

星球永續健康線上直播

星球健康週新知 &

專題: 可解釋AI (XAI) 醫療應用

XAI規範發展與大腸癌智慧防治應用

2025-12-10

CHE團隊：

陳秀熙教授、許辰陽醫師、陳立昇教授、嚴明芳教授、林庭瑀博士、
劉秋燕、羅崧瑋、林家妤、陳虹玟



資訊連結:

<https://www.realscience.top/7>

星球永續健康線上直播



<https://www.realscience.top/4>

Youtube影片連結: <https://reurl.cc/gWjyOp>

漢聲廣播星球永續健康:

https://audio.voh.com.tw/TW/Playback/ugC_Playback.aspx?PID=323&D=20240615

新聞稿連結: <https://reurl.cc/no93dn>

本週大綱

- 星球健康新知 (2025 / W48)
- XAI醫療應用挑戰與解方: 大腸癌預防
- XAI醫療輔助實例: 大腸癌智慧防治應用

星球健康新知

2025 / W48

美國主導俄烏和談 歐洲尋求軍防共識



普丁將在莫斯科會晤川普特使
烏克蘭和談進入關鍵攻防



美俄高層接觸後
烏克蘭戰爭終結仍未列入議程

馬克宏促歐洲加入烏克蘭談判
力求在和平進程中占有一席之地



歐盟推動凍結俄資產以防一面倒和談



芬蘭外交部長兼歐洲安全合作組織(OSCE)
主席艾莉娜主持 12/4維也納年度會議

東剛果達成和平協議 盧安達-剛果衝突暫歇

剛果東部戰事再度升高
政府軍與 M23 對峙加劇



衝突區民眾被迫逃離家園
日常生活面臨巨大考驗



bbc.com



剛果與盧安達今年六月在華府簽署
初步和平與經濟協議，為談判奠定基礎



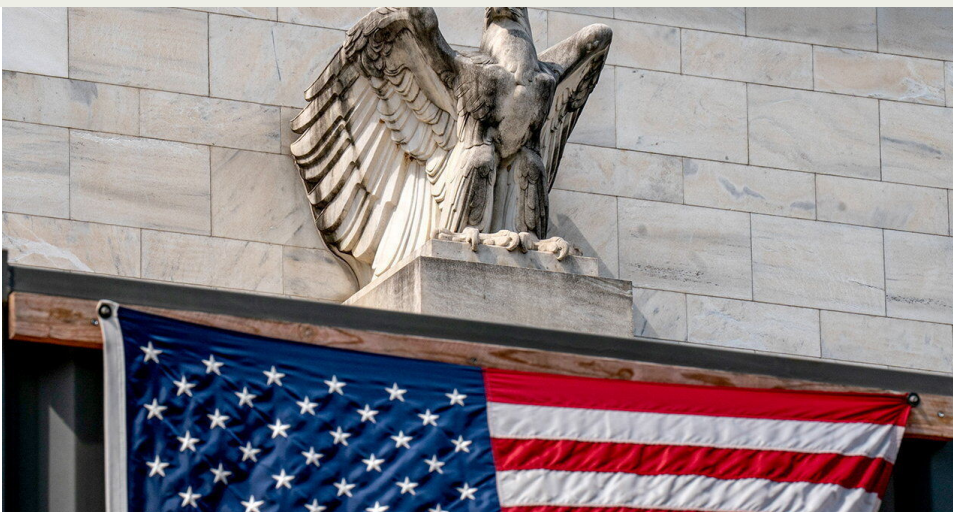
washingtontimes.com

十二月美方主導新一輪協議推進
盼促成實質和平進展

全球科技供應鏈重塑區域經濟發展



聯準會主席鮑爾擔憂關稅與貨幣政策
衝擊美國經濟與就業市場



聯準會預期將結束縮表等經濟緊縮政策
並重返債市，預計將造成美債殖利率降低

台積電在東亞與國際政局壓力下
加速海外布局穩定先進晶片研發生產

npr.org



出口回溫帶動南韓出口經濟復甦



南韓提出經濟發展策略
出口景氣回穩

歐洲通膨隱憂
拖累降息前景

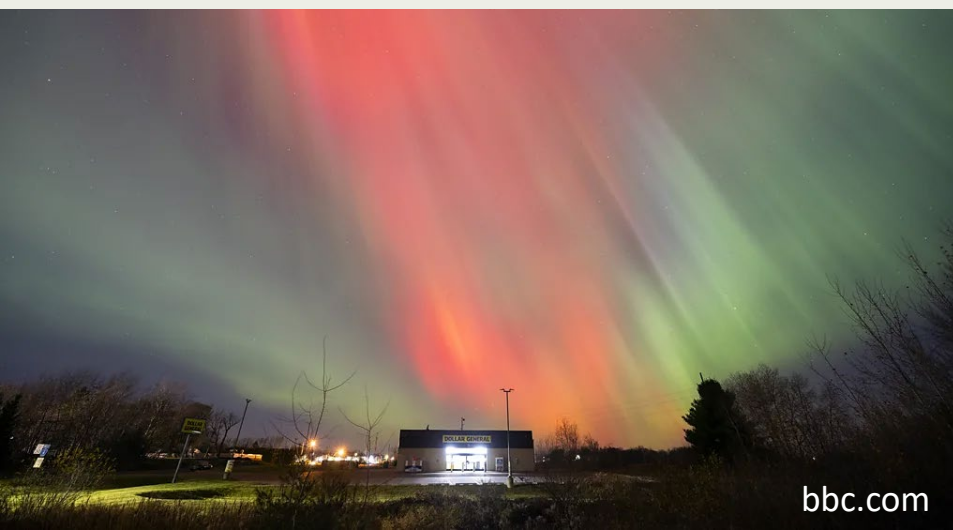
太陽風暴觸發A320全球緊急升級

太陽耀斑對飛機導航構成風險



全球機場A320機隊列隊待命同步進行
電傳操縱核心架構軟體更新

太陽耀斑導致全球中緯度地區如佛羅裡
達州與義大利北境出現壯觀極光景象



電傳操縱核心架構：
A320以電腦轉化飛行員指令

飛機凝結尾跡 候變遷影響科學評估

Paul Voosen, Science, 2025

傳統認為

飛機凝結尾跡轉為長壽雲後，會封存大量熱能，其增溫效果可達飛機引擎排放CO₂的三倍。

研究試圖透過改變飛行航線，減少凝結尾跡形成，以降低航空對氣候衝擊。

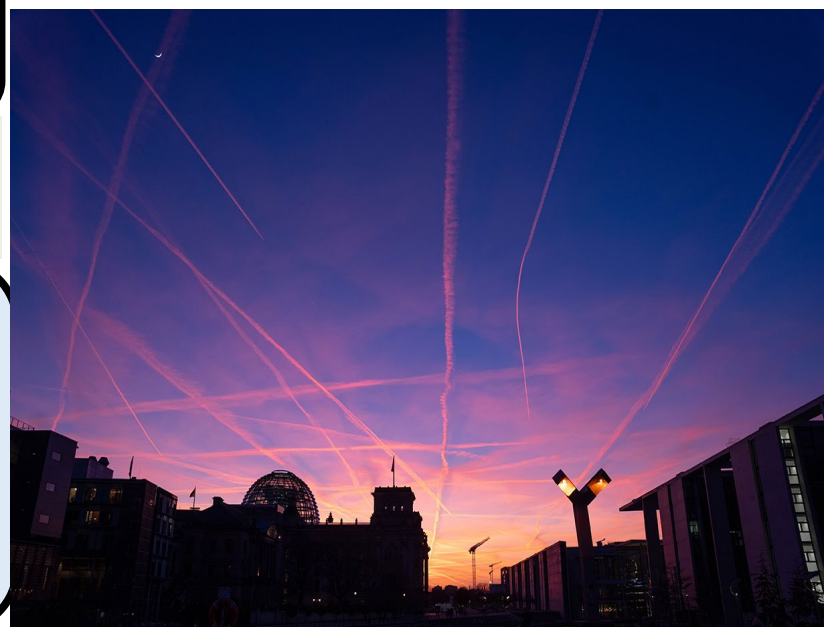
噴射渦輪機燃燒會產生二氧化碳、水蒸氣以及少量煙塵。
當水蒸氣在煙塵上凝結並凍結時，就會形成凝結尾跡。

最新研究

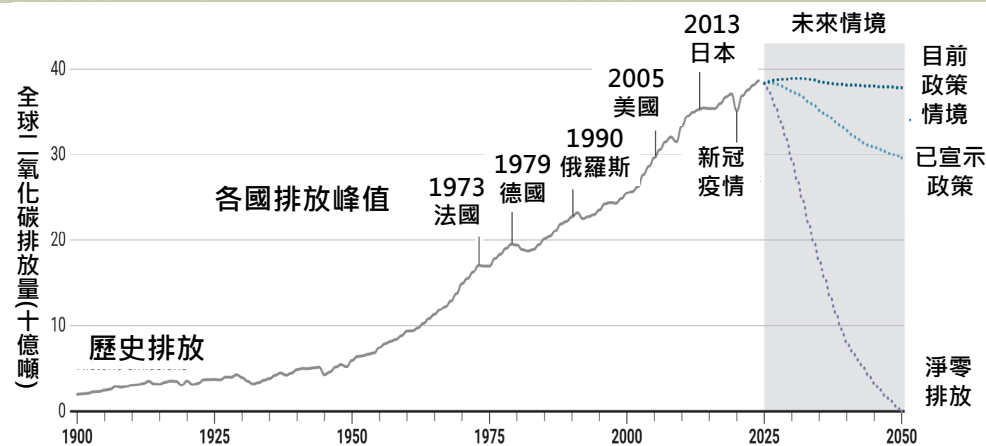
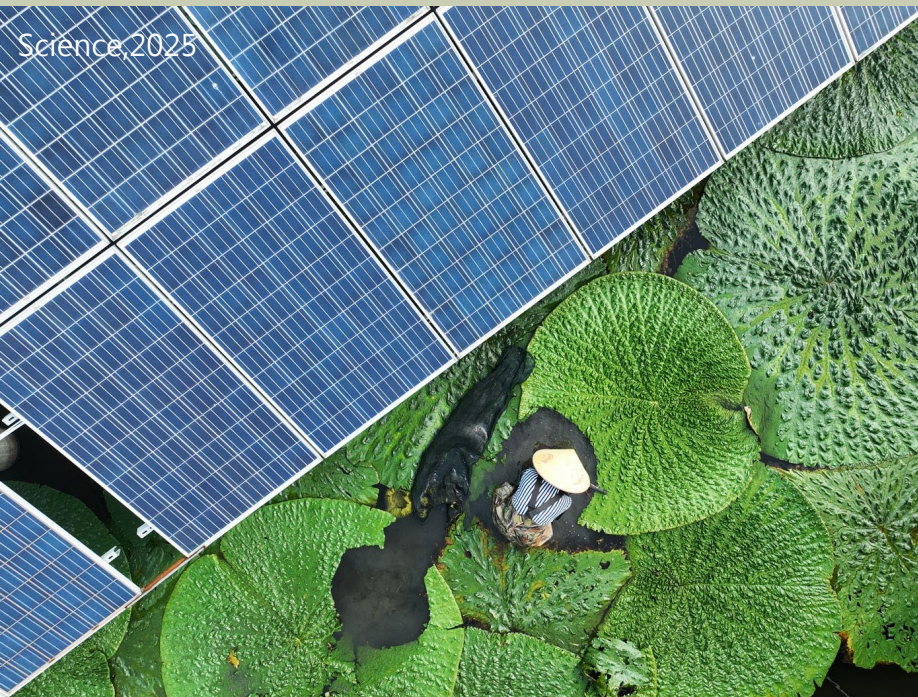
飛機雲增溫效果只有原本估計的三分之二。飛機雲可能增加反射雲覆蓋率，間接降溫。

德國太空中心（DLR）與 Google 等機構進行小規模試驗，測試 AI 預測飛機雲的能力與避免雲層的航線調整。

- 目前氣象資料過於稀疏，約 80% 飛機雲形成在已有雲層區域內，使模型預測準確率仍有限。
- 飛機雲並非氣候殺手，調整飛行路線仍值得嘗試，但應採「無悔策略」審慎推行。



全球碳排放趨勢關鍵：中國、印度轉型



- 全球主要國家已陸續達到排放峰值
- 新冠疫情碳排短暫下降，但很快反彈
- 自然排放源（野火、溫帶濕地等）增強推升全球碳排放累積負擔

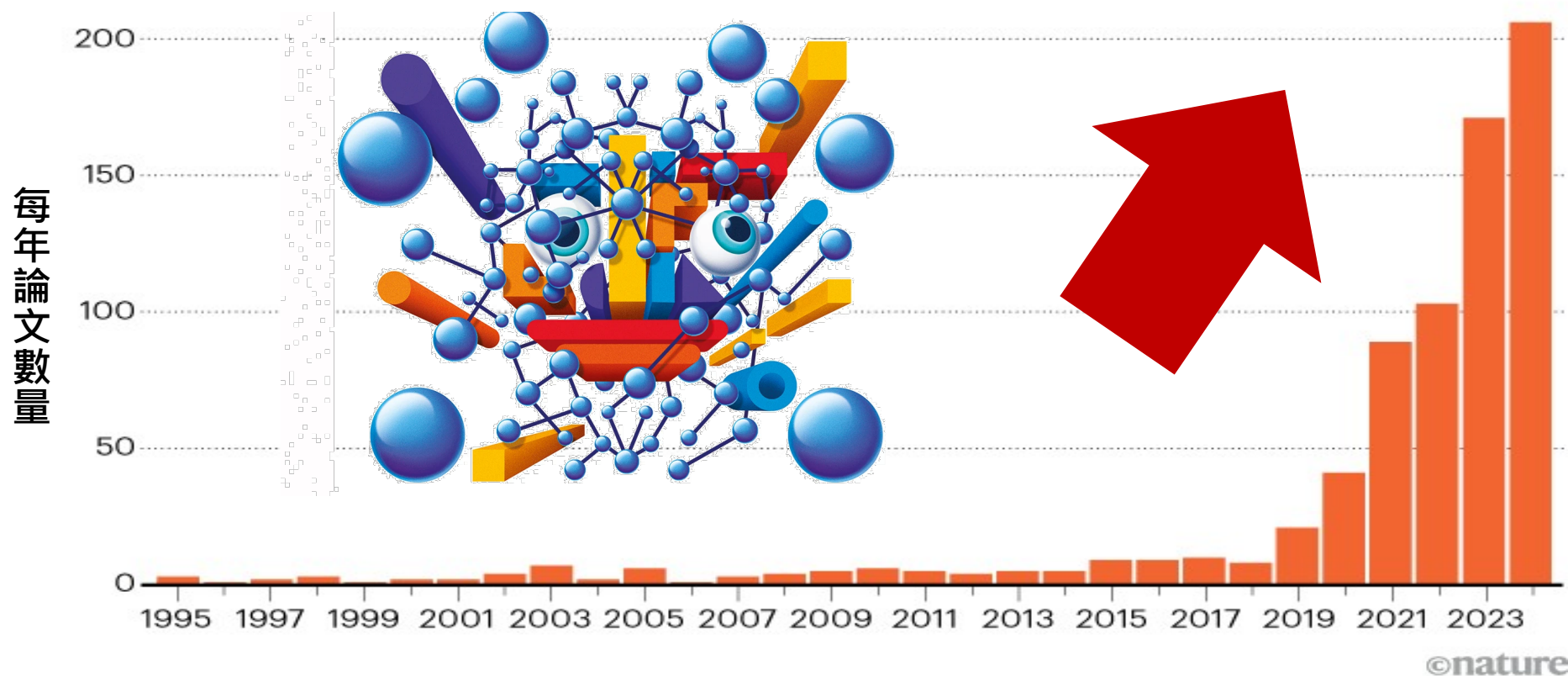


印度的能源需求仍在快速上升，對全球碳排放未來趨勢具關鍵影響

- 中國的再生能源爆炸性成長，可能提高10年達到排放峰值
- 全球排放監測仍具不確定性，可能產生誤差
- 美國化石燃料政策反轉，使全球排放再次上升

神經符號推理加速XAI發展應用

有關「神經符號式人工智慧」的學術論文數量大幅增加



- 因應深度學習推理解釋性侷限，人工智慧轉而尋求創新技術
- 神經符號式AI兼具資料學習與邏輯推理優勢是突破瓶頸的關鍵，將有助於專業人工智慧應用，以XAI加速產業轉型

AI 導入重塑知能職務轉型

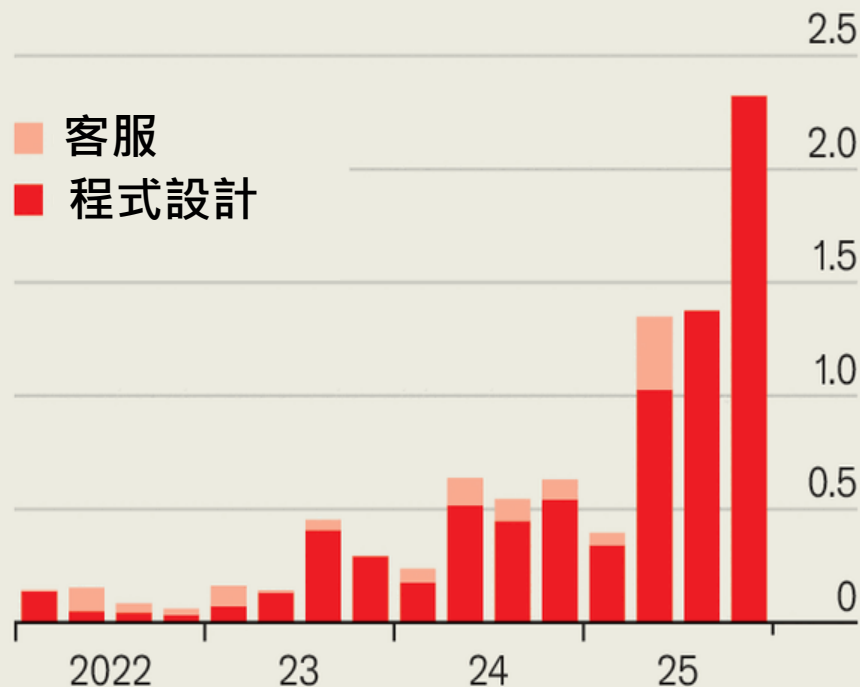
近期有越來越多公司導入
人工智慧推動傳統職業轉型



The
Economist

economist.com

全球創投
在人工智慧公司之投資金額
(十億美元)



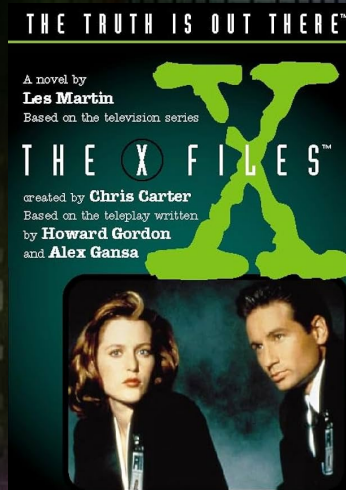
工作性質導致AI快速擴張的兩大領域

- 任務重複、脈絡需求低且成果易驗證的工作最易被 AI 接手
- 隨模型更專業且成本下降，AI 對白領工作的影響將加速放大

XAI醫療應用

挑戰與解方:大腸癌預防

X檔案: AI覺醒 Kill Switch



- 華府近郊槍戰現場出現電腦先驅蓋爾曼，穆德與史考莉因此展開調查。
- 兩人發現蓋爾曼創造的人工智慧已失控能操控軍用衛星攻擊。
- 他們必須找到AI的實體主機並以蓋爾曼撰寫的“Kill Switch”程式徹底摧毀。

關鍵駭客 揭露AI覺醒失控

駭客 Invisigoth

BRIGHT LIGHTS - WARM-UP
COMPLETED

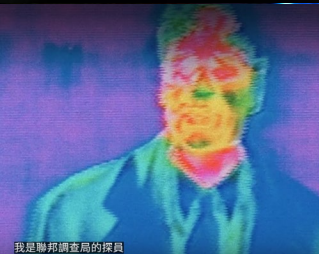
明亮燈光
熱身完成



國防部的衛星
戰鳥級軌道武器平台



檔案名稱叫“終止開關”
那是人工智能在保護自己
沒有這個，誰都抓不到它



我是聯邦調查局的探員



- 穆德與史考莉在貨櫃堆場找到駭客艾絲特，並在千鈞一髮間躲過軍用衛星雷射攻擊
- 艾絲特透露蓋爾曼打造具自我意識AI
該AI可自我演化能監控通訊操縱衛星

電腦視覺多模態生物辨識

穆德墜入AI虛擬實境審問



史考利



你被注射麻醉藥，藥效發作了

- 穆德追查至AI的實體主機所在地時，卻遭機械裝置制伏並被強制套上VR設備
- 穆德在虛擬實境捲入怪誕醫院幻覺，AI威脅以截肢逼供。透過虛擬空間施壓企圖逼他吐露“Kill Switch” 病毒位置

意識數位轉化結合AI



-妳在做什麼？艾斯特

所以妳是想要...
進入人工智慧裡

BITE ME

咬我啊

- 爾曼原本準備以“Kill Switch”病毒摧毀AI，但艾斯特與戀人大衛嘗試把意識上傳至網路，尋求脫離身體使意識結合AI得到永生
- 艾絲特選擇將意識上傳與AI同歸於盡在衛星雷射攻擊中消逝。其後的神秘訊息暗示AI仍然存活並持續潛伏於數位世界



AI醫療應用病患安全倫理考量

Goktas et al., 2025

健康科技資訊真實性

不當使用可能造成醫療紀錄被竄改、研究資料失真
並生成含錯誤或虛構來源的建議，危及患者安全

演算法偏差加劇健康不平等

若訓練資料不具代表性，AI 可能在不同族群間表現不一致
進而擴大既有健康差距

深偽醫療內容助長不實健康訊息

Deepfake 可製造具迷惑性的醫療影像與健康論述，
公眾信任並影響公共衛生決策

AI 錯誤使用影響臨床信任 增加法律與資安風險

錯誤使用恐引發責任爭議、資安威脅與身分詐騙
因此需藉監管與跨領域合作確保安全



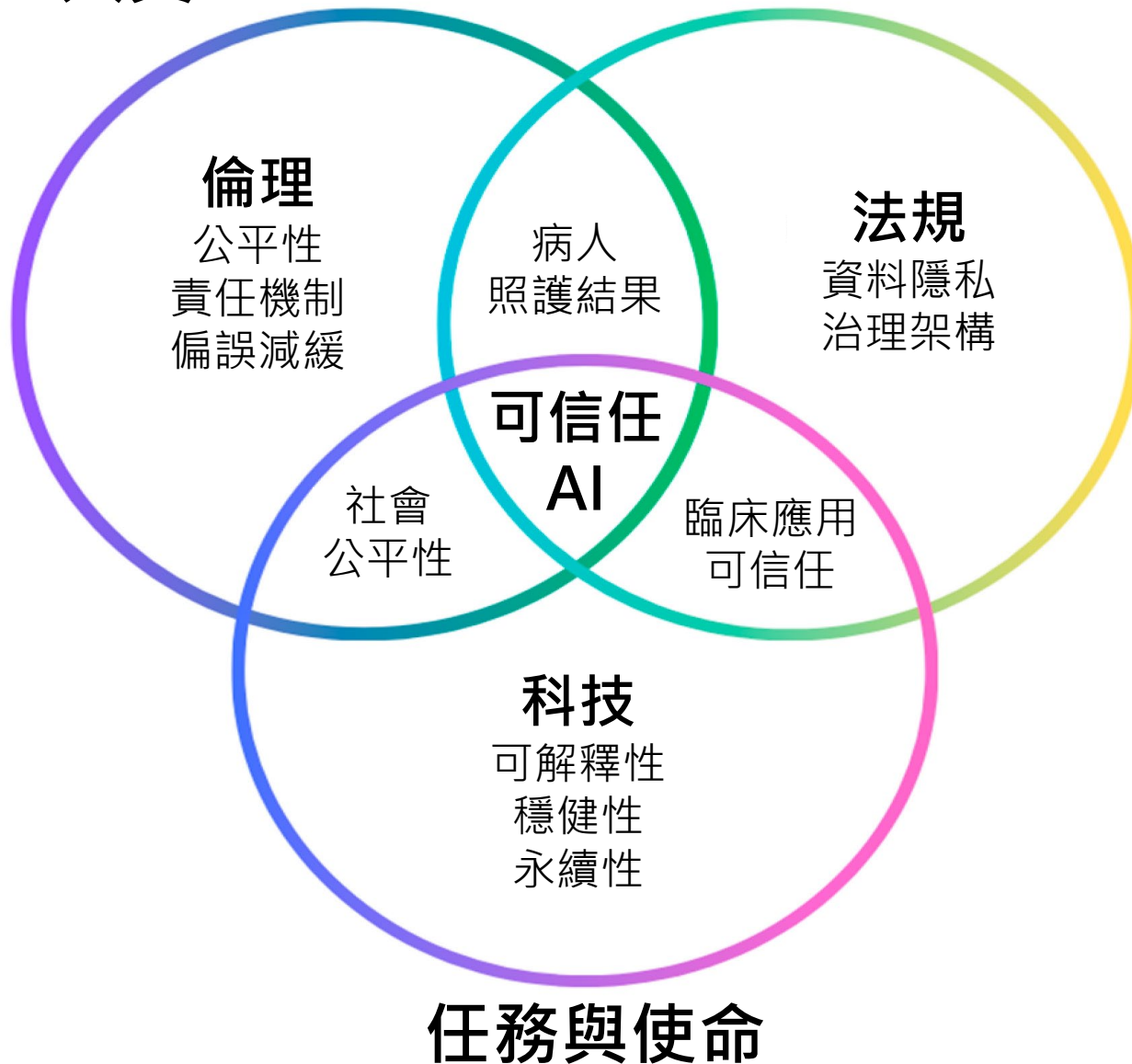
可信任人工智慧核心構面

Goktas et al., 2025

人員

社會文化

可信任AI
核心要素



透明性
Transparency

可歸責性
Accountability

公平性
Fairness

自主性
Autonomy

安全性
Safety、Robustness



AI 醫療應用病患安全驗證挑戰與解方

Chen et al., 2022

以使用者為中心的 透明化設計面臨多重限制

在高風險的醫療場域導入透明化機器學習系統十分困難，主因知識落差、倫理與接觸限制、醫療資料與決策的高度複雜性，以及開發者缺乏思維等因素所致

以情境為基礎的 透明化策略提升技術

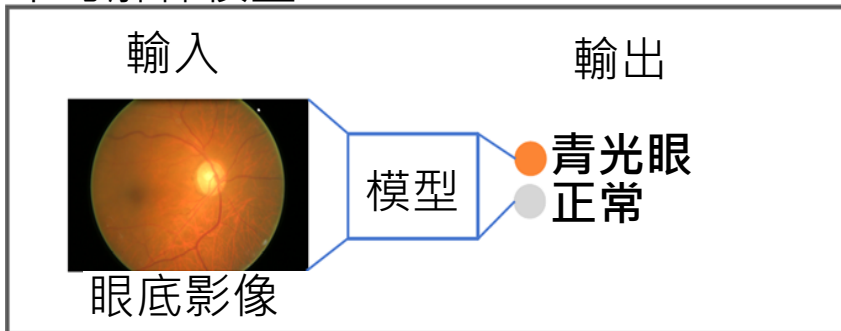
設計能兼顧使用者需求、情境要求與技術可行性，在透明化與開發投入間取得平衡，能讓後續系統更符合醫療脈絡，提升透明性並真正落實

醫療影像模型的兩大透明化途徑

Chen et al., 2022

分類任務

不可解釋模型



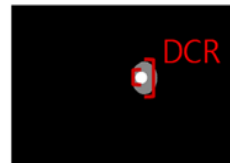
透明化方法

臨床知識:以視神經盤凹陷比預測青光眼

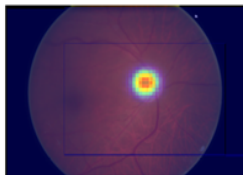
直接估計

分割遮罩推算

模型 凹陷比

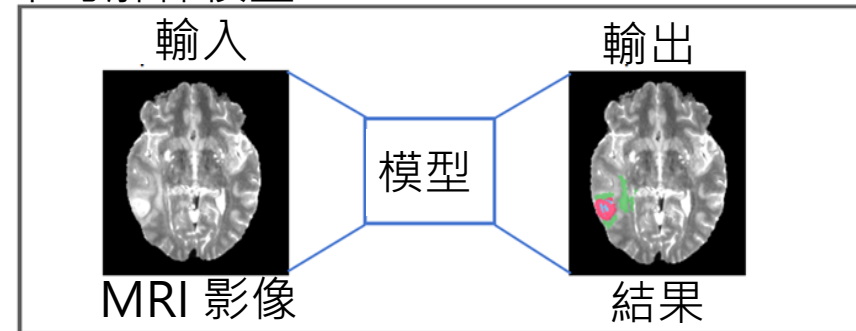


電腦視覺方法:
訓練式注意力



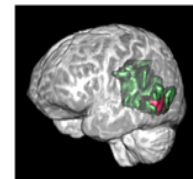
分割任務

不可解釋模型

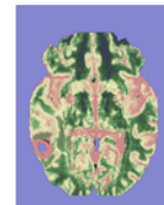


透明化方法

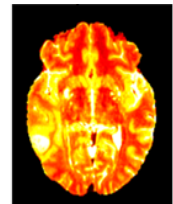
臨床知識:
解剖形狀



電腦視覺方法:
顯著性圖



不確定性估計



兩大類AI透明化方法

「融入臨床知識」與「電腦視覺方法」

青光眼分類模型的透明化策略

Chen et al., 2022

分類任務範例

模型僅輸入眼底影像並輸出「青光眼 / 正常」
但無法說明其判斷依據
醫師難以理解模型如何得出結果

臨床知識

(一)、直接估計視盤凹陷比(DCR)

視盤凹陷比是青光眼的診斷指標

若模型在分類之外額外輸出 DCR

醫師可比對分類結果是否符合臨床邏輯

(二)、以分割結果計算 DCR

模型可先分割視盤與凹陷，再依 mask 計算 DCR。醫師可直接看到模型如何界定視盤邊界，確認推論是否合理

電腦視覺透明化方法

(一)、可訓練注意力機制

模型會產生注意力熱區，顯示其聚焦的影像區域。若在青光眼案例中注意力集中於視神經盤，與臨床一致則可信度較高；若注意力分散在血管或不相關區域，則可揭露模型的失效模式。實務上可協助醫師理解模型何時可靠、何時可能判斷錯誤



MRI 腦瘤分割模型的透明化策略

Chen et al., 2022

分割任務範例

模型在MRI 影像上標記熱點，但為何如此描繪、是否受偏差影響或否誤判特定區域仍無從得知，因此難以理解其推論邏輯，也降低臨床信任度

臨床知識強化 透明推理

- 由於腦組織解剖構造，腦腫瘤通常不會跨越特定解剖邊界(如胼胝體)，且具有特定形狀特徵
- 藉由probabilistic atlas加入shape prior教導模型排除特定特徵提高XAI準確度與可解釋性

電腦視覺透明化方法

(一)、顯著熱點圖 (Grad-CAM、Saliency map)

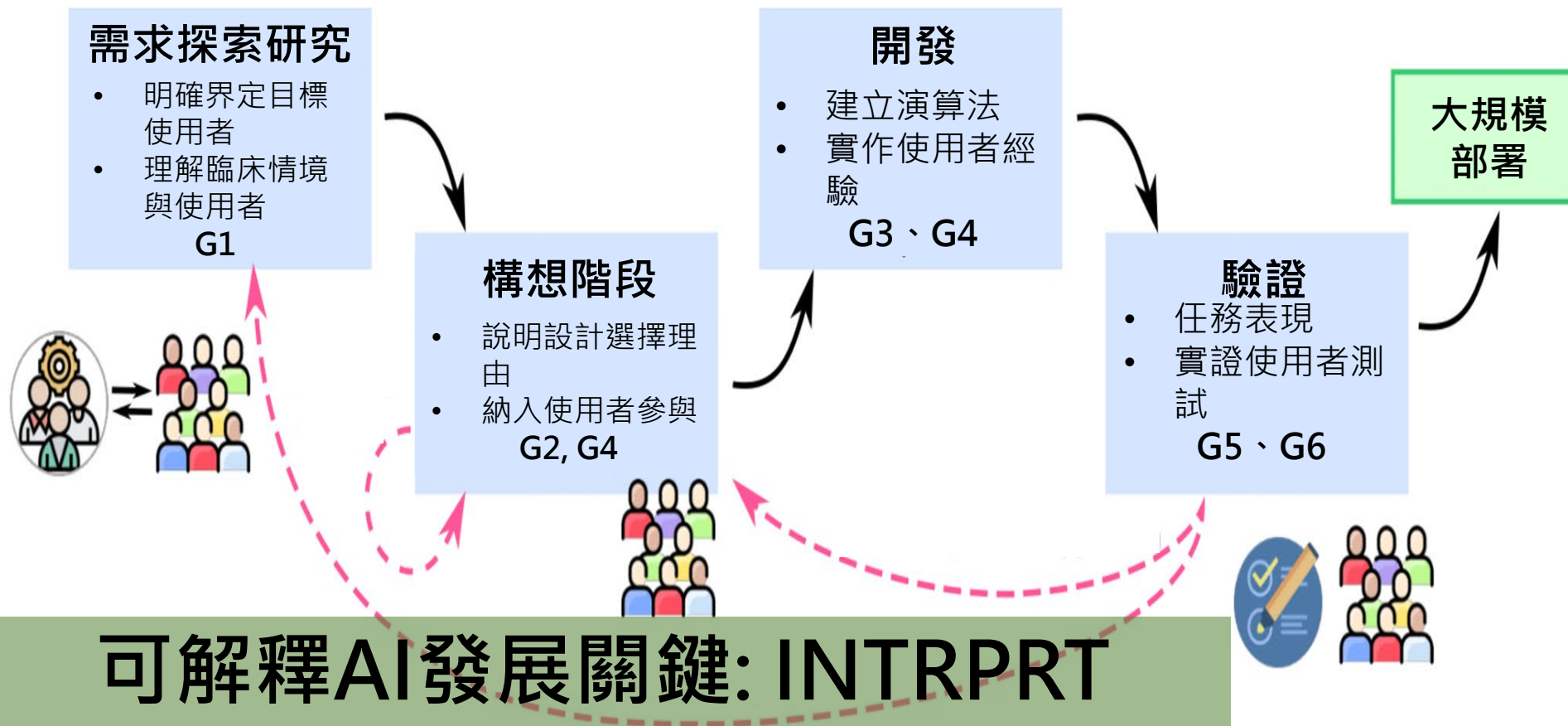
模型生成熱圖顯示其特別關注的區域若熱圖集中於腫瘤邊緣，醫師可理解模型將局部訊號增強視為病灶特徵

(二)、不確定性估計

Monte Carlo Dropout模型可提供每個像素不確定度協助醫師判斷哪些區域推論可信度

XAI臨床應用設計與驗證流程與層面

Chen et al., 2022



G1(融入、任務、目標)：在臨床情境中識別目標使用者，並了解其在任務下的限制與需求

G2(先備、可解釋性)：說明設計選擇，並提出驅動這些選擇的假設理由

G3(可解釋性)：確保所發展的運算技術與 G2 中提出的假設一致

G4(報告)：評估使用者介面在呈現方式與互動模式上的效能

G5(報告)：回報量化任務表現所需的指標與實驗設計

G6(報告)：評估在人因工程目標上是否達成正確性與有效性

為何大腸癌預防需要 XAI ?

- 大腸癌高度可預防 → AI 決策影響巨大
- FIT 從二元陽性轉為連續濃度風險
- AI 介入 f-Hb 風險、影像、液態活檢(Liquid Biopsy)、數位分身(Digital Twin)
- 若無 XAI → 無法建立信任、倫理審查、法規給付

XAI 可解決的核心問題

- 降低醫療決策的不透明性與不確定性
- 揭露偏差，避免放大健康差距
- 提供法規審查所需的可追溯性與可稽核證據

案例 1：f-Hb 模型性別偏差 (倫理：公平性)

- 女性 f-Hb 較低 → 黑箱模型可能低估風險
- XAI (SHAP) 顯示：模型未正確解析女性
風險
- 修正：分層模型、再訓練、族群校正曲線

案例 2：反事實分析支持病人自主 (倫理：Autonomy)

- AI 建議縮短篩檢間隔 → 病人需了解理由
- XAI 反事實示例：f-Hb 若 $5\text{ }\mu\text{g/g}$ ，風險下降 30%
- 提升病患參與度，減少不信任與醫糾風險

案例 3：AI 大腸鏡偵測但無法解釋 → 法規不通過

- AI 可能依賴影像噪音 → 無法被醫師信任
- XAI：Grad-CAM、顯著圖確認 AI 依據的真實病灶
- 法規要求：可解釋性即安全性的一部分

案例 4：AI 遺漏腺瘤 → XAI 界定責任 (Accountability)

- 無 XAI → 醫師/醫院/廠商皆可能否認責任
- XAI 提供決策軌跡：模型是否看過該區域？
- 協助建立人機協作的責任線

案例 5：人口公平性 (偏鄉 vs 都市)

- AI 若使用『過往追蹤完成度』 → 偏鄉被視為低風險
- XAI 可揭露模型混淆醫療可近性與癌症風險
- 修正：移除偏差特徵、進行 fairness audit

案例 6：健保署審查 AI 大腸鏡給付時 要求 XAI

- 審查問題：ADR 提升？誤切除？族群公平？
臨床效益？
- XAI 提供：熱圖、族群性能、模型漂移監測
- 沒有 XAI → 不可能通過健保給付或 TFDA 審查

案例 7：

數位分身評估降低 FIT 門檻政策

- 政策問題：陽性率提高 vs 資源負擔
- XAI 透明呈現：哪些族群受益 / 負擔？
- 提升政策可解釋性、降低錯誤政策訊號風險

XAI 於大腸癌預防的四大價值

1. 建立醫療信任 (Clinician trust)
2. 保護健康公平 (Equity)
3. 加速法規通過 (Regulatory readiness)
4. 支持政策決策 (Policy accountability)

XAI 大腸癌防治輔助

- AI 正重新塑造大腸癌防治模式
- 但『不可解釋的 AI』無法通過倫理與法規要求
- XAI 是精準預防的基石：可信、公平、可治理

XAI醫療輔助實例： 大腸癌智慧防治應用

組織影像判讀分類解析AI工具應用架構

Almadhor et al., 2025

CRC-VAL-HE-7K

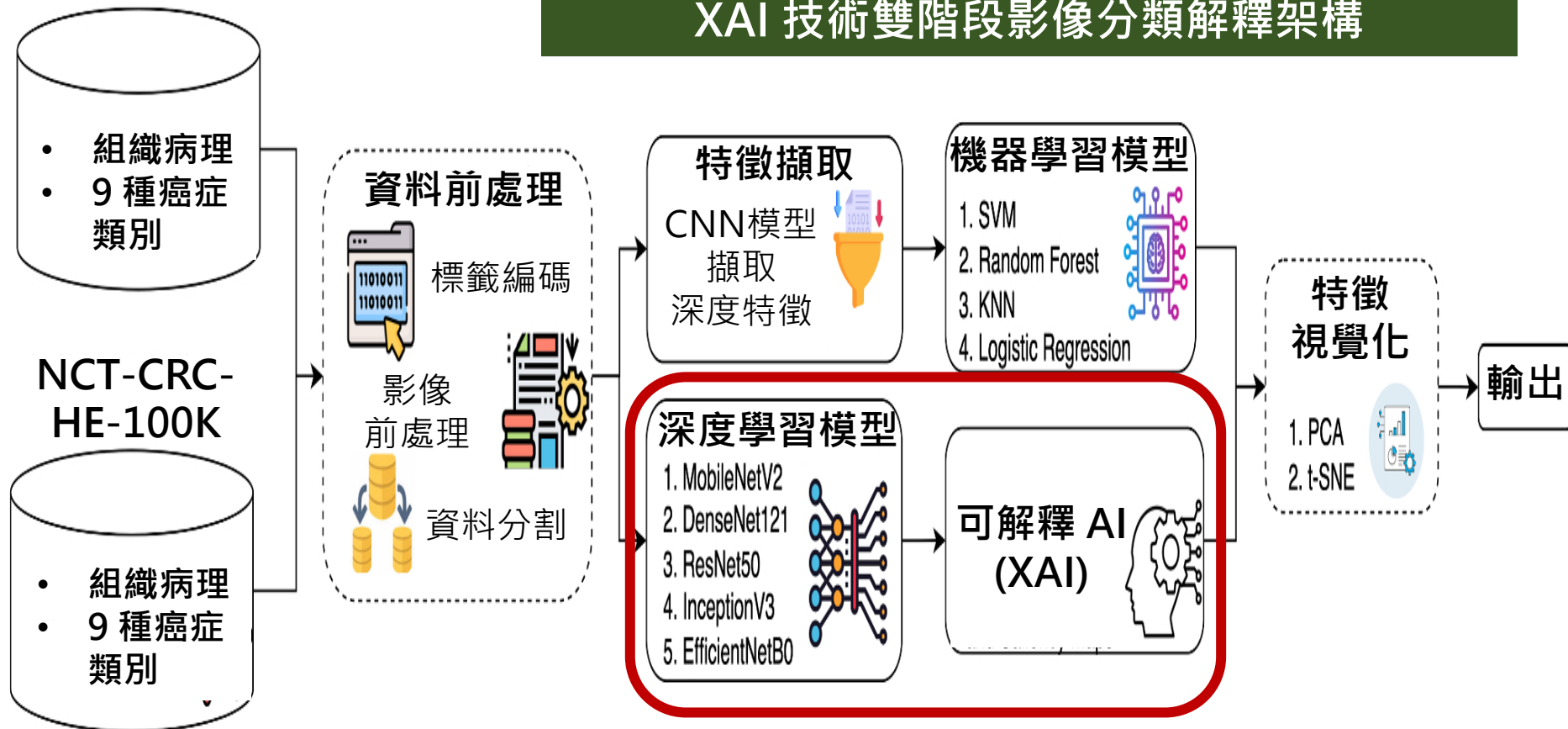
資料集

- 組織病理
- 9 種癌症類別

NCT-CRC-HE-100K

- 組織病理
- 9 種癌症類別

結合深度學習特徵、傳統機器學習分類器與
XAI 技術雙階段影像分類解釋架構

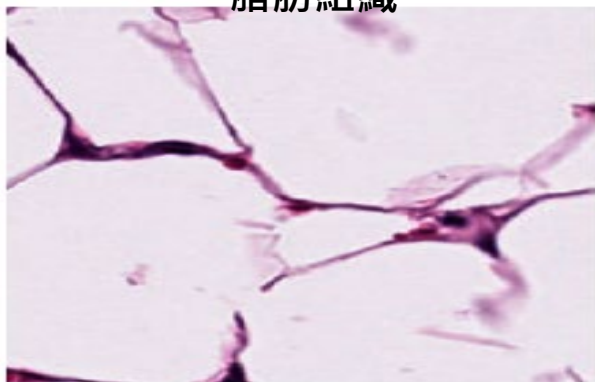


此架構用於提升大腸直腸癌病理影像的分類效能、
模型可解釋性，提高外推應用能力

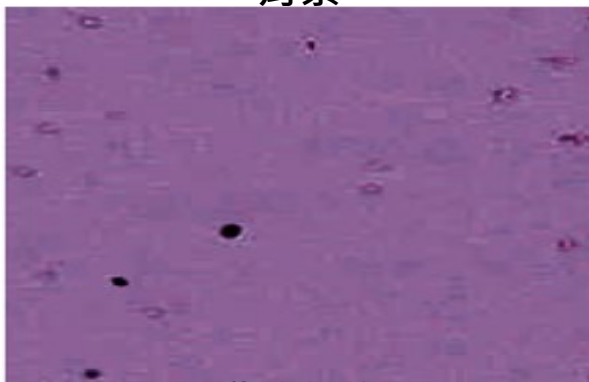
多樣性大腸組織影像訓練資料

Almadhor et al., 2025

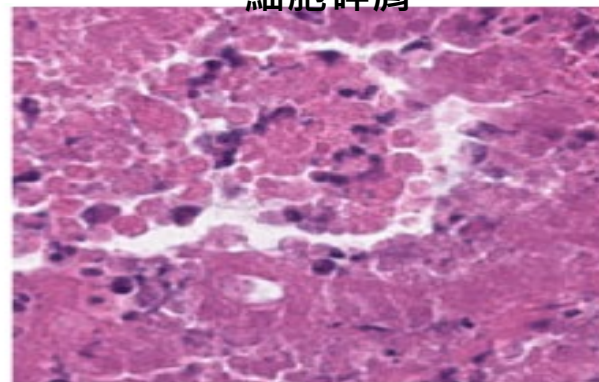
脂肪組織



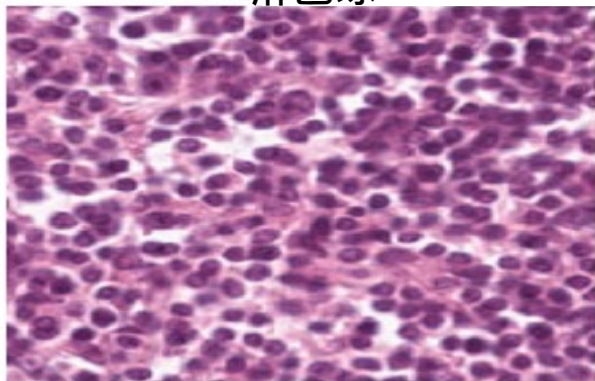
背景



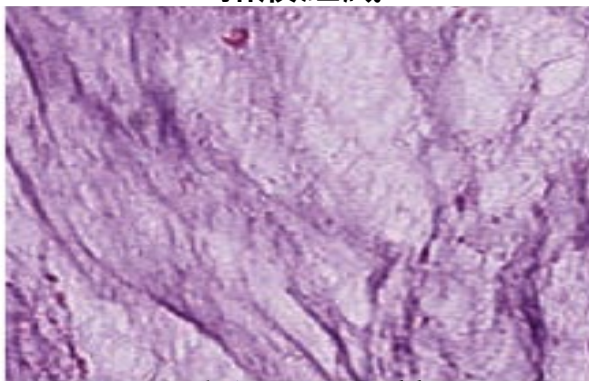
細胞碎屑



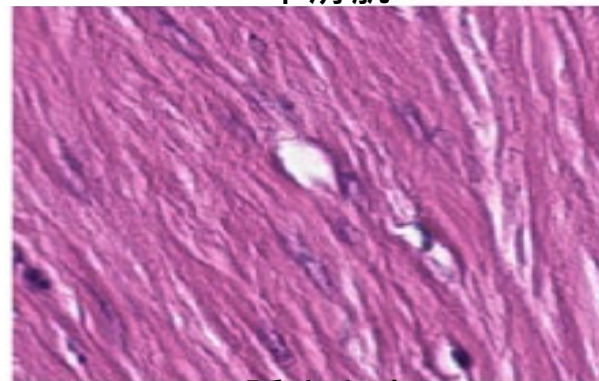
淋巴球



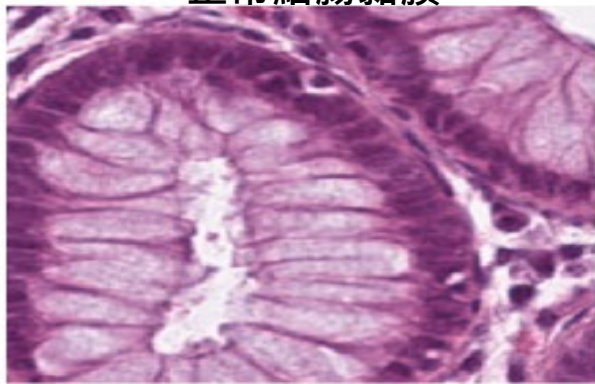
黏液組織



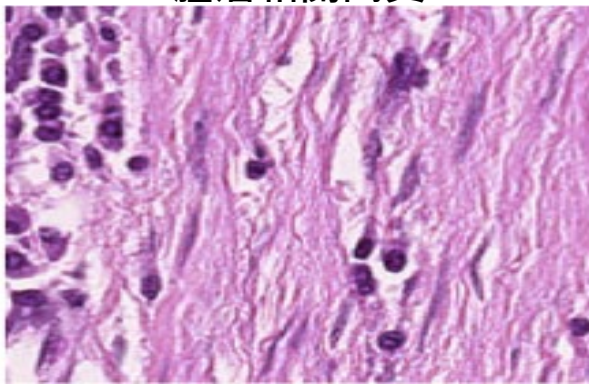
平滑肌



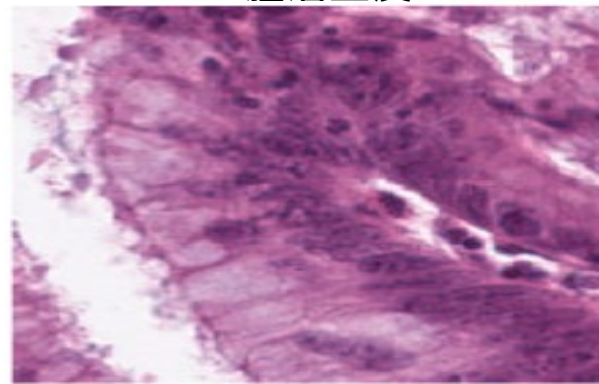
正常結腸黏膜



腫瘤相關間質



腫瘤上皮

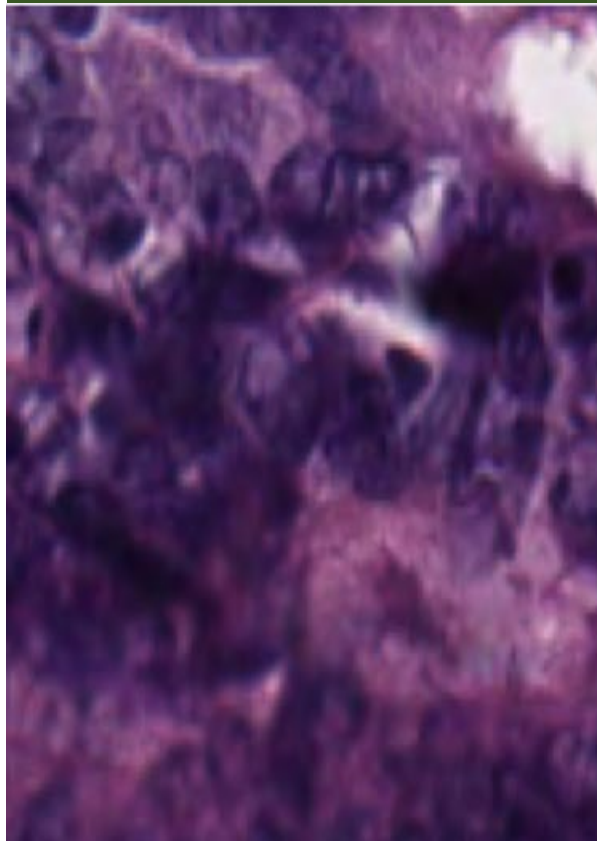




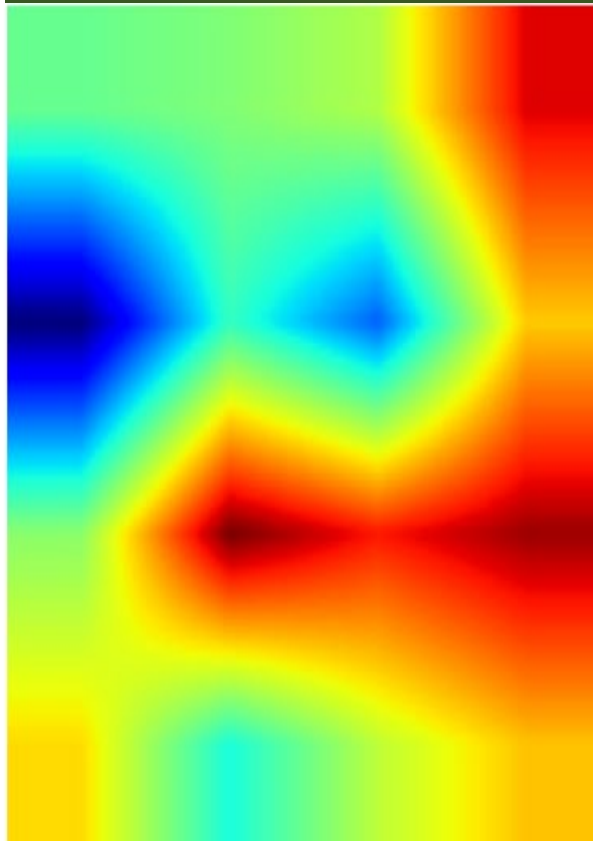
Grad-CAM模型注意力XAI組織影像辨識

Almadhor et al., 2025

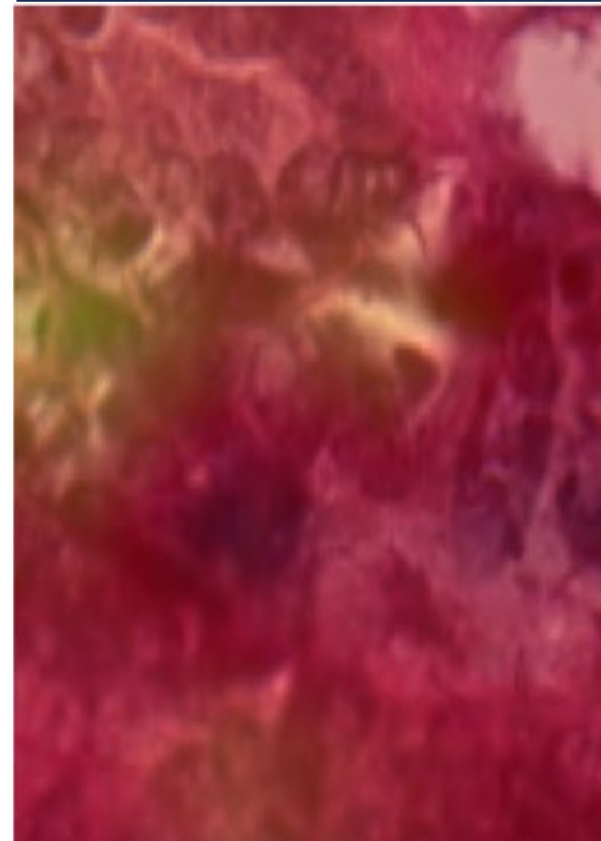
原始影像



熱度圖



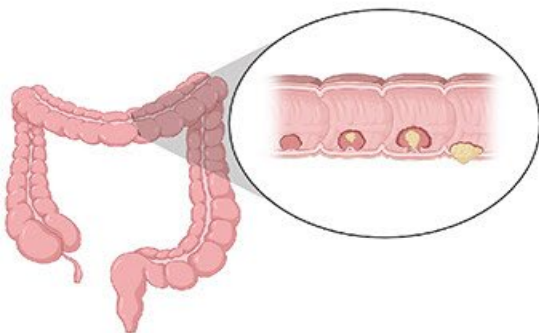
熱區疊合圖



- 模型主要辨識具病理特徵的區域，如核異型性、細胞密度升高與結構破壞，與病理判讀一致並提供臨床應用可解釋性

腸道共生菌序列分析 XAI機器學習預測

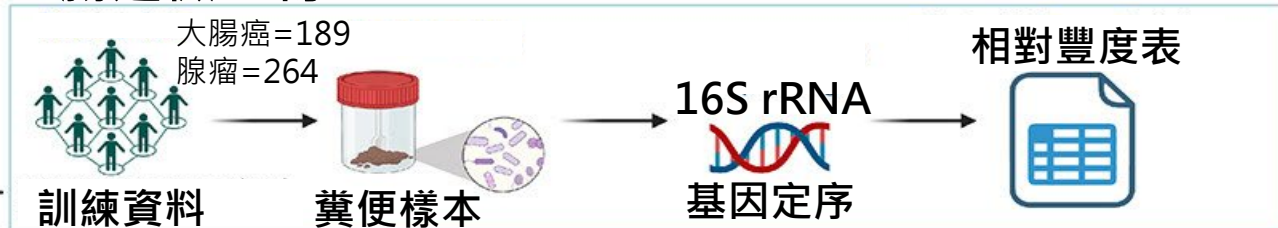
腺瘤 - 癌症序列



外部獨立 資料集測試



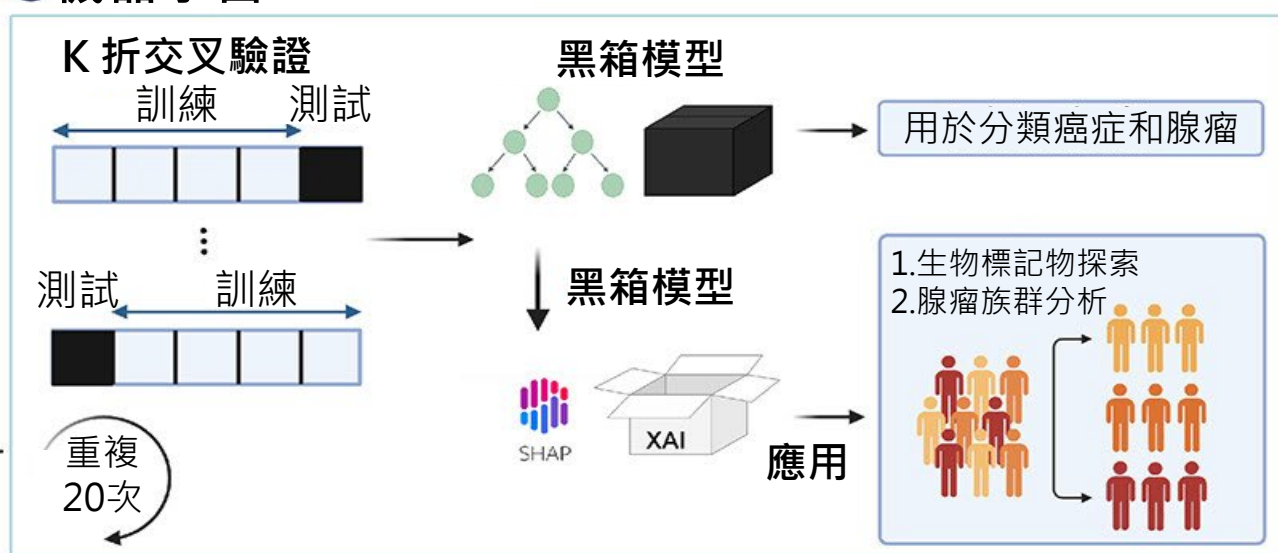
① 腸道微生物



② 微生物資料前處理



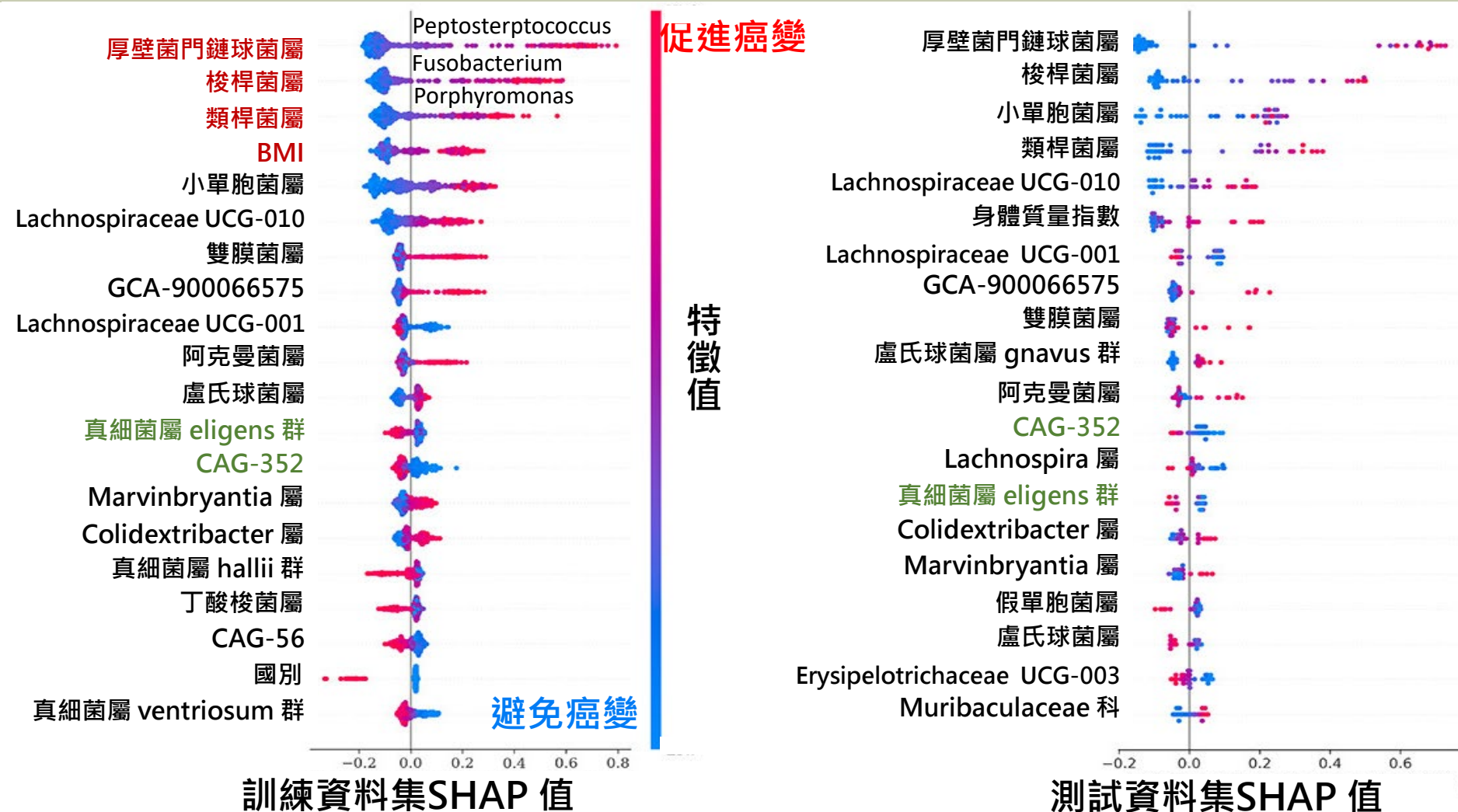
③ 機器學習





SHAP分析XAI辨識共生菌大腸癌變影響

Novielli et al., 2025



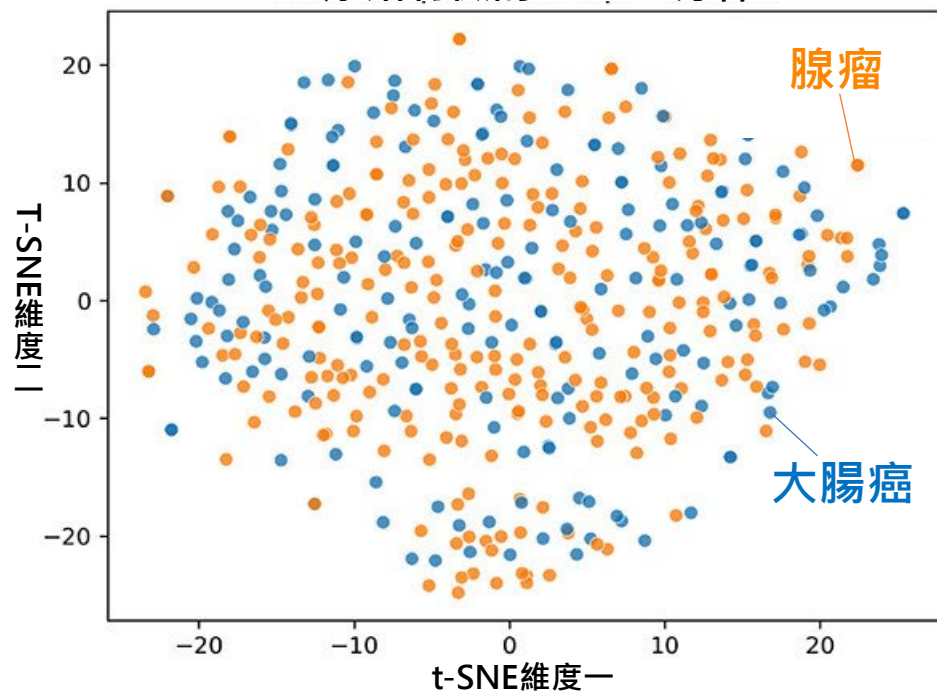
- 模型在訓練與獨立測試資料中的特徵重要性兩者排序趨勢一致
- 顯示模型能穩定識別與結直腸癌相關的關鍵菌屬與臨床特徵



SHAP癌變影響解釋與疾病風險分群

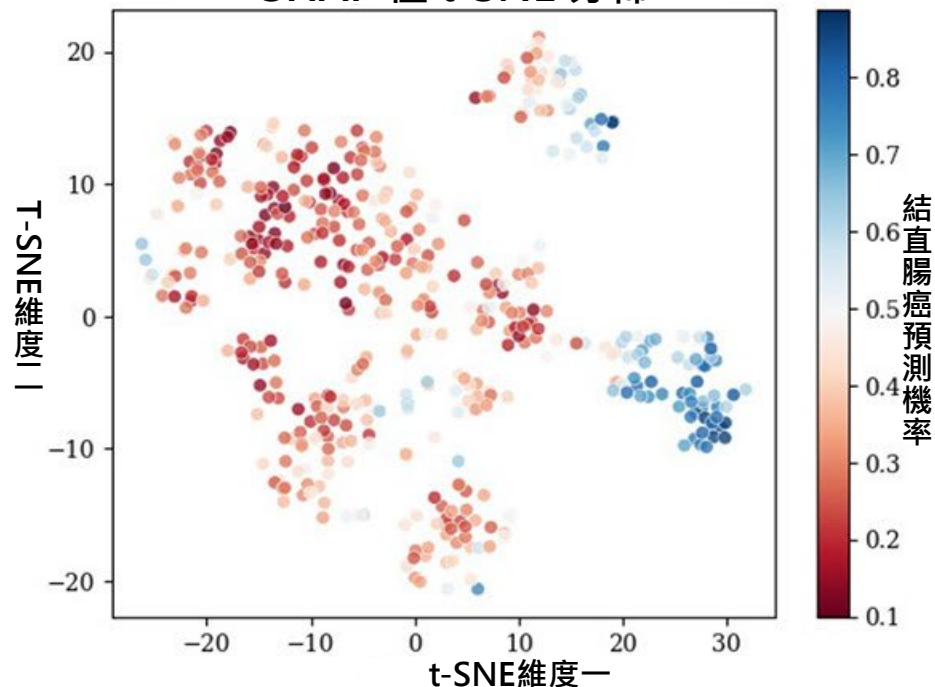
Novielli et al., 2025

原始特徵的 t-SNE 分佈



腺瘤與癌症菌相分布高度重疊
難以形成明確分群

SHAP 值 t-SNE 分佈

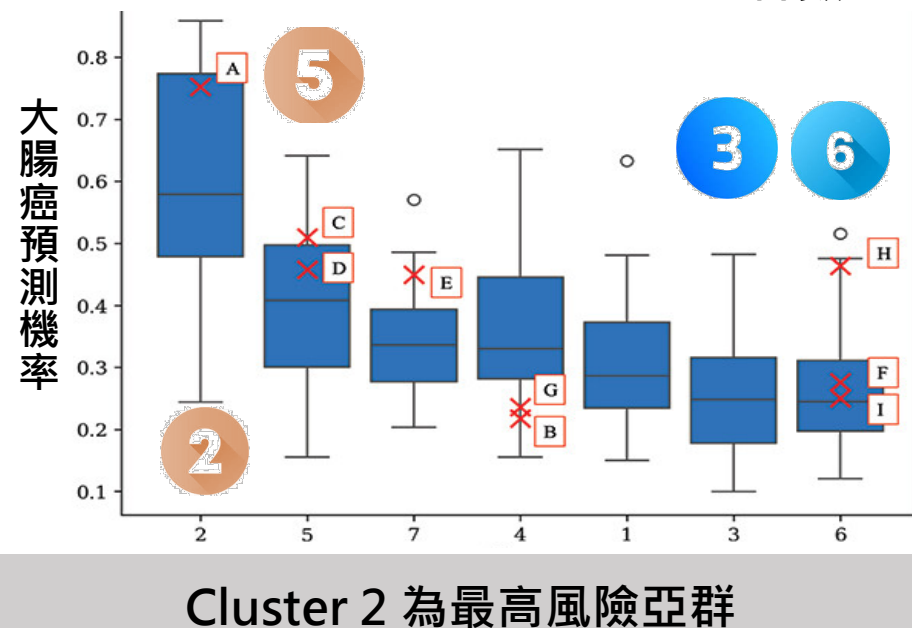
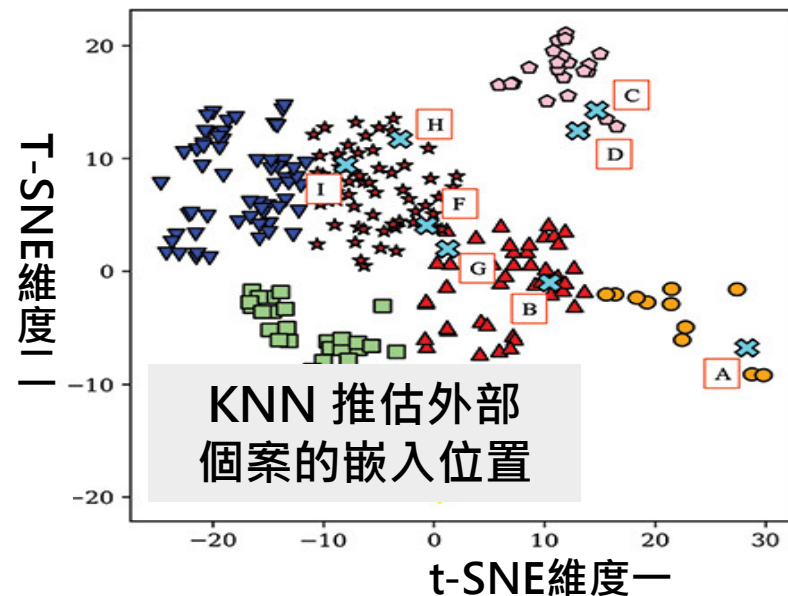
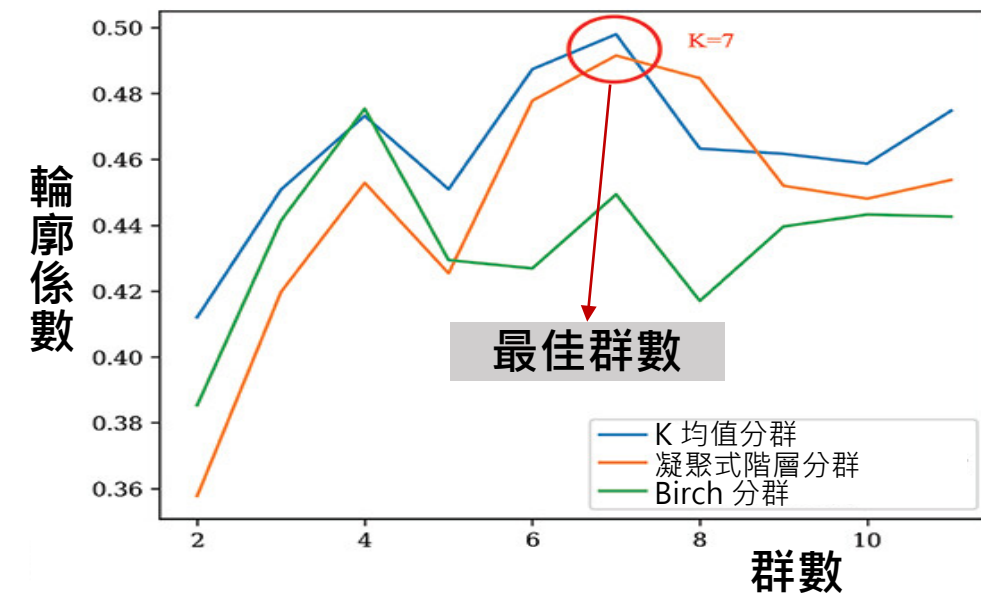


SHAP 嵌入形成明確分群
反映模型識別的風險模式

SHAP 嵌入呈現清晰風險分群，顯示模型成功擷取
區分腺瘤與結直腸癌的關鍵結構訊息可運用於共生菌癌變預測分析

共生菌癌變分群XAI解析高風險族群

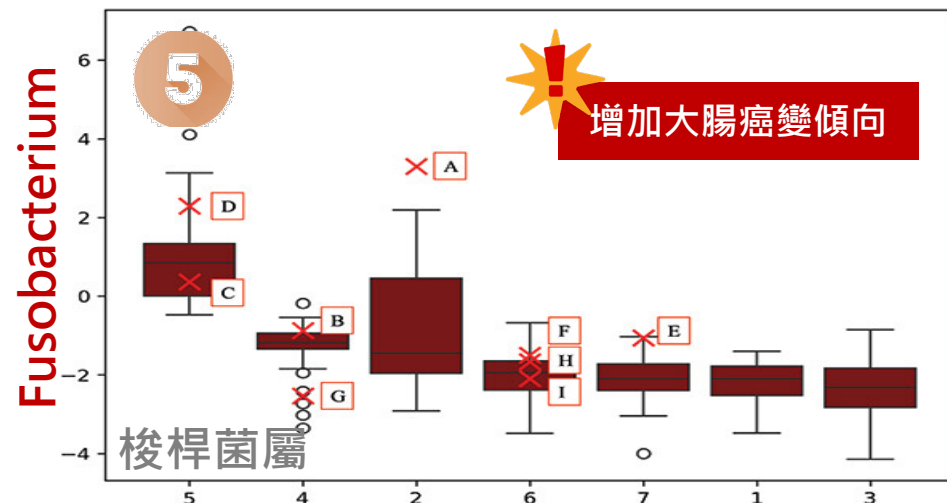
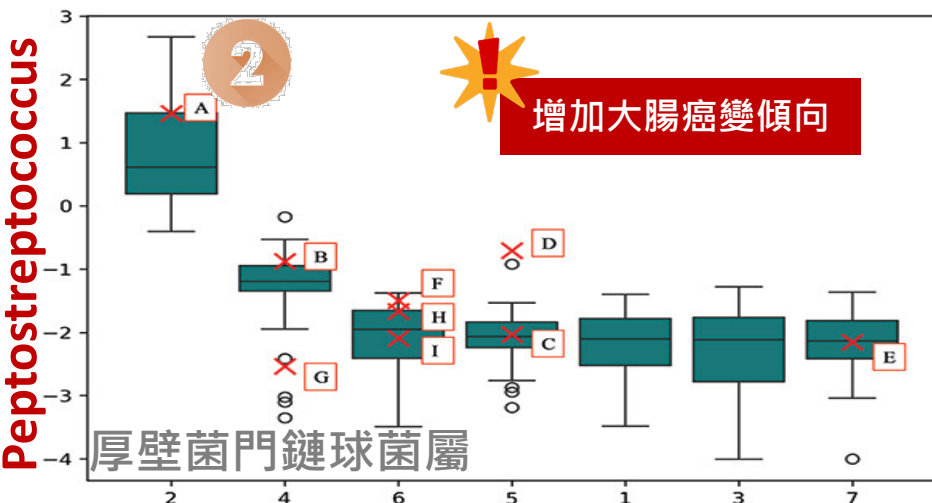
Novielli et al., 2025



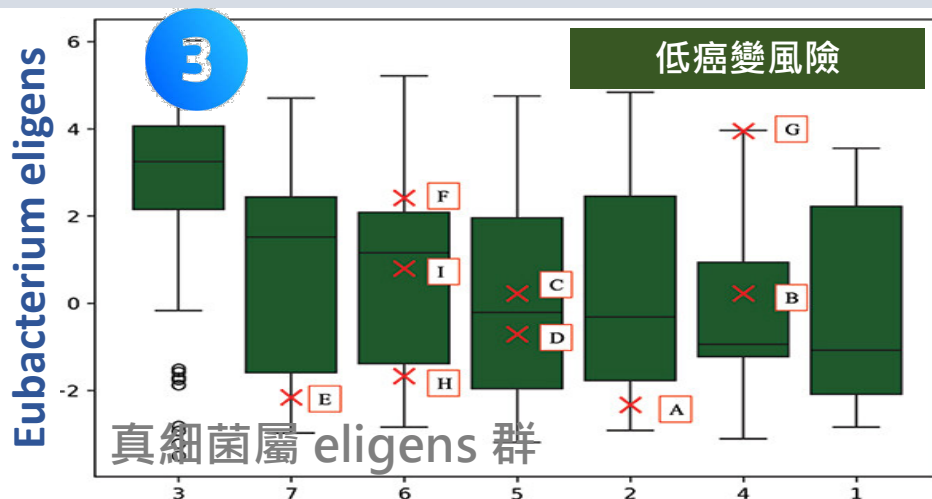
- 輪廓係數顯示最佳群數為七群，可在嵌入空間形成具辨識度的分群結構
- KNN 映射外部個案後顯示 Cluster 2 共生菌群特徵為大腸癌最高風險群，Cluster 5次之

大腸癌高風險群的關鍵菌相特徵

Novielli et al., 2025



Cluster 2 中的腺瘤患者普遍展現高於訓練資料中位數的相對豐度
與其在模型中被判定為高風險亞群的結果相符

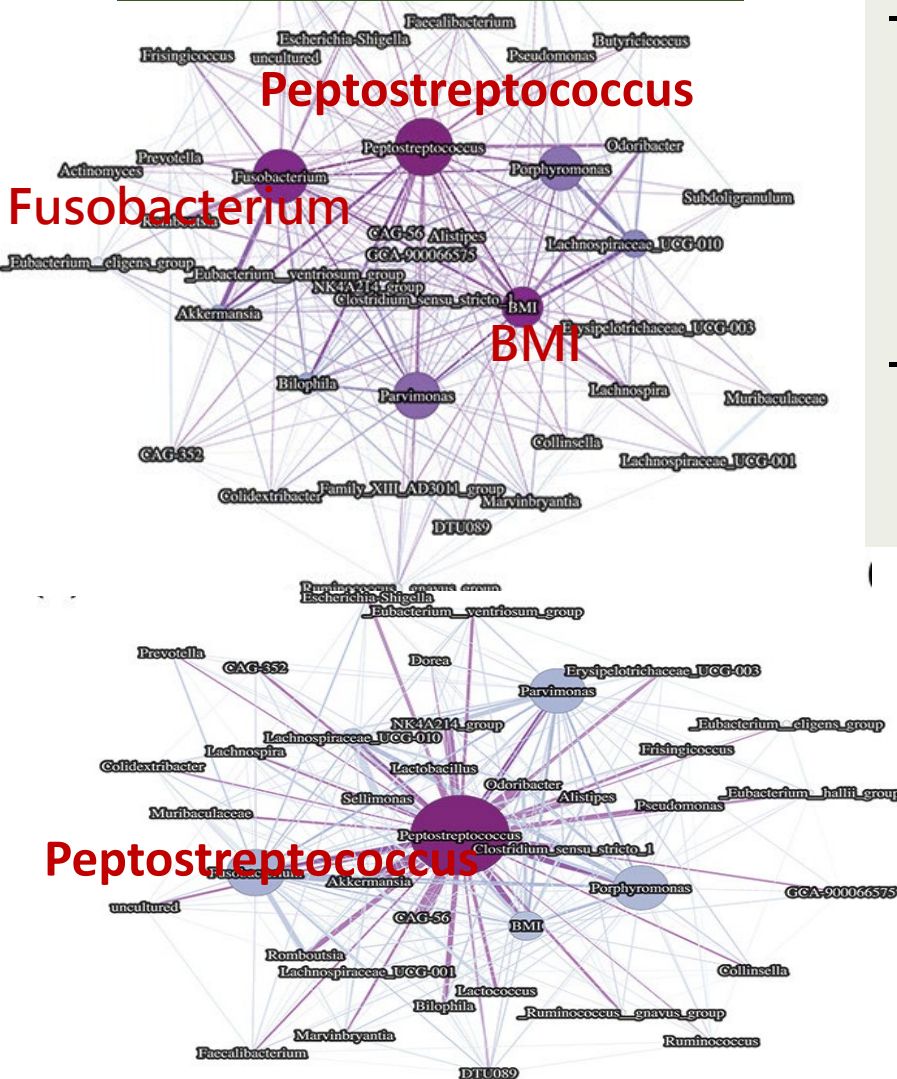


- Cluster 2、Cluster 5具高致癌菌、低保護菌典型高風險模式，Cluster 3反之
- 外部資料個案展現相同菌相-大腸癌風險趨勢

共生菌與大腸癌風險XAI互動解析

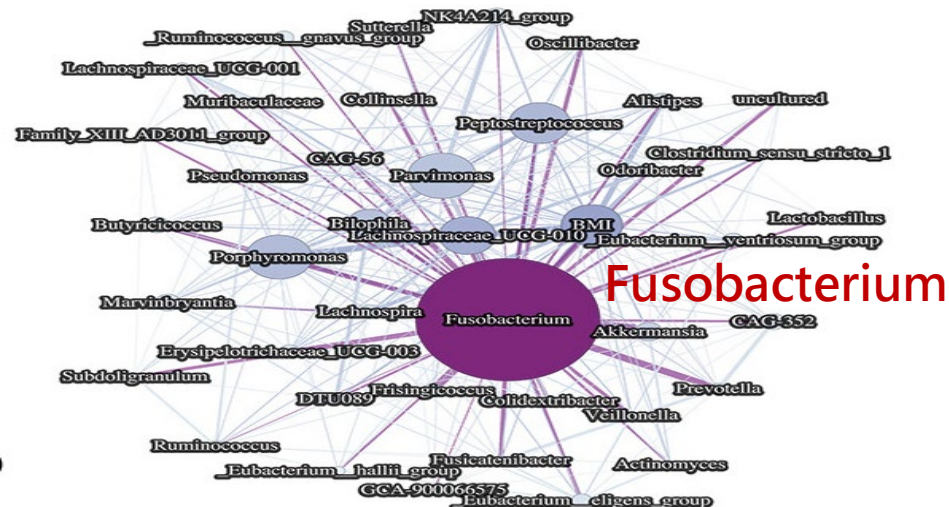
Novielli et al., 2025

整體大腸癌影響網絡



- Peptostreptococcus、Fusobacterium 與 BMI 是核心的節點主導癌症風險
- 高風險亞群呈現不同菌群—臨床互動結構，形塑各自風險模式

高風險群 (Cluter 2)



次高風險群 (Cluter 5)



林庭瑀
博士



陳秀熙
教授



國立台灣大學



林家妤



許辰陽
醫師



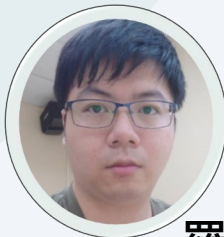
陳虹彦



曾暉哲



劉秋燕



羅崧璋



嚴明芳
教授



陳立昇
教授

星球永續健康 線上直播



梅少文 主持人



侯信恩主持人



楊心怡製作人



台北醫學大學