



資訊世界裡的穿越劇

利用轉移學習，讓電腦更聰明！

「穿越時空」一直是電影與小說非常吸引人的題材。常見的劇情是主角回到過去，運用現代科技完成一些當時看來不可能達成的任務。例如漫畫《仁醫》描述一位醫生回到了150年前的日本，用21世紀的醫學技術拯救病患。這一類劇情不一定是時間上的穿越，也可以是空間上的，例如電影「星際奇兵」(Stargate)帶出的系列科幻影集，描述人類藉由星門造訪各行星、研究外星文明與科技；或是探險家闖入尚未現代化的部落，利用較先進的知識幫助族群發展。這些穿越劇著墨之處，在於不同文化與技術交會產生的火花。

在資訊科學裡也有學者提出「轉移學習」(transfer learning)的方法，讓人工智慧系統在不同背景中穿越、學習不同知識，進而利用其他領域的知識增廣對於某個領域的理解。轉移學習是資料探勘與機器學習的旁支，主要用於解決資料探勘中某個領域資料不足的問題。

近年來，資料探勘技術已經廣泛使用在現實生活中，「推薦系統」就是典型的例子：系統利用使用者過去的購買記錄，結合其他資料，來推薦使用者可能會喜歡的商品。又例如醫學資訊上的應用，我們可以藉由提供癌症部位的X光照片，讓電腦自動學習特徵，進而自動判讀。要建構上述資料探勘系統，前提是需要有大量的資料，例如：要建構好的電影推薦系統，我們必須告知電腦大部份使用者過去對電影的評分資料；要建構自動判斷乳癌的系統，我們必須輸入某些患者的胸腔X光片，以及另一群正常人的胸腔X光片，讓電腦從中學習。

這樣的機器學習效果常受限於資料品質，當資料不足時，系統準確度容易受資料中的誤差或雜訊影響。以推薦系統的建構為例，假設現在我們成立了網路書店，精確的推薦系統在營運初期是相當重要的，這樣才能在短時間內吸引顧客購買他們有興趣的書籍。但因為書店才剛開張，

還沒有蒐集到足夠的使用者喜好資料，難以建構精準的推薦系統。要解決這樣的困局，我們可以藉由使用者在其他商場（如電影、音樂）的選購記錄，利用轉移學習的方法，把電影音樂的喜好資料用來建構書店的推薦系統。

同理，建立醫學影像分類的系統需要很多影像資料，蒐集這些資料並不容易，也牽涉到個人資料的問題。轉移學習的技術能把已取得的資料轉移來建立其他分類模型，例如利用亞洲人乳癌資料來建立適合歐洲人的乳癌分類系統，或是用乳癌的資料來建構判讀其他癌症的系統。

轉移學習的方法大致分成三類，最直覺的一類方法是「資料轉移」，目的在於把不同領域的資料利用「權重轉移」的方式加入目前學習的領域。例如，要建構適用於歐洲人的醫學影像判讀系統，

在系統學習的過程中，我們可利用其他人種的資料，但是美洲人的權重可能會高於亞洲人，因為歐洲與美洲人的相似度較歐亞兩洲人來得高。第二類方法稱為「特徵轉移」，我們可以把兩個不同領域的資料之間一些重要特徵做出連結，進而互相使用，例如把電影的類別特徵轉移到書籍系統中（喜歡看抒情電影的人，或許也喜歡看愛情小說）。第三類方法叫做「模型參數轉移」，我們可以先在某個資料較多的領域建構出一個模型，然後再把這個模型的參數轉移到另外一個領域中。例如：我們可以先建構可分辨乳癌影像的系統，然後在建立其他癌症判讀的系統時轉移乳癌影像系統的一些參數當成起始設定。

穿越劇之所以吸引人，來自不同價值、背景與技術之間互相激盪所發出的衝擊與能量，於是可以師法古今、截長補短。轉移學習的技巧，讓電腦也可以遊走於不同類型的資料中，來增廣自己的見聞，進而建構出更精確、更符合需求的系統。

林守德是台灣大學資訊工程系副教授。

SA