被製成一條連續曲線，而資料的離散化和數值分析的歐拉方法其實有著異曲同工之妙。先來說說歐拉方法，如同圖二所示，當我們想要還原藍色的連續曲線（所求），但僅有起始狀態 A_0 和微分方程式時，我們可以先找出 A_0 上的切線，並選一個固定間距定出 A_1 ，再重複前述的動作直到終點。其中，若節點越多（間距越小），誤差值就會越小，越貼近預期的結果。

事實上，歐拉方法中的每個步驟就好比神經系統中不同階層的工作，某相鄰的兩個節點可被視為某一階層的輸入和輸出。同理，在處理連續型問題時，若神經網路的隱藏層越厚，或將時間段切得越細，所得到的誤差就會越小，結果的價值也會越高—但這也表示若要達到百分之百精確，我們需要無限多的隱藏階層。

為此，Duvenaud 和他的夥伴重新設計原先的神經系統，捨棄「階層」的概念以拋開離散化的麻煩，從而選擇主打極限和無窮問題的利器—微積分。本質上來說，微積分就是一門研究變化的學問，因此若我們將輸入的資料性質轉為常微積分方程，再尋找出比歐拉方法更有效、更精準的工具估算未知的方程式，便可以提升人工智慧的效能。

這類的設計雖然已經脫離網路的概念，但基於傳統考量，研究人員還是將它取名為「常微分方程網路（ODE net）」。

比較異同

Duvenaud 作了一個比喻：一把小提琴可以拉出任何頻率的聲音，因此能夠很輕易地演奏滑音，可是鋼琴卻無法，儘管我們很盡力模擬出滑音的效果，音符之間的顆粒還是相當明顯。從神經網路轉為 ODE net 就像換鋼琴成小提琴，雖說不是所有的場合都適用 ODE net，但就連續型問題而言，ODE net 比神經系統更具有前瞻性。

另外，訓練 ODE net 的方式也有些差別。當我們在訓練傳統神經網路時，可以在開始之前指定階層的數目，並在結束後得知準確度為多少；而在訓練 ODE net 之前，我們能指定的卻是模型的準確度，並依此找出最有效率的方式進行訓練，但代價是我們無法曉得訓練時程的實際長度，這或許會是研究人員將來在選擇方法時需要慎重考量和斟酌的部分。

Duvenaud 說道：「現在還不算是這項研究的黃金時代。」和其他初步研發的技術一樣，ODE net 目前只有基礎的骨架，仍需要多些研究和進步替它添上血肉。然而，它仍是一個備受關注的潛力股，也許會在未來帶領「時序模型（time-series model）」大放光明，例如保健系統等等。

參考資料：

1. Karen Hao, "[A radical new neural network design could overcome big challenges in AI](#)", MIT Technology Review, 12 Dec 2018
2. Adrian Colyer, "[Neural Ordinary Differential Equations](#)", The Morning Paper, 9 Jan 2019