ISACA

稽核從業人員之機器學習參考指引

第一部分：科技

目錄

4 **為何人工智慧(AI)/機器學習(ML)應用程式稽核如此重要?**

6 **了解人工智慧、機器學習與深度學習**

6 / 人工智慧、機器學習與深度學習之定義

7 / 機器學習分類

7 / 監督式學習

7 / 非監督式學習

8 / 強化學習

9 **機器學習應用程式之科技風險**

9 / 資料治理

9 / 資料工程

11 / 特徵工程

13 / 模型訓練

14 / 模型評估

14 / 監督式學習模型評估

16 / 模型部署及預測

17 / 受歡迎的機器學習資料集

18 **結論**

19 **附錄：推薦資源**

20 **致謝**

摘要

隨著人工智慧（AI）和機器學習（ML）被全球企業和政府廣泛採用，現有的審計框架和資訊科技控制必須更好地適應AI和ML帶來的獨特且不斷演變的風險。風險源於ML模型的選擇、設計和開發週期。此外，風險也與遵守法規有關，例如歐盟的《一般資料保護規則》（GDPR）或美國的《加州消費者隱私保護法》（CCPA）。

本白皮書系列分為兩部分，提供了一套系統性的稽核指引，將科技風險和合規風險作為稽核人員的兩個主要關注點。第一部分探討了機器學習稽核中的科技風險，透過剖析人工智慧/機器學習演算法的典型軟體開發生命週期，並強調資訊稽核人員應調查的關鍵領域。第二部分探討了合規風險，並為資訊稽核人員提供了可採取的實際步驟，以促進遵守適用的法律、法規和行業標準。

[1] Ahmed, H.; "Auditing Guidelines for Artificial Intelligence," *@ISACA*, 21 December 2020, www.isaca.org/resources/news-and- trends/newsletters/atisaca/2020/volume-26/auditing-guidelines-for-artificial-intelligence

為何AI/ML應用程式稽核如此重要？

在嚴重特殊傳染性肺炎（COVID-19）大流行期間，人工智慧 (AI) 及其子集機器學習 (ML) 的採用率急劇上升，如哈佛商業評論的一篇近期文章和資誠的研究所示[2]。52% 的公司加速了其 AI 採用計劃，將 AI 置於其業務的核心，以提高生產力和效率，並透過新產品和服務來驅動創新。在一項最近的勤業眾信調查中，83% 的受訪公司認為 AI 已經或將會產生實際且可見的影響。隨著 AI 逐漸從新興技術轉變為普遍使用的主流技術，其提供的機會也將伴隨著挑戰。

隨著公司在產品和流程中嵌入更多 AI，不斷增加的演算法增添了不準確或有偏見決策的可能性。特別是這些複雜的模型可能會影響個人的日常生活、健康和權利，例如在申請貸款、投資決策或規劃癌症的治療方案等情形。

在過去十年中，公眾對科技風險的擔憂主要集中在個人資料的擴大蒐集和使用上。但隨著公司在產品和流程中嵌入更多 AI，不斷增加的演算法增添了不準確或有偏見決策的可能性。特別是這些複雜的模型可能會影響個人的日常生活、健康和權利，例如在申請貸款、投資決策或規劃癌症的治療方案等情形。

許多世界各地的政府和非營利組織都認為演算法所產生不準確或有偏見的決策實際上具有潛在重大的風險。因此，已經提出了保護公民的相關法規，並發布了相關指引。例如，歐盟提議的人工智慧法案[3,4]和美國通貨監理局的「模型風險管理」[5]以及環境、社會和治理 (ESG)[6]指引。

企業在採用 AI 和 ML 時需要考慮的主要挑戰可總結如下：

- **資料隱私** — AI 和 ML 模型在大量資料上進行測試和訓練，這些通常是個人資料，受全球法律法規的保護。在處理敏感資料時，使用這些模型中的資料可能會有問題。而面對這些挑戰的解決方案通常以立法形式出現。例如，個人識別資訊 (PII) 和受保護健康資訊 (PHI) 的使用受 1996 年健康保險可攜性和責任法案 (HIPAA) 的管轄，而消費者資料的其他元素則在 GDPR、CCPA 及類似法規中得到解決。
- **公平性** — 公平性與新興的資料倫理領域相關，無論模型是自行做出決策，還是與人類代理協調做出決策，均需要評估基於 AI 的產出對人們權利的影響。波士頓房價資料集是一個經典的資料公平性問題範例，其中包含以一個明確的參數來辨認一個地區

[2] McKendrick, J.; “AI Adoption Skyrocketed Over the Last 18 Months,” *Harvard Business Review*, 27 September 2021, https://hbr.org/2021/09/ai-adoption-skyrocketed-over-the-last-18-months
[3] European Institute of Public Administration (EIPA), “The Artificial Intelligence Act Proposal and its Implications for Member States,” September 2021, www.eipa.eu/publications/briefing/the-artificial-intelligence-act-proposal-and-its-implications-for-member-states/
[4] Daws, R.; “EU regulation sets fines of €20M or up to 4% of turnover for AI misuse,” AI News, 14 April 2021, https://artificialintelligence- news.com/2021/04/14/eu-regulation-fines-20m-up-4-turnover-ai-misuse/
[5] Office of the Comptroller of the Currency (OCC), “Model Risk Management,” August 2021, www.occ.treas.gov/publications-and- resources/publications/comptrollers-handbook/files/model-risk-management/pub-ch-model-risk.pdf
[6] Anand, R.; B. Greenstein; “Six AI business predictions for 2022,” 30 November 2021, PwC AI and Analytics, https://www.pwc.com/us/en/tech-effect/ai- analytics/six-ai-predictions.html

是黑人區還是白人區，以及這些種族人口統計如何影響房價。[7]

- **品牌聲譽和信任**—在過去兩到三年中，有 22% 的企業已經面臨客戶的強烈反彈，原因是 AI 系統做出的決策。客戶所擔憂的一個領域是人臉辨認系統，因為潛在的誤認可能會對敏感情境產生重大影響，例如機場安檢監控或考試監考等。
- **模型透明度**—對於簡單的 ML 模型，如線性迴歸，很容易通知用戶有關系統的邏輯和決策參數。然而，對於最先進的黑盒 ML 模型，如深度學習，目前尚無簡單的方法來解開和權衡隱藏層中的許多參數。近年來，隨著透明度和可解釋性成為 AI 使用的核心原則，如新法律、法規或行業標準所示，這些透明度差異變得尤為重要。

與隱私、公平性、信任、歧視、預測和趨勢相關的挑戰橫跨了科技與合規風險的範疇。例如，模型系統性錯誤可以通過仔細的科技風險稽核來解決。

- **模型系統性錯誤**—機器學習模型可能出現過度擬合（即模型過於複雜，傾向於「記憶」大量資料集中的雜訊，而未能掌握整體趨勢）或擬合不足（即模型過於簡單，無法對複雜資料建模）情形。如果模型沒有以科學的方式進行訓練，可能會發生過度擬合或擬合不足，進而產生系統性誤差，使產品造成錯誤。例如，使用相同的資料進行訓練和測試將導致過度擬合的情況。
- **學習和適應演算法**—雖然學習和適應的演算法可能更準確，但它們可能會以歧視的方式演變。實施持續監控程序並在開發和機器學習運營過程中強制執行是非常關鍵的。

與隱私、公平性、信任、歧視、預測和趨勢相關的挑戰橫跨了科技與合規風險的類別。例如，模型系統性錯誤可以通過仔細的科技風險稽核來解決。為了本白皮書系列的目的，對 AI 和 ML 應用程式稽核應考慮以下內容：

- **科技風險稽核**—評估與機器學習演算法及其生命週期、資料科學和網路安全相關的風險。建議從業人員根據AI和ML應用程式開發週期的六個關鍵階段，實施稽核程序。
  - 資料治理
  - 資料工程
  - 特徵工程
  - 模型訓練
  - 模型測試(評估)
  - 模型部署
- **合規風險稽核**—評估與AI/ML應用程式所影響的個人權利和自由相關的風險，這可能因產業和地區而異。資料合規監管規定的例子包括：
  - **金融業** - 美國通貨監理局「模型風險管理：新檢查手冊」(Model Risk Management: New Comptroller's Handbook Booklet)
  - **醫療保健業** - HIPAA (健康保險可攜帶性和責任法案)
  - **消費者資料隱私** - CCPA (加州消費者隱私保護法)
  - **歐洲地區規範** - GDPR (一般資料保護規則)

本白皮書系列的第一部分涵蓋了機器學習稽核的科技要素。第二部分將討論合規性方面的考量。

[7] Delve, "The Boston Housing Dataset," 10 October 1996, www.cs.toronto.edu/~delve/data/boston/bostonDetail.html
[8] Thieullent, A.-L., et al.; "AI and the Ethical Conundrum: How organizations can build ethically robust AI systems and gain trust," Capgemini Research Institute, 2020, https://www.capgemini.com/wp-content/uploads/2020/10/AI-and-the-Ethical-Conundrum-Report.pdf

了解人工智慧、機器學習與深度學習

本節將解構並釐清與 AI 相關的常見術語。尤其，儘管 AI 是一個廣泛應用的總稱，但機器學習 (ML) 和深度學習 (DL) 是 AI 的子集，對本白皮書而言有著重要的區別。

人工智慧、機器學習與深度學習之定義

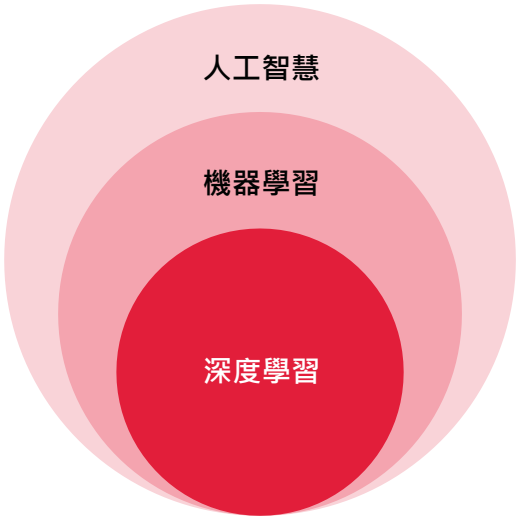
人工智慧(AI)是「電腦系統的理論與發展能夠正常地執行需要人類智慧才能完成的任務。」[9]實際上，AI 由專門從資料中「學習」的演算法所組成。有鑑於此，AI 是一個廣泛的術語，涵蓋從條件式的「如果...否則...」規則引擎到現代應用，如自動駕駛汽車與複雜的電腦程序，如 AlphaGo。

機器學習(ML)是 AI 的一個子類別，是「電腦透過學習新資料，以及在無需人類提供電腦程序指令的情形下改變執行任務方式的過程。」[10] 關鍵詞是「資料」。目前，大多數被歸類為 AI 的應用實際上是 ML 工具，它們利用統計模型「彙總」來自大量資料集的型態，這些型態爾後可用於預測新資料。

深度學習(DL)是一種基於人工神經網路的機器學習類型，其中透過使用多層處理從資料中提取更高層次的特徵。

AI 和 ML 的創新始於數十年前。然而，由於 DL 在電腦視覺、語言翻譯、語音辨識等應用中優於傳統 ML（如邏輯迴歸、支援向量機和隨機森林分類），這些技術最近引起了行業的廣泛關注。AI、ML 和 DL 之間存在層次關係，如**圖 1** 所示。

圖 1: AI、ML與 DL的層次關係



人工智慧(AI)—描述執行與人類智慧相關任務的電腦系統。

機器學習(ML)—描述電腦在執行連續或迭代任務過程中的自我修正過程，其中前一任務的輸出（如失敗、假陽性、假陰性等）作為後續任務的輸入，從而實現無需人類介入的過程修正（例如軟體修改）。

深度學習(DL)—描述一類基於人工神經網路的機器學習類別，其中多個處理層透過逐步提取的方式，從資料中獲得更高層次的特徵。

[9] Lexico, "Artificial Intelligence," www.lexico.com/en/definition/artificial_intelligence
[10] Cambridge Dictionary, "machine learning," https://dictionary.cambridge.org/us/dictionary/english/machine-learning

根據這個層次結構，本白皮書專注於機器學習，因為它是在稽核實務應用時最常見的一類演算法。

機器學習分類

機器學習可分為三大類：監督式學習、非監督式學習和強化學習。

監督式學習

「監督式學習」指的是有關訓練資料包含輸入向量的範例以及其對應目標向量的應用。由於這種學習需要使用標記資料（即範例）來訓練（即教導）電腦，因此這種學習被稱為監督式學習。

「監督式學習」指的是有關訓練資料包含輸入向量的範例以及其對應目標向量的應用。

監督式學習的主要分類包括：

- **分類**—監督式學習的一種，用於預測類別型標籤。例如，一張胸部X光影像是否顯示「有癌症」或「無癌症」？典型的演算法包括邏輯迴歸、支持向量機和隨機森林。
- **迴歸**—監督式學習的一種，用於預測數值型標籤。例如，「這棟位於郵遞區號94301的三臥室房屋價值可能為210萬美元」。典型的演算法包括線性迴歸（linear regression）和支持向量迴歸（support vector regression）。
- **混合**—監督式學習的一種，結合使用標記資料，並能從部分標記資料中進行學習，例如半監督式學習（semi-supervised learning）。

非監督式學習

非監督式學習研究如何讓系統透過未標記的資料中推斷出一個函數，以描述隱藏的結構。系統不預測正確的輸出。相反地，它透過探索資料並從資料集進行推論，以描述未標記資料中的隱藏結構。[12]

隨堂測驗：預測房地產位置

某房地產抵押貸款團隊正在嘗試建立一個機器學習模型，根據房屋面積和其他特徵來預測房產的郵遞區號。例如，一個3000平方英尺的地段搭配三房住宅，預測其位於94301郵遞區號區域；而一個8000平方英尺的地段搭配五房住宅，則預測位於75218郵遞區號區域。

問題：這個機器學習模型屬於分類還是迴歸？[13]

非監督式學習的主要分類包括：

- **分群**—自動辨認資料中的群組，例如根據動物的腿數或大小進行分組。典型演算法：K-means分群。
- **降維**—透過將資料從高維投影到低維來壓縮資料，同時保留某些特徵，例如將城市、省份及郵政編碼的多維資料簡化為地理位置。典型演算法：主成分分析（Principal Component Analysis, PCA）。
- **異常檢測**—辨認顯著偏離正常模式的異常值，例如一天內進行157次購物的信用卡帳戶，可能表明存在詐欺行為。典型演算法：中位數、z分數統計、局部離群因子（Local Outlier Factor, LOF）。
- **關聯規則**—發現資料中的相關性規則，例如購買商品X的人通常也會購買商品Y。

非監督式學習研究如何讓系統透過未標記的資料中推斷出一個函數，以描述隱藏的結構。系統不預測正確的輸出，相反地，它透過探索資料並從資料集進行推論，以描述未標記資料中的隱藏結構。

[11] Bishop, C.; *Pattern Recognition and Machine Learning*, Springer, USA, 2006, https://link.springer.com/book/9780387310732
[12] Loukas, S.; "What is Machine Learning: Supervised, Unsupervised, Semi-Supervised and Reinforcement learning methods," Towards Data Science, 10 June 2020, https://towardsdatascience.com/what-is-machine-learning-a-short-note-on-supervised-unsupervised-semi-supervised-and-aed1573ae9bb
[13] Answer: Classification. Although a ZIP code is composed of numbers, in this case it is a categorical variable to represent a geolocation.

強化學習

強化學習 (Reinforcement Learning)是學習做什麼才能最大化獎勵訊號的一種學習方式—意即如何將情境對應到行動。學習者並未被明確告知應該採取哪些行動，而是必須透過嘗試，自己發現哪些行動能帶來最多的獎勵。[14]

強化學習與監督式學習明顯有所不同。強化學習是透過與環境互動來學習，而非依賴訓練資料集。例如，在一個下棋程式中，強化學習要求代理程式進行數百萬次的對決，以摸索出策略；而監督式學習則會將過去的棋局資料作為訓練材料提供給代理程式。典型的強化學習演算法包括 Q 學習 (Q-learning) 和深度強化學習 (Deep Reinforcement Learning)。其應用範圍涵蓋自駕車、機器人控制以及 Google 的 AlphaGo 程式（擊敗世界上最強的圍棋選手）。

範例：DeepMind 科技團隊的AlphaGo及深度強化學習

AlphaGo 是第一個擊敗職業圍棋選手的電腦程式，也在 2016 年擊敗圍棋世界冠軍。它被認為是歷史上最強的圍棋玩家。圍棋可能的棋盤配置型態數量高達 10^{170}，遠超已知宇宙中原子的數量，複雜度遠高於西洋棋。AlphaGo 結合了先進的搜尋樹 (Search Trees) 與深度神經網路 (Deep Neural Networks)，例如政策網絡 (Policy Networks) 和價值網絡 (Value Networks)。DeepMind 團隊讓 AlphaGo 與眾多業餘玩家對弈，幫助程式理解人類合理的玩法。隨後，AlphaGo 與自己的不同版本進行了數千次對弈，每次都從錯誤中學習。隨著時間推移，AlphaGo 持續改進，在學習和決策能力上越來越強。這個過程被稱為深度強化學習(Deep Reinforcement Learning)[15]。

範例：銀行直接行銷活動

某家銀行機構針對超過 45,000 位客戶進行直效行銷活動。

參與這些行銷活動的客戶涵蓋數十種人口統計資料，使得人工分析變得相當複雜。

人口統計資料涵蓋下列：

- **年齡** - 數值型
- **職業** - 類別型，根據類型分類（例如：行政人員、藍領、創業家、保姆、管理階層、退休人士、自僱人士、服務人員、學生、技術人員、失業者、未知）。
- **婚姻狀況** - 類別型（例如：已婚、單身、未知）。註：單身包括從未結婚、離婚或喪偶的情況。
- **教育程度** - 類別型（例如：基礎教育4年、基礎教育6年、基礎教育9年、高中、文盲、職業課程、大學學位、未知）。
- **信用違約** - 類別型（例如：是否有信用違約？[否 | 是 | 未知]）
- **住房貸款** - 類別型（例如：是否有住房貸款？[否 | 是 | 未知]）

資料科學家使用一種快速的K眾數(K-modes)演算法（類似於K means 分群）將資料集分為兩個分群，並辨認出以下有趣的模式：

- **分群 1**：包含的特徵有行政/技術/管理類型的工作、大學學位、擁有房產。
- **分群 2**：包含的特徵有藍領工作、高中學歷、沒有房產。

透過來自分群機器學習演算法的見解，行銷長及行銷團隊可以設計下一次行銷活動，針對每個群組目標進行個別定位。[16]

[14] Sutton, R.; A. Barto; *Reinforcement Learning: An Introduction*, Second Edition, MIT Press, USA, 2018, https://mitpress.mit.edu/books/reinforcement- learning-second-edition
[15] DeepMind, "AlphaGo," https://deepmind.com/research/case-studies/alphago-the-story-so-far
[16] Moro, S.; P. Cortez; P. Rita; UCI Machine Learning Repository, "A Data-Driven Approach to Predict the Success of Bank Telemarketing," *Decision Support Systems* 62:22-31, June 2014, Elsevier, https://archive.ics.uci.edu/ml/datasets/bank+marketing

機器學習應用程式之科技風險

構建機器學習應用程式通常遵循典型的資料科學和軟體開發生命週期。本節將透過強調下列議題，為稽核相關議題提供路徑圖，並突顯關鍵風險因子：

- 資料治理
- 資料工程
- 特徵工程
- 模型訓練
- 模型評估
- 模型部署及預測

這些階段通常不是線性的，尤其是在涉及模型再訓練時，可能引入迭代過程。（本節假設已辨認機器學習應用程式的商業問題和成功標準，例如：將房價預測準確率提高至95%。）

資料治理

資料治理專注於管理企業已擁有或計劃取得的資料策略。

資料治理架構旨在建立健全的資料治理、標準化資料管理實務及在AI應用程式的誠信方面強化信任。在稽核資料治理架構時，稽核人員應考慮但不限於以下方面：

- **當責**—基於國際內部稽核協會的三道模型(IIA's three lines model)，資料治理架構文件化，且明確定義角色與職責的分配，包括高階管理層的參與。監督與監控是關鍵。
- **資料品質**—確保所使用的資料品質良好，並將任何資料品質議題適時上報給負責人員進行矯正。
- **資料安全**—存在適當的資料保護措施，以防止未經授權的存取及網路威脅（如資料中毒、AI攻擊和駭客攻擊）。保持對新興科技安全威脅的警覺至關重要。
- **資料授權**—針對開源資料集，需了解其在商業應用中的任何限制。了解熱門開源資料授權的含義（例如：創用CC相同方式分享）是必要的。

資料治理架構旨在建立健全的資料治理、標準化資料管理實務及在AI應用程式的誠信方面強化信任。

資料工程

資料工程聚焦於資料蒐集與分析的實務應用。它是構建機器學習應用程式的第一步，以因應特定商業問題。

目前，ML 和 AI 應用程式模型極度依賴訓練資料。因此，查核資料來源是 ML 稽核的關鍵首要步驟，在此階段，稽核人員應專注於資料的合規性與治理。稽核人員需查核資料來源是否符合GDPR、CCPA及其他相關法規的規定。

稽核從業人員在資料工程階段中應檢查以下關鍵領域：

- **資料蒐集**—將資料蒐集並導入目標分析環境時，需關注以下方面：
  - 資料來源的可信度
  - 所提供資料的準確性與完整性
  - 接收資料的企業是否具備適當作業流程，以檢視輸入資料的準確性與完整性
  - 資料來源，例如 Web API、其他資料庫或現有資料管道
  - 不同資料格式，如 CSV、JSON、日誌文件、原始檔、影像及音檔

- **資料探索**—探索哪些資料項目可能對構建 ML 預測模型有用。需注意原始資料來源通常較為混雜，僅有少數資料項目對商業問題真正有幫助。
- **資料清理**—清理原始資料，尤其是非結構化資料。例如，一些銀行可能使用全大寫字母記錄客戶姓名，或者區碼格式不統一，例如 (650)-123-4567 或 1234567。如果未對這些資料進行標準化，ML 模型可能因資料不準確或缺失而受影響，特別是資料來源需要整合的情形。

原始資料來源通常較為混雜，僅有少數資料項目對商業問題真正有幫助。

- **資料匿名化**—為符合 HIPAA、GDPR、CCPA 及其他法律和監管機構的要求，對受保護健康資訊（PHI）和個人識別資訊（PII）（如社會安全號碼）執行資料分析前可能須進行匿名化處理。
- **資料模式**—以資料庫管理系統（DBMS）所支援的語言描述資料庫模式，資料應用大部分將資料庫用於儲存，以便於存取和分析。
- **擷取、轉換及載入(ETL)資料管道**—擷取正確的資料項目，應用轉換函數，並將其載入資料庫（ETL）。

為符合 HIPAA、GDPR、CCPA 及其他法律和監管機構的要求，對受保護健康資訊（PHI）和個人識別資訊（PII）（如社會安全號碼）執行資料分析前可能須進行匿名化處理。

- **資料公平性**—判斷是否有任何資料項目可能導致對人們生活的不公平影響。例如，包含種族的資料項目可能導致銀行不公平地拒絕少數族裔的貸款申請。

隨堂測驗: 預測貸款申請人的信用額度

一家銀行正在構建一個 ML 應用程式，用於預測貸款申請人的信用額度，資料科學家正要求存取以下資料來源：

1. 客戶全名
2. 出生日期
3. 政府核發的身份證號碼
4. 居住國家
5. 目前居住地
6. 種族
7. 性別
8. 過往年收入
9. 與銀行往來之年數

問題：上述哪些資料來源包含可能受 GDPR 限制的 PII？[17]

進階主題：AI/ML 應用程式中的合成資料

隨著資料科學家在構建 ML 應用程式時存取敏感客戶資料（如 PHI）變得日益困難，一些企業開始生成或購買合成資料，這些資料反映了真實世界資料的重要統計特性。蒐集和使用敏感資料可能引發隱私問題，並使企業面臨資料外洩的風險；因此，《GDPR》和《CCPA》等隱私法規限制個人資料的蒐集和使用，並對違反法規的企業處以罰款。

相較於蒐集大規模資料集的成本，合成資料的成本較低，且能支持 AI/ML 的開發或軟體測試，而不影響客戶隱私。使用合成資料在訓練 ML 模型方面展現出潛力，然而，企業管理階層、風險和法律團隊應評估創建合成資料計畫的潛在效益，並確保由具備高度複雜技能的 AI 專家執行該計畫。[18]

[17] Answer: 1, 3 and 5. University of Pittsburgh, “Guide to Identifying Personally Identifiable Information (PII),” www.technology.pitt.edu/help-desk/how-to-documents/guide-identifying-personally-identifiable-information-pii.
[18] Lucini, F.; “The Real Deal About Synthetic Data,” MITSloan Management Review, 20 October 2021, https://sloanreview.mit.edu/article/the-real-deal- about-synthetic-data/

特徵工程

機器學習術語中的特徵（Feature）是指從原始資料中提取的單一可測量屬性或現象的特徵。特徵工程是指利用領域知識，在使用 ML 或統計建模創建預測模型時，選擇和轉換最相關的變數。一項優秀的特徵能使 ML 預測變得更簡單且更準確。實務上，在機器學習當中，特徵工程可能比演算法的選擇更為關鍵。[19]

以下特徵工程技術是稽核人員需要檢查的關鍵領域：

- **資料插補**—遺漏值(missing values)是準備訓練資料時最常見的問題之一。遺漏值可能由於人為錯誤、資料流中斷或隱私考量，不論原因如何，遺漏值都會對 ML 模型的性能產生負面影響。
- **異常值處理**—在統計學中，異常值是指距離母體均值超過 X 倍標準差的資料點。移除異常值的常見方法是統計學中的 Z-分數，稽核人員需要檢查異常值處理，因為它可能導致在ML 模型中引入系統性錯誤。

範例: 房價預測

一個房地產行銷團隊正在建立一個機器學習（ML）模型來預測房價，資料科學家提出一份可能具有預測性的特徵清單：

- 平方英尺數
- 臥室數量
- 浴室數量
- 房屋類型（獨棟住宅、聯排別墅或公寓）
- 郵遞區號
- 學區評分
- 地區犯罪率

要注意的是有些特徵是固定的（例如平方英尺數），而其他特徵則是情境相關的（例如學區評分）。後者可能需要從不同的資料來源進行關聯處理。

- **衍生性**—有些原始資料項目可能難以直接用於 ML 模型，例如時間戳記。然而，可以從時間戳記「衍生」多種特徵，例如年份、月份、日期、時、分或星期幾。
- **分箱與資料轉換**—在處理數值範圍極大但樣本量較少的資料時，分箱或對數轉換可能有助於模型聚焦於特徵的數量級。例如，與其詢問年收入（可能從 0 到數百萬不等），不如詢問所得稅級距，因為稅率的可能的選項較少且更有代表性。

範例：新生兒客群？

一份行銷團隊的網頁問卷蒐集銀行客戶的年齡，前端網頁表單將年齡預設為 0，這在 ML 模型的訓練資料中引入顯著異常值。在這種情況下，年齡值為 0 可以明確地從模型訓練中排除，然而，在其他情境中，異常值可能更為隱晦。例如，一家嬰兒配方奶粉公司正在蒐集嬰兒年齡時，合法的客戶資料可能顯示年齡為 0，若以 NULL 值替代所有年齡，反而會破壞資料的完整性，工程師在此可能需要對前端資料蒐集方式重新設計。因此，在特徵工程階段，稽核人員對異常值問題應特別關注。

範例：在特徵工程中進行情境編碼

在建立 ML 模型時，原始資料中的時間戳記（例如 2022 年 1 月 3 日）可能不如從該日期衍生出的「星期一」這一特徵有意義。此外，情境還可以被編碼為「該年度的第一個工作日」。相較於僅使用原始時間戳記，基於衍生特徵的 ML 模型更能精準地預測客戶行為，因為更多的情境資訊被包含在特徵工程的步驟中。

[19] Rençberoğlu, E.; "Fundamental Techniques of Feature Engineering for Machine Learning," Towards Data Science, 1 April 2019, https://towardsdatascience.com/feature-engineering-for-machine-learning-3a5e293a5114

- **縮放**—不同的原始資料可能涵蓋不同的資料範圍，例如年齡和收入。在機器學習（ML）中，如何比較這兩項資料？正規化（Normalization）和標準化（Standardization）可以將多變量特徵轉換為相似的規模。例如，年齡和收入都可以縮放到 0 和 1 之間，其中 0 表示原始資料中的最小值，1則表示最大值。

隨著深度學習（DL）演算法的興起，可以通過使用複雜的階層式神經網絡，從原始資料中自動衍生出特徵。

- **公平性**—機器學習中的公平性是近期興起的一個研究領域，目的是確保資料中的偏見和模型的不精確性不會導致模型基於種族、性別、殘疾、性取向或政治立場等特徵對個體進行不公正的處理。[20]稽核人員應對於某一特徵是否引入歧視、偏見或其他不公平的跡象進行盡職調查。
- **隱私**—特徵工程可能引發隱私疑慮，尤其是在將原始資料與外部資料來源進行關聯時。稽核人員應在此階段檢查是否引入額外的隱私疑慮。

隨著深度學習（DL）演算法的興起，可以通過使用複雜的階層式神經網路，從原始資料中自動衍生出特徵。

因此，在深度學習模型中，特徵工程步驟的重要性有所降低。然而，這引發了另一個重要議題：如何確定深度神經網路隱藏層中的黑箱ML模型編碼了哪些特徵。以下部分將探討此主題。

進階主題：白箱與黑箱 ML 模型及模型可解釋性

稽核人員經常遇到兩種類型的機器學習模型：

- **白箱模型**—例如線性迴歸模型通常容易解釋，因為它們的參數較少。然而，在預測複雜資料問題（例如檢測網路照片中的人臉）時，這種模型的效能較差（由於擬合不足）。
- **黑箱模型**—例如深度學習模型通常具有數百萬個參數和複雜的神經網路架構，這使得它們能夠更準確地處理複雜問題。然而，幾乎無法解釋神經網路隱藏層實際學習些什麼。[24]近期的研究（例如 LIME 和 XAI 演算法）旨在探討 DL 模型的可解釋性。[25]

案例探討： Netflix 百萬獎金與資料隱私的教訓(Lesson)

在機器學習時代，資料隱私的問題可能比表面上看起來更隱晦。2009 年，Netflix 被迫取消其百萬美元獎金計劃。該計畫邀請機器學習研究人員基於 48 萬名 Netflix 客戶的訓練資料集，建立一個預測更精準的電影推薦系統。雖然這些資料集於設計時有考慮保護客戶隱私，但隱私提倡者仍對此提出批評。2007 年，德克薩斯大學奧斯汀分校的兩位研究人員能夠透過比對資料集與網路電影資料庫（IMDb）上的電影評分[21]，辨認出個別用戶的身份。2009 年 12 月 17 日，四名 Netflix 用戶對 Netflix 提起集體訴訟[22]，指控 Netflix 違反美國《公平交易法》及《視訊隱私保護法》，因為他們公開這些資料集。這引發大眾關於研究參與者的隱私討論。2010 年 3 月 19 日，Netflix 與原告達成和解，原告後來自願撤回訴訟。[23]

[20] Oneto, L.; S. Chiappa; "Fairness in Machine Learning: Recent Trends in Learning From Data," *Studies in Computational Intelligence*, Springer, USA, 2020
[21] Narayanan, A.; V. Shmatikov; "Robust De-anonymization of Large Sparse Datasets," University of Texas at Austin, www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf
[22] Jane Doe, et al., v. Netflix, Class Action Complaint, US District Court for the Northern District of California, 17 December 2009, www.wired.com/images_blogs/threatlevel/2009/12/doe-v-netflix.pdf
[23] Singel, R.; "NetFlix Cancels Recommendation Contest After Privacy Lawsuit," *Wired*, 12 March 2018, www.wired.com/2010/03/netflix-cancels-contest/
[24] Gunning, D.; M. Stefik; J. Choi; T. Miller; S. Stump; G. Yang; "XAI-Explainable artificial intelligence," *Science Robotics*, 18 December 2019, https://openaccess.city.ac.uk/id/eprint/23405/
[25] Ribeiro, M.; S. Singh; C. Guestrin; "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2016, https://dl.acm.org/doi/10.1145/2939672.2939778

模型訓練

模型訓練是建立 ML 模型的核心。模型訓練的過程涉及選擇最佳超參數，使模型能夠在資料訓練上表現良好，並基於以下商業考量、使用案例及成功標準進行設計：

- **監督式學習**：最小化預測值與目標值之間的誤差。
- **深度學習**：最小化每期訓練迭代的擬合誤差。
- **非監督學習**：最大化集群緊密性指標。

在模型訓練階段進行稽核時，需重點檢查兩個關鍵領域：訓練資料及模型方法論。

- **訓練資料**—訓練資料的品質決定了 ML 模型的品質。稽核人員應該針對訓練資料提出以下問題：
  - 訓練資料是從哪裡蒐集的？明確其來源及存取權限。
  - 訓練資料的統計資料有哪些？例如，正標籤與負標籤的比例。
  - 訓練資料是如何準備的？訓練集、測試集和驗證集之間的分割比例是多少？
  - 如何處理資料不平衡的情況？例如，訓練資料中包含 10 個舞弊案例和 10,000 個正常案例，因為舞弊在母體中是一個低機率事件。是否使用上採樣或下採樣來平衡訓練資料？

模型的概念合理性對成功至關重要，ML 模型訓練的目標是找到過度擬合與擬合不足之間的最佳平衡，並理解所做的假設。

- **模型方法論**—ML 模型的概念合理性對成功至關重要，ML 模型訓練的目標是找到過度擬合與擬合不足之間的最佳平衡，並理解所做的假設。稽核人員應該針對模型方法論提出以下問題
  - 所做的分析假設有哪些？
  - 使用哪些機器學習演算法？
  - 模型選擇的標準是什麼？換句話說，為什麼使用這種演算法而不是其他替代方案？
  - 該 ML 模型的限制是什麼？
- **透明性**—這是一個白箱還是黑箱模型？如何解釋機器學習模型的輸出？

隨堂測驗：用於檢測線上詐欺的 ML 模型

一個反詐欺團隊正在建立一個 ML 模型來檢測線上詐欺。他們從法務部門蒐集過去 100 個詐欺案例。每個詐欺案例都有其獨特特徵，資料科學家希望確保 ML 模型能涵蓋所有特徵。因此，他們使用所有資料案例訓練一個深度為 50 的複雜決策樹模型，並在所有 100 個案例上達到了 98% 的準確率。

問題：這是一個可靠的 ML 模型嗎？[26]

隨堂測驗：預測房價的 ML 原型

某銀行的資料科學團隊開發兩個用於預測房價的 ML 原型：模型 1 是一個具有 5 個參數的簡單線性迴歸模型—白箱 ML 模型。

模型 2 是一個具有 10 層卷積神經網路的深度學習模型—黑箱 ML 模型。

在交叉驗證中，這兩個模型在所有指標上的表現都很好。

問題：哪個 ML 模型更適合？[27]

[26] Answer: No. This is a typical overfitting model. Without proper cross-validation, this model ends up memorizing all the cases but is unable to generalize to new examples.
[27] Answer: Model 1, because it performs just as well but is simpler.

模型評估

模型評估使用不同的指標來嚴格衡量 ML 模型的性能，對 ML 模型進行稽核可能是稽核職能中最關鍵的步驟，因為這涉及所有機器學習模型都通用的定量過程。

如前所述，ML 模型分為監督式學習、深度學習和非監督式學習，且它們的評估指標各有不同。實務上，監督式學習模型的評估較為直接，因為它們使用帶有實值的標籤化資料。

評估監督式學習模型

稽核 ML 模型時常見的陷阱是僅依賴於單一指標，這可能會誤導結果，使模型與在現實世界中的表現不一致，如以下隨堂測驗所示。

隨堂測驗：99%的精確度夠嗎？

某銀行的反詐欺資料科學團隊表示，其第一個 ML 原型在詐欺檢測中達到了驚人的99%精確度。也就是說，ML模型能夠從100筆的詐欺案內檢測出其中的99筆。

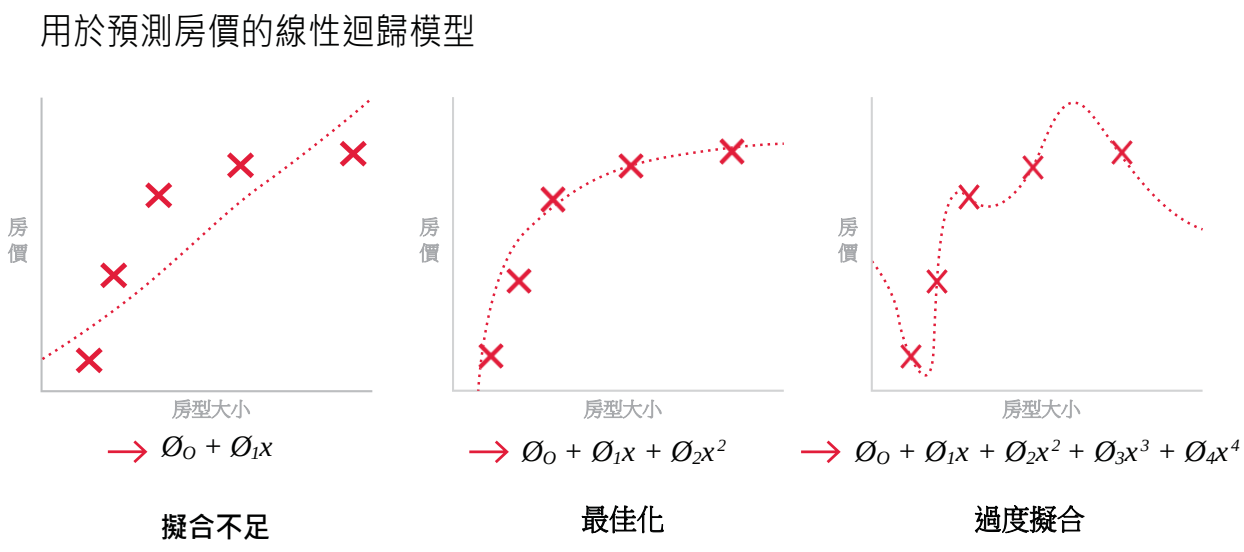
問題：這個 ML 模型夠好嗎？[28]

找到最佳 ML 模型是在「擬合不足」與「過度擬合」之間的一場平衡較量：

- **擬合不足**—對複雜資料來說，模型過於簡單。
- **過度擬合**—模型過於複雜，傾向於「記住」資料中的雜訊，但未能掌握整體趨勢。

圖 2 顯示一個用於預測房價的線性迴歸模型範例。

圖 2：擬合不足、最佳化、過度擬合的機器學習模型



來源： Ng, A.; "Machine Learning," offered by Stanford University through Coursera, https://www.coursera.org/learn/machine-learning

[28] Answer: Not necessarily. An experienced auditor would ask about the false positive rate and the specificity. That is, of the frauds detected, how many were normal transactions (i.e., false positives)? A model that predicts all transactions as fraud would achieve 100% precision, because it would successfully detect all frauds. Yet it is clearly an ineffective ML model, because it would create an excessive number of false positives and would not be practical for production.

為了全面評估 ML 模型，採用科學實驗設定至關重要。

ML 模型訓練的三個資料集

- **訓練資料集**—用於擬合模型的資料樣本。[29]
- **驗證資料集**—用於對模型在訓練資料集上的擬合進行無偏評估，並在調整模型超參數（hyperparameters）時提供回饋。當驗證資料集的性能被納入模型配置時，評估可能會變得更加偏頗。
- **測試資料集**—用於對最終模型在訓練資料集上的擬合進行無偏評估的資料樣本。

K 折交叉驗證(K-fold Cross-Validation)

交叉驗證是一種更嚴謹的評估方法，使用重抽樣程序對有限資料樣本中的 ML 模型進行評估。它將資料隨機分成k組，保留其中一組作為測試資料集，其餘 k-1 組用於訓練過程，並重複 k 次。使用K 折交叉驗證將導致偏差較低的 ML 模型。

混淆矩陣及其衍生指標

混淆矩陣以表格結構呈現，用於協助使用者將ML 模型的整體性能具象化。各種指標可以從中衍生（**圖 3**），使其成為 ML 模型評估中最具洞察力的工具。

- **準確度**—總預測數中正確預測的比例。
- **精確度**—被正確辨認的正樣本比例。
- **陰性預測值**—被正確辨認的負樣本比例。
- **靈敏度或真陽率**—被正確辨認的實際正樣本比例。
- **特異度**—被正確辨認的實際負樣本比例。

圖3：混淆矩陣

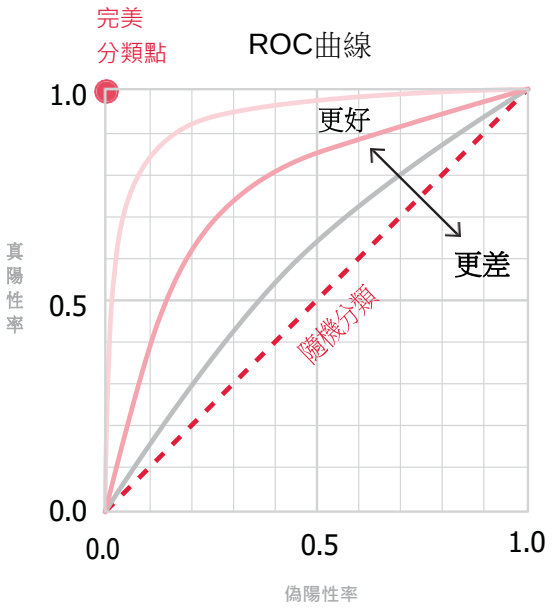
混淆矩陣		實際值			
		陽性 (P)	陰性 (N)		
預測值	陽性 (PP)	真陽性(TP)	偽陽性(FP)	精確度	TP/PP
	陰性 (PN)	偽陰性(FN)	真陰性(TN)	真陽率	TP/P
		靈敏度	特異度	準確度	
		TP/P	TN/N	(TP + TN) / (P + N)	

ROC 曲線與 AUC 分數

接收者操作特徵（ROC）曲線（**圖 4**）通過在不同閾值設置下繪製真陽性率（TPR）與偽陽性率（FPR）來創造。

ROC 分析提供工具來挑選可能的最佳模型。實務上，當 ROC 曲線接近左上角（即完美分類點）時，ROC曲線下面積（AUC）分數接近 1，表示該 ML 模型性能良好。[30]

圖4： ROC曲線



[29] Shah, T.; "About Train, Validation and Test Sets in Machine Learning," Towards Data Science, 6 December 2017, https://towardsdatascience.com/train- validation-and-test-sets-72cb40cba9e7
[30] ROC, https://en.wikipedia.org/wiki/Receiver_operating_characteristic

非監督式學習模型評估

評估非監督式學習模型的性能，例如分群（Clustering），並不像監督分類演算法簡單，因為缺乏標籤化資料。如果不使用定量測量，結果可能帶有主觀性。

直觀來看，一個良好的分群模型應能創造密集的分群（即群體內距離小）並保持群體之間的距離較大（即群體間距離大）。
大多數分群模型的評估指標是基於這些啟發式規則。

直觀來看，一個良好的分群模型應能創造密集的分群（即群體內距離小）並保持分群之間的距離較大（即群體間距離大）。大多數分群模型的評估指標是基於這些啟發式規則。

輪廓係數(Silhouette Score)

輪廓係數[31]衡量分群之間的拆分距離。它顯示每個分群內的資料點與鄰近分群的資料點之間距離有多近。該指標範圍為 [-1, 1]是一個極佳的工具，用於直觀檢查分群內的相似性和分群間的差異性。

輪廓係數的計算基於每個樣本的分群內平均距離 (i) 和最近分群的平均距離 (n)。對於每個樣本，輪廓係數的公式為：(n - i) / max(i, n)，其中 n 是樣本與最相近但不屬於該分群的樣本距離，i 是分群內的平均距離。

其他非監督式學習指標：

- 蘭德指數(Rand Index)[32]
- 調整蘭德指數(Adjusted Rand Index)
- 相互資訊(Mutual information)
- CH指標(Calinski-Harabasz Index)
- D-B指數(Davies-Bouldin Index)

模型部署及預測

一旦 ML 模型完成訓練和測試，下一步是將其部署到營運環境，並對新的即時資料流進行預測。例如，一個隨機森林模型被部署為提供網路服務，判斷支付交易是否存在詐欺行為。

現代 ML 產業正經歷一場重大變革，ML 模型不再是靜態的，模型參數和輸入資料需要根據使用案例定期更新。稽核人員需要執行持續監控。

理解服務水準協議（SLAs）對於 ML 應用程式的最終成功至關重要。模型部署和預測的 SLA 通常包括：

- **硬體需求**—需要哪些硬體（晶片）規格以滿足業務需求？
  - 是基於雲端還是邊緣設備？
  - 中央處理器（CPU）對比圖形處理器（GPU）？
  - 是否需要特定等級晶片的支持？
  - 其他硬體規格（例如 RAM、I/O、網路頻寬）有何要求？
  - 是否符合業務需求？（雲端 ML 網路應用程式將繼承雲供應商的 SLA 條款，適用於某些地區的可用性和網路延遲。）
- **軟體需求**—軟體環境將影響 ML 模型的生產效果。
  - **作業系統**：Linux、Unix、Windows、macOS、IOS、Android 或其他作業系統？
  - **軟體庫版本**：並非所有軟體庫都向後兼容，尤其是開源庫，升級可能導致之前正常運作的應用程式中斷。因此，稽核人員需檢查版本問題。

[31] Zuccarelli, E.; “Performance Metrics in Machine Learning — Part 3: Clustering,” Towards Data Science, 31 January 2021, https://towardsdatascience.com/performance-metrics-in-machine-learning-part-3-clustering-d69550662dc6
[32] Scikit-Learn, https://scikit-learn.org/stable/modules/clustering.html#rand-index

- **等待時間與吞吐量—**等待時間衡量每次預測所需的時間，吞吐量衡量同時處理的預測量。稽核人員應考量以下事項：
  - 評估等待時間，每次預測需耗時多少毫秒？在某些金融場景（如信用卡支付），交易詐欺偵測可能需要即時預測能力（如毫秒級），以避免對客戶滿意度造成負面影響。
  - 評估吞吐量，ML 應用程式在高峰期需處理多少的預測？是否支援批次處理模式？

持續性監控對於實現 ML 模型的最佳性能至關重要，特別是在資料流中可能引入 ML 模型無法準確解讀的新案例時。這在對抗性環境（如詐欺偵測和網路安全入侵偵測）中特別常見，因為對手會改變攻擊策略以繞過現有的檢測方法。稽核人員應檢查 ML 模型的營運環境是否實施持續性監控，以適應新資料場景，並確定重新訓練 ML 模型的最佳時機。

ML 和 DL 的熱門程式語言包括 Python、Julia、R 和 Java，但 Python 可能正成為機器學習的首選語言。豐富可用的程式庫和開源工具使得 Python 成為開發 ML 模型的理想選擇。

範例：深度學習與圖形處理器(GPU)

大多數深度學習（DL）演算法主要基於 GPU 獲得最佳速度，僅使用 CPU 的架構上表現不佳。在近期的 TensorFlow 基準測試中，使用 GPU 的卷積神經網路（CNN）比僅使用 CPU 的速度快超過 600%。

稽核人員應檢查是否有假設特定的處理硬體環境。[33]

熱門機器學習資料集

作為技術風險評估過程的一部分，IT 稽核人員應檢查所採用的 ML 庫。ML 和 DL 的熱門程式語言包括 Python、Julia、R 和 Java，但 Python 可能正成為機器學習的首選語言。豐富可用的程式庫和開源工具使得 Python 成為開發 ML 模型的理想選擇。[34]

持續性監控對於實現 ML 模型的最佳性能至關重要，特別是在資料流中可能引入 ML 模型無法準確解讀的新案例時。

範例：Apple M1 晶片組用於機器學習

某些 ML 應用將部署於行動裝置上，因此稽核人員應檢查是否需要特定的晶片組功能來達到預期效能。Apple 的 M1 系統晶片，結合 Apple Neural Engine，用於 ML 模型時可提供高達 15 倍的加速效能。[35]

範例：10 毫秒的 ML 預測速度是否足夠快？

一個付款詐欺 ML 模型僅需 10 毫秒即可預測交易是否可能涉及詐欺，這看似是一個相對快速的計算。然而情境十分重要。例如，全球支付網路(如Visa)每秒處理 24,000 筆交易。

在這種情況下，在一個網路應用伺服器上，ML 系統需耗時 240 秒才能處理 1 秒的交易吞吐量，這是不可接受的。因此，了解 ML 應用如何擴展以處理高吞吐量非常重要。

[33] DATAmadness, “TensorFlow 2 - CPU vs GPU Performance Comparison,” 27 October 2019, https://datamadness.github.io/TensorFlow2-CPU-vs-GPU
[34] Costa, C.; “Best Python Libraries for Machine Learning and Deep Learning,” Towards Data Science, 24 March 2020, https://towardsdatascience.com/best-python-libraries-for-machine-learning-and-deep-learning-b0bd40c7e8c
[35] Apple, “Apple Unleashes M1,” 10 November 2020, www.apple.com/newsroom/2020/11/apple-unleashes-m1/

常用的開源 ML 函式庫

廣泛使用的開源機器學習（ML）函式庫包括以下內容：

- **TensorFlow**—TensorFlow 是目前 Python 最優秀的機器學習函式庫之一。由 Google 提供，TensorFlow 使初學者和專業人員都能輕鬆建立 ML 模型。
- **PyTorch**—由 Facebook 開發的 PyTorch ，是 Python 的主要機器學習函式庫之一。除了支援 Python，PyTorch 還透過其 C++ 介面支援 C++，被認為是最佳機器學習和深度學習框架的主要競爭者之一。
- **Scikit-learn**—Scikit-learn 是一個常用的 Python 機器學習函式庫，易於整合如 NumPy 和 Pandas 等不同的 ML 程式函式庫。
- **Pandas**—Pandas 是一個 Python 資料分析函式庫，主要用於在準備訓練資料集之前進行資料操作和分析。Pandas 使得機器學習程式人員在處理時間序列和結構化多維資料時更加輕鬆。
- **Spark MLlib**—對於大規模機器學習，Apache Spark 是一個熱門選擇。Apache Spark MLlib 是一個機器學習函式庫，可以輕鬆擴展計算，簡單易用且設置快速，並提供與其他工具順利整合的解決方案。

商業機器學習函式庫

廣泛使用的商業機器學習函式庫包括以下內容：

- SAS
- H2O
- RapidMiner

稽核人員應針對機器學習函式庫提出以下問題：

- 此 ML 函式庫具有什麼軟體授權？
- 此 ML 函式庫是否為最新、穩定的版本？
- 此 ML 函式庫是否包含已知漏洞（例如在 CVE[36] 等資料庫中辨認的漏洞）？例如，NumPy（Python）曾在某些版本中報告過重大漏洞，如 DoS 溢位。
- 是否存在已知的相容性或依賴性問題？例如，Python 中的熱門常駐函式庫 Pickle 在某些版本中曾報告與 Pandas 資料框架的相容性問題。[37]

結論

本白皮書已辨認的路徑要素—資料治理、資料工程、特徵工程、模型訓練、模型評估以及模型部署及預測—是 ML 應用程式中的關鍵風險因子。這些要素提供在 ML 中的具體階段中，從業人員可用於辨認關於稽核考量事項。

對於希望熟悉 ML 的稽核人員來說，理解模型在部署之前的訓練和測試至關重要。一旦 ML 模型部署，稽核人員應檢查資料異常值並意識到持續監控的必要性，因為這些模型並非靜態。

本白皮書為稽核人員理解這些及其他 ML 稽核方面提供基礎。本白皮書也辨認 ML 實施前、實施後及持續性監控的要素，稽核人員需要評估並傳達這些要素，以辨認 ML 應用是否符合預期，以及是否有改進的空間。

[36] MITRE Corporation, “CVE® Details: The Ultimate Security Vulnerbility Data Source,” https://www.cvedetails.com/vulnerability-list/vendor_id-16835/product_id-39445/Numpy-Numpy.html
[37] GitHub, Inc., “Pickle incompatibility between 0.25 and 1.0 when saving a MultiIndex dataframe #34535,” https://github.com/pandas-dev/pandas/issues/34535

附錄: 推薦資源

機器學習是數十年來計算機科學中最活躍的學術研究領域之一，研究人員和統計學家的努力使該領域從抽象走向成熟。

儘管從業人員無需完全理解機器學習演算法的理論面，以下書籍和線上資源推薦給希望深入了解機器學習的人士。

書本

Bishop, C.; *Pattern Recognition and Machine Learning*, Springer, USA, 2006, https://link.springer.com/book/9780387310732

Domingos, P.; *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, USA, 2015, https://www.basicbooks.com/titles/pedro-domingos/the-master-algorithm/9780465061921/

Goodfellow, I.; Y. Bengio; A. Courville; *Deep Learning*, MIT Press, USA, 2016, https://mitpress.mit.edu/books/deep- learning

Hastie, T.; R. Tibshirani; J. Friedman; *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, USA, 2016, https://link.springer.com/book/10.1007/978-0-387-84858-7

課程

Ng, A.; "Machine Learning," offered by Stanford University through Coursera, www.coursera.org/learn/machine- learning

致謝

ISACA 謹此致謝：

開發領袖

Victor Fang, Ph.D.
CEO, AnChain.AI Inc., 美國

校審專家

Ibrahim Sulaiman Alnamlah
CISA, COBIT 2019, ITLv4, ISO/IEC 27001 LA
沙烏地阿拉伯

Chetan Anand
CDPSE, CCIO, CPISI, ISF IRAM2, ISO 9001 LA, ISO 22301 LA, ISO 27001 LA, ISO 27701, ISO 31000, SQAM, Lean Six Sigma Green Belt, Agile Scrum Master, Fellow of Privacy Technology, NLSIU Privacy and Data Protection Laws
印度

Shigeto Fukuda
CISA, CDPSE
日本

Kevin Fumai
CDPSE, CCSK, CEET, CIPM, CIPP/US/E, CIPT, FIP, PLS
美國

Shamik Kacker
CISM, CRISC, CDPSE, CCSK, CCSP, CISSP, ITIL Expert, TOGAF 9
美國

Harmendra Nitish Koladoo
CISA, CDPSE, CEH, CHFI, ISO 27001 LA
香港, 中國

Frank Oelker
CISA
德國

Christian Nyanor Ohene
CISA, CEH
迦納

Jian Qin, Ph.D.
CISA, CISSP, PMP, ODSC Machine Learning Certification
美國

校審專家 *(續)*

Xitij U. Shukla, Ph.D.
CISA
印度

Ioannis Vittas
CISA, CISM, COBIT 2019 Foundation, BCCLA
希臘

Wickey Wang
CISA, Six Sigma Green Belt
美國

Yap Kai Yeow
CISA, CISM, CRISC, CAMS
新加坡

ISACA 董事會

Pamela Nigro, Chair
CISA, CGEIT, CRISC, CDPSE, CRMA
Vice President, Security, Medecision, 美國

John De Santis, Vice-Chair
Former Chairman and Chief Executive Officer, HyTrust, Inc., 美國

Niel Harper
CISA, CRISC, CDPSE, CISSP
Chief Information Security Officer, Data Privacy Officer, Doodle GmbH, 德國

Gabriela Hernandez-Cardoso
Independent Board Member, 墨西哥

Maureen O'Connell
NACD-DC
Board Chair, Acacia Research (NASDAQ), Former Chief Financial Officer and Chief Administration Officer, Scholastic, Inc.,
美國

Veronica Rose
CISA, CDPSE
Senior Information Systems Auditor–Advisory Consulting, KPMG Uganda, Founder, Encrypt Africa,
肯亞

David Samuelson
Chief Executive Officer, ISACA, 美國

Gerrard Schmid
Former President and Chief Executive Officer, Diebold Nixdorf, 美國

Bjorn R. Watne
CISA, CISM, CGEIT, CRISC, CDPSE, CISSP- ISSMP
Senior Vice President and Chief Security Officer, Telenor Group, 美國

Asaf Weisberg
CISA, CISM, CGEIT, CRISC, CDPSE, CSX-P
Chief Executive Officer, introSight Ltd., 以色列

Gregory Touhill
CISM, CISSP
ISACA Board Chair, 2021-2022
Director, CERT Center, Carnegie Mellon University, 美國

Tracey Dedrick
ISACA Board Chair, 2020-2021
Former Chief Risk Officer, Hudson City Bancorp, 美國

Brennan P. Baybeck
CISA, CISM, CRISC, CISSP
ISACA Board Chair, 2019-2020
Vice President and Chief Information Security Officer for Customer Services, Oracle Corporation, 美國

Rob Clyde
CISM, NACD-DC
ISACA Board Chair, 2018-2019
Independent Director, Titus, Executive Chair, White Cloud Security, Managing Director, Clyde Consulting LLC, 美國

關於ISACA

50多年來，ISACA®（www.isaca.org）在科技領域中推動了最優秀的人才、專業知識和學習。 提供個人知識、證書、教育和社群，以促進其職業發展並轉變組織，並協助企業培訓及建立優質團隊。ISACA是一個全球專業協會及學習型組織，致力在提升150,000名會員在資訊安全、治理、確信、風險和隱私等方面的專業知識，並藉由技術推動創新。目前在188個國家及地區發展，並於全球具有220多個分會。2020年成立了慈善基金會 Tech，為資源不足、非代表性的族群提供資訊教育和職涯的支持。

ISACA

1700 E. Golf Road, Suite 400
Schaumburg, IL 60173, USA

電話: +1.847.660.5505

傳真: +1.847.253.1755

聯繫資訊: support.isaca.org

網站: www.isaca.org

提供回饋:

www.isaca.org/audit-practitioners-guide-to-ML-part-1

參加ISACA線上論壇:
https://engage.isaca.org/onlineforums

Twitter:
www.twitter.com/ISACANews

LinkedIn:
www.linkedin.com/company/isaca

Facebook:
www.facebook.com/ISACAGlobal

Instagram:
www.instagram.com/isacanews/

免責聲明

ISACA 設計並完成本創作「稽核從業人員之機器學習參考指引，第一部分：科技」（著作），主要作為專業人士的教育資源。ISACA 並不聲稱使用本著作任何內容將確保有成功的結果。 該著作不應被視為包含所有適當的資訊、程序及測試或排除合理指導獲得相同結果的其他資訊、程序與測試。 在確定任何特定資訊、程序或測試的適當性時，專業人員應將本身之專業判斷應用於特定系統或資訊科技環境所呈現的特定情況。

保留權利

中文版致謝名單

ISACA台灣分會 高進光理事長

翻譯: 林煒傑

校稿: 陳政龍

編輯排版: 游恬欣