2023資產管理金融新知暨法遵系列研討會

「生成式AI資產管理應用」活動實錄

安侯企業管理(股)公司
林大馗 執行副總經理

證基會 張麗真總經理

國立政治大學
數位金融創新實驗室
謝明華 執行長

生成式AI技術引起各界高度關注，該技術的發展有助降低成本、提高效率及推動金融市場發展，卻也可能產生資訊正確性與完整性、資訊安全與隱私保護、弱勢公平性、道德倫理四大風險。因此，金管會將著手新版金融科技發展方案，屆時將金融業使用AI納管，以因應伴隨新興科技而來的風險。

為協助資產管理業瞭解相關應用及風險，特舉辦本研討會，安排兩位實務專家擔任講席，並提供各講席和與會者交流之平台。

「生成式AI資產管理應用」議程

日期：112年6月1日(四) 14:00-17:00

地點：台北花園酒店2樓國際廳

時間	主題	講席
14:00-14:10	主辦單位致詞	張麗真 總經理 證券暨期貨市場發展基金會
14:10-15:00	專題演講1 生成式AI創新應用與衝擊	林大馗 執行副總經理 KPMG 安侯企業管理(股)公司
15:00-15:10	意見交流	
15:10-15:30	中 場 休 息	
15:30-16:50	專題演講2 ChatGPT 在金融業的可能應用	謝明華 執行長 國立政治大學 數位金融創新實驗室
16:50-17:00	意見交流	

主辦單位首長致詞

證基會 張麗真 總經理

張總經理表示，生成式AI這項新技術的導入對各領域產生廣泛影響，其高效率和低成本的特性，使其成為各組織和產業發展的重要工具。然而，該技術的蓬勃發展卻也帶來不少衝擊與挑戰，包括資料正確性、安全性和道德風險等不可忽視的問題。

金融產業亦相當關切ChatGPT的應用對該領域將產生何種影響，並應因應防範相關風險，以確保金融體系的穩定運行。

今日的研討會很榮幸邀請KPMG林大馗副總經理將特別針對金融業風險控管詳加說明，並由政治大學謝明華執行長向各位分享ChatGPT於金融業的應用。

專題演講1:
生成式AI創新應用與衝擊

KPMG 安侯企業管理(股)公司　林大馗 執行副總經理

林副總經理首先提到，2023年世界經濟論壇全球風險調查報告顯示，超過75%受調者對網路犯罪及網路安全威脅尚未做好完全準備；微軟公司調查報告亦指出，可預期組織將大量運用AI，但是高達89%的組織尚未有充分研究並且部署相對應的解決方案或控制框，以因應AI相關的安全議題。然而，人工智慧正融入到每個人的職場與生活，未來更將全面運用於生產場景，且商機持續擴大。

人工智慧發展至今已能自我生成諸如文字、圖像、音訊甚至程式等多種內容，不過儘管如此，必須強調人的價值與重要性並未就此抹滅，尤其體現於後續功能整合、測試等人為參與的部分，仍須人類的專業知識和判斷力來適應特定的需求與環境。

AI運作的基本架構分為三階段：

1. 資料蒐集：前端資料蒐集重點包含減少模糊資料、採用自動標註，避免人工標註、透過實驗推估所需資料量、減少不良資料及收集負樣本。

2. 預處理階段：資料預處理是通過修改、添加或刪除資料的方式為資料分析做準備的過程，此過程通常也稱作資料清洗(Data Cleaning)。
3. 訓練階段：透過演算法運算與深度學習來評估及訓練模型，最後部署及結果輸出。

所有生成式AI的產出都是機率推算的結果，因此想要取得「較為正確」的答案，最重要的事情就是要提出正確的問題，應避免使用「你認為...」、「你覺得...」、大哉問、前後文矛盾等缺乏精確範圍的指示，盡量採取目的優先、直接輸入完整資料和設定輸出等方式，並善用「請繼續」語句、設定提示「溫度」及適時角色重新設定，以強化提示，最後人工確認結果輸出的正確性。

如今AI雖能夠進行大數據的驅動、具備更有效能的演算法及更豐富的應用場域，卻也面臨到資訊安全風險問題。其引發的風險成因，基本可分學習偏失、人為錯誤(包含設計者或使用者，對於演算法專業疏失，或其他人為道德與管理問題)、技術缺陷、流程瑕疵(如使用者未瞭解演算法的限制，而產生的過度依賴與誤用情境)、資安攻擊(勒索軟體)、隱私侵害(演算結果濫用導致侵犯當事人隱私權)、決策誤用、演算效能不足等八類。此外，不論從服務供應商(如：OpenAI)、北大西洋公約組織(NATO)或是社群媒體(如：Meta)等不同角度、觀點，皆顯示對於生成式AI資安的疑慮。

根據2022臺灣企業資安曝險調查報告顯示：

1. 多數企業輕忽社群媒體所衍生的網路攻擊
2. 臺灣各產業資安人員能量均嚴重不足，企業資安人力亮警訊
3. 供應鏈核心產業(原物料、運輸業等產業)極需加強網路防護
4. 金融業網路防護表現仍最佳，但仍面臨高度挑戰
5. 導入並驗證資安國際標準，將顯著降低資安曝險

因此資安在AI時代將更為複雜，並形成四點趨勢：

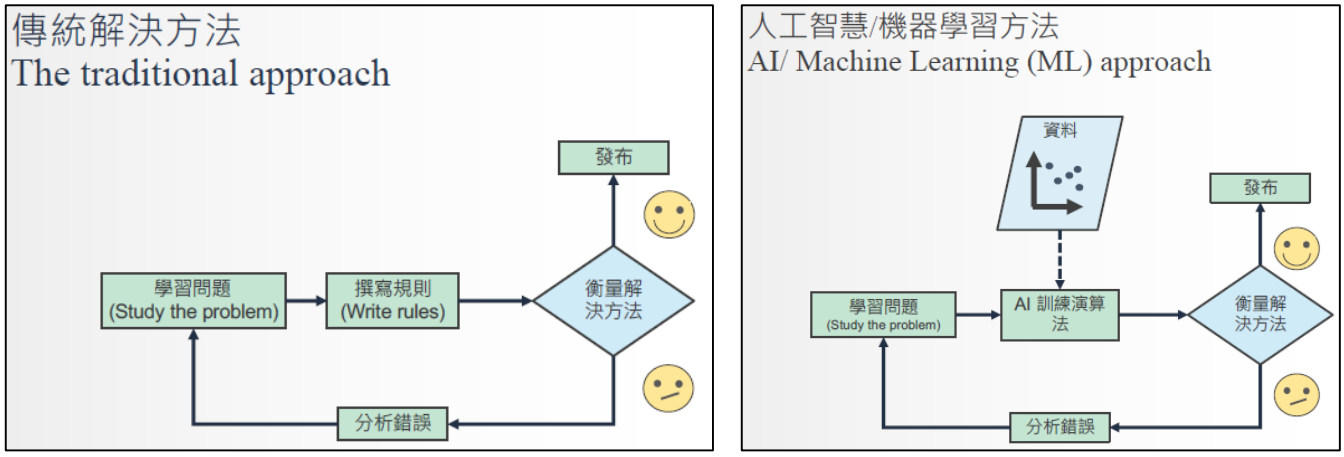
1. Deep fake深假技術將會被大量使用
2. 社交工程所引發之資安事件將會大量上升
3. 自動生成之駭客工具將讓駭客產業進入門檻急遽降低
4. 人工智能內容生成(AICG)可能造成假新聞及假消息氾濫。

AI風險下如何自處？首先仍應考量安全再決定是否託付，面對AI存在的偏見，產生誤導或不正確的資訊，也呼籲大眾要評估使用，並透過「人工智慧風險治理心法」進行修復，包含邊做邊學、感知技術的成熟、人工介入的必要性和演算法運算的速度。最後，林副總經理提及，國際間因應ChatGPT分別制定不同法規和規定，以美國為例，要求演算法資訊揭露、開發模型治理和資安與可用性的管理，以嚴格控管演算法開發過程品質。總而言之，應用AI的「智慧攻擊」，已經成為未來惡意攻擊的主流模式，以AI反制AI攻擊，是當務之急。以AI做為服務時，得以「FIRST」為優先原則：公正(Fair)、包容(Inclusive)、負責(Responsible)、安全(Safe)和透明(Transparent)。

專題演講 2：
ChatGPT 在金融業的可能應用

國立政治大學數位金融創新實驗室　謝明華 執行長

謝執行長首先說明傳統解決方法與AI之比較，如下圖：



AI的好處在於自動適應改變，但受限於資料蒐集過於昂貴及資料並非以AI訓練為目的，因此過去並不普遍採用。AI/ML系統分為「監督式學習」、「非監督式學習 」、「半監督式學習」和「強化式學習」(唯一一種不需要資料即能運作的方法)四類，其中「半監督式學習」透過現有資料即可運用，因而被許多新創公司開發並採用。

AI優劣與否取決於資料的數量及完整性，ChatGPT之所以被大量討論，

係因其改變過去對AI的做法，其內涵包括轉移式學習（TL）、基礎模型和生成式。

「轉移式學習」概念在於利用一個領域學習到的知識，將其轉移到另個領域中，目的在於解決數據蒐集不易及克服資料集需具有相同分配的限制等問題。TL模型種類採用率高至低依序為：「預訓練和微調模型」、「領域自適應」、「知識蒸餾」、「多任務學習」，謝執行長針對採用率最高及最低者進行介紹：

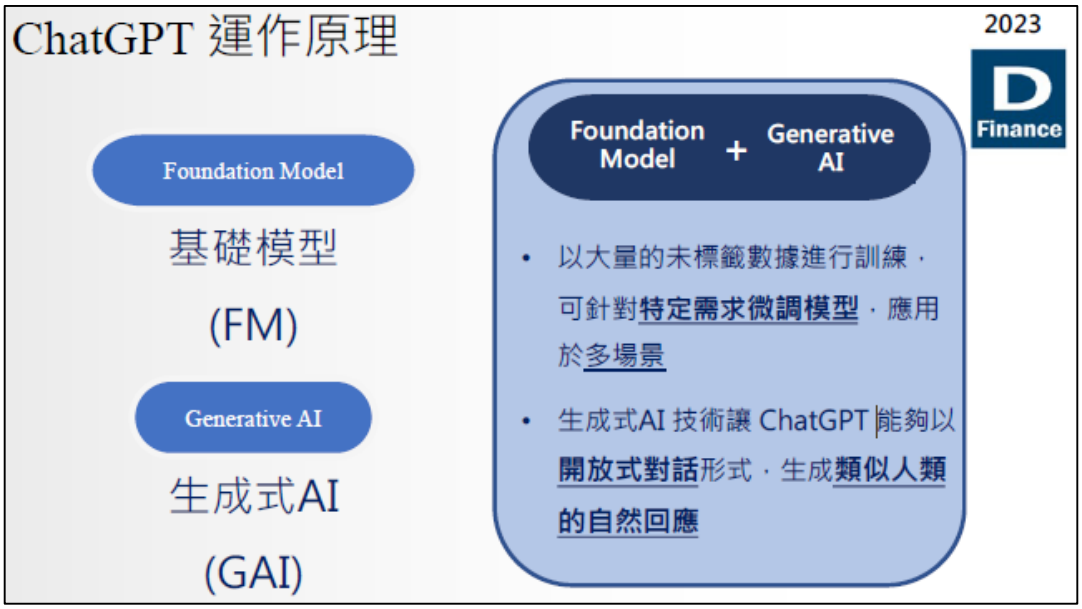
1. 「預訓練和微調模型」之「預訓練」是指事先訓練，使用通用特徵的大規模數據集進行預訓練模型學習；「微調」則以目標領域及任務的小型數據集訓練，再進行參數微調。
2. 「多任務學習」是共享特徵學習，讓模型同時對多項具相關性任務訓練，提高模型的泛化能力，以解決特定任務數據缺乏問題。

總體而言，TL具備解決數據稀少、加快訓練速度、減少碳足跡和大眾化等優點。

「基礎模型」主要是語言模型，以未標記資料集進行預訓練，微調後可運用至各領域，其具有變換器架構(學習上下文的神經網絡、專注力與自我專注力)、湧現特性、自我監督式學習(通過未詮釋的資料推導預訓練任務)及可擴展的基礎設施(高度彈性的計算、內存、儲存和網絡來進行訓練；GPU架構和資料並行性)等特性。應用領域涵蓋自然語言處理(NLP)、計算機視覺、軟體工程和一般科學等，目的是解決AI的共同問題，包含高時間成本、資源需求大、需不斷重新訓練及資料準備昂貴、耗時，而其優勢於「規模」、「準確性」、「使用預訓練模型入門成本低」及「領域適應」等四方面。

生成式AI(GAI,Generative AI)挑戰讓電腦產生新的、原創的內容，或將原有內容變得有新意，例如GANs(Generative Adversarial Networks)，能夠擴增與加速多種領域的設計，並應用於資料科學與商業分析。

ChatGPT為Open AI所開發的NLP模型，GPT-3.5主要處理文字相關問題，與GPT-3相較，其使用更多的訓練數據及更高的參數提高回應品質，提升對話自然性；GPT-4能夠處理圖片；GPT-5則是已能處理影像。如今所談的ChatGPT添加了使用者介面，以開放式對話形式回應，運作原理如下圖（圖一）：



圖一　ChatGPT運作原理

其使用方式包含「開箱即用」、「提示工程」、「自定義模型」，優缺點如圖二所示：

圖二　ChatGPT使用方式之優缺

ChatGPT輸出的可用程度取決於輸入的提示品質，需熟悉專業知識及領域，並能夠撰寫適當的提示以得到實用的回應。作為ChatGPT的使用者，需注意其輸出內容之可信度並非全然正確，甚至可能以錯誤資訊說服使用者，因此有責任透過事實查核(fact-check)驗證輸出的準確性；ChatGPT的定位應為「強化」而非取代使用者的創造力與判斷力。

最後謝執行長提醒，ChatGPT存在四種風險與限制：

1. 幻覺問題：導致創造非事實或不符邏輯的內容。
2. 訓練資料問題：可能因不足、過時或包含敏感訊息和偏見，導致回應不適當或錯誤。
3. 侵犯版權：使用受版權保護的數據來訓練模型。
4. 侵犯知識產權：輸入ChatGPT的資料會再用於訓練模型。

關於在金融業的可能應用如下(圖三)所示，謝執行長建議在使用上要有創新實驗室，並了解其發展性。

圖三　金融業可應用的領域