ZT

研百◀ 

靈進算感行機

|項與

教授李煜

新正

疾病

蛋白質/基因

蛋白質/基因

編

「AI技術超越人類智慧後,最明顯的影 響將體現在醫療領域。|入選時代雜誌「AI 百大影響力人物」的庫茲威爾在今年大膽預 測,人工智能在2032年後就能讓人類停止老 化,且未來AI與人類大腦連接後,可大幅增 進人類智力,治癒大部分疾病。

對於AI醫療發展抱有樂觀態度的專家不 在少數,2024年諾貝爾獎化學獎得主,藉自 行研發的Alpha fold 2 AI模型,攻克了預測 蛋白質複雜結構的世紀難題。三位得獎者展 現出強烈的信心, 宣稱模型的發明將會在 2036年治療人類所有的疾病,人類壽命有望 延長至150歲。

這一充滿無限想像的領域,吸引愈來愈 多的探索者加入。那麼真實情況到底如何? 《大公報》記者邀請香港高校相關的研究學 者,從藥物研發、癌症診斷、影像分析、居 家健康護理等角度,全面剖析香港學術界在 AI醫療發展方面的成績和現況。其中有學者 贊同AI醫療將給人類健康護理帶來轉折性的 改革,但認為這是一個逐步前進和變化的過 程,而非突變式的改革。

## 預測細胞形態框架圖 (MVAE) L1000 (LDM) **◆** 去噪 潛表徵 條件 →加噪 EC N(0,I)去噪UNet

▲李煜教授團隊設計的預測細胞形態框架(MorphDiff),可預測藥物在細胞上的作用,幫 助減少細胞實驗的操作部分,降低相關的時間消耗和成本

AI走進醫療①

AI時代的到來,將會為醫療行業帶來怎樣的奇跡? 美國AI企業Anthropic執行長Dario Amodei使用「壓縮

的21世紀 | 一詞來總結他對AI發展的樂觀看法。他說, 人類可能在未來5至10年取得相當於50至100年 的進展,例如治癒癌症、應付傳染病並找 出預防阿茲海默症的方法。

愈來愈多的研究似乎都在證明 着AI作為「醫療發展加速器」的 角色,香港中文大學計算機科 學與工程學系助理教授李煜 及其團隊,正與百圖生科 (BioMap) 進行相關的三 項研究,其中AI藥物協同 預測項目,可幫助預判 多個藥物組合使用的效 果,有望為醫生提供新 的治療方案,以及為新 藥研發帶來靈感。「在 藥物研發十年期間,大 概有三至四年是做藥物 篩選;而在AI的幫助下, 這個過程可縮短至一年左 右。」李煜說道。

大公報記者 湯嘉平(文) 林少權(圖)

傳統的新藥開發領域普遍存在所謂的「雙十 法則|,意思是指一款創新藥物的研發通常需耗時 十年、投入約十億美元,且成功率偏低。這使得許多罕 見疾病所需的特殊藥物,難以獲得足夠的研發資源

## 預測藥物組合效果 免人手測試

李煜教授表示,臨床上常需採用聯合用藥策略去治療疾 病,因為單一藥物往往效果有限。然而,哪些藥物的搭配能 起到一加一大於二的效果,哪些藥物的搭配又會導致效果 相互抵銷甚至變差,就需要科研人員逐一地去嘗試,「就 像鑰匙和鎖一樣,鑰匙是藥物,鎖是發生病變的蛋白 質。|科研人員需要製造出一把最適合的鑰匙去開這 個鎖,而這個鑰匙需要經過多種材料和不同的凹口 設計組合,打磨成千百種不同的模樣去開不同的 鎖,讓病變之處恢復正常。

「傳統的方法,就是我給你一萬把鑰 匙,你就一個個慢慢去試,總有一個能打 開。然後現在我們的AI藥物協同預測項 目,就是會直接告訴你,哪些藥物組合是 可行的,即哪些潛在的鑰匙是能開這個鎖 的,然後你直接去試就好。| 李煜解釋, 這就是藥物研發必須經過的藥物篩選階 段,該階段一般需要三至四年,而AI藥物 協同預測項目則有望將時間縮短至一年, 即整體的新藥研發時間,可能從原先的十 年縮短至七年左右,甚至是更短的時間。

此外,經過實驗發現,AI藥物協同預 測項目預測藥物組合效果的準確率,會比 現有技術提升10%至30%,即使是未見 過的新藥也能準確地預測效果。

談及該項目所運用的AI大模型,李煜 介紹,該AI大模型是以蛋白質語言為核心 的大模型,在預訓練階段學習多模態大規 模生物資訊(包括藥物、細胞系和疾病) 和圖結構,並進行微調,即針對特定的預 測任務進行模型參數的調整和優化。 (見 藥物組合協同作用預測模型框架圖)

## 發現逾300種胰腺癌藥物組合

至於蛋白質語言,李煜教授解釋,正 如人類擁有語言,生命世界也有一套由胺 基酸序列構成的「分子語言」,即蛋白 質。而人工智能(AI)領域的蛋白質語言 模型(PLM)展現出解碼這套語言的強 大能力,能夠精準預測蛋白質的結構與功 能。「我們根據這個蛋白質的結構,才能 去設計或者發現藥物。|

目前,該AI藥物協同預測項目已針對 胰腺癌的治療,發現300多個具有協同作 用的新組合。李煜透露,團隊現時也在用 此技術研究抗生素抗性藥物,以及包括治 療漸凍症在內的罕見病的藥物。對於疾病 數據稀缺、模型面臨「小樣本學習」的挑 戰,李煜指團隊也在想辦法去擴展數據的 來源,「不光是用這一個疾病的數據,我 們可以用很多不同疾病的數據,如果這些 疾病是相關的。比如說神經退行性疾病, 那麼不光是有漸凍症,其實還有一些其他 的跟運動神經相關的疾病,這些疾病的數 據都是有可能幫助模型去學習。 |

中大研發大模型 篩選藥物需時減至一年

藥物組合協同作用預測模型框架圖 使用「藥物─靶點相互作用模組」和「藥物─藥物相

預訓練模型 ( Pre-Trained Models ) 藥物A (新藥) **KPGT** 藥物B 細胞系 加權基因表現

疾病 RotatE (來自primeKG) 節點特性X

學 型 李 習 框 煜

,架教

準融團

一確預測2世間 開發的

周測多種藥: 用發的 AI 藥:

物規物

組模組

合的效果 合協同作

和用

圖預

結 測

醫療中出現

的錯誤,嚴重的 至上 可能影響人類的 健康性命,而用AI來主導藥物 篩選和研發,萬一誤判、或遭 何 遇算法黑箱問題,人類又要如 何甄別?李煜教授指出,藥物 澼 的開發過程中,無論是否用到 AI,都有機會遇到一些意外, 免 故在研發過程中,科研團隊還 **會通過動物實驗、臨床實驗等** 方式,想辦法解決安全性問

安全

法

\$ 7

直

驗

這些問

題。|

題。 AI算法黑箱,是指一種根 據輸入生成輸出,但不透露其 決策背後內部邏輯或推理過程 的AI系統。當我們將關鍵決策 交給AI,但它卻無法解釋「爲 什麼 | 時,就會爲人類帶來重

大安全風險。 李煜表示,若AI發現的分 子是天然界存在的,其實也並 不太存在所謂的黑箱問題 「黑箱問題,主要可能體現在 它自己去設計一些新的藥 物,如果這些新藥物在 沒有監管,以及沒有 比較詳盡的實驗之 前就使用的話, 肯定是有問題的; 但是如果進行了比 較詳盡的實驗的 話,就能避免

香港中文大學 計算機科學與工程學 系助理教授李煜及其團 隊,與百度參與發起創立 的生命科學大模型公司-百圖生科進行三項AI醫療研 究,其中AI藥物協同預測項目已 完成, AI細胞形態預測項目則於今 年9月發表,另一項研究正處於修改 論文的階段。李煜教授日前就AI細胞形 態預測項目接受《大公報》獨家專訪時 透露,該項目是由自行研發的AI大模型去 預測藥物在細胞上的作用,細胞的形態會出 現哪些變化,幫助減少細胞實驗的操作部

蛋白質/基因 藥物 疾病 原始圖結構A

蛋白質/基因

圖構建

(Graph Construction)

藥物A

疾病

損失

(Loss

藥物 疾病 優化圖結構A 多層感知器預測 嵌入向量 ( MLP Prediction ) (Embeddings) 細胞系 神經

藥物A

疾病

互作用模組」進行優化(Refine with DTI Module

& DDI Module )

經 網絡

疾病

協同作用 拮抗作用 優化 協同作用預測

藥物A 藥物B 最終嵌入向量 X

## 用AI預測藥物在細胞上的作用

降低 成本

AI與人類協作 短期內易取得突破

分,降低相關的時間消耗和成本。

一款新藥從研發到上市需經歷三個階 段:臨床前研究(藥物研究)、臨床研究及 藥物上市。臨床前研究主要是藥物篩選的階 段,包含藥物研究(例如標靶點確認、化合 物合成及篩選)、臨床前實驗(細胞實驗、 動物實驗)以及臨床實驗審批(IND)三個部 分。而AI細胞形態預測項目,可在細胞實驗階 段前,用AI去預測哪些藥物能對目標細胞達到 特定的效果,「甚至是之前還沒有見過的藥 物,也可以預測它的效果,省去傳統的實驗 環節。」李煜說:「如果通過AI已經預測到該 藥物作用於該細胞是不能產生預期效果

(hopeless)的,那可能就不做了。 與AI藥物協同預測項目所用到的模型不 同,AI細胞形態預測項目中的模型,並不涉及 預訓練和微調的階段。李煜解釋,因其總體 的數據量並不是特別大,而預訓練和微調AI模 型更適合支持大量的數據樣本。

像AI細胞形態預測項目中的AI模型,李煜 介紹,主要包含變分自編碼器模型和潛擴散 模型(見上方預測細胞形態框架圖)。當中 原理簡單而言,有點像教一個小孩怎麼去認 識世界,「比如說我們訓練一個模型,就讓AI 去識別這個物體,它是一隻狗,還是貓?給AI 一些圖片,然後給AI一些標籤,讓模型學習兩 者的映射關係。 |

問及該項目的下一步打算,李煜表示, 下一步團隊還會與其他機構一齊研究AI模型去 預測細胞基因表達的一些變化。而對於整個AI 醫療行業發展,李煜認為AI與人類的協作更容 易在短期內取得突破。

責任編輯:黃格煜 美術編輯:莫家威