

「知識蒸餾」新型態 AI 教育模式

謝佩璇

國立政治大學資訊科學系合聘教育系副教授

林國祥

國立勤益科技大學資訊工程系教授

一、前言

「人工智慧」（Artificial Intelligence，AI）或「機器學習」（Machine Learning，ML）相關主題內容，涵蓋在課程設計或課堂活動，不僅吸引資訊領域學習者的目光，也引起人文社會相關領域學習者的關注，並且開始思考學會撰寫程式的必要性，以及應該從哪個程式語法開始學習等問題。對於非資訊相關領域的開課教師，也開始思考是否跟進開始學習撰寫程式，或者盡快獲得 AI 相關技術以融合原教學內容與進程，將此技術充分應用在課堂教學中。教育行政機關、各層級學校，從上到下提供經費，積極擴充軟硬體設備與引進業界教師到課堂，讓 AI 技術應用範疇落實到各層級與各個學習領域的教育現場。教師能善用 AI 技術開發的各種應用軟體於教學中，有助於提升學生的學習動機與興趣；學生嘗試採用各種輔助學習的軟體之際，也增加學習程式撰寫的機會。家長們也因此趨之若鶩地投資，讓孩子獲得更多資訊相關技能的培訓機會。面對已經到來的 AI 時代，個人追求新技術儼然是勢在必行，新型態 AI 教育模式也因此值得探索。

二、新型態教育模式

（一）個人化學習平台

市面上已有出現各式各樣軟體與學習平台，幫助不同年齡層、不同學習階段的學習者，有效獲得資訊相關知識與技能，包括善用 AI 工具，此學習風潮成為顯學且獲得 AI 技術的門檻逐漸下降（高立芸、王俊斌，2023）。有些學習平台以會員制登入，可獲得連貫且系統性的學習計畫，手持裝置的應用程式也提供各種程式語法的講解與練習題，例如：國內的均一平台、Hahow 好學校、e 等公務園+學習平台等；國外的 Coursera、edX、Khan Academy、Udemy 等。此外，國小學生可用方塊模組來認識程式語法（賴阿福，2020），國高中學生可套用 Python 模組來訓練機器模型（黃思華、張玟慧，2020；湯維玲，2020）。還有各種功能的機器人提供學習資源或照護陪伴，例如：可與人形機器人 Pepper 對話，完成飯店登入、點餐與送餐，以及獲得參觀景點資訊；年幼或年長者有機器寵物陪伴（黃思華、張玟慧，2020）。對正接受 K-12 學校教育以及正接受高等教育而尚未進入就業市場的學習者而言，懂得採用不同 AI 技術於不同情境，甚至具備程式撰寫能力，似乎讓個人未來擁有較高的就業競爭力，甚至持續保持從業競爭優勢。

（二）自主式輔助學習

從 2019 年開始教育部為培養學生自主行動、溝通互動兩個面向的核心素養，鼓勵 K-12 各級學校採用數位學習平台，以增進教與學的品質；2021 年教育部更推出「中小學數位學習精進方案」期許四年內達到「班班有網路、生生用平板」的目標（教育部，2021）。今日，AI 技術急速成長，經過訓練所開發的各種模型，應用在不同科目的學習情境中，從而讓自主學習過程更加順暢（王俊傑，2024；洪永洲，2024；楊子奇，2024），教師教學輔助策略有賴完善的指標建立（高台茜、張維珊，2023）。尤其當學習者採用 AI 電腦輔助自主式學習，能快速且直接獲得運算最佳化結果，如同使用 ChatGPT（Generative Pre-trained Transformers，生成式預訓練轉換器）與機器聊天卻不用登入 OpenAI 帳號，可執行各種指令，例如：文字編修、繪圖著色、情緒諮商等，彙整各種來源資料提供給自主學習者，再根據學習者的反饋或多次提問而最終獲得精準的解答，此過程亦展現自主學習精神（潘乃欣，2024）。每個軟硬體背後皆有 AI 輔助自主學習，對於教師在課堂原本傳道、授業、解惑的角色會如何轉變值得探討。

三、知識蒸餾的教學代理模式

（一）知識蒸餾的模式

「知識蒸餾」（Knowledge Distillation，KD）的名詞剛開始被提出來的時間點是在 2006 年一場國際研討會（Bucilua et al., 2006）。KD 技術發展來自於 AI 神經網路（Neural Network）逐年發展而來，由於機器學習過程中需要使用大量資料來訓練與建置模型，採用 KD 技術則能節省機器學習過程需要耗費的學習資源，包括：資料量與訓練時間，也就是能有效壓縮訓練模型而仍能達到可接受的學習效果（Freitag et al., 2017；Furlanello et al., 2018；Hinton et al., 2015）。然而，當時的科技設備還有很大的發展空間，雖然有許多研究者持續探索此技術開發，卻仍未受到高度重視。此外，並非每個研究者都有優渥的資源，即便有資源不僅耗時耗力，更無法達到即時更新知識、重新快速訓練資料與修正模型的機會（Cui et al., 2017）。直到近幾年，硬體設備的記憶容量增加、大數據資料處理與模型訓練的運算速度加快，加上有越來越多研究者陸續提出能降低運算誤差的演算法，並且公開分享已訓練好的範例模型。因此，直至 Hinton 等（2014、2015）具體說明知識蒸餾技術在機器學習模型開發過程中的重要地位之後，越來越受到來自不同研究領域學者的重視。也有越來越多學者嘗試將重量級模型（例如：依排名 GPT、Claude、Gemini、Llama、Transformer、BERT）的中間層以 KD 技術達到輕量化模型後，用於評估各種任務或多任務的模型表現（Sepahvand et al., 2022）。如今，KD 技術已經廣泛應用在不同的任務或知識領域，例如：物件偵測、影像辨識、語音分析、即時翻譯等（Wang & Yoon, 2022）。

（二）知識蒸餾的教學代理

由於 KD 技術可有效壓縮重量級模型，如同「學生-教師」(Student-Teacher, S-T) 架構的師生關係，教師擁有的龐大知識能快速移轉 (Transfer) 給學生，而建置完成已壓縮的模型測試結果，仍能達到可接受或超出預期的學習成效 (Wang & Yoon, 2022)。在實際教與學的歷程，KD 可比喻為教師的傳道、授業、解惑的三個階段，如圖 1 所示。

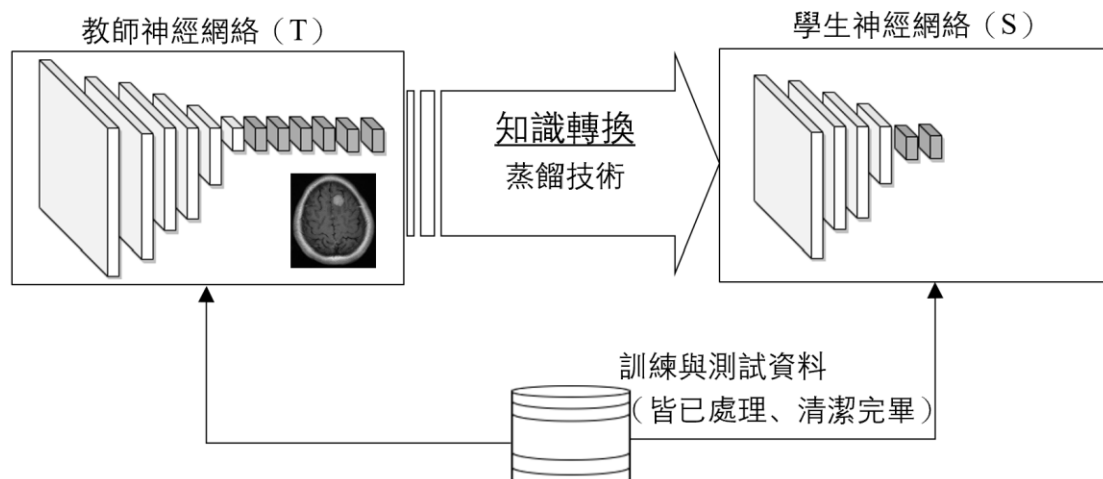


圖 1 如同師生關係的知識蒸餾技術

傳道：擁有豐富知識體的教師 (T) 極盡所能地講授知識，能在教學現場將重要的概念闡述出來，強調重要知識內涵、理論概念，這過程可說是傳道的過程。「道」在此是指關鍵的知識內涵或理論概念，可視為機器學習過程的初始階段，也就是正在訓練資料集的過程，包括參數設定與特徵擷取，可視為教師傳道過程最重要且初始的行動。

授業：在教學的現場中，因材施教才能真正達到有教無類的理想境界，因為有些學生 (S) 可充分吸收教師的知識，有些學生的準備度或理解能力不足，對抽象知識仍不清楚，此時教師需適時調整授課進度或改用基礎學習教材，先講授基礎概念，並且盡量以淺顯易懂的方式來授課，才能最終讓學生獲得教師擁有的所有知識。這些多元的授業方式，就像不同訓練資料集的方式：監督式 (supervised)、非監督式 (unsupervised)、半監督式 (semi-supervised) 或增強式學習 (reinforcement learning)。不論教師決定採用哪種教學策略，實現 KD 技術的過程，教師模型最重要任務就是訓練資料集，才能順利完成學生模型的建置。

解惑：教師傳道與授業過程中有些學生會提問，如同初步完成訓練資料集所建立的模型，可評估學生初步的學習表現或學習成效。成效不佳時，則採用不同 KD 技術進行知識移轉，或者採用不同演算法來計算函式。函式的許多參數 (parameters) 計算與使用，或者重要特徵值的「提取」(extraction 或 abstraction)，

根據不同任務的特性而有所不同，如同教師傳道與授業過程，不斷評估學生學習成效，最終目標是教師的知識移轉到學生。

知識移轉的過程，學生能充分掌握教師擁有的知識體，培養出自主學習能力並且擔任「教學代理」的角色（或稱之為助教或實習教師），肩負起傳道、授業、解惑的角色，如同機器學習過程所建置的極佳模型表現。對於評估教學代理成效的方式，則可依據上述文獻提及的教學策略指標來進行探討。

四、結語

新型態 AI 教育模式的終極目的是達到自主學習，學生有效獲得教師的知識。AI 時代的教師仍然有傳道、授業、解惑的責任，轉型之處是在教學場域多元化與擔任引導角色。從機器的角度來說，學習需要耗費許多運算資源，從資料蒐集、清理與整理，再進行資料前處理、訓練資料的模型建置、測試資料評估模型表現等。如同知識蒸餾的各種技術，提供各種能讓學生充分獲得教師知識的各種學習資源。期待學生學習成效佳，能獲得教師的智慧或甚至青出於藍、更勝於藍，能正確完成分類、偵測、診斷、辨識等各項複雜或多重的任務。

知識移轉的過程涵蓋許多技術細節，所有機器學習的演算法背後，可用師生或人與人之間的互動經驗來比喻，例如：有些厲害的教師無法用簡單語言來教導初學者，或者學生表現已經超越教師；如此描述正好是「知識蒸餾」的技術發展過程中所發生的真實狀況。還有其他人與人之間出乎意外的互動，例如：學生想從習得教師的知識領域，跨到學習另外一個知識領域，此狀況確實在「知識蒸餾」的技術發展中實現。然而，師生互動的情誼在教學現場，是無法被機器取代的。人類與機器的對話機制僅藉由一條一條的數學公式轉換為演算法，得到一行一行的程式語法。所有語法其實來自於智慧的人類，持續地堆疊彼此的智慧並且充分推論後所獲得。人類的智慧來自於生活經驗、不斷錯誤嘗試後，讓機器能正確且有效地依邏輯指令訓練各種資料，最終幫助人類獲得更多啟發。

參考文獻

- 王俊傑（2024）。師資生應具備之數位教學素養。臺灣教育評論月刊，13，43-48。
- 洪永洲（2024）。淺談校園數位內容教材與教學軟體應用在提升教師教學與學生學習的趨勢。臺灣教育評論月刊，13(5)，70-78。
- 高台茜、張維珊（2023）。促進科技輔助自主學習之教師教學輔助策略指標

建構之研究。《課程與教學季刊》，26，01-30。

- 高立芸、王俊斌（2023）。當AI遇到師資培育。《教育研究月刊》，355，53-67。
- 教育部（2019）。數位校園・智慧學習科技輔助自主學習。取自 <https://srl.ntue.edu.tw/>
- 教育部（2021）。推動中小學數位學習精進方案入口網。取自 <https://pads.moe.edu.tw/>
- 湯維玲（2020）。運算思維與程式設計教育之研究：以程式設計發展師資生運算思維為例。載於張芬芬、詹寶菁（主編），**AI時代的課程與教學：前瞻未來教育**（81-106頁）。臺北市：五南圖書。
- 黃思華、張玟慧（2020）。AI時代的機器人教育理論與實務。載於張芬芬、詹寶菁（主編），**AI時代的課程與教學：前瞻未來教育**（3-30頁）。臺北市：五南圖書。
- 楊子奇（2024）。科技輔助自主學習時代，高中職教師的信念與實踐。《臺灣教育評論月刊》，13(5)，49-52。
- 潘乃欣（2024）。AI能帶中學生自主學習、幫老師改作文？資深老師解密如何「與AI共舞」。《翻轉教育》。取自 <https://flipedu.parenting.com.tw/article/008925>
- 賴阿福（2020）。人工智慧對於國小資訊教育之影響及因應之道。載於張芬芬、詹寶菁（主編），**AI時代的課程與教學：前瞻未來教育**（227-270頁）。臺北市：五南圖書。
- Bucilua, C., Caruana, R. & Niculescu-Mizil, A. (2006). Model compression. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 535–541, August 20 - 23, 2006, Philadelphia, PA, USA. Retrieved from <https://doi.org/10.1145/1150402.1150464>
- Cui, J., Kingsbury, B., Ramabhadran, B., Saon, G., Sercu, T., Audhkhasi, K., et al. (2017). Knowledge distillation across ensembles of multilingual models for low-resource languages. *Proceedings of 2017 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 4825-4829). Retrieved from <https://doi.org/10.1109/ICASSP.2017.7953073>

- Freitag, M., Al-Onaizan, Y., & Sankaran, B. (2017). Ensemble distillation for neural machine translation. Retrieved from <https://doi.org/10.48550/arXiv.1702.01802>

- Furlanello, T., Lipton, Z., Tschannen, M., Itti, L., & Anandkumar, A. (2018). Born again neural networks. *Proceedings of the 35th International Conference on Machine Learning* (pp. 80:1607-1616). Retrieved from <https://doi.org/10.48550/arXiv.1805.04770>

- Hinton, G., Dean, J., & Vinyals, O. (2014). Distilling the knowledge in a neural network. *NIPS (28th Conference on Neural Information Processing Systems) Deep Learning and Representation Learning Workshop*, December 12, 2014, Montreal, Canada. Retrieved from <https://arxiv.org/abs/1503.02531>

- Hinton, G., Dean, J., & Vinyals, O. (2015). Distilling the knowledge in a neural network. Retrieved from <https://arxiv.org/abs/1503.02531>

- Sepahvand, M., Abdali-Mohammadi, F., & Taherkordi, A. (2022). Teacher–student knowledge distillation based on decomposed deep feature representation for intelligent mobile applications. *Expert Systems with Applications*, 202, 117474. Retrieved from <https://doi.org/10.1016/j.eswa.2022.117474>

- Wang, L. & Yoon, K. J. (2022). Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(6), 3048-3068. Retrieved from <https://doi.org/10.1109/TPAMI.2021.3055564>

