

【解牛集】— 刊於〈信報〉，2018年6月12日

AI 具解釋力演算法商機無窮

楊毅

科大商學院資訊、商業統計及營運系助理教授

歐盟《一般資料保護條例》(General Data Protection Regulation • GDPR) 於 5 月 25 日正式生效。這條法例，被視為有史以來最嚴格的個人資料保護法。別以為遠在歐洲外的香港不受影響，由於網絡無遠弗屆，不管你的公司是否位於歐盟，只要你公司的客戶或用戶中，會接觸到歐盟公民並擁有其個人資料，都受到 GDPR 的規制。譬如，網絡零售、金融、航空運輸公司恐怕便難以避免受到影響。若一不小心違規，歐盟將施以巨額罰款。這項立法，亦為人工智慧 (AI) 領域方向性發展帶來不容忽視的影響。

筆者不打算在本文詳述 GDPR 的規制，但法例當中有一項「權利解釋權」(Right to Explanation)，要求由人工智慧所制定決策的演算法，必須能夠證明決策的正確性。換句話說，人工智慧所作決策的演算法，要求具有解釋性。筆者試用 AlphaGo 下棋為例，簡單解釋一下。AlphaGo 透過極其複雜的數據分析演算法，決策下這一步棋，這步棋是否下得精確當然十分重要，但下這一步棋的演算法，還要求能夠作出解釋，何以下這一步？亦即決策要說得出推斷的理由。

演算法具解釋力大勢所趨

可以看到，由人工智慧所生出的決策，要求演算法具有解釋能力 (AI Interpretability)，或稱「機器學習解釋」(Machine Learning Explanation • ML) 能力，是目前人工智慧領域最前沿的研究，其發展趨勢和商業應用前景非常值得關注。

筆者以信用卡公司審核申請為例。究竟該不該批准 A 君的信用卡申請？銀行通過數據，包括個人資料，如年齡、學歷、入息、有沒有破產紀錄等等，透過演算法，決定不予批准。但為什麼拒絕批出？所持的理由是什麼？，AI 或 ML 演算法的模型，要求能夠給出解釋性理由；而顧客也有權知道遭拒絕的因由。這是歐盟《一般資料保護條例》制定有關「權利解釋權」的立法精神；也是在商業應用時，企業愈來愈需要重視客戶知情權的大趨勢。

看深一層，無論 AI 或 ML 演算法，其模型設計的目的，一、是作出預測，或二、

發現資料變數之間的潛在關聯特徵。隨著 AI 的技術發展與廣泛應用，演算法也愈來愈複雜，有如一個黑箱，很多演算法的預測精確能力提高，卻缺乏解釋性，令人對此龐然複雜的演算法，如墮入五裡迷霧之中，這種發展方向的可持續性確實並不理想。因此，演算法能夠對決策結果作出解釋，幫助人類對相關事件/商業行為/消費習慣.....等的演算法結果作出分析，是當前 AI 一個重要的前沿發展方向。

簡單來說，解釋性是透過分析所輸入相關變量與預測結果之間的關聯，對於所得出的預測結果，解釋究竟是哪些變量所起的作用。也可以這樣說，演算法給出的預測結果，是根據什麼而作出。

購物籃子的分析啟示

以購物籃分析 (Market Basket Analysis) 為例，電子商務巨擘亞瑪遜的銷售推薦系統值得一提，例如，當你網購一把刮鬍刀時，系統同時向你推薦 B 商品；而系統所推薦的商品，對準潛在客戶的精確性很高，而且為相應推薦提供瞭解釋——因為系統清楚告訴你，購買了這刮鬍刀的其他顧客，同時也買了 B 商品，兩者是配搭一起購買的。有研究發現，這種帶解釋性的推薦，有效地提高顧客的購買意欲。對銷售商來說，無疑有助提高營銷額。

在機器學習中，有一種稱為「監督式學習」(Supervised learning) 方法。簡單來說，是電腦程式通過一些訓練資料(如一些歷史數據)中，學到或建立一個模型，據此模型推測新的實例作出相應的預測。譬如上文提到的，銀行發不發給信用卡予這個人？保險公司是否接受某人的投保？這都屬於「監督式學習」的範疇。

在「監督式學習」中，最簡單的模型是決策樹模型 (Decision Tree Model)。該模型模仿人們日常所做決策的思路，按照一條條規則進行相應預測。用生活的例子來說明，上文提到的信用卡申請，銀行根據一些數據，包括個人資料，如單身、沒擁有物業、居住在城市的低收入地區、曾有借貸違約紀錄等，於是模型作出推斷——「拒絕申請」。決策樹模型具有很好的解釋性，但在模型精度上面表現欠佳。最近興起的神經網路深度學習(Deep Learning)，集成學習(Ensemble Learning) 可以顯著的提高決策樹模型的準確性，但由於模型過於複雜，很難具有解釋性。這樣的模型往往只能給出預測結果，不能給出預測理由。雖然模型有很高的預測能力，但如果演算法的結果缺乏解釋性，始終沒有決策說服力。對公司來說，一個 AI 演算法模型，最好能夠具高解釋性和高推測性。因此，目前 AI 或 ML 演算法，如何在不犧牲高推測性的前提下提高解釋性，是專家正在研究攻關的重點領域。

廣告投放有的放矢

誠然，大多研究都以追求更高的準確度作為演算法優良的評價標準。但是如果優秀的演算法具有解釋性，則可以更好的協助人們理解模型所推論出結果背後的原因。對公司來說，從中可以作出更好、更精確的決策。事實上，具解釋性的演算法，不僅可以讓公司找到決策錯誤的原因；也可以發現模型設計偏差之處進行補救。

再以購物籃分析為例來說明，譬如，在一個網上商貿購物平臺，究竟有那些商品經常同時在顧客的購物籃子裡出現，如何對顧客進行精準推薦？具有解釋性的演算法模型可以告訴網購商組合購物推薦的原因，幫助商家比較不同的行銷組合，發現之前忽略的因素甚至潛在的商機，於是廣告的投放，便更具針對性，吸引更多顧客購買消費。

刮鬍刀、啤酒、戶外帳幕、手電筒

在此，不妨看一個購物籃子的商品組合例子，「刮鬍刀、啤酒、戶外帳幕、手電筒」。這四個看似風馬牛不相及的商品，常常在購物籃子同時出現；再配合分析客戶的個人資料，AI 或 ML 的演算法便作出推斷解釋，喜歡喝啤酒的年青男子，多喜歡往郊外紮營旅行。於是，不僅把這四種商品放在一起，方便消費者採購，而且更把一些喜歡到郊外紮營旅行年青人可能會買的東西，也拿來放在一起試行推售，結果很有可能增加商品的銷售。由此可見，AI 或 ML 的演算法具解釋性，在銷售應用上，將是一個極具潛在增益的營銷方法。

很顯然，演算法具解釋性的應用領域十分廣泛。除了信用卡公司、保險公司、金融機構、零售商外，在美國，有一個研究相當新穎，就是利用 AI 去協助法官預測犯人保釋後，會否再犯罪。傳統上於決定批准保釋與否，法官需要閱覽很多資料，而且也不排除法官帶有感情。事實上，若然讓一個在保釋期間，會再犯罪的人保釋外出；或把一個不會重犯的人囚禁在監獄裡，都不是一個理性而有效的決策。於是便可以用 AI 構建預測模型來協助法官進行保釋的判斷。AI 模型透過輸入相關資料變量，例如，此人是否有吸毒習慣、犯罪是什麼類型，是搶劫抑或貪汙等等，分析變量與預測量的關係，若預測結果發現，此人重犯的機率為 0%至 5%低水準，而且通過輸入變量與輸出的分析，解釋到低重犯機率的原因，於是演算法給出的結論是：給犯人保釋。當演算法結果亦具解釋性，便可協助法官作出有效判斷，最終提高整個司法系統的效率。

AI 強化人類認知決策能力

值得一提，人工智慧或機器學習，並非要取代人類的智慧，而是去協助人類強化決策的認知能力和理解。目前，AI 或 ML 演算法的應用，也跨入了醫療領域。

在醫療上，使用 AI 技術去協助醫生作出診治判斷，相信會普遍起來。譬如，在一張腦掃描圖上，尋找出內裡是否有腫瘤細胞。AI 演算法方案並非去取代醫生或凌駕醫生的判斷，而是透過 AI 演算法，讓醫護人員從輸入因素中，如體溫、血壓等等，發現那幾個因素組合，更容易導致出現腫瘤病，並展示出判斷的理由，從而提高醫生的檢查成效，甚至更有效地找出預防和治療的方案。

可以說，AI 或 ML 的演算法具解釋性，發展和商業應用潛力巨大，商機無窮。很顯然，傳統通過大量數據，找到一些模態（Pattern）特徵，據此進行預測。譬如預測天氣變化、股票價格走勢、發掘潛在客戶等，在這個過程中，預測的精確性當然十分重要。但人工智慧技術發展至今日，不僅提高演算法預測的精準度，更要演算法具解釋力。筆者相信，AI 演算法預測能力與解釋能力並重，勢將成為人工智慧領域的發展主流，並為商業應用帶來更龐大的潛在收益！

〔本文由科大商學院傳訊部筆錄，楊毅教授口述及整理定稿〕