

淺談人工智能與大數據

近年來人工智能技術突飛猛進，在許多行業中開始使用，並改變其運作方式，如網頁搜索、電子商務、金融、物流等。許多媒體紛紛預測在未來一段時間，許多職業將被機械人或人工智能所替換。如電話營銷人員、股票交易員及操盤手等。人工智能擅長和局限在哪？它與最近特別流行的深度學習、大數據有何關聯？又應如何使用及發展呢？

- 人工智能非萬能：目前人工智能完成最好的任務是輸入數據 (x)，快速產生簡單回應 (y)。這種技術在機器學習領域稱為「監督學習」，由訓練和測試/應用兩個階段組成。訓練時，電腦讀入大量的 (x, y) 樣本，旨在訓練出 x 和 y 的映射關係。如人臉識別，「輸入」 x 是大量的不同人臉，「響應/輸出」 y 是對應每個人的身分。在實際應用時，電腦會根據已經算好的 $x \rightarrow y$ 映射關係，快速做出回應。如輸入人臉圖片，電腦立即給出該人的身分。顯然地， $x \rightarrow y$ 這種學習模式跟我們期望——機器人有情感，可以體會我們的感受和服務我們，還差得很遠。因此目前的人工智能技術還有很大的局限。

- 深度學習是核心技術：深度學習技術或深度神經網絡技術，最早是受大腦學習過程的啟發，通過模仿生物神經網絡的結構和功能達到表徵原數據的數學模型。由於深度學習技術可以有效表徵數據特徵，目前在許多領域中受廣泛使用。

- 深度學習需要巨量數據：訓練 $x \rightarrow y$ 對應關係需要大量的訓練樣本。比如人臉識別中需要大量的人臉圖像和對應人的身分。

- 成功的應用都具有 $x \rightarrow y$ 關係模式，如語音識別、廣告推送等。因此，若想快速證明人工智能的實用，就需要找到此類應用。該應用契合 $x \rightarrow y$ 模式，並可以很容易獲得大量具有 (x, y) 關係模式的數據。

人工智能跟我們期望達到的智能還有很大距離。然而，許多科學家和工程師正孜孜不倦地為此努力，相信在不久的將來，有更多實用技術出現在日常生活

中，並進一步改變社會。為了適應這個改變，社會需要大量具有 STEM 背景，並且能將技術和實際應用整合的複合型人才。

恒生管理學院電子計算系助理教授楊海欽博士